

НЕОДНОРОДНЫЕ ПОЛУПРОВОДНИКИ

ПОЛУКЛАССИЧЕСКОЕ РАССМОТРЕНИЕ
НЕОДНОРОДНЫХ ПОЛУПРОВОДНИКОВ
ПОЛЯ И КОНЦЕНТРАЦИИ НОСИТЕЛЕЙ
В РАВНОВЕСНОМ p — n -ПЕРЕХОДЕ
ЭЛЕМЕНТАРНАЯ КАРТИНА

ВЫПРЯМЛЯЮЩЕГО ДЕЙСТВИЯ p — n -ПЕРЕХОДА
ДРЕЙФОВЫЕ И ДИФФУЗИОННЫЕ ТОКИ
ВРЕМЕНА СТОЛКНОВЕНИЙ И РЕКОМБИНАЦИИ
ПОЛЯ, КОНЦЕНТРАЦИИ НОСИТЕЛЕЙ И ТОКИ
В НЕРАВНОВЕСНОМ p — n -ПЕРЕХОДЕ

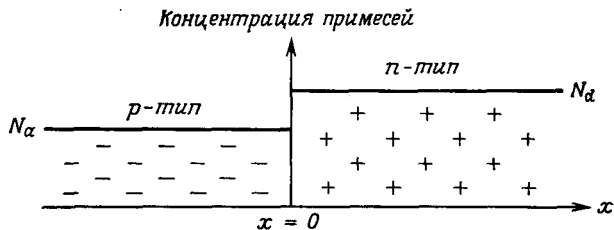
Когда слова «физика твердого тела» произносят любители музыки, то по сути дела они имеют в виду неоднородные полупроводники, и, строго говоря, именно эти слова должны украшать передние панели бесчисленных приемников и усилителей. Такое распространенное представление о физике твердого тела отражает тот факт, что наиболее широкое и яркое применение в технике она нашла благодаря своеобразным электронным свойствам полупроводниковых устройств. В этих устройствах используются полупроводниковые кристаллы, в которых специально создается неоднородное распределение концентрации донорных и акцепторных примесей. Мы не будем пытаться дать здесь описание огромного числа разнообразных полупроводниковых приборов, а остановимся только на важнейших физических принципах, на которых основана их работа. Этими принципами определяется распределение электронных и дырочных концентраций и токов в неоднородном полупроводнике как в отсутствие, так и при наличии приложенного электростатического потенциала.

Рассматриваемые нами неоднородные полупроводники представляют собой в идеальном случае монокристаллы, в которых локальная концентрация донорных и акцепторных примесей меняется от точки к точке. Один из способов приготовления таких кристаллов заключается в изменении концентрации примесей в расплаве по мере того, как из него медленно вытягивается растущий кристалл; при этом концентрация примесей меняется вдоль заданного пространственного направления. Необходимы тонкие методы выращивания кристаллов, поскольку обычно для эффективной работы устройств весьма существенно, чтобы в кристалле не происходило сильного увеличения рассеяния электронов на флуктуациях концентрации примесей.

Чтобы составить представление о физике неоднородных полупроводников, рассмотрим в качестве простейшего примера p — n -переход. Это полупроводниковый кристалл, в котором концентрация примесей изменяется только вдоль данного направления (принятого за ось x) в узкой области (вблизи $x = 0$). При отрицательных x в кристалле преобладают акцепторные примеси (т. е. он относится к p -типу), а при положительных x — донорные примеси (т. е. он относится к n -типу) (фиг. 29.1). Распределение концентраций доноров $N_d(x)$ и акцепторов $N_a(x)$ называется «профилем легирования». Термин «переход» исполь-

зается как для обозначения устройства в целом, так и в более узком смысле — для обозначения переходной области вблизи $x = 0$, где профиль легирования неоднороден.

Как мы увидим ниже, неоднородное распределение примесей вызывает неоднородное распределение концентрации $n_c(x)$ электронов в зоне проводимости и концентрации $p_v(x)$ дырок в валентной зоне, что в свою очередь приводит к появлению неоднородного потенциала $\phi(x)$. Область, где указанные концентрации носителей неоднородны, называется «обедненным слоем» (или областью пространственного заряда). Как мы увидим ниже, обедненный слой



Фиг. 29.1. Распределение концентраций примесей вдоль $p - n$ -перехода в случае «резкого перехода», в котором донорные примеси доминируют при положительных x , а акцепторные — при отрицательных x .

Доноры обозначены знаками (+), а акцепторы — знаками (-), что указывает на знак их заряда, когда они ионизованы. Чтобы переход был резким, область вблизи $x = 0$, где меняются концентрации примесей, должна быть узкой по сравнению с обедненным слоем, в котором распределение заряда неоднородно. (На фиг. 29.3 изображены типичные кривые изменения концентраций носителей, наложенные на это распределение.)

может занимать область шириной примерно от 10^2 до 10^4 Å вокруг (обычно более узкой) переходной области, где меняется профиль легирования. Везде внутри обедненного слоя, кроме непосредственной окрестности его границ, полная концентрация носителей значительно ниже, чем в однородных областях, находящихся на большем расстоянии от переходной области. Существование обедненного слоя — одно из главных свойств свойств $p - n$ -перехода. Одна из наших основных задач заключается в объяснении того, как такой слой образуется при изменении концентрации примесей и как изменяется его структура при наложении внешнего потенциала V .

