ФИЗИКА Электронных процессов

Часть І. Электронные состояния



В. А. Москалюк, Ю. С. Синекоп, Р. Кассинг

ФИЗИКА ЭЛЕКТРОННЫХ ПРОЦЕССОВ



часть Ї ЭЛЕКТРОННЫЕ СОСТОЯНИЯ

Под редакцией кандидата технических наук, профессора В. А. Москалюка

Рекомендовано Советом факультета электроники Национального технического университета Украины "Киевский политехнический институт" в качестве учебного пособия для студентов специальности "Физическая и биомедицинская электроника"

Киев 2001

УДК 621.38(07)

Москалюк В. А., Синекоп Ю. С., Кассинг Р. Физика электронных процессов, часть І: Учебное пособие. – Киев, УкрИНТЭИ, 2001. – 148 с.

В первой части учебного пособия рассмотрены свойства обобществленных и локализованных состояний электронов в твердых телах, в том числе в кристаллах низкой размерности. Описание ведется с использованием дисперсионных кривых, приведены особенности законов дисперсии для основных полупроводниковых материалов. Изложены простейшие модели дрейфового и диффузионного движения носителей; на основе описания равновесного состояния носителей рассмотрены вопросы проводимости твердых тел.

Предназначено для студентов специальности "Физическая и биомедицинская электроника", может быть полезно студентам других специальностей и аспирантам.

ISBN 966-7505-10-3 (ч. I)

© Москалюк В. А., Синекоп Ю. С., Кассинг Р., 2001

© УкрИНТЭИ, 2001

Содержание

Преди	словие	5				
1. Эле	ктронные состояния в кристаллах	7				
1.1. Обобществленные состояния						
	1.1.1. Образование энергетических зон	7				
	1.1.2. Свойства волновой функции электронов в кристалле	12				
	1.1.3. Закон дисперсии электрона в кристалле	19				
1.1.4. Эффективная масса электрона						
	1.1.5. Распределение энергетических состояний	27				
	1.1.6. Зонная структура полупроводников	31				
1.2.	Локализованные состояния электронов	42				
	1.2.1. Донорные примеси	42				
	1.2.2. Акцепторные примеси	45				
	1.2.3. Многозарядные примеси, глубокие уровни	48				
	1.2.4. Влияние концентрации примеси на зонную структуру	50				
1.3.	Электронные состояния на поверхности	53				
	1.3.1. Макроскопическая энергетическая модель поверхности	54				
	1.3.2. Таммовские поверхностные состояния	56				
	1.3.3. Адсорбированные поверхностные состояния	58				
2. Эле	ктронные состояния и проводимость твердых тел	63				
2. Эле 2.1.	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости	63 63				
2. Эле 2.1. 2.2.	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда	63 63 66				
2. Эле 2.1. 2.2. 2.3.	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда	63 63 66 72				
2. Эле 2.1. 2.2. 2.3.	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей	63 63 66 72 72				
2. Эле 2.1. 2.2. 2.3.	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда	63 63 66 72 72 78				
 Эле 2.1. 2.2. 2.3. 3. Рав 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей	63 63 66 72 72 78 84				
 Эле 2.1. 2.2. 2.3. 3. Рав 3.1. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов	63 63 66 72 72 78 84 84				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей	63 63 66 72 72 78 84 84 84				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости	63 63 66 72 72 78 84 84 88 88				
 Эле 2.1. 2.2. 2.3. Вав 3.1. 3.2. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне	 63 63 66 72 72 78 84 84 88 88 92 				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне 3.2.3. Электроны в кристаллах низкой размерности	63 63 66 72 72 78 84 84 88 88 92 96				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне 3.2.3. Электроны в кристаллах низкой размерности 3.2.4. Заполнение локализованных состояний	63 63 66 72 78 84 84 88 88 92 96 99				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 3.3. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей заряда 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне 3.2.3. Электроны в кристаллах низкой размерности 3.2.4. Заполнение локализованных состояний Проводимость металлов	63 63 66 72 72 78 84 84 84 88 88 92 96 99 101				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 3.3. 3.4. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне 3.2.3. Электроны в кристаллах низкой размерности 3.2.4. Заполнение локализованных состояний Проводимость металлов Собственная проводимость полупроводников	63 63 66 72 72 78 84 84 88 88 92 96 99 101 105				
 Эле 2.1. 2.2. 2.3. Рав 3.1. 3.2. 3.3. 3.4. 3.5. 	ктронные состояния и проводимость твердых тел Связь зонной структуры и проводимости Виды проводимости и носители заряда Перенос носителей заряда 2.3.1. Дрейф носителей 2.3.2. Диффузия носителей заряда новесное состояние носителей Основные положения статистики электронов Расчет концентрации равновесных носителей 3.2.1. Электроны в зоне проводимости 3.2.2. Дырки в валентной зоне 3.2.3. Электроны в кристаллах низкой размерности 3.2.4. Заполнение локализованных состояний Проводимость металлов Собственная проводимость полупроводников	63 63 66 72 72 78 84 84 88 88 92 96 99 101 105 109				

4		Сод	гржание
		3.5.2. Акцепторные полупроводники	120
	3.6.	Равновесные носители в электрическом поле	123
	3.7.	Термоэлектронная эмиссия	127
		3.7.1. Закон термоэлектронной эмиссии	128
		3.7.2. Влияние электрического поля на работу выхода	133
		3.7.3. Термоэлектронные катоды	139
	П1.	Модель Кронига-Пенни. Приложение к гл. 1	144
	П3.	Интегралы Ферми-Дирака. Приложение к гл. 3	147

Предисловие

Современные тенденции развития электроники вызваны неудержимым желанием человечества повышать быстродействие компьютеров и осваивать новые информационные технологии, улучшать качество связи и передачи информации, повышать надежность локационных средств и решать все новые технические проблемы с помощью электронной техники. Такие задачи требуют соответствующих скоростей работы активных электронных компонентов, что в свою очередь вынуждает использовать для их создания как новые материалы так и все более тонкие физические эффекты. Кроме того, вполне естественным и необходимым выглядит и переход от микрометровых характерных размеров приборов к нанометровым, которые осваиваются современой технологией. При этих размерах все заметнее проявляются квантовые эффекты, которые могут быть использованы для создания принципиально новых приборов.

Учебная литература, не только предназначенная для спецкурсов, но и та, которую принято относить к фундаментальной, не может оставаться в стороне от уровня и тенденций развития электроники. Предлагаемое учебное пособие представляет собой попытку отследить эти тенденции и дать будущим специалистам основы знаний, необходимых для понимания физических процессов, особенно динамических, происходящих в современных приборах.

В первой части рассмотрены обобществленные и локализованные электронные состояния в кристаллах, которые обычно изучают в курсах по физике твердого тела. Но в отличие от традиционного изложения, описание ведется не на основе зонной теории с помощью "одномерных" энергетических диаграмм, а с использованием закона дисперсии, позволяющего рассмат-

6	Предисловие
• • • • • • • • • • • • • • • • • • • •	••••••

ривать не только изменения энергии, но и импульса. Без этого невозможно понимание особенностей процессов взаимодействия электронов как с кристаллической решеткой так и с различными частицами и квазичастицами. Кроме того, знание деталей законов дисперсии позволяет понять специфические свойства различных полупроводниковых материалов. Уделено внимание также состояниям электронов в одно- и двумерных кристаллах, которые перспективны для создания твердотельных нанометровых структур.

Более подробно, чем при традиционном изложении, рассмотрены полуклассические уравнения движения носителей и механизмы проводимости, описана простейшая модель дрейфа носителей в статических полях и особенности диффузионного движения заряженных частиц. Завершается первая часть рассмотрением особенностей равновесного состояния электроннодырочных коллективов в твердых телах и их проводимости.

1. ЭЛЕКТРОННЫЕ СОСТОЯНИЯ В КРИСТАЛЛАХ

1.1. ОБОБЩЕСТВЛЕННЫЕ СОСТОЯНИЯ

1.1.1. Образование энергетических зон

Зонная теория описывает характер и особенности электронных состояний в кристаллических твердых телах. Основные свойства поведения электронов в кристаллах можно понять из общих представлений. Известно, что энергетический спектр электронов в отдельном атоме является дискретным. В твердом теле валентные электроны, которые находились в своих атомах на одинаковых уровнях, объединяя атомы в решетку, обобществляются в коллективы и должны быть равноправными, оставаясь в то же время квантовыми частицами, энергия которых квантуется. Но с другой стороны электроны одного коллектива не могут находиться в одинаковых квантовых состояниях в соответствии с принципом Паули. Примиряет эти два на первый взгляд противоречивых требования модель, в соответствии с которой энергетические состояния электронов в кристалле по сравнению с атомными состояниями несколько изменяются вследствие взаимодействия атомов. Поэтому каждый электронный энергетический уровень изолированного атома при объединении атомов в кристалл расщепляется на подуровни, которые и образуют энергетическую полосу или зону.

Степень расщепления атомного уровня на подуровни, т.е. ширина энергетической зоны, зависит от степени взаимодействия между атомами. Чем ближе расположены атомы, тем больше энергия взаимодействия между ними и тем существеннее



0 Межатомное расстояние

Рис. 1.1. Расщепление энергетических состояний атомов натрия при уменьшении межатомных расстояний. Зона 3*s* оказывается заполненной лишь наполовину.

расщепление уровня. Это показано на рис. 1.1, откуда видно, что энергетические уровни электронов в изолированных атомах превращаются в энергетические зоны по мере уменьшения расстояния между ними. Энергетические зоны, которые образовались из разноименных уровней, могут в зависимости от определенных условий перекрываться или не перекрываться. В последнем случае зоны разрешенных значений энергии разделены так называемыми запрещенными зонами, которые представляют со-

бой интервалы значений энергии, которых электроны иметь не могут. Нужно отметить, что приведенные рассуждения касаются разрешенных уровней (как заполненных так и вакантных), а характер распределения электронов в зонах определяется конкретным типом кристалла.

Количественный анализ электронных состояний в твердом теле можно провести на основе уравнения Шредингера. Но его точное решение для кристалла невозможно, а упрощенный анализ проводится с учетом некоторых предположений и приближений. Чаще всего используются адиабатическое и одноэлектронное приближения.

Адиабатическое приближение заключается в том, что не учитывается тепловое движение кристаллической решетки. Это оправдано тем, что скорость движения электронов значительно выше скорости ядер, и за время, нужное для изменения состояния электронов, положение ядер практически не изменяется.

Одноэлектронное приближение позволяет анализировать поведение электронного коллектива в твердом теле на основе решения уравнения Шредингера для одного электрона

$$\nabla^2 \psi + \frac{2m}{\hbar^2} \cdot \left[E - U(\vec{r}) \right] \psi = 0, \qquad (1.1)$$

где вид функции потенциальной энергии U(r) учитывает силовое влияние на рассматриваемый электрон некоторого усредненного электрического поля, которое создается остальными электронами и атомными остатками кристаллической решетки.

Но и эти достаточно сильные предположения еще не дают возможности решить уравнение (1.1) из-за неопределенности и сложности нахождения потенциальной функции U(r). Существуют разнообразные методы, отличающиеся в первую очередь видом аппроксимации функции U(r), которая позволяет решить уравнение Шредингера. Наиболее наглядными с методической точки зрения являются два метода, приводящие к представлению о зонном характере энергетического спектра электронов в кристалле.

В методе сильной связи принимается, что электрон сильно связан с атомом и двигается в кулоновском поле ядра и электронных оболочек (как в изолированном атоме), а действие остальных частиц решетки рассматривается как малое возмущение. При этом U(r) представляется в виде суммы потенциальной функци электрона в изолированном атоме U_a(r) и некоторой функции $\Delta U(r)$, учитывающей влияние других атомов. Естественно ожидать, что в этом случае энергетический спектр электронов остается в основном подобным спектру в изолированном атоме, а возмущение, вносимое присутствием соседних атомов, будет тем больше, чем дальше электрон находится от ядра. Качественная картина результатов такого анализа приведена на рис. 1.2, где для сравнения изображены энергетические диаграммы электрона в изолированном атоме (а) и в цепочке близко расположенных атомов (б). Возмущение, которое вызвано присутствием соседних атомов, приводит к образованию между атомами потенциальных барьеров конечной высоты, которые являются "прозрачными" для электронов из-за туннельного эффекта. Это значит, что электроны, находившиеся в изолированных атомах в одинаковых квантовых состояниях,



Рис. 1.2. Потенциальная энергия электрона в поле отдельного атома (*a*) и периодический потенциальный рельеф кристалла (*б*). Образованные из атомных уровней зоны тем шире, чем сильнее перекрываются соответствующие волновые функции, условно изображенные в верхней части рисунка.

образуют связанную систему, для которой энергетический спектр является совокупностью близко расположенных подуровней (энергетической зоной). Естественно, что это возмущение тем больше, чем сильнее перекрываются волновые функции электронов в соответствующих состояниях. Поэтому более широкими являются зоны, образованные из выше расположенных уровней. Вполне закономерно, что отмеченное приближение выполняется тем точнее, чем сильнее электрон связан с атомом, поэтому данный метод дает удовлетворительные результаты для внутренних электронных оболочек.

Метод слабой связи используется для анализа энергетического спектра валентных электронов, которые в значительной степени коллективизированы и слабо связаны со своими атомами. Потенциальная функция U(r) в этом случае представляется в виде суммы постоянного среднего значения $U_{\rm cp}$ и некоторой периодической функции $U_{-}(r)$, характеризующей периодичность поля кристаллической решетки (рис. 1.3), а уравнение Шредингера приобретает вид дифференциального уравнения с периодическими коэффициентами. В теории таких уравнений доказывается, что решением является функция, которую в физике твердого тела называют *функцией Блоха*

$$\psi(\vec{r}) = \phi(\vec{r}) \exp(ik\vec{r}) \tag{1.2}$$

и которая представляет собой плоскую волну с амплитудой $\phi(\vec{r})$ в виде *периодической функции* с тем же периодом, что и у кристаллической решетки. Но решение такого вида возможно не при всех значениях параметра уравнения *E*, а только в некоторых его интервалах, которые соответствуют разрешенным



Рис. 1.3. Периодичность поля кристаллической решетки в сечениях, проходящих через плоскость расположения ионов (сплошная линия) и между ними (штриховая линия).

зонам. В промежутке между этими интервалами решение уравнения Шредингера также возможно, но оно соответствует мнимым значениям волнового вектора k и представляется через функции $\exp(\pm \vec{k}\vec{r})$. Такие функции не удовлетворяют условию нормировки для волновых функций из-за неограниченного возрастания и не могут быть пригодными для описания *стационарного* состояния электрона в бесконечно длинном кристалле. Поэтому энергетические интервалы, соответствующие таким решениям, называют запрещенными.

Типичным примером такого уравнения, которое часто рассматривается в математике, является так называемое уравнение Матье, к которому сводится уравнение (1.1) при выборе потенциальной функции $U_{\sim}(r)$ в виде обычной синусоидальной. Другим примером, распространенным в учебной литературе, является модель Кронига-Пенни (см. приложение к гл. 1), в которой функция $U_{\sim}(r)$ принимается в виде периодической последовательности прямоугольных потенциальных барьеров.

Таким образом, используя приближения как сильной, так и слабой связи, мы приходим к представлению о *зонном характере энергетического спектра электронов* в кристалле: разрешенные значения энергии чередуются со значениями, которых электрон иметь не может.

1.1.2. Общие свойства волновой функции электронов в кристалле

Главный вывод зонной теории о зонном характере разрешенных значений энергии является достаточно плодотворным в объяснении многих физических процессов. Но такой "одномерный" вывод недостаточен для описания некоторых более тонких явлений, когда нужно привлекать знание формы *закона дисперсии*, т.е. зависимости полной энергии электрона от его волнового вектора. Этот "двумерный" результат в принципе получается и в рассмотренных методах сильной и слабой связи, но из-за принятых приближений он имеет ограниченное значение.

Весьма мощные выводы, имеющие важное значение, следуют из самых общих свойств кристалла. Если рассматривать идеальный бесконечный кристалл, то одним из наиболее важных его свойств является периодичность кристаллической решетки, или так называемая *трансляционная симметрия*, означающая, что смещение (трансляция) решетки на целое число периодов в ту или другую сторону ничего не меняет. Это свойство потенциальной энергии электрона означает, что

$$U(\vec{r}) = U(\vec{r} + \vec{a}),$$
 (1.3)

где $\vec{a} = n_x \vec{a}_x + n_y \vec{a}_y + n_z \vec{a}_z$ называется вектором трансляции, состоящим из любого целого числа (n_x, n_y, n_z) базисних векторов $\vec{a}_x, \vec{a}_y, \vec{a}_z$ данной кристаллической решетки (рис. 1.4). Это значит, что любая точка с координатой \vec{r} и трансляционно симметричная ей точка $\vec{r} + \vec{a}$ являются физически эквивалентными и неразличимыми для электрона. В свою очередь, это значит, что вероятности нахождения электрона в окрестности этих точек одинаковы, то есть $|\psi(\vec{r})|^2 = |\psi(\vec{r} + \vec{a})|^2$. Равенство модулей не означает равенства волновых функций, они могут отличаться некоторым фазовым множителем с фазой, которую естественно считать пропорциональной вектору трансляции

$$\psi(\vec{r} + \vec{a}) = \psi(\vec{r}) \exp(i\vec{k}\vec{a}). \qquad (1.4)$$

Волновую функцию электрона можно в общем виде представить как произведение некоторой пространственно меняющейся амплитуды $\phi(\vec{r})$ на волновой множитель

$$\psi(\vec{r}) = \varphi(\vec{r}) \exp(i\vec{k}\vec{r}). \quad (1.5)$$

Для выяснения свойств $\varphi(\vec{r})$ в кристалле сместимся на некоторый вектор \vec{a} . Тог-



Рис. 1.4. Базисные векторы кристаллической решетки имеют длину, равную соответствующему периоду решетки, и направления, совпадающие с ребрами элементарной ячейки.

да в соответствии с (1.4) выражение (1.5) примет вид $\psi(\vec{r} + \vec{a}) = = \phi(\vec{r}) e^{i\vec{k}\vec{r}} e^{i\vec{k}\vec{a}}$. Но с другой стороны (1.5) остается справедливым и для точки $\vec{r} + \vec{a}$, поэтому $\psi(\vec{r} + \vec{a}) = \phi(\vec{r} + \vec{a}) e^{i\vec{k}(\vec{r} + \vec{a})}$. Сравнивая последние выражения, видим, что

$$\varphi(\vec{r}) = \varphi(\vec{r} + \vec{a}), \qquad (1.6)$$

т.е. амплитуда волновой функции является периодической. Этот же результат можно получить, используя метод слабой связи, но здесь он представляется универсальным, вытекающим из самых общих свойств кристаллической решетки. Волновую функцию электрона в таком виде называют *волной Блоха*, ее условное изображение приведено на рис. 1.5.

Важность полученного результата состоит в том, что пространственная плотность вероятности $|\psi(\vec{r})|^2 = |\phi(\vec{r})|^2$ является периодической функцией, а усредненная по периоду решетки – *постоянной вдоль всего кристалла*. Это значит, что в идеальной кристаллической решетке электронная волна распространяется без рассеивания, поэтому идеальный кристалл должен иметь нулевое сопротивление. Причиной возможного рассеивания электронных волн являются любые дефекты кристаллической решетки, которые вызывают нарушение трансляционной симметрии.



Рис. 1.5. Периодичность амплитуды волны Блоха соответствует периодичности кристаллической решетки (*a*). Примерный вид волны Блоха (*б*) и квадрат ее амплитуды (*в*) показывают, что средняя плотность вероятности нахождения электрона вдоль идеального кристалла постоянна.

Особенности поведения электронов в кристалле нельзя понять без выяснения некоторых характерных свойств волнового вектора, вытекающих также из свойств симметрии кристалла. Для простоты будем рассматривать одномерный кристалл, т.е. линейную бесконечно длинную цепочку атомов, расположенных вдоль оси x с периодом a_x , и полученные результаты обобщать для других направлений.

1. Первое свойство волнового вектора состоит в том, что невозможно указать его значение однозначно, а только с точностью до некоторого слагаемого. Чтобы это понять, поставим вопрос, может ли электрон в состояниях с различными значениями k' и k'' описываться одинаковой волновой функцией. То есть предположим, что $\psi_{k'}(x) = \psi_{k''}(x)$. Рассмотрим это равенство и в точке x+a : $\psi_{k'}(x+a) = \psi_{k''}(x+a)$, где a в данном случае – одномерный вектор трансляции. Используя свойство (1.4), последнее выражение представим в виде: $\psi_{k'}(x)e^{ik'a} =$ $= \psi_{k''}(x)e^{ik''a}$, где показатели экспонент должны быть равными. Отсюда получаем, что наше допущение возможно не только для тождественного случая k' = k'', но и для более общего, когда $k'a = k''a + 2\pi n$, где n - любое целое число. Следовательно, если

$$k' = k'' + n(2\pi/a), \tag{1.7}$$

то волновые функции $\psi_{k'}$ и $\psi_{k''}$ электрона одинаковы, а его состояния с такими значениями k неразличимы.

Минимальное расстояние вдоль оси k_x между физически эквивалентными состояниями (рис. 1.6) равно $2\pi/a_x$. Область в *k*-пространстве, включаю-

цая неэквивалентные значения волнового вектора, называют зоной Бриллюэна. Принято зоны Бриллюэна располагать симметрично относительно начала координат. На рис. 1.7 изображены первые три зоны для линейной цепочки атомов



Рис. 1.6. Состояния электронов с волновыми векторами k', $k^{''}$ и $k^{'''}$, отличающимися на $2\pi/a$, описываются одинаковыми волновыми функциями и являются неразличимыми.

(a) и двумерной прямоугольной решетки (б), а также первые зоны для простой кубической (e) и гранецентрированной кубической решетки типа алмаза (z).



Рис. 1.7. Изображение первых трех зон Бриллюэна для линейной цепочки атомов (*a*), двумерной квадратной решетки (*б*) и первых зон для простой кубической (*в*) и гранецентрированной кубической решетки типа алмаза (*г*).

Важность представления о зоне Бриллюэна заключается в том, что рассмотрение всех возможных перемещений электрона по состояниям достаточно проводить только в одной зоне Бриллюэна (как правило в первой), так как другие зоны содержат повторяющиеся состояния. С отмеченной особенностью волнового вектора связано также то, что величина $\vec{p} = \hbar \vec{k}$ не имеет обычных свойств импульса и называется квазиимпульсом.

2. Второе свойство волнового вектора, дискретность его возможных значений, проявляется для кристаллов ограниченных размеров. Для подсчета числа этих значений используем следующий прием. Рассмотрим кристалл (одномерный) длиной L_r , вдоль которого укладывается $N_r = L_r/a_r$ атомов. Путем последовательных перемещений кристалла покроем кристаллической решеткой все пространство. Это дает возможность для описания электронных состояний использовать блоховские функции (1.5). Но не все эти функции, пригодные для бесконечного кристалла, могут нас удовлетворить. Из них нужно отобрать только те, для которых выполняется условие $\psi(x) = \psi(x+L_x)$, так как точки x и x+L_x являются не просто физически эквивалентными с одинаковой плотностью вероятности, но и тождественными, с одинаковыми значениями волновых функций. Используя форму (1.5) представления блоховских функций, перепишем это условие в виде $\phi(x) e^{ik_x x} = \phi(x + L_x) e^{ik_x(x + L_x)}$. Учитывая свойство (1.6) периодичности амплитуды $\phi(x)$, получаем, что волновые функции для кристалла длиной L_x должны быть такими, что $e^{ik_x L_x} = 1$, откуда

$$k_x = n_x(2\pi/L_x) = n_x(2\pi/N_xa_x), \quad n_x = 0, \pm 1, \pm 2, \dots$$
 (1.8, a)

Этот результат показывает, что для кристалла конечных размеров возможны только такие электронные состояния, которые соответствуют *дискретному* набору значений k_x (рис. 1.8), могут быть пронумерованы и количество которых, вообще говоря, бесконечно. Но если ограничиться лишь неповторяющимися значениями, то есть находящимисяся только в первой зоне Бриллюэна ($-\pi/a_x \leq k_x \leq \pi/a_x$), то нумерацию их надо проводить в пределах $-N_x/2 \leq n_x \leq N_x/2$, а число этих состояний равно числу атомов N_x в рассматриваемом кристалле.

3. Следует отметить также, что эти разрешенные состояния распределены равномерно вдоль оси k_x с плотностью $\Delta n_x/\Delta k_x = N_x a_x/2\pi$. В реальных ситуациях количество атомов в кристалле очень велико, поэтому состояния считают расположенными квазинепрерывно.



Рис. 1.8. В кристалле конечных размеров k-состояния расположены равномерно и их количество в первой зоне Бриллюэна равно количеству атомов в кристалле.



Рис. 1.9. Элементарий объем в *k*-пространстве для подсчета плотности состояний.

В трехмерном кристалле полученные результаты естественно обобщить и для остальных направлений *k*-пространства:

$$k_y = n_y(2\pi/L_y) = n_y(2\pi/N_ya_y), \ n_y = 0, \pm 1, \pm 2, \dots \pm N_y/2, \ (1.8, 6)$$

$$k_z = n_z(2\pi/L_z) = n_z(2\pi/N_za_z), \quad n_z = 0, \pm 1, \pm 2, \dots \pm N_z/2.$$
 (1.8, e)

Здесь, как и выше, число состояний в первой зоне Бриллюэна вдоль каждой из осей k-пространства равно числу атомов вдоль соответствующих направлений в кристалле. Поскольку эти состояния нумеруются независимо во всех направлениях, то полное число состояний равно произведению $N_x N_y N_z$, т.е. полному числу атомов N в кристалле. Все состояния расположены равномерно в k-пространстве (рис. 1.9) с объемной плотностью:

$$g(k) = \frac{\Delta n}{\Delta V_k} = \frac{\Delta n_x \Delta n_y \Delta n_z}{\Delta k_x \Delta k_y \Delta k_z} = \frac{L_x L_y L_z}{8\pi^3} = \frac{V}{8\pi^3}.$$
 (1.9)

Здесь $V = L_x L_y L_z$ – геометрический объем кристалла. Если принять V = 1, то получим нормированную величину плотности состояний

$$g(k) = 1/(2\pi)^3$$
. (1.10, a)

1. Электронные состояния в кристаллах

Современная технология позволяет изготавливать так называемые структуры низкой размерности: двумерные (2D) и одномерные (1D). Они представляют собой тонкие (порядка единиц или десятков межатомных расстояний) слои или нити из полупроводников, которые ограничены потенциальными барьерами, образованными более широкозонными материалами, которые их окружают. Обычно эти барьеры достаточно высоки, чтобы локализовать электроны в поперечном сечении этих структур.

Для плоской кристаллической решетки, периодической в направлениях x и y и ограниченной в направлении z, решение уравнения Шредингера дает спектр k_z -состояний с дискретным распределением k_i , в некоторой степени подобным распределению в отдельном атоме. В направлениях x и y происходит обобществление состояний, спектр которых имеет зонный характер с блоховскими волновыми функциями. Поэтому каждому k_{iz} -состоянию свойственна двумерная плотность квазинепрерывных состояний, которая в соответствии с (1.8, a и δ) имеет значение $g_i(k_x,k_y) = L_x L_y/(2\pi)^2$, или для кристалла единичной площади

$$g_i(k_x, k_y) = 1/(2\pi)^2.$$
 (1.10, 6)

Для одномерного кристалла, ограниченного в направлениях y и z, сохраняется дискретность состояний k_{ij} вдоль k_y и k_z , но в продольном направлении x происходит расщепление состояний в квазинепрерывный спектр. Для каждого k_{ij} -состояния пригодно соотношение (1.10, a), согласно которому для кристалла единичной длины

$$g_{ij}(k_x) = 1/(2\pi).$$
 (1.10, *e*)

Из последних выражений видно, что в *k*-пространстве электронные состояния распределены равномерно, с одинаковой плотностью. Отметим, что эти значения плотности нужно удвоить, учитывая возможность заполнения состояний двумя электронами с противоположными значениями спинових моментов.

1.1.3. Закон дисперсии электрона в кристалле

Явление дисперсии характерно практически для любых волновых процессов и проявляется в зависимости фазовой скорости распространения волны от ее частоты. Поскольку $v_{\phi} = \omega/k$, то закон дисперсии можно представить в виде зависимости $k(\omega)$. В физике твердого тела принято закон дисперсии выражать через обратную функцию $\omega(k)$ или (с учетом соотношения де Бройля $E = \hbar \omega$) как зависимость E(k). Важность знания закона дисперсии состоит в том, что он позволяет описывать основные особенности электронных волн в кристалле: фазовую скорость $v_{\phi} = (E/k)/\hbar$; групповую скорость $v_{\rm rp} = (dE/dk)/\hbar$; эффективную массу электрона $m^* = \hbar^2/(d^2E/dk^2)$ (см. следующий раздел) и его энергетический спектр, а также более "тонкие" детали поведения электрона при различных значениях энергии.

Нахождение законов дисперсии или дисперсионных кривых для конкретных веществ является сложной задачей и проводится совместно вычислительными и экспериментальными методами. Мы выясним только основные свойства дисперсионных кривых, вытекающих из уже известных свойств волнового вектора.

1. Дисперсионная кривая *периодична*, т.е. $E(k) = E(k+n2\pi/a)$. Это свойство следует из того, что состояния k и $k+n2\pi/a$ в кристалле неразличимы и им соответствуют одинаковые волновые функции, т.е. и одинаковые значения энергии. Период дисперсионной кривой равен одной зоне Бриллюэна.

2. Дисперсионная кривая четная функция, т.е. E(k) = E(-k). Это легко понять, так как в любом кристалле движения электронов "вперед" и "назад" совершенно равноправны и состояниям +k и -k соответствует одинаковая энергия.

3. Для кристалла конечных размеров дисперсионная кривая – *дискретна*, или точнее – определена на точечном множестве. Это свойство непосредственно следует из дискретности волнового вектора для ограниченного кристалла. Все перечисленные свойства условно изображены на рис. 1.10. Дисперсионная кривая позволяет найти *pacnpede*-



Рис. 1.10. Условное изображение основных особенностей дисперсионной кривой: периодичности, четности и дискретности.

ление энергетических состояний электронов в зоне, исходя из плотности *k*-состояний (1.10). Этот вопрос будет рассмотрен далее, а пока что ограничимся приближенными оценками.

Поскольку число разрешенных *k*-состояний равно числу атомов в кристалле, то и число энергетических состояний в разрешенной зоне имеет такой же порядок.

В кристалле объемом ~1 мм³ находится приблизительно 10^{20} атомов, поэтому в энергетической зоне шириной порядка 1 эВ энергетические состояния расположены на расстоянии ~ 10^{-20} эВ. Такие малые энергетические интервалы позволяют считать энергетический спектр электрона в пределах разрешенной зоны квазинепрерывным.



Рис. 1.11. Изображение закона дисперсии для одномерного кристалла в виде схем повторяющихся (*a*), расширенных (*б*) и приведенных (*в*) зон.

