

### Лекция 3. Полупроводники. Статистика электронов и дырок в полупроводниках.

#### 6.5. Функция плотности состояний для электронов и дырок

Из анализа формулы для концентрации электронов

$$n = \int_0^{\infty} g(E)f(E)dE \quad (6.18)$$

видно, что для вычисления концентрации электронов, помимо функции распределения электронов по состояниям, нужно знать вид функции плотности состояний  $g(E)$ . Если  $dN$  - число состояний в интервале энергий от  $E$  до  $E + dE$ , то по определению

$$g(E) = \frac{dN}{dE}. \quad (6.19)$$

Зная конкретный вид функции  $E(k)$ , можно вычислить объем  $dV_k$ , соответствующий изменению энергии от  $E$  до  $E + dE$  и заключенный между двумя изоэнергетическими поверхностями с энергиями  $E$  и  $E + dE$ . Тогда

$$g(k) = \frac{dN}{d\tau_k}, \quad (6.20)$$

где  $g(k)$  является плотностью состояний в  $k$ -пространстве, которая задается выражением:

$$g(E) = \frac{dN}{dE} = \frac{g(k)dV_k}{dE}. \quad (6.21)$$

Эта величина определяет число состояний в единице объема  $k$ -пространства, где  $d\tau_k$  - элемент объема  $k$ -пространства, равный  $dk_x dk_y dk_z$ .

Рассчитаем величину  $g(k)$ , пользуясь условием квантования состояний в  $k$ -пространстве:  $k_i = \frac{2\pi n_i}{L_i}$ ;  $n_i = 0, \pm 1, \pm 2, \dots$

Число разрешенных состояний, очевидно, будет соответствовать числу возможных значений  $n_i$ . Тогда

$$dn_i = \frac{L_i}{2\pi} dk_i; \quad (6.22)$$

$$dN_i = dn_x dn_y dn_z = \frac{L_x L_y L_z}{8\pi^3} dk_x dk_y dk_z = \frac{V}{8\pi^3} d\tau_k, \quad (6.23)$$

где  $V$  — объем кристалла. Следовательно, число состояний в единице объема кристалла, приходящееся на единичный объем  $k$ -пространства равно

$$g(k) = \frac{1}{V} \frac{dN_i}{d\tau_k} = \frac{1}{8\pi^3}. \quad (6.24)$$

Эту величину следует увеличить в два раза, так как каждой энергии, заданной набором квантовых чисел  $k_x, k_y, k_z$  соответствуют два состояния со значениями спина  $+1/2$  и  $-1/2$ . Мы учтем спиновой множитель 2 позднее. Теперь подсчитаем объем  $k$ -пространства  $dV$  в котором энергия изменяется от  $E$  до  $E + dE$ . Зная закон дисперсии  $E(k)$  в разрешенных зонах можно найти функцию плотности состояния  $g(E)$  для этих зон энергии. Для параболической зоны проводимости с экстремумом в точке  $k = 0$  и скалярной эффективной массы закон

дисперсии имеет вид в малой окрестности  $E_c$

$$E(k) = E_c + \frac{\hbar^2 k^2}{2m_n^*}, \quad (6.25)$$

где  $m_n^*$  - эффективная масса электрона (здесь скалярная величина). Такой тип закона дисперсии наблюдается для зоны проводимости в некоторых соединениях  $A^{III}B^V$  и  $A^{II}B^{VI}$ . Поверхность постоянной энергии в  $k$  - пространстве в этом случае есть сфера с радиусом  $a = \sqrt{\frac{1}{\hbar} 2m_n^*(E - E_c)}$ . Объем этой сферы  $V_k$  можно найти в виде:

$$V_k = \frac{4}{3} \pi a^3 = \frac{4}{3} \pi \left[ \frac{2m_n^*(E - E_c)}{\hbar^2} \right]^{3/2}. \quad (6.26)$$

Исходя из этого, приращение объема  $dV_k$  при изменении энергии от  $E$  до  $E + dE$  равно

$$dV_k = \frac{\partial V_k}{\partial E} dE = 2\pi \left( \frac{2m_n^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} dE. \quad (6.27)$$

Число состояний  $dN$  в объеме  $dV_k$  равно

$$dN = g(k) dV_k = \frac{1}{8\pi^3} 2\pi \left( \frac{2m_n^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} dE; \quad (6.28)$$

$$g(E) = \frac{dN}{dE} = 2\pi \left( \frac{2m_n^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} \quad (6.29)$$

с учетом того, что  $h = 2\pi\hbar$ .

Для простой сферической валентной зоны с экстремумом  $E_v$  в точке  $k = 0$  закон дисперсии имеет вид в малой окрестности  $E_v$ :

$$E(k) = E_v + \frac{\hbar^2 k^2}{2m_p^*}. \quad (6.30)$$

Проводя аналогичные вычисления, получим для простой зоны

$$g(E) = 2\pi \left( \frac{2m_p^*}{\hbar^2} \right)^{3/2} (E_v - E)^{1/2}. \quad (6.31)$$

## 6.6 Вычисление концентрации электронов в зоне проводимости и дырок в валентной зоне

Концентрацию электронов в зоне проводимости можно представить в виде:

$$n = \int_{E_{c\min}}^{E_{c\max}} g(E) f(E) dE. \quad (6.32)$$

При значениях энергии выше энергии Ферми функция  $f(E)$  быстро спадает до нуля. Поэтому, подставляя в уравнение (6.32) вид  $g(E)$ , найденный из условия

(6.29), и учитывая спиновый множитель 2, имеем:

$$n = 2 \int_{E_c}^{\infty} \frac{2\pi \left( \frac{2m^* dn}{h^2} \right)^{3/2} (E - E_c)^{1/2} dE}{e^{\frac{E-E_F}{k_B T}} + 1}. \quad (6.33)$$

Произведем замену переменных

$$\frac{E - E_c}{k_B T} = \varepsilon; \quad \frac{E_F - E_c}{k_B T} = \eta, \quad (6.33a)$$

где  $\varepsilon$  и  $\eta$  - приведенные в единицах  $k_B T$  энергия электрона в зоне проводимости и энергия Ферми, отсчитанные от дна зоны проводимости  $E_c$ . Тогда, произведя подстановку, получим:

$$n = 2 \left( \frac{2\pi m^* dn k_B T}{h^2} \right)^{3/2} \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} = N_c F_{1/2}(\eta), \quad (6.34)$$

где величина  $N_c$ , выражается как

$$N_c = 2 \left( \frac{2\pi m^* dn k_B T}{h^2} \right)^{3/2}, \quad (6.35)$$

а интеграл  $F_{1/2}(\eta)$ , называемый интегралом Ферми порядка 1/2, есть

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1}. \quad (6.36)$$

При проведении различных статистических расчетов для явлений, связанных с участием электронов зоны проводимости, могут получаться интегралы типа (6.34), но с другими значениями степени приведенной энергии  $\varepsilon$  в числителе подынтегрального выражения. Они образуют совокупность интегралов Ферми, задаваемых формулой

$$F_{1/2}(\eta) = \frac{2}{\Gamma(i+1)} \int_0^{\infty} \frac{\varepsilon^i d\varepsilon}{e^{\varepsilon - \eta} + 1}, \quad (6.37)$$

где  $\Gamma(i+1)$  - гамма-функция Эйлера.

Интегралы Ферми не вычисляются точно, однако существуют таблицы значений важнейших интегралов Ферми. Имеются асимптотические выражения для интегралов Ферми порядка 1/2, которые упрощают вид формулы (6.35) для концентрации электронов:

$$F_{1/2}(\eta) = \begin{cases} e^{\varepsilon} & \text{для } -\infty < \eta < -1 \quad (a) \\ \frac{e^{\eta}}{0.27e^{\eta} + 1} & \text{для } -\infty < \eta < 5 \quad (b) \\ \frac{4}{3\sqrt{\pi}} \eta^{3/2} & \text{для } 5 < \eta < \infty \quad (c). \end{cases} \quad (6.38)$$

Принимая во внимание указанные в формулах (6.38) пределы для приведенного уровня Ферми и учитывая условие (6.33a), можно заключить, что

формула (6.38a) пригодна для невырожденных полупроводников, а (6.38c) для сильно вырожденных полупроводников и металлов. Формула (6.38b) представляет собой приближение Эренберга—Блэйкмора и пригодна для расчетов как невырожденных полупроводников, где она дает более высокую точность, чем формула (6.38a), так и при слабом вырождении.

Для концентрации электронов в зоне проводимости в невырожденном полупроводнике в соответствии с формулой (6.34) и приближенной формулой (6.36a) имеем:

$$n = N_c e^{-\frac{E_c - E_F}{k_B T}} \quad (6.39)$$

Формула (6.39) имеет наглядный физический смысл: экспоненциальный множитель соответствует функции распределения по состояниям Максвелла - Больцмана, взятой в точке  $E = E_c$ , тогда величина  $N_c$ , имеющая значение (6.36), представляет собой *эффективное число состояний в зоне проводимости, приведенной к ее дну, т. е. к уровню  $E_c$* .

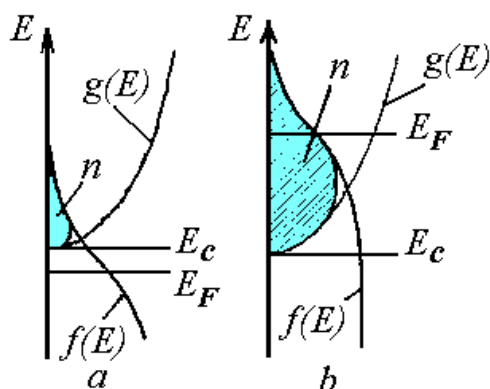
Формулу (6.35) можно представить в более удобном для расчетов виде:

$$N_c = 4.831 \cdot 10^{15} \left( \frac{m_{dn}^*}{m} \right)^{3/2} T^{3/2}. \quad (6.40)$$

Оценка для  $T \approx 300$  К и  $m_{dn}^*$  дает  $N_c \approx 2,5 \cdot 10^{19} \text{ см}^{-3}$ . Число состояний зависит от эффективной массы и температуры. В полупроводниках с небольшой эффективной массой электронов (например, в некоторых узкозонных соединениях  $A^{III}V^V$  и  $A^{II}V^{VI}$ ) эффективное число состояний сравнительно невелико. Тогда заполнение зоны проводимости быстро растет с температурой, и полупроводник может стать вырожденным.

Для сильного вырождения (6.38c) концентрация электронов в зоне проводимости получается в виде:

$$n = \frac{4N_c}{3\sqrt{\pi}} \left( \frac{E_F - E_c}{k_B T} \right) = \frac{8\pi}{3} \left( \frac{2m_{dn}^*}{h^2} \right)^{3/2} (E_F - E_c)^{3/2}. \quad (6.41)$$



**Рис. 6.6.** Концентрация электронов. а) случай невырожденного полупроводника n – типа; б) случай вырожденного полупроводника n – типа

Видно, что при сильном вырождении концентрация перестает зависеть от температуры (слабой зависимостью энергии Ферми от температуры здесь можно пренебречь).

Графическая интерпретация формулы (6.18) для концентрации электронов дана на рисунке 6.6. В случае невырожденного полупроводника уровень Ферми лежит ниже дна зоны проводимости по крайней мере, на  $k_B T$ . При вырождении уровень Ферми поднимается в зону проводимости на несколько  $k_B T$ . Площади заштрихованных областей в зонах проводимости пропорциональны концентрациям электронов. Видно, что полупроводник невырожден, когда

концентрация электронов в зоне проводимости невелика.

Для вычисления концентрации дырок  $p$  в валентной зоне можно записать аналогично формуле (6.18):

$$p = \int_{E_{v \min}}^{E_{v \max}} g(E) f_p(E) dE, \quad (6.42)$$

и, подставляя  $f_p(E) = 1 - f_n(E)$  и  $g(E)$  для валентной зоны и заменяя нижний предел на  $-\infty$ , а также учитывая быстрый спад функции распределения по состояниям  $f_p(E)$ , получим

$$p = 4\pi \left( \frac{2m_{dp}^*}{h^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{e^{\frac{E_F - E}{k_B T}} + 1}. \quad (6.43)$$

Используя подстановки (6.33) и вводя обозначения

$$\varepsilon = \frac{E - E_c}{k_B T}; \quad (6.44)$$

$$\Delta\varepsilon = \frac{E_c - E_v}{k_B T}; \quad (6.4a)$$

$$\varepsilon_p = \frac{E_v - E}{k_B T} = -\varepsilon - \Delta\varepsilon, \quad (6.44b)$$

получим:

$$p = N_v \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{(\varepsilon_p + \Delta\varepsilon + \eta)} + 1} = N_v F_{1/2}(-\Delta\varepsilon - \eta), \quad (6.45)$$

где  $N_v = 2 \left( \frac{2\pi m_{dp}^* k_B T}{h^2} \right)^{3/2}$  - эффективное число состояний в валентной зоне, а

интеграл Ферми порядка одной второй имеет аргумент, определяемый энергией Ферми, отсчитанной от потолка валентной зоны  $E_v$  в единицах  $k_B T$

$$-\Delta\varepsilon - \eta = \frac{-E_c + E_v - E_F + E_c}{k_B T} = \frac{E_v - E_F}{k_B T}. \quad (6.46)$$

Выражение (6.45) для концентрации дырок аналогично выражению (6.34) электронов в зоне проводимости.

Асимптотические выражения для интеграла Ферми  $F_{1/2}(-\Delta\varepsilon - \eta) = F_{1/2}\left(\frac{E_v - E_F}{k_B T}\right)$  аналогичны соответствующим выражениям для интеграла  $F_{1/2}(\eta)$ , приведенным в формуле (6.38)

$$F_{1/2}(-\Delta\varepsilon - \eta) = \begin{cases} e^{-\Delta\varepsilon - \eta} & \text{для } -\infty < (-\Delta\varepsilon - \eta) < -1 \quad (a) \\ \frac{e^{-\Delta\varepsilon - \eta}}{0.27e^{-\Delta\varepsilon - \eta} + 1} & \text{для } -\infty < (-\Delta\varepsilon - \eta) < 5 \quad (b) \\ \frac{4}{3\sqrt{\pi}}(-\Delta\varepsilon - \eta)^{3/2} & \text{для } 5 < (-\Delta\varepsilon - \eta) < \infty \quad (c). \end{cases} \quad (6.47)$$

Таким образом, для невырожденного полупроводника справедлива формула (6.47a), и концентрация дырок в невырожденном полупроводнике имеет вид:

$$p = N_v e^{-\Delta\varepsilon - \eta} = N_v e^{-\frac{E_v - E_F}{k_B T}}, \quad (6.48)$$

из которого наглядно определяется смысл величины как эффективного числа состояний в валентной зоне, приведенной к ее потолку. Для сильно вырожденного полупроводника справедливо выражение:

$$p = \frac{4N_v}{3\sqrt{\pi}} \left( \frac{E_v - E_F}{k_B T} \right)^{3/2} = \frac{8\pi}{3} \left( \frac{2m_{dp}^*}{h^2} \right)^{3/2} (E_v - E_F)^{3/2}. \quad (6.49)$$

Анализ этого выражения приводит нас к выводу, что при сильном вырождении концентрация дырок становится большой и перестает зависеть от температуры, как и концентрация электронов (6.41) в случае сильного вырождения в зоне проводимости. Графическая интерпретация (6.42) аналогична графическому представлению концентрации электронов в зоне проводимости (рис. 6.6).

## 6.7 Концентрация электронов и дырок в собственном полупроводнике

В собственном полупроводнике справедливо соотношение  $n = p = n_i$ . Подставляя выражение для  $n$  по формуле (6.39) и выражение для  $p$  по формуле (6.48), имеем

$$np = n_i^2 = N_v N_c e^{-\frac{E_v - E_c}{k_B T}} = N_v N_c e^{-\frac{E_g}{k_B T}}. \quad (6.50)$$

Значит, для собственной концентрации  $n_i$  в невырожденном полупроводнике справедлива формула

$$n = p = n_i = \sqrt{N_v N_c} e^{-\frac{E_g}{2k_B T}}. \quad (6.51)$$

Концентрация электронов и дырок в собственном полупроводнике оказывается независимой от положения уровня Ферми и растет с температурой по экспоненциальному закону с энергией активации  $E_g/2$ . Это понятно, если учесть, что энергия  $E_g$  затрачивается на создание пары носителей: электрона и дырки, а на каждый из носителей в этой паре приходится энергия в два раза меньшая.