Для простоты будем рассматривать только «резкие переходы», в которых переходная область столь узка, что изменения концентрации примесей ¹⁾ могут быть представлены в виде скачков при $x = 0$:

$$N_d(x) = \begin{cases} N_d, & x > 0, \\ 0, & x < 0, \end{cases} \quad (29.1)$$

$$N_a(x) = \begin{cases} 0, & x > 0, \\ N_a, & x < 0. \end{cases}$$

Скачкообразные переходы не только наиболее просты с точки зрения их описания, но и представляют наибольший практический интерес. Из проведенного ниже рассмотрения станет ясным, насколько узкой должна быть в действи-

¹⁾ Совершенно не обязательно, чтобы в области n -типа существовали только донорные примеси, а в области p -типа — только акцепторные. Достаточно, чтобы каждый из типов примесей был определяющим в своей области. В дальнейшем величину N_d можно рассматривать как избыток концентрации доноров по сравнению с концентрацией акцепторов, а N_a — как избыток акцепторов по сравнению с донорами.

тельности переходная область, чтобы условие (29.1) давало разумную модель реального $p - n$ -перехода. Мы найдем, что переход может считаться резким, если ширина переходной области в профиле легирования мала по сравнению с шириной обедненного слоя. В большинстве случаев это условие допускает существование переходных областей шириной 100 Å и более. Переход, который нельзя рассматривать как резкий, называют плавным переходом.

ПОЛУКЛАССИЧЕСКАЯ МОДЕЛЬ

Для вычисления отклика неоднородного полупроводника на приложенный внешний электростатический потенциал и даже для расчета распределения электрического заряда в отсутствие приложенного потенциала почти всегда используют полуклассическую модель, описанную в гл. 12. Когда потенциал $\phi(x)$ налагается на периодический потенциал кристалла, электроны n -й зоны можно рассматривать (в полуклассической модели) как классические частицы (т. е. волновые пакеты), описываемые гамильтонианом

$$H_n = \mathcal{E}_n(\mathbf{p}/\hbar) - e\phi(x). \quad (29.2)$$

Такое рассмотрение имеет смысл, если потенциал $\phi(x)$ изменяется достаточно медленно. Вообще говоря, очень трудно ответить на вопрос, насколько медленным должно быть это изменение. По меньшей мере надо потребовать, чтобы изменение электростатической энергии $e\Delta\phi$ на расстояниях порядка постоянной решетки было мало по сравнению с шириной запрещенной зоны E_g , но вполне возможно, что условие должно быть даже более жестким¹⁾. В случае $p - n$ -перехода потенциал ϕ претерпевает изменения практически только в пределах обедненного слоя. В этом слое энергия $e\phi$ меняется примерно на E_g на расстоянии, которое обычно составляет несколько сотен ангстрем или более (так что поле в обедненном слое может достигать 10^6 В/м). Хотя при этом выполняется минимальное необходимое условие применимости полуклассической модели (изменение $e\phi$ на межатомном расстоянии составляет не более долей процента от E_g), изменение все же достаточно сильное и нельзя исключить возможной неадекватности полуклассического описания в обедненном слое. Поэтому не следует упускать из виду того обстоятельства, что поле в обедненном слое может оказаться достаточно сильным, чтобы вызвать туннелирование электронов с уровней валентной зоны на уровни зоны проводимости, приводя к проводимости, намного превышающей значение, предсказываемое полуклассической теорией.

Сделав это предупреждение, мы, однако, будем, как обычно, считать применимым полуклассическое описание, чтобы иметь возможность использовать его следствия. Прежде чем рассматривать полуклассическую теорию токов в $p - n$ -переходе, изучим случай термодинамического равновесия, когда внешние напряжения и токи отсутствуют.

РАВНОВЕСНЫЙ $p - n$ -ПЕРЕХОД

Определим концентрации носителей и электростатический потенциал $\phi(x)$, обусловленный неоднородным легированием. Предположим, что во всем материале выполняется условие невырожденности, и поэтому концентраций носителей в каждой точке x имеют максвелловскую форму, как и концентрации (28.12)

¹⁾ Грубое рассмотрение применительно к металлам см. в приложении К. Аналогичное исследование (столь же грубое) может быть проведено и для полупроводников.

в однородном случае. В неоднородном случае полуклассический вывод выражения для концентраций носителей в точке x в области перехода при наличии потенциала $\phi(x)$ сводится к повторению соответствующего вывода для однородного случая, однако при этом надо использовать полуклассическую формулу для одноэлектронной энергии (29.2), в которой каждый уровень сдвинут на $-e\phi(x)$. Используя выражение (28.3) для энергии $\xi(\mathbf{k})$, применимое при описании уровней вблизи минимума зоны проводимости или же максимума валентной зоны, мы видим, что в результате значения констант ξ_c и ξ_v изменяются просто на $-e\phi(x)$. Поэтому вместо выражений (28.12) для равновесных плотностей носителей в этом случае мы имеем более общие выражения:

$$\begin{aligned} n_c(x) &= N_c(T) \exp \left\{ -\frac{[\xi_c - e\phi(x) - \mu]}{k_B T} \right\}, \\ p_v(x) &= P_v(T) \exp \left\{ -\frac{[\mu - \xi_v + e\phi(x)]}{k_B T} \right\}. \end{aligned} \quad (29.3)$$

Потенциал $\phi(x)$ следует определять самосогласованным образом (с помощью уравнения Пуассона) как потенциал, возникающий при концентрациях носителей, описываемых выражениями (29.3). Рассмотрим эту задачу для частного случая (снова наиболее интересного в практическом отношении), когда по обе стороны от переходной области на большом расстоянии от нее преобладает примесная проводимость, а примеси полностью «ионизованы» (стр. 205—206). Таким образом, вдали от перехода в n -области образца концентрация электронов в зоне проводимости почти равна концентрации доноров N_d , а в p -области концентрация дырок в валентной зоне почти равна концентрации акцепторов N_a :

$$\begin{aligned} N_d = n_c(\infty) &= N_c(T) \exp \left\{ -\frac{[\xi_c - e\phi(\infty) - \mu]}{k_B T} \right\}, \\ N_a = p_v(-\infty) &= P_v(T) \exp \left\{ -\frac{[\mu - \xi_v + e\phi(-\infty)]}{k_B T} \right\}. \end{aligned} \quad (29.4)$$

