

Министерство образования Республики Беларусь
Белорусский национальный технический университет
Кафедра «Техническая физика»

ЭЛЕКТРИЧЕСКИЕ СВОЙСТВА МЕТАЛЛОВ И ПОЛУПРОВОДНИКОВ

Учебное пособие
для студентов БНТУ дневной и заочной форм обучения

Электронное учебное издание

Минск 2008

Составители:

В.И. Кудин, В.А. Мартинович

Рецензенты:

А.А. Баранов, доцент кафедры физики БНТУ;

Н.М. Казючич, заведующий лабораторией высокоэнергетичной имплантации и функциональной диагностики, кафедра физики БГУ

В пособии рассматриваются электрические свойства металлов и полупроводников, а также приведены необходимые сведения о квантовой электронной теории. Рассматриваются также физические принципы работы некоторых полупроводниковых приборов. Материал изложен в объеме программы курса физики для технического университета и окажет помощь студентам в освоении лекционного курса, а также в подготовке к выполнению лабораторных работ физического практикума по данному разделу физики.

Белорусский национальный технический университет
пр-т Независимости, 65, г. Минск, Республика Беларусь
Тел.(017) 232-77-52 факс (017) 232-91-37
Регистрационный № _____

© БНТУ, 2008

© В.И. Кудин, В.А. Мартинович

Содержание

<i>1. Элементы квантовой электронной теории металлов и полупроводников...</i>	4
Пределы применимости классической электронной теории металлов	4
Металлы, полупроводники и диэлектрики	5
Электронный газ. Распределение электронов по их энергиям.....	8
Внутренняя энергия и теплоемкость электронного газа в металлах.....	13
<i>2. Электропроводность металлов.....</i>	14
Сверхпроводимость.....	20
<i>3. Электропроводность полупроводников.....</i>	22
Собственная проводимость полупроводников.....	22
Примесная проводимость полупроводников.....	25
Электронно-дырочный переход.....	27
Р–n-переход при нарушении равновесия внешним электрическим полем....	30
Вольтамперная характеристика (ВАХ) р-n-перехода.....	32
Полупроводниковые приборы.....	33
Транзисторы.....	34
<i>4. Внутренний фотоэффект в полупроводниках. Фотопроводимость полупроводников.....</i>	38
Фотоэффект в р-n-переходе.....	40
Р-n-переход при отсутствии освещения.....	40
Р-n-переход при освещении.....	43
Вольтамперные характеристики р-n-перехода при освещении.....	44
<i>Литература.....</i>	46

1. Элементы квантовой электронной теории металлов и полупроводников

Пределы применимости классической электронной теории металлов

Объяснение различных свойств вещества существованием и движением в них электронов составляет содержание электронной теории.

В классической электронной теории металлов предполагается, что движение электронов подчиняется законам классической механики Ньютона. Далее, в этой теории пренебрегается взаимодействием электронов между собой, а взаимодействие электронов с положительными ионами кристаллической решетки сводится только к соударениям. Иными словами, электроны проводимости рассматриваются как электронный газ, подобный идеальному газу молекулярной физики. Несмотря на эти допущения, которые являются только весьма приближенными, классическая электронная теория качественно объясняет многие законы электрического тока. Например, классическая электронная теория хорошо объясняет существование электрического сопротивления металлов, законы Ома и Джоуля-Ленца. Она позволяет выразить удельную электропроводность через атомарные постоянные металла, а также выяснить связь между теплопроводностью и электропроводностью металла.

Однако в некоторых вопросах классическая электронная теория приводит к выводам, находящимся в противоречии с опытом. Наиболее ярким примером является сверхпроводимость. Вторым примером несостоятельности классической электронной теории служит теория теплоемкости металлов. Имеются и другие расхождения между выводами классической электронной теории и опытом.

Одна из причин расхождений заключается в том, что движение электронов в металлах подчиняется не законам классической механики, а более сложным законам квантовой механики. А эта последняя показывает, что движение электронов в периодическом поле потенциала решетки кристалла может существенно отличаться от предсказываемого классической механикой.

Далее, в классической электронной теории предполагается, что электроны проводимости подчиняются, как и атомы газа, законам статистики Максвелла-Больцмана. В действительности же для электронов внутри металла справедлива иная, квантовая статистика, и они подчиняются другому закону распределения.

И, наконец, в классической электронной теории не учитывается взаимодействие электронов друг с другом, а их взаимодействие с решеткой металла описывается с помощью представления о кратко временных соударениях. Между тем при очень низких температурах взаимодействие между электронами может играть решающую роль.

Было бы, однако, ошибкой считать, что классическая электронная теория утратила свое значение вовсе. Она позволяет во многих случаях быстро найти правильные качественные результаты и притом в наглядной форме. Более того, оказывается, что расхождения между теорией и опытом тем меньше, чем меньше концентрация электронов проводимости и чем выше температура. В металлах, где концентрация электронов велика, эти различия имеют существенное зна-

чение. В ряде же других случаев, где концентрация электронов мала (электронные явления в газах, многие явления в полупроводниках), классическая электронная теория применима не только качественно, но и количественно.

Металлы, полупроводники и диэлектрики

Рассмотрим процессы, приводящие к образованию электронов проводимости в проводниках. Для выяснения этого вопроса большое значение имеет исследование зависимости концентрации электронов проводимости от температуры. Эти данные можно получить, измеряя, например, значение постоянной Холла при различных температурах.

Опыт показывает, что в металлах концентрация электронов практически не зависит от температуры. Даже при самых низких температурах в металлах уже имеется большое количество подвижных электронов. Это показывает, что в образовании электронов проводимости в металлах тепловое движение не играет существенной роли.

Атомы типичных металлов характеризуются тем, что в них имеется один или несколько электронов, связанных с ядром слабо. При сближении атомов металла такие электроны под действием сил взаимодействия с соседними атомами отщепляются от своих атомов. Они принадлежат уже не какому-либо отдельному атому, а всему металлу в целом и движутся по металлу в результирующем поле всех ионов и электронов. Эти отщепившиеся электроны и являются электронами проводимости.

Наряду с металлами мы встречаемся с проводниками и другого типа. Эти проводники являются так же как и металлы электронными, и в них электрический ток не сопровождается никакими химическими изменениями. Однако концентрация носителей заряда в таких проводниках чрезвычайно сильно увеличивается с увеличением температуры. Подобные проводники при низких температурах имеют весьма большое удельное сопротивление и практически являются изоляторами, но с увеличением температуры их удельное сопротивление сильно уменьшается и при достаточно высоких температурах становится весьма малым. Вещества такого типа получили название *электронных полупроводников*.

Полупроводниками являются многие элементы (кремний, германий, селен и др.), закись меди Cu_2O , сернистый свинец PbS и другие химические соединения. Так, например, по данным опыта можно заключить, что концентрация электронов в чистейшем кремнии при комнатных температурах должна быть меньше 10^{17} м^{-3} , а его удельное сопротивление должно быть больше $10^3 \text{ Ом} \cdot \text{м}$, но при температуре 700°C концентрация электронов в нем возрастает до 10^{24} м^{-3} , а удельное сопротивление падает до $0,001 \text{ Ом} \cdot \text{м}$, т. е. больше чем в миллион раз.

Сильная зависимость концентрации носителей заряда в полупроводниках от температуры показывает, что в этом случае электроны проводимости возникают под действием теплового движения. В полупроводниках атомное взаимодействие само по себе еще не достаточно для отщепления электронов от атомов и превращения их в электроны проводимости. Для этого даже наиболее слабо связанным электронам нужно сообщить некоторую добавочную энергию, которая заимствуется

из энергии теплового движения. Чем выше температура, тем большее число электронов будет существовать в полупроводнике в отщепленном состоянии, т. е. в виде электронов проводимости.

Если величина энергии отщепления велика по сравнению со средней энергией теплового движения (имеющей порядок kT) при всех температурах в области существования данного кристалла, то электроны проводимости в заметном количестве не образуются, и такой кристалл будет изолятором.

Процессы образования электронов проводимости в кристаллических твердых телах объясняются квантовой теорией твердого тела. Решение задачи о движении электрона в поле периодического потенциала решетки показывает, что существует целая система зон разрешенных энергий. Каждая из зон ограничена снизу некоторой энергией E_C («дно» зоны) и сверху энергией E_V («потолок» зоны). Эти зоны отделены друг от друга полосами запрещенных энергий. Ширина разрешенных зон энергии увеличивается по мере возрастания энергии, и широкие верхние зоны в некоторых случаях могут перекрывать друг друга, образуя единую сложную зону.

Происхождение энергетических зон имеет следующую общую причину. Положим, что мы имеем N изолированных атомов, удаленных друг от друга и поэтому не взаимодействующих между собой. В каждом из этих атомов энергия электронов может изменяться лишь скачками и поэтому характеризуется набором дискретных уровней энергии. В рассматриваемой системе невзаимодействующих атомов вместо каждого атомного энергетического уровня имеется N совпадающих уровней энергии. Представим себе теперь, что мы сближаем атомы до образования кристаллической решетки. При этом возникает взаимодействие между атомами, и уровни энергии электронов изменяются. Оказывается, что в результате такого взаимодействия первоначально совпадающие N уровней энергии становятся различными. Эта совокупность теперь уже несовпадающих уровней энергии и есть разрешенная зона энергий. Таким образом, энергетические зоны возникают в результате расщепления дискретных уровней энергии электрона в атомах под влиянием взаимодействия атомов в решетке (рис. 1.1).

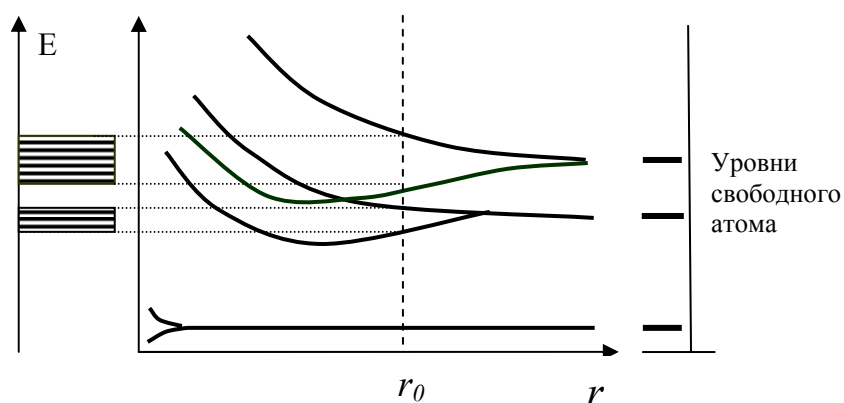


Рис. 1.1. Расщепление энергетических уровней в зависимости от расстояния r между атомами и образование разрешенных и запрещенных энергетических зон в твердом теле (заштрихованные области – разрешенные зоны, незаштрихованная область – запрещенная зона; r_0 – равновесное расстояние между атомами в твердом теле).

Число энергетических уровней в каждой зоне очень велико – порядка числа атомов в кристалле, и уровни расположены очень близко друг к другу. Поэтому во многих случаях можно считать, что внутри каждой зоны энергия электрона, как и в классической теории, изменяется непрерывно. Однако число этих уровней конечно, что имеет принципиальное значение.

Задавая тройку возможных значений компонент импульса электрона p_x, p_y, p_z и определенную ориентацию спина электрона, мы определяем некоторое возможное стационарное движение электрона, или, иначе, *квантовое состояние* электрона. Одна и та же энергия электрона может соответствовать различным значениям p_x, p_y, p_z и разным ориентациям спина, и следовательно, каждому энергетическому уровню внутри зоны соответствует целая совокупность квантовых состояний.

При заданном состоянии кристалла электроны в нем определенным образом распределяются по различным квантовым состояниям. Однако такое распределение всегда подчиняется *принципу Паули*: в любой электронной системе – атомах, молекулах, твердых телах – *в каждом квантовом состоянии не может быть более одного электрона*. Это обстоятельство приводит к тому, что число электронов, энергии которых лежат в пределах какой либо разрешенной зоны, ограничено, т. е. *вместимость каждой энергетической зоны конечна*.

Положим, что кристалл состоит из N атомов и каждый атом в изолированном состоянии имеет ξ электронов. Температуру кристалла будем считать равной абсолютному нулю. Тогда имеющиеся $N\xi$ электронов будут заполнять различные квантовые состояния в порядке возрастания их энергий. Вследствие конечной вместимости зон часть низко расположенных зон энергии окажется целиком заполненной электронами, а следующая за ними зона – либо заполненной частично, либо пустой. Более высокие зоны будут наверняка пустыми.

Конечная вместимость энергетических зон существенно проявляется в процессах переноса электрического заряда в электронном газе. Положим, что, желая создать электрический ток, мы ускорим электроны электрическим полем. При этом мы увеличиваем их энергию и переводим в новые квантовые состояния. На энергетических диаграммах это соответствует возбуждению электронов, т. е. переходу их с первоначальных энергетических уровней на другие, более высокие уровни внутри зоны. Но тогда из сказанного выше следует, что электроны целиком заполненной зоны, хотя и находятся в движении, не могут дать никакого вклада в электрический ток. Действительно, так как энергия электрона есть четная функция импульса, то каждому квантовому состоянию электрона в зоне с какой-либо компонентой импульса, скажем p_x , обязательно соответствует и другое состояние с той же самой энергией, но с компонентой импульса $-p_x$. Компоненты скорости этих электронов направлены в противоположные стороны. Поэтому электроны целиком заполненной зоны можно разбить на пары, создающие противоположно направленные токи, и следовательно, результирующий ток, создаваемый такой совокупностью электронов, всегда равен нулю. Чтобы создать отличный от нуля электрический ток, мы должны были бы возбудить электроны внутри зоны и перевести часть из них на более высокие уровни. Однако, так как в целиком заполненной зоне все имеющиеся квантовые состояния заняты, то вследствие принципа Паули это оказывается невозможным.

