

4.1. СОБСТВЕННЫЕ И ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ

Известно большое число веществ, которые по своим свойствам относятся к классу полупроводников. Что же такое полупроводник? Со школы известно историческое, но не физическое определение полупроводника основано на различной способности веществ проводить электрический ток, т.е. на различии в величинах удельной электрической проводимости. Издавна было замечено, что хорошо проводящие ток материалы обычно являются металлами. Величины их удельной проводимости σ составляют более 10^4 (Ом·см)⁻¹. Вещества, обладающие малой удельной проводимостью, порядка 10^{-7} (Ом·см)⁻¹ и менее, всегда относили к изоляторам (диэлектрикам). Все материалы, которые имели $\sigma = 10^4 - 10^7$ (Ом·см)⁻¹, принято было называть полупроводниками.

Развитие учения о полупроводниках, познание их природы показало недостаточность такого определения. Известно, что при чрезвычайно большом содержании примесей, даже у таких хорошо известных полупроводниковых кристаллов, как германий и кремний, удельная электрическая проводимость становится равной примерно 10^4 (Ом·см)⁻¹, но от этого они не перестают быть полупроводниками.

Таким образом, количественное отличие полупроводников от металлов и диэлектриков по величине удельной электрической проводимости, хотя и имеет место в большинстве случаев, но вовсе не является необходимым. Академик А.Ф.Иоффе обратил внимание на то, что от металлов полупроводники прежде всего отличаются не величиной, а характером зависимости удельной электрической проводимости от температуры (рис.4.1): у металлов она слабо уменьшается, а у полупроводников она сильно увеличивается при нагревании кристалла.

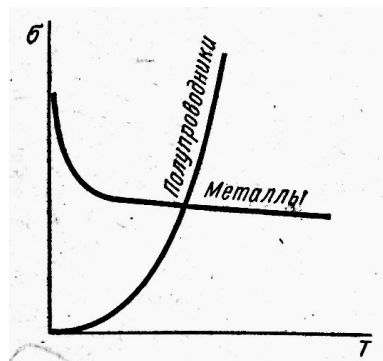


Рис. 4.1. Температурная зависимость удельной электрической проводимости металлов и полупроводников. Таким образом, по определению А.Ф.Иоффе, полупроводниками являются вещества, электропроводность которых обусловлена переносом электронов и возрастает при повышении температуры.

Еще несколько лет назад определение, данное Иоффе, было более чем достаточным, но сегодня и оно вызывает ряд возражений. Во-первых, у сильно легированных полупроводников электропроводность с повышением температуры не растет. А слабо уменьшается, почти как у металлов. Во-вторых, из приведенного выше определения следует, что у полупроводников в отличие от металлов принципиально не может существовать явление сверхпроводимости, то есть рост σ при стремлении температуры к абсолютному нулю (см. рис. 4.1), а это утверждение в настоящее время весьма обоснованно оспаривается теоретиками твердого тела.

Таким образом, дать какое-либо единое определение понятия полупроводника весьма затруднительно.

Важно, что наиболее существенной особенностью полупроводников является их способность изменять свои свойства в чрезвычайно широких пределах под влияние различных воздействий (температура, освещение, электрическое и магнитное поля, внешнее гидростатическое давление и т.д.) Под их действием характеристики полупроводников, например, электропроводность могут изменяться в 10^6-10^7 раз. Такой чувствительностью не обладают ни металлы, ни диэлектрики. Столь большая чувствительность и обусловила широкое применение полупроводников в технике. Чувствительность к температуре лежит в основе действия таких приборов, как термисторы. На большой чувствительности к свету основано действие полупроводниковых фотоприемников. Чувствительность к слабым сигналам электрического напряжения позволила создать различные полупроводниковые диоды и триоды. Чувствительность к магнитному полю используют при создании полупроводниковых приборов для измерения весьма слабых величин напряженности магнитного поля и т. д.

В чем же причина столь большой чувствительности полупроводников к внешним воздействиям? Для того, чтобы это понять, вернемся к численным значениям удельной электрической проводимости металлов и полупроводников. Примем для оценки их величины σ_m и σ_p , соответственно равными 10^5-10^6 и $10^{-3} - 10^{-2}$ (Ом·см)⁻¹. Величина σ связана с концентрацией электронов n и их подвижностью u соотношением

$$\sigma = neu. \quad (4.1)$$

Примем подвижность электронов в металле и полупроводнике одинаковыми и равными $100 \text{ см}^2/(\text{В}\cdot\text{с})$. Тогда для металла из (4.1) получаем концентрацию электронов, равную $5 \cdot 10^{21} \text{ — } 5 \cdot 10^{22} \text{ см}^{-3}$, а для полупроводника — $5 \cdot 10^{13} \text{ — } 5 \cdot 10^{14} \text{ см}^{-3}$. Таким образом, в металле число электронов, участвующих в проводимости, близко к числу атомов в 1 см^3 кристалла. Следовательно, каждый или почти каждый атом в кристаллической решетке металла отдал электрон для переноса электрического заряда. По-иному обстоит дело в полупроводниковом кристалле. Количество электронов проводимости в нем на несколько порядков по величине меньше числа атомов, т. е. далеко не каждый атом отдал электрон для участия в явлении проводимости. Следовательно ту же проводимость, например, в полупроводнике еще можно увеличить, если каким-либо способом ионизировать еще дополнительное количество атомов в кристаллической решетке. Малое число носителей заряда приводит к тому, что их увеличение или уменьшение существенно сказывается на свойствах полупроводников. Следовательно, если уметь управлять количеством свободных носителей заряда в полупроводнике, то можно управлять и свойствами полупроводника. Это управление подразумевает, во-первых, создание носителей заряда и, во-вторых, управление их движением внутри полупроводника.

4.1.1. Некоторые сведения о строении атома

Физические свойства полупроводников тесно связаны со структурой валентных оболочек атомов, из которых они образованы. Поэтому рассмотрим качественно этот вопрос.

Как известно, атом любого элемента состоит из положительно заряженного ядра с зарядом ze (где z — порядковый номер атома в системе Менделеева, e — заряд электрона) и z электронов, расположенных в виде системы оболочек с постепенно возрастающими размерами. Первая оболочка носит название K -оболочки, вторая — L -оболочки, третья — M -оболочки и т.д. Первая, K -оболочка, имеет только одну подоболочку, обозначаемую s , вторая (L) делится на s - и p -подоболочки, третья (M) — на s -, p - и d -подоболочки и т. д. В упрощенной модели электронного строения атома предполагается, что электрон подобен заряженному волчку, обладающему собственным магнитным и механическим моментами (называемыми спинами), а каждый атом подобен солнечной системе, в которой электроны движутся вокруг ядра по эллиптическим орбитам, причем номер оболочки определяет длину большой полуоси орбиты, а номер подоболочки — ее форму (эксцентриситет). Каждая подоболочка, начиная со второй, состоит из ряда орбит, отличающихся друг от друга своей ориентацией в пространстве; первая s -подоболочка, имеет только одну орбиту, на которой могут разместиться только два электрона с противоположно направленными спинами. Вторая, p -подоболочка, состоит из трех орбит, на каждой из которых может разместиться по 2 электрона; таким образом, всего во второй подоболочке может разместиться 6 электронов; третья, d -подоболочка, состоит из пяти орбит, на которых может разместиться 10 электронов, и т.д. Номер оболочки называется главным квантовым числом и обозначается n , номер подоболочки называется орбитальным квантовым числом и обозначается l . Орбитальное квантовое число определяет эксцентриситет орбиты, а следовательно, и орбитальный момент количества движения электрона, находящегося на ней. Полный момент количества движения электрона складывается из орбитального и собственного (спина) и определяется третьим квантовым числом j , которое в зависимости от ориентации спина может принимать значения $j_l = l + 1/2$ и $j_l = l - 1/2$. Наконец, ориентация орбиты в пространстве характеризуется четвертым магнитным квантовым числом m .

Согласно принципу Паули в атоме не могут существовать электроны с одинаковой четверкой квантовых чисел. Этим принципом и определяется число электронов, помещающихся на одной орбите (2), в одной подоболочке ($2l + 1$) и в одной оболочке — $2n^2$.

В таблице 1 приведены максимальные числа электронов, которые могут разместиться в каждой подоболочке и полной число электронов, которое размещается в данной оболочке.

Таблица 1

Номер оболочки (n)	Номер подоболочки					Z_n
	0(s)	1(p)	2(d)	3(f)	4(g)	
1 (K)	2	—	—	—	—	2
2 (L)	2	6	—	—	—	8
3 (M)	2	6	10	—	—	18
4 (N)	2	6	10	14	—	32
5 (P)	2	6	10	14	18	50

4.1.2. Радиусы орбит и энергия электрона в атоме

Радиусы круговых орбит и энергии электрона на них могут быть найдены из двух условий:

1 — условия классической устойчивости — равенства центробежной и центростремительной силы

$$\frac{m_0 v^2}{r} = \frac{z e^2}{r^2}, \quad (4.2)$$

2 — условия квантовой устойчивости. Согласно квантовой механики электрон, движущийся со скоростью v , описывается волной, длина которой $\lambda = h/m_0 v$, где m_0 — масса электрона и h — универсальная постоянная, так называемая постоянная Планка, равная $6.57 \cdot 10^{-27}$ эрг/сек. Для того, чтобы орбита была устойчивой, на ней должно укладываться целое число волн (для того, чтобы при многократном прохождении орбиты электронная волна находилась в фазе сама с собой, а не гасила самое себя):

$$2\pi r = n\lambda = \frac{nh}{m_0 v}. \quad (4.3)$$

Решая совместно (4.2) и (4.3), находим

$$r_n = n^2 \frac{\hbar^2}{m_0 e^2} \frac{1}{z}, \quad (4.3)$$

где $\hbar = h/2\pi$.

Величина r равна $a_0 = \hbar^2/m_0 e^2$ при $n=1$ и $z=1$, т.е. представляет собой радиус первой (самой близкой к ядру) орбиты в атоме водорода ($a_0 = 0.529 \cdot 10^{-8}$ см). Радиусы других орбит в атоме водорода или одноэлектронных ионах с зарядом ядра z выражаются через a_0 :

$$r_n = n^2 \frac{a_0}{z}. \quad (4.4)$$

Полная энергия электрона на n -ой орбите равна сумме его кинетической и потенциальной энергии:

$$E_n = T_n + U_n = \frac{1}{2} m_0 v^2 - \frac{z e^2}{r}. \quad (4.5)$$

Используя (4.2) и (4.4), получаем

$$E_n = -\frac{z e^2}{2r} = -\frac{m_0 e^4 z^2}{2\hbar^2 n^2}. \quad (4.6)$$

Одним из основных постулатов квантовой механики является предположение, что, находясь на стабильной орбите, электрон не излучает энергии, а при переходе с одной орбиты на другую излучает (или поглощает) квант энергии, величина которого равна разности энергий электрона в начальном и конечном состояниях:

$$(4.7)$$

Формула (4.7) дает точное значение частот спектра испускания и поглощения атома водорода. Этот факт был первым триумфом квантовой теории атома Бора, основанной на модели круговых орбит. Однако теория Бора, имея ряд внутренних противоречий не объясняет особенности спектров более сложных атомов, требовавших учета эллиптических орбит, релятивистских эффектов спина и магнитного момента электрона.

4.1.3. Энергетический спектр электронов в кристалле

Проследим, как изменяется характер движения и энергетический спектр валентных электронов в твердом теле. При образовании твердого тела соседние атомы настолько сближаются друг с другом, что внешние электронные оболочки не только соприкасаются, но даже перекрываются. В результате этого характер движения электронов резко изменяется: электроны, находящиеся на определенном энергетическом уровне одного атома, получают принципиальную возможность переходить без затраты энергии на соответствующий уровень соседнего атома и таким образом свободно перемещаться вдоль всего твердого тела, то есть участвовать в электропроводности при наличии электрического поля.

Вместо индивидуальных атомных орбит образуются коллективные, и подоболочки отдельных атомов объединяются в единый для всего кристалла коллектив — зону. Однако при этом число электронов, которые могут разместиться в данных подоболочках атомов, образующих кристалл, остается неизменным.

Если для изолированного атома это число было $2g$ (для s -подоболочки $g=1$, для p -подоболочки $g=3$ и т.д.), где g — число энергетических уровней в подоболочках атомов в энергетической зоне, и поскольку число атомов в кристалле N , то общее число мест в зоне по-прежнему равно $2Ng$. Однако энергетический спектр электронов при этом претерпевает существенные изменения. Если в изолированном атоме каждой

подоболочке соответствует строго определенная энергия, то зона состоит из Ng уровней, расположенных очень близко друг к другу, причем на каждом уровне, в соответствии с принципом Паули, могут разместиться только два электрона. Различным уровням одной зоны соответствует различная скорость поступательного движения электрона по кристаллу (кинетическая энергия) и различная потенциальная энергия, чем и объясняется различное положение уровней в зоне.

На рис.4.2 схематически показано, как энергетические уровни электронов в изолированном атоме трансформируются в энергетические зоны в кристалле.

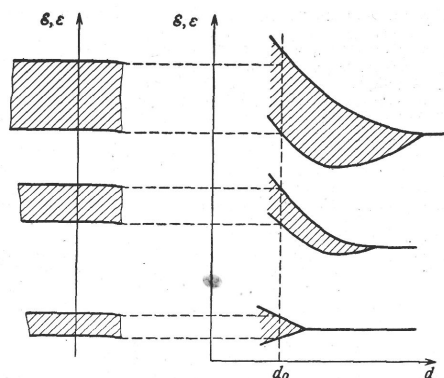


Рис.4.2. Образование энергетических зон в кристалле из атомных энергетических уровней: d — расстояние между соседними атомами; d_0 — равновесное расстояние между соседними атомами в кристалле.

Так как в нейтральном изолированном атоме внутренние энергетические уровни полностью заполнены, а не полностью заполненным может быть лишь крайний валентный уровень, то и заполнение зон в кристалле оказывается различным. Зоны, произошедшие от внутренних уровней, всегда полностью заполнены электронами.

Только зона, образованная от валентных уровней, может оказаться в ряде случаев заполненной неполностью. Зону, произошедшую от валентных уровней атомов, образующих кристалл, называют валентной зоной.

Степень заполнения валентной зоны электронами зависит прежде всего от химической природы атомов, из которых состоит кристалл, определяемой величиной Z , от кристаллической структуры и от многих других более тонких факторов.

В случае одноатомных кристаллов с простой структурой, т.е. с базисом, состоящим из одного атома, полное число электронов в кристалле будет NZ , где N — число атомов в кристалле, а Z — число электронов, приходящихся на один атом. Поэтому число квантовых состояний (энергетических уровней) в каждой отдельной разрешенной зоне равно $2N$.

В изолированном атоме выше валентного уровня находятся уровни энергии, на которых хотя электронов и нет, но они могут там находиться при возбуждении атомов. Для этого валентному электрону необходимо преодолеть энергетический барьер, равный потенциалу возбуждения атома. Точно так же и в кристалле: выше валентной зоны имеется полностью свободная от электронов зона, которую называют зоной проводимости. Заполнение зоны проводимости может начинаться только, если электроны в валентной зоне получают дополнительную энергию, достаточную для преодоления энергетического барьера, равного ширине запрещенной зоны, лежащей между валентной зоной и зоной проводимости.

На рис.4.3 изображены возможные случаи зонной структуры с учетом степени заполнения зон. Обсудим их.

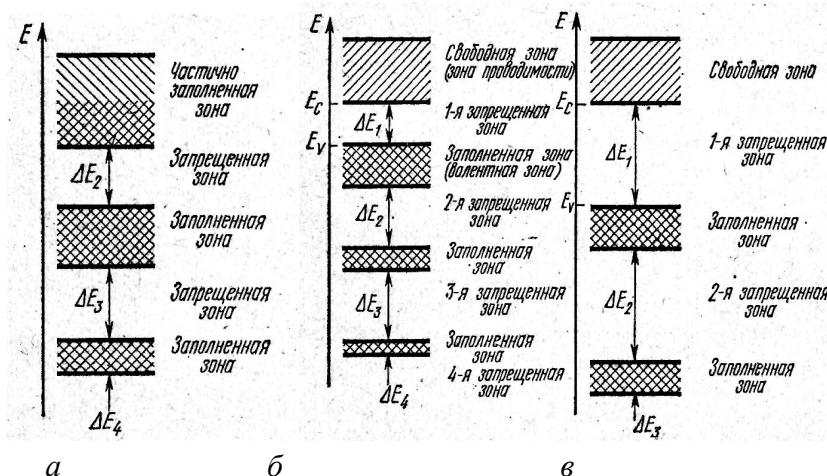


Рис.4.3. Возможные схемы зонной структуры твердых тел: a — металл, b — полупроводник, v — диэлектрик.