Наглядное изображение закона дисперсии для одномерного случая представляют графически, как на рис. 1.10. Однако, если изобразить его для нескольких разрешенных энергетических зон (рис. 1.11, a), то получим так называемую схему "повторяющихся зон", которая иногда неудобна своей многозначностью. Избежать этой многозначности можно использованием схемы "расширенных зон" (рис. 1.11, b), в которой дисперсионная кривая однозначно задана на всей оси k так, что каждой энергетической зоне соответствует только одна зона Бриллюэна. Но чаще всего дисперсионные кривые изображают только в пределах первой зоны Бриллюэна (рис. 1.11, b), т.е. при помощи схемы "приведенных зон". На том же рисунке справа (z) приведена обычная "одномерная" энергетическая диаграмма, которая показывает только энергетический спектр электрона в виде чередующихся разрешенных (заштрихованы) и запрещенных зон.

Для изображения дисперсионной кривой в виде функции составляющих волнового вектора (например, k_x и k_y) приходится использовать косоугольные проекции (рис. 1.12) или совокупность линий одинаковых энергий (изоэнергетических), представляющих собой следы пересечения поверхности $E(k_x, k_y)$ плоскостями постоянной энергии.



Рис. 1.12. Способ изображения закона дисперсии в двумерном *k*-пространстве с помощью линий равной энергии.



Рис. 1.13. Поверхности равной энергии в трехмерном *k*-пространстве вблизи экстремумов закона дисперсии имеют вид эллипсоидов.

Функция $E(k_x, k_y, k_z)$ является четырехмерной поверхностью и не может быть изображена на плоском рисунке. Для этого в справочниках обычно используют набор одномерных графиков E(k) для наиболее характерных направлений зоны

1. Электронные состояния в кристаллах

Бриллюэна. Еще одна возможность кроется в изображении $E(k_x, k_y, k_z)$ с помощью изоэнергетических поверхностей, т. е. поверхностей постоянной энергии E = const в пространстве переменных k_x , k_y и k_z (рис. 1.13).

Вопрос о форме изоэнергетических поверхностей в конкретных веществах является очень сложным. Но так как основные электрические свойства кристалла связаны с электронными состояниями вблизи края энергетических зон, некоторые общие выводы можно достаточно просто получить из разложения функции $E(\vec{k})$ в степенной ряд в окрестности экстремума (верхнего или нижнего края) энергетической зоны. Для упрощения будем считать, что этот экстремум лежит в центре зоны Бриллюэна (k = 0), т.е. $E_m = E(0)$. Поскольку закон дисперсии является четной функцией, то в разложении будут отсутствовать нечетные степени. Ограничив ряд второй степенью аргумента, получаем:

$$E(k_{x},k_{y},k_{z}) = E_{m.} + \frac{1}{2} \left(\frac{\partial^{2} E}{\partial k_{x}^{2}} \Big|_{k_{x}=0} k_{x}^{2} + \frac{\partial^{2} E}{\partial k_{y}^{2}} \Big|_{k_{y}=0} k_{y}^{2} + \frac{\partial^{2} E}{\partial k_{z}^{2}} \Big|_{k_{z}=0} k_{z}^{2} \right). \quad (1.11)$$

В это выражение не включены смешанные производные типа $\partial^2 E / \partial k_i \partial k_j$, так как их всегда можно сделать равными нулю соответствующим выбором системы координат. Целесообразно (как видно будет далее) также обозначить:

$$\frac{\partial^2 E}{\partial k_i^2} = \frac{\hbar^2}{m_i^*}, \quad i = x, y, z, \tag{1.12}$$

после чего закон дисперсии приобретает вид:

$$E(\vec{k}) = E_m + \frac{\hbar^2}{2m_x^*}k_x^2 + \frac{\hbar^2}{2m_y^*}k_y^2 + \frac{\hbar^2}{2m_z^*}k_z^2.$$
(1.13)

Уравнение изоэнергетической поверхности для фиксированного значения *E'* непосредственно получается из (1.13). Представив его в канонической форме:

$$\frac{k_x^2}{c_x^2} + \frac{k_y^2}{c_y^2} + \frac{k_z^2}{c_z^2} = 1, \qquad (1.14)$$

где $c_i = \sqrt{2m_i^*(E'-E_m)/\hbar^2}$, (i = x, y, z) получим, что в общем случае изоэнергетическая поверхность является эллипсоидом с полуосями c_x , c_y и c_z (рис. 1.13). Это утверждение, вопервых, справедливо и для экстремумов, расположенных не только при k = 0, но обязательно для абсолютных. Во-вторых, оно сохраняет свою силу не только для нижнего края энергетической зоны, где $E' > E_{\min}$ и $\partial^2 E / \partial k_i^2 > 0$. У верхнего края $E' < E_{\max}$, но так как вторые производные в окрестности максимума отрицательны, то величины c_i и в этом случае остаются действительными. Для локальных экстремумов дисперсионных кривых, которые могут иметь место внутри энергетических зон, в некоторых направлениях значения c_i^2 могут быть отрицательными, а изоэнергетические поверхности будут представлять собой другие поверхности второго порядка.

В заключение отметим, что типичными формами изоэнергетических поверхностей для многих полупроводников являются эллипсоид вращения (с двумя одинаковыми полуосями) и сфера. Форму сферы имеют изоэнергетические поверхности и для свободного электрона. Поскольку свободной частице свойственна только кинетическая энергия, то

$$E = \frac{mv^2}{2} = \frac{p^2}{2m} = \frac{\hbar^2}{2m} \left(k_x^2 + k_y^2 + k_z^2\right).$$
(1.15)

Для фиксированного значения E = E' это выражение представляет собой в *k*-пространстве уравнение сферы с радиусом $\sqrt{2mE'/\hbar^2}$. Из сравнения (1.13) и (1.15) видно, що введенные нами величины m_i^* занимают такое же место в законе дисперсии для электронов в кристалле, что и обычная масса для свободного электрона.

Приведенная аналогия позволяет при помощи так называемого *метода эффективной массы* описать поведение электронов вблизи краев энергетических зон, применяя модель свободных электронов. Естественно, что этот метод удобен в пределах квадратичности закона дисперсии, хотя иногда его 1. Электронные состояния в кристаллах

используют, когда в разложении (1.11) в ряд учитывают четвертые степени.

1.1.4. Эффективная масса электрона

На электрон в кристалле действует сила периодического поля решетки и (при ее наличии) внешняя сила, которая может быть вызвана, например, электрическим полем, созданным некоторым посторонним источником. Анализ поведения электрона при совместном действии таких сил является очень сложной задачей для квантовой механики. Но есть очень простой метод, который позволяет существенно ее упростить. Для его понимания сначала установим некоторые формальные соотношения.

Зонная теория определяет для периодического кристалла электронные состояния \vec{k}_i и соответственно закон дисперсии $E(\vec{k}_i)$, которые являются *стационарными*. Электроны в них остаются как угодно долго, если нет никакого внешнего воздействия. Изменение этих состояний возможно только под действием некоторой внешней силы \vec{F} , например, электрического поля. Для простоты рассматривая одномерный случай, имеем

$$F = \frac{dp}{dt} = \hbar \frac{dk}{dt} \,. \tag{1.16}$$

Это соотношение напоминает обычный закон Ньютона, но сюда входит только внешняя сила (периодическая сила решетки учитывается зонной теорией в определении *k*-состояний) и квазиимпульс, а не импульс.

С другой стороны скорость электрона соответствует групповой скорости электронной волны де Бройля, равной

$$v = d\omega / dk = (dE / dk) / \hbar, \qquad (1.17)$$

а ускорение электрона, учитывая неявную зависимость энергии от времени через волновой вектор k(t), можно представить в таком виде:

$$a = \frac{dv}{dt} = \frac{1}{\hbar} \cdot \frac{d}{dt} \left(\frac{dE}{dk} \right) = \frac{1}{\hbar} \cdot \frac{dk}{dt} \cdot \frac{d}{dk} \left(\frac{dE}{dk} \right) = \frac{1}{\hbar} \cdot \frac{d^2E}{dk^2} \cdot \frac{dk}{dt}.$$
 (1.18)

Сравнение выражений (1.16) и (1.18) выявляет пропорциональность между внешней силой, действующей на электрон, и его ускорением. Коэффициент пропорциональности имеет размерность массы и равен ранее введенной (1.12) нами величине m^* :

$$m^* = \frac{\hbar^2}{d^2 E / dk^2}.$$
 (1.19)

Эта величина и называется эффективной массой электрона. Таким образом, сущность такого приема, называемого *методом эффективной массы*, заключается в том, что, приписывая электрону некоторую фиктивную величину m^* , мы получаем возможность абстрагироваться от действия периодического поля кристаллической решетки и рассматривать его движение как свободной частицы с массой m^* под действием только сторонних сил. Закономерно, что в этом случае задача упрощается, так как эффективную массу считаем известной. Но это возможно лишь при знании закона дисперсии для данного кристалла.

В более общем трехмерном случае справедливо векторное соотношение:

$$\vec{F} = \mathbf{m}^* \cdot \vec{a} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} m_x^* & 0 & 0 \\ 0 & m_y^* & 0 \\ 0 & 0 & m_z^* \end{pmatrix} \cdot \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \quad (1.20)$$

где \mathbf{m}^* – тензорная величина, включающая только диагональные компоненты

$$m_i^* = \frac{\hbar^2}{\left(\partial^2 E / \partial k_i^2\right)}, \quad (i = x, y, z), \quad (1.21)$$

Внедиагональные компоненты будут равняться нулю, если оси k_x , k_y и k_z будут совпадать с главными осями эллипсоида, как показано на рис. 1.13.

В трехмерном случае групповую скорость электрона находят не из (1.17), а через градиент функции $E(\vec{k})$:

$$\vec{v} = \frac{1}{\hbar} \operatorname{grad}_{k} E(\vec{k}) = \frac{1}{\hbar} \left(\vec{e}_{x} \frac{\partial E}{\partial k_{x}} + \vec{e}_{y} \frac{\partial E}{\partial k_{y}} + \vec{e}_{z} \frac{\partial E}{\partial k_{z}} \right).$$
(1.22)

Для электронных состояний вблизи края зоны, где имеет место квадратичный закон дисперсии (1.13):

$$\vec{v} = \hbar \left(\vec{e}_x \, \frac{k_x}{m_x^*} + \vec{e}_y \, \frac{k_y}{m_y^*} + \vec{e}_z \, \frac{k_z}{m_z^*} \right).$$
 (1.23)

Последние соотношения показывают, что в кристалле вектор скорости направлен перпендикулярно к изоэнергетической поверхности $E(\vec{k}) = \text{const}$ в сторону возрастания энергии. Если изоэнергетические поверхности являются сферами (как и для свободных электронов), что возможно при равных эффективных массах, то $v(\vec{k})$ и \vec{k} направлены в одну сторону.



Рис. 1.14. Эффективная масса электрона у потолка зоны отрицательна и его ускорение направлено противоположно силе (*a*), у дна зоны положительна (*б*), а сила и ускорение совпадают.

Из-за тензорного характера эффективной массы векторы ускорения и внешней силы могут не совпадать по направлению. Более того, они могут быть противоположны друг другу, когда эффективная масса отрицательна. Это имеет место, например, у верхнего края разрешенной энергетической зоны (рис. 1.14, *a*), где из-за выпуклости дисперсионной кривой $d^2E / dk^2 < 0$.

В этом нет ничего странного, если помнить, что электрон находится в поле внешних сил и в поле кристаллической решетки, но результат их действия, т.е. результирующее ускорение электрона, мы приписываем только внешней силе. Эта ситуация $(m^* < 0)$ реализуется, когда вся работа внешних сил и кроме

того часть внутренней кинетической энергии электрона расходуется на увеличение потенциальной энергии электрона. Он в этом случае замедляется и внешней силой воспринимается как частица с отрицательной массой.

Если часть работы, совершаемой внешней силой, увеличивает кинетическую энергию электрона, направление силы совпадает с направлением ускорения и $m^* > 0$. Такое положение характерно для электронных состояний у нижнего края энергетической зоны, где $d^2E/dk^2 > 0$ (рис. 1.14, б). При этом часть работы внешних сил увеличивает потенциальную энергию электрона, и его ускорение будет меньшим, чем у свободного электрона, поэтому $m^* > m$.

Если же кинетическая энергия электрона увеличивается не только за счет работы внешних сил, но и за счет потенциальной энергии его взаимодействия с решеткой, то скорость возрастает так, будто он легче обычного, то есть $m^* < m$.

Нужно помнить, что применение понятия эффективной массы совсем не означает изменения гравитационной массы электрона, а лишь отображает факт изменения его динамических свойств, т.е. чувствительности к внешним силам за счет влияния периодического поля кристаллической решетки.

Важность проведенного рассмотрения состоит в том, что мы получаем так называемые *полуклассические уравнения движения* (1.20) и (1.22). Они по форме напоминают обычные уравнения классической механики, но учитывают квантовые особенности поведения электронов в периодическом поле решетки, и на сегодня являются основным средством для проведения анализа поведения электронов в кристалле под действием внешних сил.

1.1.5. Распределение энергетических состояний

Рассмотрим вопрос о распределении электронных состояний, который очень важен для количественного вычисления основных электрических параметров вещества. Выше мы уже выяснили, что в *k*-пространстве электронные состояния распределены равномерно с плотностью $g(\vec{k}) = 1/(2\pi)^3$. Поскольку мы не обращали внимание на спин, то этот результат нужно удвоить, так как каждое состояние может быть заполнено двумя электронами с противоположными направлениями спиновых моментов

$$g(\vec{k}) = 2/(2\pi)^3$$
. (1.24)

Во многих случаях интерес вызывает распределение не *k*-состояний, а распределение энергетических состояний, характеризуемое плотностью

$$g(E) = dn_{\rm c}/dE, \qquad (1.25)$$

которая имеет смысл количества электронных состояний в единичном энергетическом интервале.



Рис. 1.15. Элементарный объем *k*-пространства, соответствующий приращению энергии, для подсчета плотности состояний.

Функцию g(E) можно сравнительно легко найти для состояний у края энергетических зон, где закон дисперсии квадратичен (1.13), используя распределение k-состояний. Для этого рассмотрим некоторый объем k-пространства dV_k между двумя изоэнергетическими поверхностями, которые соответствуют значениям энергии E и E + dE (на рис. 1.15 изображена только восьмая часть этого объема). Число

состояний dn_c в этом слое, с одной стороны, равно $g(k)dV_k$, а с другой – g(E)dE. Так как это те же самые состояния, то

$$g(E) = g(k)dV_k/dE. \qquad (1.26)$$

Остается выяснить, как изменяется объем V_k внутри изоэнергетической поверхности при изменении значения энергии. Объем эллипсоида определяется его полуосями (1.14), поэтому

$$V_k = rac{4}{3}\pi c_x c_y c_z = rac{4}{3}\pi rac{\left(m_x^* m_y^* m_z^*
ight)^{1/2}}{\hbar^3} \left[2(E-E_m)
ight]^{3/2}.$$

Дифференцируя это выражение по E и подставляя производную в (1.26), получим после некоторых преобразований искомую функцию распределения плотности энергетических состояний:

$$g(E) = 4\pi \left(2m_d^* / h^2\right)^{3/2} \left(E - E_m\right)^{1/2}.$$
 (1.27)

В это выражение введена величина

$$m_d^* = \sqrt[3]{m_x^* m_y^* m_z^*},$$
 (1.28)

являющаяся средней из компонентов тензора эффективной массы и называемая эффективной массой плотности состояний. Отметим, что плотность энергетических состояний имеет размерность $1/M^3 \cdot Дж$.

Полученная функция распределения в принципе пригодна как для нижнего, так и для верхнего края энергетической зоны. Но для состояний у верхнего края, где $m^* < 0$ и $E < E_{\max}$ эта формула неудобна и целесообразно представить ее, поменяв знаки

$$g(E) = 4\pi \left(2m_{pd}^* / h^2\right)^{3/2} \left(E_{\max} - E\right)^{1/2}.$$
 (1.29)

Тут $m_{pd}^{*} = -m_{d}^{*}$ и (E_{\max} - E) положительные величины.

Таким образом, из выражений (1.27) и (1.29) следует, что плотность состояний в разрешенных зонах параболически возрастает с удалением от края внутрь зоны (рис. 1.16, *a*). В запрещенной зоне состояния, естественно, отсутствуют. Вдали от края зоны вопрос о нахождении функции g(E) аналитически весьма сложен и обычно решается экспериментально.

Модель двумерного (2D) кристалла представляет собой атомарную структуру, периодическую в двух направлениях (x и y), и одноатомный слой в третьем измерении (z). Локализация электрона в поперечном направлении кристалла вызывает квантование k_{iz} -состояний, которым соответствует дискретный энергетический спектр E_i . Но вдоль кристалла электронные состояния обобществляются, поэтому каждому $E_i(k_{iz})$ -состоянию



Рис. 1.16. Плотность энергетических состояний у краев зон по мере удаления от края параболически возрастает для трехмерных (3D) кристаллов (a), постоянна для двумерной (2D) решетки (δ) и спадает для одномерной (1D) цепочки атомов (s).

соответствует своя зона непрерывных (для бесконечного в *x*и *y*-направлениях кристалла) состояний, разрешенных для движения электронов в плоскости кристалла. В приближении эффективной массы у края каждой *i*-й разрешенной зоны закон дисперсии можно считать квадратичным и подобным (1.13):

$$E(k_x, k_y) = E_i(k_{iz}) + \hbar^2 k_x^2 / 2m_x^* + \hbar^2 k_y^2 / 2m_y^*. \qquad (1.30)$$

В соответствии с этим законом изоэнергетическая линия E = const имеет форму эллипса $(k_x/c_x)^2 + (k_y/c_y)^2 = 1$, где полуоси c_x и c_y имеют тот же смысл, что и в (1.14), а его площадь

$$S_k = \pi c_x c_y = 2\pi \sqrt{m_x^* m_y^*} (E - E_i) / \hbar^2$$
.

Придавая энергии приращение dE, получаем $dS_k = 2\pi(\sqrt{m_x^*m_y^*}/\hbar^2)dE$. В этом слое находится некоторое количество состояний $dn_c = g_i(E)dE = g_i(k_x, k_y)dS_k$, откуда можно получить двумерную плотность энергетических состояний в каждой *i*-зоне: $g_i(E) = g_i(k_x, k_y) dS_k/dE$. Удваивая значение (1.10, б) для плотности *k*-состояний с учетом спинового квантового числа, получим

$$g_i(E) = 4\pi m_{di}^* / h^2$$
, (1.31)

где $m_{di}^* = \sqrt{m_x^* m_y^*}$ – так называемая эффективная масса двумерной плотности состояний. Отсюда видно, что для двумерных кристаллов характерно равномерное распределение энергетических состояний в зонах (рис. 1.16, б). Отметим также, что их плотность имеет размерность $1/m^2 \cdot Дж$.

Простейшая модель одномерных (1D) кристаллов представляет собой линейную цепочку атомов, обладающую свойством трансляционной симметрии только в одном (x) направлении. Локализация электронов в y- и z-направлениях обуславливает квантование состояний k_{iy} и k_{jz} . Им соответствует дискретный энергетический спектр E_{ij} (k_{iy} , k_{jz}), в некоторой мере подобный спектру электронов в отдельном атоме. Но взаимодействие валентных электронов в продольном направлении вызывает их обобществление и соответственно зонный спектр энергий их движения вдоль кристалла. Квадратичный закон дисперсии в таких "одномерных" зонах в приближении эффективной массы можно представить в виде

$$E(k_x) = E_{ij}(k_{iy}, k_{jz}) + \hbar^2 k_x^2 / 2m_x^*, \qquad (1.32)$$

а плотность энергетических состояний определить из подобного предыдущим случаям соотношения $g_{ij}(E) = g_{ij}(k_x) dk_x/dE$, где плотность k_x -состояний задается удвоенным значением (1.10, ε). После вычисления производной и несложных преобразований получим

$$g_{ij}(E) = \sqrt{2m_x^*/h^2} (E - E_{ij})^{-1/2}$$
. (1.33)

Как видно, плотность энергетических состояний спадает при удалении от края зоны (рис. 1.16, e) и они "локализуются" вблизи уровня E_{ij} . Нужно также помнить, что плотность одномерных состояний имеет размерность $1/м \cdot Дж$.

1.1.6. Зонная структура полупроводников

Полупроводниковые свойства имеют простые вещества: бор В, углерод (графит) С, кремний Si, германий Ge, олово (серое) Sn, фосфор Р, мышьяк As, сурьма Sb, сера S, селен (красный) Se, теллур Te, йод J. К полупроводникам относятся также бинарные соединения типа $A_x B_{8-x}$, где A – элемент группы x, а B – элемент группы (8–x). Наиболее распространены и применимы из них – вещества типа $A_3 B_5$, к которым относятся соединения азота (нитриды), фосфора (фосфиды), мышьяка (арсениды), сурьмы (антимониды) с элементами третьей группы (бора, алюминия, галлия, индия). Перспективными являются соединения типа $A_2 B_6$, а также $A_4 B_4$. Полупроводниковые свойства проявляют тройные и более сложные соединения, а также некоторые органические вещества.



Рис. 1.17. Кристаллическая структура алмаза. Затушеваны атомы, образующие гранецентрированную кубическую ячейку. Незатушеванные атомы образуют вторую подрешетку.

Широкое применение в электронике имеют атомарные полупроводниковые вещества (германий и кремний). Структура верхних оболочек этих атомов, а также других атомов этой подгруппы IV группы имеет вид *ns²np²*. Из-за близкого энергетического расположения s- и p-уровней эти атомы являются четырехвалентными. Объединяясь в кристалл при помощи ковалентных связей, они образуют решетку такую же, как и у алмаза. Она содержит лве

решетки, составленные из двух грандренерироважные скубиесска ких ячеек, сдвинутых друг относительно друга на четверть диагонали куба (рис. 1.17). В таком кристалле каждый атом связан с четырьмя ближайшими, которые пространственно расположены так, как это приведено на рис. 1.18, *а*. В каждой связи участвуют два электрона, по одному от каждого атома, что является характерным для ковалентной связи. Эту ситуацию часто изображают на плоскости условно так, как это приведено на рис. 1.18, *б*.





Рис. 1.18. Пространственное расположение ближайших атомов в решетке алмаза (a) и ее условное представление на плоскости (δ).

Рис. 1.19. Образование энергетических зон при сближении атомов в ковалентных кристаллах элементов IV группы с решеткой типа алмаза. Зона проводимости и валентная зона включают гибридные *sp*³-состояния, образованные из атомных *s*- и *p*-состояний.

При образовании кристаллической решетки алмаза из N одинаковых атомов *s*-уровни атомов энергетически расщепляются на две группы (или зоны) подуровней (рис. 1.19). Верхняя группа, которая состоит из N подуровней, соответствует антисимметричным волновым функциям, которым при перекрывании свойственны силы отталкивания. Нижняя содержит также N состояний, электроны в которых имеют симметричные волновые функции, что обуславливает силы притяжения при их взаимодействии. Этим двум типам волновых функций соответствует существенно различная энергия, поэтому между двумя зонами подуровней существует запрещенная зона. Подобным образом происходит расщепление *p*-состояний на две зоны, в каждой из которых есть по 3 состояния на атом. В результате перекрытия этих зон, как это показано на рис. 1.19, образуются две гибридные sp^3 -зоны, которые состоят из 4Nсостояний каждая (из них N подобны s-, а 3N p-состояниям атомов). Естественно, что все 4N валентных электрона кристалла заполняют нижние состояния, поэтому нижняя зона, которую называют валентной, оказывается полностью заполненной, а верхняя зона (зона проводимости) свободной.

1. Электронные состояния в кристаллах



Рис. 1.20. Пространственное распределение плотности заряда электронов валентной зоны (a) и зоны проводимости (δ) .

Симметричный характер волновых функций электронов в валентной зоне приводит к тому, что плотность электронного заряда будет максимальной вдоль прямых, которые условно изображают (рис. 1.17) связи между атомами в кристаллической решетке. На плоской модели кристаллической решетки распределение заряда электронов валентной зоны приведено на рис. 1.20, а. Для электронов, которые находятся в состояниях зоны проводимости, плотность заряда (рис. 1.20, б) равна нулю вдоль линии связи, что ведет к преобладанию сил отталкивания между атомами. Поэтому переход электронов из валентной зоны в зону проводимости приводит к ослаблению связей между атомами в кристалле.

Очень полезно с точки зрения описанной картины образования энергетических зон сравнить между собой элементы четвертой группы (некоторые их параметры приведены в табл. 1.1).

№	Элемент	Валентная оболочка	Параметр решетки, Å	Запрещенная зона, эВ	Электрические свойства	Т _{плав} , К
2	Углерод С (алмаз)	$2s^{2}2p^{2}$	3,6	5,4	Диэлектрик	4300
3	Кремний Si	$3s^23p^2$	5,4	1,2	Высокоомный полупроводник	1683
4	Германий Ge	$4s^24p^2$	5,6	0,78	Полупроводник	1211
5	Олово Sn (cepoe)	$5s^25p^2$	6,5	~0,1	Низкоомный полупроводник	905
6	Свинец Pb	$6s^{2}6p^{2}$	7,0	0	Металл	601

Таблица 1.1

С увеличением порядкового номера элементов возрастает номер и радиус орбиты, на которой находятся валентные электроны. Соответственно растут "размеры" атомов и уменьшается энергия их ионизации. Это сопровождается уменьшением ширины запрещенной зоны, а значит и энергии связи между атомами в решетке (количественной мерой энергии связи может быть температура плавления, которая уменьшается с ростом номера периода). Различие в ширине запрещенной зоны приводит к тому, что алмаз является типичным диэлектриком, германий и кремний – типичными полупроводниками. Олово с решеткой типа алмаза (α -модификация или серое олово) является низкоомным полупроводником с очень узкой запрещенной зоной, близкой к 0. Часто встречается такая модификация кристаллической решетки олова (белое олово), при которой оно является типичным металлом. Замыкающий приведенную таблицу свинец является также типичным металлом.

Описанная модель энергетических зон объясняет наличие отрицательного температурного коэффициента ширины запрещенной зоны. Известно, что ширина запрещенной зоны полупроводника меняется с температурой приблизительно линейно:

$$\Delta E(T) = \Delta E(0) + \alpha T, \qquad (1.34)$$

причем для кремния и германия величина α отрицательна. Это объясняется тем, что с повышением температуры вследствие теплового расширения увеличивается расстояние между атомами и в соответствии с рис. 1.19 уменьшается ширина запрещенной зоны.

Рассмотрим особенности зонной структуры наиболее распространенных полупроводниковых материалов. Как отмечалось выше, периодичность закона дисперсии электронов в кристалле дает возможность его изображения в пределах только первой зоны Бриллюэна. Для решетки типа алмаза она имеет вид октаэдра со срезанными вершинами (рис. 1.21). Тут же показано, как такими фигурами можна заполнить все k-пространство. Вместо проекций волнового вектора на основные направления, заданные числовым набором кристаллографических индексов, часто используют другую систему обозначений. В ней характерные точки и линии симметрии в зоне Бриллюэна обозначаются буквами, а закон дисперсии задается в виде одномерной функции вдоль тех или иных направлений между


Рис. 1.21. Форма первой зоны Бриллюэна для кристаллической решетки типа алмаза и обозначение характерных точек и направлений симметрии.



Рис. 1.22. Дисперсионные кривые в зоне проводимости и валентной зоне кремния в Л- и Д-направлениях. Расстояние между абсолютными экстремумами кривых соответствует запрещенной зоне.

двумя точками. Центр зоны Брилюэна принято обозначать буквой Γ , направления эквивалентные [100] – Δ , а соответствующие им шесть точек на поверхности зоны – X. Восемь эквивалентных направлений [111] обозначают буквой – Λ , а точки их пересечения с поверхностью – L.

Кремний. Вид дисперсионных кривых кремния для двух направлений волнового вектора приведен на рис. 1.22. Наличие нескольких кривых в валентной зоне и зоне проводимости может быть связано с описанной выше картиной образования гибридных *sp*³зон и соответственно с наличием *s*- и *p*-состояний, которым свойственны различные законы дисперсии. Нужно отметить, что у раз-

личных авторов вид дисперсионных кривых иногда заметно различается, что может объясняться как отличием в методах вычисления этих кривых, так и различной степенью их детализации. Не касаясь этих различий, рассмотрим только наиболее существенные особенности законов дисперсии.

Волновые функции электронов и соответствующие им дисперсионные кривые для валентной зоны могут быть построены на основании волновых функций электронов в атомных *p*-состояниях. Им соответствуют три кривые на рис. 1.22. Дисперсионная кривая для *s*-состояний лежит около нижнего края валентной зоны и на рисунке не показана. Максимум энергии валентной зоны *вырожден* и находится в центре зоны Бриллюэна для двух ветвей, для третьей за счет спин-орбитального взаимодействия он опускается на 0,035 эВ. Разная кривизна этих ветвей в максимуме дает основание говорить о разной эффективной массе подвижных носителей: для "легких" дырок $m_{p,\pi}^* \approx 0,16m$, для "тяжелых" $m_{p,\pi}^* \approx 0,5m$. Для вычисления эффективной плотности состояний у края валентной зоны удобно пользоваться некоторой усредненной массой:

$$m_{pd}^* = \left(m_{p\pi}^{*3/2} + m_{p\pi}^{*3/2}
ight)^{2/3} \approx 0,56m.$$

Структура зоны проводимости еще более сложная, чем валентной, из-за того что здесь наблюдается перекрытие не только одной *s*- и трех *p*-подзон, но и более высоких. Абсо-

лютный минимум в зоне проводимости имеет одна из *p*-ветвей в Δ -направлении на расстоянии 0,855 отрезка Г-Х. Принимая во внимания симметрию зоны Бриллюэна (рис. 1.23), можно считать, что этот минимум состоит из шести эквивалентных минимумов, или, как часто их называют, *долин*. На рис. 1.23 эти долины изображены при помощи изоэнергетических поверхностей, которые имеют форму эллипсоида вращения. Так как полуоси эллипсоида характеризуют компоненты тензора эффективной массы, удобно говорить о *продольной* и *поперечной*



Рис. 1.23. Расположение изоэнергетических поверхностей, соответствующих ∆-долинам у дна зоны проводимости кремния, в первой зоне Бриллюэна.

эффективных массах электронов, которые для кремния соответственно имеют значения: $m_l^* \approx 0.98m$ и $m_t^* \approx 0.19m$. Иногда при определении удельной электропроводности пользуются некоторой усредненной скалярной величиной эффективной массы,

которую называют *омической* (см. раздел 2.3) и обратную величину которой находят как среднее арифметическое из обратных значений эффективных масс в трех направлениях. Поэтому $m_c^* = 3m_t^*m_l / (2m_l^* + m_t^*) \approx 0.26m$. Наличие шести эквивалентных долин у дна зоны проводимости означает шестикратное увеличение количества состояний. Это учитывают, сохраняя неизменным вид формулы (1.27) для функции распределения плотности состояний и используя эффективную массу плотности, как среднее геометрическое из компонентов тензора эффективной массы, умноженное на коэффициент $M^{2/3}$. Для кремния M = 6, поэтому $m_d^* = 6^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 1,08m$.