Температурную зависимость уровня Ферми в собственном

полупроводнике можно найти, решая так называемое уравнение электронейтральности, составленное на основании условия электронейтральности единицы объема полупроводника. Условие электронейтральности требует, чтобы концентрации отрицательных и положительных зарядов были равны. Для собственного полупроводника уравнение электронейтральности имеет очень простой вид:

$$n = p. \quad (6.52)$$

Тогда для невырожденного полупроводника можно записать

$$N_v e^{-\frac{E_c - E_F}{k_B T}} = N_c e^{-\frac{E_F - E_v}{k_B T}} \quad (6.53)$$

Логарифмируя обе части равенства (6.53) и вычисляя значение энергии Ферми, получим

$$E_F = \frac{E_c + E_v}{2} + \frac{k_B T}{2} \ln \frac{N_v}{N_c}. \quad (6.54)$$

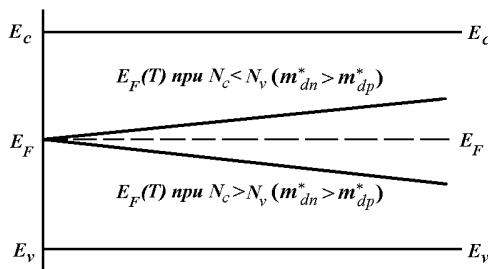
Если отсчитывать энергию Ферми от потолка валентной зоны, то

$$E_F = E_v + \frac{E_g}{2} + \frac{k_B T}{2} \ln \frac{N_v}{N_c}. \quad (6.55)$$

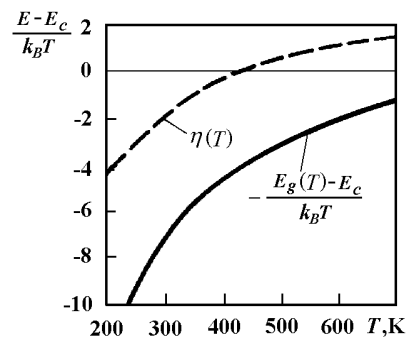
Подставляя в формулу (6.55) значения  $N_c$  и  $N_v$ , получим

$$E_F = E_v + \frac{E_g}{2} + \frac{3}{4} k_B T \ln \frac{m_{dp}^*}{m_{dn}^*}. \quad (6.56)$$

Анализируя формулы (6.55) и (6.56), можно заключить, что при  $T = 0$  уровень Ферми в собственном полупроводнике находится в середине запрещенной зоны, а с ростом температуры движется к той зоне, в которой эффективное число состояний, а значит, и эффективная масса для плотности состояний меньше. График температурной зависимости уровня Ферми в собственном невырожденном полупроводнике представлен на рис. 6.7. Здесь не



**Рис. 6.7.** Зависимость уровня Ферми в собственном полупроводнике



**Рис. 6.8.** Зависимость ширины запрещенной зоны и уровня Ферми (в единицах  $k_B T$ ) от температуры в собственном полупроводнике

учитывается температурная зависимость ширины запрещенной зоны.

В узкозонных полупроводниках с повышением температуры уровень Ферми может приблизиться к дну зоны проводимости (для соединений  $A^{III}B^V$  и  $A^{II}B^{IV}$   $m_{dn}^* < m_{dp}^*$ ) настолько, что формулы (6.55) и (6.56) станут несправедливы, так как полупроводник будет вырожденным. Иллюстрация такого вырождения в

собственном InSb с повышением температуры представлена на рис. 6.8. Здесь за начало отсчета энергии принято дно зоны проводимости. С повышением температуры ширина запрещенной зоны InSb убывает по закону  $E_g \approx (0,26 - 2,7 \cdot 10^{-4} T) \text{ эВ}$ , а уровень Ферми  $E_F$  приближается к дну зоны проводимости (расчет выполнен для  $m_{dp}^* = 10m_{dn}^*$ ), и при температуре  $\sim 400\text{К}$  уровень Ферми подходит к уровню  $E_c$  ближе чем на  $k_B T$ . При этом нужно пользоваться уравнением электронейтральности  $n = p$ , в которое подставлены значения концентраций из формул (6.34) и (6.45). При не сильном вырождении пригодно приближение Эренберга - Блэйкмора для интегралов Ферми (6.38b) и (6.47b).

Такой анализ, проведенный для InSb, показывает, что уже при температурах  $T \geq 440\text{К}$  уровень Ферми лежит внутри зоны проводимости. Быстрое температурное вырождение в области собственной проводимости связано в узкозонных полупроводниках с двумя основными причинами: малой шириной запрещенной зоны и большим различием эффективных масс электронов и дырок, приводящим к более быстрому повышению уровня Ферми с температурой. Заметим, что для гипотетического полупроводника, у которого эффективные массы для плотности состояний электронов и дырок были бы равны, не наблюдалось бы температурной зависимости уровня Ферми, и этот уровень при любой температуре находился бы посередине запрещенной зоны.

## 6.8 Нахождение ширины запрещенной зоны в собственном полупроводнике

Для собственного невырожденного полупроводника известно выражение концентрации электронов и дырок (6.51)  $n = p = n_i = \sqrt{N_v N_c} e^{-\frac{E_g}{2k_B T}}$ . Если определить экспериментально концентрацию электронов (или дырок) в области собственной проводимости, например, из измерений эффекта Холла, то, представив ее температурную зависимость в некоторых специально подобранных осях координат, можно определить ширину запрещенной зоны полупроводника. Действительно:

$$\ln n_i = \ln \sqrt{N_v N_c} - \frac{E_g}{2k_B T}, \quad (6.57)$$

и если учитывать только явную температурную зависимость, то в координатах  $\ln(n_i) = f(1/T)$  уравнение (6.57) описывает прямую линию, отсекающую на оси ординат отрезок  $\ln \sqrt{N_v N_c}$  и имеющую тангенс угла наклона к оси абсцисс,

равный  $\frac{E_g}{2k_B}$ . Такое построение представлено на рис. 6.9. Ширина запрещенной

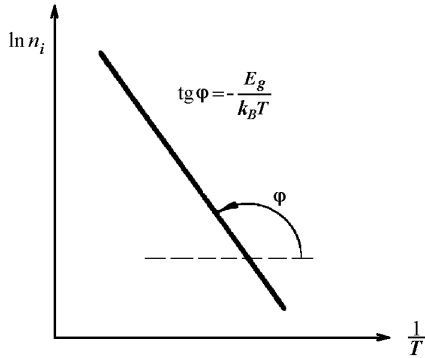
зоны может быть найдена из условия

$$E_g = -2k_B \text{tg}\varphi. \quad (6.58)$$

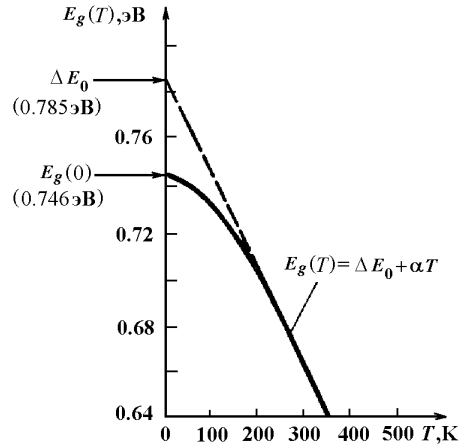
Как было замечено выше, в уравнении (6.57) мы пренебрегли неявной зависимостью концентрации  $n_i$ , от температуры. Учтем теперь более строго полную температурную зависимость  $n_i(T)$  в собственном полупроводнике. Во-



первых, от температуры зависят по закону  $\sim T^{3/2}$  обе величины  $N_c$  и  $N_v$ , во-вторых, от температуры зависит ширина запрещенной зоны  $E_g = f(T)$ . Последняя



**Рис. 6.9.** Температурная зависимость концентрации носителей заряда в собственном полупроводнике от обратной температуры



**Рис. 6.10.** Температурная зависимость  $E_g$  германия (сплошная кривая) и ее линейная экстраполяция (пунктир)

зависимость определяется, как указывалось, сложным комплексом причин, связанных с динамическими и статическими факторами. Учесть строго эти факторы пока невозможно, поэтому такую зависимость ширины запрещенной зоны от температуры находят эмпирически. Для большинства полупроводников ширина запрещенной зоны уменьшается с температурой. Зависимость ширины запрещенной зоны от температуры в некотором интервале температур имеет линейный характер:

$$E_g(T) = \Delta E_0 + \alpha T, \quad (6.59)$$

где  $\Delta E_0$  - экстраполированное по линейному закону значение ширины запрещенной зоны при  $T = 0$ ;  $\alpha$  - линейный коэффициент в зависимости ширины запрещенной зоны от температуры, отрицательный для большинства полупроводников.

Пример такой температурной зависимости (по данным оптических измерений) с ее линейризацией на интервале температур  $T > 200\text{K}$  показан на рисунке 6.10 для германия. Эта температурная зависимость при  $T > 200\text{K}$  может быть описана линейным законом:

$$E_g(T) = (0,785 - 0,0004T). \quad (6.60)$$

Параметры этой линейной зависимости  $\Delta E_0 = 0,785 \text{ эВ}$  и  $\alpha = -4 \cdot 10^{-4} \text{ эВ град}^{-1}$ . Из рис. 6.10 ясно, что значение  $\Delta E_0 = 0,785 \text{ эВ}$  не является шириной запрещенной зоны германия ни при какой температуре, а есть лишь отрезок, отсекаемый на оси ординат при продолженной до значений  $T = 0$  линейной экстраполяции  $E_g$ .

Предположим теперь, что при анализе  $n(T)$  температурной зависимости концентрации в некотором температурном интервале справедлив линейный закон температурного изменения запрещенной зоны (6.59). Тогда, учитывая, что

$$\sqrt{N_v N_c} = 4,831 \cdot 10^{15} \left( \frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} T^{3/2},$$

для  $n(T)$  можно записать:

$$n(T) = 4,831 \cdot 10^{15} \left( \frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} T^{3/2} e^{-\frac{\Delta E_0 - \alpha T}{2k_B T}}. \quad (6.61)$$

Логарифмируя обе части этого выражения, поделенного на  $T^{3/2}$ , имеем

$$\ln \frac{n}{T^{3/2}} = \ln \left[ 4,831 \cdot 10^{15} \left( \frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} \right] - \frac{\Delta E_0}{2k_B} \frac{1}{T} - \frac{\alpha}{2k_B}. \quad (6.62)$$

Теперь отрезок, отсекаемый по оси ординат прямой линией  $\ln(n) = f(1/T)$ , будет равен:

$$[\ln n]_0 = \ln \left[ 4,831 \cdot 10^{15} \left( \frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} \right] - \frac{\alpha}{2k_B}, \quad (6.63)$$

а тангенс угла наклона этой прямой к оси абсцисс составляет:

$$\operatorname{tg} \varphi = -\Delta E_0 / 2k_B, \quad (6.64)$$

тогда  $\Delta E_0 = -2k_B \operatorname{tg} \varphi$ , а истинную ширину запрещенной зоны  $E_g(T)$  для некоторой температуры  $T$  (не меньше 200 К) можно найти, если с помощью соотношения (6.63) при известных значениях  $m_{dn}^*$  и  $m_{dp}^*$  рассчитать величину коэффициента  $\alpha$  в линейной зависимости изменения ширины запрещенной зоны с температурой и подставить в формулу (6.59).

Проведенный анализ показывает, почему при расчете ширины запрещенной зоны по упрощенной схеме, разобранный в начале этого параграфа, можно получить значения  $E_g$ , отличающиеся от истинных. Следует заметить, что такой упрощенный метод обработки данных температурной зависимости концентрации носителей заряда в области собственной электропроводности дает меньшую ошибку в полупроводниках со сравнительно большой шириной запрещенной зоны.

## 6.9 Статистика примесных состояний в полупроводниках

Введение донорных и акцепторных примесей в полупроводник изменяет концентрацию носителей заряда, а значит, и положение уровня Ферми. При анализе уравнения электронейтральности в примесном полупроводнике следует учесть не только положительный и отрицательный заряд, связанный с носителями заряда в зонах, как это делалось для собственного полупроводника, но и положительный и отрицательный заряд, обусловленный и ионизированными донорными и акцепторными атомами. Тогда уравнение электронейтральности в примесном полупроводнике, в который введены доноры в концентрации  $N_d$  и акцепторы в концентрации  $N_a$ , будет иметь вид

$$n + N_a^- = p + N_d^+, \quad (6.65)$$

где  $N_a^-$  и  $N_d^+$  - концентрация ионизированных акцепторов и доноров соответственно.

Так как при ионизации атомы акцепторов принимают электроны, то

$$N_a^- = n_a, \quad (6.66)$$

где  $n_a$  - концентрация электронов, занимающих акцепторные состояния в запрещенной зоне. Аналогично для ионизированных доноров имеем:

$$N_d^+ = p_d, \quad (6.67)$$

где  $p_d$  - концентрация дырок, занимающих донорные состояния, так как формально переход электронов с донорных уровней в зону проводимости при ионизации атомов доноров эквивалентен переходу дырок из зоны проводимости на донорный уровень. Исходя из этого, уравнение (6.65) может быть записано в виде:

$$n + n_a = p + p_d. \quad (6.68)$$

Концентрацию неионизированных атомов примеси можно представить как:

$$N_d^0 = n_d; \quad (6.69)$$

$$N_a^0 = p_a, \quad (6.70)$$

где  $N_d^0$  и  $N_a^0$  - концентрации неионизированных атомов доноров и акцепторов;  $n_d$  и  $p_a$  - концентрации электронов на донорных уровнях, а дырок на акцепторных соответственно. Очевидно, что

$$N_d = N_d^0 + N_d^+; \quad (6.71)$$

$$N_a = N_a^- + N_a^0 \quad (6.72)$$

и тогда

$$n_d + p_d = N_d; \quad (6.73)$$

$$n_a + p_a = N_a. \quad (6.74)$$

Уравнение электронейтральности теперь можно записать в другом виде:

$$n + n_d - p - p_a = N_d - N_a. \quad (6.75)$$

Таким образом, для решения уравнения электронейтральности и нахождения положения уровня Ферми, а значит, и концентраций электронов и дырок в примесном полупроводнике требуется уметь находить концентрации  $n_d$  и  $p_a$  или  $p_d$  и  $n_a$  электронов и дырок, занимающих примесные состояния. Статистика примесных состояний должна отличаться от статистики состояний в разрешенных зонах энергии, так как для примесных состояний неприменим принцип Паули. Действительно, по принципу Паули на каждом энергетическом уровне может находиться два электрона (или дырки) с противоположно направленными спинами, тогда как примесный уровень для однократно ионизируемой примеси может быть занят только одним электроном (или только одной дыркой), потому что, принимая второй электрон (или вторую дырку), он становится двукратно ионизированным и изменяет энергетическое положение в запрещенной зоне. По этой причине функция распределения по состояниям Ферми - Дирака оказывается несправедливой для примесных состояний.

Для нахождения функции распределения по примесным состояниям

следует учесть, что система электронов или дырок на примесных состояниях - это система с переменным числом частиц. Как это следует из статистической физики [11], функция распределения электронов по примесным состояниям может быть записана в виде:

$$f_n = \frac{1}{\frac{1}{g} e^{\frac{E_i - E_F}{k_B T}} + 1}, \quad (6.76)$$

где  $E_i$  - энергетическое положение примесного уровня;  $g$  - фактор вырождения примесного уровня.