Таким образом, для электропроводности существенны только две энергетические зоны: та, которая при температуре абсолютного нуля заполнена лишь частично или пуста, и лежащая непосредственно под нею целиком заполненная зона. Последняя может тоже участвовать в электропроводности, так как при повышении температуры часть электронов, находящихся под потолком этой зоны, может переходить в выше расположенную незаполненную зону, и поэтому она становится уже не целиком заполненной. Обе эти зоны и представляют собой *зону проводимости* и *валентную зону*. Незанятые же электронами квантовые состояния, энергия которых лежит вблизи потолка валентной зоны, соответствуют *дыркам*.

Из сказанного следует, что металлы – это такие кристаллы, у которых при температуре абсолютного нуля одна из энергетических зон заполнена частично. Типичными же полупроводниками являются те кристаллы, энергетические зоны которых при абсолютном нуле либо заполнены целиком, либо совершенно пусты. На рис. 1.2 представлены энергетические схемы чистых кристаллов металлов, диэлектриков и полупроводников, которые отличаются только шириной запрещенной зоны ΔE . Так, у металлов запрещенная зона отсутствует, у полупроводников ширина запрещенной зоны невелика (у собственных полупроводников не более 1 эВ), а у диэлектриков ширина запрещенной зоны ΔE велика (порядка нескольких электронвольт).

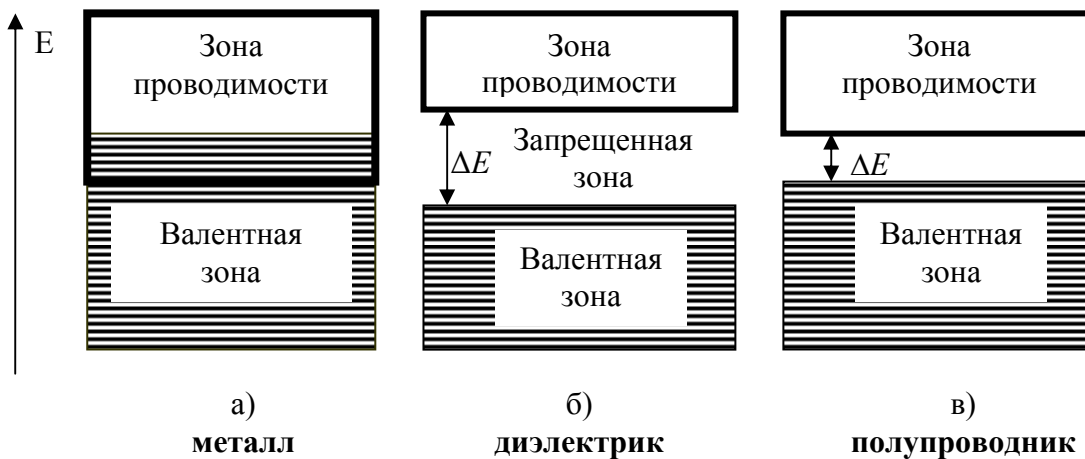


Рис. 1.2. Структура энергетических зон в металлах, диэлектриках и полупроводниках.

Электронный газ. Распределение электронов по их энергиям

Основной задачей квантовой статистики является определение числа частиц $dn(E)$ квантового газа, обладающих значениями энергий от E до $E + dE$. Это число частиц равно произведению числа квантовых состояний в интервале энергий dE на среднее число частиц, приходящихся на каждое квантовое состояние, т. е. выражение для $dn(E)$ имеет следующий вид:

$$dn(E) = g(E)f(E)dE, \quad (1.1)$$

где $g(E)$ – плотность числа квантовых состояний, которая в случае электронного газа равна

$$g(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}}, \quad (1.2)$$

$f(E) = \frac{n_i}{g_i}$ – функция распределения частиц по квантовым состояниям или среднее число частиц в квантовых состояниях с энергией от E до $E + dE$.

Вид функции $f(E)$ зависит от природы частиц газа или, другими словами, определяется их спином. Так, распределение частиц с полуполным спином описывается функцией Ферми-Дирака, с нулевым и целым спином – функцией Бозе-Эйнштейна.

Поскольку электроны обладают спином $s = 1/2$, электронный газ описывается функцией распределения Ферми-Дирака, которая имеет вид:

$$f(E) = \frac{1}{e^{\frac{E-\mu}{kT}} + 1}, \quad (1.3)$$

где μ – некоторая характерная энергия, не зависящая от переменных E и p . Она получила название **химического потенциала**, величина которого зависит от природы вещества и его состояния. Для данного вещества μ определяется полной концентрацией электронов и температурой (см. ниже).

Вид функции распределения Ферми-Дирака (1.3) существенно зависит от температуры электронного газа. При $T \rightarrow 0K$ для энергий $E < \mu$ имеем, что $\exp\left(\frac{E-\mu}{kT}\right) \rightarrow 0$ и, следовательно, $f(E) \rightarrow 1$, т. е. в каждом квантовом состоянии находится по одной частице. В случае $E > \mu$ при $T \rightarrow 0K$ имеем, что $\exp\left(\frac{E-\mu}{kT}\right) \rightarrow \infty$ и, следовательно, $f(E) \rightarrow 0$, т. е. квантовые состояния в этом случае свободны (в них нет ни одной частицы).

Вид функции Ферми-Дирака при $T = 0K$ показан на рис. 1.3, б).

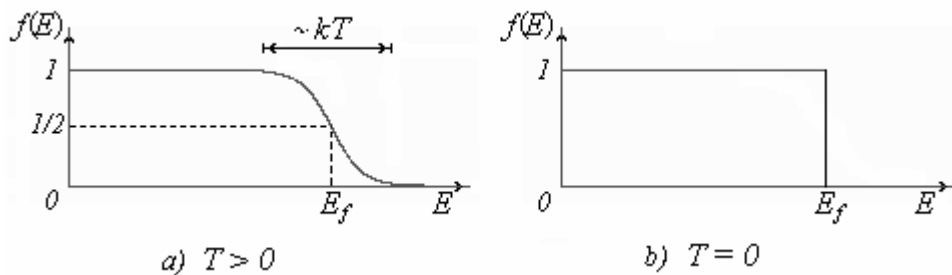


Рис. 1.3. Функция распределения Ферми-Дирака.

Такой вид функции распределение обуславливается необходимостью соблюдения двух требований. Во-первых, полная энергия должна быть минимальна и, во-вторых, должен соблюдаться принцип Паули. Поэтому электроны начинают заполнять квантовые состояния с самого нижнего энергетического уровня, последовательно заполняя их. Последний электрон занимает уровень с максимальной энергией, который называется **уровнем Ферми**, а энергия уровня – **энергией Ферми** E_f . Однако такое наглядное определение энергии Ферми имеет смысл лишь в применении к свободным электронам в металле. Например, в полупроводниках и диэлектриках энергия Ферми приходится примерно на середину запрещенной зоны, и заведомо нет электронов, которые обладают такой энергией. Поэтому в общем случае **энергией Ферми E_f называется такая энергия, при которой функция распределения Ферми-Дирака равна $\frac{1}{2}$.**

Как видно из выражения (1.3) и рис. 1.3 энергия Ферми $E_f = \mu$. Поэтому функцию распределения Ферми-Дирака можно ещё записать в виде:

$$f(E) = \frac{1}{e^{\frac{E-E_f}{kT}} + 1}. \quad (1.4)$$

При $T > 0K$ функция распределения Ферми-Дирака размывается в окрестности энергии Ферми (рис. 1.3, а). Область размывости составляет величину порядка kT .

В случае больших энергий, когда $(E - \mu) \gg kT$ (что выполняется в области “хвоста” кривой распределения) единицей в знаменателе можно пренебречь. Тогда функция распределения по состояниям с различной энергией принимает вид:

$$f(E) = e^{-\frac{(E-\mu)}{kT}} = const e^{-\frac{E}{kT}},$$

т.е. переходят в классическое распределение Больцмана.

Обратимся к распределению электронов по их энергиям, которое с учетом (1.1) – (1.3) определяется выражением:

$$dn(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \frac{E^{\frac{1}{2}} dE}{e^{\frac{E-E_f}{kT}} + 1}. \quad (1.5)$$

График этого распределения при различных температурах показан на рис. 1.4.

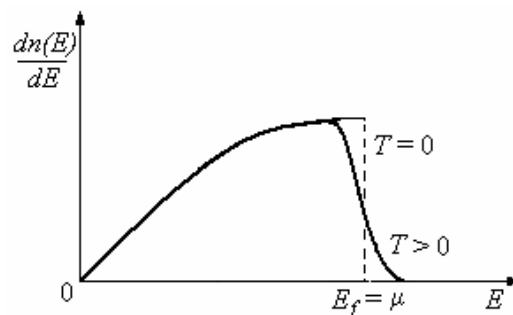


Рис. 1.4. Распределение электронов в металле по энергиям.

При $T = 0K$ выражение (1.5) принимает следующий вид:

$$dn(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} dE. \quad (1.6)$$

Данная зависимость позволяет легко вычислять значения характерных параметров электронного газа, поскольку при $T = 0K$ в распределении присутствуют только электроны с энергиями $E < E_f$ и, следовательно, интегрирование сводится к пределам от 0 до E_f .

Так, концентрация электронов в металле равна:

$$n = \int_0^{E_f} dn(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \int_0^{E_f} E^{\frac{1}{2}} dE = \frac{1}{3\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} E_f^{\frac{3}{2}}. \quad (1.7)$$

Отсюда можно получить зависимость энергии Ферми от концентрации электронов:

$$E_f = \frac{\hbar^2}{2m} (3\pi^2 n)^{\frac{2}{3}}. \quad (1.8)$$

Интегрирование выражения (1.6) позволяет также вычислить среднюю энергию электронов в металле при $T = 0K$:

$$\langle E \rangle = \frac{1}{n} \int_0^{E_f} E \cdot dn(E) = \frac{1}{\frac{1}{3\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} E_f^{\frac{3}{2}}} \int_0^{E_f} E^{\frac{3}{2}} dE = \frac{3}{5} E_f. \quad (1.9)$$

От распределения электронов по энергиям можно перейти к распределению электронов по скоростям. Для этого нужно в (1.6) подставить выражение для энергии $E = \frac{mV^2}{2}$, а также выражение для дифференциала $dE = mV dV$. В результате чего получим

$$dn(v) = \frac{m^3}{\pi^2 \hbar^3} v^2 dv. \quad (1.10)$$

При интегрировании выражения (1.10) нужно пользоваться пределами изменения скорости электронов от 0 до v_{\max} , где v_{\max} определяется энергией Ферми $E_f = \frac{mV_{\max}^2}{2}$. Полученное в результате выражение устанавливает связь между концентрацией и максимальной скоростью электронов:

$$n = \int_0^{v_{\max}} dn(v) = \frac{m^3 v_{\max}^3}{3\pi^2 \hbar^3}. \quad (1.11)$$

Тогда средняя скорость электронов в металле равна:

$$\langle v \rangle = \frac{1}{n} \int_0^{v_{\max}} v \cdot dn(v) = \frac{3}{v_{\max}^3} \int_0^{v_{\max}} v^3 dv = \frac{3}{4} v_{\max}. \quad (1.12)$$

Энергия Ферми, определяемая химическим потенциалом, не зависит от температуры только в первом приближении. При высоких температурах температурная зависимость начинает сказываться. Учет такой зависимости дается последующим членом разложения μ по степеням kT :

$$\mu = \mu_0 \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\mu_0} \right) + \dots \right] \quad (1.13)$$

Зависимость энергии Ферми от температуры становится несущественной, когда член в круглых скобках в правой части выражения (1.13) стремится к единице, т. е. при температуре

$$T_f = \frac{\mu_0}{k},$$

которая называется *характеристической температурой* или *температурой Ферми*.

Поведение электронного газа в сильной степени зависит от соотношения между температурой кристалла и температурой Ферми. Различают два предельных случая:

1. $kT \ll E_f$. В этом случае электронный газ, называется **вырожденным квантовым ферми-газом**.

2. $kT \gg E_f$. В этом случае электронный газ, называется **невырожденным квантовым ферми-газом**.

Например, для металлов температура Ферми много больше температуры их плавления. Поэтому для большинства металлов в твердом состоянии в качестве энергии Ферми можно брать μ_0 , а распределение электронного газа в них очень мало отличается от распределения Ферми-Дирака при 0 К, т. е. в металлах электронный газ является вырожденным ферми-газом.

В полупроводниках плотность свободных электронов оказывается много меньшей, чем в металлах, соответственно, энергия Ферми E_f мала ($E_f \sim n^{2/3}$) (см. (1.8)). Поэтому уже при комнатной температуре электронный газ во многих полупроводниках является невырожденным и подчиняется классической статистике.

Внутренняя энергия и теплоемкость электронного газа в металлах

Для определения средней энергии электронов при $T > 0K$ необходимо пользоваться выражением (1.5) со значением μ , задаваемым формулой (1.13). Вычисления приводят к следующему результату:

$$\langle E \rangle = \frac{3}{5} \mu_0 \left[1 + \frac{5}{12} \pi^2 \left(\frac{kT}{\mu_0} \right)^2 + \dots \right]. \quad (1.14)$$

Тогда **внутренняя энергия** электронного газа в металле равна

$$U = N \langle E \rangle = \frac{3}{5} N \mu_0 + \frac{1}{4} \frac{\pi^2 k^2 N}{\mu_0} T^2, \quad (1.15)$$

где N – общее число электронов в металле. Если N равно постоянной Авогадро N_A , то U есть внутренняя энергия моля электронов. Тогда молярная теплоемкость при постоянном объеме равна

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = \frac{\pi^2}{2} \frac{kT}{\mu_0} k N_A = \frac{\pi^2}{2} \frac{kT}{\mu_0} R. \quad (1.16)$$

По классическому закону Дюлонга и Пти для твердых тел $C_V = 3R$, что неизмеримо больше, чем значение, полученное с помощью выражения (1.16), поскольку при нормальных условиях $kT \ll \mu_0$. Это означает, что **теплоемкость металлов за счет электронов пренебрежимо мала**. Физически это обусловлено тем, что при обычных температурах в тепловом движении принимает участие лишь небольшая часть электронов, а, именно, те, которые находятся вблизи уровня Ферми. Таким образом, вырожденный электронный газ не ведет себя аналогично обычному газу, и, в частности, его вклад в теплоемкость нельзя рассчитывать, применяя к нему теорему о равномерном распределении энергии по степеням свободы.

