

Отношение числа электронов n_i с энергией ϵ_i к общему числу состояний с данной энергией g_i называется функцией распределения Ферми и обозначается f_i . Иными словами, f_i — это вероятность того, что состояние с энергией ϵ_i занято электроном. Напротив, вероятность того, что данное состояние свободно, будет

$$f'_i = 1 - f_i.$$

В статистической физике доказывается, что введенная нами постоянная μ^* связана с химическим потенциалом μ соотношением

$$\mu = \mu^* kT \quad \text{или} \quad \mu^* = \frac{\mu}{kT}. \quad (4.24)$$

В соответствии с этим величина μ^* называется приведенным химическим потенциалом (μ^* есть безразмерный химический потенциал, выраженный в единицах kT). Аналогично величина $x = \epsilon/kT$ называется приведенной энергией электрона. В дальнейшем мы будем индекс i опускать, так как f не зависит от введенной нами произвольной нумерации состояний, а целиком определяется энергией данного состояния ϵ .

Итак, мы доказали, что вероятность того, что состояние с энергией ϵ занято электроном, выражается формулой

$$f = \frac{1}{e^{\frac{\epsilon - \mu}{kT}} + 1}. \quad (4.25)^*$$

4.3. СТАТИСТИКА НЕВЫРОЖДЕННОГО ЭЛЕКТРОННОГО ГАЗА В ПОЛУПРОВОДНИКАХ

В первой главе уже упоминалось, что общее число состояний в свободной зоне твердого тела составляет примерно 10^{22} (на 1 см^3). Число же свободных электронов в полупроводниках колеблется обычно в пределах 10^{12} — 10^{18} см^{-3} . Это значит, что в отличие от металла доля занятых состояний в полупроводнике обычно ничтожно мала или, иными словами, обычно $f \ll 1$ для всех электронных состоя-

*) В тех случаях, когда необходимо подчеркнуть, что f — это равновесная функция распределения, мы будем отмечать ее индексом 0.

ний, включая самые нижние, примыкающие ко дну зоны * (мы в дальнейшем для удобства примем нуль отсчета энергии совпадающим с дном свободной зоны, т. е. для этих электронов будем считать $\varepsilon = 0$). Но если $f \ll 1$, то это значит, что знаменатель (4.25) $1 + e^{\frac{\varepsilon - \mu}{kT}} \gg 1$ или $\frac{\varepsilon - \mu}{kT} \gg 0$, т. е. уровень химического потенциала проходит значительно ниже дна зоны. В этом случае для всех состояний в зоне проводимости мы можем пренебречь единицей в знаменателе выражения (4.25) и, следовательно,

$$f \approx e^{\frac{\mu - \varepsilon}{kT}}. \quad (4.26)$$

Приближенное выражение (4.26) называется распределением Максвелла — Больцмана. Такое состояние электронного газа, в котором $f \ll 1$ и которое описывается приближенной формулой (4.26), называется невырожденным. Мы увидим несколько позже, что в этом случае в отличие от вырожденного состояния средняя энергия электронов растет пропорционально температуре (и это является одним из основных признаков отсутствия вырождения).

Замена точного выражения для функции распределения (4.25) приближением (4.26) значительно упрощает все расчеты, целый ряд вычислений в этом случае можно в общем виде довести до конца.

Подсчитаем сначала число электронов в зоне проводимости полупроводника, энергии которых находятся в интервале $\varepsilon, \varepsilon + d\varepsilon$. Для этого мы должны умножить число состояний в этом интервале $g(\varepsilon) d\varepsilon$ на вероятность их заполнения f :

$$dn = f(\varepsilon) g(\varepsilon) d\varepsilon. \quad (4.27)$$

Мы знаем, что общее число состояний в любой зоне кристалла (объемом 1 см^3) равно числу атомов N_0 , умноженному на кратность g_a атомного уровня, образовавшего данную зону; следовательно,

$$\int_{\varepsilon_1}^{\varepsilon_2} g(\varepsilon) d\varepsilon = g_a N_0, \quad (4.28)$$

*) Бывают, впрочем, случаи, когда условие $f \ll 1$ для самых нижних уровней зоны и в полупроводнике не выполняется, такие полупроводники называются вырожденными (см. ниже).