Обсудим физический смысл величины  $g$ . Примесный уровень может принять только один электрон (или дырку), но принятый электрон (или дырка) может занять состояние со спином  $+1/2$  или состояние со спином  $-1/2$ . Таким образом, состояние оказывается двукратно вырожденным при учете спинового вырождения. Тогда  $g$  - фактор для электронов на донорных состояниях или дырок на акцепторных состояниях должен быть равен двум;  $g$  - фактор для дырок на донорных состояниях или электронов на акцепторных состояниях можно вычислить, имея в виду, что любое состояние  $E$  может быть занято либо электроном, либо дыркой  $f_n + f_p = 1$ , отсюда:

$$f_n^d + f_p^d = 1; \quad (6.77)$$

$$f_n^a + f_p^a = 1 \quad (6.78)$$

(индексы  $a$  и  $d$  относятся к акцепторным и донорным состояниям),  $g$  - факторы для дырок на донорных уровнях и электронов на акцепторных уровнях будут равны  $1/2$ . В литературе термин «фактор спинового вырождения» применяют или к величине  $g$  или к величине  $g^{-1}$  в зависимости от того, какая из них больше единицы. Следует отметить, что, если учесть другие механизмы вырождения примесных состояний, кроме спинового, значения  $g$  - фактора могут получиться больше двух. Например, если примесное состояние отщепляется от разрешенной зоны сложной структуры, то вырождение состояний в экстремальных точках разрешенной зоны переносится и на примесное состояние. Например, акцепторные состояния, полученные при введении в полупроводник IV группы (Ge, Si) атомов примеси III группы, имеют фактор вырождения, равный 4, так как на спиновое вырождение накладывается двукратное вырождение экстремума валентной зоны при  $\mathbf{k} = 0$ . Для многовалентных примесных центров, которые при последовательной ионизации отдают или принимают электроны в каждом зарядовом состоянии, дающем свой уровень в запрещенной зоне, имеются свои значения факторов вырождения.

Для подсчета величин  $n_d$ ,  $p_a$ ,  $p_d$  и  $n_a$  нужно, помимо функции распределения, знать функцию плотности примесных состояний. Число состояний с энергией  $E_d$  в единице объема кристалла равно концентрации атомов доноров  $N_d$ , а для состояний  $E_a$  - концентрации  $N_a$ . Тогда:

$$n_d = N_d f_n^d = \frac{N_d}{\frac{1}{2} e^{\frac{E_d - E_F}{k_B T}} + 1}; \quad (6.79)$$

$$p_a = N_a f_p^a = \frac{N_a}{\frac{1}{2} e^{\frac{E_F - E_a}{k_B T}} + 1}; \quad (6.80)$$

$$p_d = N_d - n_d = \frac{N_d}{2e^{\frac{E_F - E_d}{k_B T}} + 1}; \quad (6.81)$$

$$n_a = N_a - p_a = \frac{N_a}{2e^{\frac{E_a - E_F}{k_B T}} + 1}. \quad (6.82)$$

Подставляя в уравнение электронейтральности (6.75) значения  $n$ ,  $p$ ,  $n_d$ ,  $p_a$ , получим общее уравнение для нахождения уровня Ферми в примесном полупроводнике:

$$N_c F_{1/2} \left( \frac{E_c - E_F}{k_B T} \right) + \frac{N_d}{2e^{\frac{E_d - E_F}{k_B T}} + 1} - N_v F_{1/2} \left( \frac{E_F - E_v}{k_B T} \right) + \frac{N_a}{2e^{\frac{E_a - E_F}{k_B T}} + 1} = N_d - N_a. \quad (6.83)$$

Для невырожденного полупроводника можно заменить интегралы Ферми асимптотическими выражениями:

$$F_{1/2} \left( -\frac{E_c - E_F}{k_B T} \right) = e^{-\frac{E_c - E_F}{k_B T}}; \quad (6.84)$$

$$F_{1/2} \left( -\frac{E_F - E_v}{k_B T} \right) = e^{-\frac{E_F - E_v}{k_B T}}. \quad (6.85)$$

Математический анализ уравнения (6.83) в этом случае существенно упрощается, так как интегралы Ферми переходят в экспоненты согласно формулам (6.38a) и (6.47a), но все-таки остаются достаточно сложными из-за двух типов примеси. Поэтому вначале разберем решение уравнения электронейтральности для невырожденного примесного полупроводника с примесью одного типа.

## 6.10 Концентрация электронов в полупроводниках с одним типом примеси

Уравнение электронейтральности (6.75) в этом случае имеет вид:

$$n + n_d - p = N_d. \quad (6.86)$$

Число членов в формуле (6.86) можно уменьшить, переписав уравнение в виде:

$$n = p - p_d. \quad (6.87)$$

Последнее уравнение имеет наглядный физический смысл. Электроны в зоне проводимости появляются за счет собственной ионизации (одновременно в

валентной зоне появляется равная концентрация дырок) и примесной ионизации (одновременно на донорных уровнях остается равная концентрация связанных дырок). Ясно, что при постепенном повышении температуры от  $T = 0$  сначала будут преимущественно ионизироваться атомы примеси, а при более высоких температурах начнется собственная ионизация. Так как  $\Delta E_d \ll \Delta E$ , то при низких температурах собственной ионизацией можно пренебречь. Тогда уравнение электронейтральности можно представить в виде:

$$n = p_d. \quad (6.88)$$

Если полупроводник невырожденный, то имеем:

$$n = N_c e^{-\frac{E_c - E_F}{k_B T}} = \frac{N_d}{g e^{\frac{E_F - E_a}{k_B T}} + 1}. \quad (6.89)$$

Вводя обычные безразмерные величины энергии и обозначая дополнительно величину  $\frac{E_c - E_d}{k_B T} = \varepsilon_d$ , получим

$$n = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \quad (6.90)$$

Заменяя в правой части  $e^\eta$  на  $n/N_c$  согласно уравнениям (6.38а) (3.39), получим квадратное уравнение относительно  $n$ :

$$n \left[ g \frac{n}{N_c} e^{\varepsilon_d} + 1 \right] = N_d. \quad (6.91)$$

Его решение имеет вид:

$$n = \frac{1}{2} g^{-1} N_c e^{-\varepsilon_d} \left\{ \left[ 1 + 4gN_d N_c^{-1} e^{\varepsilon_d} \right]^{1/2} - 1 \right\}. \quad (6.92)$$

Уравнение (6.92) с учетом формулы (6.90) можно представить как

$$n = \frac{2N_d}{1 + \sqrt{1 + \frac{4gN_d}{N_c} e^{\varepsilon_d}}}. \quad (6.93)$$

Если рассматривать область низких температур, когда выполняется условие

$$\frac{4gN_d}{N_c} e^{\varepsilon_d} \gg 1,$$

то

$$n = \sqrt{\frac{N_c N_d}{g}} e^{-\frac{\Delta E_d}{2k_B T}}. \quad (6.94)$$

Условие (6.94) означает, что при низких температурах количество электронов в зоне проводимости очень мало по сравнению с  $N_c$ . Из выражения (6.94) видно, что в области примесной ионизации энергия термической активации процесса ионизации равна  $\Delta E_d/2$ . Следовательно, уровень Ферми при низких температурах находится посередине интервала энергий между  $E_c$  и  $E_d$ . Действительно, подставляя выражение (6.94) в формулу (6.37), имеем:

$$N_c e^{\frac{E_c - E_F}{k_B T}} = \sqrt{\frac{N_c N_d}{g}} e^{\frac{E_c - E_d}{2k_B T}}; \quad (6.95)$$

$$e^{\frac{E_F}{k_B T}} = \sqrt{\frac{N_d}{g N_c}} e^{\frac{E_c + E_d}{2k_B T}}; \quad (6.96)$$

$$E_F = \frac{E_c + E_d}{2} + \frac{k_B T}{2} \ln \frac{N_d}{g N_c}; \quad (6.97)$$

$$E_F = E_c - \frac{\Delta E_d}{2} + \frac{k_B T}{2} \ln \frac{N_d}{g N_c}. \quad (6.98)$$

Таким образом, уровень Ферми при  $T = 0$  совпадает с серединой интервала между  $E_c$  и  $E_d$ , а с повышением температуры сначала приближается к  $E_c$  (пока  $N_c < N_d/g$ ), проходит через максимум и удаляется от  $E_c$  (когда  $N_c > N_d/g$ ).

Если рассмотреть в формуле (6.93) интервал более высоких температур, когда выполняется критерий

$$\frac{4gN_d}{N_c} e^{\varepsilon_d} \ll 1, \quad (6.99)$$

то

$$n = N_d. \quad (6.100)$$

В этом температурном интервале концентрация электронов в зоне проводимости не зависит от температуры и равна  $N_d$ , что соответствует области истощения донорной примеси. В этой области температурную зависимость уровня Ферми можно установить, подставляя формулу (6.100) в уравнение (6.37);

$$E_F - E_c = k_B T \ln \frac{n}{N_c} = k_B T \ln \frac{N_d}{N_c}; \quad (6.101)$$

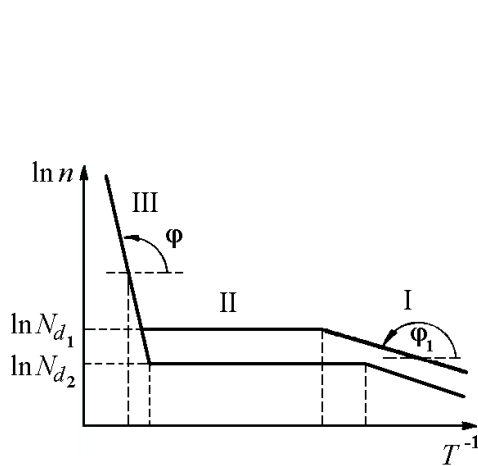
$$E_F = E_c - k_B T \ln \frac{N_d}{N_c}. \quad (6.102)$$

Таким образом, уровень Ферми с ростом температуры в области истощения донорной примеси удаляется от  $E_c$ .

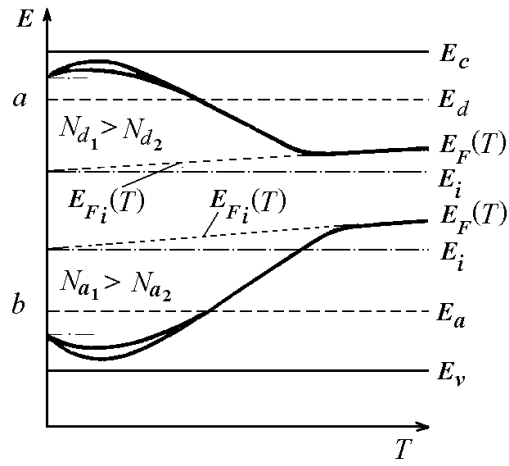
Выполнение критерия (6.99) приводит к появлению области истощения донорной примеси и постоянству концентрации электронов  $n = N_d$ . Однако при высоких концентрациях примеси  $N_d$  или относительно больших энергиях ионизации  $\Delta E_d$  этот критерий выполняется только при довольно высоких температурах, когда уже интенсивно идет процесс собственной ионизации. Тогда область истощения донорной примеси не заметна на кривой температурной зависимости концентрации электронов и отсутствует температурный интервал с постоянной концентрацией электронов.

Рассматривая уравнение электронейтральности (6.87) в области низких температур, мы пренебрегли концентрацией дырок в валентной зоне по сравнению с величиной  $p_d$ . С повышением температуры, очевидно, собственная ионизация преобладает над примесной, т. е.  $p \gg p_d$ . Тогда уравнение (6.87) переходит в уже известное нам уравнение (6.52)  $n = p$  для собственной проводимости, из которого следует температурная зависимость  $n(T)$  в виде (6.51) и температурная зависимость уровня Ферми в виде (6.54).

Объединяя результаты проведенного анализа, можно проиллюстрировать температурную зависимость концентрации в примесном полупроводнике с одним типом донорной примеси кривыми (рис. 6.11) для двух различных концентраций донорной примеси  $N_{d1} > N_{d2}$ .



**Рис. 6.11.** Концентрация электронов в зоне проводимости от обратной температуры для двух концентраций донорной примеси. I- участок возрастания примесной проводимости; II-участок истощения примеси; III-участок собственной проводимости



**Рис. 6.12.** Зависимость уровня Ферми от температуры в полупроводнике с одним типом донорной примеси (a) и одним типом акцепторной примеси (b) для двух разных концентраций. Пунктир – уровень Ферми в нелегированном полупроводнике

Из экспериментальных данных по температурной зависимости концентрации электронов  $n(T)$  в области примесной электропроводности можно найти энергию ионизации донорной примеси по наклону кривой  $n(T)$  в области I.

Действительно, в области I

$$\ln n = \ln \sqrt{\frac{N_c N_d}{2}} - \frac{\Delta E_d}{2k_B T}; \quad (6.103)$$

$$\Delta E_d = -2k_B \text{tg} \varphi_1. \quad (6.104)$$

Это определение  $\Delta E_d$  может привести к некоторой ошибке, если  $\Delta E_d > 2k_B/T$ . В этом случае необходимо учитывать неявную температурную зависимость, связанную с  $N_c$  и тогда экспериментальные данные лучше построить в осях  $\ln n/T^{3/4} = f(1/T)$ . Тангенс угла наклона полученной прямой к оси абсцисс точнее определяет  $\Delta E_d$  при подстановке в формулу (6.104), чем при анализе температурной зависимости концентрации в осях  $\ln n = f(1/T)$ .

Температурная зависимость уровня Ферми для полупроводника с донорной примесью одного типа показана на рис. 6.12. С повышением концентрации доноров уровень Ферми располагается ближе к зоне проводимости, что объясняется увеличением концентрации электронов и ростом заполнения зоны проводимости. Из рис. 6.11 видно, что с повышением концентрации примеси область истощения примеси начинается при более высоких температурах.



Температурный интервал области истощения при этом уменьшается и она при больших концентрациях примеси может не наблюдаться и тогда область примесной ионизации будет непрерывно переходить в область собственной ионизации. Наклон участка I зависимости  $\ln n = f(1/T)$  с ростом концентрации примеси не изменяется, так как он, если концентрация примеси невелика, определяется той же величиной  $\Delta E_d$ . При очень больших концентрациях примеси величина  $\Delta E_d$  может уменьшаться, как это показано экспериментально для ряда полупроводников. Все приведенные рассуждения относительно температурной зависимости концентрации и уровня Ферми в донорном полупроводнике можно применить для полупроводника с акцепторной примесью одного типа.

Очевидно, что вид зависимости  $\ln p = f(1/T)$  аналогичен приведенному на рис. 6.11 с заменой  $\Delta E_d$  на  $\Delta E_a$  и  $N_d$  на  $N_a$  соответственно. Температурная зависимость уровня Ферми в акцепторном полупроводнике с одним типом акцепторов не будет выглядеть симметричной относительно уровня  $E_i$  вследствие одинакового поведения кривой  $E_F(T)$  для донорного и акцепторного полупроводника в области собственной электропроводности. Зависимость  $E_F(T)$  для двух значений концентрации  $N_{a1} > N_{a2}$  показана на рис. 6.12b.

## 6.11 Температурная зависимость концентрации носителей заряда в реальных полупроводниках

Рассмотренный выше случай полупроводника с одним типом примеси (или только донорной, или только акцепторной) весьма идеализирован по отношению к реальным полупроводникам, в которых присутствуют и донорные, и акцепторные примеси одновременно. Специальные методы очистки могут уменьшить концентрацию нежелательных примесей, однако исключить эти примеси полностью невозможно.

При совместном присутствии донорных и акцепторных примесей в полупроводнике возникает явление компенсации примесей, изменяющее его поведение в области примесной электропроводности по сравнению со случаем одного типа примеси. Явление взаимной компенсации доноров и акцепторов состоит в следующем. Если энергетические состояния акцепторов  $E_a$  лежат в запрещенной зоне ниже энергетических состояний доноров  $E_d$  (такая ситуация чаще всего и реализуется для легирующих примесей в полупроводниках), то для ионизации донорных атомов не требуется сообщить им энергию  $\Delta E_d = E_c - E_d$ . Так как ниже заполненного электронами уровня  $E_d$  нейтральных доноров лежит свободный от электронов уровень  $E_a$  нейтральных акцепторов, то уже при  $T = 0$  электроны переходят с донорных уровней на нижележащие акцепторные. Очевидно, что если  $N_d > N_a$ , то все акцепторы ионизируются полностью, а доноры - лишь частично. Концентрация неионизированных доноров  $N_d' = N_d - N_a$  играет роль эффективной концентрации доноров, в процессе термической примесной ионизации только эти  $N_d'$  доноров могут поставлять электроны в зону проводимости. Акцепторы же, будучи уже полностью ионизированными, не могут отдавать дырки в валентную зону. В таком случае говорят о полной компенсации акцепторов и частичной компенсации доноров, и такой полупроводник является частично компенсированным донорным полупроводником. Если  $N_a > N_d$ , то происходит полная компенсация доноров и

частичная компенсация акцепторов, и полупроводник становится *частично компенсированным акцепторным полупроводником*. При  $N_d = N_a$  происходит полная компенсация как доноров, так и акцепторов, и такой полупроводник будет *полностью компенсированным полупроводником*. Термическая ионизация примесей в полностью компенсированном полупроводнике при сообщении атомам примеси энергии  $\Delta E_d$  или  $\Delta E_a$  уже невозможна.

