

В. И. СТРИХА

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

*Допущено Министерством
высшего и среднего специального образования УССР
в качестве учебного пособия для студентов
радиофизических, физических и электронных
специальностей вузов*

КИЕВ
ГОЛОВНОЕ ИЗДАТЕЛЬСТВО
ИЗДАТЕЛЬСКОГО ОБЪЕДИНЕНИЯ
«ВИЩА ШКОЛА»
1982

5
4
9
9
4
8
2
4
9
23
28
31
33
40
43
47
52
3

Контактные явления в полупроводниках. Стриха В. И.—Киев: Выща школа. Головное изд-во, 1982.— 224 с.

Изложены физические основы работы контакта металл — полупроводник, $p-n$ перехода и гетероперехода. Рассматривается физическая модель гетероперехода, из которой получаются модели $p-n$ перехода и контакта металл — полупроводник.

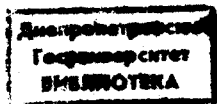
Подробно анализируются вольт-амперные характеристики контакта металл — полупроводник, $p-n$ перехода и гетероперехода с учетом всех основных эффектов, характерных для этих контактов. Рассматриваются эквивалентная схема и пробойные явления. Особое внимание уделяется омическим контактам металл — полупроводник.

Для студентов радиофизических, физических и электронных специальностей вузов, занимающихся вопросами физики полупроводников и полупроводниковой электроники. Пособие может быть полезно аспирантам, научным и инженерно-техническим работникам.

Ил. 93. Табл. 1. Библиогр.: 11 назв.

Рецензенты: д-р физ.-мат. наук П. И. Баранский (Институт полупроводников АН УССР), д-р физ.-мат. наук М. В. Пашковский (Львовский государственный университет)

Редакция литературы по кибернетике, электронике, энергетике
Зав. редакцией М. С. Хойнацкий



1315586

ОГЛАВЛЕНИЕ

Основные условные обозначения	5	3.6. Параметры вольт-амперной характеристики и их анализ	85
Предисловие	9	3.7. Экспериментальные исследования вольт-амперных характеристик	94
Глава 1. Зонные модели гетеро- и гомопереходов	11	Глава 4. Вольт-амперные характеристики $p-n$ перехода	99
1.1. Образование гетеро- и гомопереходов	11	4.1. Общее рассмотрение невырожденного $p-n$ перехода	99
1.2. Ход потенциала в области пространственного заряда полупроводника	20	4.2. Вольт-амперная характеристика $p-n$ перехода с широкой базой при малых уровнях инжекции	104
1.3. Соотношение между высотами барьеров в гетеропереходе	25	4.3. Вольт-амперная характеристика $p-n$ перехода с узкой базой	108
1.4. Влияние сил изображения на ход потенциала в гетеропереходе	30	4.4. Вольт-амперная характеристика $p-n$ перехода при больших уровнях инжекции	112
1.5. Экспериментальные исследования зонной модели гетеро- и гомопереходов	32	4.5. Вольт-амперная характеристика $p-n$ перехода с учетом генерации-рекомбинации в области пространственного заряда	114
Глава 2. Общий анализ модели гетеро- и гомопереходов при приложении напряжения	36	4.6. Избыточные токи в тонких $p-n$ переходах	119
2.1. Поведение гетеро- и гомопереходов при приложении напряжения	36	4.7. Вольт-амперная характеристика $p-n$ перехода между вырожденными полупроводниками	123
2.2. Прохождение носителей заряда через квазинейтральную область	40	4.8. Экспериментальные исследования $p-n$ переходов	128
2.3. Прохождение носителей заряда в областях пространственного заряда	45	Глава 5. Выпрямляющие свойства гетеропереходов	131
2.4. Прохождение носителей заряда в вазоре (промежуточном слое)	50	5.1. Вольт-амперные характеристики изотипных гетеропереходов	133
2.5. Соотношение между напряжениями в гетеропереходе	51	5.2. Вольт-амперные характеристики анизотипных гетеропереходов	140
Глава 3. Выпрямление в контакте металл — полупроводник	55	5.3. Влияние туннелирования через область пространственного заряда и сил изображения на вольт-амперные характеристики гетеропереходов	143
3.1. Вольт-амперная характеристика контакта с учетом одного типа носителей заряда	55	5.4. Вольт-амперные характеристики гетеропереходов для токов через поверхностные уровни при движении носителей в пределах одного типа разрешенных зон	147
3.2. Вольт-амперные характеристики контакта с учетом двух типов носителей заряда (малый уровень инжекции)	61	5.5. Вольт-амперная характеристика гетеропереходов для токов через поверхностные уровни при движении носителей с изменением вида разрешенных зон	152
3.3. Расчет вольт-амперной характеристики с учетом туннелирования через область пространственного заряда	69		
3.4. Влияние сил изображения на вольт-амперную характеристику контакта	73		
3.5. Вольт-амперная характеристика для токов через поверхностные состояния	76		

5.6. Генерационно-рекомбинационные и туннельно-рекомбинационные токи в гетеропереходе	156	7.4. Эквивалентная схема контакта для токов через поверхностные уровни	193
5.7. Параметры α и α^* вольт-амперной характеристики гетероперехода	159	7.5. Эквивалентная схема $p-n$ перехода	195
5.8. Экспериментальные исследования вольт-амперных характеристик гетеропереходов	167	7.6. Эквивалентная схема гетероперехода	199
Глава 6. Пробойные явления	171	7.7. Эквивалентная схема $p-n$ перехода на участке лавинного пробоя	201
6.1. Виды пробоя	171	7.8. Экспериментальные исследования гомо- и гетеропереходов при малом переменном напряжении	207
6.2. Лавинный пробой	173	Глава 8. Омические контакты	211
6.3. Туннельный пробой	176	8.1. Свойства и модель омических контактов	211
6.4. Тепловой пробой	177	8.2. Омические контакты для токов без участия поверхностных уровней	212
6.5. Микроплазма	179	8.3. Омические контакты для токов через поверхностные уровни	214
6.6. Экспериментальные исследования пробоя	180	8.4. Получение и исследование омических контактов	215
Глава 7. Эквивалентные схемы гетеро- и гомопереходов	183	<i>Заключение</i>	217
7.1. Общие положения	183	<i>Контрольные вопросы</i>	219
7.2. Эквивалентная схема контакта металл-полупроводник для одного типа носителей заряда без учета поверхностных уровней	186	<i>Список литературы</i>	221
7.3. Влияние поверхностных уровней на эквивалентную схему контакта	189	<i>Предметный указатель</i>	222

ОСНОВНЫЕ УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

A — постоянная Ричардсона;	$D_{n,p}$ — коэффициент диффузии для электронов и дырок;
C — емкость гомо- или гетероперехода;	ξ_g — ширина запрещенной зоны полупроводника;
C_1 — емкость зазора в контакте металл — полупроводник;	$\Delta \xi_c$ и $\Delta \xi_v$ — энергетические разрывы зоны проводимости и валентной зоны в гетеропереходе;
C_2 — емкость области пространственного заряда полупроводника в контакте металл — полупроводник;	ξ_i — энергетическое положение i -го поверхностного уровня;
C_6 — барьерная емкость гомо- или гетероперехода, связанная с максвелловскими токами смещения;	ξ, ξ_x — энергия электрона;
C_d — диффузионная емкость, связанная с существованием неосновных носителей заряда;	$E(x)$ — напряженность электрического поля;
\bar{c}_n и \bar{c}_n^m — сечения захвата электрона зоны проводимости полупроводника и металла поверхностным уровнем;	$E_{1,2,3}$ и $E_{1,2,3}^0$ — напряженность электрического поля на границах раздела в полупроводниках и зазоре при $V \neq 0$ и $V = 0$;
\bar{c}_p — сечение захвата дырки валентной зоны полупроводника поверхностным уровнем;	$M_i = e \frac{\xi_i - \mu_{s0}}{kT}$;
\bar{c}_1^i и \bar{c}_2^i — сечения захвата электрона или дырки первого и второго полупроводника;	$M_{n,p}$ — коэффициент лавинного умножения электронов и дырок;
d — толщина зазора;	$m^*, m_{n,p}^*, m_m^*$ — эффективная масса электронов и дырок в полупроводнике, зазоре и металле;
D — ширина слоя Цоттки при неравномерном распределении примесей в полупроводнике;	m_i — концентрация электронов на поверхностных уровнях;
	n_i — концентрация поверхностных уровней;
	n_d — концентрация донорных уровней;
	$n_{эд}$ — концентрация элект-

n_0, n_n, n_p — равновесные концентрации электронов в объеме полупроводника;

n_i — концентрация носителей в зоне проводимости, если уровень Ферми совпадает с энергетическим положением рекомбинационного центра;

n_k — концентрация электронов в той области полупроводника, с которой происходит электронный обмен поверхностных уровней;

Δn — избыточная концентрация электронов;

n_v, n_m — количество свободных и занятых мест в металле в области, с которой происходит электронный обмен поверхностных уровней;

N_i — плотность поверхностных состояний;

p_0, p_n, p_p — равновесная концентрация дырок в объеме полупроводника;

p_i — концентрация носителей в валентной зоне, если уровень Ферми совпадает с энергетическим положением рекомбинационного центра;

p_k — концентрация дырок в области полупроводника, с которой происходит электронный обмен поверхностных состояний;

Δp — избыточная концентрация дырок;

$\mathcal{P}(\xi)$ — коэффициент прозрачности;

$\mathcal{P}_{n,p}$ — коэффициент прозрачности для электронов и дырок;

$i_{n,p}(\xi)$ — функция распределения Ферми;

g — концентрация носителей заряда, возникающих в единицу времени под действием внешних факторов;

G — активная составляющая общей проводимости гомо- и гетероперехода;

I — общий ток;

I_s — ток, являющийся параметром вольт-амперной характеристики;

i — плотность тока;

i_n, i_p — плотности электронного и дырочного токов;

i_n^s, i_p^s, i_m^s — плотности токов, характеризующих обмен поверхностных состояний с зоной проводимости, валентной зоной полупроводника и металлом;

i_s, i_{sn}, i_{st} — плотности токов, являющиеся параметрами вольт-амперной характеристики;

Δi — амплитуда переменной составляющей плотности тока;

k — постоянная Больцмана;

K_{ni}, K_{oi}, K_{mi} — кинетические коэффициенты, характеризующие переход носителей на поверхностный уровень из зоны проводимости, валентной зоны и металла;

K_{ln}, K_{lv}, K_{lm} — кинетические коэффициенты, характеризующие обратные переходы;

$\mathcal{L}_0, \mathcal{L}$ — ширина слоя Шоттки для равномерного распределения примесей при $V = 0$ и $V \neq 0$;

$\mathcal{L}_n, \mathcal{L}_p$ — ширина слоя Шоттки в n - и p -областях p - n перехода;

l — длина свободного пробега электрона по импульсам;

$l_{n,p}$ — диффузионная длина для электронов и дырок;

$l_{n\bar{v}}, l_{p\bar{v}}$ — диффузионная длина для электронов и дырок при большом уровне инжекции;

r — сопротивление объема полупроводника;

R_T — тепловое сопротивление p - n перехода;

R — сопротивление контакта;

$S_{n,p}$ — скорость поверхностной рекомбинации электронов и дырок;

T — температура;

u_n, u_p — подвижность электронов и дырок в полупроводнике;

$\bar{v}_{др}$ — дрейфовая скорость носителей;

$\bar{v}_{n,p}$ — средняя тепловая скорость электронов и дырок в полупроводнике;

\bar{v}_n^m — средняя тепловая скорость электронов в металле;

V — напряжение, приложенное к гомо- или гетеропереходу;