Поскольку весь кристалл находится в состоянии термодинамического равновесия, химический потенциал одинаков во всех точках. В частности, в формулы (29.4) входит одно и то же значение μ . Отсюда непосредственно следует, что полное изменение потенциала на переходе описывается формулой¹⁾

$$e\phi(\infty) - e\phi(-\infty) = \xi_c - \xi_v + k_B T \ln \left[\frac{N_d N_a}{N_c P_v} \right], \quad (29.5)$$

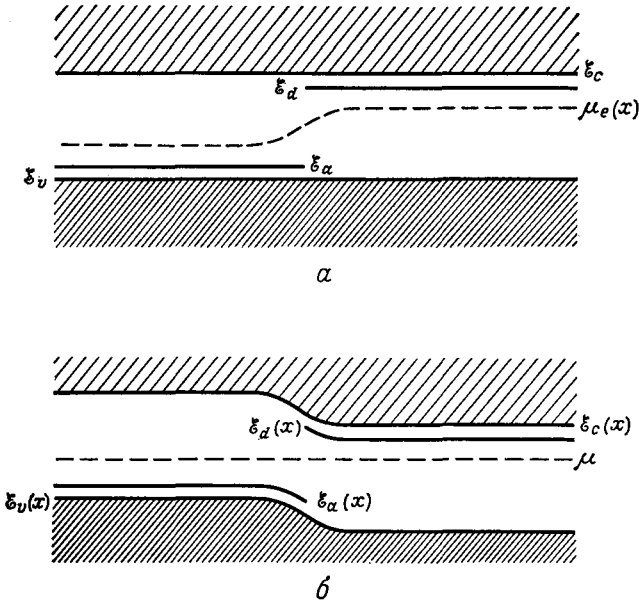
или

$$\boxed{e\Delta\phi = E_g + k_B T \ln \left[\frac{N_d N_a}{N_c P_v} \right]}. \quad (29.6)$$

Иногда полезен другой способ записи формул (29.3) и (29.6). Если ввести зависящий от пространственных координат «электрохимический потенциал»

$$\mu_e(x) = \mu + e\phi(x), \quad (29.7)$$

¹⁾ Для вывода соотношения (29.5) необходимо, чтобы соотношения (29.3) выполнялись только вдали от обедненного слоя, где потенциал ϕ изменяется медленно. Поэтому соотношение (29.5) оказывается справедливым, даже если полуклассическая модель неприменима в переходном слое.



Фиг. 29.2. Два эквивалентных способа описания влияния внутреннего потенциала $\phi(x)$ на концентрацию электронов и дырок в $p-n$ -переходе.

a — изменение электрохимического потенциала $\mu_e(x) = \mu + e\phi(x)$ вдоль $p-n$ -перехода. Концентрации носителей в любой точке x имеют такую же величину, как в однородном полупроводнике, который характеризуется фиксированными значениями энергий зон и примесных уровней ϵ_c , ϵ_v , ϵ_d и ϵ_a и химическим потенциалом $\mu_e(x)$.

b — здесь величина $\epsilon_c(x) = \epsilon_c - e\phi(x)$ представляет собой энергию локализованного вблизи x электронного волнового пакета, образованного из уровней, ближайших к дну зоны проводимости; аналогично определяется величина $\epsilon_v(x)$. Энергии локальных примесных уровней имеют значения $\epsilon_d(x) = \epsilon_d - e\phi(x)$ и $\epsilon_a(x) = \epsilon_a - e\phi(x)$. Показан также химический потенциал (имеющий постоянное значение). Концентрации носителей в любой точке x имеют такую же величину, как в однородном полупроводнике, характеризуемом значениями энергий зон и примесных уровней, равными $\epsilon_c(x)$, $\epsilon_d(x)$, $\epsilon_a(x)$ и $\epsilon_v(x)$ при фиксированном химическом потенциале μ .

то можно записать выражения для концентраций носителей (29.3) в виде

$$\begin{aligned} n_c(x) &= N_c(T) \exp \left\{ -\frac{[\epsilon_c - \mu_e(x)]}{k_B T} \right\}, \\ p_v(x) &= P_v(T) \exp \left\{ -\frac{[\mu_e(x) - \epsilon_v]}{k_B T} \right\}. \end{aligned} \quad (29.8)$$

Выражения (29.8) имеют в точности такой же вид, как и выражения (28.12) для однородных полупроводников, за исключением того, что постоянный химический потенциал μ заменяется на электрохимический потенциал $\mu_e(x)$. Таким образом, величина $\mu_e(\infty)$ представляет собой химический потенциал однородного кристалла n -типа, обладающего такими же свойствами, как и неоднородный кристалл в n -области вдали от перехода, а $\mu_e(-\infty)$ — химический потенциал однородного кристалла p -типа, аналогичного p -области неоднородного кристалла вдали от перехода. Соотношение (29.6) можно записать эквивалентным образом ¹⁾:

$$e \Delta\phi = \mu_e(\infty) - \mu_e(-\infty). \quad (29.9)$$

¹⁾ Это непосредственно следует из соотношения (29.7). Соотношение (29.9) иногда формулируют как правило, согласно которому полное изменение потенциала таково, чтобы «уровни Ферми совпадали на обоих концах перехода». Такая формулировка, очевидно, основывается на представлении, которому соответствует фиг. 29.2, б.

На фиг. 29.2, *a* электрохимический потенциал представлен в зависимости от координаты вдоль p — n -перехода. Мы предположили (ниже это будет показано), что ϕ изменяется монотонно по мере перехода от одного конца к другому. На фиг. 29.2, *b* та же информация представлена другим способом. Здесь принято, что потенциал ϕ , определяющий координатную зависимость в выражении (29.3), изменяет величину ξ_c (или ξ_v), а не μ . В любом случае смысл этих диаграмм заключается в том, что в каждой конкретной точке x вдоль перехода концентрация носителей имеет такую же величину, как и в куске однородного материала с концентрацией примесей, равной тому значению, которое она имеет в точке x , и с химическим потенциалом, положение которого по отношению к краям зоны можно определить, рассматривая вертикальное сечение диаграммы в точке x .