Однако при достаточно низких температурах электронная теплоемкость превосходит теплоемкость, обусловленную колебаниями атомов решетки твердого тела, поскольку последняя уменьшается как $\sim T^3$ и при очень низкой температуре становится пренебрежимо малой.

2. Электропроводность металлов

Классическая теория электронного газа позволяет оценить скорости свободных электронов в металле при комнатной температуре. Считая электронный газ идеальным, вычислим средне-квадратичную скорость электронов:

$$v_m = \sqrt{\frac{3kT}{m}} \approx 1,2 \cdot 10^5 \frac{m}{c},$$

где k – постоянная Больцмана, m – масса свободного электрона. Это скорость хаотического теплового движения, которое не приводит к направленному переносу зарядов и возникновению электрического тока в образце.

В металле изменять свою энергию (ускоряться электрическим полем) могут лишь электроны, имеющие энергию вблизи уровня Ферми, поэтому важно оценить тепловую скорость именно таких электронов. Так, например, для меди

$$v_m = v_f = \sqrt{\frac{2E_f}{m}} \approx 1,6 \cdot 10^6 \frac{m}{c} \quad (E_f = 7,07 \text{ эВ}).$$

Чтобы вычислить силу тока, проходящего по проводнику, необходимо знать скорость, которую приобретают электроны под действием электрического поля. Эта скорость называется **дрейфовой скоростью**, и хотя значение этой скорости намного меньше тепловой, именно она определяет силу тока в проводнике.

По определению, сила тока в проводнике I равна заряду dq , протекающему через поперечное сечение проводника S за время dt : $I = \frac{dq}{dt}$. Заряд dq равен заряду электрона e , умноженному на число электронов N , находящихся в объеме Sdx (рис. 2.1).

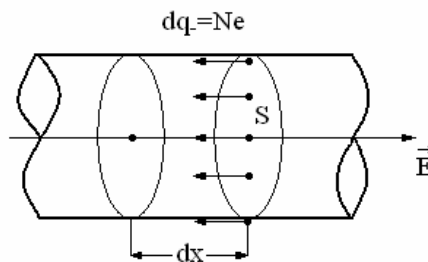


Рис. 2.1. К установлению связи между плотностью электрического тока и дрейфовой скоростью электронов.

Число электронов можно представить в виде $N = nSdx$, где n – концентрация электронов в металле. Тогда сила тока в проводнике будет равна

$$I = \frac{enSdx}{dt} = env_0S,$$

где $v_0 = \frac{dx}{dt}$ – дрейфовая скорость электронов. Разделив обе части этого равенства на S , получим выражение для плотности тока

$$j = env_0. \quad (2.1)$$

С другой стороны, согласно закону Ома в дифференциальной форме

$$j = \sigma E, \quad (2.2)$$

где E – напряженность электрического поля, σ – удельная проводимость проводника.

Сравнивая выражения (2.1) и (2.2), можно получить формулу, связывающую дрейфовую скорость электронов и напряженность приложенного электрического поля:

$$v_0 = \frac{\sigma}{en} E. \quad (2.3)$$

Оценим величину дрейфовой скорости, например, в медном проводнике, концентрация свободных электронов в котором $n = 3 \cdot 10^{22} \text{ см}^{-3}$, $\sigma = 0,6 \cdot 10^6 \text{ См/см}$ при $T = 0^\circ\text{C}$. Если приложить к проводнику длиной $L = 1 \text{ м}$ напряжение $U = 100 \text{ В}$, то $v_0 = \frac{\sigma U}{enL} \approx 1,2 \frac{\text{м}}{\text{с}}$, что намного меньше тепловой скорости.

Коэффициент пропорциональности в выражении (2.3) $b = \frac{v_0}{E}$ называется **подвижностью носителей заряда** (в нашем случае электронов) и имеет смысл скорости, которую приобретают носители заряда под действием электрического поля с единичной напряженностью. Подвижность измеряется в $\frac{\text{м}^2}{\text{В} \cdot \text{с}}$. Данная физическая величина является характеристикой условий ускорения электронов в металле. Зная подвижность и концентрацию электронов, можно вычислить удельную проводимость металлов:

$$\sigma = enb. \quad (2.4)$$

Как следует из соотношения (2.4), изменение удельной проводимости σ с увеличением или уменьшением температуры T будет определяться зависимостями $n(T)$ и $b(T)$.

Как уже указывалось, в металлах ускоряться электрическим полем могут лишь те электроны, которые имеют энергию вблизи энергии Ферми, и поскольку температура вырождения

электронного газа в металле намного выше комнатной температуры, концентрация n_f этих электронов практически не зависит от температуры:

$$n = n_f = \text{const} . \quad (2.5)$$

Поэтому далее рассмотрим температурную зависимость подвижности $b(T)$. Согласно полуклассической теории электропроводности на заряд в электрическом поле действует сила $F = eE$, которая сообщает ему ускорение $a = \frac{F}{m^*} = \frac{eE}{m^*}$, где m^* - эффективная масса электрона. Под действием этой силы электрон начинает двигаться равноускоренно, однако, возрастание скорости не может продолжаться до бесконечности.

Решетка реальных кристаллов никогда не бывает идеально правильной. Нарушения строгой периодичности решетки бывают обусловлены наличием примесей или вакансий (т. е. отсутствием атомов в узле), а также тепловыми колебаниями решетки. Столкновение электрона с атомами примеси и с атомами решетки, совершающими колебания, приводит к тому, что приобретенная скорость в направлении электрического поля падает практически до нуля. Среднее расстояние l , пролетаемое электроном от столкновения до столкновения, называется **длиной свободного пробега** электрона. Ему соответствует время τ . Таким образом, скорость, сообщаемая электрону электрическим полем, не превышает значения $v_{\max} = a\tau = \frac{eE\tau}{m^*}$. Среднее значение дрейфовой скорости, которую приобретают электроны под действием электрического поля, равно половине максимальной скорости. Более строгий вывод, учитывающий распределение свободных электронов в кристалле по энергиям приводит к аналогичному выражению для дрейфовой скорости:

$$v_0 = \frac{eE\tau}{m^*} . \quad (2.6)$$

Если для потери электроном приобретенной скорости достаточно одного столкновения, то

$$\tau = \frac{l}{v} , \quad (2.7)$$

где $v = v_m + v_0$ - полная скорость электрона, складываемая из тепловой и дрейфовой.

Таким образом, подвижность равна

$$b = \frac{e}{m^*} \frac{l}{v} . \quad (2.8)$$

Рассмотрим, как изменение температуры сказывается на процессах рассеяния электронов при их движении в кристалле. Это рассеяние может происходить вследствие взаимодействия электронов с различными дефектами решетки, однако, в наибольшей мере, на рассеяние электронов оказывают влияние колебания атомов кристаллической решетки; и электрические поля заряженных ионов примеси, всегда присутствующих в реальном кристалле.

Рассеяние электронов на колебаниях кристаллической решетки удобно описывать, рассматривая этот процесс как столкновение электронов с квантами упругих волн в кристалле – *фононами*.

Если кристаллическое тело рассматривать как систему связанных частиц, то смещение одного из атомов из положения равновесия влечет за собой смещение других соседних с ним атомов. В результате этого в кристалле возникает упругая волна. Дойдя до границы кристалла, волна отражается. При наложении прямой и отраженной волн образуется стоячая волна.

В соответствии с гипотезой Де Бройля с каждой бегущей монохроматической волной связаны энергия и импульс, определяемые соотношениями

$$\varepsilon = \hbar\omega, \quad \vec{p} = \hbar \vec{k}, \quad (2.9)$$

введенными по аналогии с теорией фотонов. Волна, несущая энергию и импульс, определяемые формулами (2.9), в каком-то отношении ведет себя как частица. Частица, уподобляемая звуковой волне в вышеуказанном смысле, называется *фононом*.

Фонон во многих отношениях ведет себя так, как если бы он был частицей с энергией и импульсом (2.9). Однако в отличие от обычных частиц (электронов, протонов, фотонов) фонон не может возникнуть в вакууме – для своего возникновения и существования фонон нуждается в некоторой среде. Подобного рода частицы называются *квазичастицами*.

Фононы хорошо приспособлены для описания слабых коллективных возбуждений атомов в кристалле. Между последовательными столкновениями фонон движется свободно, и если «длина свободного пробега» его достаточно велика по сравнению с постоянной кристаллической решетки, то возбужденное состояние кристалла можно в известном отношении рассматривать как *фононный газ*, подобно тому, как электромагнитное излучение можно представить как фотонный газ, заполняющий полость. Формально оба представления весьма схожи – и фотоны, и фононы подчиняются одной и той же статистике Бозе-Эйнштейна. Однако между фотонами и фононами имеется существенное различие: в то время как фотоны являются истинными частицами, фононы являются квазичастицами.

Так как волны, длина которых меньше удвоенного межатомного расстояния, не имеют физического смысла, то существует минимальная длина волны упругих волн в кристалле:

$$\lambda_{\min} \approx 2d,$$

где d – расстояние между соседними атомами в кристаллической решетке. Минимальной длине волны соответствует максимальная частота колебаний:

$$\omega_{\max} = \frac{2\pi\nu}{\lambda_{\min}},$$

где ν – скорость упругих волн в кристалле. Поэтому энергия фононов в кристалле ограничена сверху максимальной энергией

$$\varepsilon_{\max} = \hbar\omega_{\max}. \quad (2.10)$$

В случае фононного газа максимальной энергии фононов соответствует некоторая температура θ , определяемая условием $\varepsilon_{\max} = k\theta$, которая называется *характеристической температурой Дебая*. По определению

$$\theta = \frac{\varepsilon_{\max}}{k}. \quad (2.11)$$

Температура Дебая указывает для каждого вещества ту область, где становится существенным квантование энергии колебаний атомов кристаллической решетки.

При столкновении электронов с фононами выполняются законы сохранения энергии и импульса. Возрастание температуры означает увеличение концентрации фононов и, соответственно, повышение вероятности того, что электрон на своем пути столкнется с фононом. Длина свободного пробега электронов при этом падает ($l \sim \frac{1}{n_{\text{фон}}}$), т. е. в области высоких температур

$$l \sim \frac{1}{T}. \quad (2.12)$$

Для металлов, где электрическим полем ускоряются лишь электроны с энергией, близкой к энергии Ферми, можно записать, что $v = v_f = \text{const}$. Тогда, согласно (2.8), подвижность электронов в металле будет обратно пропорциональна температуре:

$$b \sim \frac{1}{T}. \quad (2.13)$$

При понижении температуры концентрация фононов падает, и существенным становится рассеяние на ионах примеси. При этом длина свободного пробега перестает зависеть от температуры при заданной концентрации ионов примеси.

Таким образом, удельное сопротивление $\rho = \frac{1}{\sigma}$ металлов можно представить в виде:

$$\rho = \rho_{\text{колеб.}} + \rho_{\text{прим.}}$$

где $\rho_{\text{колеб.}}$ – удельное сопротивление, обусловленное тепловыми колебаниями атомов решетки, $\rho_{\text{прим.}}$ – удельное сопротивление, обусловленное рассеянием электронов на примесных атомах. Слагаемое $\rho_{\text{колеб.}}$ уменьшается с понижением температуры и обращается нуль при $T = 0$ К. Слагаемое $\rho_{\text{прим.}}$ при небольшой концентрации примесей не зависит от температуры и образует так называемое остаточное сопротивление металла, т. е. сопротивление, которым металл обладает при $T = 0$ К. При этом речь идет о металлах, не переходящих в сверхпроводящее состояние (см. далее).

Зависимость удельного сопротивления ρ «чистого» - практически бездефектного и беспримесного – металла представлена на рис. 2.2 (кривая 1). Рассеяние на заряженных точечных дефектах в таком образце отсутствует, и подвижность электронов определяется их взаимодействием с тепловыми колебаниями решетки. В области температур T больших температуры Дебая θ для вырожденного электронного газа подвижность обратно пропорциональна температуре, а это означает, что

$$\rho = \frac{1}{enb} \sim T. \quad (2.14)$$

В области низких температур ($T < \theta$) удельное сопротивление $\rho \sim T^5$ (рис. 2.2). На этом же рисунке также показана температурная зависимость удельного сопротивления ρ металла, переходящего в сверхпроводящее состояние при $T = T_k$ (рис. 2.2, кривая 2).

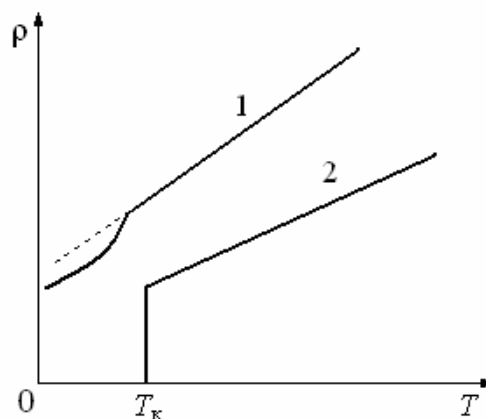


Рис. 2.2. Температурные зависимости удельного сопротивления ρ металла, не переходящего в сверхпроводящее состояние (1) и превращающегося в сверхпроводник при $T = T_k$ (2).

Сверхпроводимость

Уменьшение сопротивления металлов, не содержащих дефектов и примесей, при понижении температуры, на первый взгляд позволяет предположить, что при абсолютном нуле их сопротивление стремится к нулю. Но чистых бездефектных металлов в природе не существует, в решетке каждого из них есть (пусть даже в небольших количествах) примеси и собственные дефекты. Это означает, что даже при самых низких температурах металлы должны обладать конечным сопротивлением.

Тем удивительнее эффект, который был открыт в 1911 г. голландским физиком Х. Камерлинг-Оннесом: при понижении температуры до $T_k = 4,2$ К сопротивление очищенной ртути скачком падает до нуля. Удельная проводимость устремляется к бесконечности (рис. 5.4, кривая 2), поэтому этот эффект получил название сверхпроводимости. Если изготовить из сверхпроводящего материала кольцо, поместить его в постоянное магнитное поле, охладить до температуры T_k , а затем выключить магнитное поле, то возникающий вследствие электромагнитной индукции ток будет циркулировать по кольцу, не затухая, бесконечно долгое время. Температура T_k перехода в сверхпроводящее состояние называется **критической температурой**.