$V_{1,2,3}$ — падение напряжения в области пространственного заряда и зазоре;

V_k — контактная разность потенциалов;

$V_{1,2}^t$ — потенциал, связанный с перезарядкой поверхностных состояний при приложении напряжения;

V_b — падение напряжения в квазинейтральной области полупроводника;

$V_{пр}$ — напряжение лавинного пробоя;

$\omega, \omega_{n,p}$ — ширина квазинейтральной области полупроводника;

Z — полное сопротивление гомо- или гетероперехода;

α, α^* — множители в показателе экспоненты в выражениях вольт-амперной характеристики;

$\alpha_{n,p}$ — коэффициент ударной ионизации электронов и дырок;

γ — коэффициент инжекции;

$\epsilon_0, \epsilon, \epsilon_f, \epsilon_3, \epsilon_2$ — диэлектрические проницаемости вакуума, полупроводников и зазора;

μ — уровень химического потенциала;

μ_s и μ_{s0} — положение уровня химического потенциала на поверхности раздела при $V \neq 0$ и $V = 0$;

μ_n, μ_p — квазиуровни Ферми для электронов и дырок;

$\rho_{n,p}$ — удельные сопротивления n - и p -областей;

$\rho(x)$ — объемный заряд в области пространственного заряда полупроводника;

$\sigma_{i,j}, \sigma_{i,j}^0$ — плотности заряда на поверхностных состояниях при $V \neq 0$ и $V = 0$;

$\sigma, \sigma_n, \sigma_p$ — удельная проводимость полупроводника;

τ_n, τ_p — время жизни основных носителей заряда;
 Φ_s, Φ_m — термодинамические работы выхода полупроводника и металла;
 Φ_0, Φ_b — высота потенциального барьера области пространственного заряда полупроводника ($\Phi_b = \Phi_0 + \mu$);
 $\Phi(x)$ — потенциальная энергия электрона;
 $\Phi_1, \Phi_2, \Phi_{10}, \Phi_{20}$ — высоты потенциальных барьеров различных полу-

проводников при $V \neq 0$ и $V = 0$;
 $\Delta\phi_z$ — изменение высоты потенциального барьера вследствие действия сил изображения;
 χ_1, χ_2 — внешние работы выхода для контактирующих полупроводников;
 χ_m — разность электроотрицательностей элементов;
 ω — круговая частота переменного сигнала.

ПРЕДИСЛОВИЕ

Трудно переоценить значение полупроводниковых приборов в развитии техники сегодняшнего дня. Они обусловили принципиальные изменения в радиотехнике, сыграли огромную роль в развитии космической техники и создании совершенных ЭВМ.

Основными направлениями экономического и социального развития СССР на 1981—1985 годы и на период до 1990 года предусмотрено повысить технический уровень вычислительной техники, приборов и средств автоматизации на основе новейших достижений микроэлектроники, оптоэлектроники и лазерной техники.

В абсолютном большинстве существующих полупроводниковых приборов используются контактные явления в полупроводниках. При этом могут использоваться либо один активный контакт, например диод с барьером Шоттки, $p-n$ переход или простой гетеропереход, либо несколько контактов, например биполярный транзистор, тиристор, МДП-транзистор.

Однако как для одноконтактных приборов, так и для многоконтактных основой понимания их работы является исследование физики явлений в отдельно взятом контакте металла с полупроводником или полупроводника с другим полупроводником.

В настоящем учебном пособии и рассмотрены физические основы работы полупроводниковых приборов, использующих контактные явления. При этом рассматриваются все основные физические процессы, характеризующие работу контакта металл — полупроводник и полупроводник — полупроводник.

С точки зрения природы контактирующих тел в системе полупроводник — полупроводник можно выделить гетеропереходы, когда контактируют два полупроводника различного состава и свойств, и гомопереходы, когда контактируют одинаковые полупроводники лишь с различной концентрацией примесей. Очевидно, гетеропереход является более общим понятием, чем гомопереход, так как последний является частным случаем гетероперехода, когда свойства контактирующих полупроводников одинаковы. К частному случаю гетероперехода в известных приближениях может быть отнесен и контакт металл — полупроводник как контакт сильно вырожденного полупроводника с нулевой шириной запрещенной зоны с полупроводником с произвольной шириной запрещенной зоны.

В гл. 1 настоящего пособия анализируется общая физическая модель гетероперехода, учитывающая зазор (промежуточный слой) и поверхностные состояния, на основании которой легко получаются модели контакта металл — полупроводник и $p-n$ перехода.

В гл. 2 проводится общий анализ прохождения носителей заряда через контакт. Рассматриваются основные уравнения, описывающие работу контактов. Анализируются критерии применимости диодной и диффузионной теорий.

В гл. 3 рассматривается вольт-амперная характеристика контакта металл — полупроводник с учетом одного и двух типов носителей заряда, а также промежуточного слоя между металлом и полупроводником в предположении диодной и диффузионной теорий. Анализируется влияние на вольт-амперную характеристику таких факторов, как туннелирование и силы изображения. Подробно рассматривается участие поверхностных состояний в переносе тока через контакт.

В гл. 4 изучаются вольт-амперные характеристики $p-n$ перехода. Это рассмотрение проводится с учетом малого или большого уровня инжекции для $p-n$ переходов с широкой и узкой квазинейтральной областью (базой). Особо анализируются процессы генерации и рекомбинации в области пространственного заряда

и туннелирования носителей заряда между разрешенными зонами полупроводника (туннельные диоды).

В гл. 5 рассматриваются вольт-амперные характеристики гетеропереходов. Здесь даны выражения вольт-амперных характеристик для токов без участия промежуточных локальных уровней и для токов с участием локальных состояний.

Гл. 6 посвящена рассмотрению эквивалентной схемы контакта металл — полупроводник, $p-n$ и гетероперехода, рассматриваются барьерная и диффузионная емкости и влияние на эквивалентную схему поверхностных состояний.

В гл. 7 исследуются процессы, которые развиваются в контактах при больших отрицательных напряжениях. В этой главе анализируются лавинный, туннельный и тепловой механизмы пробоя в объеме и на поверхности полупроводника.

Гл. 8 посвящена рассмотрению омических контактов. Здесь приводятся и анализируются выражения для сопротивления омических контактов.

В конце каждой главы приводятся основные экспериментальные результаты, позволяющие проиллюстрировать справедливость рассмотренных физических моделей гомо- и гетеропереходов.

В заключении анализируются физические принципы, используемые при практическом применении контакта металл-полупроводник, $p-n$ перехода и гетероперехода, сравниваются достоинства и недостатки различных контактов.

Принятое построение книги позволяет решить несколько задач. Во-первых, рассмотреть с общей позиции явления, характерные для контакта металл — полупроводник, $p-n$ и гетероперехода. Во-вторых, провести изложение материала при рассмотрении выпрямляющих свойств, начиная от более простой системы контакта металл — полупроводник, где существенную роль играют один полупроводник и зазор, до такой сложной системы, как гетеропереход, где такую же роль могут играть два полупроводника и зазор. В-третьих, отдельно остановиться на основных явлениях, существенных для контакта металл — полупроводник, $p-n$ и гетероперехода, например, рассмотреть силы изображения, туннелирование и т. п.

Учебное пособие написано на основании курсов «Контактные явления в полупроводниках» и «Физика полупроводниковых приборов», которые читались автором в Киевском, Черновицком, Северо-Осетинском государственных университетах и Ягеллонском университете (ПНР).

Отзывы и пожелания по настоящему изданию просим направлять по адресу: 252054, Киев-54, ул. Гоголевская, 7, Головное издательство издательского объединения «Вища школа».

Глава 1. ЗОННЫЕ МОДЕЛИ ГЕТЕРО- И ГОМОПЕРЕХОДОВ

1.1. ОБРАЗОВАНИЕ ГЕТЕРО- И ГОМОПЕРЕХОДОВ

Рассмотрим образование гетеро- и гомоперехода в результате сближения и контактирования изолированных двух полупроводников или металла и полупроводника. Для простоты будем предполагать, что поверхность контактирующих тел свободна от окислов и адсорбированных газов, а также на ней нет поверхностных электронных состояний.

Как известно, плотность потока электронов, которые могут покинуть изолированное твердое тело и выйти в вакуум в результате термоэлектронной эмиссии, определяется формулой Ричардсона

$$i = \frac{4\pi m^* (kT)^2}{h^3} e^{-\frac{\varphi_{s,m}}{kT}}, \quad (1.1)$$

где m^* — эффективная масса электронов в твердом теле; $\varphi_{s,m}$ — термодинамическая работа выхода. Для полупроводника и металла термодинамические работы выхода φ_s и φ_m иллюстрируются рис. 1.1.

Отметим, что формула Ричардсона в виде (1.1) справедлива при условии $\varphi_{s,m} \gg kT$, где $T > 0$, которое дает возможность рассматривать выход электронов, подчиняющихся статистике Максвелла, а не статистике Ферми. Это важное условие необходимо всегда учитывать при использовании формулы (1.1).

Из (1.1) видно, что при постоянной температуре плотность потока электронов определяется главным образом величиной работы выхода φ_s или φ_m — чем меньше работа выхода, тем больше ток термоэлектронной эмиссии.

В случае изолированного твердого тела эмиссия электронов со временем уменьшается вследствие того, что оно заряжается положительно и возникает электрическое поле, препятствующее эмиссии.

Теперь рассмотрим, что произойдет при сближении первоначально взаимно изолированных двух твердых тел до расстояния, на котором электроны из одного твердого тела могут переходить в другое.

В этом случае согласно выражению (1.1) больше электронов будет переходить из вещества с меньшей работой выхода по направлению к веществу с большей работой выхода. Последнее будет заряжаться отрицательно, а первое положительно. В результате переход электронов вещества с меньшей работой выхода со временем

будет затрудняться, и с большей — облегчаться. Все это будет происходить до тех пор, пока потоки электронов в разных направлениях не уравниваются. Такое состояние является стационарным.

Стационарное состояние показано на рис. 1.2. Из этого рисунка видно, что заряженная область может простирается в глубь твердого тела.

Ширину заряженной области можно оценить, если принять, что основная часть разности работ выхода падает в зазоре толщиной d между твердыми телами. Тогда, используя условия постоянства поля и теорему Гаусса, напряженность поля в зазоре можно записать как

$$E \approx \frac{\Phi_1 - \Phi_2}{ed} = e\sigma,$$

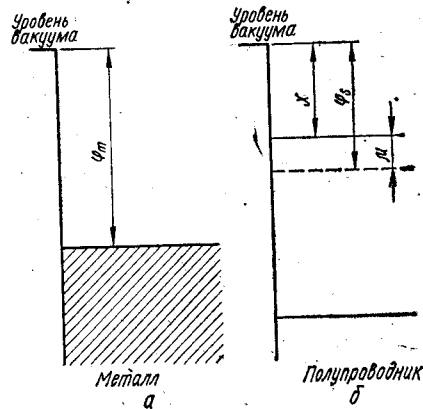


Рис. 1.1. Упрощенные зонные модели металла (а) и полупроводника (б).

Рис. 1.2. Результаты контактирования твердых тел с разной работой выхода.

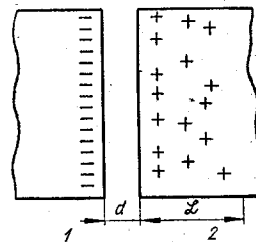


Рис. 1.2. Результаты контактирования твердых тел с разной работой выхода.

где Φ_1 и Φ_2 — работы выхода двух тел; σ — поверхностная плотность избыточных или недостающих электронов.