Соотношение (29.6) [или его эквивалентная форма (29.9)] представляет собой граничное условие для дифференциального уравнения, определяющего потенциал $\phi(x)$. Это дифференциальное уравнение есть просто уравнение Пуассона ¹⁾

$$-\nabla^2\phi = -\frac{d^2\phi}{dx^2} = \frac{4\pi\rho(x)}{\epsilon}, \quad (29.10)$$

которое связывает потенциал $\phi(x)$ с создающим его распределением заряда $\rho(x)$. Чтобы выразить $\rho(x)$ через ϕ и получить замкнутое уравнение, отметим вначале, что если (согласно нашему предположению) примеси полностью ионизованы вдали от перехода, то они полностью ионизованы и при всех x ²⁾. Следовательно, плотность заряда, обусловленная примесями и носителями, равна ³⁾

$$\rho(x) = e[N_d(x) - N_a(x) - n_c(x) + p_v(x)]. \quad (29.11)$$

Подставляя выражения (29.3) и (29.1) для концентраций носителей и примесей в формулу (29.11) для плотности заряда, а затем результат в уравнение Пуассона, получаем нелинейное дифференциальное уравнение для $\phi(x)$, для точного решения которого обычно приходится применять численные методы ⁴⁾. Однако можно получить вполне разумное описание потенциала $\phi(x)$, используя тот факт, что полное изменение $e\phi$ имеет порядок $E_g \gg k_B T$. Значение этого факта легко понять, если объединить формулы (29.3) и (29.4), записав

$$\begin{aligned} n_c(x) &= N_d e^{-[\phi(\infty) - \phi(x)]/k_B T}, \\ p_v(x) &= N_a e^{-[\phi(x) - \phi(-\infty)]/k_B T}. \end{aligned} \quad (29.12)$$

Предположим, что изменение ϕ происходит внутри области $-d_p \leq x \leq d_n$. Вне этой области ϕ достигает своего асимптотического значения, и, следова-

¹⁾ Здесь ϵ — статическая диэлектрическая проницаемость полупроводника. Использование макроскопического уравнения возможно потому, что ϕ меняется на расстоянии порядка ширины обедненного слоя, которая велика по сравнению с межатомным расстоянием.

²⁾ Если потенциал ϕ (как показано ниже) изменяется монотонно, то это утверждение вытекает из того факта, что степень ионизации примеси увеличивается при удалении химического потенциала от примесного уровня. См. фиг. 29.2 и формулы (28.32) и (28.34).

³⁾ Концентрация дырок на больших расстояниях от перехода в n -области имеет, согласно закону действующих масс, очень малое значение $p_v(\infty) = n_i^2/N_d$. Однако, концентрация электронов превышает там N_d на ту же малую величину, так что выполняется равенство $n_c(\infty) - p_v(\infty) = N_d$. Если при вычислении полной плотности заряда пренебречь [как мы сделали в (29.4)] этой малой поправкой к n_c , то мы должны также пренебречь малой компенсирующей плотностью дырок на далеких расстояниях в n -области кристалла. Аналогичные замечания справедливы и для малых концентраций электронов в p -области. Эти концентрации неосновных носителей вносят пренебрежимо малый вклад в общий баланс зарядов. Однако, как мы увидим ниже, они играют важную роль в определении токов при наличии приложенного напряжения.

⁴⁾ Более подробно это уравнение исследуется в задаче 1.

тельно, $n_c = N_d$ в n -части, $p_v = N_a$ в p -части, а $\rho = 0$. Внутри указанной области, за исключением ближайшей окрестности ее границ, отклонение величины $e\phi$ от асимптотического значения во много раз превосходит $k_B T$, поэтому $n_c \ll N_d$, $p_v \ll N_a$. Таким образом, за исключением окрестности точек $x = -d_p$ и $x = d_n$, плотность заряда в области между $-d_p$ и d_n достаточно точно описывается формулой $\rho(x) = e [N_d(x) - N_a(x)]$. При этом в указанной области концентрация носителей очень мала, так что они не могут нейтрализовать заряды «ионизованных» примесей. Точки $x = -d_p$ и $x = d_n$ определяют, следовательно, границы обедненного слоя.

Объединяя полученные выше результаты и используя формулу (29.1) для концентраций примесей, получаем, что для всех x , кроме значений, немного превышающих $-d_p$ или немного меньших d_n , уравнение Пуассона можно вполне аппроксимировать уравнением

$$\phi''(x) = \begin{cases} 0, & x > d_n, \\ -\frac{4\pi e N_d}{\varepsilon}, & d_n > x > 0, \\ \frac{4\pi e N_a}{\varepsilon}, & 0 > x > -d_p, \\ 0, & -d_p > x. \end{cases} \quad (29.13)$$

Непосредственным интегрированием получаем отсюда

$$\phi(x) = \begin{cases} \phi(\infty), & x > d_n, \\ \phi(\infty) - \left(\frac{2\pi e N_d}{\varepsilon}\right)(x - d_n)^2, & d_n > x > 0, \\ \phi(-\infty) + \left(\frac{2\pi e N_a}{\varepsilon}\right)(x + d_p)^2, & 0 > x > -d_p, \\ \phi(-\infty), & x < -d_p. \end{cases} \quad (29.14)$$

Решение (29.14) удовлетворяет граничным условиям (непрерывность потенциала ϕ и его первой производной) при $x = -d_p$ и $x = d_n$. Потребовав выполнения этих условий и при $x = 0$, можно получить два дополнительных равенства, которые определяют размеры d_n и d_p . Из непрерывности производной ϕ' при $x = 0$ следует равенство

$$N_d d_n = N_a d_p; \quad (29.15)$$

т. е. как раз условие того, что избыточный положительный заряд в n -области перехода равен избыточному отрицательному заряду в p -области. Из непрерывности потенциала ϕ при $x = 0$ следует условие

$$\left(\frac{2\pi e}{\varepsilon}\right)(N_d d_n^2 + N_a d_p^2) = \phi(\infty) - \phi(-\infty) = \Delta\phi. \quad (29.16)$$