Ширина запрещенной зоны у кремния различна в разных точках зоны Бриллюэна. В центре зоны при k = 0 она равняется $\approx 2,5$ эВ. Однако шириной запрещенной зоны, как правило, считают расстояние между абсолютными экстремумами, т.е. для кремния расстояние от максимума валентной зоны в точке Г до минимума зоны проводимости в Δ -направлении. Это расстояние равно $\approx 1,2$ эВ при абсолютном нуле и $\approx 1,1$ эВ при комнатной температуре, уменьшаясь с ростом температуры по закону (1.34), где $\alpha = -4 \cdot 10^{-4}$ эВ/К.

С точки зрения формы запрещенной зоны кремний относят к так называемым *непрямозонным* полупроводникам. Это значит, что получив энергию, равную минимальному расстоянию между зонами, валентный электрон может осуществить переход в зону проводимости, если меняется также и значение его квазиимпульса. Такой переход, обозначенный стрелкой на рис. 1.22, называется *непрямым*. Прямой переход осуществляется без изменения квазиимпульса и требует большей энергии. Эти понятия существенны при рассмотрении различных механизмов генерации и рекомбинации подвижных носителей заряда. Важным параметром с точки зрения эмиссионных свойств является энергия электронного сродства, которая определяется как расстояние от дна зоны проводимости до уровня вакуума (см. раздел 1.3.1) и для кремния равна 4,05 эВ. Германий. Законы дисперсии и структура энергетических зон германия во многом подобны кремнию. Валентная зона (рис. 1.24) имеет вырожденный максимум в центре зоны Бриллюэна и характеризуется наличием "легких" и "тяжелых" дырок ($m_{p\,\pi}^* = 0,04m$, $m_{p\,\tau}^* = 0,34m$), что соответствует эффективной массе плотности состояний $m_{pd}^* = (m_{p\,\pi}^{*\ 3/2} + m_{p\,\tau}^{*\ 3/2})^{2/3} \approx$ $\approx 0,36m$. Третья ветвь отстает от края зоны на величину энергии спин-орбитального взаимодействия, которая равна 0,28 эВ.



Рис. 1.24. Дисперсионные кривые в зоне проводимости и валентной зоне германия в Λ и Δ - направлениях. Ширина запрещенной зоны германия – расстояние между абсолютными экстремумами кривых.



Рис. 1.25. Изоэнергетические поверхности, соответствующие L-долинам у дна зоны проводимости германия. Эллипсоиды вращения лежат лишь наполовину в первой зоне Бриллюэна.

Наиболее существенное отличие от кремния состоит в том, что как минимум в центре зоны Бриллюэна, так и абсолютный минимум зоны проводимости образован *s*-ветвью, причем последний лежит на границе первой зоны Бриллюэна в точке L. Соответствующие этому минимуму изоэнергетические поверхности изображены на рис. 1.25 и имеют вид эллипсоидов вращения. Анизотропия эффективной массы у германия выражена еще больше ($m_l^*=1,58m$, $m_t^*=0,08m$), поэтому скалярное значение омической эффективной массы $m_c^* \approx 3m_t^*/2 \approx 0,12m$. Эффективная масса плотности состояний электронов в зоне про-

водимости $m_d^* = 4^{2/3} \left(m_t^{*2} m_l^* \right)^{1/3} \approx 0,56m$ учитывает наличие

восьми эквивалентных долин, каждая из которых лежит лишь наполовину в первой зоне Бриллюэна (рис. 1.25), поэтому M = 8/2 = 4.

Ширина запрещенной зоны германия в центре зоны Бриллюэна 1,1 эВ. Расстояние между абсолютными экстремумами энергетических зон, которое обычно называют запрещенной зоной, равно 0,78 эВ при нулевой температуре и 0,66 эВ при комнатной. Температурный коэффициент ширины запрещенной зоны практически такой же, как и у кремния, энергия электронного сродства у германия 4,0 эВ.

Арсенид галлия. GaAs является наиболее распространенным полупроводником из группы соединений A_3B_5 , которые иногда называются интерметаллами. Эти соединения обычно образуют кристаллическую решетку типа цинковой обманки, которая по форме подобна решетке алмаза и отличается от нее только тем, что атомы разного сорта чередуются, образуя две отдельные кубические гранецентрированные подрешетки Ga и As. При этом каждый атом одного вида связан с четырьмя ближайшими соседями другого вида по существу ковалентными связями, хотя различное количество валентных электронов, которые обобществляются от атомов Ga и As (три и пять соответственно), обуславливает заметный вклад ионной связи.

В структуре энергетических зон атомарных полупроводников IV группы и соединений A_3B_5 много общего. Так, например, характер дисперсионных кривых в валентной зоне арсенида галлия аналогичен германию и кремнию. В центре зоны (рис. 1.26) находятся максимумы двух ветвей, соответствующих "легким" ($m_{p\pi}^* \approx 0,1m$) и "тяжелым" ($m_{pB}^* \approx 0,5m$) дыркам. Третья ветвь за счет спин-орбитального взаимодействия лежит ниже на 0,33 эВ.

В отличие от германия и кремния большинство соединений A_3B_5 являются *прямозонными* полупроводниками, т.е. абсолютный минимум зоны проводимости лежит в центре зоны Бриллюэна (точка Г). Для GaAs расстояние между экстремумами составляет при нулевой температуре 1,5 эВ, уменьшаясь до 1,45 эВ при комнатной. Эффективная масса электрона вблизи абсолютного экстремума (в Г-долине) является скалярной величиной ($m^* \approx 0.067m$), так как соответствующие изоэнергетические поверхности являются практически сферическими.

Одной из наиболее ярких особенностей арсенида галлия является наличие верхней долины на высоте ≈ 0,3 эВ от края зоны проводимости в точке L. Изоэнергетические поверхности у дна этой долины представляют со-



Рис. 1.26. Характер дисперсионных кривых для арсенида галлия. Абсолютные экстремумы (Г-долины) находятся в центре зоны Бриллюэна. Изоэнергетические поверхности, соответствующие ближайшим верхним Lдолинам зоны проводимости, расположены так же, как и у германия.

бой эллипсоиды вращения, еще более вытянутые, чем у германия. Это говорит о большом различии продольной и поперечной эффективных масс, которые имеют соответственно значения: $m_l^* = 1,9m$, $m_t^* = 0,075m$. Так же, как и в германии, расчитываются омическая эффективная масса $m_c^* \approx 3m_t^*/2 \approx 0,11m$

и эффективная масса плотности состояний $m_d^* = 4^{2/3} \left(m_t^{*2} m_l^* \right)^{1/3} pprox$

≈ 0,56*m*. С этой особенностью Г- и L-долин связан эффект возникновения отрицательной дифференциальной подвижности. В зоне проводимости GaAs есть и более высокие долины, например, на расстоянии ≈ 0,5 эВ в Х-направлении, но заметное влияние на электрические свойства они имеют только при наличии очень сильного электрического поля. Энергия электронного сродства составляет 4,07 эВ.

1.2. ЛОКАЛИЗОВАННЫЕ СОСТОЯНИЯ ЭЛЕКТРОНОВ

Основные выводы зонной теории получены для идеальной кристаллической решетки. Однако практически все кристаллы, которые используются в электронике, далеки от структурного совершенства. Это обусловлено, во-первых, технологическими трудностями получения совершенных кристаллов и, во-вторых, специальным созданием дефектов для придания веществу нужных свойств. Различные дефекты (примеси, дислокации, поверхностные дефекты, межзеренные границы и другие) влияют на зонную структуру, а значит, и на электрические свойства вещества; их роль особенно возрастает при использовании пленочной технологии. Наиболее изученным является влияние примесных атомов, которые специально вводятся для целенаправленного изменения электрических свойств вещества.

1.2.1. Донорные примеси

Классическим и простейшим примером донорного полупроводника является германий или кремний с решеткой типа алмаза, в которых часть атомов полупроводника замещена пятивалентными примесными атомами. Четыре валентных электрона каждого примесного атома участвуют в образовании связей с соседними атомами решетки, пятый валентный электрон остается связанным со своим атомом. Однако энергия этой связи из-за "отвлекающего" влияния других атомов в кристалле намного меньше, чем для свободных атомов, и составляет, как будет показано далее, величины порядка 10⁻² эВ. В результате какого-либо возбуждения примесный атом может быть ионизирован и пятый электрон становится свободным. На языке энергетической модели полупроводника процесс ионизации означает появление подвижного электрона в зоне проводимости. Поэтому такой примесный атом называют донором, а полупроводник с пятивалентный примесью – донорным.

Для определения возможных энергетических состояний "лишнего" электрона необходимо решить уравнение Шредингера в поле кристаллической решетки, периодичность которой нарушена присутствием примесного атома. Характер возмущения изображен на рис. 1.27, б и может рассматриваться как наложение на периодическое поле решетки кулоновского поля однократно заряженного по отношению к электрону примесного атомного остатка. Для упрощения уравнения Шредингера в таком поле примем следующие допущения. Во-первых, учтем периодическое поле решетки, используя метод эффективной массы, рассматривая движение электрона как свободной частицы с массой *m*^{*} в центральном поле атомного остатка. Второе допущение относится к



Рис. 1.27. Пятивалентный атом примеси в четырехвалентном полупроводнике. Лишний электрон находится на расстояниях около десятка периодов кристаллической решетки (*a*). Искажение периодического поля решетки полем иона примеси (δ) и энергетический спектр лишнего электрона (ϵ). Масштаб (ϵ) увеличен по сравнению с (δ).

выбору возмущающей потенциальной функции $U(\vec{r})$ в этом уравнении. Поле атома примеси ослабляется экранирующим влиянием соседних атомов кристалла, что можно учесть введением диэлектрической проницаемости ε материала основного вещества. Основанием для использования такого макроскопического параметра, как диэлектрическая проницаемость, является тот факт, что пятый электрон примеси отдален от своего атома и радиус его орбиты составляет несколько периодов кристаллической решетки (рис. 1.27, *a*). Учитывая это ослабление, запишем выражение для потенциальной энергии электрона

$$U(\vec{r}) = -e^2/4\pi\epsilon\epsilon_0 r$$
. (1.35)

Начало отсчета этой энергии соответствует минимальной энергии свободного от атома электрона, то есть нижнему краю зоны проводимости E_c (рис. 1.27, *в*). Подставляя (1.35) в уравнение Шредингера, получим

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi - \frac{e^2}{4\pi\varepsilon\varepsilon_0 r}\psi = E\psi, \qquad (1.36)$$

что совпадает с уравнением для атома водорода с той разницей, что вместо массы электрона здесь стоит эффективная масса, а вместо диэлектрической проницаемости вакуума ε_0 – диэлектрическая проницаемость полупроводника $\varepsilon\varepsilon_0$. Естественно, что и результат решения (1.36) будет таким же, как и для атома водорода с учетом указанных изменений.

Главным итогом является вывод о том, что энергетический спектр электрона будет дискретным и определяется квантовым числом n = 1, 2, 3, ... (рис. 1.27, *в*):

$$E_n = -\frac{m^* e^4}{8\varepsilon^2 \varepsilon_0^2 h^2 n^2} = -\frac{m e^4}{8\varepsilon_0^2 h^2} \frac{m^*}{m} \frac{1}{\varepsilon^2} \frac{1}{n^2} = -13, 6\frac{m^*}{m} \frac{1}{\varepsilon^2} \frac{1}{n^2}$$
(3B). (1.37)

В основном состоянии (n = 1) электрон находится на уровне E_1 (рис. 1.27, *в*), который называют *донорным уровнем* E_{π} и который при $m^* \approx 0, 1m$ и $\varepsilon \approx 10$ лежит приблизительно на 0,01 эВ ниже дна зоны проводимости. Эта величина $E_c - E_{\pi} = \Delta E_{\pi}$ равна энергии ионизации донорного атома. Значения энергии ионизации донорных атомов в германии и кремнии приведены в табл. 1.2.

Таблица 1.2

Доноры	Энергия ионизации, эВ		Ammonth	Энергия ионизации, эВ	
	Германий	Кремний	Акцепторы	Германий	Кремний
Р	0,0120	0,044	В	0,0104	0,045
As	0,0127	0,049	Al	0,0102	0,057
Sb	0,0096	0,039	Ga	0,0108	0,065
Bi	0,0127	0,039	In	0,0112	0,160
			Ti	0,0131	

Вид волновой функции электрона в основном состоянии E_1 , как и в атоме водорода, характеризуется множителем $\exp(-r/r_m)$, где r_m имеет тот же смысл, что и радиус первой боровской орбиты $a_0 = h^2 \varepsilon_0 / \pi m e^2 = 0,53$ Å, и соответствует расстоянию, на котором радиальная плотность вероятности $4\pi r^2 |\psi|^2$ нахождения электрона в точке r максимальна. С учетом принятых допущений относительно диэлектрической проницаемости и эффективной массы

$$r_m = \frac{h^2 \varepsilon \varepsilon_0}{\pi m^* e^2} = a_0 \varepsilon \frac{m}{m^*}.$$
 (1.38)

При $m^* \approx 0,1m$ и $\varepsilon \approx 10$ радиус орбиты "лишнего" электрона $r_m \approx 53$ Å, что составляет приблизительно десять периодов кристаллической решетки. Из этих оценок следует, что область локализации электрона содержит более тысячи атомов полупроводника и использование в данной модели такого макроскопического понятия, как диэлектрическая проницаемость вполне оправдано.

Кроме основного состояния E_{π} электрона в примесном атоме существуют еще и возбужденные состояния, которые соответствуют (1.37) для n = 2,3,4...: $E_n = E_{\pi}/n^2$. Спектр этих состояний изображен на рис. 1.27. Наличие уровней возбуждения приводит к появлению дополнительных линий в экспериментальных спектрах поглощения света примесными полупроводниками.

1.2.2. Акцепторные примеси

Подобную ситуацию можно наблюдать в полупроводниках четвертой группы при замещении некоторой их части атомами

1. Электронные состояния в кристаллах

третьей группы. Три валентных электроны примесного атома участвуют в образовании связей с четырьмя соседними атомами, при этом одна валентная связь остается свободной. Эта вакансия может быть заполнена электроном из другой валентной связи под действием какого-то возмущения, вследствие чего в системе связей создается свободное состояние, которое может двигаться под действием, например, электрического поля. Это движение удобно интерпретировать как движение некоторой квазичастицы, называемой дыркой. На языке энергетической модели такой процесс соответствует переходу электрона из валентной зоны в более высокое вакантное состояние, связанное с атомом примеси. В результате примесный атом становится отрицательно заряженным ионом, а в валентной зоне образуется свободная дырка. Это объясняет, почему такой атом называют акцептором, а полупроводник с примесью атомов III группы акцепторным.

При анализе электронных состояний в акцепторе нужно иметь в виду, что в непосредственном окружении около атома все валентные связи заполнены, а вакантное состояние ("дырка") находится, как и пятый электрон в доноре, довольно далеко от центра (рис. 1.28). Такое локальное смещение электронной плотности приводит к возникновению отрицательного электрического заряда в окрестности атома примеси и вызывает нарушение периодического потенциала решетки, как это изображено на рис. 1.28, б. Это возмущение можно описать кулоновским распределением потенциала, используя понятие диэлектрической проницаемости, но потенциальная энергия дырки в поле отрицательного заряда отличается от (1.35) только знаком и началом отсчета. В результате мы приходим опять к модели атома водорода с той разницей, что поле "ядра" направлено противоположно. Волновые функции и соответствующий им спектр энергии электрона в такой системе могут быть найдены, как и в предыдущем случае, при помощи уравнения Шредингера для модели водородоподобного атома в приближении эффективной массы. Учитывая, что последняя у потолка валентной зоны является отрицательной, и меняя знаки, легко получить уравнение Шредингера для дырки, которое не будет отличаться от электронного в задаче с донорным атомом. Естественно, что и энергетический спектр дырки в поле акцепторного атома не будет отличаться (кроме обозначений) от изображенного на рис. 1.27, в. Электронную энергетическую диаграмму можно получить, лишь изменив направление отсчета энергии.

В итоге можно сделать вывод о том, что вакантные состояния недостающего в акцепторном атоме электрона имеют водородоподобный спектр, состоящий из основного состояния Е, и возбужденных E_2 , E_3 и т.д. (рис. 1.28, в), причем энергия создания отрицательного акцепторного иона ΔE_{a} = = $E_a - E_v \approx 0,01$ эВ. Значения энергии ионизации акцепторных атомов в германии и кремнии приведены в табл. 1.2.



Рис. 1.28. Трехвалентный атом примеси в четырехвалентном полупроводнике "притягивает" валентный электрон для заполнения ближайших связей, образуя отрицательный "ион". Незаполненная связь или "дырка" находятся на расстояниях около десятка периодов решетки. Диаграммы (б) и (ϵ) показывают искажение поля решетки "ионом" и возможные энергетические состояния "дырки" в области локального возмущения поля.

О пространственной локализации этой вакансии можно судить по радиусу орбиты, который, как и в донорном атоме, во много раз больше периода кристаллической решетки и со1. Электронные состояния в кристаллах

ставляет приблизительно несколько десятков ангстрем. Но нужно заметить, что в акцепторных полупроводниках, в отличие от донорных, уровни возбуждения в опытах по поглощению света не наблюдаются.

1.2.3. Многозарядные примеси, глубокие уровни

Модель водородоподобного атома удовлетворительно описывает электронные состояния примесей III и V групп в германии и кремнии. Это видно из табл. 1.2 по тому, что результаты измерений энергии ионизации, во-первых, слабо зависят от вида примеси и, во-вторых, неплохо соответствуют вычисленным значениям: ~0,01 эВ для германия (при $m^* \approx 0,2m$ и $\varepsilon = 15,8$) и ~0,05 эВ для кремния (при $m^* \approx 0,5m$ и $\varepsilon = 11,8$). Такие примесные уровни, которые расположены вблизи края соответствующих зон называют *мелкими*.

Примеси, для которых энергия ионизации является соизмеримой с шириной запрещенной зоны полупроводника, образуют *глубокие примесные уровни*. Для анализа глубоких энергетических состояний рассмотренная модель непригодна, так как у них радиус орбиты электронов значительно меньший, чем для мелких примесей, и такие понятия, как диэлектрическая проницаемость и эффективная масса, уже не могут использоваться. Даже в том случае, если радиус орбиты все же в несколько раз больше периода решетки, тяжело связать глубокий уровень со свойствами электронов в одной из разрешенных зон и использовать ту или иную эффективную массу. Несмотря на это, и для глубоких примесных состояний можно установить некоторые простые закономерности.

Глубокие уровни образуются примесными атомами, в которых количество валентных электронов отличается от четырех больше, чем на единицу. Будет атом донором или акцептором зависит от того, избыточно или недостаточно у него электронов для установления валентных связей с атомами основной решетки. Однако надо иметь в виду, что в этом случае примесные атомы могут отдавать или принимать больше одного электрона и становиться одно-, дву- или трехкратно заряженными ионами. Такие примеси иногда называют *многозарядными*. Атомы I группы для заполнения валентных связей могут принимать до трех электронов и образовывать в запрещенной зоне три акцепторных уровня (рис. 1.29) на расстояниях от валент-



Рис. 1.29. Положения энергетических уровней, образованных различными примесями в германии. Указанные значения энергий отсчитываются от края ближайшей разрешенной зоны.

ной зоны, соответствующих разной степени ионизации примесных атомов. У примесных атомов II группы незаполненными являются две валентные связи, поэтому они образуют два акцепторных уровня. Атомы VI группы имеют, кроме участвующих в валентных связях, два лишних электрона и являются донорами. Естественно, что для повторной ионизации атома нужна большая энергия и соответствующий донорный уровень лежит глубже.

Приведенные рассуждения являются далеко не полными и не могут описать всего разнообразия свойств всех примесных атомов. Например, атом золота в германии, кроме трех акцепторных образует еще и донорный уровень, который лежит приблизительно на 0,05 эВ выше потолка валентной зоны. Это значит, что единственный валентный электрон атома золота не на равных с другими участвует в образовании валентных связей и для перевода его в зону проводимости нужна меньшая энергия, чем для валентных электронов полупроводника. Такие примеси, образующие и донорные, и акцепторные уровни называются *амфотерными*. Рассмотрим еще один пример "нетипичного" поведения примесного атома. Вместо трех ожидаемых, как у меди, глубоких акцепторных уровней атом ли50 1. Электронные состояния в кристаллах

тия образует один мелкий донорный уровень. Это можно объяснить тем, что из-за своих малых размеров атом лития располагается не в узле кристаллической решетки как примесь замещения, а внедряется в междоузлие. Поэтому его валентный электрон, не участвуя в образовании валентных связей, оказывается слабо связанным с атомом и подобен в своем поведении лишнему электрону пятивалентной примеси.

Отметим, что теория глубоких примесных состояний, играющих большую роль, например, при рекомбинационных процессах в полупроводниках, еще далека от завершения.

1.2.4. Влияние концентрации примеси на зонную структуру

Энергетические электронные состояния, образованные отдельными примесными атомами, называют локальными, так как электроны в этих состояниях связаны с атомами и локализованы в их окрестности. Количественной мерой величины области локализации может быть радиус орбиты электрона, который по порядку величин равен единицам (для глубоких примесей), десяткам (для мелких примесей) и даже сотням (в некоторых соединениях A_3B_5) периодов кристаллической решетки. Совокупность локальных уровней, которые возникают при *случайном*, но более-менее равномерном распределении по объему полупроводника N примесных атомов, образует примесный уровень (рис. 1.30, *a*), который в зависимости от природы примеси может быть донорным или акцепторным.

Такое положение сохраняется, пока концентрация примесных атомов в полупроводнике незначительна. Увеличение ее приводит к сближению атомов и при некоторой их концентрации наступает положение, когда *примесные атомы начинают взаимодействовать между собой*. Потенциальная энергия лишних электронов (для доноров) уже не может быть описана кулоновским полем точечного заряда, и вместо уравнения (1.35) при расчетах энергетического спектра примеси следует использовать выражение



Рис. 1.30. При малой концентрации примеси совокупность локализованных состояний образут примесный уровень (*a*). Взаимодействие примесных атомов при их больших концентрациях приводит к расщеплению локальных состояний и образованию энергетической зоны (δ).

$$U(\vec{r}) = -\sum_{i}^{N} e^{2} / 4\pi \varepsilon \varepsilon_{0} \left| \vec{r} - \vec{R}_{i} \right|, \qquad (1.39)$$

где суммирование производится по всем примесным атомам, положение которых в решетке определяется радиус-вектором \vec{R}_i . Слагаемые этой суммы характеризуют взаимодействие рассматриваемого электрона с "чужими" атомами примеси. Так как среднее расстояние между примесными атомами порядка $N_{\rm g}^{-1/3}$, то при $N_{\rm g} \approx 10^{17}$ см⁻³ оно будет соизмеримым с радиусами орбит r_m (1.38), а энергия взаимодействия – с энергией ионизации примеси. В этом случае спектр электронных состояний не может быть описан выражением (1.37), а отдельные дискретные состояния будут расщепляться в зоны примесных состояний (рис. 1.30, δ) подобно тому, как это имело место при образовании энергетических зон в кристалле.

На рис. 1.31 изображена только одна примесная зона, которая соответствует основному состоянию E_{π} . Волновые функции электронов в зоне перекрываются и перестают быть локализованными в окрестности отдельных атомов. Резуль-



Рис. 1.31. Изменение энергетического спектра (*a*) и распределение плотности состояний (*б*) при увеличении концентрации примеси. В случае 4 наблюдается вырождение полупроводника. Случай 5 соответствует аморфной структуре, когда в энергетических зонах появляются "хвосты" плотности состояний.

тирующие волновые функции имели бы вид блоховских функций, если бы расположение примесных атомов было бы равномерным и периодическим, как в идеальном кристалле. Естественно, что ширина примесной зоны зависит от степени перекрытия волновых функций и возрастает с ростом концентрации примеси. При $N_{\rm g} \approx 10^{19}$ см⁻³ примесная зона перекрывается с зоной проводимости, образуя общую зону, нижние состояния которой заполнены электронами примесных атомов. это напоминает зонную структуру металлов, а некоторые свойства таких полупроводников, которые называются *вырожденными*, близки к свойствам металлов.

Со случайным характером распределения примеси в кристалле связана одна особенность примесных зон, которая отличает их от энергетических зон идеального кристалла. Последние имеют четкую границу, так как функция распределения плотности энергетических состояний уменьшается до нуля на краю зоны как $(E-E_m)^{1/2}$. Плотность состояний в примесной зоне, которая имеет вид дельта-функции $\delta(E-E_n)$ для дискретных локальных состояний, "размывается" и принимает вид колоколообразной кривой (рис. 1.31, б). Ширина этой кривой увеличивается с ростом концентрации примеси, но обязательно таким образом, что интеграл от функции распределения плотности состояний равен концентрации примеси. В результате энергетические зоны вырожденного полупроводника имеют так называемые "хвосты" плотности состояний. Появление "хвостов" наблюдается и в валентной зоне даже тогда, когда полупроводник донорный. Объясняется это случайным характером распределения примесных атомов, которые заметно нарушают периодичность решетки основного вещества и неизбежно вызывают изменение структуры его энергетических зон, в том числе и валентной.

Увеличение концентрации примеси до такой степени, что ее количество становится соизмеримым с количеством атомов основного вещества, приводит к ситуации, когда, во-первых, само понятие примеси теряет смысл и, во-вторых, возможно образование неупорядоченной, т.е. аморфной, структуры. Так как выводы о зонном характере энергетического спектра получены из условия периодичности кристаллической решетки, то в аморфных твердых телах, казалось бы, должен отсутствовать зонный спектр. Однако эксперименты подтверждают наличие запрещенных и разрешенных зон в некристаллических твердых телах. Ответственным за это является ближний порядок, т.е. такое взаимное расположение и соответственно взаимодействие ближних атомов, которое характерно для кристаллов. В отличие от энергетических зон кристалла зоны в аморфных телах имеют упомянутые хвосты плотности состояний, обусловленные отсутствием дальнего порядка. Особенностью таких хвостов является то, что они в значительной мере состоят из локализованных состояний, а электроны в них имеют незначительную подвижность и их вклад в электропроводность или другие кинетические явления может оказаться незначительным.

1.3. ЭЛЕКТРОННЫЕ СОСТОЯНИЯ НА ПОВЕРХНОСТИ

Энергетические состояния электронов на поверхности твердого тела должны отличаться от объемных состояний. Такой вывод можно сделать из того, что распределение последних описывается зонной теорией, результаты которой основаны на



Рис. 1.32. Нарушение периодичности потенциальной функции у края кристаллической решетки.

предположении о трансляционной симметрии кристаллической решетки и могут использоваться, строго говоря, только для бесконечного, неограниченного кристалла. Для реальных кристаллов конечных размеров условие такой симметрии приблизительно выполняется для внутренней части

объема при сравнительно ограниченных трансляциях, что дает возможность и здесь с успехом применять выводы зонной теории. На поверхности в результате обрыва решетки периодичность потенциала нарушается (рис. 1.32), и вопрос о состояниях электрона должен решаться отдельно.

В связи с этим при рассмотрении поверхности возникают три вопроса: 1) какая форма потенциального барьера на поверхности и какими параметрами он характеризуется; 2) какие особенности энергетического спектра электронов в приповерхностном слое; 3) какая роль посторонних атомов, которые часто в заметном количестве адсорбированы на поверхности.

1.3.1 Макроскопическая энергетическая модель поверхности

Потенциальная функция электрона вблизи поверхности твердого тела легко может быть определена при макроскопическом рассмотрении, когда поверхность можно считать гладкой. При выходе электрона за пределы твердого тела на него действует сила со стороны нескомпенсированного положительного заряда (рис. 1.33, *a*). Вычисление силы взаимодействия может быть проведено методом зеркального отображения, суть которого заключается в том, что распределенный по поверхности заряд заменяется точечным положительным зарядом +e, находящимся по другую сторону от поверхности на таком же расстоянии, что и электрон (рис. 1.33, *б*). Сила, действующая на электрон, в соответствии с законом Кулона равняется



Рис. 1.33. К расчету формы потенциального барьера у поверхности металла (a). Энергия взаимодействия электрона с поверхностью (a) определяется как энергия его взаимодействия с зеркально отображенным относительно поверхности зарядом + e (b).

$$F(x) = -e^2/4\pi\varepsilon_0(2x)^2 = -e^2/16\pi\varepsilon_0 x^2 . \qquad (1.40)$$

Потенциальная энергия электрона равна работе по его перемещению из данной точки на бесконечность и определяется как интеграл от силы F(x):

$$U(x) = -e^2/16\pi\varepsilon_0 x.$$
 (1.41)

Полученная потенциальная энергия равна 0 при $x \to \infty$, что соответствует минимально возможной энергии свободного электрона в вакууме, и называется *уровнем вакуума*. Функция (1.41) описывает форму потенциального барьера, существующего вблизи поверхности и препятствующего выходу электронов из твердого тела. Работа, совершаемая электроном по преодолению сил этого потенциального барьера, называется *работой выхода*.

Введенное таким образом понятие работы выхода не является определенным, так как не указывает на начальный уровень отсчета этой величины, и поэтому каждый электрон будет иметь свое "индивидуальное" значение работы выхода. Если отсчитывать ее от минимально возможной энергии свободного электрона в металле, т.е. от дна зоны проводимости E_c , то получим так называемую *полную работу выхода* $U_0 = E_0 - E_c$ (рис. 1.34). Однако в металле электроны даже при T = 0 имеют некоторую энергию, максимальное значение которой определяется уровнем Ферми (см. разд. 3.3). Для преодоления барьера им нужно выполнить меньшую работу, и ее значение, отсчитанное от уровня Ферми, называют *термодинамической* (иногда эффективной) работой выхода $\Phi = E_0 - E_{\phi}$. Естественно, что $U_0 > \Phi$ на величину энергии Ферми.



Рис. 1.34. Обозначение полной U_0 и термодинамической Φ работ выхода, а также энергии электронного сродства χ на энергетических диаграммах металла (*a*) и полупроводника (*б*).

Для полупроводников понятие термодинамической работы выхода сохраняет свой смысл, несмотря на то, что уровень Ферми может находиться в зоне запрещенных состояний (рис. 1.34, *б*) и ни один из электронов при выходе не совершает такой работы.

Энергию, которая необходима для перевода электронов со дна зоны проводимости полупроводника на уровень вакуума, называют энергией электронного сродства $\chi = E_0 - E_c$.

Выражением (1.41) можно пользоваться для описания только верхней части потенциального барьера на расстоянии от поверхности больше нескольких периодов кристаллической решетки, когда поле решетки становится гладким и возможно представление о поверхности как о плоскости. На близком расстоянии получение выражения для потенциальной функции, а значит и сопряжения уровня вакуума с каким-либо характерным энергетическим уровнем внутри кристалла, т.е. вычисление работы выхода, является сложной задачей. Поэтому обычно пользуются экспериментальными значениями работы выхода.