В настоящее время известно несколько сотен веществ, переходящих в сверхпроводящее состояние. Все эти вещества условно можно разделить на две большие группы: низкотемпературные и высокотемпературные сверхпроводники.

- 1) К низкотемпературным ($T_k < 25$ К) сверхпроводникам относятся некоторые металлы и сплавы, ряд полупроводников и интерметаллических соединений.
- 2) В 1986 г. были открыты высокотемпературные сверхпроводники, у которых T_k выше температуры жидкого азота, равной 77 К; к ним относятся сложные соединения – керамика на основе окиси меди. Теория явления в высокотемпературных сверхпроводниках пока еще не создана, поэтому далее ограничимся рассмотрением процессов, происходящих в низкотемпературных сверхпроводниках.

Следующее важное свойство сверхпроводника заключается в том, что внутри вещества в сверхпроводящем состоянии магнитная индукция магнитного поля всегда равна нулю, т. е. переход вещества в сверхпроводящее состояние сопровождается превращением его в идеальный диамагнетик (эффект Мейснера, открытый в 1933 г.)

Третья важная особенность сверхпроводимости заключается в том, что температура перехода в сверхпроводящее состояние зависит от индукции B внешнего магнитного поля: чем больше магнитная индукция, тем температура перехода меньше; T_k – это максимальная температура перехода соответствующая $B = 0$. Переходу при некоторой температуре $T < T_k$ соответствует критическая индукция $B_k(T)$, если при данной $T < T_k$ окажется, что $B > B_k$, то образец перейдет из сверхпроводящего состояния в нормальное. Таким образом, сверхпроводящее состояние может быть разрушено магнитным полем.

В основе теории сверхпроводимости, созданной в 1957 г. Дж. Бардиным, Л. Купером, Дж. Шриффером (БКШ-теории), лежит представление о том, что в металле между свободными

электронами действуют не только кулоновские силы отталкивания, но и силы притяжения, которые в сверхпроводящем состоянии преобладают над силами отталкивания.

Электрон, движущийся в металле, деформирует (поляризует) состоящую из положительных ионов кристаллическую решетку. В результате этой деформации электрон оказывается окруженным «облаком» положительного заряда, перемещающимся по решетке вместе с электроном. Электрон и окружающее его облако представляют собой положительно заряженную систему, к которой будет притягиваться другой электрон. В результате электроны проводимости объединяются в так называемые *куперовские пары* (по имени американского физика Л. Купера, впервые показавшего в 1956 г., что образование таких пар энергетически выгодно). Электроны, входящие в такую пару, имеют противоположно направленные спины. Поэтому спин пары равен нулю, и она представляет собой бозон. Бозоны склонны накапливаться в основном энергетическом состоянии, из которого их сравнительно трудно перевести в возбужденное состояние. Следовательно, куперовские пары, придя в согласованное движение, остаются в этом состоянии неограниченно долго. Такое согласованное движение пар и есть ток сверхпроводимости.

Подобная ситуация возникает при низких температурах у атомов гелия ${}^4_2\text{He}$ – они перестают взаимодействовать со стенками сосуда, внутреннее трение исчезает, наблюдается сверхтекучесть гелия. В этом смысле сверхпроводимость кристалла можно интерпретировать как сверхтекучесть бозонной жидкости – куперовских пар.

Не следует представлять себе куперовскую пару, как состоящую из одних и тех же электронов. Пары постоянно обмениваются партнерами, исчезают и появляются вновь. Кроме того, эти пары вместе с наведенным ими положительным зарядом постоянно перемещаются по кристаллу. В результате сферы действия различных пар накладываются друг на друга, а их волновые функции в большей мере перекрываются, что приводит к усилению взаимодействия между всеми электронами и решеткой. Теперь, если на пути одного из электронов, участвующего в коллективном движении пар, и положительных поляризационных зарядов возникает препятствие, то воздействие остальных членов этого коллектива частиц позволяет ему проходить это препятствие без рассеяния. Дрейфовая скорость носителя заряда, таким образом, не меняется, что и означает полное отсутствие электрического сопротивления, т. е. сверхпроводимость образца.

В куперовские пары объединяются не все электроны проводимости. При температуре T , отличной от абсолютного нуля, имеется некоторая вероятность того, что пара будет разрушена. Поэтому всегда наряду с парами имеются «нормальные» электроны, движущиеся по кристаллу обычным образом. Чем ближе T к T_k , тем доля нормальных электронов становится больше, обращаясь в 1 при $T = T_k$.

Образование куперовских пар приводит к перестройке энергетического спектра металла. Для возбуждения электронной системы, находящейся в сверхпроводящем состоянии, надо разрушить хотя бы одну пару, на что требуется энергия, равная энергии связи $E_{св}$ электрона в паре. Эта энергия представляет собой минимальное количество энергии, которое может воспринять система электронов сверхпроводника. Следовательно, в энергетическом спектре электронов, находящихся в сверхпроводящем состоянии, имеется щель ширины $E_{св}$, расположенная в области уровня Ферми (рис. 2.3). Значения энергии, принадлежащие этой щели, запрещены. Существование щели было доказано экспериментально.

Итак, возбужденное состояние электронной системы, находящейся в сверхпроводящем состоянии, отделено от основного состояния энергетической щелью ширины $E_{св}$. Поэтому квантовые переходы этой системы не всегда будут возможными. При малых скоростях своего движения (отвечающих силе тока, меньшей I_k) электронная система не будет возбуждаться, а это и означает движение без трения, т. е. без электрического сопротивления.

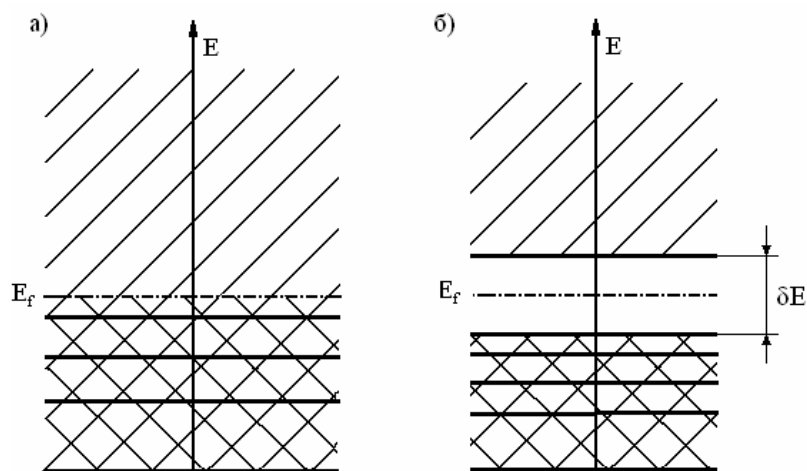


Рис. 2.3. Зонные схемы металла в обычном (а) и сверхпроводящем (б) состояниях

Ширина энергетической щели $E_{св}$ с ростом температуры уменьшается и обращается в нуль при критической температуре T_k . Соответственно, все куперовские пары разрушаются, и вещество переходит в нормальное (несверхпроводящее) состояние.

Применение сверхпроводников в мощных электрических устройствах в настоящее время наталкивается на серьезные трудности, связанные с необходимостью создания громоздких систем охлаждения до температур, которые может обеспечить лишь жидкий гелий. Поэтому использование сверхпроводников в микроэлектронике оказывается более перспективным. Типичные примеры устройств этого типа – приемники СВЧ-излучения, магнетроны и переключатели.

3. Электропроводность полупроводников

Собственная проводимость полупроводников

Собственными полупроводниками являются химически чистые полупроводники (например, Ge, Si, Se), а их проводимость называется *собственной проводимостью*. При $T=0$ К полупроводники имеют полностью заполненную электронами валентную зону и свободную зону (зону проводимости), которые разделены запрещенной зоной ΔE , ширина которой может быть от сотых долей до 3 эВ. Поэтому при $T=0$ К и в отсутствие других внешних воздействий (освещение,

облучение и т. д.) полупроводники не проводят электрического тока. Собственная проводимость возникает при повышении температуры в результате перехода электронов с верхних уровней валентной зоны в зону проводимости. При этом в зоне проводимости появляется некоторое число носителей тока – электронов, занимающих уровни вблизи дна зоны; одновременно в валентной зоне освобождается такое же число мест на верхних уровнях, в результате чего появляются дырки.

На переход электрона в зону проводимости затрачивается *энергия активации* – энергия, необходимая для создания пары электрон-дырка, равная ширине запрещенной зоны. При появлении электрона в зоне проводимости в валентной зоне возникает дырка. Следовательно, энергия, затраченная на образование пары носителей тока, должна делиться на две равные части. Начало отсчета для процессов образования электрона и дырки должно находиться в середине запрещенной зоны. Энергия Ферми в собственном полупроводнике и представляет собой энергию, от которой происходит возбуждение электронов и дырок. Поэтому уровень Ферми в собственном полупроводнике находится в середине запрещенной зоны (рис. 3.1).

Распределение электронов по уровням валентной зоны и зоны проводимости описывается функцией распределения Ферми-Дирака:

$$f(E) = \frac{1}{e^{\frac{(E-E_f)}{kT}} + 1} \quad (3.1)$$

Это распределение можно сделать очень наглядным, изобразив, как это сделано на рис. 3.1, график функции распределения совместно со схемой энергетических зон. Как уже говорилось, у собственных полупроводников отсчитанное от потолка валентной зоны значение уровня Ферми равно

$$E_F = \frac{\Delta E}{2},$$

где ΔE – ширина запрещенной зоны. Следовательно, для электронов, перешедших в зону проводимости, величина $E - E_F$ мало отличается от половины ширины запрещенной зоны. Уровни зоны проводимости лежат на хвосте кривой распределения. Поэтому в выражении (3.1) единицей в знаменателе можно пренебречь, и вероятность заполнения энергетических уровней электронами можно находить по формуле

$$f(E) = e^{-\frac{(E-E_f)}{kT}}.$$

Положив в этой формуле $E - E_F \approx \Delta E / 2$, получим, что

$$f(E) \approx e^{-\frac{\Delta E}{2kT}} \quad (3.2)$$

Количество электронов, перешедших в зону проводимости, а следовательно и количество образовавшихся дырок, будет пропорционально вероятности (3.2). Эти электроны и дырки являются носителями тока. Поскольку проводимость пропорциональна числу носителей тока, она также будет пропорциональна выражению (3.2). Следовательно, *удельная проводимость* собственных полупроводников быстро растет с повышением температуры, изменяясь по закону

$$\sigma = \sigma_0 e^{-\frac{\Delta E}{2kT}}, \quad (3.3)$$

где ΔE – ширина запрещенной зоны, σ_0 – константа, характерная для данного полупроводника.

Если на графике откладывать зависимость $\ln \sigma$ от $1/T$, то для собственных полупроводников получается прямая линия. По наклону этой прямой можно определить ширину запрещенной зоны ΔE .

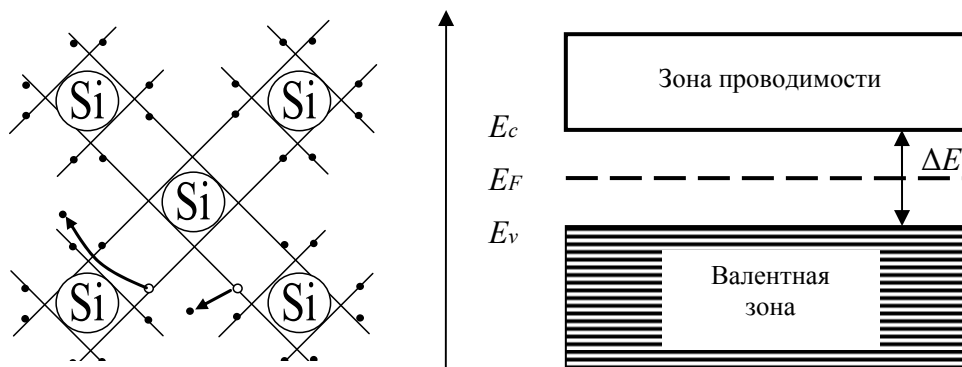


Рис. 3.1. Парно-электронные связи в кристалле кремния и энергетическая диаграмма собственного полупроводника.

Электроны проводимости и дырки свободно движутся в кристаллической решетке полупроводника. При встрече свободного электрона с дыркой они рекомбинируют (электрон занимает свободное место). Это означает, что электрон теряет свободу передвижения до тех пор, пока снова не получит от кристаллической решетки энергию, достаточную для своего высвобождения. Рекомбинация приводит к одновременному исчезновению свободного электрона и дырки. На схеме энергетических уровней (рис. 3.1) процессу рекомбинации соответствует переход электрона из зоны проводимости на один из свободных уровней валентной зоны.

Таким образом, в собственном полупроводнике идут одновременно два процесса: рождение попарно свободных электронов и дырок и рекомбинация, приводящая к попарному исчезновению электронов и дырок. Вероятность первого процесса быстро растет с повышением температуры. Вероятность рекомбинации пропорциональна как числу свободных электронов, так и числу дырок. Следовательно, каждой температуре соответствует определенная равновесная концентрация электронов и дырок, которая изменяется с температурой пропорционально выражению (3.2).

Когда внешнее электрическое поле отсутствует, электроны проводимости и дырки движутся хаотически. При включении поля на хаотическое движение накладывается упорядоченное движение: электронов против поля и дырок – в направлении поля. Оба движения – и дырок, и электронов - приводят к переносу заряда вдоль кристалла. Следовательно, собственная электропроводность обуславливается как бы носителями заряда двух знаков – отрицательными электронами и положительными дырками.

Закон Ома в дифференциальной форме для полупроводников имеет вид:

$$j = (en_n b_n + ep_p b_p)E, \quad (3.4)$$

где n_n, n_p – концентрации электронов и дырок; b_n, b_p – подвижности электронов и дырок, соответственно. В случае собственного полупроводника концентрация электронов равна концентрации дырок: $n_n = n_p = n$, и закон Ома в дифференциальной форме для собственного полупроводника примет вид:

$$j = en(b_n + b_p)E, \quad (3.5)$$

Примесная проводимость полупроводников

Примесная проводимость возникает, если некоторые атомы в узлах кристаллической решетки полупроводника заменить атомами, валентность которых отличается на единицу от валентности основных атомов.