В общем случае, схема *a*, валентная зона заполнена полностью, но запрещенная зона между валентной зоной и зоной проводимостью отсутствует. При приложении к такому кристаллу электрического поля электроны будут получать от него энергию, и такой кристалл будет хорошо проводить электрический ток, его электропроводность не будет зависеть (или почти не будет зависеть) от температуры. Таким образом, схема *a*, в которой валентная зона и зона проводимости оказываются перекрывающимися друг с другом, соответствует зонной структуре металлического кристалла.

Схемы *б* и *в* на рис.4.3 отражают случай, когда валентная зона заполнена полностью и не перекрыта с зоной проводимости. При этом перемещение электронов под действием электрического поля не может происходить до тех пор, пока каким-либо способом, например, нагревом, не будет переведена часть электронов из валентной зоны в зону проводимости. Таким образом эти схемы отражают энергетические зоны в кристаллах, электропроводность которых должна повышаться с увеличением температуры. Как было определено выше, такие кристаллы называют полупроводниками.

Различие в схемах *б* и *в* на рис.4.3 заключается лишь в том, что величины запрещенных зон разные. Понятно, что для преодоления большего энергетического зазора, рис.4.3 *в*, потребуется более высокая температура. Поэтому кристаллы, у которых проводимость возникает при очень высоких температурах, принято относить к диэлектрикам. Условно принято считать, что кристаллы такого типа будут диэлектриками, если ширина запрещенной зоны ΔE_g будет больше 2 эВ. Хотя понятно, что с физической точки зрения и такие кристаллы являются полупроводниками.

Анализ степени заполнения зон позволяет в дальнейшем исключить из рассмотрения все глубокие зоны, лежащие ниже валентной, так как электроны в этих зонах не участвуют в прохождении электрического тока через кристалл. Таким образом, зонные структуры, которые мы будем рассматривать в дальнейшем, принято изображать так, как она показана на рис.4.4. Обычно для краткости изображают лишь потолок валентной зоны, обозначаемый E_v , подразумевая, что сама валентная зона располагается ниже уровня E_v . Точно так же принято изображать дно зоны проводимости, то есть ее нижний уровень, обозначаемый E_c .

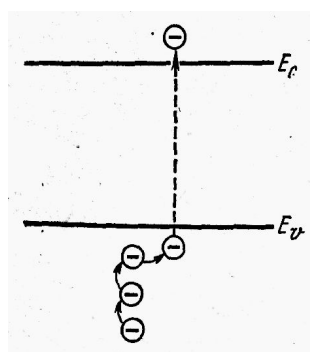


Рис.4.4. Схема возникновения движения электронов в валентной зоне.

4.1.4. Проводимость в собственных полупроводниках

Собственными полупроводниками являются такие полупроводники, в которых механизм появления зарядов (электронов и дырок) определяется характером собственного энергетического спектра кристалла и ничем не связан с содержанием примесей.

В собственных полупроводниках электроны, находясь в валентной зоне, не могут участвовать в электрическом токе. Характер теплового движения атомов в твердом теле и в газе существенно различается: в газе атомы хаотически движутся в пределах всего объема, в твердом теле — колеблются около равновесных положений, так как в этом случае атомы связаны друг с другом силами сцепления. Колебания одного атома передаются другому и вдоль всего кристалла распространяются в самых разнообразных направлениях упругие волны. Если температура всего кристалла одинакова, то одинакова и средняя кинетическая энергия тепловых

колебаний атомов, и при достаточно высоких температурах она равна $\frac{3}{2}kT$, что при комнатной температуре

составляет 0.04 эВ. Но кинетическая энергия атомов газа только в среднем равна $\frac{3}{2}kT$. Мгновенные же скорости распределяются по закону Максвелла, то есть всегда имеется некоторое число атомов, скорости которых намного больше и намного меньше средних. Вероятность того, что атом имеет энергию ΔE_g , пропорциональна $\exp(-\Delta E_g/kT)$. Это же относится и к тепловым колебаниям твердого тела; в каждый данный момент имеется небольшое число атомов, амплитуда и энергия колебания которых значительно больше средней.

Атомы при своих колебаниях взаимодействуют не только друг с другом, но и с электронами; колеблющийся атом может передать всю или часть своей энергии электрону. В результате этого энергия электрона

увеличится и он перейдет на более высокий энергетический уровень. Все это, разумеется возможно, если уровень на который должен перейти электрон в результате взаимодействия («столкновения») с атомом, свободен.

Этот акт взаимодействия называют тепловым возбуждением электрона, в результате которого электрон может «перескочить» из заполненной валентной зоны в свободную, называемую зоной проводимости и уже сможет участвовать в электрическом токе. Появившееся при этом в валентной зоне пустое место называют дыркой, которая создает в валентной зоне возможность эстафетного (дырочного) механизма проводимости: какой-либо электрон занимает освободившееся место, его место занимает другой и т.д., рис.4.4.

Число тепловых возбуждений в секунду (P) пропорционально числу электронов вблизи верхнего края валентной зоны C_v , числу пустых мест вблизи нижнего края зоны проводимости C_c и вероятности того, что какой-либо атом приобретет при своих колебаниях энергию ΔE_g , достаточную для переброса электрона через запрещенную зону (рис.1.4) в зону проводимости. Согласно изложенному эта вероятность пропорциональна экспоненциальному множителю $e^{-\Delta E_g/kT}$. Следовательно, $P = aC_cC_v e^{-\Delta E_g/kT}$, где a — некоторый коэффициент, зависящий от частоты «столкновений» атомов и электронов.

Чем больше ширина запрещенной зоны ΔE_g , тем меньше вероятность этого процесса (а следовательно, и проводимость рассматриваемого кристалла).

Оценка показывает, что при ширине запрещенной зоны $\Delta E_g \gg 2$ эВ эта вероятность остается пренебрежимо малой при всех доступных нам температурах и поэтому вещества, в которых выполняется это неравенство, не проводят электрический ток и называются диэлектриками. Разумеется, граница $\Delta E_g = 2$ эВ условна и на самом деле между диэлектриками и полупроводниками нет резкой грани.

При температуре абсолютного нуля валентная зона полупроводника заполнена целиком, зона проводимости пуста и электропроводность равна нулю. С повышением температуры экспоненциально растет число тепловых забросов электронов, следовательно, растет число электронов в зоне проводимости и дырок в валентной и соответственно электропроводность полупроводника.

Итак, число тепловых забросов в секунду, которое мы обозначим g_0 , растет экспоненциально с температурой:

$$g_0 = \alpha e^{-\frac{\Delta E}{kT}}, \quad (4.8)$$

где $\alpha = aC_cC_v$, с другой стороны, если уже имеется некоторое число электронов в свободной зоне и дырок в валентной, то одновременно идет обратный процесс — рекомбинация электронов и дырок: часть электронов падает на свободные места и сразу же исчезают один свободный электрон и одна дырка. Число рекомбинаций в секунду пропорционально вероятности «встречи», то есть произведению числа электронов n и числа дырок p :

$$r = \gamma np, \quad (4.9)$$

так как в рассматриваемом случае $n=p$, здесь γ — коэффициент пропорциональности

Если кристалл находится при данной температуре достаточно долго, то должно установиться равновесие:

$$g_0 = r. \quad (4.10)$$

Отсюда

$$n = p = \sqrt{\frac{g_0}{\gamma}}. \quad (4.11)$$

Температурная зависимость электропроводности любого материала определяется температурными зависимостями концентрации носителей тока и их подвижности. Подвижность электронов в полупроводниках, также как и в металлах, зависит от температуры обычно по степенному закону:

$$\mu = aT^b, \quad (4.12)$$

где $0 \leq a \leq +1$ и для полупроводников — $3/2 \leq b \leq 3$. Но в металле концентрация свободных электронов неизменна, и поэтому температурная зависимость электропроводности целиком определяется подвижностью. В полупроводниках же подвижность носителей n очень сильно зависит от температуры и по сравнению с этой зависимостью температурная зависимость подвижности играет слабую роль.

Поэтому можно считать, что электропроводность полупроводников в первом приближении растет с температурой примерно по тому же закону, что и концентрация электронов и дырок:

$$\sigma = nq\mu. \quad (4.13)$$

Итак, с феноменологической точки зрения полупроводники отличаются от металлов тем, что в металлах с повышением температуры электропроводность относительно медленно падает, а в полупроводниках — очень быстро растет. Температурная зависимость электропроводности полупроводника изображается обычно на графиках в так называемых полулогарифмических координатах. Если прологарифмировать выражение (4.13),

то оно примет вид

$$\ln \sigma = \ln A - \frac{\Delta E_g}{2k} \frac{1}{T} \quad (4/14)$$

Следовательно, если на графике по оси ординат откладывать $\ln \sigma$, а по оси абсцисс — $1/T$, то получим прямую с наклоном $\Delta E_g / 2k$, рис. 4.5.

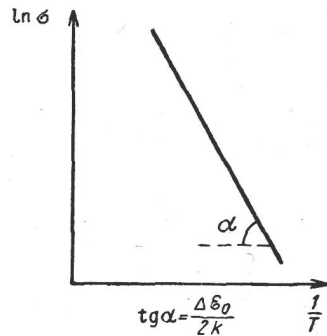


Рис.4.5. Температурная зависимость электропроводности собственного полупроводника (в полулогарифмических координатах).

Таким образом, зная наклон этой прямой, можно получить важнейшую характеристику полупроводника — ширину запрещенной зоны.

В таблице 4.2 представлены величины значений запрещенных зон наиболее важных полупроводников при комнатной температуре.

Таблица 4.2

№ группы таблицы химических элементов	IV		IV	III-V		II-VI	
	Ge	Si	α -SiC	GaP	GaAs	ZnTe	CdS
ΔE_g , эВ	0.7	1.1	3.0	2.2	1.4	0.9	2.4

4.1.5. Проводимость в примесных полупроводниках

Полупроводники, которые рассматривались в предыдущем параграфе, называются собственными полупроводниками, а их проводимость — собственной проводимостью. Название это происходит от того, что механизм появления электронов и дырок, рассмотренный выше, определяется характером собственного энергетического спектра кристалла и ничем не связан с содержанием примесей.

Наряду с этим имеется широкий класс веществ, в которых концентрация носителей определяется примесями. Эти вещества называются примесными полупроводниками.

Рассмотрим механизм образования свободных электронов и дырок в этих полупроводниках. Если в решетку кристалла входит атом другого вещества, то часть энергетических уровней этого атома попадет в запрещенный промежуток между свободной и заполненной зоной. При этом надо различать два важных случая.

A. — Энергетический уровень, занятый валентным электроном, оказался размещенным недалеко (на расстоянии) от нижнего края зоны проводимости, рис.4.6.

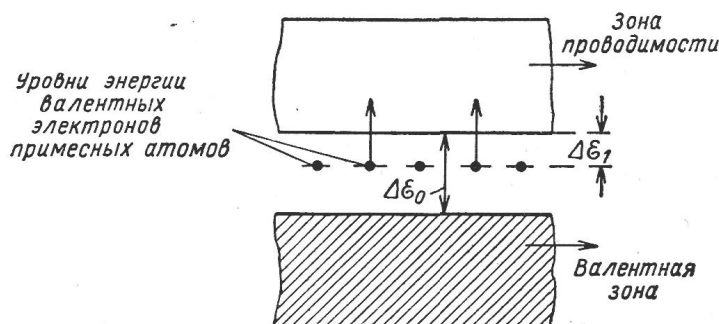


Рис.4.6. Энергетический спектр примесного электронного полупроводника.

Электрон, находясь на этом уровне, может участвовать в проводимости, но энергия теплового возбуждения, необходимая для перевода этого электрона в свободную зону, относительно мала. Поэтому такие переходы будут играть основную роль при низких температурах, при которых вероятность теплового возбуждения из заполненной зоны, а следовательно, и собственная проводимость ничтожно малы. Число электронов, попавших в зону проводимости с примесных уровней (по аналогии с выводом для собственных полупроводников), выражается формулой

$$n_1 = \alpha_1 e^{-\frac{\Delta E_1}{2kT}}, \quad (4.15)$$

где α_1 — коэффициент пропорциональный корню из концентрации примесных атомов N_1 и C_v :

$$\alpha_1 \sim \sqrt{N_1 C_v}. \quad (4.16)$$

Таким образом, при низких температурах в проводимости такого кристалла доминирующую роль играют электроны, перешедшие в зону проводимости с примесных уровней (число электронов и дырок собственной проводимости будет во много раз меньше). Такой механизм проводимости называется *примесным*, а полупроводник в этом случае — *примесным электронным полупроводником*. Атомы, отдающие электроны в зону проводимости, называются донорами, а соответствующие энергетические уровни — донорными уровнями. Для полупроводников IV группы таблицы Менделеева (Ge, Si) в качестве доноров используются атомы примесей P, As, Sb, имеющих на внешней оболочке пять валентных электронов.

Если прологарифмировать выражение (4.15), то оно примет вид

$$\ln n_1 = \ln \alpha_1 - \frac{\Delta E_1}{2k} \frac{1}{T}, \quad (4.17)$$

Соответствующий график, следовательно, имеет вид прямой с наклоном $-\Delta E_1 / 2k$.

Однако формула (4.17) справедлива лишь до тех пор, пока число электронов в зоне проводимости много меньше, чем общее число электронов на донорных уровнях. При достаточно высокой температуре все электроны с донорных уровней перейдут в зону проводимости и, следовательно, концентрация свободных электронов будет постоянна вплоть до температур, при которых начнет проявляться собственная проводимость. Таким образом, температурная зависимость концентрации свободных электронов в широком интервале температур будет иметь вид, представленный на рис.4.7. При высоких температурах примесный полупроводник становится собственным.

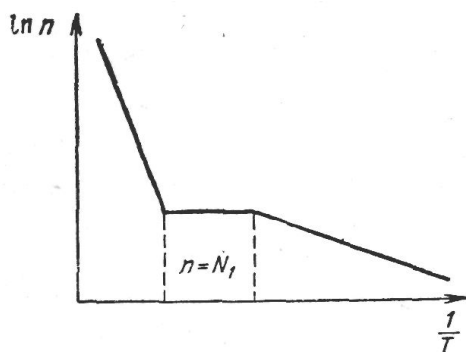


Рис.4.7. Температурная зависимость концентрации электронов в примесном полупроводнике.

Б. — Уровень, занятый валентным электроном примесного атома, расположен ниже верхнего края валентной зоны; следующий за ним свободный уровень расположен немного выше края заполненной зоны.

Валентные электроны примесных атомов, попав в общую массу валентных электронов кристалла, не будут сколько-нибудь заметно влиять на электропроводность. В этом случае оказывается весьма существенной роль первых свободных уровней примесных атомов, рис.4.8.

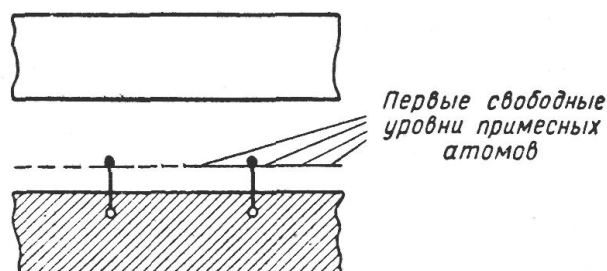


Рис.4.8. Энергетический спектр примесного дырочного полупроводника.

При температуре, отличной от абсолютного нуля, часть электронов из заполненной зоны будет забрасываться на свободные уровни примесных атомов, в заполненной зоне появляются пустые места — дырки, и поэтому становится возможной дырочная проводимость.

В электрическом поле электроны движутся от минуса к плюсу, а дырки движутся в противоположном направлении, то есть ведут себя как положительные заряды. Таким образом сложное эстафетное движение электронов можно рассматривать как перемещение фиктивной положительно заряженной частицы – дырки. В соответствии с изложенным выше, такой механизм проводимости называется *дырочным*, а этот тип полупроводников — *примесными дырочными полупроводниками*. Энергетические уровни, на которые забрасываются электроны из заполненной зоны, называются *акцепторными*. Атомы, которым принадлежат эти уровни называются *акцепторами*. Основными акцепторными примесями для полупроводников Ge и Si являются атомы примесей III группы В, Al, In.