1.3.2. Таммовские поверхностные состояния

Проблема нахождения формы потенциального рельефа вблизи поверхности возникает и при анализе энергетического спектра поверхностных состояний электрона. Несмотря на сложность количественных вычислений, общие рассуждения здесь достаточно несложные и полезные. Для электрона с энергией, выше уровня вакуума, блоховская волновая функция для внутренней области кристалла должна постепенно переходить в бегущую волну, которая описывает свободный электрон в вакууме. Если энергия электрона меньше E_0 и соответствует зоне разрешенных состояний, то его волновая функция должна иметь вид экспоненты, спадающей вглубь потенциального барьера в вакууме, а вблизи поверхности "сшивается" с волной Блоха для объемных состояний.

Особый интерес представляют поверхностные электронные состояния (рис. 1.35, *a*), которые попадают в запрещенную зону объемных. Со стороны вакуума волновая функция, как и в предыдущем случае, имеет вид экспоненты, а со стороны кристалла подобна бло-



Рис. 1.35. Условное изображение зон поверхностных состояний на энергетической диаграмме (a) и примерный вид соответствующей волновой функции (δ).

ховской функции, амплитуда которой экспоненциально затухает на глубине в несколько периодов решетки (рис. 1.35, б). Важной особенностью этих волновых функций является то, что благодаря существованию трансляционной симметрии в направлениях, параллельных поверхности кристалла, они имеют в соответствии с теоремой Блоха вид бегущих волн $\psi_{s}(x, y, z) =$ $= u(x, y, z) \exp[i(k_{u}y + k_{z}z)],$ причем амплитуда u(x, y, z) является периодической с периодом решетки в направлениях и и z. Это обстоятельство порождает, как и в трехмерном случае целую зону (или зоны) поверхностных состояний (рис. 1.35, а). Но в отличие от объемных эти зоны локализованы вблизи поверхности и постепенно сливаются с объемными зонами по мере уменьшения поверхностного возмущения и превращения потенциальной функции в периодическую. Таким образом, поверхностные состояния представляют собой не дополнительные, а модифицированные объемные состояния. Их количество соот58 1. Электронные состояния в кристаллах

ветствует количеству атомов, граничащих с поверхностью, и по порядку величин составляет $10^{14}...10^{15}$ см⁻². Поверхностные состояния, связанные с обрывом решетки, были впервые исследованы И. Е. Таммом и называются *таммовскими*. Они всегда присутствуют на поверхности и играют существенную роль в электронных процессах.

1.3.3. Адсорбированные поверхностные состояния

Состояния другого вида возникают при наличии на поверхности *адсорбированных* атомов. Поведение этих атомов, их влияние на состояние поверхности зависит от сорта атомов и может быть рассмотрено на такой несложной модели. Пусть на поверхности металла находится атом (рис. 1.36, *a*). Между кристаллической решеткой и атомом существует потенциальный



Рис. 1.36. Изменение работы выхода при адсорбции на поверхности металла электроположительных (*a*) и электроотрицательных (*б*) атомов.

барьер, сравнимый по прозрачности с межатомными барьерами кристалла. Если энергия ионизации атома меньше термодинамической работы выхода из металла, т.е. валентный уровень атома находится выше уровня Ферми в металле, то возможен переход валентного электрона в металл. В результате перехода атом становится положительно заряженным ионом, а возникающее электростатическое поле создает силы, удерживающие атом на поверхности (силы *адсорбции*). Другим не менее важным следствием этого электронного обмена является уменьшение работы выхода электронов из металла. Соответствующее изменение формы потенциального барьера показано на рис. 1.36, а для сечений, проходящих через центр атома (кривая 1) и вне него (кривая 2). Снижение уровня вакуума, естественно, заметно при достаточно большом количестве адсорбированных атомов. Такие атомы (обычно атомы щелочных или щелочноземельных металлов) по отношению к поверхности являются электроположительными.



Рис. 1.37. Зависимость эффективной работы выхода вольфрама от степени покрытия поверхности адсорбированными атомами

Интерес представляет зависимость (рис. 1.37) работы выхода от степени покрытия катода пленкой. Естественно, что при увеличении количества адсорбированных атомов работа выхода должна меняться от значения для чистой поверхности (4,5 эВ для вольфрама) до значения работы выхода вещества пленки (3,4 эВ для тория, 2,5 эВ для бария и 1,9 эВ для цезия). Однако эта зависимость не будет монотонной, что можно было ожидать, если бы атомы оставались нейтральными. Ионизация атомов и образование ускоряющего электроны поля снижает работу выхода до значений меньших, чем работа выхода покрывающего вещества. Наибольший эффект достигается, когда атомы покрывают 60...70% поверхности катода. При большей степени покрытия заметным становится взаимодействие адсорбированных атомов между собой, приводящее к уменьшению 60 1. Электронные состояния в кристаллах

вероятности их ионизации и уменьшению напряженности электрического поля.

Противоположная ситуация возникает для электроотрицательных атомов, когда работа выхода из металла меньше энергии электронного сродства адсорбированных атомов. При переходе электрона из металла на расположенный ниже свободный уровень атома образуется отрицательный ион. Электрическое поле, возникающее между ионом и поверхностью, создает, во-первых, силу адсорбции, а во-вторых, дополнительную силу, которая препятствует выходу электронов из металла. На энергетической диаграмме это выглядит как повышение потенциального барьера и увеличение работы выхода. Типичными электроотрицательными атомами являются, например, атомы кислорода, азота, хлора и др.

Возможной является также ситуация, когда работа выхода из металла меньше энергии ионизации атома, но больше его электронного сродства. Ион в этом случае не образуется, но возникающие за счет перекрытия волновых функций обменные силы (типа сил Ван-дер-Ваальса), приводят к взаимодействию металла и атома и к поляризации последнего. Наличие у адсорбированного атома дипольного электрического момента также деформирует потенциальный барьер и влияет на работу выхода.

Механизм взаимодействия адсорбированных атомов с поверхностью полупроводника принципиально такой же, как и для металлов. Но поверхностные состояния, связанные с адсорбированными атомами, могут играть также роль доноров и акцепторов, центров рекомбинации и т.д., и при достаточной концентрации ощутимо влиять на электрические свойства поверхности. На рис. 1.38, *а* изображен пример ситуации, когда энергия электронного сродства полупроводника меньше энергии ионизации адсорбированного атома. В этом случае обмен валентным электроном происходит только под действием какого-либо внешнего возбуждения, например, тепла. Такие атомы являются поверхностными донорами и при их ионизации на поверхности образуется электрическое поле, меняющее (как



Рис. 1.38. Энергетическая диаграмма полупроводника с адсорбируемым на поверхности "донорным" атомом (*a*). Изгиб зон у поверхности полупроводника при адсорбции электроположительных атомов уменьшает эффективную энергию электронного сродства (б), которая при определенных условиях может быть отрицательной (*в*).

и для металлов) высоту и форму потенциального барьера. Однако для полупроводников по сравнению с металлами наблюдаются две существенные особенности. Во-первых, вероятность ионизации поверхностных атомов, а значит и сила электрического поля зависит от вида и концентрации примеси в объеме. Например, наличие акцепторов (рис. 1.38, б) образует у нижнего края запрещенной зоны свободные состояния, которые могут заполняться валентными электронами адсорбированных атомов даже без внешнего возбуждения. Во-вторых, поле двойного электрического слоя не заканчивается на поверхности, как в металлах, а проникает в приповерхностный слой на некоторое расстояние. Потенциальная энергия этого поля на энергетических диаграммах искривляет зоны вблизи поверхности вниз для указанного направления поля. Уменьшение расстояния между уровнем вакуума и дном зоны проводимости в результате изгиба означает уменьшение эффективной энергии электронного сродства.

Этот эффект используется для создания эмиттеров с отрицательным электронным сродством (рис. 1.38, в). В таких катодах на слой сильнолегированного акцепторами материала p^+ -GaAs нанесена моноатомная пленка цезия, которая частично 62 1. Электронные состояния в кристаллах

оксидирована. Крутой изгиб зон возникает за счет сильного поля между слоем отрицательно ионизированных акцепторов и слоем положительных ионов цезия. Молекулы окиси цезия представляют собой диполи, электрическое поле которых формирует резкий спад потенциального барьера, толщина которого составляет несколько ангстрем. В результате уровень вакуума будет располагаться ниже дна зоны проводимости в глубине полупроводника, что и означает наличие отрицательного электронного сродства. Такие фотокатоды имеют повышенную чувствительность в инфракрасном диапазоне и используются в приборах ночного видения.

2. ЭЛЕКТРОННЫЕ СОСТОЯНИЯ И ПРОВОДИМОСТЬ ТВЕРДЫХ ТЕЛ

2.1. СВЯЗЬ ЗОННОЙ СТРУКТУРЫ И ПРОВОДИМОСТИ

Классификацию твердых тел с точки зрения электрических свойств проводят по различным признакам. Простейшим является деление веществ по значениям их удельной электропроводности. Вещества, которые имеют $\sigma > 10^6$ См/м, принято считать проводниками, $\sigma < 10^{-8}$ См/м диэлектриками, а промежуточные значения свойственны полупроводникам. Понятно, что такая классификация твердых тел по величине их электропроводности, меняющейся к тому же более, чем на 20 порядков, не может быть универсальной, ибо не учитывает особенностей разнообразных механизмов проводимости, существующих в твердых телах.

Другой распространенный способ классификации основан на температурной зависимости электропроводности. Известно, что электропроводность металлов с ростом температуры падает приблизительно пропорционально T^{-1} , а полупроводников – экспоненциально возрастает и в зависимости от внешних условий, как-то: температура, освещенность, давление и др., может меняться в широких пределах. Известно также, что наличие примесей в металлах обычно уменьшает их электропроводность, а в полупроводниках увеличивает.

Все эти факты хорошо объясняются зонной теорией, а энергетическая зонная структура твердых тел может быть основой их классификации по электрическим свойствам. Все разнообразие зонных структур кристаллов, которое возникает вследствие перекрывания их энергетических зон, заполненных электронами или свободных, зон для различных квантовых состояний и различных кристаллографических направлений, может быть сведено к двум видам, которые отличаются взаимным расположением зоны, заполненной валентными электронами, и ближайшей зоны свободных состояний (рис. 2.1). Энергетические зоны, которые образованы из более глубоких атомных уровней, существенно на электрические свойства не влияют и могут не приниматься во внимание.



Рис. 2.1. Диаграмма с изображением деления веществ по значениям их удельной электропроводности и его связь со структурой энергетических зон.

Вещества, у которых заполненные энергетические состояния непосредственно смыкаются с зоной свободных состояний, являются проводящими. Если зоны свободных и заполненных состояний, которые называются соответственно зоной *проводимости* и *валентной*, разделены *запрещенной* зоной, то такое вещество в нормальном состоянии является непроводящим и относится к диэлектрикам или полупроводникам. Различие между ними является количественным и не может быть четким. Если ширина запрещенной зоны, т.е. энергетическое расстояние между нижним краем зоны проводимости (дном) E_c и верхним краем валентной зоны (потолком) E_v превышает 3 эВ, такое вещество является типичным диэлектриком. Типичными полупроводниками считают вещества с запрещенной зоной меньшей, чем 2 эВ.

В чем же причина такого отличия электрических свойств кристаллов и как это связано с их зонной структурой? На первый взгляд валентные электроны, являющиеся обобществленными, т.е. не связанными с отдельными атомами, и такими, которые описываются волновыми функциями в виде волн Блоха, должны свободно передвигаться вдоль кристалла под действием электрического поля и обеспечивать проводимость. Дело в том, что поле ускоряет электроны и участие их в проводимости обязательно связано с изменением их энергетических состояний. Если же все состояния в зоне заполнены, то электрическое поле не может переводить электроны в другие состояния, и направленного движения при этом не возникает. Исключение составляют очень сильные электрические поля, под действием которых электрон может быть переведен в зону свободных состояний. Но эти случаи рассматриваются отдельно, а обычные поля с напряженностью даже в ~1 кВ/см на длине свободного пробега ~1 мкм могут сообщить электрону энергию только 10⁻¹ эВ, что конечно недостаточно для его перехода в зону проводимости. В проводящих веществах свободные состояния размещены энергетически настолько близко, что электроны будут ускоряться под действием даже очень слабого электрического поля.

Из приведенных рассуждений следует, что только электроны, которые размещены в частично заполненных зонах, могут принимать участие в проводимости. Эта ситуация реализуется в металлах даже при нулевой температуре. Полупроводники и диэлектрики могут проводить ток, если часть электронов перейдет из полностью заполненной валентной зоны в зону проводимости под действием, например, нагревания, освещения, сильного электрического поля или какого-то другого внешнего фактора.

Такое объяснение связи электрических свойств с особенностями заполнения энергетических зон, проведенное с помощью только "одномерных" энергетических диаграмм, нельзя считать достаточно аргументированным и полным: может сложиться представление, что на обобществленные электроны полностью заполненной зоны электрическое поле не действует. Для более глубокого понимания процесса проводимости необходимо, во-первых, выяснить влияние электрического поля на изменение не только энергетических состояний, но и значений волнового вектора, а рассмотрение проводить с использованием дисперсионной кривой. Во-вторых, проводимость нужно рассматривать как коллективное явление, обусловленное поведением всех электронов зоны.

2.2. ВИДЫ ПРОВОДИМОСТИ И НОСИТЕЛИ ЗАРЯДА

Действие внешнего электрического поля на движение электронов в кристалле может быть исследовано при помощи уравнения (1.16):

$$\hbar d\vec{k} / dt = \vec{F}. \tag{2.1}$$

Если направить ось *x* навстречу силовым линиям электрического поля, то вектор силы, действующей на электрон $(F = F_x = e\epsilon)$, будет совпадать с положительным направлением *x*-составляющей волнового вектора. Это дает возможность ограничиться одномерным уравнением движения $dk_x/dt = = e\epsilon/\hbar$, результат интегрирования которого имеет вид

$$k_x(t) = k_x(0) + (e/\hbar) \varepsilon t$$
 (2.2)

и показывает, что в статическом электрическом поле волновой вектор со временем линейно возрастает. В действительности этот процесс ограничивается столкновениями электрона с дефектами кристалла, а (2.2) выполняется только на промежутке между двумя столкновениями. Но для выяснения реакции электрона на внешнее поле столкновения пока не будем принимать во внимание.

Обращаясь к дисперсионной кривой (рис. 2.2), видим, что электрическое поле увеличивает волновой вектор и смещает

2.2. Виды проводимости и носители заряда

электрон, находившийся в состоянии *k*_i, в более высокие энергетические состояния. Двигаясь таким образом по дисперсионной кривой, электрон доходит до края первой зоны Бриллюэна (точка 1) и переходит во вторую, где его энергия уменьшается. Удобно воспользоваться схемой приведенных зон (рис. 1.11, в) и ограничиться только первой зоной Бриллюэна. Учитывая, что точки 1 и 1' физически тождественны, можно считать, что электрон, достигая точки 1, переходит в точку 1' и опускается вдоль левой ветви дисперсионной кривой. Оставаясь таким образом в пер-



Рис. 2.2. Внешнее электрическое поле увеличивает значения волнового вектора электронов и они смещаются по состояниям в соответствии с законом дисперсии. Однако при полном заполнении всех состояний зоны электронами их смещения не нарушают симметрии заполнения состояний.

вой зоне Бриллюэна, электрон совершает циклическое перемещение по состояниям дисперсионной кривой.

Из приведенного описания становится понятным, почему электроны полностью заполненной зоны не создают электрического тока. Если все состояния в зоне, изображенные точками на рис. 2.2, заполнены электронами и каждый из них реагирует на электрическое поле, как это описано выше, то постоянно все состояния остаются заполненными. Из-за симметрии дисперсионной кривой каждому электрону в состоянии $\vec{k_i}$ соответствует электрон в состоянии – $\vec{k_i}$, поэтому векторная сумма по всем N состояния зоны $\sum_{i=1}^{N} \vec{k_i} = 0$, и преобладающего дви-

жения электронов в каком-либо направлении не происходит. Именно поэтому проводимость, если она существует, создается электронами *только частично заполненных зон*.

При частичном заполнении электронами состояний в зонах возможны два типичных случая. Первый из них имеет место в зоне проводимости, где число заполненных состояний



Рис. 2.3. Электроны у дна частично заполненной зоны под действием электрического поля смещаются по состояниям так, что их суммарный волновой вектор или суммарная скорость направлены в сторону действия силы электрического поля, как и для свободных частиц.

п обычно намного меньше, чем свободных. В таком случае электроны заполняют нижние состояния симметрично относительно минимума дисперсионной кривой (рис. 2.3, *a*). Электрическое поле, направленное навстречу оси k_x , меняет волновой вектор электронов в соответствии с (2.2) и они смещаются в *k*-*npocm*ранстве за какое-то время в заштрихованную на рис. 2.3, *б* область. Очевидно, что результирующий волновой вектор всех n электронов $\sum_{i}^{n} \vec{k_i}$ будет отличным от нуля, что означает наличие направленного переноса заряда. Величина электрического тока состоит из элементарных токов, которые создаются каждым из n электронов зоны проводимости

$$\vec{j} = (-e) \sum_{i}^{n} \vec{v}_{i}.$$
 (2.3)

Электроны, расположенные в нижних состояниях зоны, можно рассматривать как свободные частицы с положительной

эффективной массой, поэтому в пределах параболичности закона дисперсии $m^* \vec{v} = \hbar \vec{k}$ и

$$\vec{v}_i = \hbar \vec{k}_i / m^*, \qquad (2.4)$$

где *m*^{*} имеет *одинаковое* значение для всех заполненных состояний. Учитывая это, для полного тока имеем:

$$\vec{j} = (-e\hbar/m^*)\sum_{i}^{n}\vec{k}_i. \qquad (2.5)$$

Поскольку суммарный волновой вектор положителен (рис. 2.3, δ) и противоположен вектору $\vec{\epsilon}$, то электрический ток будет, как обычно, совпадать по направлению с электрическим поля.



Рис. 2.4. Смещение под действием электрического поля электронов почти полностью заполненной зоны происходит так, что их суммарный волновой вектор направлен навстречу силе поля.

Несколько иная ситуация наблюдается, например, в валентной зоне, где количество заполненных состояний (n) больше, чем свободных (p), причем n + p = N – полному числу состояний в зоне. В этом случае электроны по-прежнему заполняют нижние энергетические состояния, так что свободные от электронов состояния будут размещены симметрично отно-

сительно максимума дисперсионной кривой (рис. 2.4, *a*). Электрическое поле действует на каждый электрон в соответствии с (2.2) и смещает их, как и в предыдущем случае, в сторону больших значений *k*. Но коллективный результат получается совсем другим. Из формы заштрихованной области на рис.2.4, *б*, обозначающей заполненные состояния, даже без детальных расчетов видно, что суммарный волновой вектор всех электронов $\Sigma_i^n \vec{k}_i$ направлен вдоль вектора $\vec{\varepsilon}$, то есть противоположно смещению каждого из электронов! Кроме этой "странности" поведения электронов также возникают трудности при расчете тока по формуле (2.5), которую нужно в этом случае записать в виде:

$$\vec{j} = -e\hbar \sum_{i}^{n} \vec{k}_{i} / m_{i}^{*}$$
 (2.5, a)

Понятие эффективной массы удобно использовать лишь у края энергетической зоны, где дисперсионная кривая параболична, а в формуле (2.5, *a*) суммирование надо производить почти по всем состояниям зоны, где m_i^* имеет *разные* значения и метод эффективной массы использовать не имеет смысла. Избежать этой трудности можно, рассматривая поведение не электронов, а свободных состояний, или *дырок*. Формально это осуществляется таким образом. Поскольку $\Sigma^N \vec{v}_i = \Sigma^n \vec{v}_i + \Sigma^p \vec{v}_i = 0$, то $\Sigma^n \vec{v}_i = -\Sigma^p \vec{v}_i$, и то же самое значение тока можно найти не из формулы (2.3), а через сумму по свободным (дырочным) состояниям:

$$\vec{j} = e \sum_{i}^{p} \vec{v}_{i}.$$
(2.6)

Отсюда видно, что рассматривая дырку как некоторую условную частицу, мы должны приписать ей положительной заряд +e. Кроме этого, электрон около верхнего края зоны имеет отрицательную эффективную массу и поэтому, аналогично (2.4), имеем

$$\vec{v}_i = \hbar \vec{k}_i / (-|m^*|),$$
 (2.7)

где m^* – одинакова для всех свободных состояний. После подстановки (2.7) в (2.6) видим, что полный ток, создаваемый электронами почти заполненной зоны, можно рассчитать как ток некоторых положительных частиц (дырок)

$$\vec{j} = \left(-e\hbar / \left|m^*\right|\right) \sum_{i}^{p} \vec{k}_i.$$
(2.8)

Но для этого еще нужно дыркам присвоить положительную эффективную массу, равную по величине эффективной массе электрона, который "оставил" эту дырку. Направление тока (2.8) совпадает с направлением электрического поля (как будто он создается положительными частицами), так как суммарный волновой вектор свободных состояний соответствует центру незаштрихованной области на рис. 2.4, б и $\Sigma_i^p \vec{k}_i > 0$.

Таким образом проводимость, создаваемая электронами почти пустой и почти заполненной зон, имеет существенно различные механизмы. Поэтому для описания этих процессов удобно рассматривать два вида квазичастиц, или два вида подвижных носителей заряда: подвижные электроны и подвижные дырки.

Иногда особенности движения дырок в реальном пространстве ошибочно переносят на k-пространство. В реальном координатном пространстве электрон, двигаясь против силовых линий электрического поля, смещается на свободное место в валентных связях и оставляет после себя опять свободное место, которое может быть занято другим электроном и так далее. Это движение электронов напоминает "эстафетное", и при этом свободные состояния (дырки) будут двигаться в направлении силовых линий поля. В k-пространстве дырки под действием поля смещаются *в том же направлении*, что и электроны, как это видно из рис. 2.4, *б* для незаштрихованной области.

Нужно помнить, что дырки являются фиктивными частицами. Перенос заряда осуществляется в обоих случаях электронами, но тогда, когда состояния в энергетической зоне почти все заполнены электронами, *удобнее* (хотя и не обязательно)
электронные процессы рассматривать, используя представление о дырках. Это удобство связано с несколько необычным поведением электронов в почти заполненной зоне, а использование дырок приводит механизм проводимости к привычным представлениям. Наличие двух механизмов проводимости и соответствующих носителей заряда подтверждается экспериментально, например, различными знаками электродвижущей силы, вызываемой различными носителями заряда, при наблюдении эффекта Холла.

2.3. ПЕРЕНОС НОСИТЕЛЕЙ ЗАРЯДА

2.3.1. Дрейф носителей

Под действием электрического поля носители заряда двигаются ускоренно и поэтому в соответствии с (2.2) и (2.5) в идеальном кристалле ток должен со временем возрастать линейно (при постоянной концентрации носителей). Но в действительности после включения поля электрический ток довольно быстро устанавливается постоянным. Причина этого явления объясняется столкновениями электронов и дырок с дефектами кристаллической решетки, которые хаотизируют движение. Если электрическое поле отсутствует (рис. 2.5, а), электроны между столкновениями двигаются прямолинейно (верхний рисунок), а их значение скорости или волнового вектора постоянно (нижний). В момент столкновения направление и скорость их движения случайно меняются и точки на нижнем графике, обозначающие волновой вектор электронов, "перескакивают" в произвольные места. Случайный характер столкновений приводит к тому, что в равновесном состоянии скорости электронов направлены с равной вероятностью во все стороны, поэтому суммарный волновой вектор и соответственно средняя скорость электронов будут равняться нулю, а какое-то преобладающее направление их движения будет отсутствовать.



Рис. 2.5. Влияние столкновений на приблизительный вид траекторий электрона и соответствующие изменения его волнового вектора (скорости) при отсутствии электрического поля (*a*) и в электрическом поле при "таких же" соударениях (б). Числа соответствуют номерам свободных пробегов.

Наличие электрического поля (рис. 2.5, *б*) существенно меняет ситуацию. Как и раньше, направление скорости носителей непосредственно после каждого столкновения будет произвольным, но во время свободного пробега между двумя последовательными столкновениями они будут испытывать действие электрического поля и их траектории становятся параболическими. Электроны будут ускоряться в направлении силы, во время свободного пробега их скорость в *x*-направлении равномерно будет возрастать и точки на нижнем рисунке будут постепенно смещаться в сторону больших значений $k_x \sim v_x$. Во время столкновений они опять принимают случайные значения в *k*-пространстве. Такое поведение носителей означает, что на их хаотическое движение будет накладываться переносное, или *дрейфовое* движение, являющееся результатом действия электрического поля.

Для количественного анализа дрейфа воспользуемся простейшей классической моделью, которая основывается на "индивидуальном" рассмотрении поведения отдельной частицы. Рассмотрим электрон, который, двигаясь хаотически, за время свободного пробега τ между двумя последовательными столкновениями в электрическом поле с напряженностью ε приобретает направленный импульс $e\varepsilon\tau$. Этот направленный против поля импульс соответствует направленной скорости v, которая равняется $e\varepsilon\tau/m$. Усредняя ее во времени по всей траектории наблюдаемой частицы, получим выражение для дрейфовой скорости:

$$v_{\rm gp} = \frac{e\overline{\tau}}{m} \mathcal{E} = \frac{e}{m} \frac{\overline{l}}{\overline{v}} \mathcal{E}.$$
 (2.9)

Использование соотношения $\overline{\tau} = \overline{l} / \overline{v}$, где \overline{l} – средняя длина свободного пробега, а \overline{v} – средняя скорость движения между столкновениями, строго говоря, неверно, правильно $\overline{\tau} = (l / v)$, то есть среднее значение отношения двух величин не равно отношению их средних значений. Но для получения основных закономерностей эта погрешность не является существенной.

Важным следствием (2.9) есть факт пропорциональности между дрейфовой скоростью и напряженностью электрического поля. Коэффициент пропорциональности

$$\mu = e\overline{\tau} / m = e\overline{l} / m\overline{\upsilon} \tag{2.10}$$

называется *подвижностью* и имеет смысл дрейфовой скорости в поле единичной напряженности. Размерность подвижности – $m^2/B\cdot c$, хотя в справочниках ее значения чаще приводят в с $m^2/B\cdot c$.

Результатом направленного дрейфового движения электронов является дрейфовый ток, плотность которого пропорциональна направленному потоку $nv_{\rm дp}$ и их заряду

$$j_{\rm ap} = env_{\rm ap} = en\mu \mathcal{E}. \tag{2.11}$$

Данное выражение является законом Ома в дифференциальной форме, а коэффициент пропорциональности между $j_{\rm дp}$ и \mathcal{E} называется удельной электропроводностью

$$\sigma = en\mu = e^2 n\overline{\tau} / m \approx e^2 n\overline{l} / m\overline{\upsilon}. \qquad (2.12)$$

Для полупроводников нужно иметь в виду, во-первых, наличие различных носителей заряда. Полный ток состоит из вкладов электронной и дырочной составляющих, поэтому

$$\sigma = \sigma_n + \sigma_p = en\mu_n + ep\mu_p. \tag{2.13}$$

Во-вторых, используя метод эффективной массы, мы должны заменить в (2.12) обычную массу на эффективную.

В непрямозонных полупроводниках, таких как кремний, германий, электроны занимают состояния нескольких долин, как это показано на рис. 2.6 для кремния (две долины, которые лежат в направлении $k_{[010]}$ и изображенные на рис. 1.23, не попали в плоскость рисунка). Вклад электронов разных долин в электропроводность не одинаков, поскольку их эффективная масса в направлении дрейфа различна и соответствует для двух долин *продольной* (m_1^*) , а для четырех – поперечной (m_{t}^{*}) эффективной массе. Однако в целом электропроводность кристалла бу-



Рис. 2.6. Заполнение состояний электронами в Х-долинах зоны проводимости кремния. При их смещении по состояниям под действием электрического поля электроны разных долин имеют разную эффективную массу.

дет изотропной, а подвижность электронов будет зависеть от некоторой средней эффективной массы, которая называется *омической* и находится из соотношения

$$\frac{1}{m_{\rm c}^*} = \frac{1}{3} \left(\frac{2}{m_t^*} + \frac{1}{m_l^*} \right), \tag{2.14}$$

которое вытекает из правила суммирования вкладов электронов разных долин в общую электропроводность:

$$\sigma_n = \sum_{i=1}^6 \sigma_{ni} = \sum_{i=1}^6 e^2 \frac{n\overline{\tau}}{6m_i^*} = e^2 n\overline{\tau} / m_c^*.$$
(2.15)

Похожая ситуация имеет место в валентной зоне: различные ветви дисперсионной кривой вблизи верхнего края валентной зоны соответствуют "легким" и "тяжелым" дыркам. Использование некоторой усредненной величины, называемой омической эффективной массой дырок, дает возможность, как и для электронов, оценивать дырочную электропроводность по простой формуле $\sigma_p = e^2 n \tau_p / m_p^*$, подобной (2.12) для "одночастичной" модели.

Из приведенных при выводе формулы (2.9) рассуждений видно, что значение подвижности определяется средним значением времени свободного пробега между столкновениями. В идеальной кристаллической решетке электронные состояния описываются волнами Блоха, имеющими в среднем постоянную амплитуду вдоль всего кристалла. Это значит, что идеальная решетка не вызывает рассеивания и длина свободного пробега носителей неограничена, а рассеивание вызывается $\partial e \phi e \kappa ma$ *ми* кристалла.

Простейшие представления о влиянии некоторых факторов на время свободного пробега можно получить, используя простое соотношение $\bar{\tau} = \bar{l} / \bar{v}$. Средняя длина свободного пробега обратно пропорциональна концентрации дефектов, или так называемых, *центров рассеивания:* $\bar{l} = 1/NS$, где коэффициент пропорциональности S, имеющий размерность площади, называется эффективным сечением рассеивания.

Отклонения от идеальной периодичности, которые всегда имеют место во всех кристаллах, вызываются тепловыми колебаниями атомов. Амплитуда этих колебаний, то есть и "дополнительная площадь" поперечного сечения атомов, возрастает пропорционально температуре решетки *T*, поэтому длина свободного пробега $\bar{l} \sim T^{-1}$. Если (в невырожденных коллективах) хаотическая скорость определяется тепловым движением носителей, то $\bar{\upsilon} = \sqrt{3k_{\rm B}T/m^*}$ и

$$\mu_{\rm T} \sim \overline{\tau} \sim \overline{l} / \overline{v} \sim T^{-3/2} \,. \tag{2.16}$$

Для количественной оценки подвижности будем считать, что рассеивание вызывается всеми атомами кристалла, но взаимодействие носителей происходит не с решеткой (это уже учтено в эффективной массе), а только с отклонением ее от состояния равновесия. Если учесть, что амплитуда тепловых колебаний составляет несколько процентов межатомного расстояния, то дополнительная площадь, которая вызывает рассеивание, приблизительно в 10³ раз меньше площади поперечного сечения атома, т.е. $S \sim 10^{-22}$ м². Поскольку концентрация атомов ~ 10^{29} м⁻³, то средняя длина свободного пробега, обратная суммарному поперечному сечению рассеивания *NS*, составляет ~ 10^{-7} м. При значении $\overline{v} = \sqrt{3k_{\rm B}T/m^*} \approx 10^5$ м/с среднее время свободного пробега имеет порядок ~ 10^{-12} с, а подвижность $\mu = e\overline{\tau}/m^* \approx 10^{-1}$ м²/В·с.