Решая последнее уравнение относительно σ , можно найти

$$\sigma = \frac{\Phi_1 - \Phi_2}{e^2 d}.$$

Ширина слоя, из которого нужно удалить или в который нужно ввести электроны, может быть выражена как $L = \frac{\sigma}{n_0}$, где n_0 — концентрация электронов в твердом теле. Теперь

$$L = \frac{\Phi_1 - \Phi_2}{e^2 d n_0}.$$

Как видно, глубина проникновения поля в твердое тело зависит от концентрации носителей заряда n_0 в нем. Например, при типичных значениях $\Phi_1 - \Phi_2 = 1$ эВ, $d = 10^{-7}$ см и $n_0 = 10^{14}$ см $^{-3}$, характерной для полупроводника, $L = 0,5 \cdot 10^{-2}$ см, а при тех же $\Phi_1 - \Phi_2$ и d , но при $n_0 = 10^{22}$ см $^{-3}$, характерной для металла, $L = 0,5 \cdot 10^{-10}$ см.

Таким образом, в случае полупроводника поле проникает далеко вглубь, а в случае металла оно сосредоточено в пределах монокристаллического слоя.

Следует отметить, что приведенный расчет является оценочным и позволяет говорить лишь о порядках величин проникновения поля в твердое тело. Ниже он будет уточнен.

Если для металлов проникновение поля на расстояние порядка межатомных расстояний практически не изменит энергетической картины металла, изображенной на рис. 1.1, то для полупроводника наличие значительного слоя с измененной концентрацией носителей заряда приведет к энергетической картине, изображенной на рис. 1.3.

Эта картина обусловлена изменением концентрации носителей заряда вследствие проникновения электрического поля в полупроводник. В случае, изображенном на рис. 1.3, а, в результате действия электрического поля из приконтактного слоя уходят электроны и притягиваются дырки.

Положительный заряд в полупроводнике, равный отрицательному заряду тела с большей работой выхода, в этом случае обуславливается главным образом для n -типа полупроводника зарядом ионизированных доноров и зарядом свободных дырок (рис. 1.3, б). В случае, изображенном на рис. 1.3, б, в результате действия электрического поля в приповерхностной области скапливаются электроны полупроводника. Их заряд компенсирует равный по величине положительный заряд тела с меньшей работой выхода (рис. 1.3, в).

В первом случае у поверхности образуется слой пространственного заряда, обедненный основными носителями заряда. Его сопротивление при пренебрежении неосновными носителями заряда больше, чем аналогичного слоя в объеме полупроводника. Такой слой называется зазорным, или обедненным, слоем.

Если неосновными носителями заряда пренебречь нельзя, то у самой поверхности может образоваться инверсный слой с малым сопротивлением, где неосновных носителей заряда больше, чем основных. При этом инверсный и зазорный слои существуют одновременно и расположены один за другим. Во втором случае у поверхности образуется слой, обогащенный основными носителями заряда. Его сопротивление меньше, чем объема полупроводника. Он называется антизазорным, или обогащенным, слоем.

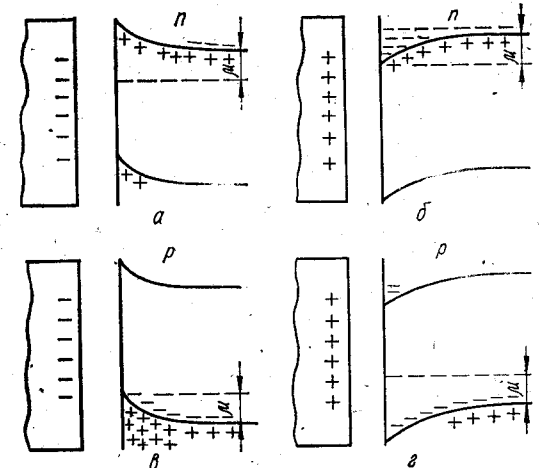


Рис. 1.3. Изгибы зон у поверхности n - и p -типа полупроводника при разном соотношении между работами выхода контактирующих тел: а, в — $\Phi_1 > \Phi_2$; б, г — $\Phi_1 < \Phi_2$.

Характер возникающего при контакте изгиба зон определяется только разностью работ выхода между контактирующими веществами. Поэтому такой же характер изгиба зон сохранится и для p -типа полупроводника (рис. 1.3, в, г). Однако поскольку основными носителями заряда для p -типа являются дырки, то в случае, когда $\varphi_1 > \varphi_2$ заряд области пространственного заряда будет обуславливаться свободными дырками и образуется антизапорный слой. При $\varphi_1 < \varphi_2$ заряд будет обуславливаться ионизированными

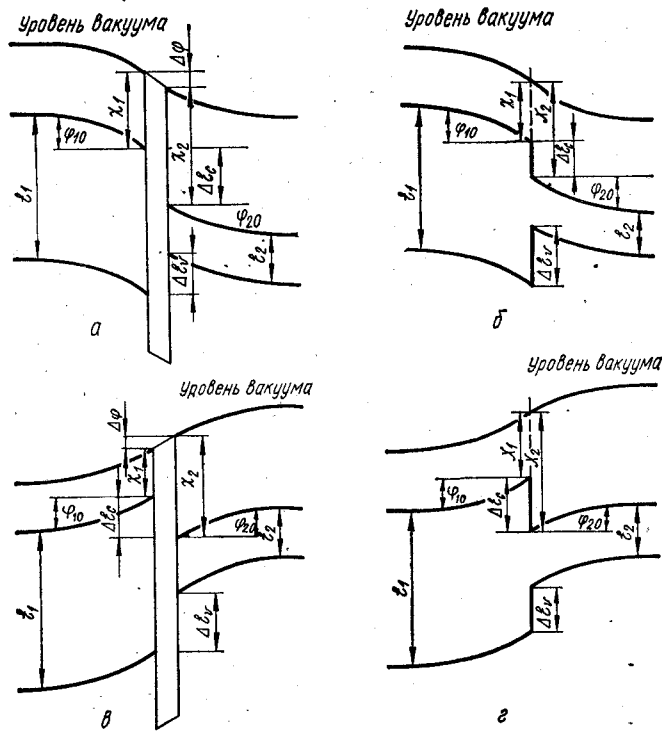


Рис. 1.4. Энергетические диаграммы гетеропереходов.

акцепторами и свободными электронами и образуется запиорный слой или инверсный и запиорный слой.

Теперь рассмотрим изгибы зон для двух контактирующих полупроводников при их сближении, учитывая, что поле проникает в оба полупроводника. Для определенности будем полагать, что ширина запрещенной зоны первого полупроводника больше, чем второго $\mathcal{E}_1 > \mathcal{E}_2$. Тогда для $\varphi_{1s} > \varphi_{2s}$ имеем картину, изображенную на рис. 1.4, а, б, а для $\varphi_{1s} < \varphi_{2s}$ — изображенную на рис. 1.4, в, г.

Как видно, направление изгиба зон в контактирующих полупроводниках различно, а величина изгиба зон зависит от расстояния между ними.

В случае рассмотренного гетероперехода выполняется соотношение

$$\varphi_{10} + \varphi_{20} + \Delta\varphi = \varphi_{1s} - \varphi_{2s}, \quad (1.2)$$

где φ_{10} и φ_{20} — части контактной разности потенциалов, падающие в полупроводниках 1 и 2, а $\Delta\varphi$ — часть контактной разности потенциалов в зазоре.

При уменьшении зазора $\Delta\varphi$ стремится к нулю и

$$\varphi_{10} + \varphi_{20} = \varphi_{1s} - \varphi_{2s}.$$

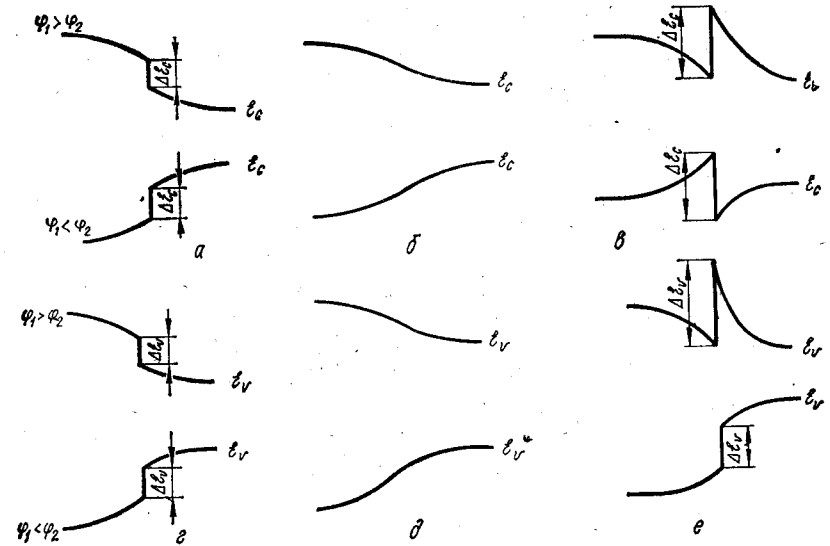


Рис. 1.5. Энергетические диаграммы для дна зоны проводимости (а—д) и валентной зоны (е—ж) в тесном гетеропереходе.

Для гетероперехода важным является существование разрыва в зонах проводимости и валентной зоне. Эти разрывы обуславливаются разными значениями внешних работ выхода χ_1 и χ_2 для контактирующих полупроводников. В самом деле, как легко видеть из рис. 1.4, а, в, разрыв в зонах проводимости

$$\Delta\mathcal{E}_c = \chi_2 - \chi_1 \pm \Delta\varphi, \quad (1.3)$$

где знак плюс при $\Delta\varphi$ относится к рис. 1.4, а, а знак минус к рис. 1.4, б.

В случае, когда зазор между полупроводниками исчезает (рис. 1.4, б, г),

$$\Delta\mathcal{E}_c = \chi_2 - \chi_1. \quad (1.4)$$

Разрыв в валентных зонах равен (рис. 1.4, а, в)

$$\Delta\mathcal{E}_v = \mathcal{E}_1 - \mathcal{E}_2 - \Delta\mathcal{E}_c = (\mathcal{E}_1 - \mathcal{E}_2) + (\chi_1 - \chi_2) \mp \Delta\varphi, \quad (1.5)$$

где знак минус при $\Delta\varphi$ относится к рис. 1.4, а, а знак плюс к рис. 1.4, б.

При отсутствии зазора $\Delta\varphi = 0$ и

$$\Delta\mathcal{E}_v = \mathcal{E}_1 - \mathcal{E}_2 + \chi_1 - \chi_2. \quad (1.6)$$

В зависимости от соотношения между χ_1 и χ_2 возможны различные значения $\Delta\mathcal{E}_v$. Три принципиально различных случая $\chi_1 > \chi_2$, $\chi_1 = \chi_2$ и $\chi_1 < \chi_2$ для разрывов в зоне проводимости тесного гетероперехода при $\varphi_{1s} > \varphi_{2s}$ показаны на рис. 1.5, а—в.

Возможные разрывы в валентных зонах для тесного гетероперехода, когда отсутствует зазор и соответственно

$$\mathcal{E}_1 - \mathcal{E}_2 > \chi_1 - \chi_2, \quad \mathcal{E}_1 - \mathcal{E}_2 = \chi_1 - \chi_2, \quad \mathcal{E}_1 - \mathcal{E}_2 < \chi_1 - \chi_2,$$

показаны на рис. 1.5, г—е.

Используя представленные на рис. 1.5 случаи, легко сконструировать любые виды гетеропереходов, как это изображено на рис. 1.6, на котором заштрихованная полоса значений, соответствующая краю валентной зоны и зоны проводимости, показывает различные значения $\Delta\mathcal{E}_v$ и $\Delta\mathcal{E}_c$.

