

до несущественного предэкспоненциального множителя

$$\mathcal{P} \sim \exp\left(-\frac{\mathcal{E}_0}{\mathcal{E}}\right), \quad (6.10)$$

где  $\mathcal{E}_0$  — постоянная, зависящая от параметров кристалла (от ширины запрещенной зоны, расположения экстремумов в зоне Бриллюэна, эффективных масс электрона в зонах проводимости и валентной и т. д.).

Пусть, например, изоэнергетические поверхности вблизи дна зоны проводимости и потолка валентной зоны суть сферы с центрами в центре зоны Бриллюэна; соответствующие эффективные массы обозначим через  $m_n$  и  $-m_p$  ( $m_p > 0$ ). Тогда

$$\mathcal{E}_0 = \frac{\pi E_g^{3/2} m_r^{1/2}}{2\hbar e}, \quad (6.11)$$

где  $m_r = \frac{m_n m_p}{m_n + m_p}$  — есть приведенная эффективная масса электрона и дырки. При  $E_g = 1$  эВ и  $m_r = 0,1 m_0$  формула (6.11) дает  $\mathcal{E}_0 \simeq \simeq 10^7$  В/см.

Видно, что в слабых полях ( $\mathcal{E} \ll \mathcal{E}_0$ ) вероятность туннельного эффекта совершенно ничтожна и им можно пренебречь. При этом достаточно рассматривать движение электрона только в одной зоне, т. е. справедливы рассуждения, изложенные в начале этого параграфа. С увеличением напряженности поля вероятность туннельного эффекта быстро возрастает, и при  $\mathcal{E} \simeq 0,1 \mathcal{E}_0$  его уже надо принимать во внимание.

Туннельный эффект может иметь место и в неоднородных электрических полях. С ним связан ряд явлений, наблюдаемых в полупроводниках. Так, он обуславливает один из возможных механизмов пробоя беспримесного образца: в достаточно сильном поле происходит спонтанное образование электронов проводимости и дырок в валентной зоне \*). Далее, на этом эффекте основано действие туннельных диодов (см. § VIII. 3).

Наконец, некоторые особенности поглощения света полупроводниками в сильном поле также объясняются туннельным эффектом (§ XVIII.10).

## § 7. Мелкие примесные уровни в гомеоплярном кристалле

Рассмотрим простейшую задачу теории реальных кристаллов — задачу о мелких уровнях, создаваемых в запрещенной зоне изолированными атомами посторонней примеси. Мелкими будем называть

\*) Этот механизм — не единственный. По-видимому, в ряде случаев более существенную роль играет механизм ударной ионизации; электрон проводимости, разгоняемый электрическим полем, приобретает энергию, достаточную для того, чтобы «вышибить» другой электрон из валентной зоны.

уровни, удовлетворяющие неравенству (4.5): расстояние от уровня до границы ближайшей к нему зоны должно быть мало по сравнению с шириной запрещенной зоны. Такие уровни возникают, например, в германии и кремнии при введении туда примесей элементов V и III групп периодической системы (§ II.9). Для расчета этих уровней можно воспользоваться уравнением (4.3) или его обобщениями (4.8) и (4.10). В качестве  $\delta U(\mathbf{r})$  здесь следует взять потенциальную энергию взаимодействия электрона с атомным остатком донора (или дырки — с акцептором). Точный вид  $\delta U$  зависит от химической природы примеси, и в общем случае его трудно определить. Для мелких уровней, однако, это обстоятельство не очень существенно, ибо, как мы увидим, главную роль в задаче играют расстояния  $r$  между электроном и ядром донора, значительно превышающие линейный размер атомного остатка. Как известно из электростатики, при этом детали, связанные с конкретным видом распределения заряда в атомном остатке, уже не играют роли, и функция  $\delta U(\mathbf{r})$  дается обычным законом Кулона:

$$\delta U = - \frac{Ze^2}{\epsilon r}. \quad (7.1)$$

Здесь  $Z$  — заряд атомного остатка в единицах  $e$  (в интересующем нас случае  $Z = 1$ ),  $\epsilon$  — диэлектрическая проницаемость решетки \*).

Мы получаем, таким образом, задачу об атоме водорода — с той лишь разницей, что в рассматриваемом случае в выражении (7.1) фигурирует диэлектрическая проницаемость среды, а оператор кинетической энергии, вместо истинной массы электрона, содержит эффективную (или, вообще, тензор обратной эффективной массы). Соответственно о модели примесного центра, основанной на использовании выражения (7.1) и метода эффективной массы, говорят как о водородной модели.