При низкой температуре тепловые колебания незначительны и преобладает рассеивание на примесных атомах, которые специально вводят в полупроводник. Поскольку уже при температурах в несколько десятков кельвинов большая часть примесных атомов термически ионизирована, рассеивание электронов вызывается электрическими полями ионов. Для этого вида взаимодействия характерно, что эффективное сечение обратно пропорционально квадрату энергии электрона, и средняя длина свободного пробега $\bar{l} \sim E^2$. Для невырожденных полупроводников ($E \sim k_{\rm B}T$, $\bar{v}_{\rm xaor} \sim T^{1/2}$) подвижность, ограниченная рассеиванием на ионизированных атомах примеси с концентрацией $N_{\rm пр}$

$$\mu_i \sim T^{3/2} / N_{\rm np}. \tag{2.17}$$

Рассеивание на нейтральных примесных атомах нужно учитывать только при очень низких температурах.

При наличии различных механизмов рассеивания суммируются соответствующие эффективные сечения, поэтому общая подвижность определяется выражением

$$1/\mu = 1/\mu_T + 1/\mu_i = aT^{3/2} + bN_{\rm mp} T^{-3/2}$$
. (2.18)

78 2. Электронные состояния и проводимость твердых тел



Рис. 2.7. Типичная температурная зависимость подвижности электронов, обусловленная рассеянием на тепловых колебаниях кристаллической решетки и ионизированных примесных атомах.

Приблизительный вид этой функции изображен на рис. 2.7, откуда видно, что при низких температурах подвижность определяется рассеиванием на ионизированных примесных атомах и возрастает при повышении температуры тем меньше, чем выше концентрация примеси. При высоких температурах увеличивается роль рассеивания на тепловых колебаниях и подвижность спадает с увеличением температуры. В типичных случаях

область комнатных температур соответствует именно этому участку, когда преобладает рассеивание на тепловых колебаниях, но еще заметен вклад примесного механизма. Более обстоятельно вопросы рассеивания рассмотрены в гл. 5.

Подобные зависимости подвижности от температуры наблюдаются также и для дырок, а количественные отличия связаны в первую очередь с различной эффективной массой электронов и дырок.

2.3.2. Диффузия носителей заряда

Во многих приборах электронной техники используется еще один, диффузионный, способ создания направленного переноса носителей заряда, который проявляется в неоднородных средах. Диффузией называется направленное движение потока частиц, которое возникает при наличии градиента концентрации под действием их хаотического движения. Именно хаотическое движение является "движущей силой" диффузии, поскольку только оно приводит к тому, что встречные потоки частиц (двигающихся с равной вероятностью во все стороны) будут различны при наличии перепада концентрации: преобладающим будет перенос частиц из области с большей концентрацией в область с ее меньшим значением. Диффузия заряженных частиц подчиняется тем же закономерностям, что и диффузия нейтральных, и описывается законами Фика. Первый из них характеризует скорость направленного диффузионного движения \vec{v}_D и имеет вид:

$$\vec{v}_D = -D \frac{1}{n} \operatorname{grad} n$$
, (2.19, *a*)

или в более удобной форме:

$$n\vec{v}_D = -D \text{ grad } n . \qquad (2.19, \, 6)$$

Поскольку левая часть представляет собой плотность потока диффундирующих частиц, т.е. их количество, которое проходит через единицу поперечного сечения за единицу времени, то смысл первого закона диффузии сводится к тому, что плотность потока диффузии пропорциональна градиенту концентрации. Коэффициент пропорциональности называется *коэффициентом диффузии* и в соответствии с кинетической теорией газов равен:

$$D = \bar{l} \cdot v_{\text{xaot}} / 3. \qquad (2.20)$$

Знак "-" в (2.19) показывает, что диффузионный поток направлен в сторону уменьшения концентрации, т.е. противоположно вектору градиента, который направлен в сторону увеличения концентрации.

Нестационарный процесс диффузии описывается вторым законом Фика, который имеет вид:

$$\frac{\partial n}{\partial t} = D\nabla^2 n , \qquad (2.21)$$

и связывает скорость изменения концентрации во времени с ее пространственным распределением $n(\vec{r}, t)$.

Несмотря на то, что диффузия заряженных частиц описывается теми же законами, что и нейтральных, она имеет ряд характерных особенностей. Прежде всего надо отметить, что диффундирующие частицы переносят заряд и создают $\partial u \phi dy$ зионный ток $\vec{j}_D = -en\vec{v}_D$, величина которого пропорциональна потоку частиц и их заряду. При наличии частиц разного вида (например, электронов или дырок в полупроводнике) полный ток имеет различные составляющие, которые с учетом направления (рис. 2.8) выражаются соотношениями:



Рис. 2.8. Направления диффузионных потоков и электрических токов электронов (*a*) и дырок (*б*) относительно направления градиента их концентрации.

$$egin{aligned} &ec{j}_{Dn} = eD_n \, ext{grad} \, n, \ &ec{j}_{Dp} = -eD_p \, ext{grad} \, p. \end{aligned}$$

Вторая особенность диффузии связана с перераспределением зарядов в результате диффузионного переноса. Если нейтрализации диффундирующих зарядов не происходит, то возникает пространственный заряд, который создает тормозящее электрическое поле, или на языке энергетических моделей – потенциальный барьер, сдерживающий процесс диффузии (рис. 2.9). Это явление можно также описать при помощи дрейфового тока, вызванного электрическим полем. Пусть диффузионный поток электронов (рис. 2.9, а) направлен справа налево. В результате диффузионного переноса зарядов слева возникает избыток электронов, а справа - недостаток, и соответственно создается пространственный заряд в виде двойного электрического слоя, поле которого направлено, как указано на рисунке. Сила этого поля препятствует диффузионному переносу, возвращая обратно те из электронов, энергии которых недостаточно для преодоления тормозящего действия поля. Это создает дрейфовый ток, направление которого противоположно диффузионному. Аналогичная ситуация возникает и при



Рис. 2.9. Направления диффузионных и внутренних дрейфовых токов электронов (*a*) и дырок (*б*). Создаваемое диффузионным перераспределением зарядов электрическое поле вызывает дрейфовые токи, направленные навстречу диффузионным.

рассмотрении диффузии дырок (рис. 2.9, б). В результате полный ток определяется суммой дрейфовой и диффузионной составляющих: $\vec{j} = \vec{j}_{\rm дp} + \vec{j}_D$. Выражения для полного тока электронов и дырок соответственно имеют вид:

$$\vec{j}_n = en\mu_n \vec{\varepsilon} + eD_n \operatorname{grad} n$$
, (2.23, a)

$$\vec{j}_p = ep\mu_p\vec{\epsilon} - eD_p \text{ grad } p.$$
 (2.23, 6)

Эти выражения пригодны также и в том случае, если электрическое поле создано не внутренним перераспределением зарядов, а внешним источником, или их совместным действием.

Еще одна особенность диффузии заряженных частиц проявляется, если в описанной перед этим ситуации наступает *равновесие* между встречными диффузионными и дрейфовыми процессами и полный ток их равен нулю. Равенство и взаимная компенсация токов j_{dp} и j_D определяет и наличие связи между параметрами этих процессов, т.е. между подвижностью и коэффициентом диффузии. Установить эту связь, например, для электронов можно, приравняв полный ток в (2.23, *a*) нулю. Рассматривая для упрощения одномерный случай, и заменяя |grad n| = dn/dx, а $\mathcal{E} = -dV/dx$, получим:

$$D_n dn/dx = \mu_n n dV/dx,$$

82 2. Электронные состояния и проводимость твердых тел откуда

$$\frac{dn}{n} = \frac{\mu_n}{D_n} dV \,. \tag{2.24}$$

Интегрирование этого уравнения дает закон изменения концентрации электронов в поле с потенциальным рельефом U(x) = -eV(x):

$$n(x) = n(0) \exp[-\mu_n U(x)/eD_n],$$
 (2.25)

где n(0) – концентрация в точке с нулевым потенциалом.

Это выражение совпадает с распределением Больцмана $n \sim \exp(-E/k_{\rm B}T)$, если считать $eD_n = \mu_n k_{\rm B}T$. Аналогичное соотношение может быть получено и для дырок, а учитывая то, что каждый коллектив может характеризоваться своей температурой, запишем:

$$D_n = \mu_n k_{\rm B} T_n / e, \qquad (2.26, a)$$

$$D_{p} = \mu_{p} k_{\rm B} T_{p} / e.$$
 (2.26, 6)

Данные выражения, которые называются *соотношениями* Эйнштейна, устанавливают связь между коэффициентом диффузии заряженных частиц и их подвижностью. Следует также помнить, что они пригодны для невырожденных коллективов (см. гл. 3), находящихся в равновесном состоянии, а температура, которая входит в них, подчеркивает тепловую природу механизма диффузии.

Отмеченные особенности описывают процесс диффузии частиц одного типа, так называемую монополярную диффузию. Еще одна особенность проявляется в среде, которая содержит подвижные носители разного знака, и заключается в том, что их диффузия не может протекать независимо друг от друга. Это явление называется амбиполярной диффузией. В этом случае носители различного типа, которые обычно имеют различную скорость диффузии, образуют пространственный заряд в виде двойного электрического слоя. Локальное электрическое поле направлено таким образом, что будет тормозить диффузию "быстрых" частиц и ускорять "медленные". В стационарном состоянии суммарные скорости диффузии разных частиц выравниваются, и полный ток будет отсутствовать, то есть $\vec{j}_n + \vec{j}_p = 0$. Если градиенты концентраций носителей не очень отличаются друг от друга (grad $n \approx \text{grad } p$), то, исключая из (2.23, *a*) и (2.23, *b*) внутреннее электрическое поле, можно получить

$$\vec{j}_n = -\vec{j}_p \approx e \frac{\sigma_p D_n + \sigma_n D_p}{\sigma_n + \sigma_p} \operatorname{grad} n.$$
 (2.27)

Сравнивая это выражение с (2.22), нетрудно увидеть, что коллективы различных частиц будут диффундировать как одно целое, а дробь в (2.27) является коэффициентом амбиполярной диффузии.

$$D_{\rm A} = (\sigma_p D_n + \sigma_n D_p) / (\sigma_n + \sigma_p). \qquad (2.28)$$

Полезно отметить одну особенность амбиполярной диффузии, которая проявляется, когда преобладает концентрация носителей какого-то типа. Например, в донорном полупроводнике n >> p, поэтому $\sigma_n >> \sigma_p$, и если D_n одного порядка с D_p , то $D_A \approx D_p$. В акцепторном полупроводнике p >> n и $D_A \approx D_n$. В обоих случаях скорость амбиполярной диффузии определяется скоростью диффузии *неосновных* носителей.

3. РАВНОВЕСНОЕ СОСТОЯНИЕ НОСИТЕЛЕЙ

Функционирование любого электронного прибора связано с процессами взаимодействия подвижных носителей заряда с электромагнитными полями в межэлектродном пространстве. Для получения (генерации) достаточного количества подвижных носителей заряда используется какой-либо внешний источник, энергия которого расходуется на разрыв электрических связей, существующих между зарядами в веществе.

3.1. ОСНОВНЫЕ ПОЛОЖЕНИЯ СТАТИСТИКИ ЭЛЕКТРОНОВ

Кристалл представляет собой коллектив, состоящий из огромного количества микрочастиц. При постоянной температуре и отсутствии другого внешнего влияния коллектив электронов твердого тела будет находиться в термодинамическом равновесии с кристаллической решеткой. При этом непрерывно происходит обмен энергией между электронами и кристаллической решеткой. Электроны, поглощая тепловую энергию решетки (фононы), переходят в возбужденное состояние с более высокой энергией. Одновременно другие электроны, которые нашли незанятые состояния с более низкой энергией, переходят в них, возбуждая колебания кристаллической решетки. В состоянии термодинамического равновесия эти встречные процессы уравновешивают друг друга, и в любой момент полученная и переданная коллективами частиц энергия в среднем одинаковы. Такое состояние коллектива частиц называется равновесным.

Для описания свойств большого коллектива нет смысла, да и практически невозможно знать детальное поведение каждой частицы отдельно. Поведение коллектива как целого подчиняется статистическим закономерностям, основной особенностью которых является их вероятностный характер. Вероятностное состояние коллектива частиц может быть описано статистической функцией распределения

$$N(E) = dn/dE, \qquad (3.1)$$

которая соответствует энергетической плотности заполнения частицами состояний вблизи данного значения энергии E. Поэтому функцию N(E) удобно представить как произведение плотности возможных состояний $g(E) = dn_c/dE$ на среднее число частиц $f(E) = dn/dn_c$ в данном состоянии, то есть

$$N(E) = g(E) \cdot f(E). \tag{3.2}$$

Удобство такого представления заключается в том, что задача поиска функции распределения распадается на две: квантовую задачу определения характера распределения энергетических состояний и статистическую задачу нахождения вероятности заполнения этих состояний частицами.

Для описания коллектива электронов в кристалле при условии термодинамического равновесия используется *статистика* $\Phi epmu$ -*Дирака*. В соответствии с этой статистикой при низких температурах электроны кристалла занимают энергетические состояния с минимумом энергии в соответствии с принципом Паули, а их распределение по состояниям описывается функцией распределения f(E, T), которая определяет вероятность того, что энергетическое состояние E при температуре T занято электроном. Эта функция имеет вид

$$f(E,T) = \frac{1}{\left(\exp\frac{E-E_{\phi}}{k_{\rm B}T}+1\right)},\tag{3.3}$$

где $E_{\phi}(T)$ – параметр функции распределения, называемый электрохимическим потенциалом или энергией Ферми, которая также может зависеть от температуры. Из выражения (3.3)

следует, что $f(E_{\phi}) = 1/2$ при любой температуре; поэтому иногда пользуются формальным смыслом величины E_{ϕ} , которая соответствует такому энергетическому состоянию, вероятность заполнения которого электронами равна 1/2.



Рис. 3.1. Распределение Ферми-Дирака для электронов проводимости в серебре при нулевой температуре, комнатной температуре и температуре плавления.

График функции f(E) показан на рис. 3.1. При T = 0 с вероятностью 1 электронами заполнены все энергетические состояния ниже величины E_{ϕ} и нет ни одного электрона с энергией, выше E_{ϕ} . При повышении температуры часть электронов приобретает тепловую энергию и полная их энергия может стать выше уровня Ферми. Ступенчатость функции f(E) "размывается", и тем сильнее, чем выше температура. Однако тепловое возбуждение влияет только на те электроны, энергия которых незначительно отличается от E_{ϕ} (на величину порядка $k_{\rm B}T$ в обе стороны).

Знание функции распределения позволяет определить основные статистические параметры коллектива электронов. Например, интегрируя функцию распределения в некотором интервале энергий, можно найти концентрацию электронов, которые имеют энергию в этом интервале:

$$n = \int_{\mathbf{E}_1}^{\mathbf{E}_2} N(E) dE . \qquad (3.4)$$

Интеграл от произведения энергии на функцию распределения равен полной энергии всех электронов, энергия которых лежит в границах интегрирования

$$E_{\Sigma} = \int_{E_1}^{E_2} E \cdot N(E) dE. \qquad (3.5)$$

Отношение (3.5) к (3.4) определяет среднюю энергию одного электрона

$$\overline{E} = E_{\Sigma} / n = \int_{E_1}^{E_2} E \cdot N(E) dE / \int_{E_1}^{E_2} N(E) dE.$$
(3.6)

Вычисление указанных интегралов в явном виде возможно только в отдельных случаях, поэтому обычно приходится использовать приближенные методы. Однако решение подобного рода задач существенно упрощается, если рассматриваемая область энергий лежит выше уровня Ферми хотя бы на несколько единиц $k_{\rm B}T$. Это связано с тем, что при $(E - E_{\rm D}) > k_{\rm B}T$ можно пренебречь единицей в знаменателе формулы (3.3) и функция распределения Ферми-Дирака приобретает вид

$$f_{\Phi-\Pi}(E) \approx \exp[(E_{\Phi} - E)/k_{\rm B}T] = A \exp(-E/k_{\rm B}T) = f_{\rm M-B}(E).$$
 (3.7)

Выражение (3.7) характерно для распределения Максвелла-Больцмана, которое описывает поведение классических частиц. Тот факт, что при больших энергиях функция f(E) << 1 означает, что количество состояний во много раз больше, чем количество заполняющих их электронов. Электроны слабо взаимодействуют между собой и поэтому особенности, связанные с квантовой природой частиц (в частности, с их спином), при распределении по состояниям не проявляются. Именно поэтому такой коллектив частиц может быть охарактеризован классическим распределением Максвелла-Больцмана и называется невырожденным. Если указанное приближение использовать нельзя, то коллектив частиц называется вырожденным, и тогда он описывается квантовой статистикой Ферми-Дирака.

С распределением электронов по состояниям тесно связано распределение дырок. Поскольку любое состояние E может быть занятым или электроном, или дыркой, то $f_n(E) + f_p(E) = 1$. Отсюда функция распределения дырок будет иметь вид

$f_p(E) = 1 - f_n(E) = 1 / \{ \exp[(E_{\phi} - E) / k_{\rm B}T] + 1 \}.$ (3.8)



Рис. 3.2. Функции Ферми-Дирака для электронов (заштрихована) и дырок дополняют друг друга до единицы и совпадают при замене направления отсчета энергии дырок на обратное. Поскольку энергия дырок по знаку противоположна энергии электронов, то увеличение энергии электронов соответствует перемещению дырок вниз на электронной энергетической диаграмме. Если изменить знак энергии в (3.8), то функции распределения для электро-

нов и дырок по форме будут тождественны (рис. 3.2). Полезно также иметь в виду одну интересную особенность функции Ферми-Дирака $f(\varepsilon)$

$$\frac{df(\varepsilon)}{d\varepsilon} = \frac{d}{d\varepsilon} \left(\frac{1}{1+e^{\varepsilon}} \right) = \frac{e^{\varepsilon}}{\left(1+e^{\varepsilon}\right)^2} = \frac{1}{1+e^{\varepsilon}} \cdot \frac{1}{1+e^{-\varepsilon}} = f(\varepsilon) f_p(\varepsilon) = f(\varepsilon) \left[1 - f(\varepsilon) \right].$$
(3.9)

Используя это свойство, можно интегралы вида $\int f(E) [1-f(E)] \times dE = \int f(E) f(-E) dE$ приводить к виду $\int df(E)$.

3.2. РАСЧЕТ КОНЦЕНТРАЦИИ РАВНОВЕСНЫХ НОСИТЕЛЕЙ

3.2.1. Электроны в зоне проводимости

Рассчитать концентрацию электронов в зоне проводимости можно в соответствии с формулой (3.4), зная функцию распределения плотности электронов N(E) = g(E)f(E) по состояниям в зоне проводимости. Плотность состояний g(E) вблизи дна зоны проводимости описывается параболической зависимостью (рис. 3.3, *a*), полученной в разделе 1.1.5

$$g(E) = 4\pi (2m_d^*/h^2)^{3/2} (E-E_c)^{1/2}.$$
(3.10)

Вид функции распределения Ферми-Дирака (рис. 3.3, б) определяется уровнем Ферми, положение которого подлежит



Рис. 3.3. Функции распределения плотности состояний (*a*), вероятности их заполнения (б) и энергетической плотности электронов (*b*) у дна зоны проводимости полупроводника. Масштаб функции N(E) увеличен по сравнению с g(E).

уточнению, и температурой. Концентрация электронов в зоне проводимости равна площади под кривой N(E) на рис. 3.3, *в* и может быть вычислена как интеграл (3.4) в пределах от нижнего края зоны проводимости до верхнего. Однако, для удобства интегрирования целесообразно расширить верхний предел до бесконечности, что почти не скажется на точности вычислений, ибо подынтегральная функция под влиянием сомножителя f(E) очень быстро спадает с ростом энергии. Поэтому, подставляя в (3.4) выражения (3.10) и (3.3), получим

$$n = \int_{E_c}^{\infty} g(E) f(E) dE = \int_{E_c}^{\infty} \frac{4\pi (2 m_d^* / h^2)^{3/2} (E - E_c)^{1/2}}{\exp[(E - E_{\oplus}) / k_{\rm B} T] + 1} dE.$$
(3.11)

Произведем замену переменных

$$(E-E_c)/k_{\rm B}T = \varepsilon, \quad (E_{\rm p}-E_c)/k_{\rm B}T = \xi,$$
 (3.12)

где є и ξ – приведенные к единицам $k_{\rm B}T$ энергия электрона в зоне проводимости и энергия Ферми, отсчитанные от дна зоны проводимости. Подставляя (3.12) в (3.11) и вводя величину

$$N_c = 2 \left(2\pi m_d^* k_{\rm B} T/h^2 \right)^{3/2} = 4,83 \cdot 10^{21} (m_d^*/m)^{3/2} T^{3/2} \, ({\rm M}^{-3}), \tag{3.13}$$

которая иногда называется эффективной плотностью состояний в зоне проводимости, получим

$$n = N_c \Phi_{1/2}(\xi). \tag{3.14}$$

Здесь

$$\Phi_{1/2}(\xi) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1}, \qquad (3.15)$$

есть так называемый интеграл Ферми порядка 1/2. Этот интеграл не вычисляется аналитически, а задается в виде таблиц или графиков. В приложении к гл.3 представлена его зависимость от аргумента ξ . Также удобны асимптотические выражения для этого интеграла:

$$\Phi_{1/2}(\xi) = egin{cases} {\rm e}^{\xi}, & -\infty < \xi < -1, & a) \ {\rm e}^{\xi} / (0, 27 \, {\rm e}^{\xi} + 1), & -1 < \xi < 5, & \delta) \ {\rm 4}\xi^{3/2} \ / \ 3\sqrt{\pi}, & 5 < \xi < \infty, & s) \end{cases}$$

Формула (3.16, *a*) пригодна для невырожденных полупроводников, a (3.16, *b*) для сильно вырожденных полупроводников и металлов. Формулу (3.16, *b*) можно использовать для расчетов как невырожденных коллективов, где она дает более высокую точность, чем (3.16, *a*), так и при слабом вырождении. Таким образом, выражение (3.14) является универсальным и пригодным для расчета концентрации электронов в зоне проводимости как металлов, так и полупроводников.

Суммарная энергия всех электронов в зоне проводимости (отсчитанная от нижнего края зоны) может быть найдена по формуле (3.5), которая при подстановках (3.12) приобретает вид

$$E_{\Sigma} = \int_{E_c}^{\infty} (E - E_c) g(E) f(E) dE = (3 k_{\rm B} T/2) N_c \Phi_{3/2}(\xi). \quad (3.17)$$

Здесь

$$\Phi_{3/2}(\xi) = \frac{4}{3\sqrt{\pi}} \int_{0}^{\infty} \frac{\varepsilon^{3/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1}$$

является интегралом Ферми порядка 3/2, который для невырожденных коллективов ($\xi < -1$) вычисляется аналитически и равен е^{ξ}. Это означает, что средняя энергия (3.6) электрона в зоне проводимости

$$\overline{E} = 3k_{\rm B}T/2, \qquad (3.18)$$

имеет тепловую природу и вычисляется, как и для молекул идеального газа. Часто это же соотношение используют и для неравновесных состояний, используя так называемую электронную температуру T_e , которая может заметно отличаться от температуры кристаллической решетки T.

Многодолинные полупроводники, например, GaAs, InP, GaSb (рис. 1.26), имеют в зоне проводимости несколько минимумов дисперсионной кривой. Практически важное значение имеют случаи, когда верхние подзоны (обычно L- или X- долины) находятся от абсолютного минимума зоны проводимости E_c (Г-долины) на расстоянии в несколько десятых долей электронвольта. Для расчета концентраций электронов в подзонах, используя этот же метод для двухдолинных моделей зонной структуры, можно получить

$$n = n_{\Gamma} + n_{L} = N_{c\Gamma} \Phi_{1/2}(\xi) + N_{cL} \Phi_{1/2}(\xi - \Delta_{\Gamma L}/k_{\rm B}T), \quad (3.19)$$

где $N_{c\Gamma} = 2(2\pi m_{d\Gamma}^* k_{\rm B}T/h^2)^{3/2}$, $N_{c\rm L} = 2(2\pi m_{d\rm L}^* k_{\rm B}T/h^2)^{3/2}$ – эффективные плотности состояний, а $m_{d\Gamma}^*$ и $m_{d\rm L}^*$ – эффективные массы плотности состояний в Г- и L-долинах соответственно.

Интересная особенность наблюдается в тех материалах, где $m_{d\Gamma}^* < m_{dL}^*$ и соответственно $N_{c\Gamma} < N_{cL}$. Благодаря большей плотности состояний концентрация электронов в верхней долине при некоторых условиях может быть выше, чем в нижней. Заселенность верхней долины удобно характеризовать

функцией $b_{\rm L}(T)=n_{\rm L}/n=(n_{\rm L}/n_{\rm \Gamma})/(1+n_{\rm L}/n_{\rm \Gamma})$, где для невырожденного полупроводника

 $n_{\rm L}/n_{\rm \Gamma} = (m_{d\,\rm L}^*/m_{d\,\rm \Gamma}^*)^{3/2} \exp(-\Delta_{\rm \Gamma L}/k_{\rm B}T).$



Рис. 3.4. Зависимость относительной заселенности L-долины в GaAs от электронной температуры или средней энергии электронов.

На рис. 3.4 представлена зависимость $b_{\rm L}(T)$ для GaAs ($m_{d\Gamma}^* = 0,067m, m_{dL}^* =$ = 0,56m, $\Delta_{\Pi} = 0,3$ эВ), откуда видно, что при комнатной температуре практически все электроны находятся в нижней долине, а при T > 1500 К, преобладают электроны, заполняющие L-долину. Нужно также иметь в виду, что такие температуры обычно достигаются не нагреванием материала полупроводника, а "разогревом" электронного газа, например, в сильном электрическом поле. В этом

(3.20)

случае в (3.20) нужно использовать электронную температуру T_{e} .

3.2.2. Дырки в валентной зоне

Метод расчета концентрации дырок в валентной зоне полупроводника по существу такой же, как и для электронов в зоне проводимости. Плотность состояний вблизи потолка валентной зоны описывается функцией (рис. 3.5, *a*)

$$g(E) = 4\pi (2m_{pd}^{*}/h^2)^{3/2} (E_v - E)^{1/2}.$$
(3.21)

Вероятность заполнения состояний электронами в валентной зоне описывается функцией Ферми-Дирака (рис. 3.5, δ). Там же на вставке изображен фрагмент этой функции для дырок $f_p(E) = 1 - f(E)$. Интегрирование функции N(E) (рис. 3.5, ϵ)



Рис. 3.5. Функции распределения плотности состояний (*a*), вероятности их заполнения (*б*) и энергетической плотности электронов (*s*) и дырок (*c*) у потолка валентной зоны полупроводника. Масштаб функции $N_p(E)$ увеличен по сравнению с N(E).

в пределах всей валентной зоны не имеет смысла, поскольку функция плотности состояний (3.21) определена только для верхних состояний зоны. Удобнее рассматривать свободные состояния, где плотность распределения дырок изображена на рис. 3.5, *в* заштрихованной частью, или на рис. 3.5, *г* в увеличенном масштабе.

Концентрация дырок в валентной зоне равна интегралу от функции распределения $N_p(E) = g(E) f_p(E)$ по всем состояниям валентной зоны:

$$p = \int_{-\infty}^{E_v} g(E) f_p(E) dE . \qquad (3.22)$$

Здесь, подобно предыдущему интегралу, нижний предел интегрирования расширен до бесконечности, что практически не влияет на результат. При этом интеграл (3.22) по структуре совпадает с (3.11) с точностью до направления отсчета энергии. Полная тождественность достигается подстановками, аналогичными (3.12):

$$(E_v - E)/k_{\rm B}T = \varepsilon, \qquad (E_v - E_{\rm p})/k_{\rm B}T = \xi_p, \qquad (3.23)$$

при которых результат интегрирования принимает вид

$$p = N_v \Phi_{1/2}(\xi_p), \qquad (3.24)$$

где

$$N_{v} = 2(2\pi m_{pd}^{*} k_{\rm B} T/h^{2})^{3/2} = 4,83 \cdot 10^{21} (m_{pd}^{*}/m)^{3/2} T^{3/2} \, ({\rm m}^{-3}) \quad (3.25)$$

– эффективная плотность состояний в валентной зоне, а интеграл Ферми $\Phi_{1/2}(\xi_p)$ отличается от (3.15), тем что здесь ξ_p представляет собой уровень Ферми, отсчитанный от потолка валентной зоны и выраженный в единицах $k_{\rm B}T$. Асимптотические выражения для интеграла Ферми $\Phi_{1/2}(\xi_p)$, аналогичны соответствующим выражениям (3.16) для интеграла $\Phi_{1/2}(\xi)$.

Полученные соотношения (3.14) и (3.24) для расчета концентрации подвижных носителей в приближении эффективной массы являются универсальными, ибо при их выводе принималось только одно слабое допущение о расширении пределов интегрирования. Наиболее распространенным является приближение (3.16, *a*), согласно которому

$$n = N_c \exp \left[(E_{\phi} - E_c) / k_{\rm B} T \right],$$
 (3.26)

$$p = N_v \exp \left[(E_v - E_{\phi}) / k_{\rm B} T \right].$$
 (3.27)

Эти выражения пригодны для невырожденных полупроводников, так как получены при условии ξ , $\xi_p <-1$, которое означает, что уровень Ферми удален от края соответствующей разрешенной зоны более, чем на $k_{\rm B}T$. При этом в интеграле Ферми можно пренебречь единицей в знаменателе, что означает фактически переход к статистике Максвелла-Больцмана.

Из полученных выражений следует два полезных вывода. Взяв отношение последних выражений, учитывая (3.13) и (3.25), получим

$$p/n = (m_{pd}^{*}/m_{d}^{*})^{3/2} \exp\{2 \left[(E_{c} + E_{v})/2 - E_{\phi}\right]/k_{\rm B}T\}.$$
 (3.28)

Отсюда мы видим, что соотношение между концентрациями подвижных электронов и дырок зависит от отношения их эффективных масс и особенно сильно от положения уровня Ферми относительно середины запрещенной зоны $(E_c + E_v)/2$. Концентрация подвижных носителей больше в той зоне, ближе к которой расположен уровень Ферми. Последнее утверждение пригодно для равных эффективных масс электронов и дырок, но они отличаются друг от друга не более, чем на порядок, тогда как за счет даже небольшого смещения уровня Ферми можно получить изменение концентраций на несколько порядков. Это перераспределение концентраций подвижных носителей между зонами может быть осуществлено введением в полупроводник примесей, так как положение уровня Ферми существенно зависит от вида и концентрации вводимой примеси. Однако произведение концентраций подвижных электронов и дырок не зависит от положения уровня Ферми и при данной температуре зависит только от ширины $\Delta E = E_c - E_n$ запрещенной зоны полупроводника и эффективных масс электронов и дырок:

$$n \ p = N_c N_v \exp(-\Delta E/k_B T) = n_i \ p_i = n_n \ p_n = n_p \ p_p. \quad (3.29)$$

Первое равенство в этом соотношении, которое непосредственно следует из умножения (3.26) на (3.27), называют законом действующих масс. Продолжение этого равенства означает, что произведение концентраций носителей не зависит от вида и концентрации примесей (индексы i, n и p обозначают принадлежность к собственному, донорному и акцепторному полупроводнику соответственно). Поэтому введение в полупроводник, например, донорных примесей увеличивает в зоне проводимости количество электронов и автоматически в соответствии с (3.29) уменьшает количество дырок в валентной зоне.