Например, при внесении в четырехвалентный кремний пятивалентных атомов фосфора для образования ковалентных связей с соседями атому фосфора достаточно четырех электронов (см. рис. 3.2). Следовательно, пятый валентный электрон оказывается как бы лишним и легко отщепляется от атома за счет энергии теплового движения, образуя свободный странствующий электрон. При этом образование свободного электрона не сопровождается нарушением ковалентных связей, т. е. образованием дырки. Таким образом, в полупроводнике с примесью, валентность которой на единицу больше валентности основных атомов, имеются избыточные носители тока – электроны. Соответственно говорят, что такой полупроводник обладает электронной проводимостью или является полупроводником *n-типа*. Атомы примеси, поставляющие электроны проводимости, называются *донорами*.

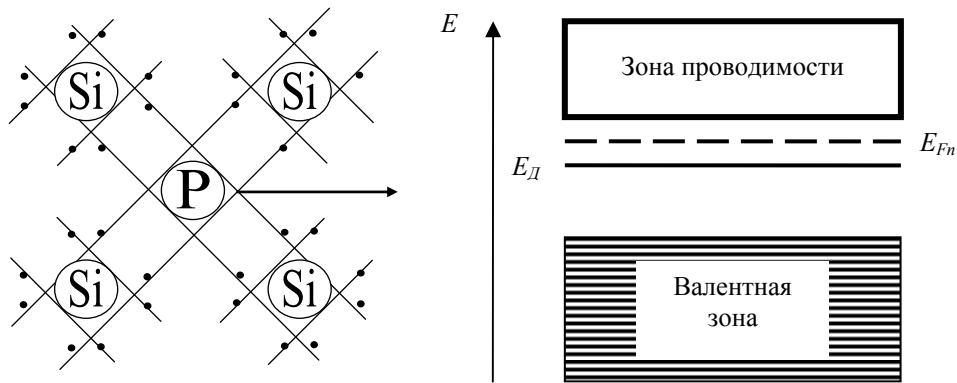


Рис.3.2. Атомы фосфора в решетке кремния и энергетическая диаграмма полупроводника n -типа. E_D – донорный уровень; E_{Fn} – уровень Ферми в полупроводнике n -типа.

Теперь рассмотрим случай, когда валентность примеси на единицу меньше валентности основных атомов, например, кристаллический кремний с примесью трехвалентных атомов бора (см. рис. 3.3). Трех валентных электронов атома бора недостаточно для образования связей со всеми четырьмя соседями. Поэтому одна из связей окажется не укомплектованной и будет представлять собой место, способное захватить электрон. При переходе на это место электрона из одной из соседних пар валентных электронов возникнет дырка, которая будет кочевать по кристаллу. Таким образом, в полупроводнике с примесью, валентность которой на единицу меньше валентности основных атомов, возникают избыточные носители тока – дырки. Проводимость в этом случае называют дырочной, а о полупроводнике говорят, что он принадлежит к p -типу. Примеси, вызывающие возникновение дырок, называются **акцепторными**.

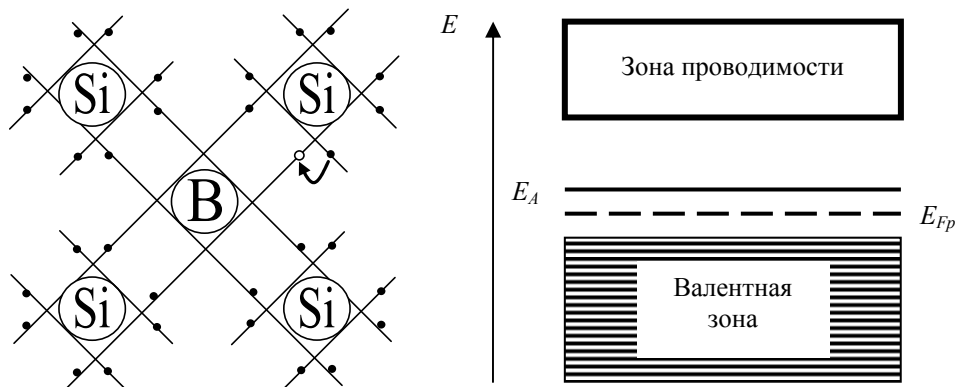


Рис. 3.3. Атомы бора в решетке кремния и энергетическая диаграмма полупроводника p -типа. E_A – донорный уровень; E_{Fp} – уровень Ферми в полупроводнике p -типа.

Примеси искажают поле кристаллической решетки, что приводит к возникновению на энергетической схеме примесных уровней, расположенных в запрещенной зоне кристалла. В случае

полупроводников *n*-типа примесные уровни называются **донорными** (рис. 3.2), в случае полупроводников *p*-типа **акцепторными** (рис. 3.3).

При $T=0$ К уровень Ферми в полупроводниках *n*-типа располагается в верхней половине запрещенной зоны, а в полупроводниках *p*-типа – в нижней половине запрещенной зоны. При повышении температуры уровень Ферми в полупроводниках обоих типов смещается к середине запрещенной зоны. Донорные уровни, расположенные недалеко от дна зоны проводимости могут существенно повлиять на электрические свойства кристалла. В этом случае энергия теплового движения даже при обычных температурах оказывается достаточной для того, чтобы перевести электрон с донорного уровня в зону проводимости (см. рис. 3.2). Этому процессу соответствует отщепление пятого валентного электрона от атома примеси.

Акцепторные уровни, расположенные недалеко от потолка валентной зоны (см. рис. 3.3), оказывают также существенное влияние на электрические свойства кристалла. В этом случае образованию дырки отвечает переход электрона из валентной зоны на акцепторный уровень.

При повышении температуры концентрация примесных носителей тока быстро достигает насыщения. Это означает, что практически освобождаются все донорные или заполняются электронами все акцепторные уровни. Вместе с тем по мере роста температуры все в большей степени начинает сказываться собственная проводимость полупроводника, обусловленная переходом электрона непосредственно из валентной зоны в зону проводимости. Таким образом, при высоких температурах проводимость полупроводника будет складываться из примесной и собственной проводимостей. При низких температурах преобладает примесная проводимость, а при высоких – собственная проводимость.

Электронно-дырочный переход

Электронно-дырочный или *p-n-переход* – это контакт двух полупроводников с различными типами проводимости (*n*- и *p*-типов). Для получения контакта с хорошо контролируемыми и постоянными свойствами необходимо создавать его в виде внутренней границы раздела, на которой полупроводник одного типа переходил бы непрерывно в полупроводник другого типа. Это достигается путем легирования (введения) примеси в соответствующие места во время выращивания кристалла либо путем диффузии или имплантации примеси уже в готовый кристалл. *P-n*-переход может быть образован в одном полупроводниковом материале, например, германии (Ge), кремнии (Si), арсениде галлия (GaAs), а также быть создан между полупроводниками с различной шириной запрещенной зоны: Ge-Si, Ge-GaAs и др.

Рассмотрим физические процессы, происходящие в *p-n*-переходе. В полупроводнике *n*-типа основными носителями свободного заряда являются электроны, их концентрация значительно превышает концентрацию дырок – неосновных носителей заряда. В полупроводнике *p*-типа основными носителями являются дырки, а неосновными – электроны. Уровни Ферми в полупроводниках при комнатной температуре до контакта расположены на разной высоте: в полупроводнике *n*-типа – вблизи дна зоны проводимости, а в полупроводнике *p*-типа – у потолка ва-

лентной зоны. Оба полупроводника электрически нейтральны в любом участке своего объема, т. к. заряды ионизированных атомов полностью нейтрализованы зарядами противоположного знака свободных носителей заряда.

Пусть полупроводник p -типа приводится в контакт с полупроводником n -типа (рис. 3.4). Из-за градиента концентрации носителей заряда возникает их диффузия в области с противоположным типом проводимости через границу раздела: из n -области в p -область диффундируют электроны, а навстречу им – дырки. Такой установившийся направленный перенос основных носителей заряда приводит к возникновению тока, который называется *диффузионным* (I_d). Диффузионный ток, создаваемый переходом электронов, складывается с диффузионным током, создаваемым встречным потоком дырок, т. е.

$$I_d = I_n + I_p.$$

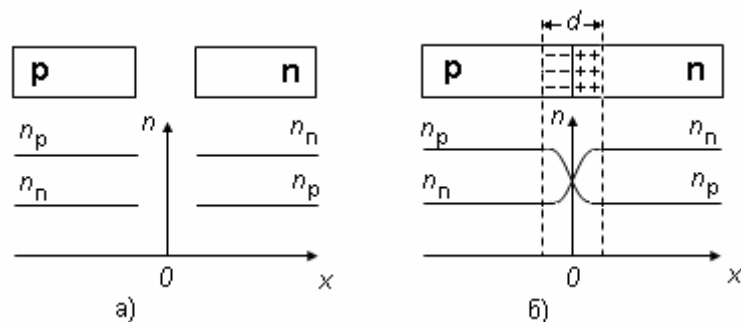


Рис. 3.4. а) Концентрации электронов и дырок в полупроводниках p -типа и n -типа до создания p - n -перехода; б) образование контактного слоя при создании p - n -перехода.

Электроны, перешедшие из n -области в p -область, рекомбинируют здесь с дырками. Рекомбинация происходит в тонком слое этой области, прилегающем к границе раздела. В результате рекомбинации отрицательный заряд ионов акцепторной примеси в этом слое не компенсируется теперь положительным зарядом дырок и весь слой заряжается отрицательно. В n -области в результате рекомбинации электронов с перешедшими сюда из p -области дырками пограничный слой заряжается положительно благодаря не скомпенсированному положительному заряду ионов донорной примеси. Таким образом, в области контакта образуется двойной электрический слой объемных зарядов (d), а, следовательно, *контактное электрическое поле* (E_k), напряженность которого направлена от n -области к p -области.

Приконтактная область (d) обладает большим электрическим сопротивлением, т. к. концентрация свободных носителей заряда в ней очень мала. Эта область обычно достигает толщины порядка десятков и сотен межатомных расстояний. Контактное электрическое поле препятствует дальнейшей диффузии основных носителей заряда навстречу друг другу.

На рис. 3.5 показана зонная схема p - n -перехода, находящегося в термодинамическом равновесии. Диффузионное перетекание основных носителей, приводящее к появлению не скомпенсированных зарядов в приконтактных областях, вызывает смещение энергетических уровней. В n -области, зарядившейся положительно, уровни понижаются, а в p -области – смещаются вверх.

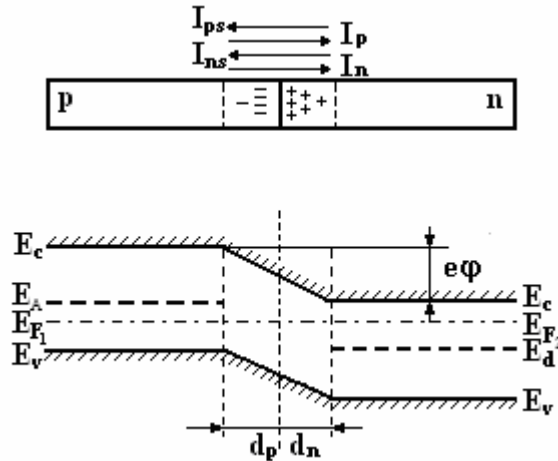


Рис. 3.5. Энергетические зоны при создании p - n -перехода в случае равновесного состояния при отсутствии внешнего электрического поля.

Перемещение уровней происходит до тех пор, пока уровни Ферми (E_{Fn} и E_{Fp}) не расположатся на одной высоте, что соответствует установлению состояния равновесия. Сдвиг энергетических зон соответствует образованию в области контакта *потенциального барьера*, высота которого равна $e\phi$, где ϕ - *контактная разность потенциалов*. Как видно из рисунка, высота потенциального барьера

$$e\phi = E_{F_p} - E_{F_n}.$$

Положение уровня Ферми зависит от степени легирования полупроводника: чем сильнее легирован полупроводник, тем ближе к соответствующей разрешенной зоне располагается уровень Ферми. Таким образом, максимальное значение потенциального барьера на границе двух полупроводников разных типов равно ширине запрещенной зоны ΔE . В реальных случаях высота потенциального барьера обычно меньше. Так, для германиевых p - n -переходов $e\phi$ приблизительно равно 0,35 эВ, для кремниевых - 0,6 эВ, что составляет примерно половину ширины запрещенной зоны ($\Delta E = 0,67$ эВ для германия и $\Delta E = 1,1$ эВ для кремния).

Контактное электрическое поле является регулятором перехода носителей через p - n -переход. Преодолеть потенциальный барьер могут лишь те основные носители, энергия которых больше потенциального барьера. Эти основные носители создают *диффузионный ток*, который имеет две составляющие: электронную и дырочную.

$$I = I_n + I_p.$$

Но высокоэнергетичных электронов и дырок мало, т. к. основная часть электронов находится вблизи дна зоны проводимости, а дырок – вблизи потолка валентной зоны. Поэтому потоки основных носителей через область перехода невелики, хотя общее число этих носителей огромно.

Вместе с тем, контактное электрическое поле способствует переходу через границу раздела неосновных носителей, создающих дрейфовый ток (I_s), который направлен навстречу диффузионному. Дрейфовый ток имеет также две составляющие: электронную и дырочную.

$$I_s = I_{ns} + I_{ps}.$$

Величина дрейфового тока невелика, т. к. определяется концентрацией неосновных носителей, и практически не зависит от значения контактной разности потенциалов φ .

Итак, в состоянии равновесия через p - n -переход проходят четыре тока: I_n – диффузионный ток основных носителей – электронов, имеющих достаточную энергию для совершения работы против сил контактного электрического поля и проникающих из n -области в p -область; I_p – диффузионный ток основных носителей – дырок, из p -области в n -область; I_{ns} – дрейфовый ток неосновных для p -области носителей – электронов, возвращаемых в n -область контактным полем; I_{ps} – дрейфовый ток неосновных для n -области носителей – дырок, возвращаемых в p -область. Сумма этих токов в условиях равновесия равна нулю.