где \mathcal{E}_1 и \mathcal{E}_2 — энергия нижнего и верхнего края зоны, но конкретный вид самой функции $g(\epsilon)$ нам неизвестен.

Нахождение точного вида этой функции является очень сложной задачей. Однако если учесть, что функция распределения f очень круто спадает по мере удаления (вверх) от дна зоны проводимости, то мы можем удовлетвориться знанием величины $g(\epsilon)$ вблизи дна зоны. Для таких значений энергии интересующая нас функция относительно легко вычисляется, как мы это уже делали для фононов:

$$dN = g(\epsilon) d\epsilon = \frac{8\pi p^2 dp}{h^3}. \quad (4.29)$$

Выражая в (4.29) импульс через энергию $p = \sqrt{2m_n\epsilon}$ и $p dp = m_n d\epsilon$, получаем

$$dN = \frac{4\pi}{h^3} (2m_n)^{3/2} \epsilon^{1/2} d\epsilon.$$

Таким образом, в случае одной простой сферической зоны получаем

$$g(\epsilon) = \frac{4\pi}{h^3} (2m_n)^{3/2} \epsilon^{1/2}. \quad (4.30)$$

Точно так же можно показать, что в случае анизотропной эффективной массы

$$g(\epsilon) = \frac{4\pi}{h^3} 2^{3/2} \sqrt{m_1 m_2 m_3} \epsilon^{1/2}, \quad (4.31)$$

где m_1, m_2, m_3 — главные значения тензора эффективной массы.

Наконец, в случае нескольких (N) эллипсоидов

$$g(\epsilon) = \frac{4\pi}{h^3} 2^{3/2} N \sqrt{m_1 m_2 m_3} \epsilon^{1/2} \quad (4.32)$$

можно все эти случаи объединить в один, если ввести понятие эффективной массы плотности состояний ($m_{\text{пс}}$) и определить эту массу соотношением

$$m_{\text{пс}} = N^{2/3} \sqrt[3]{m_1 m_2 m_3}. \quad (4.33) *$$

*) Все приведенные выше формулы для плотности состояний верны лишь для параболических зон, т. е. таких, в которых энергию можно представить в виде $p^2/2m$, где эффективная масса $m = \text{const}$ не зависит от энергии. За последнее время исследован целый ряд полупроводников (с узкой запрещенной зоной), в которых это условие не выполняется.

Тогда для всех случаев, рассмотренных выше,

$$g(\varepsilon) = \frac{4\pi}{h^3} (2m_{nc})^{3/2} \varepsilon^{1/2}. \quad (4.34)$$

Зная обе функции, $g(\varepsilon)$ и $f(\varepsilon)$, мы можем теперь получить общее число электронов в зоне проводимости

$$n = \int_{\mathcal{E}_1}^{\mathcal{E}_2} f(\varepsilon) g(\varepsilon) d\varepsilon \approx \int_0^{\infty} \frac{4\pi}{h^3} (2m_n)^{3/2} \varepsilon^{1/2} e^{\frac{\mu - \varepsilon}{kT}} d\varepsilon$$

(верхний предел интегрирования мы вправе продлить до бесконечности, так как $f(\varepsilon)$ с ростом ε очень быстро убывает).

Для того чтобы вычислить n , вынесем из-под интеграла все величины, не зависящие от энергии, и превратим подынтегральное выражение в безразмерное, для чего поделим и умножим его на $(kT)^{3/2}$. Тогда выражение для n примет вид

$$n = \frac{4\pi}{h^3} (2m_n kT)^{3/2} e^{\frac{\mu}{kT}} \int_0^{\infty} e^{-x} x^{1/2} dx = \frac{2 (2\pi m_n kT)^{3/2} e^{\frac{\mu}{kT}}}{h^3}, \quad (4.35)^*$$

откуда

$$\mu^* = \frac{\mu}{kT} = \ln \frac{h^3 n}{2 (2\pi m_n kT)^{3/2}}. \quad (4.35a)$$

Аналогичные вычисления дают число дырок p в заполненной зоне

$$p = \frac{2 (2\pi m_p kT)^{3/2} e^{\frac{\mu'}{kT}}}{h^3}, \quad (4.36)$$

где m_p — эффективная масса дырки и μ' — расстояние от уровня химического потенциала до верхнего края заполненной зоны:

$$\mu + \mu' = -\Delta\mathcal{E}_0 \quad \text{или} \quad \mu' = -\Delta\mathcal{E}_0 - \mu. \quad (4.37)$$

*) Так как $\int_0^{\infty} e^{-x} x^{1/2} dx = \frac{1}{2} \sqrt{\pi}$.