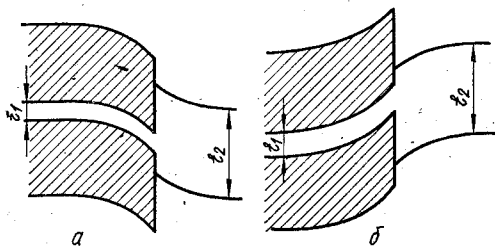


Рис. 1.6. Возможное энергетическое строение тесных гетеропереходов:
а — $\varphi_{1s} > \varphi_{2s}$; б — $\varphi_{1s} < \varphi_{2s}$.

Следует отметить, что энергетическая картина изгибов зон определяется только разностью работ выхода двух полупроводников и одинакова для анизотипных гетеропереходов между полупроводниками с разным типом проводимости ($p-n$ или $n-p$) и изотипных гетеропереходов между полупроводниками с одним типом проводимости ($n-n$ или $p-p$). Однако при одинаковом характере изгиба зон для анизотипных и изотипных гетеропереходов образуются различные области пространственного заряда. Так, для $p-n$ перехода при $\varphi_{1s} > \varphi_{2s}$ в p - и n -полупроводниках образуются запиорные, а при $\varphi_{1s} < \varphi_{2s}$ антизапиорные слои. Для изотипного $n-n$ перехода образуется при $\varphi_{1s} > \varphi_{2s}$ в полупроводнике 1 антизапиорный слой и в полупроводнике 2 антизапиорный и при $\varphi_{1s} < \varphi_{2s}$ в полупроводнике 1 — запиорный, а в полупроводнике 2 — антизапиорный. В случае $p-p$ перехода при $\varphi_{1s} > \varphi_{2s}$ для полупроводника 1 характерен запиорный изгиб зон и для полупроводника 2 — антизапиорный, а при $\varphi_{1s} < \varphi_{2s}$ — наоборот.

Как видно, в $p-n$ гетеропереходах в обоих полупроводниках образуются запиорные или антизапиорные слои, а в изотипных гетеропереходах в одном полупроводнике образуется запиорный, а в другом — антизапиорный слой.

Частным случаем рассмотренной здесь картины являются контакт металл — полупроводник и классический $p-n$ переход. В самом деле, при $\mathcal{E}_1 \rightarrow 0$ и пренебрежении падением напряжения в левом твердом теле имеем энергетическую картину, которая характер-

на для контакта металл — полупроводник. При $\varphi_m > \varphi_s$ для n -типа полупроводника это показано на рис. 1.7.

Величина изгиба зон в полупроводнике в этом случае равна

$$\varphi_0 = \varphi_m - \varphi_s - \Delta\varphi.$$

Изгиб зон максимален и равен $\varphi_0 = \varphi_m - \varphi_s$, когда между металлом и полупроводником не остается зазора и контакт тесный. Для двух одинаковых полупроводников $\chi_2 = \chi_1$ и $\mathcal{E}_1 = \mathcal{E}_2$. При этом разрывы в разрешенных зонах исчезают и мы имеем картину,

изображенную для $\varphi_{1s} > \varphi_{2s}$ на рис. 1.8 и характерную для $p-n$ перехода.

Рассматривая рис. 1.4—1.6, можно видеть, что для гетероперехода высоты потенциального

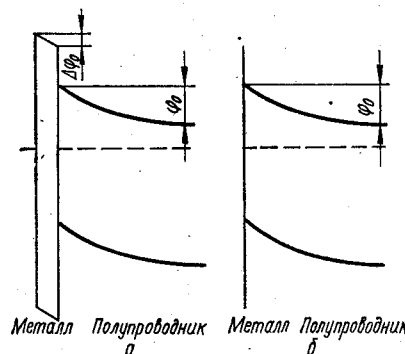


Рис. 1.7. Энергетическая диаграмма контакта металл — полупроводник.

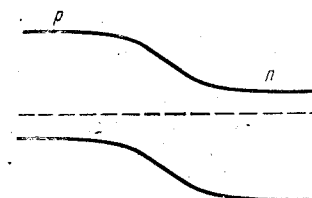


Рис. 1.8. Схема $p-n$ перехода.

барьера для электронов n -полупроводника и для дырок p -полупроводника в общем случае различны и равны только для классического $p-n$ перехода. Это различие высот барьера является одним из основных достоинств гетероперехода и позволяет существенно изменять соотношения между электронным и дырочным токами.

Однако, как уже отмечалось выше, рассмотренная здесь картина образования гетеро- и гомоперехода является идеализированной. Дело в том, что обычно на поверхности полупроводников существуют поверхностные электронные состояния, на которых может накапливаться заряд. Кроме того, поверхность полупроводника или металла, приготовленная стандартным образом, обычно до контактирования покрыта окисным слоем сложного состава или слоем адсорбированных газов.

Если поверхностными состояниями можно пренебречь, то все рассмотренное выше остается справедливым. В этом случае нельзя только получить тесного контакта двух тел, а свойства зазора определяются свойствами окисных и адсорбированных слоев.

Сложнее обстоит дело, если учитывать электронные поверхностные состояния на поверхностях изолированных твердых тел. В этом случае еще до контактирования в полупроводниках существует изгиб зон, как показано, например, на рис. 1.9, где для конкретности учтена зона поверхностных состояний. Уровни поверхностной зоны, расположенные ниже уровня Ферми, заряжены отрицательно, а уровни, расположенные выше уровня Ферми,

нейтральны. Отрицательный заряд на уровнях обуславливает изгиб зон вверх, что соответствует запирающему слою для n -типа полупроводника (рис. 1.9, а) и антизапирающему слою для p -типа полупроводника (рис. 1.9, б).

Рассмотрим теперь, что будет, если сближать такой полупроводник, на поверхности которого есть электронные состояния, с металлом или полупроводником с другой работой выхода. Пусть для определенности $\phi_{2s} < \phi_{1s}$ или ϕ_m , где индекс 2 относится к рассматриваемому правому полупроводнику. Тогда, как видно из уравнения (1.1), большее количество электронов будет переходить из рассматриваемого полупроводника в соседнее вещество, которое будет заряжаться отрицательно. Перешедший заряд отталкивает электроны и притягивает дырки в полупроводнике 2. Это должно привести к увеличению изгиба зон в n -типе и уменьшению в p -типе. Уровень Ферми будет опускаться по отношению к поверхностной зоне, что приведет к уменьшению отрицательного заряда на поверхностных состояниях. Такое уменьшение само по себе приводит к уменьшению изгиба зон для n -типа и увеличению для p -типа, т. е. приводит к противоположному действию по сравнению с характером изменения, вызываемым разностью работ выхода. Это приводит к тому, что при наличии поверхностных состояний изменение изгиба зон при контактировании будет меньшим, чем для рассмотренного выше случая, когда поверхностных состояний нет. Для контакта металл — n -полупроводник это иллюстрируется рис. 1.10, а, б, а для контакта полупроводник — полупроводник — рис. 1.10, в, г.

Рис. 1.9. Изгиб зон на свободной поверхности, обусловленной поверхностными состояниями.

Как видно из рисунка, контактная разность потенциалов падает в зазоре между металлом и полупроводником, а заряд перераспределяется между контактирующими веществами и поверхностными состояниями так, что заряд в области пространственного заряда изменяется мало (в частном случае, изображенном на рис. 1.10, предполагается, что он постоянен).

Если сближать полупроводник с металлом или другим полупроводником, для которых выполняется обратное соотношение $\phi_{2s} > \phi_{1s}$ или $\phi_{2s} > \phi_m$, то большее количество электронов будет переходить в рассматриваемый полупроводник. Изгиб зон при этом для полупроводника должен уменьшаться. Но уровень Ферми при таком процессе должен перемещаться относительно поверхностной зоны вверх. Это в свою очередь приводит к увеличению заряда на поверхностных состояниях и уменьшению изменения изгиба зон.

При достаточно большой плотности поверхностных состояний изгиб не будет изменяться при изменении разности работ выхода твердого тела.

Все другие возможные энергетические структуры контактов можно рассматривать, используя описанную выше методику. При этом следует учитывать, что изгибы зон для гетеропереходов определяются не только значением работ выхода, но и поверхностными состояниями. Поэтому количество возможных энергетических структур для гетеропереходов увеличивается. Например, для рассмотренного на рис. 1.10 примера в изотипном гетеропереходе существуют два запирающих слоя.

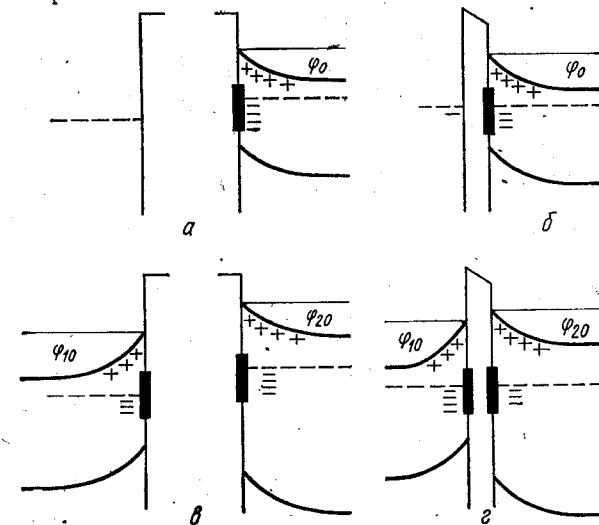


Рис. 1.10. Примеры энергетических структур контакта металл — полупроводник и гетероперехода при учете поверхностных состояний.

Однако рассмотренный здесь путь получения гетеро- и гомопереходов прямым контактированием твердых тел применяется редко. Обычно применяются другие технологические методы их получения. Это жидкостная эпитаксия, осаждение из газообразных соединений, химическое и электрохимическое осаждение, сплавление, диффузия, ионная имплантация и т. д. При этом полупроводник или металл наносятся на предварительно подготовленную поверхность выбранного другого полупроводника. Для такой подготовки применяются механическая обработка, химическое, электрохимическое или газовое травление.

С точки зрения рассмотренного выше подхода все методы получения гетеро- и гомопереходов можно разбить на две группы в зависимости от степени изменения свойств полупроводника, на котором образуется гетеро- или гомопереход.

К первой группе можно отнести методы, которые существенно не нарушают свойств поверхности полупроводника, на который наносится металл или другой полупроводник. Сюда может быть, например, отнесено термическое напыление при небольших температурах подложки.

Ко второй группе могут быть отнесены методы, когда первоначальная поверхность полупроводника существенно изменяется или образование перехода происходит в глубине полупроводника, например происходит образование $p-n$ перехода при диффузии или ионной имплантации.

При применении методов, существенно не изменяющих свойств первоначальной поверхности, физическая модель перехода должна отражать эти свойства. В этом случае остается справедливым все сказанное выше относительно образования гетеро- и гомоперехода.

Если же в результате получения перехода существенно изменяются свойства первоначальной поверхности, картина оказывается сложнее. Но все же и в этом случае для общности нужно учитывать зазор, который теперь является скорее промежуточным слоем между первым и вторым твердым телом, а также поверхностные состояния, которые обычно существуют на границе раздела любых двух фаз. Если ни промежуточного слоя, ни поверхностных состояний нет, это частный случай более общей картины, проанализированной выше.

Поэтому рассмотренная выше картина образования гетеро- и гомоперехода в результате простого контактирования является общей модельной картиной, пригодной для всех случаев. Из нее следует, что характерным для гетеро- или гомоперехода является образование потенциальных барьеров для электронов и дырок, переходящих из одной части контакта в другую. Именно эти барьеры и определяют основные свойства контактов. Кроме того, в ряде случаев существенную роль играют зазор или переходный слой и поверхностные электронные состояния.