Температурная зависимость концентрации дырок дырочного полупроводника имеет тот же вид, что и электронного, рис.4.7. При низких температурах экспоненциально растет концентрация дырок, а затем, когда все акцепторные уровни уже забиты, она делается постоянной вплоть до той температуры, при которой начинает проявляться собственная проводимость.

4.1.6. Эффективная масса

В предыдущих параграфах мы часто использовали термин «свободные электроны», имея в виду при этом электроны, которые могут перемещаться под действием электрического поля, т. е. принимать участие в электрическом токе. Не следует, однако, понимать этот термин буквально. Свободной мы с полным правом можем называть частицу, которая настолько удалена от других тел, что не испытывает практически никакого воздействия и находится, в согласии с первым законом Ньютона, в состоянии покоя или прямолинейного и равномерного движения. Если же частица движется в поле других частиц и испытывает их притяжение, то свободной мы ее можем назвать лишь в том случае, если ее кинетическая энергия больше энергии связи с частицами и поэтому она движется по незамкнутой траектории, т. е. с течением времени удалится на сколько угодно большое расстояние от частиц, в поле которых она находится.

Совершенно в иных условиях находится электрон проводимости в кристалле. Электрон, по сути дела, при своем движении все время связан с каким-либо атомом (причем энергия связи больше его кинетической энергии) и лишь благодаря тому, что решетка построена симметрично и из одинаковых атомов и энергия электрона не зависит от того, с каким именно атомом он связан, электрон и получает возможность перемещаться от одного атома к другому. При этом энергия связи периодически меняется: при приближении к положительно заряженному ядру она очень сильно возрастает, при удалении — убывает; поэтому периодически меняется и скорость движения. Лишь средняя скорость остается постоянной. Это следует из закона сохранения энергии: полная энергия электрона остается постоянной, так как нет внешних сил, которые бы совершали работу. В аналогичных точках кристалла потенциальная энергия электрона одинакова, следовательно, должна быть одинаковой и кинетическая энергия. Поэтому, если не обращать внимания на периодические колебания скорости, а говорить о средней, то можно считать, что электрон в кристалле в отсутствие внешнего поля движется прямолинейно и равномерно, в известном смысле, как свободная частица в отсутствие внешних сил.

Когда к кристаллу прикладывается разность потенциалов, то характер движения электрона должен измениться. Если бы он был «по-настоящему свободным», то вопрос бы решался просто согласно закону сохранения энергии: работа силы F на пути S должна равняться изменению кинетической энергии:

$$F \cdot S = \frac{1}{2} m_0 v^2 - \frac{1}{2} m_0 v_0^2, \quad (4.18)$$

или

$$F = m_0 a, \quad (4.19)$$

где t — время, за которое электрон пройдет путь S ; отсюда, сократив левую и правую части на S и поделив на t , сразу же получаем второй закон Ньютона:

$$F = m_0 a. \quad (4.20)$$

Разумеется, закон сохранения энергии соблюдается и для электронов в кристалле, но дело осложняется тем, что в данном случае при изменении траектории движения электрона под действием внешнего поля меняется одновременно его потенциальная и кинетическая энергия и только одна часть работы сил поля переходит в кинетическую энергию, а другая — в потенциальную.

При этом движение электрона может изменяться под действием поля таким образом, что потенциальная энергия его уменьшится, а, следовательно, увеличение его кинетической энергии будет больше работы сил

поля (в кинетическую энергию перейдет также часть потенциальной); в этом случае электрон будет вести себя как очень легкая частица, т. е. частица с массой, меньшей массы свободного электрона. Может случиться, что почти вся работа пойдет на увеличение потенциальной энергии, скорость электрона увеличится так незначительно, как будто его масса очень велика. Может быть и так, что увеличение потенциальной энергии будет больше работы внешних сил, т. е. в потенциальную энергию перейдет и часть кинетической — скорость электрона уменьшится и он будет вести себя как частица с отрицательной массой.

Ускорение электрона, вообще говоря, может быть не пропорционально силе поля, а связано с ним более сложной функциональной зависимостью.

Все эти возможности действительно реализуются в различных кристаллах для электронов, находящихся на различных уровнях в зоне проводимости и валентной зоне; из этого можно заключить, насколько сложен характер движения электрона в кристалле. Однако существуют и здесь некоторые общие закономерности, которые позволяют упростить анализ механизма электропроводности.

Для электронов, находящихся вблизи нижнего края зоны проводимости, ускорение (на длине свободного пробега) пропорционально приложенной силе

$$F = ma$$

или

$$a = \frac{F}{m}, \quad (4.21)$$

где m — некоторый положительный коэффициент пропорциональности, различный для различных кристаллов.

В силу формальной аналогии между выражением (4.21) и законом Ньютона (4.20) m называют эффективной массой электрона в кристалле.

Из всего сказанного следует, что эффективная масса может быть и меньше и больше массы свободного электрона, что действительно имеет место в различных кристаллах. Из сказанного выше следует также, что эффективная масса может зависеть от направления движения электрона, что также наблюдается в большинстве полупроводниковых кристаллов. В этих случаях зависимость энергии электрона ε от его импульса изображается не сферой $\varepsilon = p^2/2m$, как для свободного электрона, а эллипсоидом

$$\varepsilon = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{p_3^2}{2m_3}. \quad (4.22)$$

Во многих случаях зонная структура носит еще более сложный характер. Так, в зоне проводимости германия имеется четыре одинаковых эллипсоида вращения, в зоне проводимости кремния — три. В ряде кристаллов валентные и свободные электроны размещаются в нескольких зонах, имеющих различные параметры (т. е. абсолютные значения и анизотропию эффективных масс).

В последние годы было проведено весьма большое число работ, и теоретических и экспериментальных, посвященных исследованию энергетического спектра электронов в полупроводниках. Эти исследования показали, что при определенных условиях, даже вблизи нижнего края зоны, масса не остается постоянной, а растет с ростом энергии электрона. Эффективная масса зависит от температуры, от давления, может зависеть от характера и числа примесей и от концентрации электронов.

Более того, скорость электронов обращается в нуль у верхнего края зоны. Электроны, находящиеся вблизи верхнего края зоны, двигаясь в направлении поля, увеличивают свою энергию, т. е. поднимаются на энергетической диаграмме вверх и таким образом приближаются к «потолку» зоны. Но это значит в свете сказанного выше, что скорость их при этом уменьшается; двигаясь в направлении приложенной силы, они не ускоряются, а замедляются, т. е. ведут себя как частицы с отрицательной массой. Пустые места (дырки), напротив, ведут себя здесь как нормальные положительно заряженные частицы с положительной массой. Для дырок вблизи верхнего края зоны точно так же, как для электронов, вблизи нижнего края зоны ускорение пропорционально приложенной силе

$$(4.23)$$

где m — коэффициент, называемый эффективной массой дырок.

Все сказанное выше об эффективной массе электронов относится и к эффективной массе дырок; поэтому мы не будем здесь повторяться.

Для электронов, находящихся в середине зоны, обычно не существует простой пропорциональности между силой и ускорением. Мы не будем на этом случае останавливаться подробнее, так как он наиболее сложен и сравнительно редко встречается в полупроводниках.

Уже упоминалось, что общее число электронных состояний в любой зоне $2Ng$ равно сумме числа состояний на атомарных уровнях, из которых она образовалась. Однако ширина валентных и свободных зон в разных

кристаллах колеблется в очень широких пределах — от долей электрон-вольта до нескольких электрон-вольт. Чем шире зона, тем меньше число электронных состояний dG , приходящихся на определенный интервал энергии $d\varepsilon$. Величину $g(\varepsilon) = dG/d\varepsilon$ называют плотностью электронных состояний. Теория показывает, что вблизи края зоны $g(\varepsilon)$ определяется эффективной массой m и выражается через нее следующим образом:

$$g(\varepsilon) = 4\pi(2m)^{\frac{3}{2}}\varepsilon^{\frac{1}{2}}, \quad (4.24)$$

где m — соответственно эффективная масса электрона или дырки.

4.1.7. Электроны в тепловом равновесии

При своем движении в зоне проводимости полупроводника электроны все время испытывают соударения с атомами, совершающими тепловые колебания, и обмениваются с ними кинетической энергией. Именно поэтому кинетическая энергия этих электронов в среднем, так же как кинетическая энергия атомов, равна $\frac{3}{2}kT$ (при комнатной температуре приблизительно равна 0,04 эВ).

Таким образом, энергия теплового движения электронов так же, как энергия частиц обычного газа, растет пропорционально температуре, скорость — пропорционально квадратному корню из температуры. Совершенно по-иному ведут себя валентные электроны в металле. Вернемся опять к примеру щелочных металлов. Как уже упоминалось, в этом случае валентная зона кристалла оказывается наполовину заполненной. Ширина валентной зоны составляет несколько электрон-вольт, и, таким образом, все энергетические уровни в ней до уровня с энергией по крайней мере в 1 эВ заполнены уже при абсолютном нуле. Поэтому тепловые колебания атомов не могут оказать значительного влияния на распределение электронов по энергиям.

Действительно, электроны, находящиеся не на самых верхних уровнях, а в нижней половине зоны, не могут изменить своей энергии под действием столкновений с атомами, так как при этом они должны перейти на другой уровень, а все соседние уровни заняты. Электроны же, находящиеся на верхних уровнях, имеют энергию порядка 1 эВ, в то время как средняя энергия тепловых колебаний атомов составляет при комнатной температуре всего лишь 0,04 эВ, при температуре 1000° К (700° К) — 0,1 эВ, т. е. во много раз меньше. Поэтому при столкновении с атомом электрон не может увеличить свою энергию; с другой стороны, он не может также и передать свою энергию атому, так как при этом он должен был бы перейти в энергетически более низкое состояние, а все нижние энергетические уровни заняты.

Впрочем, все сказанное относится лишь к столкновениям электронов с атомами со средней энергией. Наряду с этим всегда имеется некоторое число атомов, энергия которых значительно больше средней; под действием столкновений с такими атомами электроны могут переходить в верхнюю половину зоны. Так как число таких атомов очень мало, вероятность этих переходов тоже очень мала, но все же благодаря им при температуре, отличной от абсолютного нуля, всегда имеется некоторое число электронов в верхней половине валентной зоны и некоторое число пустых мест в нижней.

Итак, в силу того, что в щелочных металлах половина валентной зоны заполнена электронами, средняя энергия электронов уже при абсолютном нуле значительно выше, чем энергия тепловых колебаний атомов при любых температурах вплоть до точки плавления металла; поэтому энергия электронов не может значительно изменяться под действием соударений с атомами. Это относится и к щелочно-земельным металлам, так как в них s -зона заполнена целиком и они проводят электрический ток лишь благодаря тому, что верхний край s -зоны перекрывается нижним краем следующей, пустой p -зоны.

Такое состояние свободных электронов, когда их энергия почти не зависит от температуры, называется *вырожденным*. В полупроводниках энергия электронов растет пропорционально температуре; такое состояние электронного газа называется *невырожденным*.

Из сказанного выше ясно, что степень вырождения зависит от числа электронов в зоне проводимости. В полу проводниках это число обычно мало, и все электроны могли бы разместиться на самых низких уровнях в виде тонкого и плотного слоя «электронной жидкости». Но под действием теплового движения электроны «испаряются», т. е. забрасываются на более высокие уровни, расположенные в энергетическом интервале порядка kT , и так как число уровней в этом интервале во много раз больше, чем число электронов, то «электронная жидкость» превращается в разреженный газ. В металлах же зона заполнена до половины плотной «электронной жидкостью», над которой при температуре, отличной от нуля, имеется тонкий слой «электронного пара». Для того чтобы вся эта «жидкость» превратилась в пар (или газ), нужна очень высокая температура (в зависимости от ширины зоны — от 5 000 до 10 000°), и фактически сам металл испарится раньше, чем это должно было бы произойти.

В 1 см³ металла содержится приблизительно 10²² свободных электронов; в полупроводниках концентрация свободных электронов колеблется обычно в зависимости от чистоты, ширины запрещенной зоны и температуры в пределах от 10¹² до 10¹⁷ см⁻³. Наконец, в полуметаллах и в очень «грязных» полупроводниках мы имеем дело с промежуточным случаем: в них число свободных электронов в 1 см⁻³ меняется в зависимости

от содержания Примесей от 10^{18} до 10^{21} . Предположим для определенности, что оно порядка 10^{19} . Такое число электронов при плотной упаковке и эффективной массе $m = m_0$ заняло бы все уровни в зоне приблизительно вплоть до энергий $0,03 \text{ эВ}$. Как мы уже упоминали, при комнатной температуре энергия тепловых колебаний атомов составляет примерно столько же. Поэтому в данном случае при температурах значительно ниже комнатной электронный газ будет находиться в вырожденном состоянии, при температурах значительно выше комнатной — в невырожденном. В интервале температур приблизительно от 100 до 400° К электроны будут находиться в состоянии частичного вырождения, т. е. постепенно переходить из вырожденного состояния в невырожденное. В соответствии с этим энергия электронов при низких температурах будет очень слабо зависеть от температуры, в промежуточной области зависимость постепенно усилится и, наконец, при дальнейшем нагревании средняя энергия электронов начнет расти пропорционально температуре.

Все сказанное выше о полуметаллах с концентрацией электронов 10^{19} см^{-3} будет относиться ко всем полуметаллам (в том числе и дырочным) с той лишь разницей, что в зависимости от концентрации электронов или дырок переходная область будет сдвигаться соответственно в сторону либо более высоких, либо более низких температур.

Примесный полупроводник с большой концентрацией примесей и малой энергией активации также может при низких температурах находиться в вырожденном состоянии.

Качественно это можно объяснить следующим образом (на примере электронного примесного полупроводника). При температуре, близкой к абсолютному нулю, число электронов в зоне проводимости ничтожно и поэтому они находятся в невырожденном состоянии.

При увеличении температуры начинают конкурировать два процесса: заброс электронов в зону проводимости и «расселение» их по зоне в силу растущего теплового движения этих электронов. Если первый процесс будет преобладать, то в некотором интервале температур электроны в зоне могут «забить» все нижние уровни и наступит вырождение. При дальнейшем увеличении температуры вырождение всегда снимается.

Более того, вырожденное состояние может иметь место и в собственном полупроводнике, если плотность состояний в валентной зоне во много раз больше, чем в свободной, и ширина запрещенной зоны невелика.

При этих условиях относительно небольшое число электронов, перешедших из валентной зоны в свободную, может заполнить последнюю до высокого уровня и, таким образом, создать вырождение. Это действительно имеет место в антимониде индия

4.1.8. Функция распределения Ферми.

В собственном полупроводнике при абсолютном нуле вся валентная зона заполнена, а зона проводимости пуста; при температуре, отличной от нуля, часть электронов «испаряется» из валентной зоны в зону проводимости. При этом отличие от ситуации в металле заключается в том, что валентная зона и зона проводимости разделены запрещенным промежутком, а верхняя и нижняя половины зоны металла непосредственно примыкают друг к другу.

В примесном электронном полупроводнике в зону проводимости переходят электроны с донорных уровней, а на них появляются неподвижные дырки. В дырочном полупроводнике электроны забрасываются на акцепторные уровни из валентной зоны и «застревают» там, а в зоне появляются подвижные дырки.

Для анализа подавляющего большинства явлений недостаточно только этих качественных представлений и даже знания того, сколько электронов при данной температуре находится в зоне проводимости полупроводника (или в верхней половине валентной зоны металла), а необходимо знать, как эти электроны распределены по энергиям. Для этого вводится так называемая функция распределения $f_0(E)$ (названная функцией распределения Ферми), характеризующая вероятность того, что состояние с данной энергией E занято:

$$(4.25)$$

Как показывает теория, $f_0(E)$ является универсальной функцией; плотность состояний также может быть более или менее точно вычислена для полупроводников и металлов; таким образом, распределение электронов по энергиям $n(E)$ может быть получено из (4.24) и (4.25):

$$(4.26)$$

Вычисление плотности состояний $g(E)$ и вывод функции распределения $f_0(E)$ являются отдельными задачами. Здесь же приведем выражение для $n(E)$ без вывода:

$$(4.27)$$

где μ — величина, зависящая от энергетического спектра электронов в данном материале, числа электронов в нем и температуры.