Вместе с (29.15) оно определяет расстояния d_n и d_p :

$$d_{n,p} = \left\{ \frac{(N_a/N_d)^{\pm 1} \varepsilon \Delta\phi}{(N_d + N_a) 2\pi e} \right\}^{1/2}. \quad (29.17)$$

Чтобы оценить значения этих величин, мы можем записать (29.17) в форме, более удобной для численных расчетов:

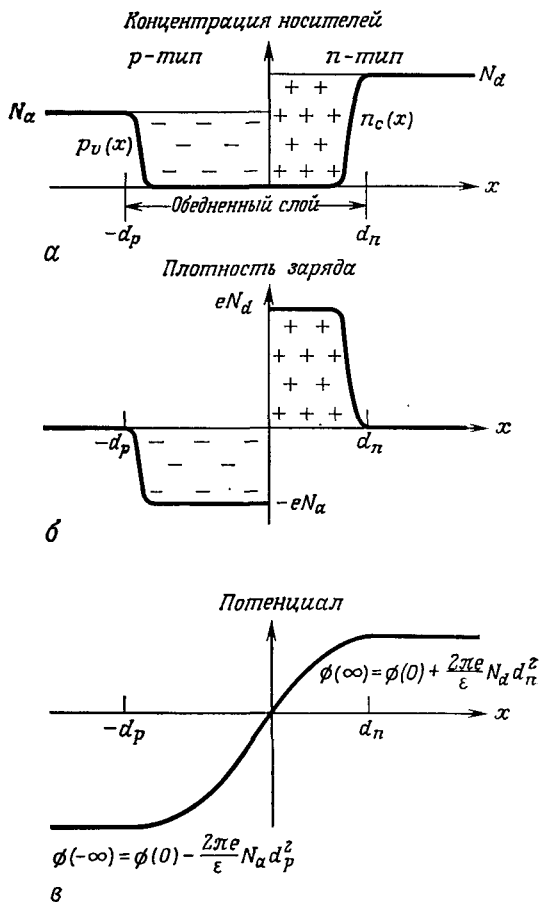
$$d_{n,p} = 33 \left\{ \frac{(N_a/N_d)^{\pm 1}}{10^{-18}(N_d + N_a)} [\varepsilon e \Delta\phi]_{\text{эВ}} \right\}^{1/2} \text{ \AA}. \quad (29.18)$$

Величина $\varepsilon e \Delta\phi$ имеет обычно порядок 1 эВ, и, поскольку типичные концентрации примесей составляют от 10^{14} до 10^{18} атомов в 1 см^3 , размеры d_n и d_p ,

определяющие ширину обедненного слоя, как правило, лежат в пределах от 10^4 до 10^2 Å. Поле внутри обедненного слоя имеет порядок $\Delta\phi/(d_n + d_p)$, и, следовательно, для указанных значений d оно имеет величину от 10^5 до 10^7 В/м, если ширина запрещенной зоны равна 0,1 эВ.

Получающаяся в результате схема обедненного слоя представлена на фиг. 29.3. Потенциал ϕ , как указывалось выше, изменяется в слое монотонно. Везде внутри слоя, за исключением его границ, концентрации носителей пренебрежимо малы по сравнению с концентрациями примесей, поэтому плотность заряда обусловлена ионизованными примесями. Вне обедненного слоя концентрации примесей и носителей сбалансированы, и плотность заряда равна нулю.

Механизм возникновения области с резко уменьшенной концентрацией носителей относительно прост. Предположим, что в начальный момент мы сумели задать такие концентрации носителей, чтобы в каждой точке выполнялось условие электронейтральности. Подобная конфигурация не может оставаться неизменной, так как электроны начнут диффундировать из n -области (где их концентрация высока) в p -область (где их концентрация очень мала), а дырки — диффундировать в противоположном направлении. Диффузия будет сопровождаться переносом заряда, что приведет в результате к возникновению электрического поля, препятствующего диффузионным токам. В конце концов будет достигнута равновесная конфигурация, в которой действие поля на токи в точности компенсирует эффекты, обусловленные диффузией. Поскольку носители весьма подвижны, в равновесной конфигурации концентрации носителей очень малы там, где поле имеет заметную величину. Именно такое положение дел и отражено на фиг. 29.3.

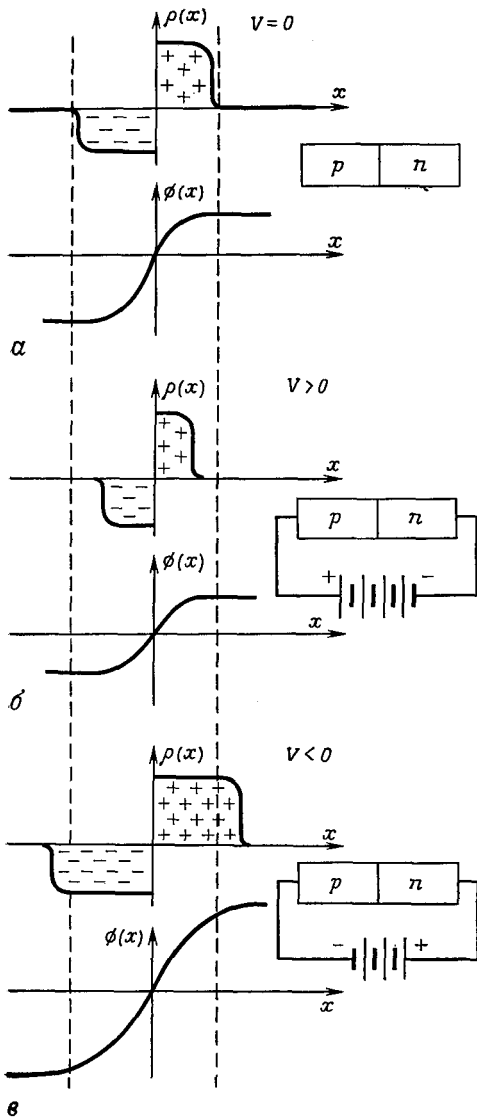


Фиг. 29.3. Концентрация носителей (а), плотность заряда (б), потенциал $\phi(x)$ (в), представленные в зависимости от положения точки относительно резкого $p-n$ -перехода.