Концентрация свободных носителей заряда, т. е. электронов и дырок в зоне проводимости и валентной зоне соответственно, может возрасть за счет собственной ионизации или путем термических перебросов электронов с заполненных уровней акцепторов в зону проводимости или из валентной зоны на свободные уровни доноров. Последние два процесса по энергии активации близки к процессу собственной ионизации, так как для легирующих примесей  $\Delta E_{a,d} \ll E_g$ , поэтому по величине концентрации свободных носителей заряда полностью компенсированные полупроводники близки к собственным при тех же температурах. Заметим, однако, что подвижности носителей заряда в полностью компенсированном полупроводнике оказываются значительно ниже, чем в собственном, за счет дополнительного рассеяния носителей на заряженных ионах доноров и акцепторов. Следовательно, более высокое удельное сопротивление полупроводника еще не свидетельствует о более высокой степени очистки его от примесей.

Рассмотрим явление компенсации примесей в полупроводниках на наиболее простой модели полупроводника с одним типом доноров с концентрацией  $N_d$  и энергетическим положением уровней в запрещенной зоне  $E_d$  и одним типом акцепторов с концентрацией  $N_a$  и энергетическим положением уровней  $E_a$ . Запишем уравнение электронейтральности, воспользовавшись формой его записи в виде (6.75). Если предположить, что  $N_d > N_a$  (случай частично компенсированного донорного полупроводника), то в уравнении (6.75) член  $p_a = 0$ , так как акцепторы полностью компенсированы и нейтральных атомов акцепторов нет. Предположим также, что рассматривается температурный диапазон, где собственная ионизация мала, т. е. участок примесной проводимости. В этих условиях  $p < n$ ,  $n_d$ ,  $N_d$ ,  $N_a$  и можно пренебречь концентрацией дырок  $p$  и тогда уравнение (6.74) приобретает вид

$$n + n_d = N_d - N_a \quad (6.105)$$

а, так как  $N_d - n_d = p_d$ , то

$$n + N_a = p_d \quad (6.106)$$

Проанализируем это уравнение для невырожденного полупроводника, когда для концентрации электронов  $n$  можно воспользоваться выражением (6.39). Значение  $p_d$  подставим из формулы (6.81), заменяя множитель 2 перед экспонентой в знаменателе более общим обозначением  $g$ . Получим

$$N_c e^{-\frac{E_c - E_F}{k_B T}} + N_a = \frac{N_d}{g e^{\frac{E_F - E_d}{k_B T}} + 1} \quad (6.107)$$

Вводя безразмерные энергии  $\eta = \frac{E_c - E_F}{k_B T}$ , получим

$$\eta = \ln \left\{ \frac{2g^{-1}(N_d - N_a)}{g^{-1}N_c + N_a e^{\varepsilon_d}} + \left[ \frac{g^{-1}N_c + N_a e^{\varepsilon_d}}{g^{-1}N_c + N_a e^{\varepsilon_d}} + 4g^{-1}N_c(N_d - N_a)e^{\varepsilon_d} \right]^{1/2} \right\}. \quad (6.108)$$

Приведенное выражение сложно, и мы будем рассматривать его в сравнении с более частными выражениями для определенных температурных диапазонов. Величину  $\eta$  можно найти из выражения  $\eta = \ln n/N_c$  на основании формулы (6.39). Решая уравнение (6.107) относительно концентрации электронов  $n$ , получим

$$n = 2(N_d - N_a) / \left\{ \left[ 1 + (gN_a/N_c)e^{\varepsilon_d} \right] + \left\{ \left[ 1 + (gN_a/N_c)e^{\varepsilon_d} \right]^2 + (4g/N_c)(N_d - N_a)e^{\varepsilon_d} \right\}^{1/2} \right\}. \quad (6.109)$$

Если концентрация компенсирующих центров  $N_a$  стремится к нулю, то формула (6.109) переходит в выражение (6.93) для концентрации электронов в чисто донорном полупроводнике. Сравним теперь выражения (6.93) и (6.109). При высоких температурах, когда все примеси ионизированы, согласно (6.93)  $n = N_d$ , а по выражению (6.109)  $n = N_d - N_a = N_d'$ , что соответствует истощению донорной примеси в первом случае и истощению эффективной ионизации доноров во втором. По мере понижения температуры произведения  $(4g/N_c)(N_d - N_a)e^{\varepsilon_d}$  и  $(gN_a/N_c)e^{\varepsilon_d}$  в формуле (6.109) сначала сравнимы с единицей, а затем становятся больше единицы. Если степень компенсации очень низкая, т. е.  $N_a \ll N_d$ , то можно найти такую область температур, где  $N_a \ll n \ll N_d$ , а значит, будет выполняться соотношение

$$(gN_a/N_c)e^{\varepsilon_d} \ll 1 \ll [4g(N_d - N_a)/N_c]e^{\varepsilon_d}. \quad (6.110)$$

Температурная зависимость концентрации электронов в этом диапазоне температур определяется соотношением

$$n = \left[ g^{-1}N_c(N_d - N_a)e^{-\varepsilon_d} \right]^{1/2} \quad \text{для} \quad N_a \ll n \ll N_d \quad (6.111)$$

или, переходя от безразмерных обозначений энергии к обычным, получим

$$n = \sqrt{\frac{N_c(N_d - N_a)}{g} e^{-\frac{\Delta E_d}{2k_B T}}} \quad \text{для} \quad N_a \ll n \ll N_d \quad (6.112)$$

что аналогично выражению (6.94) для концентрации электронов в чисто донорном полупроводнике при замене величины  $N_d$  на  $N_d - N_a$ .

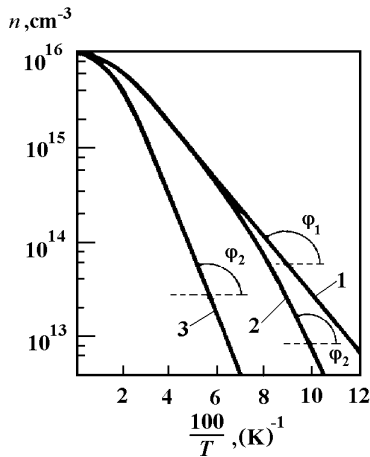
При дальнейшем понижении температуры величина  $n$  должна уменьшаться до значений, близких к  $N_a$ , а затем стать еще меньше. Тогда в знаменателе выражения (6.109) начнут преобладать оба члена вида  $(gN_a/N_c)e^{\varepsilon_d}$  и при достаточно низких температурах

$$n = \left[ g^{-1} N_c (N_d - N_a) / N_a \right] e^{-\varepsilon_d} \quad (6.113)$$

или в обычных обозначениях энергии

$$n = \frac{N_c (N_d - N_a)}{g N_a} e^{-\frac{\Delta E_d}{2k_B T}}; \quad n \ll N_a \ll N_d \quad (6.114)$$

Сравнивая выражения (6.114) и (6.112), видим, что при низких температурах для компенсированного полупроводника величина  $\ln n$  меняется в зависимости от  $1/T$  в два раза быстрее, так как для области, определяемой формулой (6.114),

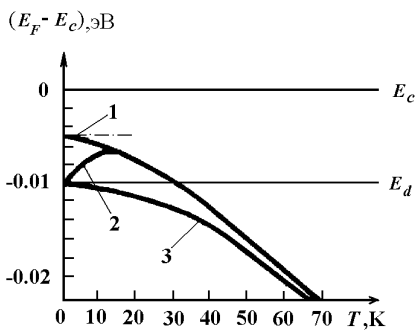


**Рис. 6.13.** Температурная зависимость концентрации электронов в донорном полупроводнике.  $N_d - N_a = 10^{16} \text{ см}^{-3}$ ,

$\Delta E_d = 0.01 \text{ эВ}$ ;  $m_c^* = 0.25m$ ;  $y = 2$ ; 1-  $N_a = 0$ ; 2-  $N_a = 10^{14} \text{ см}^{-3}$ ; 3-  $N_a = 10^{16} \text{ см}^{-3}$

тангенс угла наклона линейного участка  $\ln n = f(1/T)$  к оси абсцисс равен  $-\Delta E_d/k_B$ , а по формуле (6.112)  $-\Delta E_d/2k_B$ .

Переходная область от выражения (6.114) к формуле (6.112) соответствует достижению соотношения  $n \approx N_a$ . При более высокой степени компенсации, т. е.



**Рис. 6.14** Зависимость уровня Ферми от температуры в донорном полупроводнике с различными степенями компенсации доноров

при более высоких значениях  $N_a$ , область температур, описываемая зависимостью (6.112), исчезает. Теперь от наклона, определяемого величиной  $\Delta E_d/2k_B$ , мы перейдем к участку истощения эффективной концентрации донорной примеси, где  $n = N_d - N_a$  не будет зависеть от температуры. На рис. 6.13 представлена температурная зависимость концентрации электронов для различной компенсации донорных примесей при сохранении одинаковой эффективной концентрации доноров  $N_d' = N_d - N_a$ . На рис. 6.14 даны температурные зависимости уровня Ферми для области примесной электропроводности при концентрациях, соответствующих рис. 6.13. Кривые 1 на рис.

6.13 и 6.14 соответствуют чисто донорному полупроводнику, и поэтому наклон в области примесной ионизации на рис. 6.13 определяется уравнением (6.104). Кривые 2 и 3 соответствуют частично компенсированным полупроводникам. Для образца, характеризуемого кривой 2, степень компенсации невысока и в области примесной ионизации на рис. 6.13 можно выделить два участка с различными наклонами, где  $\text{tg}\varphi_2 = -\Delta E_d/k_B$  при низких температурах и  $\text{tg}\varphi_1 = -\Delta E_d/2k_B$  при некотором повышении температуры. При достаточно

высокой степени компенсации (кривая 3) область с наклоном, определяемым  $-\Delta E_d/2k_B$ , практически исчезает. Поведение кривых 2 и 3 на рис. 6.14 в области низких температур определяется выражением

$$E_F = E_c - \Delta E_d - \ln \frac{gN_a}{N_d - N_a}, \quad (6.115)$$

которое получается подстановкой уравнения (6.114) в формулу (6.37).

При анализе формулы (6.115) видно, что при  $N_a > \frac{N_d}{1+g}$  уровень Ферми

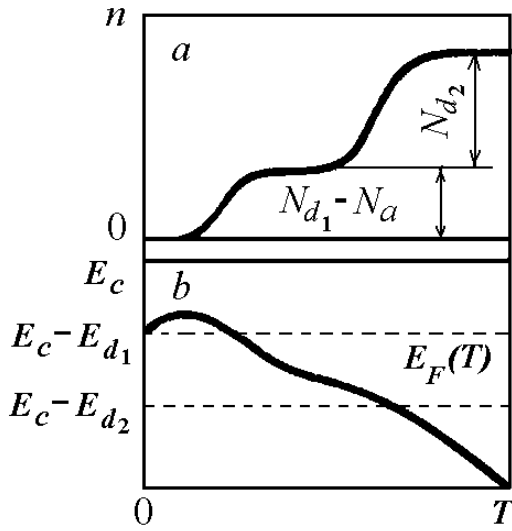
понижается от  $E_d$  (кривая 3), а при  $N_a < \frac{N_d}{1+g}$  уровень Ферми повышается от

величины  $E_d$  (кривая 2). Заметно, что даже весьма малая добавка компенсирующей примеси заставляет уровень Ферми при  $T = 0$  переместиться из точки  $E_c - \Delta E_d/2$  в точку  $E_c - \Delta E_d$ . Поведение концентрации носителей в частично компенсированном полупроводнике в области собственной проводимости, очевидно, аналогично поведению чисто донорного полупроводника с одним типом доноров. Весь анализ, проведенный нами для частично компенсированного донорного полупроводника, можно совершенно аналогично провести для частично компенсированного акцепторного полупроводника с той разницей, что величины  $E_d$  и  $N_d$  перейдут в  $E_a$  и  $N_a$ , а величины  $(N_d - N_a)$  и  $N_d$  перейдут в  $(N_a - N_d)$  и  $N_a$  соответственно. Кроме этого, при вычислениях необходимо учесть другое значение  $g$  - фактора для распределения электронов по акцепторным состояниям в сравнении с распределением по донорным уровням.

При анализе температурной зависимости концентрации реальных примесных полупроводников помимо явления возможной компенсации примесей следует учитывать еще некоторые особенности. Рассматривая энергетическую диаграмму примесного полупроводника, мы считали, что доноры дают только один уровень в запрещенной зоне  $E_d$ , а акцепторы -  $E_a$ . Однако для примесных состояний, помимо основного состояния, которое мы характеризовали значением энергии  $E_d(E_a)$ , существуют возбужденные состояния. Например, для мелких водородоподобных примесных состояний, модельные представления которых рассмотрены ранее, показано существование водородоподобной серии возбужденных состояний с соотношением энергии как  $1/n^2$  (где  $n$  - главное квантовое число).

При низких температурах электроны занимают лежащее ниже основное состояние  $E_d$ , а с ростом температуры они могут распределяться по возбужденным примесным состояниям, что должно отразиться на изменении концентрации свободных носителей в зоне проводимости с повышением температуры, так как возбужденные состояния имеют другие энергии активации и другие значения  $g$ -фактора. Вклад переходов с возбужденных состояний примеси в зону проводимости может изменить температурную зависимость концентрации на участке, описываемом формулой (6.114), т. е. при более высоких температурах, когда большее число донорных электронов будет распределяться по возбужденным состояниям. Если учесть вклад возбужденных состояний в температурную зависимость концентрации носителей заряда в области примесной проводимости, величина концентрации свободных носителей несколько уменьшится. Без учета этого фактора при определении значений  $\Delta E_d$  и  $g$ -фактора по зависимости  $n(T)$  будет допускаться некоторая систематическая ошибка.

Однако здесь мы не будем останавливаться подробнее на учете вклада возбужденных примесных состояний, так же как и на других факторах, влияющих



**Рис.6.15.** Температурные зависимости концентрации электронов (a) и уровня Ферми (b) в донорном полупроводнике с двумя типами донорной примеси при частичной компенсации доноров

на температурную зависимость концентрации носителей заряда (расщепление основного состояния примеси в поле кристалла, влияние давления и внешних полей и т. п.). К отклонению вида зависимости  $n(T)$  в области примесной электропроводности от выражений (6.112) и (6.114) может привести также учет процессов ионизации других донорных уровней, если их энергетическое положение близко к данному уровню  $E_d$ . Если уровни энергии нескольких различных доноров лежат достаточно далеко один от другого в запрещенной зоне, то на кривой температурной зависимости  $n(T)$  в области примесной электропроводности будет несколько горизонтальных участков в виде ступеней (рис. 6.15). В этом случае переходы в зону проводимости с уровня  $E_{d_1}$  к моменту начала переходов с уровня  $E_{d_2}$ , уже

закончились, и вклады переходов с различных уровней можно легко разделить. Если уровни различных примесей расположены близко один от другого, то такое разделение невозможно.