Вычисление μ также является отдельной задачей. Здесь же ограничимся качественным рассмотрением вопроса.

Как видно из (4.27), μ имеет размерность энергии, поэтому на энергетической диаграмме μ характеризует определенный уровень, который называют уровнем химического потенциала или уровнем Ферми.

Исследуем графически, как зависит функция распределения от энергии и температуры. Для этого будем откладывать, как это мы обычно делаем, по оси ординат энергию E , а по оси абсцисс — $f_0(E)$ (рис. 4.9). Нуль отсчета энергии, как всегда, произволен; уровень μ на рис. 4.9 обозначен горизонтальной штрих-пунктирной линией.

Рассмотрим сначала вид $f_0(E)$ при абсолютном нуле. При $T \rightarrow 0$ и при $E > \mu$ значение $(E - \mu)/kT \rightarrow \infty$ и величина $f_0(E) \rightarrow 0$; при $E < \mu$ $(E - \mu)/kT \rightarrow -\infty$ и $f_0(E) \rightarrow 1$; при $E = E = \mu$ $f_0(E)$ скачком переходит от значения $f_0(E) = 1$ к значению $f_0(E) = 0$ и, следовательно, изображается ломаной кривой, представленной на рис. 4.9, а. Таким образом, все состояния ниже уровня Ферми заполнены, а выше него пусты. В соответствии с изложенным это означает, что при абсолютном нуле уровень Ферми проходит в металле как раз посередине валентной зоны.

При температуре, отличной от абсолютного нуля, согласно (4.27) будет по-прежнему $f_0(E) = 0$ при $(E - \mu)/kT \rightarrow \infty$ и $f_0(E) = 1$ при $(E - \mu)/kT \rightarrow -\infty$, но для значений энергий, близких к μ , положение существенно изменяется (рис. 4.9, б). Теперь при $E = \mu$ значение $f_0(E) = 1/2$, при $E < \mu$ будет выполняться $1/2 < f_0 < 1$ и при $E > \mu$

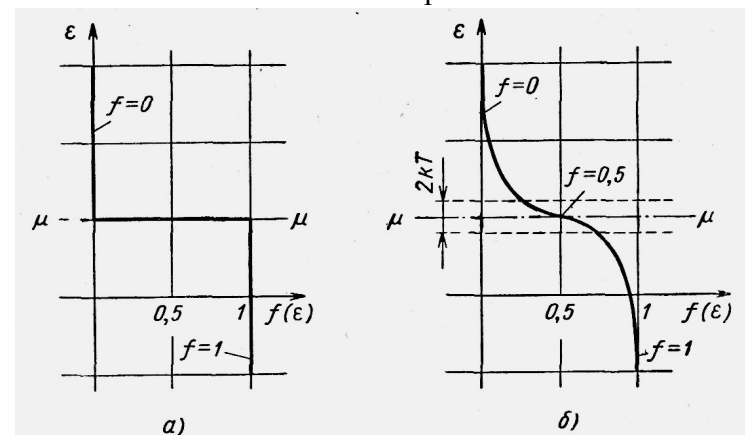


Рис. 4.9. Функция распределения Ферми при двух различных температурах: а) $T = 0^\circ \text{K}$; б) $T > 0^\circ \text{K}$.

имеет место неравенство $1/2 < f_0 < 1$ функция распределения теперь уже плавно переходит от значения $f = 1$ к значению $f = 0$, причем интервал энергий, в котором происходит этот постепенный переход, порядка $2kT$ (рис. 1.9, б). Чем выше температура, тем больше «расплывается» функция распределения Ферми. Качественно это означает, что часть электронов забрасывается из нижней половины зоны в верхнюю. Так как все это происходит в интервале энергий порядка $4kT$, что при комнатной температуре составляет всегда лишь $0,1 \text{ эВ}$, то в среднем энергия всех свободных электронов в металле почти не изменится. Такое состояние электронов, в котором их энергия почти не зависит от температуры, называется вырожденным. Рассмотрим теперь, как обстоит дело в полупроводниках. Общее число состояний в свободной зоне полупроводника равно $2Ng$, где N — число атомов и g — кратность уровня, из которого образовалась зона. Число атомов в 1 см^3 твердого тела $N \approx 10^{22}$, следовательно, такого же порядка будет и число состояний. Число же свободных электронов в полупроводниках колеблется обычно в пределах $10^{12} - 10^{18} \text{ см}^{-3}$. Это значит, что доля занятых состояний ничтожно мала или, иными словами, $f \ll 1$ для всех состояний, включая самые нижние, примыкающие ко дну зоны.

Но если $f \ll 1$, то это значит, что $(E - \mu)/kT \gg 0$ или $\mu \ll E$, т. е. уровень химического потенциала должен проходить ниже дна зоны. В этом случае для всех состояний в зоне проводимости мы можем пренебречь единицей в выражении для (4.27) и, следовательно,

$$f_0(E) = e^{\frac{\mu-E}{kT}}. \quad (4.28)$$

Приближенное выражение (4.28) называется распределением Максвелла — Больцмана. Состояние электронного газа, в котором $f_0(E) \ll 1$ и которое описывается поэтому приближенной формулой (4.28), называется невырожденным. Мы увидим ниже, что в этом случае энергия в отличие от вырожденного состояния растет пропорционально температуре. Анализ положения и температурной зависимости уровня Ферми позволяет сделать следующие выводы.

1. В области собственной проводимости уровень химического потенциала обычно (если эффективные массы электронов и дырок примерно одинаковы) проходит посередине зоны.

2. В примесном (некомпенсированном) электронном полупроводнике при температуре абсолютного нуля уровень химического потенциала проходит посередине между дном зоны проводимости и примесными уровнями, а затем начинает подниматься. При этом конкурируют два процесса: заброс электронов в зону проводимости и расселение (или «испарение» со дна зоны) заброшенных электронов.

До тех пор пока первый процесс преобладает, плотность электронного газа в зоне проводимости растет и уровень Ферми поднимается. Когда начинает преобладать второй процесс, уровень Ферми начинает опускаться; когда половина примесных электронов перешла в зону, уровень Ферми совпадает с примесными уровнями, затем пересекает их и приближается к середине зоны — начинается область собственной проводимости.

3. Аналогично все происходит в дырочном (нескомпенсированном) полупроводнике. При абсолютном нуле уровень Ферми проходит посередине между верхним краем валентной зоны и акцепторными уровнями, затем сначала опускается и приближается ко дну валентной зоны, потом начинает подниматься, пересекает акцепторные уровни и приближается к середине запрещенной зоны.

В некоторых случаях мы еще более упростим вопрос и будем считать, что уровень химического потенциала во всей области примесной проводимости проходит посередине между примесными уровнями и краем соответствующей зоны.

4.1.9. Электропроводность полупроводников. Формулы для подвижности носителей и электропроводности.

В отсутствие поля скорости свободных электронов имеют всевозможные направления. Под действием электрического поля E электроны получают ускорение $a = -eE/m_n$ и направленную по полю добавку к скорости $v = at$, где t — время, в течение которого действует ускорение a . Если бы t не было ничем ограничено, то и скорость электрона в направлении поля неограниченно возрастала бы. В действительности, однако, это не так. Теория показывает, что в строго периодическом поле, т. е. в идеальной кристаллической решетке с покоящимися атомами, электроны действительно двигались бы ускоренно на протяжении очень больших промежутков времени. Но идеальных кристаллов не существует. В реальных же кристаллах всегда существуют нарушения периодического потенциала, обусловленные, во-первых, тепловыми колебаниями атомов и, во-вторых, всевозможными дефектами решетки:

- так называемыми точечными дефектами: пустыми узлами и атомами в междоузлиях, нейтральными и заряженными примесными атомами (которые также могут размещаться в узлах и междоузлиях);
- линейными дефектами — краевыми и винтовыми дислокациями;
- плоскими дефектами — границами зерен, блоков, двойников;
- объемными дефектами — трещинами, полостями, включениями (в том числе зародышами) другой фазы.

В силу этого электрон лишь на сравнительно небольшом отрезке пути l , называемом длиной свободного пробега, движется ускоренно, затем испытывает соударение, теряет при этом свою направленную скорость, и весь процесс начинается сначала. При высоких температурах длину свободного пробега ограничивают тепловые колебания атомов, при низких — примеси и другие дефекты кристаллов.

Время t , входящее в выражение для v , которое называется временем свободного пробега, мы можем, зная длину свободного пробега l , вычислить по формуле

$$t = l/v \quad (4.29)$$

Опыт показывает, что длина свободного пробега электрона обычно очень мала, $l \approx 10^{-8}$ см, и в силу этого направленная добавка к скорости электрона $v = at$ при не очень сильных полях мала по сравнению со скоростью хаотического теплового движения электронов $v_{th} = \sqrt{2kT/m_n}$ в отсутствие поля, поэтому в (4.29) вместо v может быть

поставлено v_0 :

$$\tau_n \approx \frac{l_n}{v_0}. \quad (4.30)$$

Средняя скорость, с которой электрон будет двигаться вдоль поля,

$$\bar{v} = \frac{\overline{\Delta v}}{2} = \frac{e}{2m_n} \frac{l_n}{v_0} E = u_n E, \quad (4.31)$$

где отношение скорости дрейфа электронов к электрическому полю обозначено буквой u :

$$u_n = \frac{e}{2m_n} \frac{l_n}{v_0}. \quad (4.32)$$

Это отношение называется подвижностью электрона. Иными словами, подвижность — это скорость дрейфа электронов в поле напряженностью 1 В/см . Более строгий расчет дает в два раза большее значение подвижности:

$$u_n = \frac{e}{m_n} \frac{l_n}{v_0} = \frac{e}{m_n} \tau_n. \quad (4.33)$$

Если в 1 см^3 содержится n свободных электронов и все они движутся в направлении поля со скоростью \bar{v} , то через площадку в 1 см^2 за секунду пройдут все электроны, находящиеся в объеме параллелепипеда длиной \bar{v} ; число таких электронов будет $n\bar{v}$; заряд, перенесенный этими электронами, называется плотностью тока

$$j = en\bar{v} = \frac{ne^2 \bar{l}_n}{m_n v_0} E. \quad (4.34)$$

Следовательно, электропроводность материала

$$y_n = \frac{j}{E} = \frac{ne^2 \bar{l}_n}{m_n v_0} = enu_n. \quad (4.35)$$

Совершенно аналогично вычисляется подвижность и электропроводность дырочного полупроводника; поэтому мы приведем здесь соответствующие выражения, не повторяя проделанных элементарных вычислений:

$$u_p = \frac{e}{m_p} \frac{\bar{l}_p}{v_0} \quad (4.36)$$

и

$$(4.37)$$

В области собственной и смешанной (т. е. переходной от примесной к собственной) проводимости электронная и дырочная проводимости складываются:

$$(4.38)$$

при наличии нескольких сортов носителей (дырок и электронов) с различными массами и подвижностями выражение (4.38) легко обобщается:

$$(4.39)$$

где суммирование проводится по всем сортам носителей (и дыркам и электронам).

4.1.10. Зависимость длины свободного пробега электрона от температуры. Температурная зависимость подвижности.

Выше уже упоминалось, что длина свободного пробега, а следовательно, и подвижность электрона в реальных кристаллах ограничивается нарушениями периодического потенциала, обусловленными тепловыми колебаниями и дефектами решетки. Дефекты играют основную роль при низких температурах, а тепловые колебания — при высоких. Рассмотрим теперь каждый механизм рассеяния электронов подробнее.

Рассеяние на дефектах решетки. Выше мы перечислили ряд причин, нарушающих периодичность потенциала в реальных кристаллах.

Для того чтобы электрон изменил направление своего движения в результате взаимодействия с нейтральным дефектом, необходим акт столкновения в буквальном смысле этого слова — траектория электрона должна

проходить через место расположения дефекта либо через непосредственно примыкающую к нему область решетки, в которой им вызваны искажения.

Электрическое поле ионизированного примесного атома распространяется значительно дальше, и электрон, проходя на значительном расстоянии от него, изменит под действием этого поля направление своего движения. Поэтому из всех перечисленных выше дефектов основную роль в рассеянии электронов играют ионизированные примеси и только их мы сейчас будем рассматривать.

Чем больше тепловая скорость движения электрона, тем меньше времени он находится в непосредственной близости от примесного иона и тем меньшее отклонение он испытывает; поэтому длина свободного пробега электрона очень резко растет с увеличением его скорости. Теория показывает, что в этом случае длина свободного пробега пропорциональна четвертой степени скорости или, что то же самое, квадрату кинетической энергии:

$$\bar{l} \propto v_0^4 e^{-2} \quad (4.40)$$

Из (4.40) следует два важных вывода:

1. При рассеянии на ионах примеси доля участия электронов с различными скоростями в электрическом токе различна — основную роль играют быстрые электроны, так как их длина свободного пробега больше. Отсюда также следует, что средняя энергия в потоке (электрическом токе) электронов будет отлична (в большую сторону) от средней тепловой энергии.

2. Согласно (4.36) подвижность электрона пропорциональна отношению длины свободного пробега к средней скорости: $\mu \propto \bar{l}/v_0$, и согласно формуле (4.40) $\mu \propto v_0^3$. Но для невырожденного полупроводника средняя кинетическая энергия пропорциональна температуре, а средняя тепловая скорость пропорциональна квадратному корню из температуры $v_0 \sim T^{1/2}$; следовательно, в рассматриваемом нами случае рассеяния электронов на ионах примеси

$$\mu \propto T^{3/2} \quad (4.41)$$

Таким образом, здесь осуществляется очень редкий случай, когда подвижность электронов растет с температурой.

Рассеяние электронов на тепловых колебаниях решетки.

С повышением температуры увеличиваются тепловые скорости электронов и уменьшается вероятность их рассеяния на примесных ионах. Одновременно при этом растут амплитуды тепловых колебаний атомов и растет роль этого механизма рассеяния. Поэтому при высоких температурах рассеяние на тепловых колебаниях играет решающую роль, а рассеянием на ионах примеси можно пренебречь. В этом случае зависимость длины свободного пробега электрона от температуры обусловлена двумя причинами: во-первых, как уже упоминали, с ростом температуры увеличиваются амплитуды тепловых колебаний атомов, что непосредственно ведет к уменьшению длины свободного пробега; во-вторых, длина свободного пробега

электрона зависит от его кинетической энергии ϵ (или, что то же самое, от v_0), которая, в свою очередь, зависит от температуры.

Теория показывает, что в этом случае выражение для длины свободного пробега имеет вид

$$\bar{l} \propto T^{-1} \quad (4.42)$$

причем показатель степени n и температурная зависимость $\mu(T)$ различны в кристаллах различного типа и в различных интервалах температур.

В ионных полупроводниках при рассеянии электронов на оптических колебаниях в области высоких температур (выше температуры Дебая) $\bar{l} \propto T^{-1}$ и

$$\mu \propto T^{-1/2} \quad (4.43)$$

где a — некоторая константа; таким образом,

$$\mu \propto T^{-1/2} \quad (4.44)$$

и, следовательно, подвижность

$$\mu \propto T^{-1/2} \quad (4.45)$$

В невырожденных полупроводниках $\bar{l} \propto T^{-1}$ и, следовательно,

$$\mu \propto T^{-1/2} \quad (4.46)$$

В тех же кристаллах, но при температурах ниже температуры Дебая $r = \frac{1}{2}$, $l(T) \propto e^{\theta/T}$, где θ — температура Дебая; следовательно,

$$\bar{l} \propto \frac{1}{e^{\frac{1}{2} \frac{\theta}{T}}}, \quad (4.47)$$

время релаксации $\bar{\tau} = \bar{l}/v_0$ не зависит от энергии и

$$\text{и } \propto e^{\theta/T}. \quad (4.48)$$

В ковалентных полупроводниках при рассеянии на акустических колебаниях $r = 0$

$$\bar{l} \propto T^{-1} \quad (4.49)$$

и не зависит от энергии; таким образом,

$$\text{и } \propto \frac{1}{e^{\frac{1}{2} T^{-1}}} \propto T^{\frac{3}{2}}. \quad (4.50)$$

Теория показывает также, что подвижность и прямо [см. формулу (4.33)] и косвенно (через v_0 и l_0) зависит от эффективной массы по двум причинам: так как тепловая средняя скорость зависит от эффективной массы

$\frac{1}{2} m v_0^2 = \frac{3}{2} kT$ и так как вероятность столкновений, а следовательно, и длина свободного пробега l зависят от плотности состояний $g(\epsilon)$, а плотность состояний зависит от эффективной массы.