Закон действующих масс пригоден для невырожденных полупроводников, находящихся в состоянии термодинамического равновесия. Иногда (3.26) и (3.27) можно использовать также в тех случаях, когда кроме теплового существует и некоторый другой способ возбуждения носителей, нарушающий равновесие. В этих случаях статистика Ферми-Дирака непри3. Равновесное состояние носителей

годна, но выражения (3.26) и (3.27) можно использовать, если вместо уровня Ферми формально ввести так называемые *квази*уровни $E_{\phi n}$ и $E_{\phi p}$. Поскольку для электронов и дырок они, как правило, различны, то

$$np = n_0 p_0 \exp \left[(E_{\phi n} - E_{\phi p}) / k_{\rm B} T \right].$$
 (3.30)

Здесь через n_0 и p_0 обозначены равновесные концентрации электронов и дырок, рассчитанные в соответствии с (3.26) и (3.27). Последнее выражение показывает, что закон действующих масс не выполняется для неравновесных носителей заряда. Кроме того, полезно отметить, что в случае, когда неравновесные концентрации электронов и дырок превышают их равновесные значения, то $E_{\phi n} > E_{\phi} > E_{\phi p}$. Это наблюдается, если внешнее возбуждение, вызывая нарушение термодинамического равновесия, приводит к появлению избыточных носителей. Если же нарушение равновесные концентрации уменьшаются и электронный квазиуровень Ферми снижается, а дырочный повышается относительно E_{ϕ} .

3.2.3. Электроны в кристаллах низкой размерности

Для одномерных и двумерных кристаллов наблюдается обобществление электронов и соответственно зонный спектр лишь в тех направлениях, где выполняется условие трансляционной симметрии. Для других направлений электронные состояния остаются локализованными и сохраняется дискретный спектр. Вычисление количества электронов в соответствующих зонах может быть проведено в принципе тем же методом, что и для объемных (3D) кристаллов. Но произведение функции (1.31) или (1.33) для плотности энергетических состояний на функцию Ферми-Дирака дает энергетическую плотность электронов в расчете на единицу площади в двумерных кристаллах и на единицу длины в одномерных. Поэтому интегрирование этих плотностей по всей разрешенной зоне позволяет рассчитывать значения соответствующих концентраций. Так, в 2*D*-кристаллах интегрирование

$$\int_{E_i}^{\infty} g_i(E) f(E) dE = n_{is}$$
(3.31)

определяет значение поверхностной концентрации электронов в двумерной зоне, соответствующей каждому дискретному состоянию E_i . Подстановка выражений (1.31) для энергетической плотности двумерных состояний и (3.3) для функции Ферми-Дирака, а также замена переменных $(E - E_i)/k_{\rm B}T = \varepsilon$, $(E_{\rm \phi} - E_i)/k_{\rm B}T = \xi$, подобная (3.12), приводит выражение (3.31) к форме

$$n_{is} = (4\pi m_{xy}^* k_{\rm B} T / h^2) \int_0^\infty \frac{d\varepsilon}{\exp(\varepsilon - \xi) + 1}.$$
 (3.32)

Такой интеграл называется интегралом Ферми нулевого порядка $\Phi_0(\xi)$. Хотя его значения и табулированы, как и для подобных интегралов других порядков (приложение к гл. 3), он может быть вычислен в явном виде

$$\Phi_0(\xi) = \int_0^\infty \frac{d\varepsilon}{\exp(\varepsilon - \xi) + 1} = \ln \frac{\exp \xi}{\exp(\varepsilon - \xi) + 1} \bigg|_0^\infty = \ln(1 + e^{\xi}). \quad (3.33)$$

Обозначая $N_{c2}=4\pi m_{xy}^{*}k_{\rm B}T\,/\,h^{2}$, получаем выражение для поверхностной концентрации

$$n_{is} = N_{c2} \ln\{1 + \exp[(E_{\phi} - E_i)/k_{\rm B}T]\}.$$
(3.34)

При малых значениях аргумента ($\xi < -1$) интеграл $\Phi_0(\xi)$, как и подобные (рис. П.З.1), аппроксимируется функцией e^{ξ} . Поэтому для невырожденного коллектива в данной зоне удобнее пользоваться выражением

$$n_{is} = N_{c2} \exp[(E_{\phi} - E_i)/k_{\rm B}T].$$
 (3.35)

В одномерных (1*D*) кристаллах концентрация электронов n_l на единицу длины в зоне, лежащей над состоянием E_{ij} , квантованным по двум направлениям, определяется интегрированием функции распределения по всей зоне:

$$\int_{E_{ij}}^{\infty} g_{ij}(E)f(E)dE = n_{ijl}.$$
(3.36)

Подстановка одномерной плотности состояний (1.33) и подобная замена переменных $(E - E_{ij})/k_{\rm B}T = \varepsilon$, $(E_{\rm \Phi} - E_{ij})/k_{\rm B}T = \xi$ приводят к выражению

$$n_{ijl} = N_{c1} \Phi_{-1/2}(\xi), \qquad (3.37)$$

где

$$\Phi_{-1/2}(\xi) = \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \frac{\varepsilon^{-1/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1}$$
(3.38)

есть интеграл Ферми порядка -1/2 (см. приложение к гл. 3), а $N_{c1} = (2\pi m_x^{\ *} k_{\rm B} T/h^2)^{1/2}$ – так называемая эффективная плотность состояний в одномерной зоне. Как и для других размерностей, для невырожденных электронов в данной зоне ($\xi < -1$) интеграл $\Phi_{-1/2}(\xi)$ аппроксимируется функцией e^{ξ} и для расчета концентрации удобнее пользоваться выражением

$$n_{ijl} = N_{c1} \exp[(E_{\phi} - E_{ij})/k_{\rm B}T].$$
(3.39)

Часто интерес представляет вопрос о средней энергии электронов в зонах кристаллов низкой размерности. Используя для этого выражение (3.6), после подстановки соответствующих функций распределения нетрудно получить для кристалла *n*-й размерности

$$\overline{E} = \frac{n}{2} k_{\rm B} T \frac{\Phi_{n/2}(\xi)}{\Phi_{n/2-1}(\xi)}.$$
(3.40)

Для невырожденных полупроводников интегралы Ферми равны e^{ξ} и средняя энергия определяется только тепловым движением и равна $nk_{\rm B}T/2$, по $k_{\rm B}T/2$ на каждую степень свободы движения электронов вдоль кристалла, что находится в соответствии со статистикой Максвелла-Больцмана.

3.2.4. Заполнение локализованных состояний

Общие принципы статистики Ферми-Дирака, которые используются для описания заполнения непрерывных или квазинепрерывных электронных состояний в разрешенных зонах, могут быть использованы и для локализованных состояний с дискретным энергетическим спектром. Строго говоря, в этом случае распределение Ферми-Дирака непригодно, так как оно получено при условии, что каждое состояние может быть занято только двумя электронами с противоположными спинами. Но, например, в донорном энергетическом состоянии, образованном пятивалентным примесным атомом, может находиться только один электрон с любым направлением спина, поэтому такое состояние считается двукратно вырожденным. Подобная ситуация имеет место и для других примесей, в том числе многозарядных, и может быть описана при помоши распределения Гиббса, которое учитывает кратность вырождения локализованных состояний. Однако для понимания основных механизмов генерации носителей заряда ограничимся более простой функцией Ферми-Дирака с небольшой потерей точности в количественных расчетах.

В соответствии с физическим смыслом функции f(E) вероятность заполнения электроном какого-либо локализованного энергетического состояния E_t равна значению функции при этом значении энергии:

$$f(E_t) = 1 / \{ \exp\left[(E_t - E_{\oplus}) / k_{\rm B} T \right] + 1 \}.$$
(3.41)

Если концентрация таких центров равна $N_{\scriptscriptstyle t}$, то из них заполненными будут

$$N_t^{n} = N_t f(E_t) = N_t / \{ \exp \left[(E_t - E_{\phi}) / k_{\rm B} T \right] + 1 \}, \qquad (3.42)$$

а остальные (N_t - N_tⁿ) - свободными от электронов

$$N_t^{p} = N_t - N_t^{n} = N_t [1 - f(E_t)] = N_t / \{ \exp \left[(E_{\phi} - E_t) / k_{\rm B} T \right] + 1 \}.$$
 (3.43)

В том случае, если локализованные состояния образованы пятивалентной донорной примесью с концентрацией N_{π} , то

(3.43) определяет количество положительно ионизированных доноров

$$N_{_{\rm A}}^{^+} = N_{_{\rm A}} / \{ \exp \left[(E_{_{\rm \Phi}} - E_{_{\rm A}}) / k_{_{\rm B}} T \right] + 1 \},$$
 (3.44)

а остальные примесные атомы будут удерживать свои пятые "лишние" электроны и оставаться нейтральными. Если же в полупроводнике присутствуют трехвалентные акцепторные атомы с концентрацией $N_{\rm a}$, то в соответствии с (3.42) отрицательно заряженными из них будут

$$N_{\rm a}^{-} = N_{\rm a} / \{ \exp \left[(E_{\rm a} - E_{\rm p}) / k_{\rm B} T \right] + 1 \}, \qquad (3.45)$$

а остальные (N_a-N_a⁻) будут оставаться нейтральными.

При этом нужно помнить, что ионизация донорных атомов сопровождается возбуждением "лишнего" электрона в зону проводимости, а ионизация акцепторов – возбуждением валентного электрона в вакантное состояние и соответственно образованием вакансии в системе валентных связей, т.е. дырки. Это справедливо при наличии в полупроводнике только одного вида примеси.

В заключение отметим, что полученные выражения для расчета концентраций подвижных носителей и ионизированных атомов являются в пределах принятых допущений универсальными. Но воспользоваться ими в той или иной конкретной ситуации можно, только рассчитав предварительно положение уровня Ферми. В этом и состоит задача определения статистических свойств подвижных носителей заряда для данного типа металла или полупроводника. Найти положение уровня Ферми можно, если на полученные выражения наложить некоторое дополнительное условие. Для металлов таким условием является известное значение концентрации электронов в зоне проводимости. Для полупроводников оно формулируется как уравнение электронейтральности

$$n + N_{\rm a}^{-} = p + N_{\rm \pi}^{+},$$
 (3.46)

суть которого заключается в том, что процессы генерации и рекомбинации подвижных носителей не нарушают электриче-

ской нейтральности вещества. Это физически простое условие справедливо для равновесных состояний при отсутствии электрических полей. В дальнейшем мы увидим, что наличие поля может вызывать локальное нарушение электронейтральности вещества за счет пространственного перераспределения подвижных носителей.

3.3. ПРОВОДИМОСТЬ МЕТАЛЛОВ

В металлических кристаллах все обобществленные электроны энергетически расположены в зоне проводимости, а их концентрация определяется концентрацией атомов и количеством обобществляемых валентных электронов от каждого атома. Распределение электронов по энергетическим состояниям зоны проводимости описывается статистикой Ферми--Дирака, поэтому концентрацию электронов можно определить по формуле (3.11). Она значительно упрощается при нулевой температуре, так как при T = 0 электронами заняты все состояния только до уровня Ферми, т.е. f(E) = 1 при $E < E_{\phi}$ и f(E) = 0 при $E > E_{\phi}$. Поэтому формулу (3.11), заменяя верхний предел интегрирования на $E_{\phi}(0)$, можно записать в виде

$$n = \int_{0}^{\infty} g(E)f(E)dE = \int_{0}^{E_{\phi}(0)} g(E) \cdot 1 \cdot dE = \frac{8\pi}{3} \Big[2m^{*}E_{\phi}(0) / h^{2} \Big]^{3/2}$$

Отсюда при известной концентрации электронов *n* находим энергию Ферми при нулевой температуре

$$E_{\phi}(0) = \frac{h^2}{2m^*} \left(\frac{3n}{8\pi}\right)^{2/3}.$$
 (3.47)

Это есть наибольшее значение энергии, которую может иметь электрон в металле при T = 0. Такой же результат для $E_{\phi}(0)$ можно получить из (3.14), если использовать асимптотическое выражение для интеграла Ферми (3.16, *в*).

Оценивая энергию Ферми, нужно иметь в виду, что для типичных металлов эффективная масса электронов близка к обычной, поэтому 2. 3. Равновесное состояние носителей

$$E_{\rm b}(0)~({\rm yB}) = 0.365 \cdot 10^{-18} \, n^{2/3} \, ({\rm M}^{-3}), \qquad (3.47, a)$$

что для присущих металлам концентраций электронов порядка $10^{28}...10^{29}$ м⁻³ составляет несколько электронвольт.

При температурах, отличных от нуля, приближенное выражение для уровня Ферми $E_{\phi}(T)$ можно получить, воспользовавшись разложением интеграла Ферми $\Phi_{1/2}(\xi)$ в формуле (3.15) в степенной ряд для большого положительного параметра ξ . Такое приближение возможно, поскольку для существующих в металлах концентраций электронов параметр $\xi = E_{\phi}/k_{\rm B}T$ является достаточно большим и можно ограничиться двумя первыми членами сходящегося ряда:

$$E_{\phi}(T) = E_{\phi}(0) \left[1 - \frac{\pi^2}{12} \left(\frac{k_{\rm B}T}{E_{\phi}(0)} \right)^2 \right].$$
(3.48)

Численные расчеты показывают, что с ростом температуры энергия Ферми снижается относительно $E_{\phi}(0)$ на очень малую величину. Так, при $T \approx 1000$ К уровень Ферми уменьшается всего на $10^{-3}...10^{-4}$ эВ, и только при температурах порядка $10^4...10^5$ К уровень Ферми смещается в область отрицательных значений, а электронный газ становится невырожденным.



Рис. 3.6. Функции распределения плотности состояний (a), вероятности их заполнения (δ) и энергетической плотности электронов (b) в зоне проводимости металлов.

На рис. 3.6 показаны функции распределения плотности состояний (*a*), вероятности заполнения состояний (*б*) и плотности электронов (*в*) в зоне проводимости металла при T = 0 и T > 0.

Нужно отметить, что для температур, при которых металлы находятся в твердом состоянии, электронный газ в металле является вырожден-

ным и распределение электронов по энергиям определяется не тепловым движением, а принципом Паули. Подтвержде-

нием этого положения может служить значение средней энергии одного электрона, которое можно получить из формулы (3.6). Для этого воспользуемся разложением подынтегральных функций в ряды и ограничимся первыми членами ряда.

В результате получим

$$\overline{E} = E_{\phi}(0) \cdot \frac{3}{5} \left[1 - \frac{5\pi^2}{12} \left(\frac{k_{\rm B}T}{E_{\phi}(0)} \right)^2 \right].$$
(3.49)

Первое слагаемое в этом выражении характеризует среднюю энергию электрона при нулевой температуре и соответственно называется "*нулевой*" *энергией*, которая имеет чисто квантовую природу:

$$\overline{E}_{0} = 3E_{0}(0)/5.$$
 (3.50)

Имея в виду случайный характер движения электронов, данную энергию рассматривают как кинетическую энергию хаотического движения, которое существует и при T = 0, со средней скоростью

$$v_0 = \sqrt{2\overline{E}_0 / m^*} = \sqrt{6E_{\oplus}(0)/5m^*}$$
. (3.51)

Второе слагаемое, которое зависит от температуры, описывает среднюю *тепловую энергию* электрона

$$\overline{E}_{T} = \pi^{2} (k_{\rm B} T)^{2} / 4E_{\phi}(0).$$
 (3.52)

Сравнение \bar{E}_0 и \bar{E}_T показывает, что даже при температурах, близких к температуре плавления, тепловая энергия электрона составляет $10^{-3}...10^{-4}$ от "нулевой" энергии, а соответствующая тепловая скорость – единицы процентов от "нулевой" скорости. Такая слабая восприимчивость электронов в металлах к нагреванию следует из самого характера функции распределения электронов. Это обстоятельство объясняет, в частности, малый вклад теплоемкости электронного газа в общую теплоемкость металлов.

По классической теории удельная теплоемкость металлов должна быть выше теплоемкости диэлектриков, так как она, кроме решеточной теплоемкости, включает в себя еще и теплоемкость электронного газа. Но экспериментальные исследования не обнаружили существенной разницы в значениях удельной теплоемкости металлов и диэлектриков. Объяснение этому было дано только при использовании квантовой статистики. Поведение функции распределения Ферми-Дирака (рис. 3.1) при изменении температуры показывает, что тепловая энергия порядка $k_{\rm B}T$ приобретается очень малой частью всех электронов, составляющей величину порядка $k_{\rm B}T/E_{\rm ob}$ от общего их количества, и в пересчете на один электрон имеет значение порядка $(k_{\rm E}T)^2/E_{\rm d}$, что значительно ниже их полной энергии (3.49). Именно поэтому электронный газ, который не является теплоемким, очень слабо влияет на общую теплоемкость металлов.

Несостоятельность классических представлений проявилась прежде всего в вопросе о температурной зависимости электропроводности металлов, для которой эксперимент показал, что в широком температурном диапазоне $\sigma \sim T^{-1}$. По классической теории причиной рассеивания электронов в металлах являются их столкновения с атомами кристаллической решетки, поэтому и длина свободного пробега электронов \overline{l} должна быть близкой к величине межатомного расстояния и составлять несколько ангстрем, в то время как численные оценки дают значения порядка сотен периодов решетки. Применение для электронного газа классической статистики Максвелла-Больцмана, в соответствии с которой хаотическое движение определяется только тепловым, а его скорость $v_{xaor} \sim T^{1/2}$.

Квантовая теория, применяя для электронного коллектива статистику Ферми-Дирака, предполагает, что при обычных условиях хаотическое движение практически не зависит от температуры, а его скорость $v_{xaor} \approx v_0$ определяется по (3.51). Второе принципиальное утверждение квантовой теории свя-

104

зано с тем, что столкновения электронов происходят не непосредственно с атомами кристаллической решетки, а с ее дефектами, которые возникают в том числе и в результате тепловых колебаний. Количественная оценка, проведенная в разд. 2.3, показывает, что в этом случае средняя длина свободного пробега, во-первых, составляет величину порядка 10^{-7} м и, во-вторых, $\bar{l} \sim T^{-1}$. Поэтому $\sigma \sim \bar{\tau} = \bar{l} / v_0 \sim T^{-1}$, что соответствует экспериментальным фактам.

Следует обратить внимание на то, что энергетическое распределение электронов в зоне проводимости металлов подобно распределению плотности молекул жидкости по высоте. Процесс перехода электронов на более высокие энергетические уровни при нагревании аналогичен процессу испарения молекул жидкости. С этой точки зрения более правильным для электронов в металле было бы название "электронная жидкость" (уровень Ферми электронов соответствует высоте уровня поверхности жидкости). Такая аналогия характерна для электронного коллектива в целом, а та часть электронов, энергия которых превышает уровень Ферми на несколько единиц $k_{\rm B}T$, аналогична пару над поверхностью жидкости и может быть названа "электронным газом", к которому с достаточной точностью может быть применима статистика Максвелла-Больцмана.

3.4. СОБСТВЕННАЯ ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКОВ

В полупроводниках обобществленные электроны принимают участие в создании ковалентных связей между атомами и, будучи связанными обменными силами с кристаллической решеткой, не могут принимать участие в электропроводности. Чтобы электроны стали подвижными в электрическом поле, они должны преодолеть силы связей, получив определенную энергию от внешних источников, и стать свободными. Здесь мы рассматриваем образование (генерацию) подвижных носителей, которое возникает под действием только теплового движения.

Процессы тепловой генерации в полупроводниках на языке энергетических диаграмм можно описать следующим образом. При нулевой температуре все состояния валентной зоны полностью заполнены электронами, а с повышением температуры часть электронов, энергия которых становится больше ширины запрещенной зоны, переходит в зону проводимости. В валентной зоне создается такое же количество дырок. Процессы генерации сопровождаются обратными переходами электронов в валентную зону и последующим их исчезновением как подвижных носителей, т.е. рекомбинацией с дырками. В условиях термодинамического равновесия скорость тепловой генерации электронно-дырочных пар равна скорости их рекомбинации, а концентрации электронов и дырок в соответствующих зонах устанавливаются статистически постоянными или *равновесными*.

Поскольку в беспримесных (собственных) полупроводниках образование и исчезновение носителей всегда происходит парами, то концентрации электронов и дырок равны между собой, и $n_i = p_i$ (индекс *i* обозначает принадлежность к собственному полупроводнику). Это равенство и является *уравнением электронейтральности*, суть которого в том, что процессы генерации подвижных носителей не нарушают электрической нейтральности вещества. Тогда для невырожденного полупроводника можно записать

$$N_c \exp \left[(E_{\phi} - E_c) / k_{\rm B} T \right] = N_v \exp \left[(E_v - E_{\phi}) / k_{\rm B} T \right].$$
(3.53)

Логарифмируя обе части этого равенства и вычисляя значение энергии Ферми, получим

$$E_{\oplus} = \frac{E_c + E_v}{2} + \frac{k_{\rm B}T}{2} \ln \frac{N_v}{N_c} = E_i + \frac{3}{4} k_{\rm B}T \ln \frac{m_{pd}^*}{m_d^*}, \qquad (3.54)$$

где $E_i = (E_c + E_v)/2$ – середина запрещенной зоны.

Анализ формулы (3.54) показывает, что с ростом температуры уровень Ферми линейно смещается от середины запрещенной зоны при T = 0 в сторону зоны с более легкими носителями (рис. 3.7). Поскольку эффективные массы плотности состояний m_d^* и m_{pd}^* не отличаются более чем на порядок, то это смещение не превышает $(1...2)k_{\rm B}T$. Например, при T = 300 К уровень Ферми расположен в германии на ~8 мэВ,



Рис. 3.7. Изменение положения уровня Ферми в собственном полупроводнике в зависимости от температуры. Изображено также температурное изменение ширины запрещенной зоны.

а в кремнии на ~12 мэВ ниже середины запрещенной зоны.

С изменением температуры также меняется и ширина запрещенной зоны по линейному закону (1.34). В таких распространенных полупроводниках, как Ge, Si и GaAs, температурный коэффициент α является отрицательным и равен приблизительно 10^{-4} эВ/К. Поэтому температурные смещения уровня Ферми и краев запрещенной зоны соизмеримы. В узкозонных полупроводниках с очень различными эффективными массами носителей с ростом температуры уровень Ферми может приблизиться к дну зоны проводимости или даже оказаться в ней. Так, например, в InSb $\Delta E(T) = 0.26 - 2.7 \cdot 10^{-4}T$ (эВ), $m_{pd}^* \approx 10m_d^*$ и уже при $T \ge 440$ К уровень Ферми лежит в зоне проводимости. Это приводит к вырождению газа более легких носителей, и для расчета концентраций *n* и *p* необходимо пользоваться общими выражениями (3.14) и (3.24).

Если полупроводник остается невырожденным, то остаются в силе выражения (3.26) и (3.27), которые при найденном положении уровня Ферми (3.53) упрощаются и принимают вид

$$n_i = p_i = \sqrt{N_c N_v} \exp(-\Delta E/2k_{\rm B}T). \qquad (3.55)$$

На рис. 3.8 представлена зависимость концентраций электронов и дырок в германии от температуры, изображенная в


Рис. 3.8. Зависимость концентрации электронов и дырок в чистом германии от температуры, изображенная в разных координатах.

различных координатах: *a*) линейный масштаб наиболее нагляден с точки зрения степени влияния температуры на концентрацию; *б*) логарифмический позволяет на одном графике изображать величины, которые отличаются на несколько порядков; *в*) зависимость логарифма концентрации от обратной температуры 1/T описывается функцией:

$$\ln n_i = \ln(N_c N_v)/2 - \Delta E/2k_{\rm B}T \approx \text{const} - (\Delta E/2k_{\rm B}) \cdot (1/T). \quad (3.56)$$

Если учитывать только указанную температурную зависимость, то равенство (3.56) описывает прямую линию с угловым коэффициентом $\Delta E/2k_{\rm B}$. Преимущество последнего способа представления зависимости (3.55) состоит в том, что, вопервых, линейность построенного по результатам эксперимента графика в координатах $\ln n_i = f(1/T)$ подтверждает правильность расчетной формулы (3.55) и, во-вторых, позволяет по наклону прямой определить ширину запрещенной зоны полупроводника.

При более строгом анализе температурной зависимости $n_i(T)$ нужно учитывать, что N_c и N_v зависят от температуры по закону $\sim T^{3/2}$, и что от температуры зависит ширина запрещенной зоны $\Delta E = f(T)$.

Зависимость удельной электропроводности от температуры определяется в основном концентрационной зависимостью и имеет вид

$$\sigma_i = en_i\mu_n + ep_i\mu_p = \sigma_0 \exp(-\Delta E/2k_{\rm B}T), \qquad (3.57)$$

где коэффициент $\sigma_0 = e \sqrt{N_c N_v} (\mu_n + \mu_p) = \text{const}(T)$, так как зависимость подвижности от температуры при рассеивании электронов и дырок на тепловых колебаниях решетки пропорциональна $T^{-3/2}$ и компенсирует температурную зависимость N_c и N_v .

3.5. ПРИМЕСНАЯ ПРОВОДИМОСТЬ

В примесных полупроводниках подвижные электроны и дырки могут появляться не только в результате нарушения валентных связей в основном полупроводнике, но и за счет ионизации примесных атомов. Ионизация донорных атомов

приводит к появлению в зоне проводимости свободных электронов, а ионизация акцепторных к созданию дырок в валентной зоне.

Ha энергетической лиаграмме (рис. 3.9) эти процессы выглядят следующим образом. При нулевой температуре в донорном полупроводнике электронами заняты все состояния валентной зоны и примесного уровня, в а акцепторном - все состояния в валентной зоне. С повышением лействием температуры под возбуждения теплового



Рис. 3.9. Механизмы тепловой генерации подвижных носителей в примесных полупроводниках. Светлыми кружками условно обозначено энергетическое положение дырок, темными – электронов. Знаки + и – обозначают, что в результате генерации носителей примесные атомы превращаются соответственно в ионизированные доноры и акцепторы.

пронизк сдаюн оцерекходюжежнаровновуровней в зону проводимости (1), переход электронов из валентной зоны на акцепторный уровень (2), а также межзонные переходы (3), при которых генерируются как электроны, так и дырки.

Как и для собственного полупроводника, эти переходы сопровождаются обратными переходами. Так, в донорном полупроводнике электроны из зоны проводимости возвращаются на донорные уровни и в валентную зону, а в акцепторном из зоны проводимости и с акцепторного уровня – в валентную зону, где происходит их рекомбинация с дырками. При постоянной температуре процессы генерации и рекомбинации подвижных носителей уравновешиваются и их концентрации устанавливаются статистически постоянными.



Рис. 3.10. Типичные температурные зависимости концентрации электронов и дырок в полупроводнике, содержащем донорные (акцепторные) примеси. Обозначено три характерных участка этих зависимостей.

Интенсивность различных переходов зависит от температуры и требуемого изменения энергии. Поскольку энергия ионизации примесных атомов ($\Delta E_{\pi} \sim 0,01$ эВ) значительно меньше энергии, необходимой для разрыва валентных связей собственного полупроводника (ΔE – десятые доли, единицы электронвольт), то при температурах в единицы и десятки кельвинов заметной будет только ионизация примесных атомов, а межзонные переходы, интенсивность которых зависит от ширины запрещенной зоны, особенно сущест-

венны при сотнях кельвинов. Это позволяет относительно просто описать качественную картину изменения концентраций с ростом температуры для полупроводника с одним видом примеси (рис. 3.10).

Рассмотрим, например, донорный полупроводник или полупроводник типа *n* (в дальнейшем все величины для донорного полупроводника будем обозначать индексом *n*). При низких температурах происходит *ионизация донорных атомов*, в результате чего появляются подвижные носители только одного вида – электроны в зоне проводимости. Их концентрация будет равна концентрации неподвижных положительно заряженных доноров: $n_n = N_{\pi}^+$. Иногда участок 1 на рис. 3.10 называют областью *вымораживания* примесей, что оправдано при описании процессов, которые происходят при охлаждении полупроводника.

С ростом температуры возрастает вероятность ионизации доноров и при температуре в несколько десятков кельвинов практически все доноры будут ионизированы. Если при этих температурах количество примесей $N_{\rm g}$ заметно превышает количество подвижных носителей в собственном полупроводнике, то концентрация электронов в зоне проводимости будет равной концентрации примеси $n_n = N_{\rm g}$. Такая ситуация называется истощением примеси.

Только при температуре в несколько сотен кельвинов становятся заметными межзонные переходы (3 на рис. 3.9). Такие переходы приводят к образовнию пары подвижных носителей: электрона и дырки. Поэтому концентрация электронов в зоне проводимости определяется не только количеством примесей, но и количеством созданных электронно-дырочных пар: $n_n = N_{\pi} + p_n$.

Дальнейшее повышение температуры выше некоторого значения T_i (рис. 3.10) приводит к интенсивной генерации электронно-дырочных пар, и их количество может заметно превысить величину $N_{\rm g}$. В этом случае $n_n \approx p_n \approx n_i$ и изменения концентрации носителей будут такими же, как и в собственном полупроводнике.

Такая качественная оценка процессов в донорном полупроводнике позволяет в области низких и средних температур определить электроны как *основные* носители, а дырки как *неосновные*.

В акцепторном полупроводнике типа *p* основными носителями будут дырки, которые создаются в валентной зоне за счет перехода электронов в примесные акцепторные состояния (ионизация примесей) и в зону проводимости (межзонные переходы), а неосновными – электроны в зоне проводимости. 3. Равновесное состояние носителей

Качественная температурная зависимость концентрации основных и неосновных носителей в акцепторном полупроводнике будет такой же, как и в донорном. В области высоких температур распределение носителей на основные и неосновные не имеет смысла.