P–n-переход при нарушении равновесия внешним электрическим полем

При обычных температурах контактный слой является *запирающим*. Сопротивление запирающего слоя можно изменить с помощью внешнего электрического поля. Если полупроводник с p - n -переходом подключен к источнику тока так, что положительный полюс источника соединен с n -областью, а отрицательный – с p -областью, то возникающее поле совпадает по направлению с контактным электрическим полем (рис. 3.6). Так как область перехода обладает большим сопротивлением по сравнению с остальной частью полупроводника, то приложенная внешняя разность потенциалов практически вся будет падать на запирающем слое и падением напряжения на остальных участках полупроводников можно пренебречь.

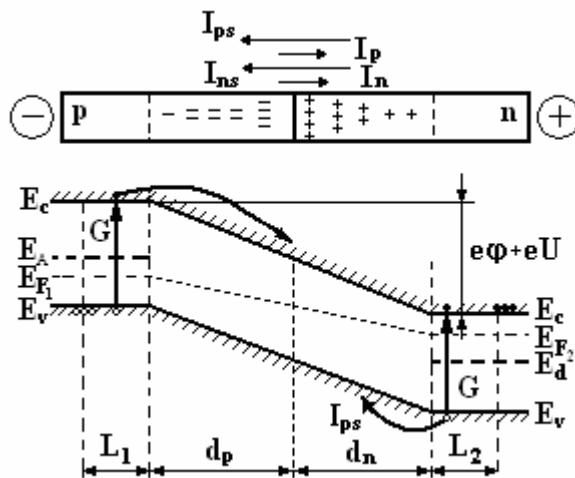


Рис. 3.6. Энергетические зоны при нарушении равновесия внешним электрическим полем (обратное направление).

Внешнее напряжение U при этом смещает уровни энергии в контактирующих областях на величину, равную eU . Сдвигаются и уровни Ферми, что указывает на нарушение состояния равновесия. Высота потенциального барьера возрастает и становится равной $e(\varphi + U)$. Приложенное напряжение препятствует диффузионному движению основных носителей. Основные носители тока как бы "отстают" от p - n -перехода, что приводит к увеличению толщины (d) и возрастанию сопротивления контактного слоя. Однако незначительному количеству неосновных носителей зарядов из n -области и p -области удастся пройти p - n -переход, и в цепи будет протекать весьма небольшой ток, называемый *обратным* током. Напряжение, поданное на p - n -переход в этом случае, называют *обратным*, и его принято считать отрицательным.

Если p - n -переход соединить с источником так, чтобы положительный полюс источника был соединен с p -областью, а отрицательный с n -областью, то высота потенциального барьера уменьшается на величину приложенного напряжения и становится равной $e(\varphi - U)$ (рис. 3.7).

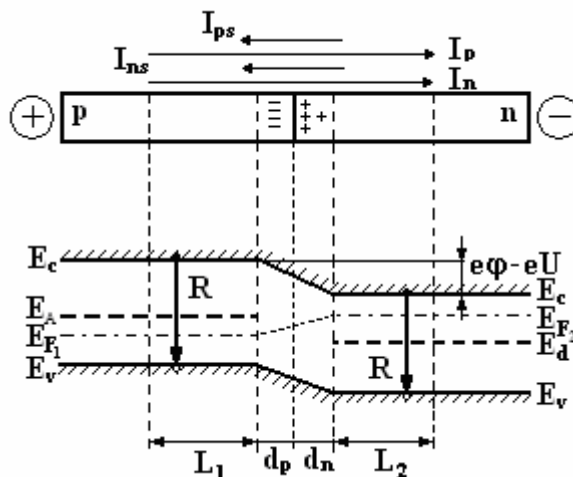


Рис. 3.7. Энергетические зоны при нарушении равновесия внешним электрическим полем (прямое направление).

Снижение потенциального барьера приводит к увеличению числа электронов n -области, обладающих энергией, необходимой для его преодоления, что обуславливает быстрый рост диффузионного тока. Кроме того, в этом случае внешнее и контактное поля в области перехода имеют противоположные направления. Поэтому результирующее поле оказывается ослабленным, из-за чего уменьшается толщина (d) и сопротивление контактного слоя, что также способствует росту диффузионного тока. Поскольку сила дрейфового тока мала, то можно сказать, что ток, текущий через p - n -переход, полностью обусловлен потоками основных носителей (этот ток называют *прямым* током). Сила тока в этом случае будет возрастать при увеличении напряжения источника. Напряжение, поданное на p - n -переход в этом случае, называют *прямым* и считают положительным. Таким образом, p - n -переход обладает *односторонней проводимостью*.

Вольтамперная характеристика (ВАХ) p - n -перехода

Вольтамперная характеристика (ВАХ) p - n -перехода представлена на рис. 3.8. Полный ток через переход при прямом напряжении, равный разности между диффузионным и дрейфовым токами, быстро растет по мере увеличения напряжения благодаря увеличению диффузионного тока при неизменном дрейфовом токе.

ВАХ описывается уравнением:

$$I = I_s (e^{\frac{eU}{kT}} - 1),$$

где U – внешнее напряжение, приложенное к p - n -переходу с учетом знака, I_s – значение, к которому стремится обратный ток при увеличении обратного напряжения (дрейфовый ток), k – постоянная Больцмана, T – температура. При комнатной температуре ($kT=0,025$ эВ), при обратном напряжении $U=-0,1B$, $e^{\frac{eU}{kT}} = e^{-4} < 0,02$ и поэтому $I = -I_s$. При прямом напряжении $U=0,1B$, $e^{\frac{eU}{kT}} = e^4 > 50$, единицей в скобках можно пренебречь, что приводит к экспоненциальному росту прямого тока.

Обратный ток через переход всегда намного меньше прямого, поэтому при построении обратной ветви ВАХ используются обычно другие масштабы. Обратный ток, равный разности дрейфового и диффузионного токов, на начальном участке довольно быстро растет за счет резкого уменьшения диффузионного тока с увеличением потенциального барьера. При дальнейшем увеличении обратного напряжения диффузионный ток падает практически до нуля, и обратный ток становится равным дрейфовому току, который не зависит от приложенного напряжения.

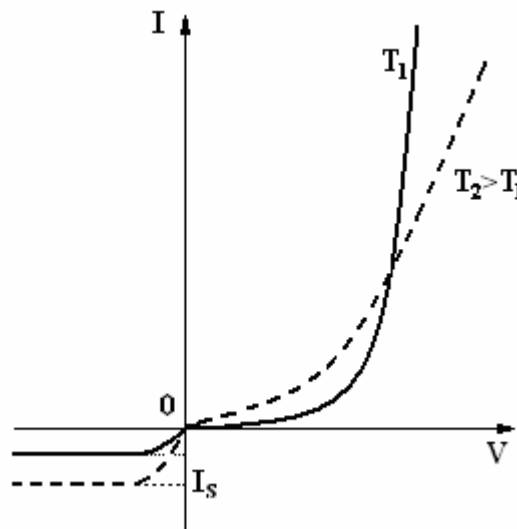


Рис. 3.8. ВАХ p - n -перехода.

Таким образом, ВАХ $p-n$ -перехода нелинейная. $P-n$ -переход обладает очень большим сопротивлением при обратном направлении внешнего электрического поля и малым сопротивлением при прямом направлении. Величину $R_d = \frac{dU}{dI}$ называют *дифференциальным сопротивлением*, характеризующим изменение напряжения при изменении тока через $p-n$ -переход. Эта величина содержит важную информацию о физических процессах в полупроводнике.

Если обратное напряжение непрерывно повышать, то при некотором значении напряжения, называемого *напряжением пробоя*, наблюдается резкое увеличение силы тока. Известны следующие виды пробоя полупроводниковых $p-n$ -переходов: *тепловой, туннельный и лавинный*.

Полупроводниковые приборы

В настоящее время полупроводниковые приборы используются практически во всех областях электроники и радиотехники. Несмотря на чрезвычайное разнообразие этих приборов, в основе их, как правило, лежит работа $p-n$ -перехода или системы из нескольких $p-n$ -переходов.

Полупроводниковый диод содержит один $p-n$ -переход, к каждой из областей которого подведены с помощью омических контактов металлические вводы. Диоды используются для выпрямления и преобразования переменных токов. ВАХ диода в значительной мере определяется свойствами его $p-n$ -перехода и зависит от характеристик материалов, на базе которых они выполнены. Так, с увеличением прямого напряжения прямой ток вначале экспоненциально возрастает, а затем, начиная с некоторого напряжения $U=\varphi$, растет почти линейно (рис. 3.8). Это объясняется тем, что с ростом прямого напряжения происходит уменьшение потенциального барьера, равного $e(\varphi-U)$ (рис. 3.7), что приводит к росту силы тока, текущего через переход. Когда внешнее напряжение превысит контактную разность потенциалов, потенциальный барьер совсем исчезнет. Вместе с ним исчезнет и обедненная носителями область, после чего напряжение окажется распределенным по всему образцу. Дальнейшее увеличение напряжения обуславливает рост тока в соответствии с законом Ома.

Зависимость обратного тока для диодов можно представить в виде:

$$I_s = C e^{-\frac{\Delta E}{kT}},$$

где C – множитель, слабо зависящий от температуры, ΔE – ширина запрещенной зоны. Экспоненциальный множитель определяет сильную зависимость обратного тока, как от температуры, так и от ширины запрещенной зоны. При увеличении ΔE , например, при замене германия кремнием, обратный ток уменьшается на несколько порядков, и, как следствие, изменяется ВАХ при прямом включении. Причем линейный рост прямого тока для германиевых диодов происходит

при меньших значениях напряжения, чем для кремниевых, что определяется различной контактной разностью потенциалов.

Стабилитроны. При электрическом пробое обратного смещенного $p-n$ -перехода сила тока возрастает очень круто. Этот участок ВАХ может быть использован для стабилизации напряжения. Диод, работающий при обратном напряжении и предназначенный для стабилизации напряжения, называется *стабилитроном*. Наилучшими для этой цели являются кремниевые диоды, у которых при достижении напряжения пробоя незначительные изменения напряжения вызывают изменения силы тока через диод в очень широких пределах (рис. 3.9). Физическим механизмом, обуславливающим наличие практически линейного участка на обратной ветви ВАХ, является лавинный или туннельный пробой $p-n$ -перехода.

Основными параметрами стабилитрона являются: напряжение стабилизации $U_{ст}$ и дифференциальное сопротивление R_d на рабочем (линейном) участке. Дифференциальное сопротивление характеризует изменение стабилизированного напряжения при изменении тока через диод. Чем меньше R_d на линейном участке, тем выше качество стабилизации.

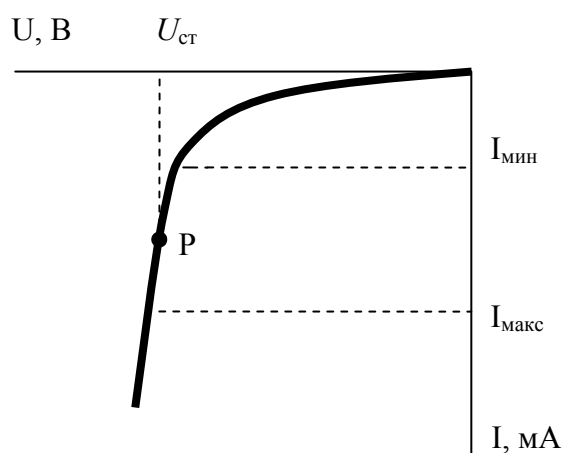


Рис. 3.9. ВАХ стабилитрона.

Транзисторы

Одними из наиболее известных представителей класса полупроводниковых приборов являются транзисторы, получившие свое название от английского словосочетания *transfer resistance* (управляемое сопротивление). Транзисторы разработаны в США Дж. Бардиным, У. Браттейном и У. Шокли в 1948-1949 г.г.

Транзисторы наиболее распространенной конструкции содержат два $p-n$ -перехода, включенных навстречу друг другу. Различают транзисторы $p-n-p$ -типа и $n-p-n$ -типа. Средняя часть транзистора называется его **базой**. Прилегающие к базе с обеих сторон области с иным, чем у нее, типом проводимости образуют **эмиттер** и **коллектор** (см. рис. 3.10).

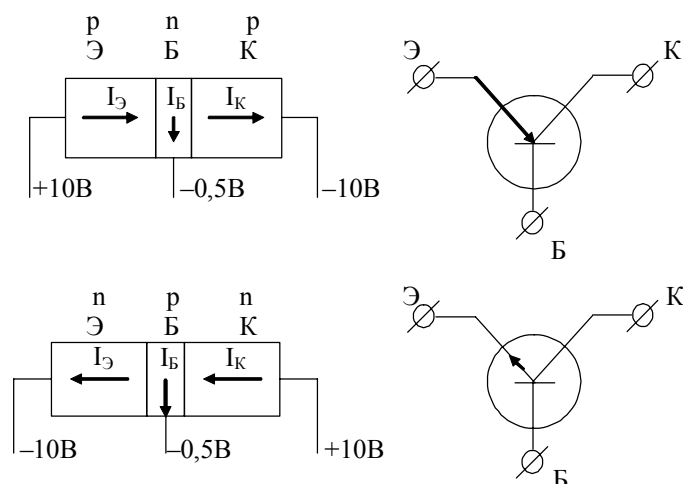


Рис. 3.10. Блок-схема транзисторов $p-n-p$ -типа и $n-p-n$ -типа и обозначение их на принципиальных электрических схемах.

В настоящее время в силу технологических особенностей изготовления одним из основных элементов интегральных микросхем являются транзисторы $n-p-n$ -типа, поэтому остановимся на рассмотрении принципа работы прибора именно этого типа. Процессы, происходящие в других типах транзисторов, во многом аналогичны.

Существует несколько способов включения транзистора в электрическую цепь: схемы с общей базой, общим эмиттером или с общим коллектором. Рассмотрим работу транзистора в схеме включения с общей базой (рис. 3.11), когда на переход “эмиттер – база” подается прямое смещение, а на переход “база – коллектор” – обратное смещение.