Экспериментальные и теоретические исследования, проведенные за последние годы, показали, что в целом ряде полупроводников (в первую очередь в полупроводниках с узкой запрещенной зоной) эффективная масса зависит от температуры.

Причины этого заключаются в следующем. Эффективная масса в полупроводниках с узкой запрещенной зоной растет по мере роста расстояния от дна зоны. Так как с ростом температуры средняя энергия электронов растет и они удаляются от дна зоны, то соответственно растет и средняя эффективная масса, входящая в выражение (4.33). Эффективная масса зависит также от межатомных расстояний: с ростом межатомных расстояний она, как правило, увеличивается; таким образом, она возрастает из-за теплового расширения кристалла. В некоторых случаях на эффективную массу оказывают также прямое влияние тепловые колебания.

В силу температурной зависимости эффективной массы температурная зависимость подвижности в области высоких температур (т. е. рассеяния на тепловых колебаниях решетки) оказывается часто более крутой, чем

это предсказывают формулы, приведенные выше. В ряде полупроводников ; и даже Для дальнейшего важно, однако, что длина свободного пробега в большинстве случаев является степенной функцией от энергии электрона:

$$(4.51)$$

где показатель степени зависит от механизма рассеяния электронов; в атомной решетке при расстоянии на акустических колебаниях $\alpha = 0$; $\alpha = 1$ при $T >$ и $\alpha = 1/2$ при в ионной решетке; при рассеянии электронов на ионах примеси $\alpha = 2$.

За последние годы обнаружен довольно широкий класс полупроводников (окислы переходных металлов и др.), в которых при низких температурах перемещение электронов носит характер активационных скачков и подвижность экспоненциально растет с повышением температуры. Этот механизм электропроводности теоретически еще мало изучен, несмотря на то, что на этих материалах основан важный класс полупроводниковых приборов — терморезисторы.

4.2. РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ТОКА В ПРОВОДНИКАХ

4.1. Основные определения

До сих пор рассматривались концентрации электронов и дырок, устанавливающиеся при термодинамическом равновесии полупроводника. Эти концентрации обозначаются индексом нуль: n_0 и p_0 . Но в полупроводнике можно создать концентрацию носителей, превышающую равновесную. Например, воздействуя на кристалл

светом частоты , с энергией кванта $h\nu$, большей чем , световой квант может передать свою энергию электрону и перевести его с примесного уровня в зону проводимости. Аналогичным образом свет может перевести электрон на акцепторный уровень из валентной зоны и образовать в ней дырку. Если энергия кванта превысит ширину запрещенной зоны E_g , то свет будет возбуждать одновременно и электроны и дырки, так как

при этом часть электронов валентной зоны, получив достаточную энергию, преодолет потенциальный барьер E_g и перейдет в зону проводимости.

Очевидно, что число неравновесных носителей заряда, возбуждаемых светом (или каким-либо другим способом), не может расти до бесконечности. Этому будут препятствовать обратные переходы электронов в исходные состояния. Совершенно ясно, что чем больше создано свободных электронов и дырок, тем чаще оказываются встречи их друг с другом, т. е. тем вероятнее оказывается заполнение электроном свободного уровня и, следовательно, количество обратных переходов становится больше.

Процесс образования избыточных носителей называют процессом генерации электронно-дырочных пар. Если генерация пар происходит в полупроводнике n -типа, то относительное приращение концентрации электронов в зоне проводимости $\Delta n/n_0$ будет невелико, а приращение концентрации дырок в валентной зоне $\Delta p/p_0$, наоборот, будет значительным. При генерации пар в кристалле p -типа будет иметь место обратная картина. Обобщая, можно говорить о значительном относительном увеличении неосновных носителей при генерации пар в полупроводнике с ярко выраженным тем или иным типом электропроводности. В этом случае говорят об *инжекции неосновных носителей*.

Обратный процесс сваливания электронов из зоны проводимости в валентную носит название процесса *рекомбинации носителей*. Почти все электронные полупроводниковые приборы основаны на явлении инжекции неравновесных носителей при воздействии на кристалл внешними силами (световое излучение, электромагнитное излучение радиочастотного диапазона, облучение ядерными частицами и т.д.). Поэтому от скорости рекомбинации носителей, например, будет зависеть готов или не готов кристалл (точнее, прибор) к повторному приему внешнего воздействия, т. е. пришли ли носители заряда в разрешенных зонах к равновесному состоянию в промежутке между импульсами внешнего воздействия — сигнала. совершенно очевидно, что большей скорости рекомбинации будет соответствовать способность к приему сигналов, следующих с большей частотой.

Процессы генерации и рекомбинации можно уподобить (как делалось ранее с равновесными носителями) некой *химической* реакции:

$$h\nu \frac{K_1}{K_2} n + p. \tag{4.52}$$

Тогда эта реакция будет характеризоваться двумя скоростями скоростью генерации K_1 и скоростью рекомбинации K_2 .

Скорость генерации определим как число носителей, генерируемых в 1 см^3 за 1 с , а скорость рекомбинации, наоборот, как число рекомбинируемых носителей в 1 см^3 за 1 с .

Таким образом, в 1 см^3 кристалла в 1 с генерируется g_0 равновесных электронно-дырочных пар обычным тепловым способом и при воздействии, например, светом — g пар неравновесных носителей.

Таким образом, скорость генерации K_1 будет

$$\tag{4.53}$$

За это же время будет рекомбинировать K_2 электронно-дырочных пар, причем в случае прямой рекомбинации, т. е. такой, — которая происходит непосредственно между электронами зоны проводимости и дырками валентной зоны

$$\tag{4.54}$$

где — коэффициент пропорциональности, зависящий от свойств полупроводника, но не зависящий от n и p .

В состоянии теплового равновесия, когда свет выключен ($g = 0$), $n_0 p_0$, а так как $K_1 = K_2$, то

$$\tag{4.55}$$

Когда свет включен ($g \neq 0$), то в течение некоторого времени n и p будут убывать до нуля т. е. до тех пор, пока число рекомбинируемых пар не уравнивает число генерируемых

пар. Таким образом, n и p есть не что иное, как изменение концентрации носителей во времени t или τ . Знак минус показывает, что число носителей во времени уменьшается. Таким образом, используя (4.53)—(4.55), запишем

$$\frac{dn}{dt} = \frac{dp}{dt} = K_2 - K_1 = g - \gamma(np - n_0p_0). \quad (4.56)$$

Если избыточная концентрация мала по сравнению с равновесной

$$\Delta n \ll n_0 + p_0 \quad (4.57)$$

то (4.56) можно упростить. Для этого запишем равенство

$$np - n_0p_0 = \Delta n [(n_0 + p_0) + \Delta n]. \quad (4.58)$$

Из-за (4.57) величиной Δn в квадратной скобке (4.58) можно пренебречь:

$$np - n_0p_0 \approx \Delta n (n_0 + p_0)$$

Уравнение (4.56) теперь упрощается;

$$\frac{dn}{dt} = g - \gamma(n - n_0)(n_0 + p_0). \quad (4.59)$$

После включения света генерация пар прекращается и $g = 0$.

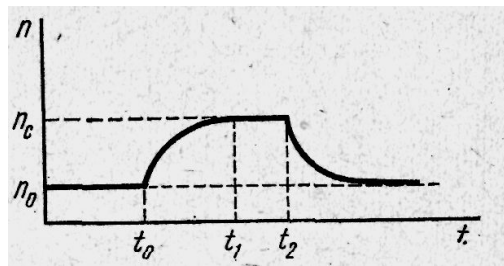


Рис. 4.10. Нарастание и спад концентрации неравновесных носителей во времени

В этих условиях рекомбинацию будет описывать уравнение

$$\frac{dn}{dt} = -\gamma(n - n_0)(n_0 + p_0) = -\frac{n - n_0}{\tau}, \quad (4.60)$$

где величина $\gamma(n_0 + p_0)$, зависящая только от свойств полупроводника, и не зависящая от Δn обозначена через

τ . Из (4.60) видно, что τ имеет размерность времени.

Рассмотрим физический смысл этой величины. Для этого обратимся к рис. 4.10, на котором изображено

изменение концентрации носителей в кристалле до и после освещения его светом. До времени t_0 свет не

включен и в кристалле установилась равновесная концентрация n_0 . Как только включается свет ($t = t_0$),

начинают генерировать избыточные носители, причем скорость генерации превышает скорость

рекомбинации. Так будет происходить до тех пор, пока обе эти скорости не уравновесят друг друга ($dn/dt = 0$).

В кристалле установится стационарная концентрация электронов n_c . Затем, если мы выключим свет в момент

времени t_2 , то в результате рекомбинации концентрация носителей начнет уменьшаться до тех пор, пока

она снова не достигнет первоначального равновесного значения n_0 . Спад концентрации $n(t)$ на участке

описывается интегрированием (4.59):

$$(4.61)$$

где C — постоянная интегрирования, равная n_c .

Отсюда видно, что за время τ число избыточных носителей убывает в e раз. Таким образом, величина

характеризует быстроту спада кривой рис. 4.10. Или, что то же самое, величина τ определяет время существования избыточных носителей, или точнее *среднее время жизни электронно-дырочных пар*.

Совершенно очевидно, что в общем случае (например, когда рекомбинируют электроны на примесном уровне и дырки в валентной зоне) вовсе не обязательно равенство времен жизни электронов и дырок, так же как

необязательны и принятые выше соотношения $\tau_n = \tau_p$ и $\tau_n = \tau_p$

Поэтому вводят в рассмотрение разные времена τ_n и τ_p которые по определению равны (см. (4.59):

$$\tau_n = \frac{\Delta n}{R_n}; \quad \tau_p = \frac{\Delta p}{R_p}, \quad (4.62)$$

где R_n и R_p — полные скорости рекомбинации электронов и дырок в кристалле.

Эти скорости можно представить в виде

$$\left. \begin{aligned} R_n &= r_n(n, p) - g_{n0}(n_0, p_0), \\ R_p &= r_p(n, p) - g_{p0}(n_0, p_0). \end{aligned} \right\} \quad (4.63)$$

Здесь r_n и r_p — скорости захвата электронов и дырок в единице объема кристалла; g_n и g_p — скорости генерации электронов и дырок.

Индекс нуль относится к равновесному случаю, так как в равновесии скорости рекомбинации и генерации равны нулю. Таким образом (4.63), есть уравнение баланса скоростей. Если преобладает первое слагаемое, то результирующая скорость рекомбинации отлична от нуля.

Если определение времени жизни осуществляется в нестационарных условиях после выключения источника генерации (света), то скорости рекомбинации R_n и R_p будут совпадать с dn/dt и dp/dt (в частном случае прямой рекомбинации $dn/dt = dp/dt$).

Тогда в (4.62) величины R_n и R_p можно заменить на $\left. \frac{dn}{dt} \right|_{t \rightarrow t_2}$ и $\left. \frac{dp}{dt} \right|_{t \rightarrow t_2}$, где t_2 — начальный момент выключения света (см. рис. 4.10). Таким образом, времена жизни носителей будут определяться соотношением

$$\left. \frac{dn}{dt} \right|_{t \rightarrow t_2} = \frac{\Delta n}{\tau_n}; \quad \left. \frac{dp}{dt} \right|_{t \rightarrow t_2} = \frac{\Delta p}{\tau_p}. \quad (4.64)$$

Следует различать *стационарные* и *нестационарные времена жизни носителей*. Первые характеризуют рекомбинацию в условиях непрерывного внешнего воздействия, т. е. при постоянной скорости генерации носителей: $K_1 = g \text{ см}^{-3} \cdot \text{с}^{-1}$.

В этом случае концентрации неравновесных носителей будут возрастать до тех пор, пока скорости рекомбинации K_n и K_p не станут равны скорости генерации g . Таким образом, стационарные времена будут определяться как $\tau_n = \Delta n / g$ и $\tau_p = \Delta p / g$. Абсолютные значения τ_n и τ_p в этом случае могут значительно

отличаться друг от друга, если $K_n \neq K_p$. Нестационарные времена характеризуют рекомбинацию в отсутствие генерации, т. е. после выключения внешнего воздействия на кристалл.

Существуют случаи, при которых стационарные и нестационарные времена жизни оказываются неравными.

Это происходит, когда имеет место *эффект прилипания*, при котором

В общем случае, значения τ_n и τ_p зависят от вида и механизма рекомбинации, состава полупроводника, температуры, а также от условий, в которых производится их определение.

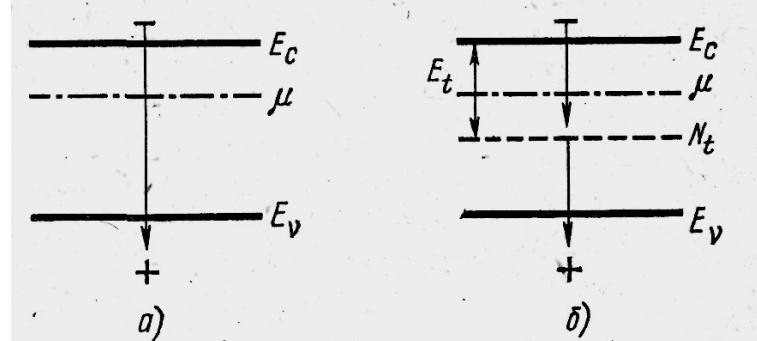


Рис. 4.11. Различные виды рекомбинации:

а — рекомбинация зона — зона; б — рекомбинация через глубокие локальные центры

Имеется два вида рекомбинации. Первый вид — зона — зона, при котором избыточные электроны из зоны проводимости непосредственно «сваливаются» в валентную зону, и второй вид — такой, при котором рекомбинация идет через глубокие уровни в запрещенной зоне полупроводника (рис. 4.11).

При генерации носители получают энергию от внешнего воздействия, в рассматриваемом случае — освещения. Естественно, что при обратном акте рекомбинации эта энергия должна выделиться. При этом

вовсе не обязательно, чтобы поглощенная в виде светового кванта энергия выделялась бы также в виде света. Именно поэтому классификацию механизмов рекомбинации принято осуществлять по способу отвода энергии, выделяющейся при акте захвата носителей при рекомбинации.

Наиболее вероятными механизмами -рекомбинации являются:

1. *Излучательная рекомбинация*, при которой энергия выделяется в виде квантов электромагнитного излучения;
2. *Фононная рекомбинация*, связанная с непосредственной передачей выделяющейся энергии колебаниям атомов кристаллической решетки;
3. *Ударная рекомбинация (эффект Оже)*. В этом случае энергия вначале передается ближайшему свободному электрону (или дырке), который затем отдает свою избыточную энергию, либо колебаниям атомов решетки, либо другим носителям.

Все три механизма могут осуществляться как при рекомбинации зона — зона, так и при рекомбинации через локальные центры. Кроме того, на обоих этапах рекомбинации через локальные центры (при захвате электрона и при захвате дырки) могут иметь место различные механизмы рекомбинации.

Таким образом, могут представиться девять случаев рекомбинации.

Анализ всех девяти случаев и вычисление соответствующих им вероятностей элементарных актов захвата составляет квантово-механическую часть задачи о рекомбинации носителей заряда в полупроводниках. В каждом из этих случаев вероятности захвата оказываются различными функциями концентрации равновесных носителей n_0, p_0 , температуры, ширины запрещенной зоны, энергетического положения глубоких уровней и других характеристик полупроводника. По этим различиям и определяется экспериментально преобладание того или иного случая из девяти перечисленных.

4.2.2. Межзонная излучательная рекомбинация

Рассмотрим рекомбинацию электронов зоны проводимости и дырок валентной зоны, когда каждый элементарный рекомбинационный акт сопровождается излучением фотона.

Представим себе сначала полупроводник в термодинамическом равновесии. Это означает, что в разрешенных зонах содержится n_0 и p_0 свободных электронов и дырок. Кроме того, в 1 см^3 в 1 с количество испускаемых фотонов при рекомбинации будет равно количеству поглощенных фотонов. Эта скорость рекомбинации r_0 должна быть равна скорости генерации g_0 и пропорциональна произведению $n_0 p_0$:

$$r_0 = g_0 = \gamma_i n_0 p_0 = \gamma_i n_i^2, \quad (4.65)$$

где γ_i — коэффициент (вероятность) захвата носителей заряда с излучением фотона, называемый коэффициентом излучательной рекомбинации, n_i — концентрация носителей тока в собственном полупроводнике.