1.2. ХОД ПОТЕНЦИАЛА В ОБЛАСТИ ПРОСТРАНСТВЕННОГО ЗАРЯДА ПОЛУПРОВОДНИКА

Для количественного рассмотрения хода потенциала $V(x)$ в области пространственного заряда любого из полупроводников можно воспользоваться уравнением Пуассона, которое для одномерного случая имеет вид

$$\frac{d^2V(x)}{dx^2} = -\frac{\rho(x)}{\epsilon_0 \epsilon}, \quad (1.7)$$

где $\rho(x)$ — плотность объемного заряда, ϵ_0 и ϵ — диэлектрические проницаемости вакуума и полупроводника.

Уравнение Пуассона может быть использовано и для нахождения хода энергии электронов у краев разрешенных зон ϕ , если учесть, что $\phi(x) = -eV(x)$ (благодаря отрицательному заряду электрона кривые для энергии и электростатического потенциала имеют противоположный ход и появляется знак минус). Тогда

$$\frac{d^2\phi(x)}{dx^2} = \frac{e\rho(x)}{\epsilon_0 \epsilon}. \quad (1.8)$$

Используя уравнение Пуассона, найдем зависимость потенциальной энергии $\phi(x)$, соответствующей дну зоны проводимости или верху валентной зоны полупроводника от расстояния. Для этого нужно знать $\rho(x)$. Вначале ограничимся равномерным распределением примесей. Будем для определенности рассматривать n -типа полупроводник, в области пространственного заряда которого можно пренебречь дискретностью заряда ионизированных примесей (только в этом случае зависимость $\phi(x)$ носит монотонный характер).

Плотность заряда в любой точке n -типа полупроводника

$$\rho(x) = e(n_d - n_{э,д} - n + p),$$

где $n_d - n_{э,д}$ — концентрация ионизированных доноров (n_d — концентрация донорных уровней; $n_{э,д}$ — концентрация электронов на них).

В объеме полупроводника число свободных электронов равняется сумме ионизированных доноров и числа дырок в валентной зоне:

$$p_0 + (n_d - n_{э,д}) - n_0 = 0.$$

Подставляя найденное отсюда значение $n_d - n_{э,д}$ в (2.1), имеем

$$\rho(x) = e[(p - p_0) - (n - n_0)].$$

Если отсчитывать $\phi(x)$ от дна зоны проводимости в глубине полупроводника, то в термодинамически равновесном состоянии, когда через систему не идет ток и нет вырождения и, следовательно, для электронов выполняется распределение Максвелла, концентрация электронов для равномерного распределения примесей в области пространственного заряда

$$n(x) = n_0 e^{-\frac{\phi(x)}{kT}}. \quad (1.9a)$$

В этих же предположениях концентрация дырок в валентной зоне

$$p(x) = p_0 e^{\frac{\phi(x)}{kT}}. \quad (1.9b)$$

Все это позволяет записать суммарный заряд в виде

$$\rho(x) = e[p_0(e^{\frac{\phi(x)}{kT}} - 1) - n_0(e^{-\frac{\phi(x)}{kT}} - 1)]. \quad (1.10)$$

Графически изменение с расстоянием положительного, отрицательного и суммарного заряда представлено на рис. 1.11.

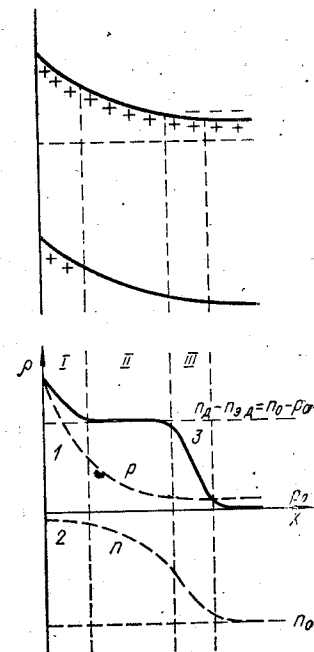


Рис. 1.11. Зависимость плотности зарядов от расстояния.

Однако уравнение (1.10) с $\rho(x)$, задаваемым выражением (1.10), аналитически в общем случае не решается. Поэтому в большинстве случаев для полупроводниковых приборов используют приближение, при котором область пространственного заряда является слоем Шоттки, приближение, при котором уравнение (1.10) имеет аналитическое решение.

Область пространственного заряда, называемая слоем Шоттки, реализуется, когда в этой области можно пренебречь зарядом свободных электронов и дырок по сравнению с зарядом ионизированных примесей. Для n -типа полупроводника это означает, что везде в области пространственного заряда

$$n_0 \gg n_0 e^{-\frac{\varphi(x)}{kT}}; \quad (1.11)$$

$$n_0 \gg p_0 e^{\frac{\varphi(x)}{kT}}. \quad (1.12)$$

Первое неравенство, строго говоря, выполнить при $\varphi(x) \rightarrow 0$ нельзя. Поэтому условно принимают, что слой Шоттки начинается не при $\varphi = 0$, а при $\varphi(x) = 2,3 kT$, когда $e^{\frac{\varphi(x)}{kT}} = 0,1$. Это значение $\varphi(x)$ и ограничивает минимальную высоту барьера, при которой реализуется слой Шоттки.

Второе неравенство будет выполняться для значений $\varphi(0) = \varphi_0$, ограниченных сверху. Действительно, когда $n_0 \gg p_0 e^{\frac{\varphi_0}{kT}}$, для значений, меньших $\varphi(x)$, чем φ_0 , оно выполняется автоматически.

Неравенство (1.12) можно переписать в виде

$$\frac{n_0}{p_0} = \frac{n_0 p_0}{p_0^2} \gg e^{\frac{\varphi_0}{kT}}.$$

Если использовать соотношения $n_0 p_0 = N_c N_v e^{-\frac{\xi_g}{kT}}$ и $p_0 = N_v e^{-\frac{\xi_g - \mu}{kT}}$, то последнее неравенство преобразуется к виду:

$$\left(\frac{m_n^*}{m_p^*}\right)^{3/2} e^{\frac{\xi_g - 2\mu}{kT}} \gg e^{\frac{\varphi_0}{kT}}.$$

Отсюда должно выполняться неравенство для показателей экспонент

$$\varphi_0 < \xi_g - 2\mu + \frac{3}{2} kT \ln \frac{m_n^*}{m_p^*}.$$

Теперь окончательно условие реализации слоя Шоттки может быть записано в виде

$$2kT < \varphi_0 < \xi_g - 2\mu + \frac{3}{2} kT \ln \frac{m_n^*}{m_p^*}, \quad (1.13)$$

где нижний предел обуславливается пренебрежением заряда свободных электронов, а верхний предел — пренебрежением заряда свободных дырок по сравнению с зарядом ионизированных доноров. Как видно из последнего неравенства, слой Шоттки не может быть получен для собственного полупроводника, в котором $\mu \simeq \frac{\xi_g}{2}$. Кроме того, максимальное значение высоты барьера Шоттки для невырожденного полупроводника, когда $\mu \geq 0$, не может превы-

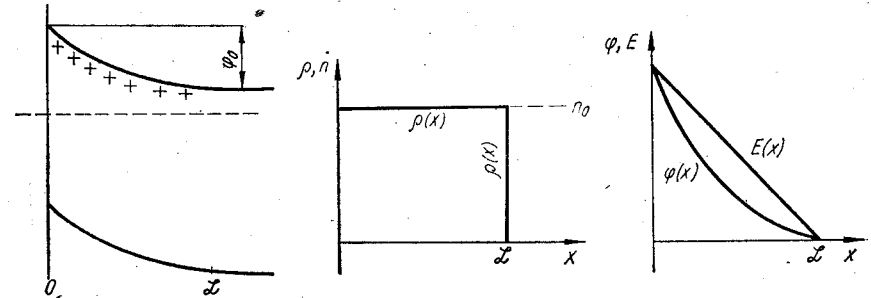


Рис. 1.12. Слой Шоттки при равномерном распределении примеси.

шать ширины запрещенной зоны, т. е. $\varphi_0 \leq \xi_g$. Следует особо отметить, что рассмотренные здесь критерии для существования слоя Шоттки годятся только для термодинамически равновесных условий, когда к гетеро- или гомопереходу не приложено внешнее напряжение. Таким образом, в случае n -типа полупроводника, когда $n_0 \gg p_0$, для слоя Шоттки при равномерном распределении примесей $\rho(x) = en_0$ и для нахождения $\varphi(x)$ необходимо решать уравнение

$$\frac{d^2\varphi}{dx^2} = \frac{e^2 n_0}{\epsilon_0 \epsilon} \quad (1.14)$$

при граничных условиях:

$$\begin{aligned} x=0, \quad \varphi(0) &= \varphi_0, \quad E(0) = E_0; \\ x=\mathcal{L}, \quad \varphi(\mathcal{L}) &= 0, \quad E(\mathcal{L}) = 0. \end{aligned} \quad (1.15)$$

Интегрируя дважды уравнение (1.14), получаем

$$\varphi(x) = \frac{e^2 n_0}{\epsilon_0 \epsilon} \left(\frac{x^2}{2} + Ax + B \right).$$

Используя граничное условие $E(\mathcal{L}) = 0$ и $\varphi(\mathcal{L}) = 0$, находим $A = -\mathcal{L}$, $B = \frac{\mathcal{L}^2}{2}$. В результате можно записать:

$$\varphi(x) = \frac{e^2 n_0}{2\epsilon_0 \epsilon} (\mathcal{L} - x)^2; \quad (1.16)$$

$$E(x) = \frac{1}{e} \cdot \frac{d\varphi}{dx} = -\frac{en_0}{\epsilon_0 \epsilon} (\mathcal{L} - x). \quad (1.17)$$

Изгиб зон, соответствующий слою Шоттки, зависимость $\rho(x)$ и $n(x)$, а также $\varphi(x)$ и $E(x)$ для слоя Шоттки при равномерном распределе-

нии примесей представлены на рис. 1.12. На этом рисунке значения φ_0 и E_0 для простоты совмещены. А вообще между ними существует простая аналитическая связь

$$E_0 = -\frac{2\varphi_0}{e\mathcal{L}} = -\left(\frac{2\varphi_0 n_0}{\varepsilon_0 e}\right)^{1/2},$$

которую легко получить, если положить в (1.16) и (1.17) $x = 0$.

Значение ширины слоя Шоттки \mathcal{L} может быть найдено, если положить в (1.16) $x = 0$:

$$\mathcal{L} = \left(\frac{2\varepsilon_0 e \varphi_0}{e^2 n_0}\right)^{1/2}. \quad (1.18)$$

При $\varepsilon = 16$, $\varphi_0 = 0,3$ эВ и $n_0 = 10^{14}$ см $^{-3}$ $\mathcal{L} = 2,3 \cdot 10^{-11}$ см, а при тех же значениях ε и φ_0 , но $n_0 = 10^{18}$ см $^{-3}$ $\mathcal{L} = 2,3 \cdot 10^{-6}$ см. Таким образом, типичные значения \mathcal{L} лежат в пределах $10^{-4} \dots 10^{-6}$ см. Значения \mathcal{L} больше для полупроводников, характеризующихся меньшим значением n_0 .