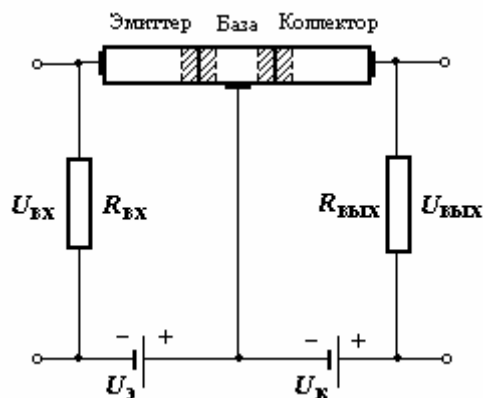


Рис. 3.11. Схема включения транзистора $n-p-n$ -типа с общей базой.

На рис. 3.12 приведена зонная схема транзистора $n-p-n$ -типа в случае, когда смещающие напряжения и входной сигнал отсутствуют.

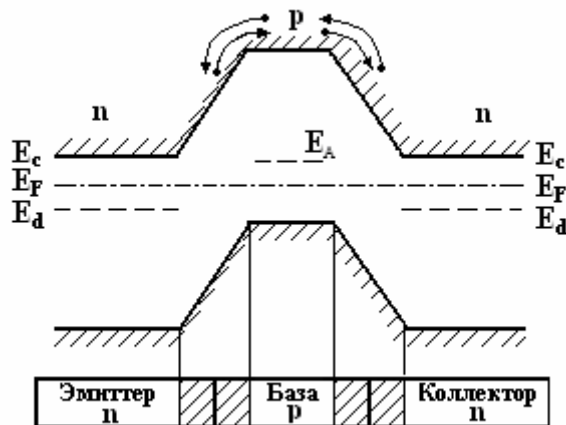


Рис. 3.12. Зонная схема транзистора $n-p-n$ -типа в отсутствие напряжения.

Подключение прямого напряжения $U_э$ понижает потенциальный барьер на первом переходе, а подключение обратного напряжения $U_к$ повышает потенциальный барьер на втором переходе (рис. 3.13). В результате концентрация электронов в p -области вблизи эмиттера возрастает, а вероятность попадания электронов из коллектора в базу становится малой. Возникает перепад концентраций неосновных носителей заряда от одного края базы к другому, и, следовательно, начнется диффузия электронов сквозь базу транзистора от эмиттера к коллектору.

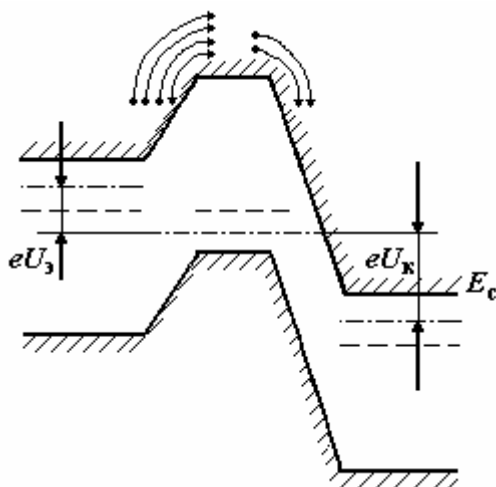


Рис. 3.13. Зонная схема транзистора $n-p-n$ -типа, работающего по схеме с общей базой.

Если база легирована равномерно, электрическое поле в ней практически отсутствует, так как оно сосредоточено лишь в барьерных областях транзистора. При достаточно тонкой базе электроны успевают дойти до коллекторного перехода, не рекомбинируя в ней с дырками. И хотя рекомбинация все же присутствует, в реальных транзисторах ею можно пренебречь, и тогда оказывается, что ток “эмиттер – база“ $I_э$ примерно равен току $I_к$.

Ток $I_к$, проходя через сопротивление нагрузки $R_н$, создает выходное напряжение $U_{вых} = I_к R_н$. В схеме включения с общей базой переход “эмиттер – база“ смещен в прямом направлении, и его сопротивлением можно пренебречь. Если же при этом сопротивление нагрузки $R_н$ много

больше сопротивления $R_{вх}$ на входе устройства, то и напряжение сигнала на выходе транзистора $U_{вых}$ окажется гораздо выше входного $U_{вх} = I_{э} R_{вх}$.

Таким образом, в схеме с общей базой даже небольшое изменение напряжения на входе транзистора приводит к сильному изменению напряжения на нагрузке. Одновременно возрастает и энергия, выделяемая в единицу времени на $R_{н}$. Дополнительная энергия черпается при этом из источника напряжения в цепи коллектора.

Усиления по напряжению и току можно добиться, используя схему включения с общим эмиттером (рис. 3.14). В такой схеме входным током прибора является уже не ток эмиттера $I_{э}$, а ток базы $I_{б}$.

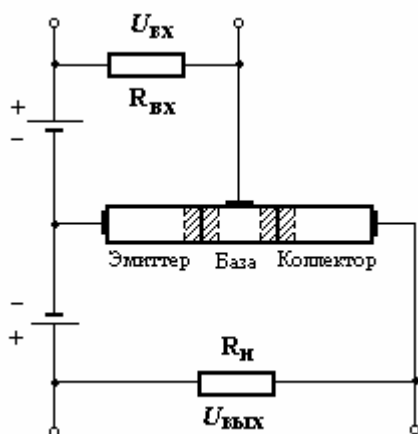


Рис. 3.14. Схема включения транзистора $n-p-n$ -типа с общим эмиттером.

Варьируя напряжение на входе, можно менять как ток базы, так и остальные токи; при этом, очевидно, должно выполняться условие:

$$\Delta I_{э} = \Delta I_{к} + \Delta I_{б}. \quad (3.6)$$

С учетом того, что часть электронов, диффундирующих сквозь базу транзистора от эмиттерного перехода к коллекторному, все же рекомбинирует с поступающими из цепи базы дырками, можно записать:

$$\Delta I_{к} = \alpha \Delta I_{э}. \quad (3.7)$$

При этом, поскольку $\Delta I_{э} > \Delta I_{к}$, $\alpha < 1$. Коэффициент α выражает отношение изменения тока коллектора к изменению тока эмиттера при постоянном напряжении $U_{к}$ на коллекторном переходе.

Из соотношений (3.6) и (3.7) следует:

$$\frac{\Delta I_{к}}{\Delta I_{б}} = \frac{\alpha}{1 - \alpha} \gg 1. \quad (3.8)$$

Полученное соотношение говорит о том, что даже небольшое изменение тока I_B в схеме включения транзистора с общим эмиттером приводит к резкому изменению коллекторного тока.

Еще большего усиления по току можно добиться в схеме включения с общим коллектором, когда входной сигнал подают в цепь базы, а выходной получают в цепи эмиттера.

Очевидно, что для повышения коэффициента усиления по току α число рекомбинаций в базе транзистора должно быть сведено к минимуму. Чем большее время затрачивается электроном на проход через базу, тем выше вероятность встречи электрона с дыркой. Поскольку движение электрона носит диффузионный характер, то средняя скорость его перемещения оказывается невысокой, и поэтому базу приходится делать тонкой.

4. Внутренний фотоэффект в полупроводниках. Фотопроводимость полупроводников

Поглощение электромагнитного излучения твердым телом осуществляется различными путями: 1) энергия излучения расходуется на перевод электронов в более высокое энергетическое состояние; 2) энергия электромагнитного поля передается кристаллической решетке и превращается в тепло.

Явлением *внутреннего фотоэффекта* в полупроводниках называются переходы электронов при поглощении света в более высокие энергетические состояния, приводящие к изменению электропроводности полупроводника. На рис. 4.1 показаны возможные переходы электронов в полупроводнике p -типа под действием света, приводящие к внутреннему фотоэффекту. Переход 1 приводит к появлению электрона в зоне проводимости и дырки в валентной зоне, он возможен при энергии фотонов $h\nu \geq \Delta E$ и соответствует собственному поглощению. При энергии фотонов $h\nu < \Delta E$ могут происходить переходы электронов с примесных уровней в зону проводимости (переход 2) или из валентной зоны на эти уровни (переход 3).

В результате переходов 1,2,3 в разрешенных зонах появляются избыточные носители, которые в присутствии электрического поля принимают участие в переносе заряда. Эту дополнительную проводимость называют *фотопроводимостью*.

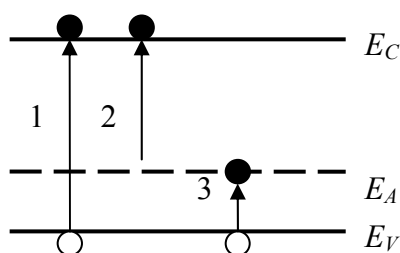


Рис. 4.1. Переходы электронов в полупроводнике p -типа при поглощении света.

На явлении внутреннего фотоэффекта основана работа многих фотоприемников. Максимальная длина волны, при которой свет еще вызывает фотопроводимость, называется длинноволновой границей внутреннего фотоэффекта. В зависимости от механизма поглощения света эта длина волны приближенно рассчитывается по формулам:

$$\lambda_{\max} = \frac{hc}{\Delta E} \quad \text{или} \quad \lambda_{\max} = \frac{hc}{E_n},$$

где ΔE – ширина запрещенной зоны, E_n – энергия активации примесных атомов (E_n соответствует энергии переходов 2 и 3 на рис. 4.1).

Подбором полупроводника и его примесного состава можно создавать фотоэлементы (фотосопротивления или фоторезисторы), чувствительные к свету от ультрафиолетовой до далекой инфракрасной области спектра. В частности, наиболее чувствительные фотосопротивления в видимой области спектра изготавливают из сульфида кадмия (фотопроводимость может в 10^5 – 10^6 раз превышать темновую проводимость). Для инфракрасного диапазона широко используются фотосопротивления из сернистого свинца.

Если к фотосопротивлению приложить напряжение U , то через него будет протекать фототок

$$I_{\phi} = A\Delta\sigma U,$$

где A – коэффициент, определяемый геометрическими размерами фоточувствительного слоя, $\Delta\sigma$ – фотопроводимость.

Фотопроводимость $\Delta\sigma$ вычисляется по формуле:

$$\Delta\sigma = eb_n\Delta n + eb_p\Delta p,$$

где Δn и Δp – концентрации фотоэлектронов и фотодырок, соответственно; b_n , b_p – их подвижности; e – заряд электрона.

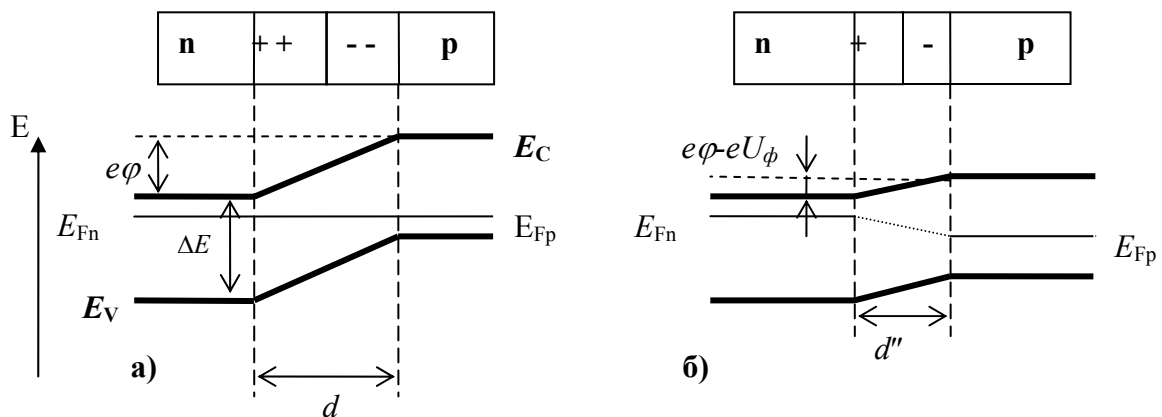
Отношение фототока I_{ϕ} к потоку излучения Φ , падающего на фотосопротивление, называется *токовой чувствительностью*

$$S = \frac{I_{\phi}}{\Phi}.$$

Другими важными параметрами фотоматериала являются: *квантовый выход* внутреннего фотоэффекта η , определяемый числом носителей (пар), которые создаются каждым поглощенным фотоном, и *скорость генерации* G , которая представляет собой число носителей, создаваемых светом в единице объема за единицу времени.

Фотозффект в p - n -переходе

Рассмотрим работу фотоэлементов с запирающим слоем (диодов с p - n -переходом). *Электронно-дырочный* или p - n -переход - это контакт двух полупроводников с различными типами проводимости (n - и p -типов). На рис. 4.2 показана зонная схема p - n -перехода, находящегося в состоянии термодинамического равновесия и в отсутствие внешнего напряжения: (а) без освещения p - n -перехода; (б) при освещении p - n -перехода.



E_C, E_V – дно зоны проводимости и потолок валентной зоны, соответственно;
 ΔE – ширина запрещенной зоны;
 E_{Fn}, E_{Fp} – уровни Ферми в полупроводниках n - и p -типа;
 d – толщина контактного слоя.

Рис. 4.2. p - n -переход в состоянии термодинамического равновесия и в отсутствие внешнего напряжения: (а) без освещения p - n -перехода; (б) при освещении p - n -перехода.

P - n -переход при отсутствии освещения

Рассмотрим физические процессы, происходящие в p - n -переходе при отсутствии освещения. Напомним, что в полупроводнике n -типа *основными* носителями свободного заряда являются электроны, их концентрация значительно превышает концентрацию дырок – *неосновными* носителями заряда. В полупроводнике p -типа *основными* носителями являются дырки, а *неосновными* – электроны. Уровни Ферми в полупроводниках при комнатной температуре до контакта расположены на разной высоте: в полупроводнике n -типа – вблизи дна зоны проводимости, а в полупроводнике p -типа – у потолка валентной зоны.