Из (4.65) имеем

$$(4.66)$$

Если полупроводник не находится в термодинамическом равновесии, то вместо n_0 и p_0 в зонах будет n электронов и p дырок. Но в зоне проводимости и в валентной зоне равновесные и неравновесные носители неразличимы, поэтому процесс рекомбинации тех и других должен быть одинаков. Следовательно, и в отсутствие термодинамического равновесия излучательная рекомбинация должна характеризоваться коэффициентом γ_i :

$$(4.67)$$

По определению (4.62)), с учетом (4.66) и (4.67), времена жизни неравновесных электронов и дырок будут

$$(4.68)$$

Подставляя сюда

$$(4.69)$$

получим в общем виде

$$\tau_n = \tau_p = \frac{1}{r_0} \cdot \frac{n_i^2}{n_0 \frac{\Delta p}{\Delta n} + p_0 + \Delta p}. \quad (4.70)$$

Обычно имеет место равенство $\Delta n = \Delta p$. Тогда (4.70) перейдет в

$$\tau_n = \tau_p = \frac{1}{r_0} \cdot \frac{n_i^2}{n_0 + p_0 + \Delta n}. \quad (4.71)$$

Теперь получим более простые выражения для ряда частных случаев.

Могут быть два случая: малого и большого уровня инжекции.

При малом уровне инжекции $\Delta n \approx n_0 + p_0$ с учетом (4.67) выражение (4.71) перейдет в

$$\tau_n = \tau_p = \frac{1}{r_0} \cdot \frac{n_i^2}{n_0 + p_0} = \frac{1}{\gamma_u (n_0 + p_0)}. \quad (4.72)$$

При большом уровне инжекции, наоборот, $\Delta n \approx n_0 + p_0$ и (4.71) переходит в

$$\tau_n = \tau_p = \frac{1}{\gamma_u \Delta n_0}. \quad (4.73)$$

Полученные выражения справедливы как для собственного, так и для примесного полупроводников. Обычно полупроводниковые материалы содержат существенно различные количества донорных и акцепторных уровней, так что $n \neq p$.

При малом уровне инжекции ($\Delta n \approx n_0 + p_0$) для полупроводника n-типа ($n_0 \gg p_0$) из (4.71) получим

$$\tau_n = \frac{1}{\gamma_u n_0}, \quad (4.74)$$

а для полупроводника p-типа ($p_0 \gg n_0$)

$$\tau_p = \frac{1}{\gamma_u p_0}. \quad (4.75)$$

При большом уровне инжекции для кристаллов n- и p-типа найдем выражение для $\tau_n = \tau_p$ опять-таки в виде (4.73).

Из выражений (4.71)—(4.75) вытекают важные следствия.

Во-первых, видно, что *только при малых уровнях инжекции излучательное время жизни зависит от степени легирования полупроводника* (через n_0 и p_0). Во-вторых, *с увеличением степени легирования величина излучательного времени жизни падает*.

4.2.3. Межзонная безызлучательная рекомбинация

Безызлучательный механизм межзонной рекомбинации предполагает передачу энергии, освобождающейся в процессе рекомбинации, тепловым колебаниям решетки.

Для элементарного акта такого процесса скорость генерации не должна зависеть от концентрации избыточных носителей заряда, а скорость рекомбинации пропорциональна (в невырожденном полупроводнике) произведению np . Таким образом, здесь имеем условия, аналогичные тем, которые рассматривались и при излучательной рекомбинации. Следовательно, и безызлучательная рекомбинация описывается уравнениями (4.65) и (4.66). Тогда можно считать, что безызлучательная межзонная рекомбинация даст некоторую поправку

(слагаемое) к эффективной величине τ в уравнениях предыдущего параграфа.

Из общих соображений можно считать эту поправку малой. Действительно, для того чтобы вероятность безызлучательной рекомбинации была не малой, необходимо предполагать, что, по крайней мере, один из носителей заряда сильно взаимодействует с решеткой. Но это противоречит тому, что до рекомбинации и электрон, и дырка являлись свободными. Кроме того, нет и экспериментальных данных, указывающих на безызлучательные переходы зона — зона.

В настоящее время считается установленным, что подавляющая роль безызлучательных рекомбинационных переходов связана с локальными уровнями.

4.2.4. Межзонная ударная рекомбинация

Процесс ударной рекомбинации предполагает участие трех носителей заряда: рекомбинирующую пару электрон — дырка и еще один электрон или дырку, которым передается энергия, выделяющаяся при рекомбинации. Таким образом, этот процесс будет характеризоваться не одним сечением захвата, а двумя; γ_{mn} — вероятность захвата при передаче энергии электрону и γ_{pp} — при передаче энергии дырке. Кроме того, скорость захвата r будет пропорциональна не квадрату, а кубу концентрации носителей заряда из-за участия трех носителей в процессе рекомбинации. Учитывая эти особенности, можно для межзонной ударной рекомбинации записать r в виде

$$\left. \begin{aligned} r_n &= \gamma_{mn} n^2 p, \\ r_p &= \gamma_{pp} n p^2. \end{aligned} \right\} \quad (4.76)$$

А полный коэффициент захвата будет

$$r = r_n + r_p. \quad (4.77)$$

В условиях термодинамического равновесия $r_n = r_{n0}$ и $r_p = r_{p0}$. Подставляя эти величины в (4.76), найдем

$$\gamma_{mn} = \frac{r_{n0}}{n_0^2 p_0} \quad \text{и} \quad \gamma_{pp} = \frac{r_{p0}}{p_0^2 n_0}. \quad (4.78)$$

Соответственно полный коэффициент захвата будет

$$r = \frac{r_{n0}}{n_0^2 p_0} + \frac{r_{p0}}{p_0^2 n_0}. \quad (4.79)$$

Теперь запишем выражение для скорости генерации. Будем считать, что ударной рекомбинации соответствует обратный процесс — ударная ионизация. Скорость генерации посредством ударной ионизации пропорциональна числу быстрых избыточных электронов:

$$\left. \begin{aligned} g_n &= \gamma_{mn}^r n, \\ g_p &= \gamma_{pp}^r p. \end{aligned} \right\} \quad (4.80)$$

Из условий термодинамического равновесия можно выразить γ_{mn}^r и γ_{pp}^r как

$$\gamma_{mn}^r = g_{n0} / n_0 \quad \text{и} \quad \gamma_{pp}^r = g_{p0} / p_0. \quad (4.81)$$

С учетом равенства в равновесных условиях скоростей генерации и рекомбинации (;) можно записать полную скорость рекомбинации (4.63) в виде

что после простых преобразований дает

$$(4.82)$$

Переходя посредством (4.62) к выражению для времени жизни, получим

$$(4.83)$$

Или, подставляя сюда выражение (4.58), получим

$$(4.84)$$

Обозначив

$$(4.85)$$

и введя теперь величину — время жизни при ударной рекомбинации и собственном полупроводнике (при малом уровне инжекции () и дополнительном условии

$$\beta(p_0 + \Delta n) \approx n_0 + \Delta n, \quad (4.86)$$

получаем

$$\tau_0 = \frac{2\tau_i n_i^2}{(n_0 + p_0)(n_0 + \beta p_0)}. \quad (4.87)$$

При большом уровне инжекции ($\Delta n \gg n_0, p_0$) по мере увеличения Δn величина τ стремится к

$$\tau_\infty = 2\tau_i \left(\frac{n_i}{\Delta n}\right)^2 \frac{1}{1 + \beta}. \quad (4.88)$$

Так как обычно $\beta < 1$, то

$$\tau_\infty \approx 2\tau_i (n_i / \Delta n). \quad (4.89)$$

Из приведенных выражений следует, что времена жизни существенно зависят от температуры и от ширины запрещенной зоны. Можно сделать вывод, что в полупроводниках с узкой запрещенной зоной вероятность ударной рекомбинации возрастает. Вероятность ударной рекомбинации увеличивается с ростом температуры.

Время жизни при ударной рекомбинации соответственно будет обратно пропорционально вероятности рекомбинации.

Следует отметить, что среди механизмов рекомбинации зарядов называются рекомбинация через однозарядные локальные центры и механизмы рекомбинации через локальные центры.

В первом случае предполагается, что в полупроводнике имеется один тип локальных центров. Соответствующий энергетический уровень расположен в запрещенной зоне между зонами проводимости и валентной. Процесс рекомбинации сводится к захвату этими уровнями электронов из зоны проводимости, а затем к захвату дырки из валентной зоны.

Во втором случае экспериментально рекомбинационное излучение при переводе носителей из соответствующих зон на энергетические уровни локальных центров наблюдалось в германии и кремнии, а затем и в арсениде галлия. Однако строгой теории этого механизма до настоящего времени нет.

4.2.5. Термоэлектрические явления.

Термоэлектрическими явлениями называются явления термо-э. д. с. (Зеебека), Пельтье и Томсона.

Сущность явления термо-э. д. с. (открытого Зеебеком в 1821 г.) заключается в том, что в электрической цепи, состоящей из последовательно соединенных разных материалов, возникает электродвижущая сила (термо-э. д. с.), если места контактов поддерживаются при различных температурах, рис. 4.13.

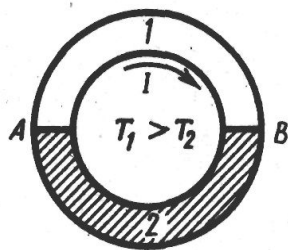


Рис. 331

Рис.4.13. Возникновение электрического тока в замкнутой цепи, состоящей из двух материалов при разной температуре контактов.

В простейшем случае, когда такая цепь состоит из двух различных материалов, она носит название термоэлемента или термопары. Экспериментальные исследования показали, что термоэлектродвижущая сила термопары зависит от температуры горячего () и холодного () спаев и от состава материалов, образующих термопару.

В небольшом интервале температур величину термо-э. д. с. E можно считать, с достаточной для практических целей точностью, пропорциональной

разности температур и некоторому коэффициенту , называемому коэффициентом термо-э. д. с.:

$$(4.90)$$

более точно, . Коэффициент термо-э. д. с. определяется в первую очередь материалами ветвей

термоэлемента, для обозначения которых мы ввели индексы 1 и 2, но зависит также и от интервала температур, в котором используется или исследуется термопара; во многих случаях с изменением температуры он даже меняет знак.

Явление Зеебека используется для измерения температуры. Для этого применяются термоэлементы или термопары, состоящие из двух соединенных между собой разнородных материалов. Термопары применяются как для измерения ничтожно малых разностей температур, так и для измерения очень высоких и очень низких температур (например, внутри доменных печей или жидких газов. Точность определения температуры с помощью термопар составляет, как правило, несколько кельвин, а у некоторых термопар достигает $\approx 0.01K$. Явление Зеебека может быть использовано для генерации электрического тока. Так в настоящее время к.п.д. полупроводниковых термобатарей достигает $\approx 18\%$. Следовательно, совершенствуя полупроводниковые термоэлектрогенераторы, можно добиться эффективного прямого преобразования солнечной энергии в электрическую.

Явление Пельтье (открытое в 1834 г.) представляет собой эффект, обратный явлению Зеебека; сущность его состоит в том, что при прохождении тока в цепи, состоящей из различных проводников, в местах контакта в дополнение к теплоте Джоуля выделяется или поглощается (в зависимости от направления тока) некоторое количество тепла Q_n , пропорциональное прошедшему через контакт количеству электричества (т. е. силе тока и времени) и некоторому коэффициенту Π , зависящему от природы находящихся в контакте материалов, названному коэффициентом Пельтье:

$$Q_n = \Pi It \tag{4.91}$$

Физический смысл коэффициента Пельтье можно определить как количество тепла, выделяющегося (или поглощающегося) на контакте при прохождении через него единицы заряда.

Рассмотрим замкнутую цепь, состоящую из двух разнородных проводящих материалов 1 и 2, рис. 4.14, по которым пропускают ток I' (его направление в

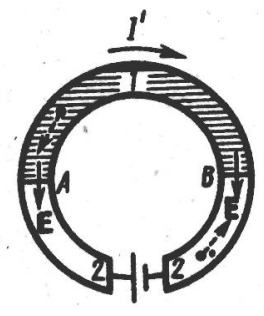


Рис. 4.14. Демонстрация явления Пельтье при пропускании тока через замкнутую цепь, состоящую из двух материалов.

данном случае выбрано совпадающим с направлением термотока ($T_1 > T_2$). Согласно наблюдениям Пельтье, спай A , который при явлении Зеебека поддерживался при более высокой температуре, будет теперь охлаждаться, а спай B — нагреваться. При изменении направления тока I' спай A будет нагреваться, спай B — охлаждаться.

Объяснить явление Пельтье можно следующим образом. Электроны по разную сторону спаев обладают различной средней энергией (полной — кинетической плюс потенциальной). Если электроны (направление их движения задано на рис. 4.14 пунктирными стрелками) пройдут через спай B и попадут в область с меньшей энергией, то избыток своей энергии они отдадут кристаллической решетке и спай будет нагреваться. В спае A электроны переходят в область с большей энергией, забирая теперь недостающую энергию у кристаллической решетки, и спай будет охлаждаться.

Явление Пельтье используется в термоэлектрических полупроводниковых холодильниках и в некоторых электронных приборах.

В. 1856 г. Томсон, применив к термоэлектрическим явлениям первое и второе начала термодинамики, вывел соотношение между коэффициентом термо-э. д. с. и коэффициентом Пельтье :

$$\tag{4.92}$$

и предсказал существование третьего явления, названного в дальнейшем явлением Томсона.

Явление Томсона (установленное экспериментально Леру в 1867 г.) заключается в следующем: если вдоль проводника, по которому протекает электрический ток, существует перепад температур, то в дополнений к теплоте Джоуля в объеме проводника выделяется или поглощается, в зависимости от направления тока,

некоторое количество тепла Q_τ (теплота Томсона), пропорциональное силе тока I , времени t , перепаду температур $(T - T_0)$ и коэффициенту (названному коэффициентом Томсона), зависящему от природы материала:

$$Q_\tau = \tau(T - T_0)It. \quad (4.93)$$

Согласно теории Томсона коэффициент термо-э. д. с. пары проводников связан с коэффициентами Томсона соотношением

$$\frac{d\alpha_{1,2}}{dT} = \frac{\tau_1 - \tau_2}{T}. \quad (4.94)$$

Оба соотношения Томсона (4.92) и (4.94) многократно подтверждались экспериментальными исследованиями. Качественно термоэлектрические явления можно объяснить следующим образом. Энергия свободных невырожденных электронов растет пропорционально температуре. Если вдоль проводника существует перепад температур, то электроны на горячем конце приобретают более высокие энергии и скорости, чем на холодном; в полупроводнике концентрация свободных электронов также растет с температурой и, следовательно, будет на горячем конце больше, чем на холодном; в результате всего этого возникает поток электронов от горячего конца к холодному и на холодном конце накапливается отрицательный заряд, а на горячем остается нескомпенсированный положительный заряд.

Этот процесс накопления заряда будет продолжаться до тех пор, пока возникшая таким образом разность потенциалов не создаст поток электронов, движущийся в обратном направлении, равный первичному, и благодаря этому не установится стационарное состояние. Алгебраическая сумма таких разностей потенциалов в цепи дает одну составляющую термоэлектродвижущей силы.

Вторая составляющая есть следствие температурной зависимости контактной разности потенциалов.

Если оба спая термоэлемента находятся при одной и той же температуре, то контактные разности потенциалов равны и направлены в противоположные стороны; в этом случае они компенсируют друг друга. Если же

температура спаев различна, то разность ΔV_k на контактах дает свой «контактный» вклад в термо-э. д. с., который может быть сравним с объемной термо-э. д. с. .

В металлах скорости электронов из-за вырождения очень слабо зависят от температуры, уровень химического потенциала и работа выхода также почти не зависят от температуры. Поэтому термо-э. д. с. металлов очень мала.