Все рассмотренное выше относится к случаю, когда распределение примеси в полупроводнике равномерное. Если примесь распределена неравномерно, то необходимо учитывать связь $\rho(x)$ с распределением примеси. Например, для случая, когда примесь распределена по линейному закону

$$n = ax = \frac{n_0}{A} x \quad (1.19)$$

и слой Шоттки лежит в области изменения n , нахождение $\varphi(x)$ сводится к решению уравнения Пуассона.

$$\frac{d^2\varphi}{dx^2} = \frac{e^2 a}{\varepsilon_0 e} x$$

при граничных условиях:

$$\begin{aligned} x = 0, \quad \varphi(0) = \varphi_0, \quad E(0) = E_0; \\ x = D, \quad \varphi(D) = 0, \quad E(D) = 0, \end{aligned}$$

где D — ширина слоя Шоттки при неравномерном распределении примесей.

Следует отметить, что граничные условия при $x = D$ не являются строгими, так как при $\frac{dn}{dx} \neq 0$ и φ отлично от нуля. Такая запись по существу означает, что $\varphi(0) \gg \varphi(D)$ и $E(0) \gg E(D)$.

В результате интегрирования уравнения Пуассона при линейном распределении примесей получаем:

$$\varphi(x) = \frac{e^2}{\varepsilon_0 e} a \left(\frac{x^3}{6} - \frac{x D^3}{2} + \frac{D^3}{3} \right); \quad (1.20)$$

$$E(x) = -\frac{e}{2\varepsilon_0 e} a (D^2 - x^2). \quad (1.21)$$

Как видно из этих выражений, $\varphi(x)$ и $E(x)$ иначе зависят от расстояния, чем в случае равномерного распределения примесей.

Положив $x = 0$, можно найти $\varphi_0 = \frac{e^2}{\varepsilon_0 e} \cdot \frac{a}{3} D^3$. Отсюда

$$D = \left[\frac{3\varphi_0 \varepsilon_0 e}{e^2 a} \right]^{1/3}. \quad (1.22)$$

Из последнего выражения видно, что ширина слоя Шоттки D при линейном распределении примесей изменяется с φ_0 как корень кубический, а не квадратный.

Для неравномерного распределения примесей изменяется связь $\varphi(x)$, $E(x)$ и D с другими параметрами полупроводника.

Величины $\varphi(0)$ и $E(0)$ для неравномерного распределения примесей могут быть легко получены на основе уравнений (5.5) и (5.6).

1.3. СООТНОШЕНИЕ МЕЖДУ ВЫСОТАМИ БАРЬЕРОВ В ГЕТЕРОПЕРЕХОДЕ

Используя результаты предыдущего параграфа, можно рассмотреть численные соотношения между высотами барьеров в гетеропереходах. Для общности будем учитывать зазор между различными полупроводниками и поверхностные состояния на границах полупроводник — зазор, как это показано на рис. 1.13.

Для нахождения связи φ_{10} и φ_{20} в гетеропереходе учтем условия непрерывности вектора электрической индукции на границах раздела первого и второго полупроводников с зазором и запишем систему уравнений:

$$\varepsilon_0 \varepsilon_1 E_1 = \varepsilon_0 \varepsilon_3 E_3 + \sigma_1; \quad (1.23)$$

$$\varepsilon_0 \varepsilon_2 E_2 = \varepsilon_0 \varepsilon_2 E_2 + \sigma_2; \quad (1.24)$$

$$\varphi = \varphi_{10} + \varphi_{20} + \Delta\varphi, \quad (1.25)$$

где

$$E_3 = \frac{\Delta\varphi}{ed}$$

и предполагается, что величины плотности поверхностных зарядов σ_1 и σ_2 известны.

Записанная система уравнений позволяет найти три неизвестных φ_{10} , φ_{20} и $\Delta\varphi$, если использовать выражения для E_1 и E_2 из предыдущего параграфа.

Аналитически эта система может быть решена для случая, когда в обоих полупроводниках область пространственного заряда является слоем Шоттки.

Подставив в выражения (1.23) и (1.24)

$$E_1 = -\left(\frac{2\varphi_{10} p_1}{\varepsilon_0 \varepsilon_1}\right)^{1/2}; \quad E_2 = -\left(\frac{2\varphi_{20} n_2}{\varepsilon_0 \varepsilon_2}\right)^{1/2},$$

можно получить из первых двух уравнений (1.23) и (1.24) φ_{10} и φ_{20} , а используя соотношение (1.25), $\Delta\varphi$:

$$\varphi_{10} = \frac{1}{2\varepsilon_0 \varepsilon_1 p_1} \left(\sigma_1 + \frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi \right)^2; \quad (1.26)$$

$$\varphi_{20} = \frac{1}{2\varepsilon_0 \varepsilon_2 n_2} \left(\frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi - \sigma_2 \right)^2; \quad (1.27)$$

$$\Delta\varphi = \frac{-b + \sqrt{b^2 - 4ac}}{a}, \quad (1.28)$$

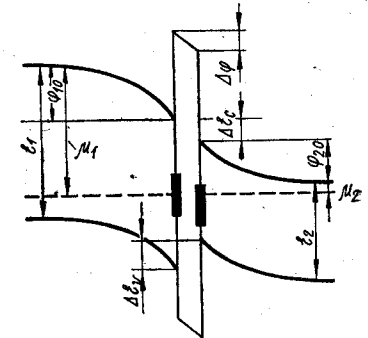


Рис. 1.13. Энергетическая диаграмма гетероперехода.

где

$$a = \frac{1}{2} \left(\frac{\epsilon_0 \epsilon_3}{ed} \right)^2 \left[\frac{1}{\epsilon_0 \epsilon_1 \rho_1} + \frac{1}{\epsilon_0 \epsilon_2 n_2} \right];$$

$$b = 1 + \frac{\epsilon_0 \epsilon_3}{ed} \left[\frac{\sigma_1}{\epsilon_0 \epsilon_1 \rho_1} + \frac{\sigma_2}{\epsilon_0 \epsilon_2 n_2} \right];$$

$$c = \frac{\sigma_1^2}{2\epsilon_0 \epsilon_1 \rho_1} + \frac{\sigma_2^2}{2\epsilon_0 \epsilon_2 n_2} - \varphi.$$

Как видно из полученных выражений, величины φ_{10} , φ_{20} и $\Delta\varphi$ зависят от ширины зазора и его диэлектрической постоянной, от концентраций носителей заряда в обоих полупроводниках и их диэлектрических проницаемостей, а также от заполнения поверхностных состояний.

При отсутствии поверхностных состояний $\sigma_1 = \sigma_2 = 0$ и полученные выражения упрощаются:

$$\varphi_{10} = \frac{1}{2} \left(\frac{\epsilon_0 \epsilon_3}{ed} \right)^2 \frac{(\Delta\varphi)^2}{\epsilon_0 \epsilon_1 \rho_1} = \frac{\epsilon_2 n_2 (\varphi - \Delta\varphi)}{\epsilon_1 \rho_1 + \epsilon_2 n_2}; \quad (1.29)$$

$$\varphi_{20} = \frac{1}{2} \left(\frac{\epsilon_0 \epsilon_3}{ed} \right)^2 \frac{(\Delta\varphi)^2}{\epsilon_0 \epsilon_2 n_2} = \frac{\epsilon_1 \rho_1 (\varphi - \Delta\varphi)}{\epsilon_1 \rho_1 + \epsilon_2 n_2}; \quad (1.30)$$

$$\Delta\varphi = -A + \sqrt{A^2 + 4A\varphi}, \quad (1.31)$$

где

$$A = 2 \left(\frac{ed}{\epsilon_3} \right)^2 \frac{\epsilon_1 \epsilon_2 n_2 \rho_1}{\epsilon_0 \epsilon_1 \rho_1 + \epsilon_0 \epsilon_2 n_2}.$$

Как показывает анализ этих выражений, отношение $\frac{\varphi_{10}}{\varphi_{20}} = \frac{\epsilon_2 n_2}{\epsilon_1 \rho_1}$ и, следовательно, φ_{10} больше, а φ_{20} меньше для случая, когда $\epsilon_1 \rho_1 > \epsilon_2 n_2$, т. е. основное падение потенциала происходит в области с меньшей концентрацией носителей заряда. Величина $\Delta\varphi$ тем меньше, чем тоньше зазор и меньше концентрация носителей заряда в обоих полупроводниках.

Наличие поверхностных уровней приводит к уменьшению φ_{10} и φ_{20} и увеличению $\Delta\varphi$ при постоянных остальных параметрах.

В случае контакта металл — полупроводник $\rho_1 \gg n_2$ и $\varphi_{10} = 0$, а $\varphi_{20} = \varphi_0 = \varphi - \Delta\varphi$, где при $\sigma_2 = \sigma_1 = 0$, $\Delta\varphi$ определяется выражением (1.31) при

$$A = 2 \left(\frac{\epsilon_0 \epsilon_3 ed}{\epsilon_0 \epsilon_3} \right)^2 \frac{n_2}{\epsilon_0 \epsilon_1}.$$

Для тесного гетероперехода при $d = 0$, $\Delta\varphi = 0$:

$$\varphi_{10} = \frac{\epsilon_2 n_2 \varphi}{\epsilon_1 \rho_1 + \epsilon_2 n_2}; \quad (1.32)$$

$$\varphi_{20} = \frac{\epsilon_1 \rho_1 \varphi}{\epsilon_1 \rho_1 + \epsilon_2 n_2}. \quad (1.33)$$

Значение $\varphi = \varphi_{10} + \varphi_{20} + \Delta\varphi$, которое рассматривалось выше, может быть найдено также из разности между положениями уровней

энергии, соответствующих дну зоны проводимости или валентной зоны для различных полупроводников.

В самом деле, легко показать (см. рис. 1.13), что

$$\varphi = \xi_1 - \mu_2 - \mu_1 - \Delta\xi_c - \Delta\varphi, \quad (1.34)$$

или

$$\varphi = \xi_2 - \mu_2 - \mu_1 + \Delta\xi_v - \Delta\varphi,$$

где положительные значения $\Delta\xi_c$ и $\Delta\xi_v$ увеличивают разность между уровнями, соответствующими краям разрешенных зон в различных полупроводниках.

Учитывая, что $\rho_1 = N_v e^{-\frac{\mu_1}{kT}}$, а $n_2 = N_c e^{-\frac{\mu_2}{kT}}$, получаем

$$\varphi = \xi_1 - kT \ln \frac{N_{v1} N_{c2}}{n_2 \rho_1} - \Delta\xi_c - \Delta\varphi =$$

$$= kT \ln \frac{n_2 \rho_1}{n_{i1}^2} \left(\frac{m_2^*}{m_1^*} \right)^{3/2} - \Delta\xi_c - \Delta\varphi, \quad (1.35)$$

или

$$\varphi = \xi_2 - kT \ln \frac{N_{v1} N_{c2}}{n_2 \rho_1} - \Delta\varphi - \Delta\xi_v =$$

$$= kT \ln \frac{n_2 \rho_1}{n_{i2}^2} \left(\frac{m_1^*}{m_2^*} \right) - \Delta\xi_v - \Delta\varphi.$$

Здесь использовано соотношение $n_i^2 = N_v N_c e^{-\frac{\xi_g}{kT}}$, а значки «1» и «2» относятся к p - и n -типу полупроводника.

Из выражений (1.35) видно, что φ будет тем больше, чем больше величины n_1 и ρ_1 , и меньше значение n_{i1} и n_{i2} . В результате φ оказывается больше для $p-n$ гетеропереходов при сильном легировании p - и n -областей, а также для полупроводников, у которых шире запрещенная зона.

В случае тесного гомо- $p-n$ перехода $\Delta\varphi = 0$, $\Delta\xi_c = \Delta\xi_v = 0$ и выражение (1.35) переходит в такое:

$$\varphi = kT \ln \frac{\rho_1 n_2}{n_i^2}, \quad (1.36)$$

из которого можно рассчитывать φ , зная концентрации носителей заряда в p - и n -частях полупроводника.