В анализе, проведенном в тексте, мы приближенно принимали, что концентрации носителей и плотность заряда имеют постоянную величину всюду, кроме точек $x = -d_p$ и $x = d_n$, где они испытывают скачки. Более точное рассмотрение (см. задачу 1) показывает, что эти величины испытывают резкие изменения внутри области, отношение ширины которой к полной ширине обедненного слоя имеет порядок $(k_B T / E_g)^{1/2}$. Типичная ширина обедненного слоя составляет $10^2 - 10^4$ Å.

ЭЛЕМЕНТАРНОЕ РАССМОТРЕНИЕ ВЫПРЯМЛЯЮЩЕГО ДЕЙСТВИЯ $p-n$ -ПЕРЕХОДА

Рассмотрим теперь поведение $p-n$ -перехода, когда к нему приложено внешнее напряжение V . Будем считать напряжение V положительным, если оно увеличивает потенциал в p -области и уменьшает его в n -области. Как мы



Фиг. 29.4. Плотность заряда ρ и потенциал ϕ в обедненном слое в отсутствие смещающего напряжения (а); при $V > 0$ (прямое смещающее напряжение на переходе) (б); при $V < 0$ (обратное смещающее напряжение) (в).

Границы обедненного слоя, расположенные при $V = 0$ в точках $x = d_n$ и $x = -d_p$, изображены штриховыми линиями. Ширина обедненного слоя и величина изменения потенциала ϕ уменьшаются при прямом и увеличиваются при обратном смещающем напряжении.

видели ранее, вокруг той точки, в которой легирование p -типа переходит в легирование n -типа, при $V=0$ существует обедненный слой шириной примерно $10^2 - 10^4 \text{ \AA}$, где концентрация носителей значительно меньше, чем в однородных областях, расположенных дальше. Поскольку концентрация носителей в обедненном слое значительно меньше, чем в однородных областях, он обладает гораздо более высоким сопротивлением, а всю систему можно рассматривать как цепь с последовательно соединенными сопротивлениями, в которой относительно большое сопротивление помещено между двумя малыми (сэндвич). Если к такой цепи приложено напряжение V , то почти все падение напряжения приходится на область с высоким сопротивлением. Поэтому даже при наличии внешнего напряжения V потенциал $\phi(x)$ будет меняться заметным образом только в обедненном слое. При $V = 0$ мы нашли, что увеличение потенциала $\phi(x)$ при переходе из p -области в n -область обедненного слоя [которое мы теперь обозначим как $(\Delta\phi)_0$] описывается формулой (29.6). Поэтому можно сделать вывод, что при $V \neq 0$ изменение потенциала в обедненном слое будет равно

$$\Delta\phi = (\Delta\phi)_0 - V. \quad (29.19)$$

Изменение разности потенциалов на обедненном слое связано с изменением размеров слоя. Величины d_n и d_p , т. е. размеры слоя в n - и p -области, определяются формулами (29.15) и (29.16). В эти формулы входит только полное изменение потенциала в слое; кроме того, при их выводе использовалось предположение, что концентрации носителей практически во всем слое значительно меньше, чем вне его. Мы покажем ниже, что это предположение остается верным и при $V \neq 0$, поэтому формула (29.17) для d_n и d_p остается правильной, если заменить $\Delta\phi$ на $(\Delta\phi)_0 - V$. Поскольку в соответствии с формулой (29.17) d_n и d_p меняются пропорционально $(\Delta\phi)^{1/2}$, можно сделать вывод, что при $V \neq 0$

$$d_{n, p}(V) = d_{n, p}(0) \left[1 - \frac{V}{(\Delta\phi)_0} \right]^{1/2}. \quad (29.20)$$

Поведение ϕ и ширины обедненного слоя показано на фиг. 29.4.

Чтобы вывести зависимость величины тока через $p-n$ -переход от внешнего смещающего напряжения V , мы должны рассмотреть отдельно электронные и дырочные токи. В дальнейшем будем обозначать символом J плотность потока частиц, а символом j — плотность электрического тока; тогда

$$j_e = -eJ_e, \quad j_h = eJ_h. \quad (29.21)$$

При $V = 0$ как J_e , так и J_h обращаются в нуль. Это означает, конечно, не отсутствие движения отдельных носителей через переход, а только то, что в обоих направлениях движутся равные количества электронов (или дырок). При $V \neq 0$ баланс нарушается. Рассмотрим, например, дырочный ток через обедненный слой. Он включает следующие две компоненты.

1. *Ток генерации*, т. е. дырочный ток, текущий из n -области в p -область перехода. Как видно из названия, этот ток обусловлен дырками, генерируемыми непосредственно в n -области обедненного слоя при тепловом возбуждении электронов с уровней валентной зоны. Хотя концентрация таких дырок (неосновных носителей) в n -области чрезвычайно мала по сравнению с концентрацией электронов (основных носителей), они играют важную роль в переносе тока через переход. Это происходит потому, что каждая дырка, попадающая в обедненный слой, тут же перебрасывается в p -область под действием сильного электрического поля, которое имеется внутри слоя. В результате величина возникающего тока генерации не зависит от значения изменения потенциала в обедненном слое, поскольку любая дырка, оказавшаяся в слое, перебрасывается из n -области в p -область¹).