## 6.12 Вырожденные примесные полупроводники

При анализе температурной зависимости уровня Ферми как в чисто донорном, так и в частично компенсированном полупроводнике, проведенном в двух предыдущих параграфах, мы пользовались формулами для невырожденного полупроводника. Однако мы видели, что при повышении концентрации доноров в чисто донорном полупроводнике  $N_d$  или эффективной их концентрации  $N_d - N_a$  в компенсированном полупроводнике уровень Ферми тем ближе подходит ко дну зоны проводимости, чем больше концентрация  $N_d$  или  $(N_d - N_a)$ . Для некоторой концентрации примеси критерий невырожденного полупроводника  $E_c - E_F \gg k_B T$  может не выполняться, и полупроводник становится вырожденным и тогда при решении уравнения электронейтральности нужно использовать выражения для концентрации носителей через интегралы Ферми (6.34) и (6.45). Прделанный таким образом расчет показывает, что с ростом концентрации примеси уровень Ферми, действительно, может приблизиться к зоне проводимости ближе, чем на  $k_B T$ , и даже зайти в зону проводимости. Проанализируем, от каких факторов, кроме концентрации примеси, зависит наступление вырождения электронов в зоне проводимости на основе модели полупроводника с одним типом одновалентных доноров. Зависимость уровня Ферми от температуры (6.97),

полученная нами для этого случая ранее, приводит к выводу, что величина  $E_F(T)$  проходит через максимум при некоторой температуре. Обозначим эту температуру  $T_{\max}$ , а соответствующее ей значение  $E_F(T_{\max}) = E_{F\max}$ . Дифференцируя уравнение (6.97) по температуре и приравнявая нулю производную, найдем условие максимума:

$$\frac{dE_F}{dT} = \frac{k_B}{2} \ln \frac{N_d}{gN_c} - \frac{k_B T}{2} \cdot \frac{N_d}{gN_c} \cdot \frac{2}{N_d} \cdot \frac{dN_c}{dT} = 0, \quad (6.116)$$

откуда

$$\ln \frac{N_d}{gN_c} - \frac{T}{N_c} \frac{dN_c}{dT} = 0, \quad (6.117)$$

Учитывая, что

$$\frac{dN_c}{dT} = \frac{3}{2} \frac{N_c}{T}, \quad (6.118)$$

получим условие экстремума в виде:

$$\ln \frac{N_d}{gN_c} + \frac{3}{2} N_c T_{\max} = \frac{N_d}{ge^{3/2}}, \quad (6.119)$$

которое с учетом формулы (6.40) может быть для температуры  $T_{\max}$  представлено в виде:

$$T_{\max} = 8.15 \left( \frac{m}{m_{dn}^*} \right) \left( \frac{N_d}{10^{18}} \right)^{2/3}. \quad (6.120)$$

Отметим, что температура  $T_{\max}$  растет с концентрацией примеси.

Подставляя выражение (6.120) для  $T_{\max}$  в формулу (6.97) для температурной зависимости уровня Ферми, получим значение  $E_{F\max}$  в виде:

$$\begin{aligned} E_{F\max} &= \frac{E_c + E_d}{2} + \frac{k_B T_{\max}}{2} \ln \frac{N_d}{gN_c(T_{\max})} + \frac{E_c + E_d}{2} + \\ &+ \frac{3}{4} k_B T_{\max} = \frac{E_c + E_d}{2} + 5.3 \cdot 10^{-4} \left( \frac{m}{m_{dn}^*} \right) \left( \frac{N_d}{10^{18}} \right)^{2/3}. \end{aligned} \quad (6.121)$$

Здесь энергия  $E_{F\max}$  выражена в электрон-вольтах. Так как значение  $E_{F\max}$  оказалось линейно связанным со значением  $T_{\max}$ , то влияние концентрации примеси  $N_d$  и эффективной массы для плотности состояний на  $E_{F\max}$  аналогично влиянию этих же величин на  $T_{\max}$ . Величины  $E_{F\max}$  и  $T_{\max}$  тем больше, чем выше концентрация примеси  $N_d$  и чем меньше эффективная масса  $m_{dn}^*$  электронов в зоне проводимости. Назовем *критической концентрацией вырождения*  $N_{d(\text{крит})}$  величину концентрации примеси, для которой уровень Ферми совпадает с дном зоны проводимости  $E_c$ . Оценим ошибку при вычислении критической концентрации вырождения  $N_{d(\text{крит})}$  с помощью формулы (6.121), справедливой, строго говоря, только для невырожденного полупроводника. Расчет показывает, что если уровень Ферми совпадает с дном зоны проводимости  $E_c$ , то относительная ошибка в определении по формулам для невырожденного полупроводника составляет  $\sim 20\%$ .

Считая, что ошибка в определении критической концентрации вырождения

также составит величину  $\sim 20\%$ , используем формулу (6.121), подставив в нее значение  $E_{F\max} = E_c$  для нахождения  $N_{d(\text{крит})}$ . Тогда получим

$$E_{F\max} = E_c = \frac{E_c + E_d}{2} + 5.3 \cdot 10^{-4} \left( \frac{m}{m_{dn}^*} \right) \left( \frac{N_{d(\text{крит})}}{10^{18}} \right)^{2/3}, \quad (6.122)$$

откуда выведем значение критической концентрации с точностью порядка 20% в виде:

$$N_{d(\text{крит})} (\text{см}^{-3}) = 10^{22.5} \left( \frac{m_{dn}^*}{m} \right)^{3/2} [\Delta E_d (\text{эВ})]^{3/2}. \quad (6.123)$$

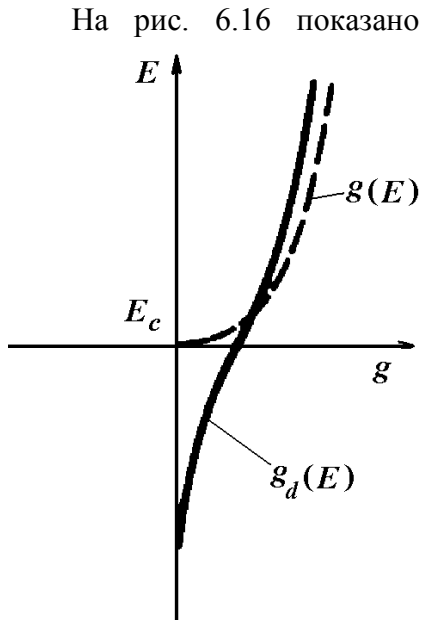
Из формулы (6.123) видно, что критическая концентрация вырождения чувствительна к эффективной массе и энергии ионизации примеси. Оценка  $N_{d(\text{крит})}$  для  $m_{dn}^* = 3m$  и  $\Delta E_d = 0,03$  эВ дает значение  $N_{d(\text{крит})} \approx 2 \cdot 10^{19} \text{ см}^{-3}$ . При уменьшении величин  $m_{dn}^*$  до  $0,001m$  и  $\Delta E_d$  до  $0,0001$  эВ получается для  $N_{d(\text{крит})}$  оценка, примерно равная  $10^{12} \text{ см}^{-3}$ . Последнее оценочное значение соответствует критической концентрации вырождения в узкозонных полупроводниках  $A^{III}B^V$ .

Зависимость концентрации вырождения от эффективной массы носителей заряда можно качественно объяснить путем следующих рассуждений. С повышением концентрации примеси растет концентрация свободных носителей заряда при температуре  $T_{\max}$ . Если эффективная масса для плотности состояний невелика, то мала и плотность состояний в зоне проводимости. Зона с малой плотностью состояний быстрее заполняется свободными электронами, и вырождение наступает при меньших концентрациях примеси.

Однако термин «вырождение» здесь применяется в связи с выбором расчетных формул для вычисления концентрации носителей заряда и положения уровня Ферми. Для сильно вырожденного электронного газа концентрация электронов в зоне проводимости перестает зависеть от температуры. Но расчет, проведенный по формулам, аппроксимирующим интегралы Ферми для сильного вырождения (6.38с) и (6.47с), дает столь большие концентрации примеси, что теряет физический смысл введение самого понятия «энергетический уровень примесного состояния». Например, для германия, начиная с концентрации доноров  $\sim 10^{18}$ , эффект экранирования свободными носителями кулоновского взаимодействия между примесными электронами и ионами примеси приводит к исчезновению примесных уровней. В этих условиях следует говорить не о возникающей благодаря перекрытию примесных волновых функций примесной зоне, а об образовании в запрещенной зоне области значений энергии, для которой функция плотности состояний  $g(E)$  отлична от нуля.

Большой вклад в современную теорию примесных состояний внес В. Л. Бонч-Бруевич, решив методом функций Грина трехмерную задачу о хаотическом распределении примеси для предельно высоких концентраций. Было показано, что при высокой концентрации примеси плотность состояний отлична от нуля во всей запрещенной зоне. Вблизи уровня Ферми она мало отличается от функции плотности состояний идеального Ферми-газа, но в окрестности экстремальных точек разрешенных зон она изменяется существенно, быстро убывая с уменьшением энергии в глубь запрещенной зоны.





**Рис. 6.16.** *Функция плотности состояний в чистом  $g(E)$  и сильно легированном  $g_d(E)$  полупроводнике*

На рис. 6.16 показано возникновение «хвоста» функции плотности состояний  $g_d(E)$  при сильном легировании. Для сравнения приведена функция  $g(E)$  нелегированного полупроводника. Однако явный вид функции  $g_d(E)$  в глубине запрещенной зоны установить трудно, так как он сильно зависит от выбранного вида потенциала взаимодействия электронов с ионами примеси.

Методика исследования вольтамперной характеристики туннельного диода позволяет экспериментально определить вид «хвоста» функции плотности состояний.

Уменьшение энергии ионизации примесей и ширины запрещенной зоны при больших уровнях легирования было установлено экспериментально для значительного числа полупроводников. Однако удовлетворительная теоретическая интерпретация экспериментальных данных такого рода пока еще не достигнута.

## 6.13 Кинетические явления в полупроводниках

Ранее были рассмотрены элементарные представления об электропроводности полупроводников и теории рассеяния носителей заряда. Здесь же будут представлены более строгие понятия и представления, учитывающие, что направленное движение носителей заряда не соответствует равновесному состоянию и, следовательно, система носителей заряда описывается некоторой неравновесной функцией распределения  $f(x, y, z, t)$ . С помощью этой функции строят теорию кинетических явлений в полупроводниках. Основным уравнением, характеризующим процессы переноса, является кинетическое уравнение Больцмана.

### 6.13.1 Кинетическое уравнение Больцмана для стационарного состояния в слабом электрическом поле

Для однородного кристалла в отсутствие градиента температуры изменение функции распределения под действием поля напряженностью  $\mathcal{E}$ , направленного по оси  $x$ , можно представить в виде

$$\left(\frac{df}{dt}\right)_{\mathcal{E}} = \frac{df}{dt} + \frac{df}{dp_x} \cdot \frac{dp_x}{dt} \quad (6.124)$$

Так как производная по времени от  $p_x$  равна  $F_x$ , то для электронов имеем

$$\frac{dp_x}{dt} = F_x = -e\mathcal{E} \quad (6.125)$$

Следовательно, из (6.124) и (6.125) получаем

$$\left(\frac{df}{dt}\right)_\varepsilon = \frac{df}{dt} - e\varepsilon \frac{df}{dp_x}. \quad (6.126)$$

Внешнее поле  $\varepsilon$  обычно много меньше внутреннего поля кристалла и поэтому вызывает относительно небольшое изменение функции распределения по сравнению с равновесной. Таким образом, можно записать

$$f = f_0 + f_1. \quad (6.127)$$

Здесь  $f_0 = f_\Phi$  или  $f_{M-B}$  для вырожденного и невырожденного газа соответственно, а  $f_1$  — малая добавочная величина, которая, однако, определяет процессы переноса (кинетические явления).

Если в некоторый момент времени, который можно принять за начало отсчета, выключить внешнее поле, то вследствие столкновений электронов с рассеивающими центрами будет восстанавливаться равновесное состояние.

При условии, что отклонение системы от равновесия является небольшим, можно считать скорость изменения функции распределения  $(df/dt)_{cm}$ , вызванного столкновениями, пропорциональной величине отклонения:

$$\left(\frac{df}{dt}\right)_{cm} = -\frac{f_1}{\tau} = -\frac{f - f_0}{\tau}, \quad (6.128)$$

где  $\tau$  — время релаксации. Поскольку  $f_0$  от времени не зависит,

$$\frac{d(f - f_0)}{f - f_0} = -\frac{dt}{\tau}, \quad (6.129)$$

отсюда

$$f_1 = (f_1)_0 e^{-t/\tau}. \quad (6.130)$$

Здесь  $(f_1)_0$  есть значение  $f_1$  в момент  $t = 0$ , т. е. в момент выключения поля. В общем случае  $\tau = \tau(\mathbf{k})$  вид функции зависит от механизма рассеяния. Следовательно, в полупроводнике, где создано электрическое поле  $\varepsilon$ , протекают два процесса: изменение функции распределения носителей заряда по состояниям под действием поля со скоростью  $(df/dt)_\varepsilon$  и релаксационный процесс, стремящийся вернуть систему в равновесное состояние, со скоростью  $(df/dt)_{cm}$ .

Стационарное состояние устанавливается при условии, когда указанные процессы уравниваются, т. е.

$$(df/dt)_\varepsilon = (df/dt)_{cm}. \quad (6.131)$$

В частности, для направления  $x$  в этом случае по (6.125) и (6.128) получим

$$-\frac{df}{dt} + e\varepsilon \frac{df}{dp_x} = \frac{f - f_0}{\tau(\mathbf{k})}. \quad (6.132)$$

Поскольку рассматривается стационарное состояние, то изменение  $f$  с течением времени равно нулю:

$$\frac{df}{dt} = 0. \quad (6.133)$$

Следовательно, из (6.132) и (6.133) имеем

$$f = f_0 + e\varepsilon \frac{df}{dp_x}. \quad (6.134)$$

Поскольку  $f$  мало отличается от  $f_0$ , то (6.134) можно переписать в виде

$$f = f_0 + \frac{e\epsilon\tau}{m_n} \frac{df_0}{dp_x}. \quad (6.135)$$

Равенство (6.135) есть один из вариантов записи кинетического уравнения Больцмана для электронного газа.

Переходя от дифференцирования  $f_0$  по импульсу к дифференцированию по энергии  $E$ , получим

$$f = f_0 + e\epsilon\tau v_x \frac{df_0}{dE}. \quad (6.136)$$

Рассмотрим процессы переноса с помощью кинетического уравнения Больцмана на примере вывода формулы для удельной электропроводности.

### 6.13.2 Электропроводность электронных полупроводников

Найдем вначале концентрацию электронов, энергия которых заключена в интервале от  $E$  до  $E + dE$ . Расчет проводим для сферической изоэнергетической поверхности с центром при  $k=0$ , так что имеем

$$dn = N(E)2f(E)dE = \frac{4\pi(2m_n)^{3/2}}{h^3} (E - E_c)^{1/2} f(E)dE, \quad (6.137)$$

где  $m_n$  — эффективная скалярная масса;  $m_n = m^* = m_{on}$ . Далее удобнее принять за начало отсчета энергии дно зоны проводимости, так что положим  $E_c = 0$ .

Заряд, который переносят  $dn$  электронов, дрейфующих вдоль оси  $x$ , за единицу времени через единицу площади,

$$-e v_x dn = -\frac{4\pi e(2m_n)^{3/2}}{h^3} E^{1/2} f(E) dE. \quad (6.138)$$

Следовательно, плотность тока в направлении  $x$

$$j = -\frac{4\pi e(2m_n)^{3/2}}{h^3} \int_0^\infty v_x E^{1/2} f(E) dE. \quad (6.139)$$

При подстановке в (6.139) выражения для  $f(E)$  по (6.136) получим сумму двух интегралов, первый из которых равен нулю, так как он выражает ток в равновесных условиях, а второе слагаемое дает

$$j = -\frac{4\pi e(2m_n)^{3/2}}{h^3} \epsilon \int_0^\infty v_x^2 \tau E^{1/2} \frac{df_0}{dE} dE. \quad (6.140)$$

Следовательно, ток связан с неравновесной добавкой  $f_1$  к функции распределения  $f_0$ , вызванной электрическим полем. Далее полагают

$$v_x^2 \approx v_y^2 \approx v_z^2 \approx \frac{2}{3} \frac{E}{m_n}. \quad (6.141)$$

Тогда по (6.139) получим

$$j = -\frac{8\pi e^2 (2m_n)^{3/2}}{3h^3 m_n} \varepsilon \int_0^{\infty} \tau E^{3/2} \frac{df_0}{dE} dE. \quad (6.142)$$

Для невырожденного электронного газа  $f_0 = f_{\text{М-Б}}$  и соответственно

$$j = \frac{8\pi e^2 (2m_n)^{3/2}}{3h^3 m_n kT} e^{\frac{E_\phi}{kT}} \varepsilon \int_0^{\infty} \tau E^{3/2} e^{-\frac{E_\phi}{kT}} dE. \quad (6.143)$$

Учитывая, что  $n_0 = N_c e^{-\frac{E_c - E_\phi}{kT}}$ , при  $E_c = 0$  получим

$$e^{\frac{E_\phi}{kT}} = \frac{n_0}{N_c} = \frac{n_0 h^3}{2(2\pi m_n kT)^{3/2}}. \quad (6.144)$$

Умножая и деля (6.143) на  $\int_0^{\infty} E e^{-\frac{E}{kT}} E^{1/2} dE$ , с учетом (6.144) получим

$$j = \frac{n_0 e^2 \langle \tau \rangle}{m_n} \varepsilon. \quad (6.145)$$

Величина  $\langle \tau \rangle$  представляет собой среднее время релаксации, полученное путем усреднения его по энергии электронов:

$$\langle \tau \rangle = \frac{\int_0^{\infty} \tau E e^{-\frac{E}{kT}} E^{1/2} dE}{\int_0^{\infty} E e^{-\frac{E}{kT}} E^{1/2} dE} \quad (6.146)$$

Кроме того, произведена замена:

$$\int_0^{\infty} E e^{-\frac{E}{kT}} E^{1/2} dE = \frac{3}{4} \sqrt{\pi} (kT)^{5/2}, \quad (6.147)$$

т. е. найден указанный интеграл.