Поскольку обычно экспериментально исследуют зависимость высот барьера от разности работ выхода между различными полупроводниками или металлом и полупроводником (крп), представляет также интерес зависимость φ_{10} , φ_{20} и $\Delta\varphi$ от величины контактной разности потенциалов V_k . Так как при существовании поверхностных состояний изгибы зон могут быть еще до контактирования (см. п. 1.1 § 11), V_k связано с φ_{10} , φ_{20} и $\Delta\varphi$ следующим соотношением:

$$eV_k = \varphi_{20} - \varphi_2^0 + \varphi_{10} - \varphi_1^0 + \Delta\varphi = \varphi - \varphi_2^0 - \varphi_1^0, \quad (1.37)$$

где φ_2^0 и φ_1^0 — высоты барьера до контактирования. Из последнего соотношения следует, что $\frac{dV_k}{d\varphi} = 1$. Это позволяет использовать полученные выше выражения для нахождения изменения φ_{10} , φ_{20} и $\Delta\varphi$ с V_k , дифференцируя их по φ .

Дифференцируя выражения (1.20) — (1.28), можно получить:

$$\frac{d\varphi_{10}}{d\varphi} = \frac{2\varphi_1 \left(\frac{d\sigma_1}{d\varphi} + \frac{\varepsilon_0 \varepsilon_3}{ed} \frac{d\Delta\varphi}{d\varphi} \right)}{\sigma_1 + \frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi}; \quad (1.38)$$

$$\frac{d\varphi_{20}}{d\varphi} = \frac{2\varphi_2 \left(\frac{\varepsilon_0 \varepsilon_3}{ed} \frac{d\Delta\varphi}{d\varphi} - \frac{d\sigma_2}{d\varphi} \right)}{\frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi - \sigma_2}; \quad (1.39)$$

$$\frac{d\Delta\varphi}{d\varphi} = \frac{1}{a} \left(-\frac{db}{d\varphi} + \frac{b \frac{db}{d\varphi} - 2a \frac{dc}{d\varphi}}{\sqrt{b^2 - 4ac}} \right), \quad (1.40)$$

где

$$\frac{db}{d\varphi} = \frac{\varepsilon_0 \varepsilon_3}{ed} \left(\frac{1}{\varepsilon_0 \varepsilon_1 \rho_1} \cdot \frac{d\sigma_1}{d\varphi} + \frac{1}{\varepsilon_0 \varepsilon_2 n_2} \cdot \frac{d\sigma_2}{d\varphi} \right);$$

$$\frac{dc}{d\varphi} = \frac{\sigma_1}{\varepsilon_0 \varepsilon_1 \rho_1} \frac{d\sigma_1}{d\varphi} + \frac{\sigma_2}{\varepsilon_0 \varepsilon_2 n_2} \frac{d\sigma_2}{d\varphi} - 1.$$

При отсутствии поверхностных состояний $\frac{d\sigma_1}{d\varphi} = 0$, $\frac{d\sigma_2}{d\varphi} = 0$,

$\frac{db}{d\varphi} = 0$ и $\frac{dc}{d\varphi} = -1$. При этом:

$$\frac{d\varphi_{10}}{d\varphi} = 2\varphi_1 \frac{d \ln \Delta\varphi}{d\varphi}; \quad (1.41)$$

$$\frac{d\varphi_{20}}{d\varphi} = 2\varphi_2 \frac{d \ln \Delta\varphi}{d\varphi}; \quad (1.42)$$

$$\frac{d\Delta\varphi}{d\varphi} = \frac{2}{\sqrt{b^2 - 4ac}}. \quad (1.43)$$

Анализ выражений (1.41) — (1.43) показывает, что $\frac{d\varphi_{10}}{d\varphi}$ и $\frac{d\varphi_{20}}{d\varphi}$ больше при меньших толщинах зазора и меньших концентрациях ρ_1 или n_2 . Отношение $\frac{d\varphi_{10}}{d\varphi} / \frac{d\varphi_{20}}{d\varphi} = \frac{\varphi_{10}}{\varphi_{20}}$ и чем больше высота барьера, тем больше его изменение с φ .

Существование поверхностных состояний приводит к уменьшению величин $\frac{d\varphi_{10}}{d\varphi}$ и $\frac{d\varphi_{20}}{d\varphi}$ и увеличению величины $\frac{d\Delta\varphi}{d\varphi}$.

Для контакта металл — полупроводник $\frac{d\varphi_{10}}{d\varphi} = 0$, $\frac{d\varphi_{20}}{d\varphi} = \frac{d\varphi_0}{d\varphi} = 1 - \frac{d\Delta\varphi}{d\varphi}$.

Для гетероперехода при $d = 0$ и $\sigma_1 = \sigma_2 = 0$:

$$\frac{d\varphi_{10}}{d\varphi} = \frac{1}{1 + \frac{\varepsilon_1 \rho_1}{\varepsilon_2 n_2}}; \quad (1.44)$$

$$\frac{d\varphi_{20}}{d\varphi} = \frac{1}{1 + \frac{\varepsilon_2 n_2}{\varepsilon_1 \rho_1}}. \quad (1.45)$$

Изменение φ_{20} , как уже и отмечалось, больше в полупроводнике с меньшей концентрацией носителей заряда.

В заключение проведенного выше анализа отметим, что хотя полученные выше аналитические выражения для величин φ_{10} , φ_{20} , $\Delta\varphi$, $\frac{d\varphi_{10}}{d\varphi}$ и $\frac{d\varphi_{20}}{d\varphi}$ относились к случаю, когда области пространственного заряда в обоих контактирующих полупроводниках являются слоями Шоттки, качественные выводы из этих выражений остаются справедливыми для любых областей пространственного заряда.

Например, если в результате контакта в полупроводнике 1 образуется антизапорный слой, а во 2 — запорный, то $\varphi_{10} \ll \varphi_{20}$ и $\frac{d\varphi_{10}}{d\varphi} \ll \frac{d\varphi_{20}}{d\varphi}$, поскольку концентрация носителей заряда в антизапорном слое значительно больше, чем в запорном.

Найденные выше соотношения между высотами барьера в гетеропереходе позволяют также найти суммарную ширину области пространственного заряда для случая, когда области пространственного заряда в полупроводниках 1 и 2 являются слоями Шоттки.

Учтем, что

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + d, \quad (1.46)$$

где \mathcal{L}_1 и \mathcal{L}_2 — ширина слоя Шоттки в полупроводниках 1 и 2.

Для этого домножим выражение (1.26) на $\frac{2\varepsilon_0 \varepsilon_1}{e^2 \rho_1}$, а выражение (1.27) — на $\frac{2\varepsilon_0 \varepsilon_2}{e^2 n_2}$. Тогда из выражения (1.26) имеем

$$\mathcal{L}_1 = \frac{1}{e\rho_1} \left(\sigma_1 + \frac{\varepsilon_0 \varepsilon_3 \Delta\varphi}{ed} \right),$$

а из (1.27)

$$\mathcal{L}_2 = \frac{1}{en_2} \left(\frac{\varepsilon_0 \varepsilon_3 \Delta\varphi}{ed} - \sigma_2 \right).$$

Подставляя \mathcal{L}_1 и \mathcal{L}_2 в (1.46), имеем

$$\mathcal{L} = \frac{1}{e\rho_1} \left(\sigma_1 + \frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi \right) + \frac{1}{en_2} \left(\frac{\varepsilon_0 \varepsilon_3}{ed} \Delta\varphi - \sigma_2 \right) + d. \quad (1.47)$$

Анализ последнего выражения показывает, что \mathcal{L} будет большим для больших значений σ_1 и σ_2 и меньших ρ_1 и n_2 . В частном случае тесного гетероперехода без поверхностных состояний, когда $d = 0$

и $\sigma_1 = \sigma_2 = 0$,

$$\mathcal{L} = \left[\frac{2\varepsilon_0\varepsilon_1\varepsilon_2\varphi}{e^2(\varepsilon_1\rho_1 + \varepsilon_2\rho_2)} \frac{(n_2 + \rho_1)}{n_2\rho_1} \right]^{1/4}. \quad (1.48)$$

Для классического $p - n$ перехода последнее равенство принимает вид

$$\mathcal{L} = \left[\frac{2\varepsilon_0\varepsilon(n_2 + \rho_1)\varphi}{e^2n_2\rho_1} \right]^{1/4}. \quad (1.49)$$

При $n_2 \gg \rho_1$

$$\mathcal{L} = \left[\frac{2\varepsilon_0\varepsilon\varphi}{e^2\rho_1} \right]^{1/2}$$

и область пространственного заряда распространяется в p -область полупроводника.

При $\rho_1 \gg n_2$

$$\mathcal{L} = \left(\frac{2\varepsilon_0\varepsilon\varphi}{e^2n_2} \right)^{1/2}$$

и область пространственного заряда распространяется в n -область полупроводника.

Последние два выражения описывают также ширину области пространственного заряда тесного контакта металл — полупроводник для p - и n -полупроводников.

1.4. ВЛИЯНИЕ СИЛ ИЗОБРАЖЕНИЯ НА ХОД ПОТЕНЦИАЛА В ГЕТЕРОПЕРЕХОДЕ

Строго говоря, если электрон находится у границы раздела двух сред, необходимо учитывать его взаимодействие с соседней средой. Это взаимодействие обычно описывается с помощью сил изображения. В этом случае можно определить эффективную потенциальную энергию электрона у краев разрешенных зон как сумму энергии, определяемой силами изображения, и энергии, найденной из решения уравнения Пуассона

$$\varphi_{\Sigma} = \varphi + \varphi^*.$$

Если силы изображения не изменяют ход $\rho(x)$, как, например, в случае слоя Шоттки, для определения φ можно воспользоваться результатами, полученными выше, и нахождение φ_{Σ} сводится к нахождению φ^* .