Следовательно, мы можем также записать (4.36) в виде

$$p = \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{-\frac{\Delta \mathcal{E}_0}{kT}} e^{-\frac{\mu}{kT}}. \quad (4.36a)$$

Значительно проще вычисляется число электронов n_d на донорных уровнях; согласно сказанному выше оно равно числу этих уровней, умноженному на функцию распределения:

$$n_d = \frac{N_d}{1 + e^{\frac{-\Delta \mathcal{E}_1 - \mu}{kT}}}, \quad (4.38)$$

где $\Delta \mathcal{E}_1$ — энергия активации для отрыва электрона с этого уровня.

Выведенные выше выражения показывают, что если мы будем знать μ , то сможем определить число электронов в зоне проводимости на локальных уровнях и число дырок в заполненной зоне. Но и сама величина μ вычисляется из того условия, что общее число электронов в системе нам известно. Действительно, электроны в свободной зоне появляются по двум причинам: за счет перебросов с донорных уровней и за счет перебросов из заполненной зоны. Следовательно, число свободных электронов должно равняться сумме числа дырок в заполненной зоне и числа свободных мест на донорных уровнях (числа «связанных дырок») ($N_d - n_d$):

$$n = p + (N_d - n_d) \text{ или } n - p + n_d = N_d. \quad (4.39)$$

Подставляя в (4.39) выражения для n и p из (4.35), (4.37) и (4.38), получаем уравнение, из которого может быть определено μ :

$$\begin{aligned} \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}} - \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{-\frac{\Delta \mathcal{E}_0}{kT}} e^{-\frac{\mu}{kT}} + \\ + \frac{N_d}{e^{\frac{-\Delta \mathcal{E}_1 - \mu}{kT}} + 1} = N_d. \end{aligned} \quad (4.40)$$

Однако легко убедиться, что (4.40) представляет собой уравнение третьей степени относительно $e^{-\mu/kT}$. Поэтому обычно его не решают в общем виде, а ограничиваются рассмотрением частных случаев *).

ОБЛАСТЬ ПРИМЕСНОЙ ПРОВОДИМОСТИ

Предположим, что температура недостаточно высока для того, чтобы в заметном количестве наблюдались перебросы электронов из заполненной зоны, т. е. что электроны в зоне проводимости появляются в основном за счет переходов с примесных уровней, а число дырок в заполненной зоне мы можем пренебречь. Тогда уравнение (4.39) заменится более простым:

$$n + n_d = N_d \quad (4.41)$$

или согласно (4.35) и (4.38)

$$\frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}} + \frac{N_d}{e^{-\frac{\Delta \mathcal{E}_1 - \mu}{kT}} + 1} = N_d. \quad (4.42)$$

Уравнение (4.42) — уже квадратное уравнение относительно $e^{\frac{\mu}{kT}}$. Решая его, получаем

$$\mu = -\Delta \mathcal{E}_1 + kT \ln \frac{1}{2} \left[\sqrt{1 + \frac{2e^{\frac{\Delta \mathcal{E}_1}{kT}} N_d h^3}{(2\pi m_n kT)^{3/2}}} - 1 \right]. \quad (4.43)$$

Выражение (4.43) все еще слишком сложно для его непосредственной интерпретации, поэтому рассмотрим его предельные значения в различных интервалах температур.

Область низких температур:

$$e^{\frac{\Delta \mathcal{E}_1}{kT}} \gg \frac{2(2\pi m_n kT)^{3/2}}{N_d h^3}. \quad (4.44)$$

*) Еще более сложный вид принимает выражение для определения химического потенциала при наличии нескольких сортов донорных уровней (с различной энергией активации) или доноров и акцепторов.