Отметим сразу же одно качественное следствие из этой модели, справедливое при любой анизотропии зон: расположение уровней не должно зависеть от химической природы примеси (разумеется, в пределах данной группы периодической системы — например, V группы для доноров и III группы для акцепторов в германии и кремнии). Таким образом, относительная разность между энергиями ионизации различных примесей этого типа в данном кристалле может служить мерой точности водородной модели. Данные таблицы 2.1 показывают, что для примесей III и V групп в Ge водородная модель дает неплохое приближение.

\*) В рассматриваемом нами случае гомеополлярных кристаллов вся поляризуемость безынерционная. Для мелких уровней область дисперсии  $\epsilon$  соответствует частотам, значительно превышающим характерное значение  $\mathcal{I}/\hbar$ , где  $\mathcal{I}$  — энергия ионизации примеси. Это позволяет считать  $\epsilon$  просто константой, равной своему статическому значению.

Конкретный расчет особенно прост в случае изотропной эффективной массы. Подставляя выражение (7.1) (при  $Z = 1$ ) в уравнении (4.3), мы имеем

$$-\frac{\hbar^2}{2m} \nabla^2 \chi - \frac{e^2}{\epsilon r} \chi = E \chi. \quad (7.2)$$

Введем характерные единицы длины и энергии:

$$a_B = \frac{\epsilon \hbar^2}{me^2}, \quad E_B = \frac{me^4}{2\epsilon^2 \hbar^2}. \quad (7.3)$$

Величины  $a_B$  и  $E_B$  обычно называют боровским радиусом и боровской энергией в кристалле.

Будем измерять длину в единицах  $a_B$ , а энергию — в единицах  $E_B$ . Тогда уравнение (7.2) примет вид (мы сохраняем прежние обозначения для безразмерных переменных)

$$-\nabla^2 \chi - \frac{2}{r} \chi = E \chi. \quad (7.2')$$

Решение уравнения (7.2') можно найти в курсах квантовой механики.

Имеются два участка спектра, соответствующие тому или другому знаку энергии  $E$ .

а) Непрерывный спектр:  $E > 0$ .

Здесь все значения  $E$  дозволены; функция  $\chi$  есть кулоновская волновая функция непрерывного спектра.

Согласно сказанному в § 4 эта область энергий соответствует зоне проводимости. Отличие функции  $\chi$  от плоской волны (4.4) описывает влияние примеси на движение свободных носителей заряда.

б) Дискретный спектр:  $E < 0$ .

Дозволенные значения энергии даются выражением

$$E_n = -\frac{E_B}{n^2}, \quad n = 1, 2, \dots \quad (7.4)$$

«Сглаженная» функция  $\chi$  в данном случае есть кулоновская волновая функция дискретного спектра. В частности, при  $n = 1$  (низший, или основной, уровень) мы имеем (в справедливости этого решения можно убедиться и простой подстановкой)

$$\chi = (\pi a_B^3)^{-1/2} \exp(-r/a_B). \quad (7.5)$$

Постоянный множитель здесь выбран в соответствии с обычным условием нормировки

$$\int |\chi|^2 d\mathbf{r} = 1.$$

Поскольку начало отсчета энергии у нас совмещено с дном зоны проводимости  $E_c$ , дискретные уровни (7.4) лежат в запрещен-

ной зоне. Электрон, занимающий один из таких уровней, локализован вблизи атома примеси (соответственно чему эти уровни иногда называют локальными). Наибольшая степень локализации соответствует основному состоянию ( $n = 1$ ). Тогда, согласно (7.5), «радиус орбиты», т. е. наиболее вероятное расстояние между атомом примеси и электроном, равен  $a_B$ . При больших значениях  $n$  «радиус орбиты» есть  $na_B$ . Из предыдущего ясно, что весь расчет имеет смысл лишь при достаточно слабой локализации электрона. Именно, формула (7.1) оправдана, если

$$a_B \gg a. \quad (7.6)$$

Это же условие обеспечивает и возможность использовать метод эффективной массы и представление о диэлектрической проницаемости.