2. *Ток рекомбинации*, т. е. дырочный ток, текущий из p -области в n -область²). Электрическое поле в обедненном слое препятствует этому току, и только те дырки, которые попадают на границу обедненного слоя, имея достаточную кинетическую энергию, чтобы преодолеть потенциальный барьер, вносят вклад в ток рекомбинации. Число таких дырок пропорционально $e^{-\Delta\phi/k_B T}$ и, следовательно³),

$$J_h^{\text{rec}} \sim e^{-e[(\Delta\phi)_0 - V]/k_B T}. \quad (29.22)$$

В отличие от тока генерации ток рекомбинации чрезвычайно чувствителен к величине приложенного напряжения V . Мы можем сравнить величины этих двух токов, заметив, что при $V = 0$ суммарный ток через переход отсутствует:

$$J_h^{\text{rec}}|_{V=0} = J_h^{\text{gen}}. \quad (29.23)$$

Отсюда и из формулы (29.22) следует, что

$$J_h^{\text{rec}} = J_h^{\text{gen}} e^{eV/k_B T}. \quad (29.24)$$

¹ Концентрация дырок, от которой зависит ток генерации, нечувствительна также к величине V , если величина eV мала по сравнению с E_g , так как, согласно закону действующих масс, эта концентрация полностью определяется концентрацией электронов. Последняя лишь немного отличается от значения N_c вне обедненной области, если значение eV мало по сравнению с E_g , как это видно из более подробного рассмотрения, проведенного ниже.

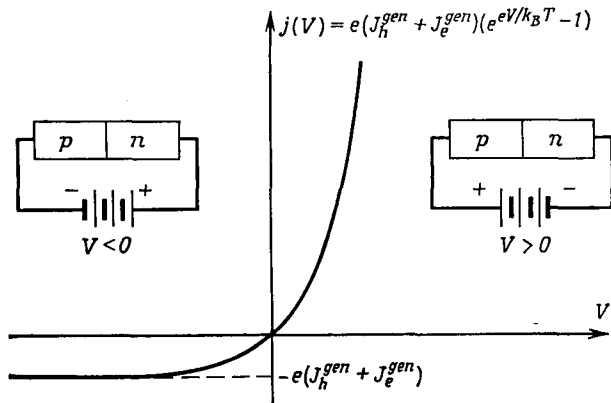
² Он назван так из-за судьбы дырок, попадающих в n -область перехода, где какой-нибудь из избыточных электронов в конце концов заполняет пустой уровень, каковым является дырка.

³ Утверждая, что (29.22) описывает основную часть зависимости дырочного тока рекомбинации от V , мы предполагаем, что концентрация дырок вблизи границы p -области обедненного слоя только немного отличается от N_a . Как будет показано, и это предположение выполняется, если величина eV мала по сравнению с шириной запрещенной зоны E_g .

Полный дырочный ток, текущий из p -области в n -область, представляет собой разность между токами рекомбинации и генерации:

$$J_h = J_h^{\text{rec}} - J_h^{\text{gen}} = J_h^{\text{gen}} (e^{eV/k_B T} - 1). \quad (29.25)$$

Аналогичное рассмотрение применимо к компонентам электронного тока с тем только изменением, что токи генерации и рекомбинации электронов направлены противоположно соответствующим дырочным токам. Поскольку,



Фиг. 29.5. Зависимость тока через p — n -переход от приложенного напряжения V . Эта зависимость имеет место, когда величина eV мала по сравнению с шириной запрещенной зоны E_g . Ниже будет показано, что ток насыщения ($eJ_h^{\text{gen}} + eJ_e^{\text{gen}}$) изменяется с температурой как $e^{-E_g/k_B T}$.

однако, электроны имеют противоположный заряд, электрические токи генерации и рекомбинации электронов совпадают по направлению с электрическими токами генерации и рекомбинации дырок. Поэтому полная плотность электрического тока есть

$$j = e (J_h^{\text{gen}} + J_e^{\text{gen}}) (e^{eV/k_B T} - 1). \quad (29.26)$$

Эта зависимость имеет весьма несимметричный вид, характерный для выпрямителей (фиг. 29.5).

ОСНОВНЫЕ ФИЗИЧЕСКИЕ ЧЕРТЫ НЕРАВНОВЕСНОГО СЛУЧАЯ

Рассмотрение, проведенное в предыдущем разделе, не дает возможности оценить величину $e (J_h^{\text{gen}} + J_e^{\text{gen}})$, входящую в формулу (29.26). Кроме того, в неравновесном случае ($V \neq 0$) локальные концентрации носителей, вообще говоря, не выражаются через локальный потенциал ϕ с помощью простых равновесных максвелловских соотношений (29.3). В такой ситуации требуется дополнительный анализ, чтобы построить столь же детальную картину концентраций носителей в окрестности переходной области, как и в равновесном случае.

При таком более детальном подходе разделение электронных и дырочных токов через переход на токи генерации и рекомбинации не дает особых преимуществ. Вместо этого запишем для каждой точки x (как вне, так и внутри

обедненного слоя) уравнения, связывающие полные электронные и дырочные токи $J_e(x)$ и $J_h(x)$, концентрации электронов и дырок $n_c(x)$ и $p_v(x)$ и потенциал $\phi(x)$ [или, что то же самое, электрическое поле, $E(x) = -d\phi(x)/dx$]. Мы получим пять уравнений, которые позволят в принципе определить эти пять величин. Такой метод представляет собой непосредственное обобщение подхода, который мы применяли при анализе равновесного случая ($V = 0$). При равновесии электронные и дырочные токи обращаются в нуль и имеются только три неизвестные величины; в этом случае используются три уравнения: уравнение Пуассона и два уравнения (29.3), связывающие при термодинамическом равновесии $n_c(x)$ и $p_v(x)$ с $\phi(x)$. Итак, можно считать, что неравновесная задача заключается в нахождении подходящих уравнений, заменяющих равновесные соотношения (29.3), когда $V \neq 0$ и существуют токи.