Описанные выше явления в дырочных полупроводниках протекают аналогично тому, как в электронных, с той лишь разницей, что на холодном конце скапливаются положительно заряженные дырки, а на горячем остается нескомпенсированный отрицательный заряд. В силу этого в термопаре, состоящей из дырочного и электронного полупроводника, термоэлектродвижущие силы ветвей складываются.

Наряду с примесными полупроводниками, в которых ток переносится носителями одного знака (либо только дырками, либо только электронами), существует большое количество веществ (и полупроводников, и металлов) со смешанной проводимостью, т. е. веществ, в которых электрический ток переносится одновременно и электронами и дырками. В последнем случае к холодному концу одновременно диффундируют и электроны и дырки, и их заряды компенсируют друг друга. Если концентрации и подвижности электронов и дырок равны, то электрическое поле вовсе при этом не образуется и термо-э. д. с. равна нулю; таким материалом с нулевой термо-э. д. с. является, в частности, свинец. Если же числа свободных электронов и дырок либо их подвижности отличаются, то термо-э. д. с. возникает как разностный эффект, и величина ее обычно относительно мала.

Если вдоль проводника, по которому протекает электрический ток, существует перепад температур, причем направление тока соответствует движению электронов от горячего конца к холодному, то, переходя из более горячего участка в более холодный, электроны передают избыточную энергию окружающим атомам, чем вызывают нагрев проводника (выделение тепла); при обратном направлении тока электроны, проходя из более холодного участка в более горячий, пополняют свою энергию за счет окружающих атомов (поглощение тепла). Этим и объясняется в первом приближении явление Томсона. Для более точного описания явления следует учесть также, что в первом случае электроны тормозятся, а во втором — ускоряются полем термоэлектродвижущей силы, что изменяет значение коэффициента Томсона, а в некоторых случаях приводит даже к перемене его знака.

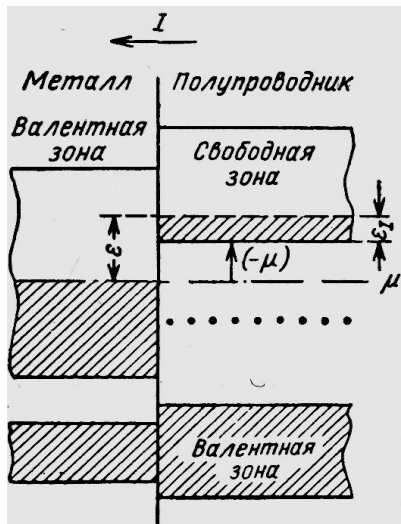


Рис. 4.15. Схема расположения энергетических уровней на контакте полупроводника с металлом (ε — энергия, выделяющаяся в виде тепла при переходе одного электрона из полупроводника в металл).

Причина возникновения явления Пельтье заключается в том, что средняя энергия электронов, участвующих в переносе тока в двух материалах, находящихся в контакте, различна, несмотря на то, что уровни их электрохимического потенциала совпадают. Это особенно наглядно видно на контакте электронного полупроводника и металла (рис. 4.15). Предположим, что направление поля соответствует направлению перехода электронов из полупроводника в металл.

Если бы электроны, находящиеся на примесных уровнях, могли бы точно так же перемещаться под действием электрического поля, как электроны в зоне проводимости, то в среднем энергия электронов, участвующих в электрическом токе, в полупроводнике равнялась бы энергии электронов на уровне Ферми в металле. При таких условиях переход электронов в металл не нарушил бы теплового равновесия в последнем; но электроны на примесных уровнях неподвижны, а энергия свободных электронов в зоне проводимости значительно выше энергии на уровне химического потенциала. Перейдя в металл, эти электроны опускаются до уровня Ферми, отдавая при столкновениях свою избыточную энергию атомам металла. Выделяющаяся при этом теплота и есть теплота Пельтье. Так как электроны приходят в тепловое равновесие в результате нескольких десятков соударений в непосредственной близости от контакта, то и вся теплота Пельтье выделяется почти на самом контакте.

При противоположном направлении тока весь процесс идет в обратном направлении. В полупроводник из металла могут перейти электроны с энергетических уровней, расположенных выше дна зоны проводимости, т. е. значительно выше уровня Ферми. Тепловое равновесие в металле при этом нарушается и восстанавливается за счет энергии тепловых колебаний решетки. При этом поглощается теплота Пельтье.

Основные формулы

Получив выражение для одного (любого) из термоэлектрических коэффициентов, остальные можно вывести, воспользовавшись соотношениями Томсона (4.92) и (4.94).

Рассмотрим для этого количественно энергетический баланс при переходе электрона из полупроводника в металл. Как видно из рис. 4.15, каждый электрон при таком переходе будет переносить избыточную энергию, состоящую из двух частей:

1. — кинетической энергии в зоне проводимости. Значок $\frac{1}{2}mv^2$ означает, что это энергия электрона, участвующего в электрическом токе. Следует сразу же подчеркнуть, что среднее значение этой энергии не равно средней тепловой энергии:

Объясняется это тем, что доля участия электронов с различной энергией в электрическом токе зависит от механизма рассеяния. Так, например, при рассеянии на ионах примеси длина свободного пробега пропорциональна четвертой степени скорости. Это значит, что в данном случае в электрическом токе (или в любом другом, например, диффузионном потоке электронов) будут играть основную роль быстрые электроны и средняя энергия в потоке будет намного больше средней тепловой. В случае невырожденного электронного газа теория дает следующее выражение для $\langle \varepsilon \rangle$:

$$(4.95)$$

где r — показатель степени в зависимости длины свободного пробега l электрона от энергии.

2. Второе слагаемое (ε_2) в энергии, переносимой электроном из полупроводника в металл, как видно на рис. 4.15, измеряется расстоянием от дна зоны проводимости до уровня Ферми. Если за нуль отсчета энергии взять дно зоны проводимости полупроводника, то

$$\varepsilon_2 = |\mu| = -\mu, \quad (4.96)$$

так как μ в данном случае лежит ниже дна зоны и, следовательно, отрицательно, а ε_2 — положительный вклад в выделяющуюся теплоту.

Таким образом, согласно (4.95) и (4.96)

$$\bar{\varepsilon} = \bar{\varepsilon}_j + (-\mu) = \bar{\varepsilon}_j - \mu. \quad (4.97)$$

Опускаясь в металле до уровня Ферми, электрон передает избыточную энергию атома и металла, и она, таким образом, выделяется в виде тепла вблизи контакта.

Точно такое же количество тепла поглотится на контакте при противоположном направлении тока. Коэффициент Пельтье, представляющий собой отношение количества выделившегося тепла к перешедшему через контакт заряду, выразится, следовательно, формулой

$$\Pi_{PM} = \frac{\bar{\varepsilon}}{-e} = -\frac{1}{e}(\bar{\varepsilon}_j - \mu) \quad (4.98)$$

и согласно (4.92) коэффициент термо-э. д. с. полупроводника по отношению к металлу равен

$$\alpha_{PM} = \frac{\Pi}{T} = -\frac{k}{e} \left(\frac{\bar{\varepsilon}_j}{kT} - \mu \right), \quad (4.99)$$

где k — постоянная Больцмана.

Так как термо-э. д. с. металлов очень мала, то мы вправе в дальнейшем индексы при α опускать и всю термо-э. д. с. относить к полупроводнику.

Величина μ , для невырожденного электронного полупроводника с простой параболической зоной определяется выражением

$$(4.100)$$

Таким образом, получаем окончательно

$$(4.101)$$

Совершенно аналогичный, с точностью до знака, вид имеет выражение для термо-э. д. с. дырочного полупроводника

$$(4.102)$$

Как уже упоминалось, в случае смешанной проводимости, когда электрический ток переносится электронами и дырками, термо-э. д. с. как правило, значительно ниже. В этих условиях уровень химического потенциала обычно находится приблизительно посередине запрещенной зоны и почти не смещается с изменением температуры (за исключением того сравнительно редкого случая, когда эффективные массы электронов и дырок сильно различаются); поэтому контактный член в термо-э. д. с. близок к нулю. Что касается объемного члена, то он обычно тоже очень мал; если подвижности и концентрации электронов и дырок одинаковы, то объемное поле вообще не возникает, так как оба рода носителей диффундируют от горячего участка к холодному в равном количестве.

Если же концентрация или подвижность носителей одного знака, например, электронов, больше, чем другого, то они диффундируют на холодный конец в большом количестве до тех пор, пока возникшее вследствие этого поле (тормозящее электроны и ускоряющее дырки) не уравнивает оба потока.

Формула для термо-э. д. с. выглядит в этом случае следующим образом:

$$\alpha = \frac{\alpha_p u_p p - \alpha_n u_n n}{u_p p + u_n n}, \quad (4.103)$$

где u_p, u_n, p и n — соответственно подвижности и концентрации дырок и электронов; α_p и α_n — «парциальные» термоэлектродвижущие силы, вычисленные согласно (4.101) и (4.102).

4.2.6. Явления в сильных электрических полях

Все выводы относительно подвижности и электропроводности, сделанные выше, основывались на двух предположениях: 1) направленная добавка к скорости мала по сравнению с тепловой скоростью v_0 и 2) концентрация носителей электричества не зависит от поля и остается равновесной, определяемой статистикой Максвелла — Больцмана или Ферми. В действительности оба эти условия соблюдаются в полупроводниках до определенных для каждого материала и температуры критических полей; новые эффекты, которые происходят при нарушении хотя бы одного из этих условий, называются явлениями в сильных полях. На них мы и остановимся коротко в этом параграфе.

В первую очередь — это отступления от закона Ома, т. е. нарушение линейной зависимости тока от напряжения. Согласно формулам

$$j = \sigma E \quad \text{и} \quad \sigma = en\mu \quad (4.104)$$

это будет иметь место, когда либо подвижность, либо концентрация носителей начнут зависеть от поля.

Рассмотрим каждое из этих явлений в отдельности.

Согласно выражению

$$u = \frac{e l}{m v} \quad (4.105)$$

подвижность начинает зависеть от поля с того момента, как скорость v в (4.83) перестает быть постоянной, т.

е. когда добавкой Δv к скорости за счет поля нельзя пренебречь по сравнению с тепловой скоростью v_0 , иными словами, критерием слабого поля является условие

$$\Delta v \ll v_0. \quad (4.106)$$

Добавку к тепловой скорости Δv можно оценить, исходя из закона сохранения энергии (в дальнейшем мы увидим, что эта оценка не точна):

$$(4.107)$$

или, так как (поскольку направлено беспорядочно, а — по полю), то

$$(4.108)$$

Таким образом, сформулированный выше критерий слабых полей будет иметь вид

$$(4.109)$$

При комнатной температуре ≈ 300 К, полагая тепловую скорость электрона $v_0 \approx 10^8$ см/сек и время между соударениями $\tau \approx 10^{-14}$ сек, получаем

$$E \approx 10^4 \text{ В/см}. \quad (4.110)$$

В этом случае критическое поле приблизительно равно 10^4 В/см.

При понижении температуры средняя кинетическая энергия электронов падает пропорционально температуре, длина свободного пробега, если исключить рассеяние на ионах примеси (на котором мы остановимся отдельно), либо остается постоянной, либо растет с понижением температуры; поэтому согласно (4.109) критическое поле, как правило, довольно резко падает с понижением температуры. Так, например, в достаточно чистом германии подвижность при гелиевых температурах начинает зависеть от поля при полях меньше 10^4 В/см.

Здесь следует подчеркнуть, что критические поля в неоднородных полупроводниках могут проявляться при очень малых разностях потенциала. Так, например, в кристалле, в котором имеется p - n переход, на него, как

правило, падает все напряжение (так как сопротивление p - n перехода много больше сопротивления толщи полупроводника). Толщина p - n перехода d меняется в различных приборах от 10^{-3} до 10^{-7} см. При $d = 10^{-7}$ см и разности потенциалов 1 В поле достигнет колоссальных размеров:

$$E = \frac{V}{l} = 10^7 \frac{B}{\text{см}}. \quad (4.111)$$

Следовательно, в тех случаях, когда эффекты, связанные с сильными полями, нежелательны, p - n переход следует делать достаточно толстым. Это, в частности, относится к силовым выпрямителям; одним из условий, необходимых для того, чтобы они выдерживали большое обратное напряжения (порядка 10^3 В/см), является достаточная толщина p - n перехода ($d \sim 10^{-4}$ см). При толстом p - n переходе прямая ветвь вольтамперной характеристики хуже, чем при тонком, однако на это приходится идти, чтобы обеспечить высокое обратное напряжение. Напротив, в приборах, основанных на явлениях в сильных полях (туннельных диодах и приборах, основанных на лавинном эффекте), p - n переход должен быть тонким.

В поликристаллических образцах также очень часто сопротивление прослоек между зернами бывает во много раз больше сопротивлений самих зерен. Размеры зерен обычно колеблются в пределах от 10^{-2} до 1 мм, а толщина прослоек — в пределах 10^{-3} — 10^{-6} см. Следовательно, и в этом случае поле в прослойках может быть в 10 — 10^6 раз больше, чем поле, приложенное к образцу. На этом «усилении поля» основан целый ряд полупроводниковых приборов: действие нелинейных сопротивлений из карбида кремния и других материалов, фоточувствительных пленок из сернистого свинца и т. д.

После этого краткого отступления вернемся к влиянию поля на подвижность — проанализируем качественно, как это влияние зависит от механизма рассеяния носителей.

Рассеяние на тепловых колебаниях акустической ветви атомной решетки.

Почто во всех явлениях переноса длина свободного пробега электрона l зависит от его энергии ε , так как она определяет среднюю энергию электронов в потоке. Как известно, эта зависимость носит степенной характер

$$l \propto \varepsilon^r. \quad (4.112)$$

При рассеянии на ионах примеси $r = 2$, при рассеянии на акустических колебаниях $r = 0$, при рассеянии на оптических колебаниях ионных кристаллов $r = 1/2$ при $T < \theta$ и $r = 1$ при $T > \theta$.

Согласно формуле (4.112) можно записать

$$\mu \propto \frac{l}{v} \propto \varepsilon^{r-1/2}. \quad (4.113)$$

В рассматриваемом нами случае, следовательно,

$$(4.114)$$

Согласно (4.108) в сильном поле, следовательно,

$$(4.115)$$

то есть уменьшается с ростом поля.

Рассеяние на ионах примеси ()

$$\text{и} \quad (4.116)$$

Следовательно, в этом случае подвижность должна расти с ростом поля, что действительно наблюдалось в германии при температурах порядка 20° К и поле, приблизительно равном 100 В/см. Однако при дальнейшем росте поля подвижность начинала падать. Это странное на первый взгляд явление имеет простое качественное объяснение.

Действительно, при температурах, отличных от абсолютного нуля, наряду с примесным рассеянием имеет место рассеяние на тепловых колебаниях решетки, однако с понижением температуры интенсивность тепловых колебаний падает, одновременно с этим падает и средняя тепловая скорость электронов, а следовательно, соответственно растет эффективность рассеяния на ионах примеси, и при низких температурах это рассеяние начинает играть преобладающую роль. Поэтому при наличии ионов примеси температурная зависимость подвижности в широком интервале температур имеет вид, схематически представленный на рис. 4.16. В области рассеяния на тепловых колебаниях решетки подвижность падает с ростом температуры по

закону

при низких температурах преобладает рассеяние на ионах примеси и подвижность растет пропорционально

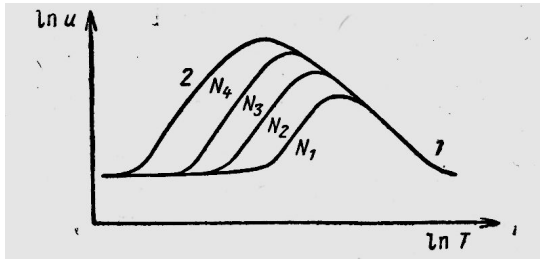


Рис. 4.16. Температурная зависимость подвижности для различного количества ионов примеси:

1 — область высоких температур (т. е. рассеяния на акустических тепловых колебаниях) $\mu \propto T^{\frac{3}{2}}$; 2 — область низких температур рассеяния на ионах примеси.

При этом излом в зависимости подвижности от температуры проявляется при тем более высокой температуре, чем больше содержание ионов ($N_1 > N_2 > N_3 > N_4$). Наконец, при дальнейшем понижении температуры наступает вырождение: скорость электронов перестает зависеть от температуры и поэтому подвижность перестает зависеть от температуры и содержания примесей.