3.5.1. Донорные полупроводники

Воспользоваться выражениями (3.14) и (3.24), или (3.26) и (3.27) для расчета концентрации носителей при данных условиях возможно только при известной энергии Ферми. Ее величина может быть определена при подстановке в уравнение электронейтральности (3.36), которое при наличии только донорной примеси принимает вид

$$n_n = p_n + N_{\pi}^{+},$$
 (3.58)

соответствующих выражений (для невырожденного полупроводника (3.26), (3.27) и (3.44)) и решении его относительно E_{ϕ} . Минимальные упрощения, которые допускают аналитическое решение трансцендентного уравнения, которое при этом получается, возможны для двух на первый взгляд крайних режимов:

а) в области низких температур можно пренебречь неосновными носителями ($p_n << N_{\rm g}^+$) и считать, что электроны в зону проводимости возбуждаются только за счет ионизации примесей, то есть

$$n_n = N_{\pi}^{+},$$
 (3.59)

б) в области высоких температур можно принять, что все примесные атомы ионизированы ($N_{a}^{+} = N_{a}$), а дополнительные носители образуются парами в результате межзонных переходов:

$$n_n = p_n + N_{\mathfrak{A}}. \tag{3.60}$$

На самом деле, как будет показано дальнейшими количественными расчетами, если энергия ионизации и ширина запрещенной зоны существенно отличаются, то в области истощения примеси выполняются оба эти условия: температуры уже достаточно высокие для полной ионизации всех донорных атомов и еще не такие высокие, чтобы учитывать межзонные переходы.

Воспользуемся этими допущениями для невырожденного полупроводника:

а) *Область низких температур*. Для упрощения вида получаемых результатов удобно ввести величину

$$n_1 = N_c \exp \left[(E_{\rm g} - E_c) / k_{\rm B} T \right] = N_c \exp \left(-\Delta E_{\rm g} / k_{\rm B} T \right),$$
 (3.61)

которую называют эффективной плотностью состояний, приведенной к примесному уровню, приписывая ей "физический" смысл, который вытекает из сравнения (3.26) и (3.61): эта величина будет равна концентрации электронов в зоне проводимости, если уровень Ферми будет совпадать с донорным уровнем. При этом уравнение электронейтральности (3.59), записанное в предположении несущественности процессов генерации электронно-дырочных пар, принимает вид

$$n_n = N_{\pi} / (n_n / n_1 + 1).$$
 (3.62)

Учитывая, что n_n связана с E_{ϕ} соотношением (3.26), из (3.62) можно получить значение энергии Ферми

$$E_{\phi} = E_{\pi} + k_{\rm B}T \ln \left(\sqrt{N_{\pi}/n_1 + 1/4} - 1/2\right)$$
 (3.63)

и концентрации электронов

$$n_n = N_{\rm A} / \left(\sqrt{N_{\rm A} / n_1 + 1/4} + 1/2 \right).$$
 (3.64)

При необходимости концентрация неосновных носителей, дырок в валентной зоне, может быть найдена из закона действующих масс:

$$p_n = n_i^2 / n_n,$$
 (3.65)

где величина n_i предварительно определяется из (3.55).

б) Область высоких температур. В этой области все примесные атомы ионизированы ($N_{\pi}^{+} = N_{\pi}$), а дополнительные носители образуются за счет межзонных переходов и концентрация электронов всегда превышает концентрацию дырок на величину N_{π} . Меняя в уравнении электронейтральности (3.60) концентрацию дырок на n_i^2/n_n в соответствии с законом действующих масс, получим уравнение $n_n^2 - N_{\pi}n_n - n_i^2 = 0$, решение которого для $n_n > 0$ будет иметь вид

$$n_{n} = \frac{N_{\pi}}{2} \left[1 + \sqrt{1 + \left(\frac{2n_{i}}{N_{\pi}}\right)^{2}} \right].$$
(3.66)

Из равенств (3.26) и (3.55) находится положение уровня Ферми:

$$E_{\phi} = E_c + k_{\rm B}T \ln\left\{\frac{N_{\pi}}{2N_c}\left[1 + \sqrt{1 + \left(\frac{2n_i}{N_{\pi}}\right)^2}\right]\right\} = E_c + k_{\rm B}T \ln\left(\frac{n_n}{N_c}\right). \quad (3.67)$$

При использовании последних соотношений для расчетов необходимо применять выражение (3.55) для предварительного находжения n_i . Расчет концентрации неосновных носителей может быть выполнен при известных n_i , n_n и E_{ϕ} различными способами: из уравнения электронейтральности $p_n = n_n - N_{a}$; из закона действующих масс $p_n = n_i^2/n_n$; из общего выражения (3.27); из выражения:

$$p_n = N_{\pi} \left[\sqrt{1 + (2n_i / N_{\pi})^2} - 1 \right] / 2,$$
 (3.68)

которое может быть получено так же, как и (3.66).

Формулы, полученные для низких и высоких температур, пригодны для расчетов во всем температурном диапазоне, так как в области истощения примеси выполняются оба принятых допущения, и результаты расчетов перекрываются (рис. 3.11, *a*): функция (3.64) стремится к $N_{\rm g}$ со стороны низких температур, а (3.66) – со стороны высоких.



Рис. 3.11. Использование двух приближений (3.59) и (3.60) приводит к формулам, которые дают результаты, перекрывающиеся в режиме насыщения примеси (*a*). В спрямляющих координатах (*б*) удобно использовать аппроксимацию в виде трех линейных участков.

Иногда может оказаться удобной менее точная, но более простая аппроксимация исследуемых зависимостей на трех характерных участках:

а) в области самых низких температур, где $k_{\rm B}T << \Delta E_{\rm g}$, или $n_1 << 4N_{\rm g}$, из (3.63) и (3.64) получаем соотношения

$$n_n = (N_c N_{\pi})^{1/2} \exp(-\Delta E_{\pi}/2k_{\rm B}T),$$
 (3.69)

$$E_{\phi} = (E_c + E_{\pi})/2 + k_{\rm B}T \ln (N_{\pi}/N_c)^{1/2}, \qquad (3.70)$$

подобные по форме (3.55) и (3.53) для собственного полупроводника, и функция (3.69) в координатах $\ln n_n = f(1/T)$ принимает вид прямой линии (участок 1 на рис. 3.11, б) с наклоном, пропорциональным энергии ионизации примеси ΔE_n ;

б) в области истощения примеси (участок 2) из (3.63) и (3.64) при $n_1 << N_{\rm g}$, и из (3.66), (3.67) при $2n_i << N_{\rm g}$ следует, что

$$n_n = N_{\rm g}, \qquad (3.71)$$

$$E_{\phi} = E_c + k_{\rm B}T \ln (N_{\rm A}/N_c),$$
 (3.72)

а из закона действующих масс получаем

$$p_n = n_i^2 / N_{\rm g};$$
 (3.73)

в) в области высоких температур преобладают межзонные переходы ($n_i >> N_{\pi}$), роль примеси становится несу3. Равновесное состояние носителей

щественной, полупроводник ведет себя как собственный, а участок 3 на рис. 3.11, б, так называемая область собственной проводимости, соответствует рис. 3.8, в.

Переходную область между участками 2 и 3 при необходимости можно описать более простыми соотношениями, чем (3.66) и (3.68), если в них ограничиться двумя членами разложения $\sqrt{1 + (2n_i/N_{\pi})^2}$ в степенной ряд:



$$n_n = N_{\rm g} + n_i^2 / N_{\rm g}, \quad p_n = n_i^2 / N_{\rm g},$$
 (3.74)

Рис. 3.12. Зависимости концентраций электронов и дырок в различных координатах (*a*, *б*, *в*), а также уровня Ферми (*z*) от температуры в германии ($\Delta E = 0,785$ эВ), легированном донорными примесями ($\Delta E_{\pi} = 0,01$ эВ) с концентрацией 10¹⁶ см⁻³ (1) и 10¹⁷ см⁻³ (2). Температурное изменение ширины запрещенной зоны не учитывалось. В круге изображен фрагмент зависимости $E_{\phi}(T)$ для самых низких температур в увеличенном масштабе.

На рис. 3.12 представлены в различных координатах результаты расчета температурных зависимостей концентрации основных и неосновных носителей и энергии Ферми для германия, легированного донорной примесью с энергией ионизации 0,01 эВ, по формулам, соответствующим двум допущениям (3.59) и (3.60). При расчете не принималось во внимание температурное изменение ширины запрещенной зоны, а также изменение энергетических донорных состояний при увеличении концентрации примеси.

Вид концентрационных зависимостей соответствует общей картине физических процессов, описанных в начале раздела. Надо также обратить внимание на одну особенность поведения неосновных носителей, которая наглядно проявляется в начале роста концентраций за счет собственной проводимости. Концентрация дырок в валентной зоне всегда меньше, чем ее значение в собственном полупроводнике, и это тем заметнее, чем выше концентрация доноров. Формально этот факт описывается законом действующих масс, а физически объясняется тем, что при увеличении концентрации примеси при прочих равных условиях увеличивается концентрация электронов, то есть заполненных состояний в зоне проводимости, что уменьшает вероятность межзонных переходов. Полезно обратить внимание на порядок значений концентрации неосновных носителей при комнатной температуре: для концентрации доноров (то есть и электронов в зоне проводимости) $N_{\pi} = 10^{15}$ см⁻³ концентрация дырок составляет $2 \cdot 10^{11}$ см⁻³ в германии и $1 \cdot 10^4$ см⁻³ в кремнии, а для $N_{\pi} = 10^{16}$ см⁻³ эти значения на порядок ниже: $2 \cdot 10^{10} \text{ cm}^{-3}$ и $1 \cdot 10^{3} \text{ cm}^{-3}$ соответственно.

Для температурной зависимости энергии Ферми (рис. 3.12, *г*) нужно отметить такие особенности. При T = 0 в соответствии с (3.70) уровень Ферми находится посредине между дном зоны проводимости и донорным уровнем. При повышении температуры из-за увеличения $N_c \sim T^{3/2}$ уровень Ферми приближается к зоне проводимости, потом ln $(N_{\rm A}/N_c)$ меняет знак, уровень Ферми опускается и, как следует из (3.63) при $E_{\rm p} = E_{\rm A}$, при температуре

$$T'_{s} = \Delta E_{\pi} / k_{\rm B} \ln \left(2N_{c} / N_{\pi} \right) \tag{3.75}$$

пересекает донорный уровень. Это значение часто принимают в качестве *температуры истощения примеси*. Однако надо помнить о условности этой границы, так как при температуре T'_s только половина примесных атомов ионизирована, что следует из (3.64) при $n_1(T'_s) = N_{\pi}/2$.

При использовании трехкусочной аппроксимации для такой температуры выбирают значение

$$T_s'' = \Delta E_{\pi} / k_{\rm B} \ln \left(N_c / N_{\pi} \right), \qquad (3.76)$$

что соответствует точке "излома" на рис. 3.11, б и следует из (3.69) при $n_n = N_{\pi}$.

При дальнейшем повышении температуры уровень Ферми отдаляется от дна зоны проводимости. Формально это подтверждается общим выражением (3.28), в котором соотношение p/n, оставаясь очень малым, растет за счет увеличения количества дырок при увеличении интенсивности межзонных переходов и практически постоянном значении концентрации электронов в режиме истощения примеси. При комнатной температуре уровень Ферми расположен ниже дна зоны проводимости на 0,23 эВ в Ge и на 0,26 эВ в Si для значения $N_{\rm g} = 10^{15}$ см⁻³ и соответственно на 0,17 эВ и на 0,20 эВ для концентрации примеси 10^{16} см⁻³.

В области высоких температур, где преобладают межзонные переходы, уровень Ферми практически совпадает с $E_{\Phi i}$ в собственном полупроводнике.

Верхней границей области истощения примеси считают такую температуру T_i , при которой значение n_i становятся соизмеримыми с $N_{\rm g}$, т.е. $n_i = \alpha N_{\rm g}$, где $\alpha \sim 1$. С учетом выражения (3.55) получим

$$T_i = \Delta E / k_{\rm B} \ln \left[N_c N_v / (\alpha N_{\rm g})^2 \right]. \tag{3.77}$$

Эта температура при $\alpha = 1$ соответствует точке "излома" между участками 2 и 3 на рис. 3.11, δ , а значения концентраций носителей, как следует из (3.66) и (3.68), равняются $n_n \approx 1.6$ N_{π} и $p_n \approx 0.6 N_{\pi}$. Ширина температурного интервала от T_s до

 T_i , в котором концетрация электронов остается постоянной, имеет значение при выборе материала для полупроводниковых приборов для уменьшения зависимости их парметров от температуры. Это относится в первую очередь к таким приборам, где решающую роль играют основные носители, ибо концентрация неосновных носителей в этом интервале хотя и является небольшой, но сильно зависит от температуры. Как следует из выражений для T_s и T_i , расширение этого температурного интервала возможно за счет: а) увеличения ширины запрещенной зоны, б) уменьшения энергии ионизации примеси и в) увеличения ее концентрации.

Например, при концентрации примеси 10^{15} см⁻³ в германии этот диапазон имеет границы от $T_s = 16$ К до $T_i = 360$ К при $\alpha = 0,1$ или до $T_i = 430$ К при $\alpha = 1$, а в кремнии от $T_s =$ = 34 К до $T_i = 540$ К ($\alpha = 0,1$) или до $T_i = 640$ К ($\alpha = 1$).

Проводимость примесных полупроводников, как и собственных, в общем случае обеспечивается носителями двух видов, поэтому

$$\sigma = en_n\mu_n + ep_n\mu_p. \tag{3.78}$$

Наличие примеси приводит к тому, что в отличие от собственных полупроводников к рассеиванию носителей, вызванному тепловыми колебаниями кристаллической решетки, прибавляется рассеивание на ионизированных и нейтральных (при самых низких температурах) атомах примеси, а температурная зависимость подвижности электронов и дырок описывается функциями типа (2.18). Степенной характер этих функций намного слабее, чем экспоненциальный для зависимостей $n_n(T)$ и $p_n(T)$ в областях ионизации примеси и собственной проводимости. Поэтому зависимость $\ln \sigma$ от 1/T, которая представлена на рис. 3.13, на этих участках практически прямолинейна. И только в области истощения примеси, когда концентрация основных носителей постоянна и равна концентрации доноров, а неосновных пренебрежительно мала, $\sigma = eN_n\mu_n$, и подвижность является главным фактором, кото-



Рис. 3.13. Зависимость удельной электропроводности кремния от обратной температуры при значениях концентрации донорной примеси: $1 - 5 \cdot 10^{23} \text{ м}^{-3}$, $2 - 3 \cdot 10^{24} \text{ м}^{-3}$.

рый определяет влияние температуры на электропроводность. В этой области температур примесное рассеивание играет второстепенную роль, а основным является механизм рассеивания электронов на тепловых колебаниях кристаллической решетки, при котором с ростом температуры подвижность, а значит и электропроводность падает изза увеличения интенсивности тепловых колебаний.

3.5.2. Акцепторные полупроводники

В полупроводнике, который содержит только акцепторные примеси, как уже отмечалось, процессы генерации в значительной мере подобны процессам в донорном полупроводнике с той разницей, что основными подвижными носителями здесь будут дырки в валентной зоне, которые образуются в первую очередь в результате перехода электронов из валентной зоны на свободные акцепторные уровни (переходы 2 на рис. 3.9). Акцепторные атомы в этом случае становятся отрицательно заряженными ионами. При высоких температурах дырки образуются и при межзонных переходах. Одновременно с этим генерируются подвижные электроны в зоне проводимости, которые для акцепторного полупроводника будут неосновными носителями.

Концентрацию подвижных носителей в зонах можно рассчитать по общим выражениям (3.14) и (3.24), а для невырожденных полупроводников по (3.26) и (3.27). Положение уровня Ферми может быть определено из уравнения электронейтральности, которое в данном случае имеет вид

$$p_p = N_a^{-} + n_p. \tag{3.79}$$

Здесь $N_{\rm a}^{-}$ – концентрация отрицательно ионизированных акцепторных атомов, которая определяется из (3.45), а индекс *p* около концентраций показывает на принадлежность к акцепторному полупроводнику.

Повторять решение уравнения электронейтральности нет необходимости. Удобнее воспользоваться отмеченной ранее симметрией между зоной проводимости и валентной зоной, а также между донорными и акцепторными состояниями. Действительно, изменение во всех функциях распределения направления отсчета энергии и замена соответствующих обозначений и индексов (*n* на *p* и наоборот, E_c на E_v , $N_{\rm g}$ на $N_{\rm a}$ и т.п.) не меняет их. Поэтому выражения для уровня Ферми и концетрации подвижных носителей в зонах для невырожденного акцепторного полупроводника могут быть получены указанной заменой из соответствующих выражений для донорного полупроводника.

Область низких температур:

$$E_{\phi} = E_{\rm a} - k_{\rm B}T \, \ln\left(\sqrt{N_{\rm a}/p_1 + 1/4} - 1/2\right),$$
 (3.80, a)

$$p_p = N_a / \left(\sqrt{N_a / p_1 + 1/4} + 1/2 \right),$$
 (3.80, 6)

$$p_1 = N_v \exp [(E_v - E_a)/k_B T] = N_v \exp (-\Delta E_a/k_B T)$$
. (3.80, *s*)

Область высоких температур:

$$E_{\phi} = E_{v} - k_{\mathrm{B}}T \ln\left\{\frac{N_{\mathrm{a}}}{2N_{v}}\left[1 + \sqrt{1 + \left(\frac{2n_{l}}{N_{\mathrm{a}}}\right)^{2}}\right]\right\} = E_{v} - k_{\mathrm{B}}T \ln\left(\frac{p_{p}}{N_{v}}\right), \quad (3.81, a)$$

$$p_p = N_{\rm a} \left(1 + \sqrt{1 + (2n_i/N_{\rm a})^2} \right) / 2$$
, (3.81, 6)

$$n_p = p_p - N_a = n_i^2 / p_p.$$
 (3.81, *s*)



Рис. 3.14. Те же зависимости, что и на рис. 3.12, для германия, легированного акцепторными примесями.

Эти выражения перекрываются, как и для донорного полупроводника, в области истощения примеси и достаточны для расчета концентраций во всем температурном диапазоне (рис. 3.14). Более простые выражения, пригодные для трех характерных областей, подобных (рис. 3.11), могут быть получены из (3.80) и (3.81) как частные случаи.

В заключение нужно отметить, что полупроводники, которые содержат только один вид примеси, практически не встречаются. Наряду со специально вводимыми примесями одного вида в полупроводнике из-за несовершенной очистки обязательно присутствуют и примеси другого вида. Иногда те и другие примеси вводят специально для придания полупроводнику нужных свойств. Наличие примесных атомов различного вида приводит к тому, что даже при отсутствии внешнего возбуждения происходит ионизация примесей из-за того, что электроны переходят из донорных состояний в энергетически более низкие акцепторные. При этом происходит так называемая компенсация примеси с меньшей концентрацией. При равенстве концентраций доноров и акцепторов происходит полная компенсация примесей и поведение такого полупроводника с точки зрения концентрации носителей и положения уровня Ферми во многом подобно поведению собственного полупроводника. При различных концентрациях примесей, например, при $N_{\rm q} > N_{\rm a}$, происходит частичная компенсация, а температурные изменения концентрации подвижных носителей и уровня Ферми в первом приближении можно считать такими же, как и в донорном полупроводнике, содержащем примеси с эффективной концентрацией, равной $N_{\rm q} - N_{\rm a}$.

3.6. РАВНОВЕСНЫЕ НОСИТЕЛИ В ЭЛЕКТРИЧЕСКОМ ПОЛЕ

Если в кристалле каким-либо внешним источником создано электрическое поле, характеризующееся некоторым потенциальным рельефом $V(\vec{r})$, то строго говоря зонной теорией пользоваться нельзя, так как наличие внешнего поля нарушает периодичность поля в кристалле. Однако ситуация существенно упрощается, если внешние поля намного слабее поля кристаллической решетки, имеющего напряженность порядка 10¹⁰ В/м. В таком случае внешнее поле хотя и нарушает периодичность поля решетки в целом, но локально в пределах некоторой области кристалла, включающей достаточно много периодов, это возмущение можно рассматривать как слабое, а потенциальную энергию электрона в поле $U(\vec{r}) = -eV(\vec{r}) -$ постоянной. Поэтому естественно считать, что решение уравнения Шредингера в суммарном поле будет приводить практически к такому же энергетическому спектру, что и в идеальной решетке, но смещенного локально на величину $U(\vec{r})$.



Рис. 3.15. Искривления потенциального рельефа и энергетических зон и изменение концентраций подвижных электронов и дырок при наличии внешнего электрического поля в полупроводнике.

Это означает, что в слабых внешних полях можно сохранить представление о энергетических зонах, но все уровни этих зон будут искривляться соответственно созданному в кристалле потенциальному рельефу $V(\vec{r})$ (рис. 3.15). Это искривление может быть описано пространственными зависимостями для краев зон:

$$E_{c}(\vec{r}) = E_{c0} + U(\vec{r}) = E_{c0} - eV(\vec{r}),$$

$$E_{v}(\vec{r}) = E_{v0} + U(\vec{r}) = E_{v0} - eV(\vec{r}),$$
(3.82)

где E_{c0} – дно зоны проводимости, а E_{v0} – потолок валентной зоны при отсутствии внешнего электрического поля, или, как говорят, для случая *плоских зон*.

Все сказанное выше относится также к внутренним *мак*роскопическим полям, которые могут быть созданы в результате перераспределения пространственного заряда. Характерными являются контактные поля, которые возникают в месте контакта двух различных веществ. В полупроводниковых материалах они имеют достаточно большую протяженность, поэтому будут выполняться условия, при которых можно использовать модель искривленных зон. В металлах они локализованы на расстоянии порядка периода решетки, их напряженность соизмерима с периодическим полем кристалла, и поэтому данная модель непригодна.

Если в таких полях носители не создают электрического тока, что возможно при отсутствии электрического контакта с источником внешнего напряжения или при наличии макроскопического внутреннего поля, то состояние электроннодырочного коллектива остается равновесным и значение энергии Ферми, как параметра функции распределения, также остается постоянным для всего объема. Но локальное расстояние между уровнем Ферми и границами зон меняется (рис. 3.15) из-за искривления энергетических зон по закону (3.82). Соответственно происходит и перераспределение подвижных носителей в зонах, их концентрация будет пространственно неоднородной и может быть описана выражениями (3.14) и (3.24), измененными в соответствии с (3.82):

$$n(\vec{r}) = N_c \Phi_{1/2}[\xi_0 + eV(\vec{r})/k_{\rm B}T],$$

$$p(\vec{r}) = N_v \Phi_{1/2}[\xi_{p0} - eV(\vec{r})/k_{\rm B}T].$$
(3.83)

Здесь ξ_0 и ξ_{p0} – приведенные уровни Ферми ξ и ξ_p в точке, где $V(\vec{r}) = 0$.

Для невырожденных полупроводников при использовании сооотношений (3.26) и (3.27), модифицированных в соответствии с (3.82), нетрудно получить

$$n(\vec{r}) = n_0 \exp[eV(\vec{r})/k_{\rm B}T],$$

$$p(\vec{r}) = p_0 \exp[-eV(\vec{r})/k_{\rm B}T],$$
(3.84)

где n_0 и p_0 – значения концентраций электронов и дырок в точке, где потенциал принимается равным 0, или, как иногда говорят, в режиме "плоских зон". Полученные выражения показывают, что подвижные электроны накапливаются во "впадинах" потенциального рельефа, т.е. в областях с повышенным потенциалом (рис. 3.15), а в "выпуклостях" их концентрация понижается. Поведение подвижных дырок, естественно, будет обратным, причем, как вытекает непосредственно из (3.84), произведение концентраций электронов и дырок во всех точках остается постоянным:

$$n(\vec{r}) \cdot p(\vec{r}) = n_0 \cdot p_0. \tag{3.85}$$

Это соответствует закону действующих масс и является одним из формальных признаков равновесного состояния.

Соотношение (3.84) для электронов было фактически получено в разделе 2.3.2 из условия равновесия диффузионных и дрейфовых токов при выполнении соотношения Эйнштейна. Действительно, в каждой точке пространства из-за наличия электрических полей и градиента концентраций носителей возникают их локальные дрейфовые и диффузионные потоки, но, как нетрудно показать это на основании выражений (3.84), они компенсируют друг друга, что и приводит к равновесному состоянию носителей. Интересно отметить, что это равновесие выполняется отдельно как для электронов так и для дырок.



Рис. 3.16. Контактная разность потенциалов, потенциальный барьер и распределение подвижных носителей вблизи контакта электронного и дырочного полупроводников.

Характерной и практически важной иллюстрацией подобной ситуации является распределение носителей вблизи контакта электронного и дырочного полупроводников. (рис. 3.16). Возникающая контактная разность потенциалов V_к порядка десятых долей вольт в приконтактной области $10^{-1}...10$ толшиной мкм создает электрическое поле с напряженностью порядка 10⁴...10⁶ В/м. Это электрическое поле создается в первую очередь ионизированными примесными атомами: положительно заряженными

донорами в n- и отрицательно заряженными акцепторами в p-полупроводнике. Распределение подвижных носителей в приконтактной области описывается выражениями (3.84), где n_0 и p_0 соответствуют равновесным концентрациям n_p и p_p в

126

акцепторном полупроводнике, если выбрать здесь начало отсчета потенциала. В *n*-области, где концентрации носителей n_n и p_n , потенциал равняется V_{κ} , поэтому

$$n_n = n_p \exp(eV_{\rm g}/k_{\rm B}T), \ p_n = p_p \exp(eV_{\rm g}/k_{\rm B}T).$$
 (3.86)

Отсюда могут быть получены два равноправных выражения для контактной разности потенциалов

$$eV_{\rm K} = k_{\rm B}T \ln (n_n/n_p) = k_{\rm B}T \ln (p_p/p_n).$$
 (3.87)

В приконтактной области, как это следует из (3.84), подвижных носителей обеих знаков будет меньше, чем ионизированных примесей, поэтому эта область, которую называют электронно-дырочным переходом, считается обедненной подвижными носителями и, как это будет показано далее, имеет нелинейные свойства, характерные для диодных структур.

3.7. ТЕРМОЭЛЕКТРОННАЯ ЭМИССИЯ

Термоэлектронной эмиссией называется выход электронов из вещества (твердого или жидкого) под действием нагрева. Сущность этого явления состоит в том, что с повышением температуры увеличивается энергия тепловых колебаний атомов вещества и, соответственно, тепловая энергия электронов. Несмотря на то, что среднее значение тепловой энергии невелико и составляет десятые доли электронвольт при температурах в тысячи кельвинов, все же очень малая часть электронов приобретает энергию, которая во много раз превышает среднюю и может оказаться достаточной для преодоления сил, препятствующих выходу электронов из вещества. Хотя таких электронов очень мало, но именно они обуславливают процесс эмиссии. По своей физической сущности термоэлектронная эмиссия во многом аналогична процессу испарения молекул из жидкости.

3.7.1. Закон термоэлектронной эмиссии

Природа сил, существующих у поверхности твердого тела и препятствующих эмиссии электронов, рассматривалась в первой главе. Эти силы создают у границы твердого тела с вакуумом потенциальный барьер, изображенный на энергетических диаграммах металла и полупроводника (рис. 3.17) и



Рис. 3.17. Энергетические диаграммы и функции распределения электронов в зоне проводимости полупроводника (*a*) и металла (*б*). На вставке приведены эти функции для вышедших и внутренних электронов в состояниях выше уровня вакуума. Заштрихованные области равны при термодинамическом равновесии.

форма которого может быть рассчитана по методу зеркального изображения. На этом же рисунке с энергетической диаграммой совмещена функция распределения электронов по энергиям, из которой следует, что при температуре абсолютного нуля все электроны заполняют состояния, лежащие в металлах ниже уровня Ферми, а в полупроводниках - в валентной зоне. В состояния выше уровня вакуума E_0 , из ко-

торых возможен их выход, электроны переходят при повышении температуры, и их количество, пропорциональное заштрихованной области, возрастает с температурой. Однако не все электроны, имеющие энергию $E > E_0$, могут принимать участие в эмиссии, так как скорости их направлены в разные стороны. Необходимо, чтобы энергия их движения к поверхности была достаточной для преодоления сил зеркального изображения ($v_x \ge v_{xmin}$). Значения параллельных поверхности составляющих скорости v_y и v_z не играют роли и могут быть любыми. Поскольку условия выхода формулируются для значений скоростей, то количество электронов, участвующих в эмиссии, не может быть определено по энергетической функции распределения. Для этого надо воспользоваться распределением по составляющим скорости. Однако неквадратичность закона дисперсии вдали от края зоны проводимости затрудняет нахождение этой функции распределения в состояниях вблизи E_0 .

Для вывода основного закона эмиссии воспользуемся следующим приемом. Рассмотрим электронный газ эмитированных электронов над поверхностью катода. Если внешнее электрическое поле не отводит электроны, то электронный газ в вакууме будет находиться в термодинамическом равновесии с газом в объеме полупроводника и уровни Ферми для них будут совпадать (рис. 3.17). При этом потоки электронов, эмитированных катодом и возвращающихся обратно, будут равны между собой, и плотность тока эмиссии легко найти, зная плотность потока возвращающихся частиц. Число частиц, попадающих на единицу поверхности за единицу времени, определяется в соответствии с известными результатами кинетической теории газов и равно $nv_{\rm r}/4$, где n – концентрация электронов в вакууме, а $v_{\rm r} = \sqrt{8kT/\pi m}$ их средняя тепловая скорость. Поэтому

$$j_s = env_{\rm T}/4 = en\sqrt{kT/2\pi m}$$
 . (3.88)

Использовать результаты кинетической теории газов в данном случае можно, так как уровень Ферми находится ниже уровня вакуума E_0 на величину $\Phi = E_0 - E_{\phi}$, обычно во много раз превышающую $k_{\rm B}T$, и электронный газ над поверхностью невырожден. По этой же причине его концентрацию легко рассчитать, как и концентрацию подвижных электронов в зоне проводимости невырожденного полупроводника, заменив в выражении (3.26) уровень E_c на минимальную энергию электронов в вакууме E_0 , а в (3.13) – эффективную массу на массу свободного электрона:

$$n = 2(2\pi m k T/h^2)^{3/2} \exp[-(E_0 - E_{\phi})/kT].$$
 (3.89)

Подставляя
 $\Phi = E_0 - E_{\phi}$ и (3.84) в (3.83), приходим к выражению

$$j_s = AT^2 \exp(-\Phi/kT), \qquad (3.90)$$

где $A = 4\pi mek^2/h^3 = 1,2\cdot 10^6 \text{ A/M}^2 \cdot \text{K}^2$. Выражение (3.90) называется формулой Ричардсона-Дешмана и является основным законом термоэлектронной эмиссии. Он характеризует зави-



Рис. 3.18. Зависимость плотности тока термоэлектронной эмиссии от температуры в естественных (*a*) и спрямляющих (*б*) координатах. Наклон прямой тем больше, чем больше работа выхода.

симость тока эмиссии от температуры и термодинамической работы выхода (рис. 3.18). График $j_s(T)$ показывает, что с температуры ростом плотность эмиссионного тока возрастает очень резко, так как эта зависимость определяется в основном экспоненциальным множителем. Ее, как и зависимость концентрации в собстполупроводнике венном (рис. 3.8, в), удобно представить в

логарифмическом масштабе от обратной температуры. Действительно, так как $\ln(j_s/T^2) = \ln A - \Phi/k_{\rm B}T$, то функция (3.90) в координатах рис. 3.18, б имеет вид прямой линии, угловой коэффициент которой равен $-\Phi/k_{\rm B}$, а точка ее пересечения с осью ординат – $\ln A$. Эта особенность используется для экспериментального определения термодинамической работы выхода Φ и постоянной A в уравнении (3.90).