Отметим прежде всего, что при наличии электрического поля и градиента концентрации носителей плотность потока носителей можно представить в виде суммы двух членов, один из которых пропорционален полю (*дрейфовый ток*), а другой — градиенту концентрации (*диффузионный ток*):

$$\begin{aligned} J_e &= -\mu_n n_c E - D_n \frac{dn_c}{dx}, \\ J_h &= \mu_p p_v E - D_p \frac{dp_v}{dx}. \end{aligned} \quad (29.27)$$

Положительные ¹⁾ коэффициенты пропорциональности μ_n и μ_p , входящие в (29.27), называются соответственно *подвижностями* электронов и дырок. Мы ввели подвижности, а не записали дрейфовый ток через проводимость, чтобы выделить явно зависимость дрейфового тока от концентрации носителей. Если имеются только электроны, концентрация которых однородна, то $\sigma E = j = -eJ_e = e\mu_n n E$. Используя формулу Друде (1.6) для проводимости $\sigma = ne^2\tau/m$, находим

$$\mu_n = \frac{e\tau_n^{\text{coll}}}{m_n}, \quad (29.28)$$

и аналогично

$$\mu_p = \frac{e\tau_p^{\text{coll}}}{m_p}, \quad (29.29)$$

где m_n и m_p — соответствующие эффективные массы, а τ_n^{coll} и τ_p^{coll} — средние времена между столкновениями носителей ²⁾.

Положительные ³⁾ коэффициенты пропорциональности D_n и D_p , входящие в (29.27), называются *коэффициентами диффузии*. Они связаны с подвижностями

¹⁾ Знаки в (29.27) выбраны так, чтобы подвижности были положительны; дырочный дрейфовый ток направлен по полю, а электронный — против поля.

²⁾ В полупроводниках имеется еще одно характерное время, играющее фундаментальную роль, — *время рекомбинации* (см. ниже). Индекс «coll» (collision — столкновение) введен, чтобы отличить средние времена между столкновениями от времен рекомбинации.

³⁾ Они положительны, так как диффузионный ток направлен из области высокой в область низкой концентрации носителей. При нулевом поле соотношения (29.27) иногда называют законом Фика.

соотношениями Эйнштейна ¹⁾):

$$\mu_n = \frac{eD_n}{k_B T}, \quad \mu_p = \frac{eD_p}{k_B T}. \quad (29.30)$$

Соотношения Эйнштейна следуют непосредственно из того факта, что при термодинамическом равновесии электронные и дырочные токи должны обращаться в нуль. Только если подвижности и коэффициенты диффузии удовлетворяют равенствам (29.30), токи, определяемые формулами (29.27), обратятся в нуль при равновесных значениях концентраций носителей (29.3) ²⁾ [что легко проверить прямой подстановкой (29.3) в (29.27)].

Соотношения (29.27), выражающие токи через поля и градиенты концентраций, и формулы (29.28) — (29.30) для подвижностей и коэффициентов диффузии можно получить также непосредственно при помощи простого кинетического рассмотрения типа использованного в гл. 1 (см. задачу 2).

Отметим, что при термодинамическом равновесии соотношения (29.27) и условия $J_e = J_h = 0$ содержат всю информацию, необходимую для определения концентраций носителей, поскольку в отсутствие токов мы можем проинтегрировать уравнения (29.27) и (с помощью соотношений Эйнштейна) независимым образом вывести выражения для равновесных концентраций (29.3). При $V \neq 0$ и наличии токов необходимы еще уравнения, которые можно рассматривать как обобщение на неравновесный случай равновесного условия отсутствия токов. Если бы число носителей сохранялось, то необходимое обобщение свелось бы просто к уравнениям непрерывности

$$\begin{aligned} \frac{\partial n_c}{\partial t} &= -\frac{\partial J_e}{\partial x}, \\ \frac{\partial p_v}{\partial t} &= -\frac{\partial J_h}{\partial x}. \end{aligned} \quad (29.31)$$

Эти уравнения отражают тот факт, что изменение числа носителей в некоторой области полностью определяется скоростью, с которой носители втекают в эту область и вытекают из нее. Однако число носителей не сохраняется. Электроны зоны проводимости и дырки валентной зоны могут *генерироваться* в результате теплового возбуждения электронов с уровней в валентной зоне. Помимо этого, электрон зоны проводимости может *рекомбинировать* с дыркой валентной зоны (т. е. электрон может занять пустой уровень, который и является дыркой), что приводит к исчезновению обоих носителей. В уравнения непрерывности следует добавить члены, описывающие все процессы, которые изменяют число частиц в выделенной области:

$$\begin{aligned} \frac{\partial n_c}{\partial t} &= \left(\frac{dn_c}{dt} \right)_{g-r} - \frac{\partial J_e}{\partial x}, \\ \frac{\partial p_v}{\partial t} &= \left(\frac{dp_v}{dt} \right)_{g-r} - \frac{\partial J_h}{\partial x}. \end{aligned} \quad (29.32)$$

Чтобы определить вид производных $(dn_c/dt)_{g-r}$ и $(dp_v/dt)_{g-r}$, отметим, что, когда концентрации носителей отклоняются от равновесных значений, генерация и рекомбинация ведут к восстановлению термодинамического равновесия. В областях, где n_c и p_v превышают равновесные значения, рекомбинация происходит быстрее, чем генерация, приводя к уменьшению концентрации носителей;

¹⁾ Соотношения Эйнштейна имеют весьма общий характер; они возникают при рассмотрении любых заряженных частиц, подчиняющихся статистике Максвелла — Больцмана, например ионов в электролитах.

²⁾ Обобщение соотношений (29.30) на вырожденный случай дано в задаче 3.