Пусть полупроводник n -типа приводится в контакт с полупроводником p -типа. Из-за градиента концентрации носителей заряда возникает их диффузия в области с противоположным типом проводимости через границу раздела: из n -области в p -область диффундируют электроны, а навстречу им – дырки. Такой установившийся направленный перенос основных носителей заряда приводит к возникновению тока, который называется *диффузионным* (I_d). Диффузионный ток, создаваемый переходом электронов, складывается с диффузионным током, создаваемым встречным потоком дырок, т.е.

$$I_d = I_n + I_p.$$

Электроны, перешедшие из n -области в p -область, рекомбинируют здесь с дырками. Рекомбинация происходит в тонком слое этой области, прилегающем к границе раздела. В результате рекомбинации отрицательный заряд ионов акцепторной примеси в этом слое не компенсируется теперь положительным зарядом дырок и весь слой заряжается отрицательно. В n -области в результате рекомбинации электронов с перешедшими сюда из p -области дырками пограничный слой заряжается положительно благодаря нескомпенсированному положительному заряду ионов донорной примеси. Таким образом, в области контакта образуется двойной электрический слой объемных зарядов (d), а, следовательно, *контактное электрическое поле*, напряженность которого направлена от n -области к p -области.

Приконтактная область обладает большим электрическим сопротивлением, т.к. концентрация свободных носителей заряда в ней очень мала. Эта область обычно достигает толщины порядка десятков и сотен межатомных расстояний. Контактное электрическое поле препятствует дальнейшей диффузии основных носителей заряда навстречу друг другу.

На рис.6а показана зонная схема p - n перехода, находящегося в термодинамическом равновесии. Диффузионное перетекание основных носителей, приводящее к появлению нескомпенсированных зарядов в приконтактных областях, вызывает смещение энергетических уровней. В n -области, зарядившейся положительно, уровни понижаются, а в p -области – смещаются вверх. Перемещение уровней происходит до тех пор, пока уровни Ферми (E_{Fn} и E_{Fp}) не расположатся на одной высоте, что соответствует установлению состояния равновесия. Сдвиг энергетических зон соответствует образованию в области контакта *потенциального барьера*, высота которого равна $e\varphi$, где φ - *контактная разность потенциалов*. Высота потенциального барьера

$$e\varphi = E_{F_p} - E_{F_n}.$$

Контактное электрическое поле является регулятором перехода носителей через p - n -переход. Преодолеть потенциальный барьер могут лишь те основные носители, энергия которых больше потенциального барьера. Но высокоэнергетичных электронов и дырок мало, т.к. основная часть электронов находится вблизи дна зоны проводимости, а дырок – вблизи потолка валентной зоны. Поэтому потоки основных носителей через область перехода невелики, хотя общее число этих носителей огромно.

Вместе с тем, контактное электрическое поле способствует переходу через границу раздела неосновных носителей, создающих *дрейфовый* ток (I_s), который направлен навстречу диффузионному. Дрейфовый ток имеет также две составляющие: электронную и дырочную.

$$I_s = I_{ns} + I_{ps}.$$

Величина дрейфового тока невелика, т.к. определяется концентрацией неосновных носителей и практически не зависит от значения контактной разности потенциалов φ . Сумма токов диффузионного и дрейфового токов в условиях равновесия равна нулю.

При обычных температурах контактный слой является *запирающим*. Сопротивление запирающего слоя можно изменить с помощью внешнего электрического поля. Если полупроводник с p - n -переходом подключен к источнику тока так, что положительный полюс источника соединен с n -областью, а отрицательный – с p -областью, то возникающее поле совпадает по направлению с контактным электрическим полем. Так как область перехода обладает большим сопротивлением по сравнению с остальной частью полупроводника, то приложенная внешняя разность потенциалов практически вся будет падать на запирающем слое и падением напряжения на остальных участках полупроводников можно пренебречь. Внешнее напряжение U смещает уровни энергии в контактирующих областях на величину, равную eU . Сдвигаются и уровни Ферми, что указывает на нарушение состояния равновесия. Высота потенциального барьера возрастает и становится равной $e(\varphi+U)$. Приложенное напряжение препятствует диффузионному движению основных носителей. Основные носители тока как бы "отстают" от p - n -перехода, что приводит к увеличению толщины и возрастанию сопротивления контактного слоя. Однако незначительному количеству неосновных носителей зарядов из n - и p -области удастся пройти p - n переход, и в цепи будет протекать весьма небольшой ток, называемый *обратным* током. Напряжение, поданное на p - n переход в этом случае, называют *обратным*, и его принято считать отрицательным.

Если p - n -переход соединить с источником так, чтобы положительный полюс источника был соединен с p -областью, а отрицательный с n -областью, то высота потенциального барьера уменьшается на величину приложенного напряжения и становится равной $e(\varphi-U)$. Снижение потенциального барьера приводит к увеличению числа электронов n -области, обладающих энергией, необходимой для его преодоления, что обуславливает быстрый рост диффузионного тока. Кроме того, в этом случае внешнее и контактное поля в области перехода имеют противоположные направления. Поэтому результирующее поле оказывается ослабленным, из-за чего уменьшается толщина и сопротивление контактного слоя, что также способствует росту диффузионного тока. Поскольку сила дрейфового тока мала, то можно сказать, что ток, текущий через p - n -переход, полностью обусловлен потоками основных носителей (этот ток называют *прямым*). Сила тока в этом случае будет возрастать при увеличении напряжения источника. Напряжение, поданное на p - n переход в этом случае, называют *прямым* и считают положительным. Таким образом, p - n переход обладает *односторонней проводимостью*.

P-n-переход при освещении

Если на *p-n*-переход направить поток фотонов с энергией большей, чем ширина запрещенной зоны, то по обе стороны от контактного слоя и в самом слое возникают пары неравновесных носителей «электрон-дырка». Пары, генерируемые в запирающем слое, сразу разделяются контактным электрическим полем, причем электроны движутся в область с *n*-типом проводимости, а дырки – в *p*-область. Кроме того, вследствие диффузии в контактный слой попадают электроны, созданные светом в *p*-области, и дырки из *n*-области. Неосновные носители заряда, созданные светом, создают обратный ток (*фототок* I_ϕ).

С другой стороны, разделение зарядов при отсутствии внешнего напряжения приводит к понижению потенциального барьера на величину eU_ϕ , где U_ϕ – *фото-ЭДС*, и уменьшению толщины запирающего слоя (d'') (рис. 4.2 б). В результате через переход пойдет прямой ток I , который, как известно из теории *p-n*-перехода, определяется выражением:

$$I = I_s \left(e^{\frac{eU}{kT}} - 1 \right),$$

где U – внешнее напряжение, приложенное к *p-n*-переходу, I_s – значение, к которому стремится обратный ток при увеличении обратного напряжения (дрейфовый ток), k – постоянная Больцмана, T – температура. В состоянии равновесия встречные токи I_ϕ и I равны, и тогда

$$I_\phi = I_s \left(e^{\frac{eU_\phi}{kT}} - 1 \right).$$

Отсюда можно определить фото-ЭДС, которая возникает в освещенном *p-n*-переходе, который не включен в электрическую цепь (*вентильный режим работы фотодиода*).

Также существует *фотодиодный режим работы* (при наличии внешнего напряжения). Если освещенный диод включен в электрическую цепь и к нему подведено внешнее напряжение U , то в цепи возникнет ток

$$I = I_s \left(e^{\frac{eU}{kT}} - 1 \right) - I_\phi.$$

При коротком замыкании фотодиода ($U=0$) ток во внешней цепи $I = -I_\phi$, т.е. обусловлен потоком только неравновесных неосновных носителей заряда, созданных светом и разделенных контактным электрическим полем. Обычно к фотодиоду прикладывается обратное напряжение ($-U$), причем, достаточно большое, так что в этом случае

$$I = -I_s - I_\phi.$$

Вольтамперные характеристики p - n -перехода при освещении

На рис. 4.3 представлены вольтамперные характеристики (ВАХ) освещенного p - n -перехода для кремниевого диода при различных световых потоках ($\Phi_2 > \Phi_1$), а также темновая ВАХ фотодиода ($\Phi = 0$).

Вольтамперная характеристика описывается уравнением:

$$I = I_s (e^{\frac{eU}{kT}} - 1) - I_\phi.$$

Величина фототока I_ϕ прямо пропорциональна световому потоку.

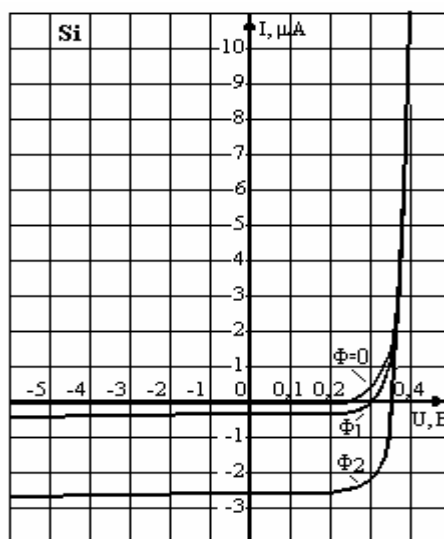


Рис. 4.3. ВАХ p - n -перехода при различных световых потоках ($\Phi_2 > \Phi_1$) и темновая ВАХ ($\Phi=0$).

При прямом напряжении, приложенном к p - n -переходу, полный ток через переход быстро растет по экспоненциальному закону по мере увеличения напряжения за счет диффузионного тока основных носителей заряда при неизменном дрейфовом токе неосновных носителей. При увеличении напряжения величина прямого тока очень быстро становится больше величины фототока I_ϕ , и поэтому все ВАХи, начиная с некоторого значения прямого напряжения, сливаются в одну кривую, которая быстро стремится вверх. Тогда как при меньших значениях прямого напряжения вольтамперные характеристики будут существенно отличаться друг от друга (рис. 4.3). При определенных значениях прямого напряжения полный ток через переход меняет свое направление и становится отрицательным. Кривые переходят в IV квадрант координатных осей, и ход их определяется величиной фототока, который прямо пропорционален световому потоку

Φ . Эта часть кривых соответствует режиму генерации фото-ЭДС (вентильный режим работы p - n -перехода).

Обратный ток через p - n -переход сначала довольно быстро растет (по модулю) за счет резкого уменьшения диффузионного тока основных носителей заряда с увеличением потенциального барьера. При дальнейшем увеличении обратного напряжения диффузионный ток падает практически до нуля, и обратный ток становится равным дрейфовому току и фототоку, которые не зависят от приложенного напряжения. При этом величина полного обратного тока практически равна величине фототока I_{ϕ} . Это уже видно из того, что обратный ток неосвещенного p - n -перехода на несколько порядков меньше обратного тока при освещении. Поэтому на рис. 4.3 обратная ветвь ВАХ неосвещенного p - n -перехода практически совпадает с осью координат x .

Если к освещенному p - n -переходу приложено обратное напряжение, то кривые ВАХ переходят в III квадрант, сохраняя величину обратного тока практически неизменной. Эта область ВАХ соответствует фотодиодному режиму работы p - n -перехода и используется для измерения освещенности, т.к. величина обратного тока, равного фототоку, прямо пропорциональна световому потоку, падающему на фотодиод.

Таким образом, ВАХ освещенного p - n -перехода (также как и неосвещенного) существенно нелинейная. p - n -переход обладает очень большим сопротивлением в случае обратного тока через переход и малым сопротивлением в случае прямого тока. Величину $R_{\text{д}} = \frac{dU}{dI}$ называют *дифференциальным сопротивлением*, характеризующим изменение напряжения при изменении тока через p - n -переход. Дифференциальное сопротивление диода при обратном токе очень велико – от 10^4 до 10^8 Ом.

Так как на ВАХ p - n -перехода имеются области с существенно различными дифференциальными сопротивлениями, то необходимая точность измерения ВАХ может быть достигнута при определенных условиях измерений. На прямой ветви ВАХ, где дифференциальное сопротивление мало, небольшие изменения прямого напряжения приводят к большим изменениям тока. Поэтому при измерении прямой ветви целесообразно задавать прямой ток через диод и измерять прямое напряжение. Требование «задавать ток через диод» означает, что внутреннее сопротивление источника питания должно быть больше сопротивления диода. Это условие распространяется и на измерение напряжения на всех участках ВАХ, где дифференциальное сопротивление мало.

На обратной ветви ВАХ, где дифференциальное сопротивление велико, изменение напряжения приводит к незначительному изменению тока. Поэтому при измерении обратной ветви ВАХ надо задавать напряжение, приложенное к диоду, и измерять ток.

Как уже указывалось, существует два режима работы фотоэлементов с электронно-дырочным переходом:

1. *Фотогальванический (вентильный) режим*, т.е. режим генерирования фото-ЭДС, - без внешнего источника питания.

2. *Фотодиодный режим*, - когда на p - n -переход подано обратное напряжение.

Фотогальванический режим соответствует IV квадранту вольтамперной характеристики. Участок ВАХ, расположенный в IV квадранте рис. 4.3, представляет собой вольтамперную ха-

рактеристику фотогальванического элемента, т.е. элемента, который является преобразователем световой энергии в электрическую. Несколько десятков таких элементов, соединенных последовательно, образуют «солнечную» батарею. Такие батареи применяются для питания радиоаппаратуры на космических станциях и спутниках Земли.

Фотодиодный режим соответствует III квадранту рис. 4.3. При изменении светового потока вольтамперные характеристики смещаются параллельно оси абсцисс на величину, пропорциональную световому потоку. Зависимость тока фотодиода от светового потока при постоянном обратном напряжении называется *световой характеристикой* фотодиода. Поскольку фотодиоды имеют линейную световую характеристику, они используются для фотометрии. Фотодиоды имеют преимущества перед другими фотоэлементами: малый вес и габариты, высокая чувствительность, небольшие рабочие напряжения.

Литература

1. Савельев И.В. Курс общей физики. – М.: Наука, т. 2,3, 1989-2003.
2. Калашников С.Г. Электричество. - М.: Наука, 1977.
3. Матвеев А.Н. Курс общей физики. - М.: Высшая школа, т. 2, 1976-1989.
4. Верещагин И.К., Кокин С.М., Никитенко В.А., Селезнев В.А., Серов Е.А. Физика твердого тела. - М., Высшая школа, 2001.