Итак, для невырожденного электронного газа получен закон Ома в виде (6.145), так что удельная электропроводность

$$\sigma = \frac{n_0 e^2}{m_n} \langle \tau \rangle. \quad (6.148)$$

Соответственно подвижность электронов

$$\mu_n = \frac{e}{m_n} \langle \tau \rangle. \quad (6.148a)$$

Для вырожденного электронного газа электропроводность характеризуется лишь

той частью электронов, которые располагаются вблизи уровня Ферми.

Практически считают, что время релаксации этих электронов можно считать постоянной величиной  $\tau_{\phi}$ , соответствующей электронам, занимающим уровень Ферми. В этом случае все значительно проще, и в формулы (6.148), (6.148а) следует подставлять постоянную  $\tau_{\phi}$ .

### 6.13.3 Рассеяние электронов на тепловых колебаниях

При взаимодействии электрона (или дырки) с колебаниями решетки — фононами может наблюдаться передача энергии решетке и рождение нового фонона либо обратный процесс. При этом справедливы законы сохранения энергии и квазиимпульса. При рождении фонона имеем

$$E(\mathbf{k}) = E_1(\mathbf{k}_1) + \hbar\omega_q; \quad (6.149)$$

$$\mathbf{k} = \mathbf{k}_1 + \mathbf{q}, \quad (6.150)$$

где  $E(\mathbf{k})$  — энергия электрона до столкновения (до взаимодействия с колебанием решетки);  $\mathbf{k}$  — соответствующий волновой вектор;  $E_1(\mathbf{k}_1)$  — энергия электрона после рождения нового фонона с волновым вектором  $\mathbf{q}$  и энергией  $\hbar\omega_q$ .

При поглощении электроном фонона имеем

$$E_1(\mathbf{k}_1) = E(\mathbf{k}) + \hbar\omega_q; \quad (6.151)$$

$$\mathbf{k}_1 = \mathbf{k} + \mathbf{q}, \quad (6.152)$$

где  $E_1(\mathbf{k}_1)$  — энергия электрона после поглощения фонона,  $\mathbf{k}_1$  — волновой вектор этого электрона.

Описанный механизм рассеяния называют однофононным. Так как число фононов определяется температурой, то и рассеяние электронов на тепловых колебаниях решетки зависит от температуры.

В кристалле возможен, но менее вероятен многофононный механизм рассеяния, когда наблюдается рождение или поглощение более чем одного фонона. Для случая рассеяния электронов на продольных акустических фононах в атомных полупроводниках время релаксации находят в виде

$$\tau(\mathbf{k}) = \frac{\tau_0}{(m^*)^{3/2} E^{1/2} T}, \quad (6.153)$$

где  $\tau_0$  — постоянная величина для данного кристалла;  $m^*$  — эффективная масса электронов;  $E$  — энергия электронов. Соответственно подвижность выражается формулой

$$\mu_{TK} = \frac{4e\tau_0}{3\sqrt{\pi} k^{1/2} (m^*)^{5/2} T^{3/2}} = bT^{-3/2}. \quad (6.154)$$

Зависимость, близкая к (6.154), наблюдается нередко в довольно широких интервалах температуры. Но кроме того одновременно происходит рассеяние на оптических колебаниях решетки, двухфононное рассеяние и другие типы рассеяния и поэтому точного выполнения зависимости типа (6.154) нельзя ожидать.

### 6.13.4 Термоэлектрические явления

К термоэлектрическим явлениям относятся три эффекта: Зеебека, Пельтье, Томсона. Явлением Зеебека называют возникновение термо – э.д.с. в цепи, составленной из двух различных твердых тел, при наличии разности температур спаев. Явление Пельтье состоит в нагревании или охлаждении спая двух материалов при протекании через него постоянного тока. Этот эффект не зависит от выделения тепла Ленца — Джоуля, т. е. имеет иную природу. Явление Томсона состоит в выделении или поглощении тепла, добавочно к теплу Ленца—Джоуля, при протекании постоянного тока по однородному полупроводнику, в котором имеется градиент температуры.

Явления Зеебека и Пельтье обычно наблюдаются в металлах. Однако в полупроводниках эти явления выражены значительно ярче, например соответствующие величины данных явлений могут на несколько порядков превышать таковые в металлах. Поэтому явления Зеебека и Пельтье в полупроводниках находят большое практическое применение. В частности, полупроводниковые пары с большой термо-э.д.с. можно использовать как источники электропитания, а явление Пельтье — в холодильных установках.

Явление Томсона представляет главным образом теоретический интерес. Градиент температуры создает градиент концентрации и соответствующий диффузионный ток, вследствие чего возникают объемные заряды вдоль полупроводника. Если поле объемных зарядов направлено против внешнего поля, то последнее совершает работу против внутреннего поля, и выделяется дополнительное тепло. В случае совпадения направления обоих полей внутреннее поле совершает часть работы по созданию дрейфа носителей заряда, что осуществляется в конечном итоге за счет тепловой энергии полупроводника, поэтому последний охлаждается.

Наибольший практический интерес представляет явление возникновения термо-э.д.с., т. е. явление Зеебека. Остановимся на нем более подробно, кратко затрагивая явление Пельтье.

#### Явление Зеебека в невырожденных полупроводниках

Дифференциальную термо- э.д.с. определяют в виде

$$\alpha = \frac{d\varepsilon}{dT}, \quad (6.155)$$

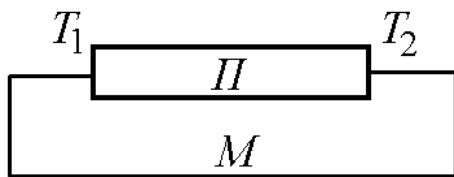


Рис. 6.17. Контакт металл – полупроводник

где  $dT$  - разность температур спаев;  $d\varepsilon$  - термо-э.д.с., соответствующая  $dT$ .

При конечной разности температур имеем

$$\varepsilon = \int_{T_1}^{T_2} \alpha(T) dT. \quad (6.156)$$

Часто определяют  $\alpha$  полупроводника по отношению к металлу, т. е. рассматривают контакты полупроводника с металлом.

На рис. 6.17 полупроводник замыкается с двух концов металлом. Контакты имеют температуру  $T_1$  и  $T_2$ . В цепи протекает ток, величина которого определяется значением термо-э.д.с. Основной вклад в термо-э.д.с. в данном случае вносит разность потенциалов, возникающая непосредственно в полупроводнике. Контактные же разности потенциалов оказываются относительно малыми.

Равновесные концентрации носителей заряда в невырожденном примесном полупроводнике являются экспоненциальными функциями температуры, т. е. для концов полупроводника имеем

$$n_{01} = N_c e^{-\frac{E_c - E_f}{kT_1}} ; \quad (6.157)$$

$$p_{01} = N_v e^{-\frac{E_f - E_v}{kT_1}} ; \quad (6.158)$$

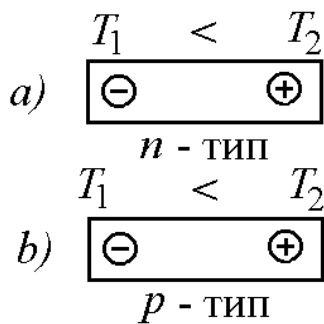
$$n_{02} = N_c e^{-\frac{E_c - E_f}{kT_2}} ; \quad (6.159)$$

$$p_{02} = N_v e^{-\frac{E_f - E_v}{kT_2}} . \quad (6.160)$$

Ввиду небольшой разницы температур концов, которая здесь предполагается, можно считать  $E_f$  константой.

При  $T_2 > T_1$  имеем  $n_{02} \gg n_{01}$  и  $p_{02} \gg p_{01}$ . Следовательно, как только создается разность температур ( $T_2 - T_1$ ), сразу же начинается процесс диффузии носителей заряда от более горячего конца образца полупроводника к более холодному. Указанный преимущественно направленный перенос приводит к изменению равновесных концентраций носителей заряда и к возникновению объемных зарядов в полупроводнике. Последние создают встречное поле, и через некоторое время процесс достигает динамического равновесия.

При этом диффузионный ток носителей заряда будет уравниваться дрейфовым током, создаваемым возникшим внутренним полем (полем объемных зарядов), так что суммарный ток будет равен нулю.



**Рис. 6.18.** Объемный заряд в полупроводниках *n*-типа (а) и *p*-типа (б) при наличии градиента температур

Если полупроводник *n*-типа, т. е. равновесная концентрация электронов много больше равновесной концентрации дырок, то можно принимать во внимание лишь диффузию электронов. В полупроводнике *p*-типа, наоборот, можно рассматривать лишь диффузию дырок.

В электронном полупроводнике указанная диффузия приводит к возникновению на холодном конце (спае) отрицательного объемного заряда, а на горячем конце - положительного объемного заряда. Соответственно величина  $\alpha$  считается отрицательной (рис. 6.18 а). В дырочном полупроводнике диффузия носителей заряда приводит к противоположному результату - холодный конец (спай) заряжается положительно, горячий - отрицательно. Соответственно величина  $\alpha$  считается положительной (рис. 6.18, б).

Таким образом, знак  $\alpha$  определяется по полярности термо-э.д.с., т. е. по знаку холодного или горячего конца спае (контакта) полупроводника с металлом. Положительно заряженный холодный контакт наблюдается для дырочных полупроводников, отрицательно заряженный - для электронных.

На этом основан метод определения типа примесной электропроводности с помощью термозонда (рис. 6.19). Один из контактов - зондов подогревается, а другой имеет температуру окружающей среды, как и полупроводник. Направление тока в цепи определяется электроизмерительным прибором. Если

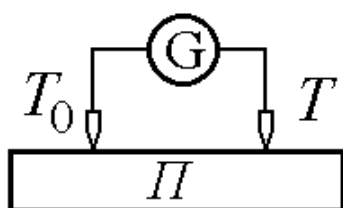


Рис.6.19. Термозонд

Для вывода этой формулы воспользуемся связью между коэффициентом термо-э.д.с. и коэффициентом Пельтье, установленной в термодинамике. Теплота Пельтье  $Q_{\Pi}$  определяется в виде

$$Q_{\Pi} = \Pi It, \quad (6.161)$$

причем

$$\Pi = \alpha T. \quad (6.162)$$

Здесь  $I$  - сила постоянного тока, протекающего через спай в течение  $t$ ;  $\Pi$  - коэффициент Пельтье. Теплота Пельтье выделяется на спае вследствие передачи части энергии носителей заряда решетке при переходе их из одного материала в другой.

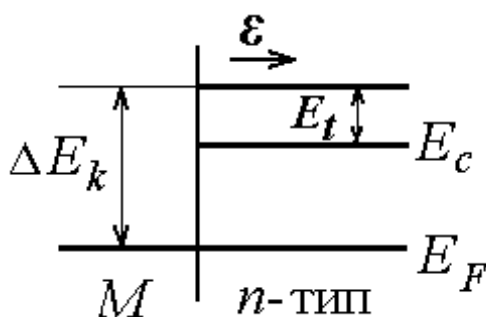


Рис. 6.20. Зонная диаграмма контакта металл – полупроводник

При обратном направлении тока наблюдается обратный процесс – решетка передает часть своей энергии носителям заряда и охлаждается.

На рис. 6.20 изображена зонная диаграмма для контакта металл-полупроводник  $n$ -типа. В условиях равновесия уровни Ферми всегда устанавливаются одинаковыми. При наложении электрического поля  $\epsilon$  указанного направления электроны из зоны проводимости полупроводника переходят в металл. Полагая, что поле  $\epsilon$  слабое и поэтому сдвигом уровней Ферми можно пренебречь, так что зоны остаются практически горизонтальными, а также пренебрегая выпрямительными эффектами на контакте полупроводник - металл, получим уменьшение энергии электрона (охлаждение электрона) на величину  $\Delta E_k$ . Эта энергия и есть теплота Пельтье, выделяющаяся при переносе количества электричества в виде  $e$ , т. е. одного элементарного заряда. Таким образом,

$$\Delta E_k = Q_{\Pi} = Pe = \alpha Te. \quad (6.163)$$

Если пропускать ток в противоположном направлении, т. е. поднимать электроны с уровня Ферми до некоторого уровня в зоне проводимости полупроводника, то требуется нагревать электроны на соответствующую величину  $\Delta E_k$ . Такое нагревание осуществляется за счет передачи решеткой энергии и, следовательно, охлаждения спая.

Здесь рассмотрены переходы на уровень Ферми и с уровня Ферми металла, поскольку в электропроводности, как и в тепловом движении, в металле участвуют лишь электроны, располагающиеся вблизи уровня Ферми. Кроме того, как уже отмечалось ранее, условием термодинамического равновесия любых

подогреваемый зонд оказывается положительным электродом, то полупроводник  $n$ -типа, если же отрицательным, то полупроводник -  $p$ -типа.

Связь  $\alpha$  с концентрацией носителей заряда в примесных полупроводниках выражают с помощью формулы Н. Л. Писаренко, выведенной им в начале 40-х годов нашего века, когда теория полупроводников только начинала развиваться.



контактирующих систем является равенство их химических потенциалов (уровней Ферми), так что в невырожденном электронном полупроводнике носители заряда располагаются значительно выше  $E_{\phi}$ , общего для полупроводника и контактирующего с ним металла.