Приведенный вывод закона термоэлектронной эмиссии в принципе пригоден как для металлов, так и полупроводников. Однако, несмотря на формальное совпадение этих законов, температурная зависимость тока эмиссии для полупроводников обычно намного сложнее, чем для металлов. И вызвано это тем, что термодинамическая работа выхода полупроводников может сильно зависеть от температуры. Для сравнительно простых случаев, например, для донорных и акцепторных полупроводников, рассмотренных в разделе 3.5, изменение уровня Ферми, а значит, и работы выхода может быть рассчитано. Однако для многих реальных полупроводниковых катодов, имеющих сложный состав примесей и структуру, такой расчет представляет трудности, и первостепенное значение приобретают экспериментальные методы исследования термоэлектронных свойств полупроводниковых катодов.

Интересно обратить внимание на тот факт, что в полупроводниках, как и в металлах, ток эмиссии определяется расстоянием между уровнем Ферми и уровнем вакуума. Это может на первый взгляд показаться странным, учитывая, что уровень Ферми лежит обычно в запрещенной зоне полупроводника и ни один электрон при выходе не совершает работы, равной Ф. Но если вспомнить, что энергия Ферми соответствует *средней* энергии электронов, участвующих в тепловом движении, и это среднее значение может находиться в запрещенной зоне, то станет понятным, почему в полупроводниках термодинамическая работа выхода отсчитывается от уровня Ферми.



Рис. 3.19. Влияние вида примеси на работу выхода. Донорные полупроводники (a) являются лучшими эмиттерами, чем акцепторные (δ).



Рис. 3.20. Влияние энергии электронного сродства на работу выхода. При равном расстоянии между уровнем вакуума и валентной зоной меньшую работу выхода имеет полупроводник с большей запрещенной зоной.

Рассмотрим, от каких факторов зависит работа выхода полупроводниковых материалов. Существенное влияние на работу выхода оказывают вид и концентрация примеси. Если рассмотреть энергетическую диаграмму примесного полупроводника (рис. 3.19), то станет ясным, что при прочих равных 3. Равновесное состояние носителей

условиях меньшей работой выхода обладают донорные полупроводники, так как уровень Ферми у них расположен выше, чем у акцепторных. По этой же причине увеличение концентрации донорной примеси, а также уменьшение энергии ионизации примесных атомов снижает работу выхода полупроводников.

Важным параметром полупроводниковых эмиттеров является энергия электронного сродства χ , то есть расстояние между уровнем вакуума и дном зоны проводимости. Полупроводник с меньшей энергией электронного сродства имеет и меньшую работу выхода, если, конечно, уровень Ферми одинаково расположен относительно дна зоны проводимости. Интересно также отметить, что из двух собственных полупроводников с одинаковым расстоянием между уровнем вакуума и потолком валентной зоны лучшим термоэлектронным эмиттером будет тот (рис. 3.20, *a*), который имеет меньшее электронное сродство. Второй (рис. 3.20, *б*) будет иметь лучшую электропроводность, так как при меньшей ширине запрещенной зоны концентрация носителей в нем будет выше.



Рис. 3.21. Наличие акцепторной примеси повышает вероятность ионизации адсорбируемой примеси.

На работу выхода полупроводников влияет присутствие адсорбированных атомов. Как и в металлах, пленка электроположительных атомов снижает работу выхода, создавая у поверхности электрическое поле, способствующее выходу электронов, а наличие электроотрицательных атомов "отравляет" эмиттер, повышая работу выхода. Надо отметить, что на свойства адсорби-

руемых атомов влияет характер примеси в полупроводнике. На рис. 3.21, *а* изображена ситуация, когда энергетическое состояние валентного электрона атома попадает в запрещенную зону и атом остается нейтральным, а возможна только его по3.7. Термоэлектронная эмиссия

ляризация за счет взаимодействия о поверхностными (например, таммовскими) состояниями полупроводника. Введение акцепторной примеси в полупроводник (рис. 3.21, δ) приводит к появлению свободных состояний, лежащих ниже уровня валентного электрона, и соответственно к ионизации адсорбируемого атома. При этом увеличивается электрическое поле у поверхности и снижение работы выхода будет более значительным, чем в случае беспримесного полупроводника.

Влияние адсорбированных атомов не исчерпывается изменением формы и высоты потенциального барьера "снаружи" полупроводника. Не меньшую роль играет также искривление энергетических уровней "внутри" под действием образующегося электрического поля. Причина этого влияния будет рассмотрена ниже.

3.7.2. Влияние электрического поля на работу выхода

Силы зеркального отображения, существующие у поверхности твердого тела и препятствующие выходу из него электронов, могут быть ослаблены внешним электрическим полем, если поле является ускоряющим, т. е. если его силовые линии направлены к поверхности катода. Это приводит к снижению потенциального барьера и соответственно к уменьшению работы выхода.

Для количественного анализа рассмотрим плоский металлический катод (рис. 3.22, *a*), у поверхности которого действует сила зеркального отображения, равная согласно (1.40) – $e^2/16\pi\varepsilon_0 x^2$, и сила электрического поля *e*E. Эти силы направлены навстречу друг другу и в точке

$$x_0 = \sqrt{e/16\pi\varepsilon_0 \varepsilon} . \qquad (3.91)$$

взаимно компенсируются, что непосредственно следует из равенства этих сил. Из рис. 3.22, б видно, что силы кристаллической решетки препятствуют выходу электрона только до



Рис. 3.22. Изменение силы (б) и потенциальной энергии (в) у поверхности металла при ускоряющем внешнем поле (а). В точке x_0 силы поля и зеркального изображения уравновешиваются, что соответствует максимуму потенциального барьера и его снижению на величину $\Delta \Phi$.

точки x_0 . На расстояниях, больших x_0 , поле будет ускоряющим. Потенциальная энергия электрона в таком поле будет определяться суммой энергий в поле решетки (1.41) и во внешнем поле $-e \varepsilon x$ (рис. 3.22, *в*):

$$U(x) = -e^2/16\pi\varepsilon_0 x - e\varepsilon x.$$
 (3.92)

Точка x_0 соответствует максимуму потенциальной энергии и определяет наибольшую высоту потенциального барьера. Так как значение $U(x_0)$ отсчитывается от уровня вакуума при отсутствии внешнего поля, то оно равно уменьшению работы выхода:

$$\Delta \Phi = U(x_0) = -e \sqrt{e \mathcal{E}/4\pi \varepsilon_0} . \qquad (3.93)$$

Это снижение барьера приводит к тому, что его сможет преодолеть большее число электронов, т. е. к увеличению тока термоэлектронной эмиссии:

$$j_s = j_{s0} \exp[(e/kT) \sqrt{e \mathcal{E}/4\pi \varepsilon_0}]. \qquad (3.94)$$

Здесь j_{s0} – ток эмиссии, определяемый по (3.90) при отсутствии внешнего поля. Такое явление называется эффектом Шоттки.

Подтверждением правильности полученных результатов служат представленные на рис. 3.23 экспериментальные результаты в координатах $\ln j_s$ и $\mathcal{E}^{1/2}$, спрямляющих зависимость (3.94). Отклонение экспериментальных точек от прямой

при слабых полях может быть объяснено тем, что пространственный заряд эмитируемых электронов экранирует катод от внешнего поля, а также более сложными эффектами, связанными с эмиссионной



Рис. 3.23. Влияние внешнего ускоряющего поля на ток термоэлектронной эмиссии. Единицы тока произвольны.

неоднородностью поверхности катода.

Присутствие у катода *тормозящего* электрического поля (рис. 3.24) не приводит к возрастанию потенциального барьера у поверхности, как это имело место для эффекта Шоттки. Ко-

личество электронов, достигающих анода, будет определяться в этом случае не силой зеркального отображения, локализованной у поверхности катода, а силой внешнего электрического поля, или высотой потенциального барьера $e|V_a|$, который создается в промежутке анодкатод отрицательным анодным потенциалом (рис. 3.24). Так как при энергии $E < e|V_a|$ эмитированные электроны возвращаются полем обратно на ка-



Рис. 3.24. Потенциальный барьер, созданный тормозящим полем анода, преодолевают электроны с энергией, превышающей высоту барьера.

тод, анодный ток будет создаваться только наиболее быстрыми электронами, количество которых равно интегралу от функции их распределения по энергиям в пределах $e|V_a| < E < \infty$. Поэтому

$$j_{a} \sim \int_{e|V_{a}|}^{\infty} N(E) dE = f(V_{a}), \qquad (3.95)$$

и вид вольт-амперной характеристики промежутка катоданод, получившей название кривой задержки, при задерживающих анодных потенциалах определяется видом функции распределения N(E). Так как $dj_a/dV_a \sim N(E)$, то, дифференцируя экспериментально измеренную кривую задержки, нетрудно определить вид функции распределения эмитированных электронов по энергиям. Этот метод, широко используемый при исследовании эмиссионных явлений, называется методом задерживающего потенциала.

Необходимо уточнить, что условие $E \ge e|V_a|$ пригодно только при сферической конструкции, когда радиус катода много меньше радиуса анода. В этом случае независимо от направления вылета из катода электроны летят по радиальным траекториям. В системе плоских электродов электрон сможет преодолеть тормозящее поле, если будет достаточной составляющая его скорости v_x в направлении силовых линий поля, или $m v_x^2/2 \ge e|V_a|$. Очевидно, что и дифференцирование



Рис. 3.25. Вольтамперная характеристика диода с металлическим катодом. При задерживающих анодных напряжениях ток обусловлен начальными скоростями вылета электронов. При ускоряющих наблюдается увеличение тока из-за эффекта Шоттки.

кривой задержки даст информацию только о распределении электронов по составляющей скорости v_x . Подобным образом можно показать, что в цилиндрической системе электродов вид кривой задержки зависит от распределения электронов только по радиальным составляющим скорости.

В итоге вольт-амперная характеристика вакуумного диода с металлическим катодом (рис. 3.25) при отрицательных анодных потенциалах должна

иметь вид возрастающей кривой, так как при увеличении $V_{\rm a}$ все большее число электронов преодолевает тормозящее поле. При положительных анодных напряжениях анодный ток воз-

растает за счет эффекта Шоттки. Надо помнить, что такой вид характеристика будет иметь, если движение отдельных электронов не зависит друг от друга. При больших плотностях тока на вид характеристики сильно влияет величина пространственного заряда, создаваемого самими электронами.

Эффект Шоттки проявляется и для полупроводниковых катодов. Однако картина физических явлений и их количественный анализ гораздо сложнее, так как в этом случае, кроме снижения потенциального барьера, наблюдается проникновение электрического поля в приповерхностные слои полупроводника. Из-за сравнительно малой концентрации подвижных носителей в полупроводниках заряды, компенсирующие внешнее электрическое поле (рис. 3.26, а), располагаются не у самой поверхности, как в металлах, а в приповерхностном слое толщиной в десятки и сотни периодов решетки. Соответственно и напряженность внешнего поля равна нулю не на поверхности



Рис. 3.26. Работа выхода из полупроводника во внешнем поле уменьшается за счет изгиба зон у поверхности и снижения потенциального барьера из-за эффекта Шоттки.

полупроводника, а постепенно убывает внутри этого слоя. Такое положение меняет не только форму потенциального барьера, но и, как отмечалось в разделе 3.6, вызывает искривление энергетических зон (рис. 3.26, *в*). Поэтому результирующее уменьшение работы выхода $\Delta \Phi$ определяется двумя составляющими: снижением барьера по Шоттки $\Delta \Phi_{\rm m}$ и величиной изгиба зон – $\Delta \Phi_{\rm m} = eV_s$.



Рис. 3.27. Адсорбированные электроположительные атомы деформируют потенциальный барьер и искривляют энергетические зоны у поверхности. В результате эффективная энергия электронного сродства уменьшается.

Подобное явление может быть вызвано не только внешним ускоряющим полем, но и наличием поверхностных состояний. Как уже отмечалось, для металлов присутствие на поверхности электроположительных адсорбированных атомов меняет форму потенциального барьера, снижая его. В полупроводниках к этому надо добавить искривление энергетических зон у поверхности, вызванное электрическим полем ионизированного

слоя адсорбированных атомов. Изменение работы выхода в таком случае учитывают обычно через изменение энергии эффективного электронного сродства $\chi_{\rm эф\phi}$, определяемого как расстояние от дна зоны проводимости в объеме до уровня вакуума (рис. 3.27). Искривление энергетических зон может быть вызвано не только адсорбированными атомами, но и таммовскими поверхностными состояниями, которые могут заполняться электронами из объема или отдавать их, образуя слой поверхностного заряда.

Величина эффективного электронного сродства зависит не только от характера поверхностных состояний, но и от вида примеси в полупроводнике, а также от ее концентрации. В полупроводниках *n*-типа более вероятным является такой обмен, когда электроны с донорных уровней переходят в лежащие ниже свободные поверхностные состояния и образуют двойной электрический слой, поле которого искривляет энергетические зоны вверх (рис. 3.28, *a*), повышая эффективное электронное сродство.

В акцепторных полупроводниках (рис. 3.28, *б*), как правило, наблюдаются обратные переходы электронов из запол-



Рис. 3.28. Влияние поверхностных состояний на энергию эффективного электронного сродства в зависимости от вида и концентрации примеси.

ненных поверхностных состояний на свободные акцепторные уровни, что вызывает изгиб зон вниз и снижает эффективное электронное сродство. Важно иметь в виду, что степень изгиба зон и глубина проникновения поля зависят от плотности пространственного заряда. Увеличение концентрации примеси приводит к большему изменению электронного сродства и одновременно уменьшает толщину слоя поверхностного пространственного заряда. При этом эффективное электронное сродство может статьи *отрицательным* (рис. 3.28, *в*). Стоит отметить также, что при большой плотности поверхностных состояний слой пространственного заряда весьма эффективно экранирует объем полупроводника от внешнего электрического поля. В таких случаях при ускоряющем поле эффект Шоттки проявляется так же, как и в металлах.

3.7.3. Термоэлектронные катоды

Явление термоэлектронной эмиссии используется для создания катодов в различных электронных приборах. Для изготовления катодов наибольшее распространение получил вольфрам, реже используются молибден, тантал, рений, цирконий и некоторые другие. Благодаря большой тугоплавкости рабочая температура вольфрамового катода может достигать 2500 К, при которой, несмотря на большую работу выхода (4,5 эВ), плотность тока достигает 0,3...0,5 А/см². Металлические катоды обладают малой эффективностью, так как требуют больших затрат для нагрева до высокой температуры. Поэтому область их использования ограничивается мощными лампами, работающими при высоких напряжениях, и приборами, требующими продолжительной стабильности эмиссионных свойств, которая не может быть обеспечена другими типами катодов.

Значительно большую эффективность имеют так называемые *пленочные катоды*. Они представляют собой металлическую поверхность, покрытую слоем адсорбированных атомов. Физические процессы, происходящие в такой структуре, были рассмотрены в гл. 1 и сводятся к следующему. В результате обмена электронами между пленкой адсорбированных атомов и металлом возникает двойной электрический слой, поле которого деформирует потенциальный барьер. Если адсорбированные атомы электроположительны по отношению к металлу, то силы электрического поля способствуют выходу электронов из металла, снижая потенциальный барьер и уменьшая работу выхода. Хоропими эмиссионными свойствами обладают пленки тория, бария, цезия и некоторых других веществ на поверхности вольфрама (табл. 3.1).

Таблица	3.1
---------	-----

Материал	Работа выхода, эВ	Постоянная <i>А</i> ·10 ⁻⁶ , А/м ² ·К ²	Рабочая темпе- ратура, К	Плотность тока эмиссии, A/cm^2
Вольфрам	4,5	0, 50, 6	2500	0,4
Молибден	4,3	0, 50, 6	2300	0,2
Тантал	4,1	0,40,6	2100	0,15
Торий на вольфраме	2,6	0,03	1800	0,4
Барий на вольфраме	1,6	0,15	1000	0,02
Оксидный катод	1,0	$10^{-5}10^{-4}$	1100	0,110

Из данных таблицы видно, что у пленочных катодов существенно уменьшается и постоянная *A* в уравнении (3.90). Это связано, во-первых, с сильной "пятнистостью" пленочных катодов, так как адсорбированные атомы образуют на поверхности катода пленки в виде отдельных групп атомов или пятен. Во-вторых, наличие атомов сильно меняет форму, а значит,

140

и прозрачность потенциального барьера и способствует тем самым уменьшению тока эмиссии.

Серьезным недостатком пленочных катодов, ограничивающим их использование, является невысокая стабильность эмиссии. При перегреве или при бомбардировке ионами остаточных газов поверхностный слой легко разрушается и увеличивается работа выхода.



Рис. 3.29. Структура (а) и энергетическая диаграмма (б) оксидного катода.

Наиболее распространенным видом термоэлектронных эмиттеров является оксидный катод. Его изготавливают обычно из смеси карбонатов бария, кальция и стронция, которые при нагреве в процессе изготовления превращаются в соответствующие окислы. Готовый оксидный катод представляет собой (рис. 3.29, а) пористый слой мелкодисперсной смеси окислов, нанесенный на металлическую подложку (керн). Между подложкой и оксидом в результате химического взаимодействия окислов и веществ, добавляемых в керн для придания катоду определенных свойств, образуется тонкий слой, который обладает очень высоким сопротивлением и поэтому называется запорным слоем. Основные процессы в оксидном катоде определяются окисью бария, энергетическая диаграмма которой представлена на рис. 3.29, б. Окись бария является полупроводником с шириной запрещенной зоны более 2 эВ и энергией электронного сродства около 1 эВ. Наличие акцепторных состояний, природа которых обычно связывается с неизбежными 3. Равновесное состояние носителей

в мелкозернистой структуре дефектами решетки, приводиг к тому, что работа выхода такого катода не может быть малой и составляет величину около 3 эВ. Для улучшения эмиссионных свойств еще на этапе изготовления прибора такой катод активируют. Активировка состоит в нагреве катода до температуры примерно в 1300 К. При разложении оксида образуются свободные атомы бария, которые по отношению к окиси бария являются донорными. Появление донорных атомов на начальном этапе активировки, во-первых, нейтрализует акцепторные состояния и, во-вторых, смещает уровень Ферми ближе к зоне проводимости. Кроме того, атомы бария на поверхности катода образуют пленку положительных ионов, что также снижает работу выхода. У готового активированного катода работа выхода равна примерно 1 эВ, что при рабочей температуре оксидного катода около 1100 К обеспечивает плотность тока эмиссии более 1 А/см².

Рассмотрим некоторые элементарные процессы, происходящие в оксидном катоде. В процессе термоэлектронной эмиссии катод покидают электроны, энергия которых превышает среднее значение. При этом равновесное максвелловское распределение термоэлектронов в зоне проводимости оксида нарушается, и электронный газ, а вместе с ним и катод, охлаждается. Одновременно происходит и обратный процесс разогрева проходящим анодным током. Этот разогрев неоднородный и наибольший там, где имеются участки с большим омическим сопротивлением (запорный слой и места контактов зерен оксида) или участки повышенной плотности тока (выступы на шероховатой поверхности катода).

В процессе работы оксидного катода количество атомов бария уменьшается в результате их испарения, что приводит прежде всего к образованию обедненного поверхностного слоя, обладающего повышенным по сравнению с объемом удельным сопротивлением. Восстановление бария происходит за счет термического разложения окиси бария, а также в результате электролитических процессов под действием электрического поля, возникающего в катоде при протекании тока. Ионы бария при этом движутся к подложке, там нейтрализуются и диффундируют по всему объему. Ионы и атомы кислорода, образующиеся при разложении оксида, а также остаточные газы (особенно галогены, пары воды) могут образовывать на поверхности катода слой отрицательно заряженных ионов, повышающих работу выхода и отравляющих катод. Отравление катода приводит к уменьшению тока эмиссии в процессе работы. Этому же способствует и втягивание ионов бария электрическим полем внутрь катода.

При больших отбираемых токах и пониженной температуре катода на поверхности можно наблюдать искрение, сопровождающееся отрывом частичек оксида. Это явление может быть связано с локальным перегревом отдельных участков катода, возникающим из-за неравномерности эмиссии с поверхности. Действительно, легко получающееся из (3.90) соотношение $dj_s/j_s = -d\Phi/k_{\rm B}T$ означает, что при рабочей температуре локальное уменьшение работы выхода даже на 0,1 эВ вызывает изменение тока с этого участка на величину $\Delta j_s \approx j_{s \, {\rm cp}}$, т. е. вдвое. Уменьшение температуры усиливает эту неравномерность, поэтому работа катода при недокале может привести к разрушению оксидного слоя.

Интересные особенности оксидного катода проявляются в импульсном режиме. Если на анод прибора подавать высокое напряжение в виде коротких (порядка микросекунд) и сравнительно редких импульсов, то ток эмиссии может возрастать в сотни и даже тысячи раз по сравнению с непрерывным режимом. Увеличение плотности тока вызывается сильным электрическим полем, возникающим в слое оксида при протекании большого тока и приводящим к такому изменению максвелловского распределения электронов, которое соответствует "разогреву" электронного газа. Немаловажную роль играет, вероятно, и освобождение пространственного заряда, накапливающегося в порах оксида.

Приведенное рассмотрение некоторых процессов, происходящих в оксидном катоде, не описывает всех его свойств, а лишь подчеркивает сложность явлений и отчасти объясняет, по-
чему до настоящего времени не существует точной теории катода, несмотря на многочисленность проводившихся исследований.

Приложение к гл. 1

П1. Модель Кронига-Пенни

Модель Кронига-Пенни является одной из поучительных моделей, позволяющих провести некоторые количественные оценки особенностей электронных состояний в периодическом поле. В качестве потенциального рельефа выбирается периодическая последовательность чередующихся прямоугольных барьеров и ям (рис. П1.1,a). При этом уравнение Шредингера с периодической функцией U(x) может быть представлено в виде уравнений с постоянными коэффициентами для отдельных областей:



циальный рельеф в модели Кронига-Пенни (*a*) и его предельное

представление (б).

I) в области ямы U = 0

$$\frac{d^2\psi}{dx^2} + \alpha^2 \psi = 0 , \qquad (\Pi 1.1)$$
$$\alpha^2 = 2mE / \hbar^2 ,$$

II) в области барьера $U = U_0 > E$

$$\frac{d^2\psi}{dx^2} - \beta^2 \psi = \mathbf{0}, \qquad (\Pi \mathbf{1.2})$$

 $\beta^2 \equiv 2m(U_0 - E) / \hbar^2 \, .$

Решения этих уравнений можно представить в виде

$$\psi_1(x) = A\cos\alpha x + B\sin\alpha x,$$

$$\psi_2(x) = C \mathrm{ch} \beta x + D \mathrm{sh} \beta x.$$
 (II1.3)

Но с другой стороны в периодическом поле функция $\psi(x)$ должна иметь вид волны Блоха $\psi(x) = \varphi(x)e^{ikx}$, где амплитуда $\varphi(x)$

обладает свойством периодичности $\varphi(x) = \varphi(x+a+s)$. Поэтому

$$\varphi_1(x) = (A\cos\alpha x + B\sin\alpha x)e^{-ikx}, \qquad \varphi_2(x) = (Cch\beta x + Dsh\beta x)e^{-ikx}. \qquad (\Pi 1.4)$$

На границах областей эти функции и их производные должны быть непрерывны. Из условия $\varphi_1(0) = \varphi_2(0)$ получаем, что A = C, а из условия

 $\phi'_1(0) = \phi'_2(0)$ следует $\alpha B = \beta D$. "Сшивание" на другой границе $\phi_1(a) = \phi_2(-b)$, $\phi'_1(a) = \phi'_2(-b)$ соответствует условию периодичности амплитуд и приводит к соотношениям:

$$\beta(A\cos\alpha a + B\sin\alpha a) = (\beta A ch\beta b - \alpha B sh\beta b)e^{ik(a+b)}, \qquad (\Pi 1.5, a)$$

$$\alpha(A\sin\alpha a - B\cos\alpha a) = (\beta A \sinh\beta b - \alpha B \cosh\beta b) e^{ik(a+b)}, \qquad (\Pi 1.5, \delta)$$

которые являются системой уравнений для нахождения констант A и B. Она имеет решение, если ее определитель равен 0, или

$$\frac{\beta^2 - \alpha^2}{2\alpha\beta} \sin\alpha a \cdot \mathrm{sh}\beta b + \cos ka \cdot \mathrm{ch}\beta b = \cos k(a+b). \tag{II1.6}$$

Данное выражение представляет собой *дисперсионное уравнение*, решение которого позволяет найти *закон дисперсии*, т.е. связь между энергией электрона, входящей в параметры α и β , и волновым числом *k*.

Упростить уравнение (П1.6) можно без потери общности результатов, увеличивая высоту барьеров ($U_0 \to \infty$) и одновременно их сужая ($b \to 0$), но так, чтобы произведение $U_0 b$, которое характерзует их прозрачность, оставалось бы постоянным (рис. П1.1,6). При этом $\beta b \to 0$ и дисперсионное уравнение можно представить в форме

$$\cos ka = f(\alpha(E)) \equiv P(\sin \alpha a / \alpha a) + \cos \alpha a, \qquad (\Pi 1.7)$$

где $P \equiv mabU_0/\hbar^2$. Представление о особенностях решения данного уравнения можно из рис П1.2, где изображена функция $f(\alpha(E))$ для P = 4. Поскольку $1 \ge \cos ka \ge -1$, то действительные решения возможны лишь в отмеченных областях значений *ka*.



Рис. П1.2. Графическое определение областей решения трансцендентного уравнения (П1.7).

Главным выводом из данной модели является утверждение о том, что энергетический спектр электронов представляет собой чередующиеся области (зоны) разрешенных и запрещенных состояний. В интервалах, где $|f(\alpha(E))| \leq 1$, существуют действительные значения k, которым соответствуют решения уравнения Шредингера в виде блоховских волн. В промежутках между ними решения (П1.7) будут мнимыми, а волновые функции становятся экспоненциально нарастающими и не удовлетворяют условиям нормировки для бесконечно длинного кристалла. Поэтому такие состояния образуют "запрещеные" зоны.



Рис. П1.3. Дисперсионные кривые в модели Кронига-Пенни(*a*). Верхние края разрешенных зон (*δ*) совпадают с уровнями квантования в бесконечно глубокой яме (*в*).

Более полно результаты модели Кронига-Пенни представлены на рис. $\Pi 1.3, a$ в виде дисперсионных кривых E(k) для каждой разрешенной энергетической зоны (б). Наличие в (П1.7) функции coska приводит к тому, что, изменяя значения k на целое число $2\pi/a$, получим те же значения энергии. Это означает, что дисперсионные кривые для каждой разрешенной зоны описываются периодическими и четными функциями. Верхние края зон соответствуют значениям $\alpha a = n\pi$ (n = 1, 2, ...). При этом $\cos ka = 1, k =$ $= n\pi/a$, а значения энергии $E_n = n^2 \hbar^2 \pi^2/2ma^2$ будут такими же, как и в бесконечно глубокой потенциальной яме шириной а (рис. П1.3, в). В этих состояниях электроны описываются волновыми функциями $\sin(n\pi x/a)$, характерными для стоячих волн. Значениям $\cos ka = -1$ соответствуют нижние края разрешенных зон и волновые функции типа $\cos(n\pi x/a)$, также описывающие стоячие волны. Это объясняется тем, что при этом на периоде решетки *а* укладывается целое число полуволн $a = n\lambda/2$ и отраженные от каждого барьера волны будут синфазно складываться. Серединам разрешенных зон соответствуют случаи, когда $f(\alpha(E))$ проходит через 0, а $ka = (2n - 1)\pi/2$, т.е. $a = (2n - 1)\lambda/4$ и на длине *а* укладывается нечетное число четвертей длин волн. В итоге отраженные от барьеров волны, будучи противофазными, взаимно гасятся и волновые функци будут иметь вид exp(*ikx*), характерный для бегущих волн.

На рис. П1.3 штриховой линией изображена также параболическая дисперсионная кривая для свободного электрона. Если прозрачность барьеров бесконечно велика (P = 0), то (П1.7) принимает вид $\cos ka = \cos \alpha a$. При этом разрешены любые значения $k = \alpha$ и $E = \hbar^2 k^2/2m$, как и для электрона в свободном пространстве.

Приложение к гл. 3

ПЗ. Интегралы Ферми-Дирака

Наиболее удобная и часто встречающаяся форма представления интегралов Ферми-Дирака имеет вид

$$\Phi_n(\xi) = \frac{1}{\Gamma(n+1)} \int_0^\infty \frac{\varepsilon^n d\varepsilon}{\exp(\varepsilon - \xi) + 1} \,. \tag{II3.1}$$

Здесь n – порядок интеграла, а $\Gamma(n+1)$ – так называемая гамма-функция, которая является обобщением понятия факториала $\Pi(n) = n! = \Gamma(n+1)$ для нецелых значений n. Для целых положительных значений n, как обычно, $\Gamma(n+1) = 1 \cdot 2 \cdot 3 \cdot ... \cdot n$; для $n = 3/2 - \Gamma(5/2) = 3\sqrt{\pi}/4$; для $n = 1/2 - \Gamma(3/2) = \sqrt{\pi}/2$; для $n = -1/2 - \Gamma(1/2) = \sqrt{\pi}$. Полезными являются свойство гамма-функции: $\Gamma(n+1) = n \cdot \Gamma(n)$, а также свойство интеграла Ферми-Дирака: $d\Phi_{n+1}(\xi)/d\xi = \Phi_n(\xi)$.

Графики интегралов Ферми-Дирака для некоторых значений n, рассматриваемых в гл. 3, представлены на рис. ПЗ.1. Из графиков видно, что в области достаточно больших отрицательных значений ξ интегралы Ферми-Дирака для всех порядков достаточно хорошо аппроксимируются функцией e^{ξ} .

Более подробно со свойствами интегралов Ферми-Дирака можно ознакомиться, например, в книге: Дж. Блекмор, "Статистика электронов в полупроводниках", пер. с англ., изд. Мир", Москва, 1964 г., 392 с. Там же приведены таблицы значений функций $\Phi(\xi)$.



Рис. II3.1. Зависимость интегралов Ферми-Дирака *n*-порядка от приведенного уровня Ферми. Для невырожденных коллективов удобна их аппроксимация функцией $\exp(\xi)$.

Учебное пособие

Москалюк Владимир Александрович Синекоп Юрий Степанович Кассинг Райнер

ФИЗИКА ЭЛЕКТРОННЫХ ПРОЦЕССОВ

часть I

электронные состояния

Под редакцией кандидата технических наук, профессора В. А. Москалюка

Компьютерная верстка С. Л. Волкович

Подписано к печати 17.11.2001. Формат 60×84 1/16. Гарнитура SchoolBook. Офсетная печать. Усл. печ. л. 9,25. Тираж 500 экз. Заказ 1487.

Издательско-полиграфический филиал УкрИНТЭИ, 03680, Киев, ул. Горького, 180