Пренебрегая единицами под знаком логарифма, получаем

$$\mu \approx -\frac{\Delta\mathcal{E}_1}{2} + kT \ln \left[\frac{N_d h^3}{2 (2\pi m_n kT)^{3/2}} \right]^{1/2}. \quad (4.45)$$

Выражение (4.45) показывает, что при абсолютном нуле температуры уровень химического потенциала проходит строго посередине между дном запрещенной зоны и донорными уровнями. Второй член в (4.45) при низких температу-

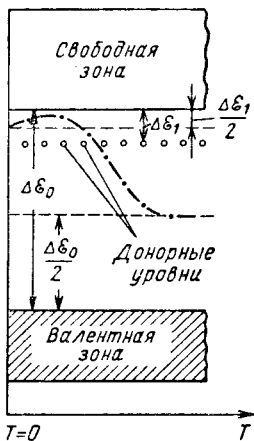


Рис. 4.1. Температурная зависимость уровня химического потенциала в электронном полупроводнике.

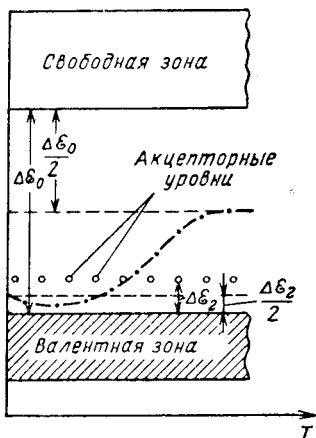


Рис. 4.2. Температурная зависимость уровня химического потенциала в дырочном полупроводнике.

рах [т. е. пока $(2\pi m_n kT)^{3/2} < N_d h^3$] положительный, а при более высоких становится отрицательным. Поэтому уровень химического потенциала сначала поднимается, а затем начинает опускаться (рис. 4.1). Подставляя (4.45) в (4.35), получаем концентрацию свободных электронов:

$$n = (2N_d)^{1/2} \frac{(2\pi m_n kT)^{3/4}}{h^{3/2}} e^{-\frac{\Delta\mathcal{E}_1}{2kT}} = A e^{-\frac{\mathcal{E}_1}{2kT}}, \quad (4.46)$$

где $A = (2N_d)^{1/2} \frac{(2\pi m_n kT)^{3/4}}{h^{3/2}}$ — величина, слабо зависящая от температуры (по сравнению со вторым экспоненциаль-

ным множителем). Выражение (4.46) показывает, что (в этом интервале температур) концентрация электронов в зоне проводимости растет экспоненциально с ростом температуры. Прологарифмировав (4.46), получим

$$\ln n = \ln A - \frac{\Delta \mathcal{E}_1}{2k} \cdot \frac{1}{T}. \quad (4.47)$$

Следовательно, если мы будем на графике откладывать по оси ординат $\ln n$, а по оси абсцисс — обратную температуру ($1/T$), то мы получим прямую с наклоном $\Delta \mathcal{E}_1/2k$.

Область «высоких» температур:

$$e^{\frac{\Delta \mathcal{E}_1}{kT}} \ll \frac{2(2\pi m_n kT)^{3/2}}{h^3 N_d}. \quad (4.48)$$

Термин «высокие температуры» здесь следует принимать лишь в смысле выполнения неравенства (4.48), но мы по-прежнему будем предполагать, что собственная проводимость еще не играет существенной роли. Учитывая (4.48) и разлагая радикал в выражении (4.43) в ряд, получаем

$$\mu \approx kT \ln \frac{N_d h^3}{2(2\pi m_n kT)^{3/2}}. \quad (4.49)$$

Согласно (4.48) и (4.49) в этом интервале температур уровень химического потенциала уже спустился ниже донорных уровней, и, следовательно, почти все они должны быть пусты.

Действительно, подставляя (4.49) в (4.35), получаем

$$n = N_d,$$

т. е. все (или точнее почти все, так как мы все время пользуемся приближенными формулами) электроны перешли в зону проводимости. Концентрация электронов проводимости в этом интервале температур будет оставаться постоянной, что соответствует горизонтальному участку кривой рис. 1.7.