Энергия ионизации  $\mathcal{I}$ , т. е. работа, которую надо совершить, чтобы перевести электрон на дно зоны проводимости, есть, очевидно,  $-E_1$ :

$$\mathcal{I} = E_B = \frac{e^2}{2\epsilon a_B} = 13,7 \frac{m}{m_0} \epsilon^{-2} \text{ эВ}. \quad (7.7)$$

Для германия  $\epsilon = 16$ . Полагая еще  $m = 0,2m_0$ , мы получаем из (7.7)  $\mathcal{I} = 1,07 \cdot 10^{-2}$  эВ, что близко к экспериментальным данным для мелких доноров (табл. 2.1). По этой причине для ориентировочной порядковой оценки величин, характеризующих электроны проводимости в германии, иногда пользуются изотропной моделью с эффективной массой  $m = 0,2m_0$ . Согласно (7.3) при этом  $a_B = 40 \text{ \AA}$ , что удовлетворяет неравенству (7.6).

Строгий расчет мелких донорных и акцепторных уровней в германии и кремнии должен основываться на уравнениях (4.8) и (4.10). Здесь приходится прибегать к приближенным методам расчета, детали которых можно найти в оригинальных статьях [4, 5, 6].

До сих пор мы явно предполагали, что носитель заряда взаимодействует лишь с одним примесным атомом. Фактически атомов примеси всегда много, и, строго говоря, вместо (7.1) надо было бы писать

$$\delta U(\mathbf{r}) = - \sum_i \frac{Z_i e^2}{\epsilon |\mathbf{r} - \mathbf{R}_i|}, \quad (7.1')$$

где индекс  $i$  нумерует атомы примеси, а  $\mathbf{R}_i$  есть радиус-вектор  $i$ -го примесного атома. Переход от (7.1') к (7.1) оправдан лишь при достаточно малой концентрации примеси  $N_i$ . Именно, пусть носитель заряда локализован на  $n$ -м уровне около какого-то примесного атома. Тогда энергия его взаимодействия со всеми остальными атомами примеси должна быть (по абсолютной величине) мала по сравнению с величиной  $|E_n|$  (7.4). По порядку величины названная энергия составляет  $e^2 N_i^{1/3} \epsilon^{-1}$  (так как среднее расстояние

между примесными атомами — порядка  $N_i^{-1/3}$ ). Таким образом, должно выполняться неравенство  $\varepsilon E_B n^2 \gg e^2 N_i^{1/3}$  или, в силу (7.3)\*,

$$a_B n^2 \ll N_i^{-1/3}. \quad (7.8)$$

При нарушении неравенства (7.8) взаимодействие между различными атомами примеси приводит к сдвигу и расщеплению уровней (7.4).

Особый интерес представляют явления, происходящие при перекрытии волновых функций носителей, локализованных у различных атомов примеси. Такое перекрытие становится существенным, когда среднее расстояние между атомами оказывается сравнимым с радиусом соответствующей орбиты:

$$N_i^{-1/3} \sim n a_B. \quad (7.9)$$

Рис. 4.9. Примесная зона доноров заштрихована. Двойной штриховкой показаны заполненные уровни в примесной зоне.

возможность непосредственного (без участия зоны проводимости) туннельного перехода электрона (дырки) с одного донора (акцептора) на другой. В результате дискретные примесные уровни «размываются», образуя дозволённую область, лежащую в запрещённой зоне (подобно тому как «размываются» уровни регулярных атомов). Эта область называется *примесной зоной* (рис. 4.9).

Заметим, однако, что аналогия с образованием «собственных» зон кристалла не полна, ибо атомы примеси, вообще говоря, не образуют периодической решетки. По этой причине было бы, например, неправильно приписывать электрону в примесной зоне строго определённый квазиимпульс.

При образовании примесной зоны становится возможным новый механизм электропроводности — по примесной зоне. Для этого примесная зона не должна быть целиком заполненной. Так обстоит дело, если кроме основной примеси, скажем, мелких доноров (концентрация  $N_d$ ), в полупроводнике имеются еще и компенсирующие акцепторы (с концентрацией  $N_a < N_d$ ), уровни энергии которых расположены ниже уровней доноров. Тогда при низких температурах все акцепторы будут заполнены электронами, доноры же частично (с концентрацией  $N_a$ ) окажутся пустыми и станут

\* Фактически неравенство (7.8) может быть несколько ослаблено благодаря учету взаимодействия электронов друг с другом. Этот эффект приводит, в частности, к экранированию взаимодействия электронов с донорами, в результате чего поля последних эффективно ослабляются (см. гл. XVII).

возможными переходы электронов с заполненных доноров на пустые. Так как среднее расстояние между примесными центрами больше расстояния между атомами основной решетки, то и вероятности туннельных переходов между атомами примеси меньше, чем между атомами основной решетки. Поэтому подвижность электронов (дырок) в примесной зоне гораздо меньше, чем в «собственных» зонах.