Посмотрим теперь, что происходит в достаточно сильных полях, когда скорость электрона начинает увеличиваться с ростом поля. С этой точки зрения (т. е. в связи с ростом скорости электронов) увеличение поля эквивалентно росту температуры электронного газа до некоторой эффективной (T_3). Поэтому эффективность рассеяния на ионах падает и подвижность растет до тех пор, пока не начинает сказываться

рассеяние на тепловых колебаниях решетки, после этого она начинает падать по закону $\mu \propto E^{-\frac{1}{2}}$. Поэтому зависимость подвижности от поля должна иметь вид, аналогичный температурным зависимостям подвижности (рис. 4.16). Эти закономерности, как мы уже упоминали, качественно подтвердились на германии при низких температурах.

Рассеяние на оптических колебаниях ионной решетки.

При температурах ниже температуры Дебая $r = 1$ и

$$\mu \propto \frac{l}{v} \propto E^{\frac{1}{2}}, \quad (4.117)$$

т. е. подвижность должна расти с увеличением поля; при температуре выше температуры Дебая и подвижность не должна, если исходить из этих качественных соображений, зависеть от поля:

$$(4.118)$$

Однако более строгий анализ, проведенный учеными, показал, что в этом случае в сильных полях (т. е. при росте) нарушаются предпосылки, необходимые для правомерности выражения для подвижности (4.84), и все выводы надо пересмотреть.

В заключение этого раздела уточним определение сильного поля. В действительности *полевые приращения скорости* , входящие в неравенство (4.105) и равенство (4.107), разные. Дело в том, что большинство столкновений, которые испытывает электрон, почти упругие, а следовательно, при них скорость почти не меняется по абсолютной величине и требуется большое число таких столкновений, чтобы электрон пришел в тепловое равновесие с решеткой; поэтому добавка к скорости электрона в стационарном состоянии, входящая в неравенство (4.105), много больше добавки, приобретаемой на *одной длине свободного пробега*, определяемой соотношением (4.107), эту добавку мы обозначим .

Оценим отношение на примере столкновений электронов с ионами. В этом случае при одном столкновении электрон теряет следующую часть своей энергии:

$$(4.119)$$

Следовательно, энергия электрона в электрическом поле будет возрастать до тех пор, пока потеря энергии при

одном столкновении не будет равна энергии, приобретаемой за время одного свободного пробега. Таким образом, условие стационарности будет иметь вид

$$eE\bar{l} - \frac{m}{M}\bar{\varepsilon} = 0,$$

$$\bar{\varepsilon} = eE\bar{l}\frac{M}{m} = eE\bar{l}\alpha,$$
(4.120)

откуда
где

$$\alpha = \frac{M}{m}$$

и

$$\Delta v = \Delta v_1 \sqrt{\frac{M}{m}} = \Delta v_1 \sqrt{\alpha}.$$

Полагая $M = 10^4 m$, получаем

$$\Delta v = 100 \Delta v_1.$$
(4.121)

Согласно (2П70) явления в сильных полях могут начинаться тогда, когда приращение энергии электрона на длине свободного пробега в 10^4 раз меньше его средней тепловой энергии. Аналогичные оценки параметра α для рассеяния на тепловых колебаниях дают $\alpha = 100$; эта цифра ближе к истине, так как в сделанном выше расчете мы не учли тепловых колебаний, столкновения с которыми являются менее упругими и роль которых будет расти по мере роста энергии электрона.

Влияние поля на концентрацию носителей

Существует несколько механизмов влияния поля на концентрацию носителей:

- электростатическая ионизация — за счет туннельного эффекта,
- термоэлектронная ионизация,
- ударная ионизация.

Рассмотрим каждое из этих явлений в отдельности.

Электростатическая ионизация.

На рис. 4.17 представлена зонная структура беспримесного полупроводника при наличии сильного электрического поля.

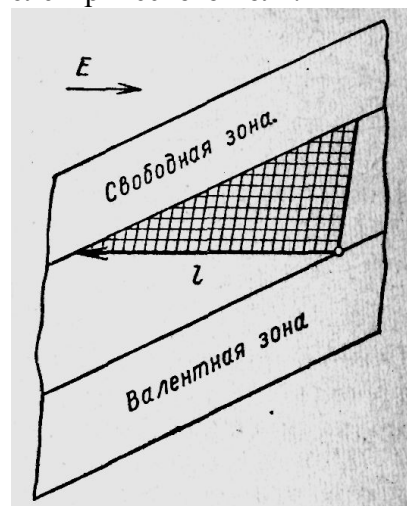


Рис. 4.17. Схема зонной диаграммы беспримесного полупроводника при наличии сильного поля, иллюстрирующая электростатическую ионизацию.

Заштрихована площадь барьера.

Как видно из рисунка, в этом случае возможен переход электрона из валентной зоны в свободную. Оценим вероятность такого перехода. Волновая функция зонного электрона может быть преобразована к виду, более удобному для оценки этой вероятности; для этого выразим волновое число через энергию электрона:

$$k = \frac{p}{h} \sqrt{2Tm} = \frac{1}{h} \sqrt{2m(E-U)}, \quad (4.122)$$

и тогда

$$\psi = A \exp \left[\frac{2\pi i}{h} \sqrt{2m(E-U)} x \right], \quad (4.123)$$

где T — кинетическая; U — потенциальная и E — полная энергия электрона.

Формула (4.123) показывает, что при попадании электрона в область отрицательной полной энергии, т. е. в область потенциального барьера, волновая функция начинает затухать по закону

$$\psi(x) = A e^{-\alpha x},$$

где

$$\alpha = \frac{2\pi}{h} \sqrt{2\Delta E m}.$$

Здесь $\Delta E = (U - E)$ — высота барьера; вероятность просачивания через барьер w определяется соотношением

$$w \approx e^{-\int_0^l \alpha^2 dx}. \quad (4.124)$$

В рассматриваемом случае ΔE_0 — ширина запрещенной зоны, барьер имеет форму треугольника, основание которого l , как видно из рис. 2.7, определяется условием $eEl = \Delta E_0$, откуда

$$w \approx \exp \left(-\frac{\pi^2 m a}{h^2} \frac{\Delta E_0^2}{lE} \right). \quad (4.125)$$

Оценка по формуле (4.125) показывает, что при $\Delta E \approx 1$ эВ вероятность туннельного эффекта становится заметной при полях порядка 10^7 В/см.

При электростатической ионизации доноров вероятность туннельного эффекта значительно возрастает, так как при этом и высота, и ширина барьера становятся меньше; в данном случае этот эффект может быть замечен при полях порядка 10^6 В/см.

Еще более вероятен туннельный эффект на контакте полупроводника и металла, если при этом ширина барьера не увеличивается областью объемного заряда. На этом эффекте основано действие туннельных диодов.

Особенностью электростатической ионизации является тот факт, что ее вероятность не зависит от температуры; по этому признаку такой механизм ионизации можно отличить от двух других, к рассмотрению которых мы сейчас перейдем.

Термоэлектронная ионизация.

На рис. 4.18 представлен энергетический спектр донорного примесного атома при наличии сильного электрического поля. Как видно из рисунка, в этом случае действие поля проявляется в двух аспектах: 1) появляется возможность безактивационного просачивания в зону проводимости путем туннельного эффекта (этот механизм мы рассмотрели выше) и 2) уменьшается на величину энергия активации, необходимая для теплового заброса электрона в зону проводимости. Так как величина (как видно из рисунка) растет с ростом поля, то соответственно будет расти число электронов в свободной зоне и электропроводность.

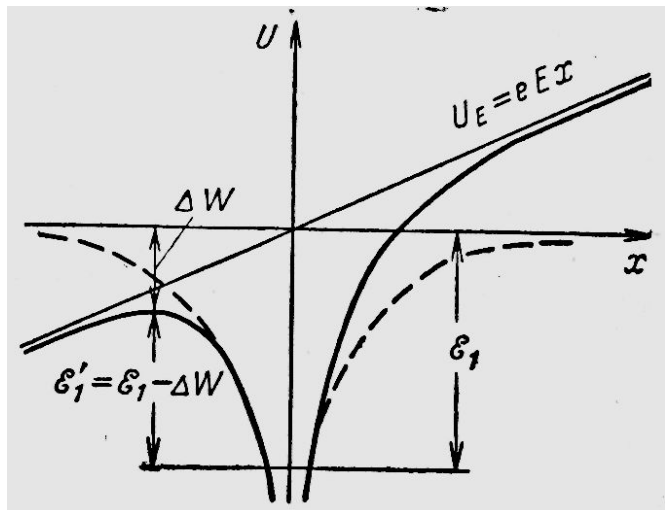


Рис. 4.18. Схема термоэлектронной ионизации.

Этот механизм ионизации называется термоэлектронным. Нетрудно оценить этот эффект количественно. В отсутствие поля число электронов в зоне проводимости

$$n_0 = A e^{-\Delta E_1 / 2kT}, \quad (4.126)$$

где A — коэффициент, слабо зависящий от температуры и ΔE_1 — энергия ионизации (активации) примесных атомов (см. рис. 4.16).

При наличии поля величина E_1 снижается на ΔW и

$$E'_1 = E_1 - \Delta W, \quad (4.127)$$

$$n = A e^{\frac{\Delta E'_1}{2kT}} = n_0 e^{\frac{\Delta W}{2kT}}. \quad (4.128)$$

Величина ΔW численно равна потенциальной энергии электрона в той точке x_0 , где энергия достигает максимума (см. рис. 4.18).

Согласно рис. 4.18 при наличии поля ход потенциальной энергии вблизи примесного атома в направлении поля (x) будет выражаться соотношением

$$(4.129)$$

Высота потенциального барьера будет максимальна в точке , определяемой уравнением

$$(4.130)$$

откуда

$$(4.131)$$

Высота барьера в точке

$$(4.132)$$

и следовательно

$$(4.133)$$

Оценка n согласно формуле (4.133) показывает, что рост концентрации носителей за счет термоионной ионизации должен наблюдаться при относительно слабых полях, порядка 10^2 — 10^3 В/см, что в

действительности не имеет места. Для того чтобы согласовать теорию Я. И. Френкеля (изложенную выше) с опытом, А. И. Губанов внес в нее два уточнения.

1. Если учесть экранировку заряда примесного атома свободными электронами, то ход потенциала вблизи него будет выражаться формулой

$$U = \frac{e^2}{x\mathcal{J}} e^{-\frac{x}{r_0}}, \quad (4.134)$$

где r_0 — радиус экранировки Дебая, равный

$$r_0 = \sqrt{\frac{\mathcal{J}kT}{8\pi ne^2}}. \quad (4.135)$$

Согласно (4.134) U быстрее стремится к нулю по мере роста r , чем без учета экранировки, и, следовательно, при наличии поля «горб» потенциала будет ближе к ядру и выше; поэтому эффект поля будет меньше.

2. Согласно развитой теории электрон каждый раз покидает примесный атом, когда его энергия достигает значения (4.127).

В действительности это неточно, так как только в одной точке орбиты, x_0 , энергии электрона достаточно, чтобы оторваться от примесного атома (во всех остальных точках орбиты высота потенциального барьера выше).

Для того чтобы электрон достиг этой точки, не растратив по пути свою энергию, приобретенную в поле, необходимо, чтобы длина орбиты была меньше длины свободного пробега электрона $2\pi x_0 < l$ или, более грубо,

$$l > x_0. \quad (4.136)$$

Учет этих двух поправок позволяет лучше согласовать изложенную выше теорию с опытом.

Согласно теории Я. И. Френкеля эффект поля должен экспоненциально расти с температурой.

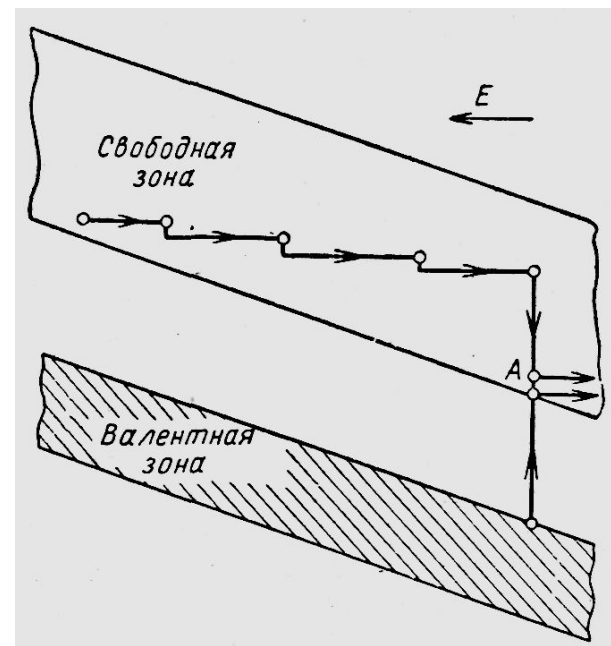


Рис4.19. Схема ударной ионизации.

Ударная ионизация.

На рис. 4.19 схематически представлено движение электрона в зоне проводимости при наличии сильного поля. Как видно из рис. 4.19, в точке А его кинетическая энергия становится достаточной для того, чтобы перебросить электрон из валентной зоны в свободную, после чего весь процесс может начаться сначала. Поля, при которых начинается ударная ионизация, можно грубо оценить согласно формуле (4.126)

$$(4.137)$$

где ΔE — ширина запрещенной зоны; недостатком формулы (4.137) является то, что в нее входит длина свободного пробега электрона, которая, как мы видели, сама зависит от поля.

Однако соотношение (4.137) правильно указывает, что вероятность ударной ионизации должна зависеть от ΔE , т. е. от чистоты материала и от температуры, что действительно наблюдается на опыте. Теоретические оценки

и опыт показывают, что ударная ионизация начинает играть существенную роль при полях $10^4 - 10^6$ В/см. Впервые предположение об ударной ионизации было высказано А. Ф. Иоффе для объяснения механизма пробоя твердых диэлектриков. Теория ударной ионизации была развита Фрейлихом, Гиппелем, Келдышем и др.

Существует два механизма ударной ионизации. *Первый из них — низковольтный*, т. е. такой, при котором электрон на одной длине свободного пробега приобретает лишь незначительную энергию, но все же превышающую ту, которую он теряет при соударении.

При этом накопление энергии электроном происходит постепенно и, так как при каждом или в результате небольшого числа столкновений электрон теряет направленную по полю составляющую скорости (v_x), она (т. е. направленная составляющая скорости) остается малой ($v_x \ll v$), а растет скорость хаотического движения, т. е. тепловая скорость. Таким образом, при этом нарушается тепловое равновесие между решеткой и электронным газом; последний приобретает свою собственную температуру, которая может значительно превосходить температуру решетки.

Так как в этом случае рост энергии электрона происходит постепенно, то такой пробой должен развиваться более эффективно в толстом слое (в противоположность туннельному эффекту). Эту закономерность подтвердил опыт: пробивное напряжение *p-n* перехода падает с ростом толщины запирающего слоя L .

Разумеется, эта закономерность наблюдается до определенного предела: если толщина слоя становится настолько большой, что все электроны, созданные полем, не успевают из него выйти, а рекомбинируют в нем, то критическое поле перестает зависеть от L .

Второй механизм ударной ионизации, в отличие от первого, можно условно назвать **высоковольтным**; в этом случае электрон уже на одной длине свободного пробега приобретает скорость, достаточную для вырывания другого электрона.

В этом случае становится велика направленная составляющая скорости v_x , а хаотическая скорость может оставаться близкой к нормальной, если напряженность поля такова, что он (электрон) при неупругом столкновении всю приобретенную скорость теряет.

Интересно отметить, что, хотя поле в этом случае больше, чем в первом, температура электронов может быть значительно ниже, так как начинает более эффективно действовать новый (неупругий) механизм столкновений.

В данном случае может наблюдаться периодический рост температуры электронов с ростом поля (так как она определяется разностью между энергией, приобретенной на длине свободного пробега, и потерей энергии при одном акте ионизации).

Отличие этого типа ионизации от рассмотренного выше заключается в том, что материал при этом сохраняет свой тип проводимости: в полупроводнике *n*-типа электрон ионизует донорные уровни и образец остается электронным, в полупроводнике *p*-типа дырки забрасывают электроны из валентной зоны на примесные уровни и сохраняется дырочная проводимость.

В *p-n* переходе ионизация производится, как правило, и дырками и электронами и при каждом акте создается пара электрон — дырка.