Заметим в заключение, что совершенно аналогичные выводы можно получить для примесного дырочного полупроводника. В этом случае в области низких температур

$$\mu' \approx \frac{\Delta \mathcal{E}_2}{2} - kT \ln \frac{N_d h^3}{2(2\pi m_p kT)^{3/2}} \quad (4.50)$$

где $\Delta\mathcal{E}_2$ — энергия, необходимая для заброса электрона из заполненной зоны на акцепторный уровень, т. е. уровень химического потенциала проходит примерно посередине между верхним краем валентной зоны и акцепторным уровнем (рис. 4.2). При этом концентрация дырок меняется экспоненциально:

$$p = (2N_a)^{1/2} \frac{2 (2\pi m_p kT)^{3/4}}{h^{3/2}} e^{-\frac{\Delta\mathcal{E}_2}{kT}}. \quad (4.51)$$

В области более высоких температур уровень химического потенциала поднимается выше акцепторных уровней, и все они в силу этого оказываются забитыми электронами, а число дырок в заполненной зоне становится постоянным:

$$p = N_a.$$

При дальнейшем подъеме температур и в случае электронного и в случае дырочного полупроводника основную роль начинает играть собственная проводимость.

ОБЛАСТЬ СОБСТВЕННОЙ ПРОВОДИМОСТИ

Рассмотрим теперь второй предельный случай, т. е. предположим, что число перебросов электронов из заполненной зоны в свободную очень велико и по сравнению с этим примесные электроны играют второстепенную роль. Тогда мы можем принять

$$n = p. \quad (4.52)$$

Подставляя (4.35) и (4.36а) в (4.52), получаем

$$\mu = -\frac{\Delta\mathcal{E}_0}{2} + \frac{3}{2} kT \ln \frac{m_p}{m_n}, \quad (4.53)$$

т. е. если $m_n \approx m_p$ уровень химического потенциала проходит приблизительно посередине запрещенной зоны. Подставляя (4.53) в (4.35) и (4.36а), найдем концентрацию электронов и дырок:

$$n = p = \frac{2 (2\pi \sqrt{m_n m_p} kT)^{3/2}}{h^3} e^{-\frac{\Delta\mathcal{E}_0}{2kT}}. \quad (4.54)$$

В полулогарифмическом масштабе

$$\ln n = f\left(\frac{1}{T}\right), \quad (4.55)$$

$$\ln n = \ln A - \frac{\Delta \mathcal{E}_0}{2kT}, \quad (4.56)$$

где A приблизительно постоянно, таким образом, (4.56) изобразится прямой с наклоном $\Delta \mathcal{E}_0/2k$.

4.4. ЭНЕРГИЯ ЭЛЕКТРОНОВ В ЗОНЕ ПРОВОДИМОСТИ, ВЫРОЖДЕНИЕ

Остановимся коротко на термине «вырождение», который в теоретической физике имеет свой особый смысл.

Вырожденными в квантовой механике называются системы, энергия которых не зависит от какого-либо параметра. Если при каких-либо изменениях в условиях существования системы ее энергия начинает зависеть от этого параметра, то мы говорим, что «вырождение снимается».

ВЫРОЖДЕНИЕ В АТОМЕ

В водородном атоме, т. е. при кулоновской зависимости потенциальной энергии от расстояния до ядра ($U = -e^2/r$), энергия электрона не зависит от эксцентриситета орбиты (орбитального квантового числа l), а зависит только от размера ее большой полуоси — главного квантового числа n . Все подоболочки: p , d , f , s , внутри одной оболочке вырождены, т. е. имеют одну и ту же энергию. В более сложных атомах поле других электронов экранирует поле ядра; поэтому по мере удаления от ядра потенциальная энергия убывает быстрее, чем в атоме водорода. В непосредственной близости от ядра $U = -ze^2/r$ постепенно убывает до $U' = -e^2/r$ на больших расстояниях. Поэтому энергия валентных электронов тем меньше (или энергия связи тем больше), чем ближе данная орбита подходит к ядру. Таким образом, в результате взаимодействия электронов друг с другом вырождение снимается, энергия подоболочек в порядке возрастания располагается в следующей последовательности: s , p , d , f .

Второй тип вырождения, имеющегося в любом атоме, так называемое ориентационное. Благодаря шаровой симметрии атомного потенциала энергия электронов на всех