Существование проводимости по примесной зоне можно обнаружить экспериментально, исследуя электрические свойства кристаллов при низких температурах. На рис. 4.10\*) даны температурные зависимости удельного сопротивления  $\rho$  и постоянной Холла  $R$  в кристаллах дырочного германия с примесями элементов III и V групп. При понижении температуры  $\rho$  и  $R$  сначала увеличиваются по экспоненциальному закону. Однако в дальнейшем зависимость  $\rho$  от  $T$  становится слабой, а  $R$  проходит через максимум (в рассматриваемом примере около 5,5 К).

Такое поведение  $\rho$  и  $R$  объясняется тем, что полная концентрация электронов на донорах ( $N_d$  —  $N_a$ ) (для определенности мы говорим о полупроводнике  $n$ -типа) распределяется между зоной проводимости (концентрацию в которой обозначим через  $n_1$ ) и примесной зоной (концентрация  $n_2$ ). Тогда  $\rho = 1/e (n_1\mu_1 + n_2\mu_2)$ , а постоянная Холла выражается формулой (1.4.5). При сравнительно высоких температурах  $n_1\mu_1 > n_2\mu_2$  и доминирует электропроводность в зоне проводимости. Если бы примесные уровни были дискретными ( $\mu_2|_{T=0} = 0$ ), то при  $T \rightarrow 0$  мы имели бы  $\rho \rightarrow \infty$ ,  $R \rightarrow \infty$ , так как  $n_1 \rightarrow 0$  (пунктир на рис. 4.10). При очень низких температурах, когда  $n_1\mu_1 \ll n_2\mu_2$ , главный вклад в электропроводность вносит примесная зона, концентрация электронов (дырок) в которой практически не зависит от температуры. При  $T \rightarrow 0$  величины  $\rho$  и  $R$  должны стремиться к конечным значениям:

$$\rho = \frac{1}{e(N_d - N_a)\mu_2}, \quad R = -\frac{\mu_2 H}{\mu_2 c e(N_d - N_a)}.$$

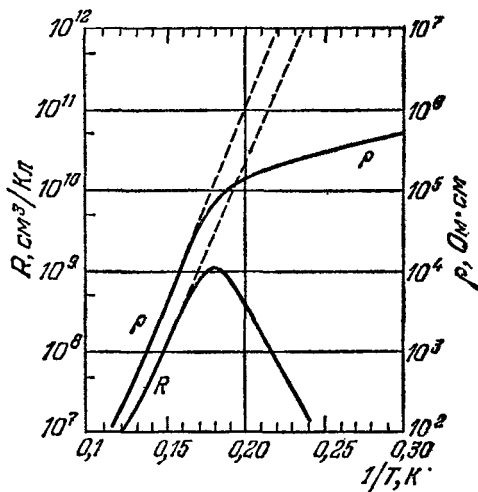


Рис. 4.10. Удельное сопротивление  $\rho$  и постоянная Холла  $R$  при низких температурах в германии.  $N_a = 4,4 \cdot 10^{15} \text{ см}^{-3}$ ,  $N_d = 1,0 \cdot 10^{14} \text{ см}^{-3}$ .

\*) По данным работы *H. Fritzsche*, Phys. Rev. 99, № 2, 406 (1955).

Существование максимума постоянной Холла следует из формулы (I.4.5) при учете условия  $n_1 + n_2 = \text{const.}$  Так, например, если бы подвижности не зависели от температуры, а холловские факторы в обеих зонах были равны:  $\mu_{1H}|\mu_1 = \mu_{2H}|\mu_2$ , то выражение (I.4.5) достигало бы максимума при  $n_1\mu_1 = n_2\mu_2$ .

Примесная зона возникает сначала в результате перекрытия волновых функций возбужденных состояний. По мере увеличения концентрации примеси «размываются» все более и более низколежащие уровни — вплоть до основного. Наконец, при достаточно больших значениях  $N_f$  примесная зона перекрывается с ближайшей «собственной» зоной кристалла — зоной проводимости или валентной. В германии, легированном донорами V группы, такое перекрытие наступает при  $N_f \approx 10^{17} \text{ см}^{-3}$ . При этом деление на примесную и собственную зоны теряет смысл — возникает единая область непрерывного спектра, и вещество ведет себя в электрическом отношении как плохой металл. Полупроводники такого типа называют сильно легированными. Они рассматриваются в гл. XIX.