Величину  $\Delta E_{\kappa}$  входящую в соотношение (6.163), можно представить в виде

$$\Delta E_{\kappa} = (E_c - E_{\phi}) + E_T, \quad (6.164)$$

где  $E_T$  - средняя тепловая энергия невырожденного электронного газа, так как для такого газа большинство частиц имеет тепловые скорости, близкие к средней. Дно зоны проводимости принимают за уровень, где кинетическая энергия электронов проводимости равна нулю, т. е. полная энергия электронов на дне зоны состоит лишь из потенциальной энергии. Следовательно,  $E_T$  отсчитывается вверх от  $E_c$ , и представляет собой кинетическую энергию электронов проводимости. При таком отсчете

$$E_T = (r + 2)kT. \quad (6.165)$$

Величина  $r$  определяется из формулы

$$l = CE^r, \quad (6.166)$$

где  $l$  — средняя длина свободного пробега;  $E$  — полная энергия носителя заряда;  $r$  - показатель степени. Через  $C$  обозначены все остальные величины, которые входят в эту формулу. Формулу (6.166) выводят согласно теории рассеяния, где показано, что в зависимости от механизма рассеяния интервал изменения  $r$  следующий:

$$0 \leq r \leq 2, \quad (6.167)$$

причем при рассеянии на акустических колебаниях решетки  $r = 0$ ; при рассеянии на оптических колебаниях ионной решетки  $r = 1$ ; при рассеянии на ионах примеси  $r = 2$ . Согласно (6.163)-(6.165) получим

$$\alpha Te = E_c - E_{\phi} + (r + 2)kT. \quad (6.168)$$

Величину  $(E_c - E_{\phi})$  определим из следующей формулы

$$n_0 = N_c e^{-\frac{E_c - E_{\phi}}{kT_1}} \quad (6.169)$$

в виде

$$E_c - E_{\phi} = kT \ln \frac{N_c}{n_0}, \quad (6.170)$$

тогда с учетом (6.168)

$$\alpha = \frac{k}{e} \left( r + 2 + \ln \frac{N_c}{n_0} \right). \quad (6.171)$$

Если учесть условие  $\alpha < 0$  для электронного полупроводника, то перед правой частью (6.171) следует поставить знак минус. То же самое можно получить, если считать, что в формулу для  $\alpha$  входит заряд частицы, взятый по абсолютному значению (для электронов  $e < 0$ , а в формуле подразумевается  $e > 0$ ). Итак, для электронного полупроводника формула Писаренко имеет вид

$$\alpha_n = -\frac{k}{e} \left( r + 2 + \ln \frac{N_c}{n_0} \right); \quad (6.172)$$

для дырочного полупроводника

$$\alpha_p = \frac{k}{e} \left( r + 2 + \ln \frac{N_v}{p_0} \right). \quad (6.173)$$

Формулы (6.172), (6.173) достаточно хорошо согласуются с опытом, причем значения  $\alpha$  могут достигать мВ/град, тогда как для металлических пар  $\alpha$  приблизительно мкВ/град. Согласно этим формулам величина  $\alpha$  в полупроводнике не зависит от металла, с которым осуществляется контакт. Поэтому можно говорить о термо-э.д.с. самого полупроводника без указания металла, по отношению к которому она определяется. При выводе формулы (6.172) был использован только один контакт полупроводника с металлом, что оказалось возможным в связи с зависимостью (6.173). Для эффекта Пельтье нет необходимости рассматривать второй контакт, а связь коэффициента Пельтье с коэффициентом термо-э. д. с. (6.162) использована в предположении, что  $T$  есть средняя температура полупроводника  $\left( T = \frac{T_1 + T_2}{2} \right)$  и разность  $dT = T_2 - T_1$  является малой величиной (температуру спая можно также принять за величину  $T$ ).

Таким образом,  $\alpha_n$  и  $\alpha_p$  являются функциями концентрации носителей заряда и функциями температуры полупроводника (последнее через  $N_c$ ,  $N_v$ ,  $n_0$ ,  $p_0$ ). Как  $\alpha_p$ , так и  $\alpha_n$  с ростом концентрации носителей заряда по абсолютному значению убывают. При наличии заметной составляющей неосновных носителей заряда, т. е. при смешанной, а также при собственной электропроводности получают:

$$\alpha = \frac{\alpha_p \mu_p p_0 + \alpha_n \mu_n n_0}{\mu_p p_0 + \mu_n n_0}. \quad (6.174)$$

$$\alpha_i = \frac{\alpha_p \mu_p + \alpha_n \mu_n}{\mu_p + \mu_n}. \quad (6.175)$$

Поскольку согласно (6.172), (6.173) имеем  $\alpha_n < 0$  и  $\alpha_p > 0$ , то числитель (6.174) есть разность соответствующих величин, т. е. может оказаться очень малым (по абсолютному значению). Знак же  $\alpha$  определяется в зависимости от абсолютных величин, слагаемых в числителе (6.174).

При подходе к области собственной электропроводности для полупроводника, содержащего акцепторы, наблюдается инверсия знака  $\alpha_p$ , если  $\mu_n > \mu_p$ , т. е. если подвижность электронов больше подвижности дырок. В полупроводнике, содержащем доноры, с ростом  $T$  уменьшается абсолютная величина  $\alpha_n$ .

При достаточно высокой температуре  $\alpha_i$  не зависит от наличия акцепторов и доноров и по абсолютному значению является малой величиной (при  $\mu_n > \mu_p$  она имеет отрицательный знак). Следует иметь в виду, что для каждого значения  $T$  (средней температуры полупроводника) величина  $\alpha$  экспериментально определяется при малом  $dT$  (например, порядка нескольких градусов).

#### **Вывод формулы для коэффициента термо-э.д.с. с помощью уравнения Больцмана**

Выше приведен один из вариантов (6.135) записи уравнения Больцмана. Другой вариант уравнения Больцмана для стационарного случая можно записать в виде

$$(\mathbf{v}_n \nabla_r f) + \frac{1}{\hbar} (\mathbf{F} \nabla_k f) = -\frac{f_1}{\tau}, \quad (6.176)$$

где  $\mathbf{v}_n$  - скорость электрона;  $\mathbf{F}$  - сила, действующая на электрон;  $\nabla_r$  и  $\nabla_k$  - операторы градиента в пространстве координат и в  $k$ -пространстве.

Напомним, что по определению

$$\text{grad}_r = \nabla_r = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k}; \quad (6.177)$$

$$\text{grad}_k = \nabla_k = \frac{\partial}{\partial k_x} \mathbf{i} + \frac{\partial}{\partial k_y} \mathbf{j} + \frac{\partial}{\partial k_z} \mathbf{k}, \quad (6.178)$$

где  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  — единичные векторы, направленные по осям координат.

Скалярные произведения, входящие в уравнение, можно раскрыть следующим образом:

$$\begin{aligned} (\mathbf{v}_n \nabla_r f) &= v_{nx} \frac{\partial f}{\partial x} + v_{ny} \frac{\partial f}{\partial y} + v_{nz} \frac{\partial f}{\partial z}; \\ \frac{1}{\hbar} (\mathbf{F} \nabla_k f) &= \left( \frac{dk}{dt} \nabla_k f \right) = \frac{\partial k_x}{\partial t} * \frac{\partial f}{\partial k_x} + \frac{\partial k_y}{\partial t} * \frac{\partial f}{\partial k_y} + \frac{\partial k_z}{\partial t} * \frac{\partial f}{\partial k_z}. \end{aligned}$$

При малой величине отклонения  $f_1 = f - f_0$  в левой части уравнения (6.176)  $f$  можно заменить на  $f_0$ .

Следовательно, имеем

$$(\mathbf{v}_n \nabla_r f_0) + \frac{1}{\hbar} (\mathbf{F} \nabla_k f_0) = -\frac{f_1}{\tau}. \quad (6.179)$$

Сила в данном случае выражается формулой

$$\mathbf{F} = -e\varepsilon = e\nabla_r \varphi, \quad (6.180)$$

где  $\varepsilon$  - напряженность электрического поля объемных зарядов;  $\varphi$  - соответствующий электростатический потенциал. Функция распределения для невырожденного электронного газа

$$f_0 = e^{-\frac{E - E_\phi}{kT}} = f_0(E, E_\phi, T), \quad (6.181)$$

Полная энергия электрона

$$E = \frac{\hbar^2 k^2}{2m^*}. \quad (6.182)$$

Полагаем для простоты, что  $m^*$  скаляр, т. е. имеем дело со сферической изоэнергетической поверхностью. Скорость электрона можно выразить в виде

$$\mathbf{v}_n = \frac{1}{\hbar} \nabla_k E(\mathbf{k}) = \frac{\hbar}{m^*} \mathbf{k}. \quad (6.183)$$

При дифференцировании  $f_0$  в пространстве координат следует учитывать, что в данном случае  $T = T(x, y, z)$  и  $E_\phi = E_\phi(x, y, z)$ .

Соответственно получим

$$\nabla_r f_0 = \frac{\partial f_0}{\partial T} \nabla_r T + \frac{\partial f_0}{\partial E_\phi} \nabla_r E_\phi = \frac{\partial f_0}{\partial E} \left( \frac{E_\phi - E_0}{T} \nabla_r T - \nabla_r E_\phi \right); \quad (6.184)$$

$$\nabla_k f_0 = \frac{\partial f_0}{\partial E} \nabla_k E = \hbar \frac{\partial f_0}{\partial E} \mathbf{v}_n. \quad (6.185)$$

Подставляя эти выражения в (6.179), определим

$$f_1 = -\tau(k) \frac{\partial f_0}{\partial E} \left( \frac{E_\phi - E_0}{T} \nabla_r T - \nabla_r (E_\phi - e\varphi) \right) \mathbf{v}_n. \quad (6.186)$$

Можно провести аналогичные выкладки для свободных дырок и получить соответствующую формулу для  $f_1$ .

Плотность тока вычисляют, например, для случая, когда рассеяние носителей заряда осуществляется на акустических колебаниях кристаллической решетки, и длина свободного пробега не зависит от энергии ( $r = 0$ ), так что можно записать

$$\tau = \frac{l}{v_n} = \frac{m^* l}{\hbar k}. \quad (6.187)$$

В таком случае электронная составляющая плотности тока

$$j_n = \frac{e\hbar l}{3\pi^2 m^*} \left( \frac{E_\phi}{T} \nabla T - \nabla_r (E_\phi - e\varphi) \right) \int_0^\infty \frac{\partial f_0}{\partial E} k^3 dk - \frac{e\hbar l}{3\pi^2 m^* T} \nabla T \int_0^\infty \frac{\partial f_0}{\partial E} E k^3 dk.$$

Далее находят дырочную составляющую плотности тока и полный ток приравнивают нулю, поскольку термо-э. д. с. определяется как разность потенциалов для разомкнутой цепи.

Дифференциальная термо - э. д. с.

$$\alpha = \frac{\left| \nabla \left( \varphi - \frac{E_\phi}{y} \right) \right|}{|\nabla T|}. \quad (6.188)$$

При рассеянии на акустических колебаниях решетки по (6.154), выражая подвижность через длину свободного пробега, находят

$$\mu_{TK} = \frac{4el}{3(2\pi k m^* T)^{1/2}}. \quad (6.189)$$

Используя уже применявшиеся формулы для  $N_c$ ,  $N_v$ ,  $n_0$ ,  $p_0$  и других величин, в результате расчета получим

$$\alpha = \frac{k}{e} * \frac{\mu_p p_0 \left( 2 + \ln \frac{N_c}{n_0} \right) - \mu_n n_0 \left( 2 + \ln \frac{N_v}{p_0} \right)}{\mu_p p_0 + \mu_n n_0}, \quad (6.190)$$

т. е. формулу (6.174) при  $r = 0$ .

Для собственного полупроводника, принимая  $N_c = N_v$ ,

$$\alpha = -\frac{k}{e} * \frac{\mu_n / \mu_p - 1}{\mu_n / \mu_p + 1} \left( 2 + \frac{\Delta E_0}{2kT} \right). \quad (6.191)$$

Таким образом, в собственном полупроводнике величина  $\alpha$  определяется отношением подвижности электронов и дырок, шириной запрещенной зоны  $\Delta E_0$  и температурой  $T$ . Заметим, что при выводе соотношений (6.190), (6.191) не учитывалось увлечение носителей заряда фононами. При наличии градиента температуры в полупроводнике наблюдается направленное движение фононов от горячих областей к холодным. При каждом столкновении электрона с фононом, которое сопровождается уничтожением фонона, его импульс передается электрону. Поскольку фононов с импульсом, направленным от горячих областей к холодным больше, чем с противоположным направлением импульса, часть импульса передается электронам.

Вследствие указанных столкновений наблюдается дополнительный дрейф электронов — увлечение электронов фононами. Этот дрейф создает дополнительную термо - э. д. с. Однако эффект увлечения наблюдается лишь при низких температурах, так как при повышении температуры рассеяние фононов на фононах быстро восстанавливает равновесное распределение фононов, и эффект увлечения исчезает.

### 6.13.5 Гальваномагнитные эффекты. Эффект Холла.

При совместном действии электрического  $\mathbf{E}$  и магнитного  $\mathbf{B}$  полей в веществе возникает ряд кинетических явлений, приводящих к изменению в продольном направлении, вдоль которого приложено электрическое поле  $\mathbf{E}$ , электропроводности и теплопроводности, а также возникновению поперечной по отношению к направлениям электрического и магнитного полей разности потенциалов и температур. Такие явления обычно называют *гальваномагнитными эффектами*. Среди многочисленных гальваномагнитных эффектов хорошо изучены эффект Холла, эффект магнетосопротивления или изменения продольного сопротивления в магнитном поле, эффект Эттингсгаузена или возникновение поперечной разности температур (поперечный гальванотермомагнитный эффект) и эффект Нернста или возникновение продольной разности температур (продольный гальванотермомагнитный эффект).

Термины *поперечный* и *продольный* относятся к направлению градиента температур. По отношению к магнитному полю гальваномагнитные эффекты делятся на четные и нечетные. Эффект называется *нечетным*, если он не меняет знак при изменении направления магнитного поля на противоположный. Если знак эффекта меняется, то он называется *четным*. К нечетным эффектам относятся эффекты Холла и Эттингсгаузена. К четным относятся эффект магнетосопротивления и эффект Нернста. Гальваномагнитные явления можно рассматривать в адиабатических, когда образец не обменивается энергией с окружающей средой, и в нем возникает градиент температур, и изотермических условиях, когда происходит обмен энергией с окружающей средой, и,

следовательно,  $\nabla T = 0$ . В первом случае эффект будет адиабатическим, во втором - изотермическим. Эффекты Холла и магнетосопротивления обычно рассматриваются в изотермических условиях.

Качественную картину гальваномагнитных эффектов можно понять, рассматривая движение заряженной частицы в скрещенных магнитном и электрическом полях. Как известно из электродинамики, сила, действующая на электрический заряд в электромагнитном поле, является силой Лоренца:

$$\mathbf{F} = e\mathbf{E} + e[\mathbf{v}, \mathbf{B}]. \quad (6.192)$$

В скрещенных полях ( $\mathbf{E} \perp \mathbf{B}$ ) носитель заряда будет двигаться по циклоиде, или некоторой другой траектории, получающейся в результате сложения двух видов движения частицы: вращения по окружности радиуса

$$r_0 = m^* \mathbf{E} / e\mathbf{B}^2 \quad (6.193)$$

под действием электрического и магнитного поля и перемещения под действием электрического поля со скоростью дрейфа

$$v_d = \mu\mathbf{E}. \quad (6.194)$$

Таким образом, наложение магнитного поля на проводник, по которому течет электрический ток, будет изменять траекторию носителей заряда, «закручивая» их по круговой орбите тем сильнее, чем больше магнитное поле.

Критерием величины магнитного поля можно считать соотношение между радиусом кривизны  $r_0$  (6.193) и длиной свободного пробега  $\lambda$ . Если радиус кривизны траектории движения электрона много больше его средней длины свободного пробега:

$$r_0 \gg \lambda, \quad (6.195)$$

то движение электрона мало искажается, и такие поля называются *слабыми*. Очевидно, если

$$r_0 \ll \lambda, \quad (6.196)$$

то поле является *сильным*, так как оно закручивает носители заряда уже так сильно, что они резко меняют свою траекторию и, следовательно, механизм рассеяния. Условия (6.195) - (6.196) можно заменить эквивалентными при сравнении времени релаксации  $\tau$  с периодом обращения носителя по круговой орбите под действием магнитного поля  $T_{цикл}$ :

$$\text{поле слабое, если } \tau \ll T_{цикл}, \quad (6.197)$$

$$\text{поле сильное, если } \tau \gg T_{цикл}. \quad (6.198)$$

Величина  $T_{цикл}$  может быть найдена из условия для циклотронного резонанса:

$$\omega_{цикл} = 2\pi / T_{цикл} = e\mathbf{B} / m^*. \quad (6.199)$$

где величина  $\omega_{цикл}$  называется циклотронной частотой. Для слабого поля из формул (6.197) и (6.198):

$$\mu\mathbf{B} \ll 1; \quad (6.200)$$

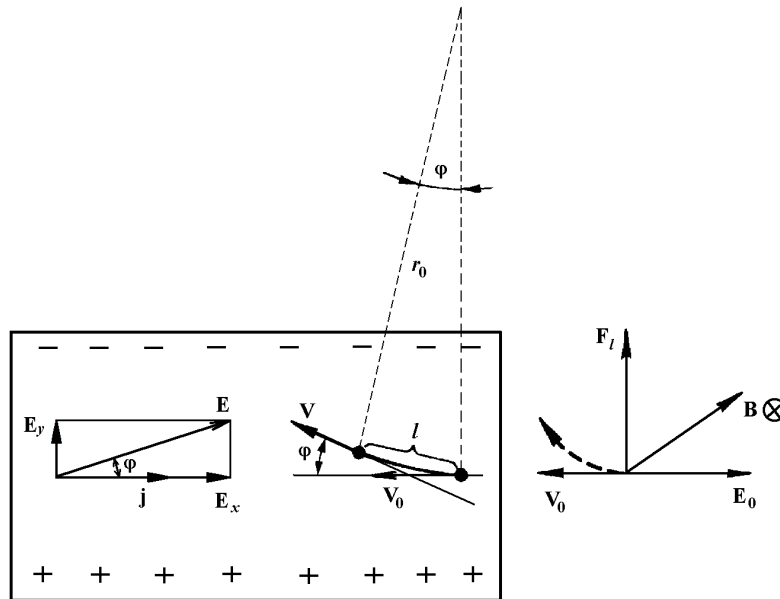
для сильного поля

$$\mu\mathbf{B} \gg 1. \quad (6.201)$$

Из таких критериев величины магнитного поля (6.200) и (6.201) видно, что понятие слабого или сильного поля зависит не только от «внешнего» фактора — величины индукции магнитного поля  $\mathbf{B}$ , но и от подвижности носителей в

кристалле, т. е. от свойств вещества. При этом может оказаться, что одно и то же поле для одного значения подвижности будет слабым, а для другого значения подвижности - сильным. Например, для германия при высоких температурах подвижность электронов  $\mu_n \approx 3000 \text{ см}^2/\text{В}\cdot\text{с}$ . Магнитное поле напряженностью 10 кЭ будет удовлетворять критерию слабого поля (6.200). То же поле при низких температурах ( $\sim 10 \text{ К}$ ) и подвижности электронов  $10^4 - 10^5 \text{ см}^2/\text{В}\cdot\text{с}$  будет удовлетворять условию (6.201) и, следовательно, считаться сильным. При рассмотрении совместного действия магнитного и электрических полей необходимо также учитывать разброс носителей заряда по скоростям и энергиям.

Рассмотрим основные гальваномагнитные эффекты при выполнении критерия слабого магнитного поля (6.200). Заметим, что возникновение поперечного по отношению к электрическому  $\mathbf{E}$  и магнитному  $\mathbf{B}$  полям холловского поля  $\mathbf{E}_x$  как раз и обусловлено закручивающим действием магнитной составляющей силы Лоренца  $e[\mathbf{v}_d, \mathbf{B}]$ . Холловское поле  $\mathbf{E}_x$  растет до тех пор, пока оно не скомпенсирует магнитную составляющую силы Лоренца. После этого носители заряда двигаются как бы только под действием приложенного к образцу электрического поля  $\mathbf{E}$  в направлении  $x$  (если  $\mathbf{E} = \mathbf{E}_x$ ,  $\mathbf{B} =$



**Рис. 6.21.** Искривление траектории движения носителя заряда в магнитном поле и возникновение угла между  $\mathbf{j}$  и  $\mathbf{E}$  в эффекте Холла

$\mathbf{B}_y$ ,  $\mathbf{E}_x = \mathbf{E}_z$ ), и траектория движения носителей представляет собой снова прямую линию вдоль оси  $x$ . Следовательно, вектор плотности тока  $\mathbf{j}$  также направлен вдоль оси  $x$ , однако суммарное электрическое поле  $\mathbf{E}$ , учитывающее возникшее поперечное холловское поле  $\mathbf{E}_x$  приложенное внешнее электрическое поле  $\mathbf{E}_y$ , при этом повернуто на некоторый угол  $\phi$  относительно  $x$  или  $\mathbf{j}$ . Таким образом, за счет накопления зарядов на поперечных гранях образца при действии магнитного поля между вектором плотности тока  $\mathbf{j}$  и напряженностью поля в образце  $\mathbf{E}$  возникает угол  $\phi$ , называемый *углом Холла*.

Угол Холла с учетом того, что возникшее поперечное поле  $E_z = v_x B_y$ ,  $\mathbf{j} = en\mathbf{v}_d$ , так что  $E_z = j_x B_y / (en) = R[\mathbf{j}\mathbf{B}]$  и  $\sigma = en\mu$ , а также рис. 6.21 определяется как

$$\operatorname{tg}\varphi = \mathbf{E}_X / \mathbf{E} = R[\mathbf{j}\mathbf{B}] / \mathbf{E} = R\mathbf{B}\sigma\mathbf{E} / \mathbf{E} = \mu\mathbf{B}. \quad (6.202)$$

При описании эффекта Холла не учитывались процессы рассеяния. Более строгое рассмотрение эффекта Холла на основе кинетического уравнения Больцмана показывает, что коэффициент Холла  $R$  должен учитывать механизм рассеяния носителей заряда через некоторую константу  $A$ . Тогда:

$$R = \pm A / en, \quad (6.203)$$

где  $n$  — концентрация электронов или дырок в зависимости от знака правой части формулы (6.203). С учетом соотношения (6.202) и определения проводимости  $\sigma = en\mu$  можно записать

$$\mu_d = |R|\sigma / A, \quad (6.204)$$

откуда следует, что из измерений коэффициента Холла и электропроводности можно с точностью до константы  $A$  определить дрейфовую подвижность. Величину  $A\mu_d$ , имеющую, очевидно, размерность подвижности, называют *холловской подвижностью*:

$$\mu_X = A\mu_d. \quad (6.205)$$

Именно она определяет угол Холла  $\varphi$ , и вычисляется как произведение коэффициента Холла на электропроводность. Соотношение (6.205) позволяет находить истинную дрейфовую подвижность при известном механизме рассеяния, определяющем величину  $A$ .

Как мы видели, механизмы рассеяния достаточно хорошо описываются степенной зависимостью времени релаксации или длины свободного пробега от энергии (6.153). Поэтому величина  $A$  принимает различные значения в зависимости от характера механизма рассеяния. Для ковалентных кристаллов при рассеянии на акустических фононах длина свободного пробега не зависит от энергии ( $r = 0$ ):

$$A = 3\pi/8 \quad \text{и} \quad R = 3\pi/8 \cdot 1/en \approx 1.17 \frac{1}{en}. \quad (6.206)$$

В полупроводниках с ионной решеткой при рассеянии на оптических фононах и температурах ниже температуры Дебая ( $T \ll \theta_D$ ),  $r = 1/2$ :

$$A \approx 0.99 \quad \text{и} \quad R \approx 0.99 \frac{1}{en}; \quad (6.207)$$

при температурах  $T > \theta_D$ ,  $r = 1$ :

$$A \approx 1.11 \quad \text{и} \quad R \approx 1.11 \frac{1}{en}; \quad (6.208)$$



при рассеянии на ионах примеси  $r = 2$ :

$$A \approx 315/512 \approx 1.93 \quad \text{и} \quad R \approx 1.93 \frac{1}{en}; \quad (6.209)$$

при рассеянии электронов на дырках также:

$$A \approx 315/512 \quad \text{и} \quad R \approx 1.93 \frac{1}{en}; \quad (6.210)$$

при рассеянии носителей заряда на нейтральных примесных атомах:

$$A = 1 \quad \text{и} \quad R = \frac{1}{en}. \quad (6.211)$$

Величина  $A$  получена из теоретических расчетов на основании кинетического уравнения Больцмана. В вырожденных полупроводниках, как и в металлах, в электропроводности принимают участие только электроны, находящиеся на самых высоких уровнях энергии (вблизи поверхности Ферми для металлов). Следовательно, в этом случае можно не учитывать распределения электронов по энергиям. Тогда время релаксации можно считать постоянным. В этом случае  $A = 1$  и холловская подвижность равна дрейфовой

$$\mu_X = \mu_d. \quad (6.212)$$

При смешанной электропроводности, когда в процессе переноса участвуют и электроны, и дырки, анализ эффекта Холла усложняется. Как известно, сила Лоренца для электронов и дырок направлена в одну сторону. Следовательно, электроны и дырки отклоняются в одну сторону. Учитывая разные подвижности и концентрации электронов и дырок, в данном случае нельзя говорить о взаимной компенсации объемных зарядов электронов и дырок. Тогда на одной поперечной грани будет скапливаться отрицательный заряд, а на другой - положительный некомпенсированный заряд. Таким образом, и в этом случае возникает поперечное холловское поле, тормозящее движение носителей одного знака (например, электронов) и ускоряющее движение других (дырок). Холловское поле будет возрастать до тех пор, пока оно не уравнивает потоки электронов и дырок. Применяя для анализа эффекта Холла кинетическое уравнение, можно получить для смешанной электропроводности:

$$R = \frac{A}{e} \frac{\mu_p^2 p - \mu_n^2 n}{(\mu_p p + \mu_n n)^2}. \quad (6.213)$$

Постоянная  $A$  определяется теми же механизмами рассеяния (7.206) - (6.211). Из формулы (6.212) следует, что, если в собственном полупроводнике ( $n = p$ ) подвижности электронов и дырок одинаковы, то холловское поле равно нулю. Можно заметить, рассматривая уравнения  $\sigma = \sigma_n + \sigma_p$  и (6.212), что для смешанной проводимости измерение электропроводности и постоянной Холла не дает достаточных данных для нахождения подвижности и концентрации электронов и дырок. В этом случае к уравнениям  $\sigma = \sigma_n + \sigma_p$  и (6.212) нужно добавить независимые дополнительные уравнения, которые могут быть получены из измерений других кинетических эффектов.

Магнитное поле приводит не только к появлению угла Холла между  $\mathbf{j}$  и  $\mathbf{E}$ , но и влияет на величину электропроводности, т. е. изменяет сопротивление образца в магнитном поле (*эффект магнетосопротивления*). Количественно этот

эффект определяется отношением изменения сопротивления в магнитном поле  $\Delta\rho = \rho(\mathbf{B}) - \rho(0)$  к сопротивлению в отсутствие поля  $\rho(0)$ :

$$H = \frac{\rho(\mathbf{B}) - \rho(0)}{\rho(0)} = \frac{\Delta\rho}{\rho(0)}, \quad (6.214)$$

где  $H$  - кинетический коэффициент магнетосопротивления.

Физической причиной эффекта магнетосопротивления является распределение носителей заряда по скорости и анизотропия времени релаксации и эффективной массы носителей, когда изоэнергетические поверхности несферичны. Действительно, если бы все носители заряда двигались с одной средней дрейфовой скоростью, имели одинаковую длину свободного пробега и эффективную изотропную массу, то в результате компенсации магнитной силы Лоренца  $e[\mathbf{v}_d, \mathbf{B}]$  холловским полем, носители двигались бы прямолинейно, без изменения длины свободного пробега. Так как при неизменной длине пробега скорость носителя не изменяется, то, следовательно, не меняется и электропроводность образца. Однако электроны имеют определенное распределение по энергиям и скоростям, и поэтому при включении магнитного поля траектория их движения уже не будет прямолинейной, поскольку указанная компенсация будет выполняться только для носителей со средней скоростью. На носители со скоростью меньшей, чем средняя, больше действует электрическое холловское поле  $\mathbf{E}_x$ , в то же время на носители со скоростями большими, чем средняя скорость, сильнее действует магнитная составляющая силы Лоренца  $e[\mathbf{v}_d, \mathbf{B}]$ , отклоняющая носители в противоположную сторону.

И в том, и другом случае отклонения скорости движения носителя от средней величины эффективная длина свободного пробега в магнитном поле уменьшится. Действительно, если носитель заряда в отсутствие магнитного поля ( $\mathbf{B} = 0$ ) имел траекторию  $l$  (см. рис. 6.21), то при включении магнитного поля и возникновении угла Холла между скоростью носителя без поля  $\mathbf{v}_0$  и с полем  $\mathbf{v}$ , его траектория изменяется. Следовательно, за время свободного пробега носитель вдоль поля  $\mathbf{E}$  пройдет меньший путь, чем  $l$ .

$$l_x = l \cos \varphi \approx l \left(1 - \varphi^2 / 2\right) = l \left(1 - \mu^2 B^2 / 2\right). \quad (6.215)$$

Если учесть, что подвижность пропорциональна длине пути на длине свободного пробега в направлении поля  $\mathbf{E}$ , а электропроводность определяется подвижностью, то уменьшение пути  $l$  вдоль поля снизит электропроводность.

Такой разброс носителей по скоростям уменьшит вклад быстрых и медленных носителей заряда в электропроводность в магнитном поле. Для примесных полупроводников с одним типом носителей и учетом статистического распределения по скоростям теория дает следующее выражение для кинетического коэффициента магнетосопротивления:

$$H = C \mu_{n,p} B^2, \quad (6.216)$$

где  $C$  — коэффициент, зависящий от механизма рассеяния, и, следовательно, также определяемый степенной зависимостью времени релаксации (6.120) или длины свободного пробега (6.122) от энергии.

Например, при рассеянии носителей заряда на акустических фонах ( $p = -1/2$ )  $C = \pi/10$ , а при рассеянии на ионах примеси ( $p = 3/2$ )  $C = 1$ .

Как следует из формулы (6.215), измерение сопротивления в магнитном поле позволяет найти подвижность носителей, если известен механизм рассеяния.

*Эффект Эттингсгаузена* состоит в возникновении в образце градиента температуры в направлении, перпендикулярном магнитному полю  $\mathbf{E}$  и вектору плотности тока  $\mathbf{j}$ , текущего в направлении приложенного электрического поля  $\mathbf{E}$ :

$$\nabla T = A_E [\mathbf{j}\mathbf{B}]. \quad (6.217)$$

или

$$\nabla_z T = A_E \mathbf{B}_x \mathbf{j}_x. \quad (6.218)$$

Для невырожденного полупроводника  $n$ -типа кинетический коэффициент эффекта Эттингсгаузена имеет вид:

$$A_E = \frac{k_0 T}{e} \frac{2r-1}{2} \frac{R_\sigma}{\chi_q + \chi_e}. \quad (6.219)$$

Физический смысл эффекта Эттингсгаузена обусловлен теми же причинами, что и магнетосопротивление: разным действием холловского поля и магнитной силы Лоренца на быстрые и медленные носители заряда. Как уже мы рассматривали, для носителей, скорость которых больше средней, магнитная составляющая силы Лоренца  $e[\mathbf{v}_d, \mathbf{B}]$  превышает действие поперечного электрического холловского поля  $E_x$ , и они будут отклоняться к одной из граней образца в направлении  $z$ . Для носителей заряда, скорость которых меньше средней, холловское поле будет действовать сильнее, чем сила Лоренца, и они отклонятся к противоположной грани вдоль направления  $z$ . Очевидно, что более быстрые носители, скапливаясь у одной из граней, будут отдавать свою энергию кристаллической решетке, и эта грань будет нагреваться. Противоположная грань, на которой скапливаются медленные носители, будет охлаждаться, так как, приходя к термодинамическому равновесию, носители будут забирать энергию у кристаллической решетки. Таким образом, вдоль направления  $z$  поперек магнитного и электрического полей наряду с поперечным полем Холла возникнет поперечный градиент температур  $\nabla_z T$ .

*Эффект Нернста* также связан с различным действием холловского поля и магнитной силы Лоренца на носители заряда, движущиеся с разной скоростью. Более медленные носители будут сильнее закручиваться магнитным полем, чем быстрые. Вследствие разного отклонения «горячих» и «холодных» носителей будет изменяться их вклад в энергию, переносимую вдоль электрического поля в продольном направлении  $x$ . Более быстрые («горячие») носители будут накапливаться у той грани вдоль направления  $x$ , к которой они движутся, и нагревать ее. Более медленные носители из-за замедления их магнитным полем будут собираться у противоположной грани вдоль направления, вызывая ее охлаждение. Таким образом, при отсутствии теплового потока вдоль направления электрического поля  $\mathbf{E}$  возникает продольный градиент температур

$$\nabla_x T = A_N \mathbf{B}_y \mathbf{j}_x, \quad (6.220)$$

где  $A_N$  — кинетический коэффициент эффекта Нернста.

Заметим, что знак продольного градиента температур не зависит от направления магнитного поля, но меняется при изменении направления тока.

Гальваномагнитные эффекты рассматриваются более строго на основе кинетического уравнения Больцмана.