Нахождение методом последовательных изображений приводит к следующим значениям φ^* :

$$\varphi_1^* = -\frac{e}{4\varepsilon_1} \left\{ \tilde{\varepsilon}_{23} (1 - \tilde{\varepsilon}_{12}^2) \sum_{n=0}^{\infty} \frac{(\tilde{\varepsilon}_{12}\tilde{\varepsilon}_{32})^n}{nd+x} + \frac{\tilde{\varepsilon}_{12}}{x} \right\}; \quad (1.50)$$

$$\varphi_2^* = -\frac{e}{2\varepsilon_{12}} \left\{ \sum_{n=1}^{\infty} \frac{(\tilde{\varepsilon}_{21}\tilde{\varepsilon}_{23})^n}{nd} + \varepsilon_{21} \sum_{n=0}^{\infty} \frac{(\tilde{\varepsilon}_{21}\tilde{\varepsilon}_{23})^n}{2(nd+x)} + \right.$$

$$\left. + \tilde{\varepsilon}_{23} \sum_{n=1}^{\infty} \frac{(\varepsilon_{21}\varepsilon_{23})^{n-1}}{2(nd-x)} \right\}; \quad (1.51)$$

$$\varphi_3^* = -\frac{e}{4\varepsilon_3} \left\{ \tilde{\varepsilon}_{21} (1 - \tilde{\varepsilon}_{32}^2) \sum_{n=0}^{\infty} \frac{\tilde{\varepsilon}_{12}\tilde{\varepsilon}_{32}}{nd+x} + \frac{\tilde{\varepsilon}_{32}}{x-d} \right\}. \quad (1.52)$$

Здесь φ_1^* , φ_2^* , φ_3^* , а также ε_1 , ε_3 и ε_2 относятся к первому полупроводнику, промежуточному слою и второму полупроводнику; x — расстояние от границы раздела первого полупроводника с промежуточным слоем;

$$\tilde{\varepsilon}_{21} = -\tilde{\varepsilon}_{12} = \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_2 + \varepsilon_1}; \quad \tilde{\varepsilon}_{32} = -\tilde{\varepsilon}_{23} = \frac{\varepsilon_3 - \varepsilon_2}{\varepsilon_3 + \varepsilon_2}.$$

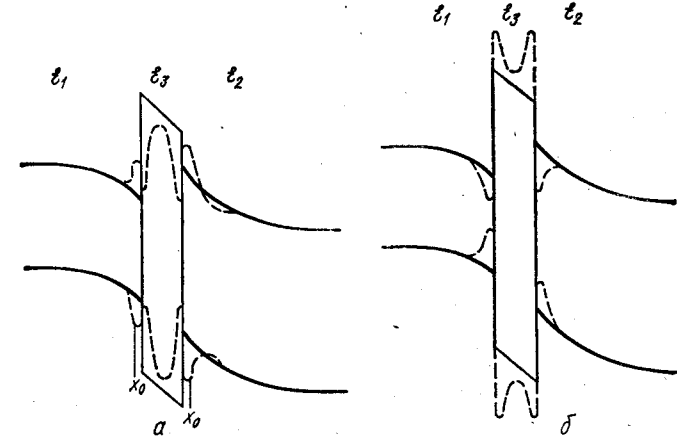


Рис. 1.14. Энергетическая диаграмма гетероперехода с учетом сил изображения: а — $\varepsilon_1 > \varepsilon_3$, $\varepsilon_2 > \varepsilon_3$; б — $\varepsilon_1 < \varepsilon_3$, $\varepsilon_2 < \varepsilon_3$.

Используя выражения (1.50) — (1.52), можно проанализировать частные случаи для $\varphi_{1,2,3}$ у границ раздела сред:

$$\text{при } |x| \ll d \quad \varphi_1^* \simeq \frac{\varepsilon_1 - \varepsilon_2}{(\varepsilon_2 + \varepsilon_1)x};$$

$$\text{при } x \ll d \quad \varphi_2^* \simeq \frac{\varepsilon_2 - \varepsilon_1}{(\varepsilon_2 + \varepsilon_1)x};$$

$$\text{при } x \simeq d \quad \varphi_3^* \simeq \frac{\varepsilon_2 - \varepsilon_1}{(\varepsilon_3 + \varepsilon_2)(d-x)}.$$

Как видно из этих выражений, знак $\varphi_{1,2,3}^*$ изменяется в зависимости от соотношения диэлектрических проницаемостей соседних сред. Если электрон находится в среде с меньшим значением диэлектрической постоянной, то $\varphi_{1,2,3}^* < 0$, а если с большим, то $\varphi_{1,2,3}^* \gg 0$. Это означает, что φ_2 в первом случае будет большим, а во втором меньшим по сравнению с φ , найденной из решения уравнения Пуассона.

Все изложенное иллюстрируется рис. 1.14. На этом рисунке также учтено, что использованный здесь метод изображений не справедлив при $x \rightarrow 0$ и $x \rightarrow d$. Из-за этого выбрано предельное значение x_0 , для которого еще справедлив применяемый метод. Для меньших значений силы изображения считаются постоянными и $\varphi_{1,2,3} = \text{const}$.

Как видно, силы изображения могут существенно изменять ход потенциальной энергии электрона и дырки у границы раздела двух сред. Кроме того, наличие сил изображения может приводить к следующим двум эффектам: во-первых, изменяется ширина запрещенной зоны у границы раздела двух сред (сужается в среде с меньшим значением диэлектрической постоянной и расширяется в среде с большим значением), во-вторых, на границе раздела двух сред могут возникать потенциальные ямы для носителей заряда.

Отметим, что несмотря на то, что силы изображения существенно изменяют ход потенциальной энергии, в принятых приближениях они не изменяют соотношения между потенциальными барьерами в разных частях гетероперехода. В самом деле, если $\varphi_{1,2,3} = \text{const}$ при $x \rightarrow 0$ или $x \rightarrow d$, то $\frac{d\varphi^*}{dx} = 0$ на границах раздела фаз (в малой окрестности вблизи $x = 0$ и $x = d$) и при использовании условия непрерывности вектора электрической индукции, куда входит напряженность поля у границы раздела, силы изображения не играют никакой роли.

1.5. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ЗОННОЙ МОДЕЛИ ГЕТЕРО- И ГОМОПЕРЕХОДОВ

Экспериментальному изучению зонной модели гетеро- и гомопереходов посвящено большое число исследований. Они обычно основаны на анализе вольт-амперных, вольт-емкостных зависимостей и зависимости фотоответа от длины волны падающего света.

В результате подробных исследований контактов металл — полупроводник установлено, что высота потенциального барьера в полупроводнике зависит от удельного сопротивления и типа проводимости полупроводника, обработки поверхности полупроводника, материала металлического электрода и метода его нанесения.

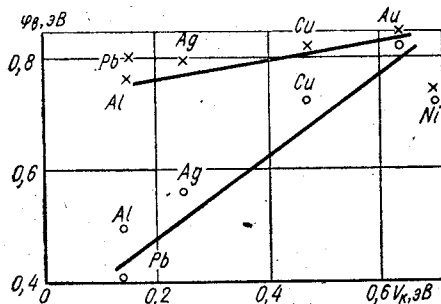


рис. 1.15. Зависимость величины $\varphi_b = \varphi_0 + \mu$ от контактной разности потенциалов: \circ — химически травленая поверхность; \times — скол в высоком вакууме.

В качестве примера на рис. 1.15 представлена экспериментальная зависимость величины $\varphi_b = \varphi_0 + \mu$ от контактной разности потен-

циалов V_k для напыленных контактов различных металлов с кремнием n -типа. Как видно из этого рисунка, для напыленных контактов на сколотую поверхность в вакууме 10^{-8} мм рт. ст. φ_b больше, чем для напыленных на химически травленную поверхность. Зависимость φ_b от V_k для металлов, напыленных на сколотую в высоком вакууме поверхность, слабее, чем для напыленных на химически травленную поверхность.

Анализ данных, приведенных на рисунке, показывает, что экспериментальные значения φ_b значительно слабее зависят от V_k , чем для случая, когда зазора и поверхностных состояний нет. В самом деле, используя результаты п. 1.3, можно показать, что

$$\frac{d\Delta\varphi}{d\varphi} = \frac{d\Delta\varphi}{d\varphi_b} \frac{d\varphi_b}{d\varphi} = \frac{d}{\varepsilon_0 \varepsilon_3} \left[\frac{\varepsilon_0 \varepsilon_2}{\mathcal{L}} + \frac{d\sigma_2}{d\varphi_b} \right] \frac{d\varphi_b}{d\varphi}. \quad (1.53)$$

Подставляя это выражение в $\frac{d\varphi_b}{d\varphi} = 1 - \frac{d\Delta\varphi}{d\varphi}$, получаем

$$\frac{d\varphi_b}{d\varphi} = \frac{1}{1 + \frac{d}{\varepsilon_0 \varepsilon_3} \left(\frac{\varepsilon_0 \varepsilon_2}{\mathcal{L}} + \frac{d\sigma_2}{d\varphi_b} \right)}. \quad (1.54)$$

Как видно из последнего выражения, только при $d \neq 0$ $\frac{d\varphi_b}{d\varphi} < 1$. Кроме того, оценки показывают, что для концентрации примесей в полупроводнике $n_2 < 10^{16}$ см $^{-3}$ при $d \approx 10^{-7}$ см, член $\frac{\varepsilon_2 d}{\varepsilon_3 \mathcal{L}} \ll 1$. А это означает, что если для полупроводников с такими параметрами $\frac{d\varphi_b}{d\varphi}$ существенно меньше единицы, то значительную роль должны играть поверхностные состояния.

Если аппроксимировать экспериментальные зависимости φ_b от V_k прямой с $\frac{d\varphi_b}{d\varphi} = \text{const}$, то это означает, что $\frac{d\sigma_2}{d\varphi_b}$ также постоянно. Решая уравнение (1.54) относительно $\frac{d\sigma_2}{d\varphi_b}$, можно получить

$$N_i = \frac{d\sigma_2}{ed\varphi_b} \approx \frac{\left(1 - \frac{d\varphi_b}{d\varphi}\right) \varepsilon_0 \varepsilon_3}{ed \frac{d\varphi_b}{d\varphi}}, \quad (1.55)$$

где N_i — постоянная плотность поверхностных состояний.

Как видно из последнего выражения, для контакта с меньшим зазором (сколотая поверхность) роль поверхностных состояний, стабилизирующих высоту барьера при изменении V_k , больше, чем для химически травленной поверхности, что свидетельствует о большей плотности поверхностных состояний в этом случае. Сплошные кривые, совпадающие с экспериментальными данными на рис. 1.15, построены с использованием теоретического соотношения (1.55) при предположении, что в контактах существует зазор с $d = 8 \text{ \AA}$ и $d = 20 \text{ \AA}$ и поверхностные состояния имеют плотность $N_i = 4 \times$

$\times 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$ и $N_i = 4 \cdot 10^{12} \text{ см}^{-2} \cdot \text{эВ}^{-1}$ для скелотой и химически травленной поверхности соответственно.

Аналогичные результаты получены и для контактов металлов с другими полупроводниками, причем для контактов, изготовленных в чистых условиях, когда зазор между металлом и полупроводником минимален, наблюдалась зависимость наклона $S = \frac{d\varphi_b}{dV_K}$ от материала полупроводника. Это показано на рис. 1.16, где дана зависимость s от разности электроотрицательностей элементов кон-

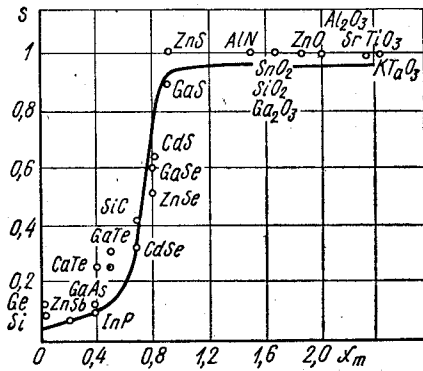


Рис. 1.16. Зависимость $s = \frac{d\varphi_b}{dX_m}$ от разности электроотрицательностей элементов X_m .

такта. Как видно из рисунка, экспериментально наблюдается четко выраженный переход от ионных полупроводников к ковалентным. Для ионных полупроводников $s \approx 1$, что соответствует малой плотности поверхностных состояний, а для ковалентных $s < 1$, что соответствует большой плотности поверхностных состояний. Причем для полупроводников с ковалентной связью максимальная плотность поверхностных состояний сосредоточена примерно на расстоянии $\frac{1}{3} E_g$ от верха валентной зоны, о чем свидетельствует зависимость φ_b от ширины запрещенной зоны E_g для контактов золота — полупроводник, представленная на рис. 1.17.

Из приведенных экспериментальных данных видно, что в общем случае для контакта металл — полупроводник нельзя пренебрегать зазором и поверхностными электронными состояниями.

Хотя энергетическая зонная модель гетеропереходов исследовалась не меньше, чем энергетическая модель контакта металл — полупроводник, полученные данные здесь не являются законченными. Основные результаты, полученные при исследовании гетеропереходов, сводятся к следующему. Если постоянные решетки контактирующих полупроводников согласованы недостаточно хорошо, то на границе раздела возникают нарушенные связи, которые являются причиной появления поверхностных электронных состояний.

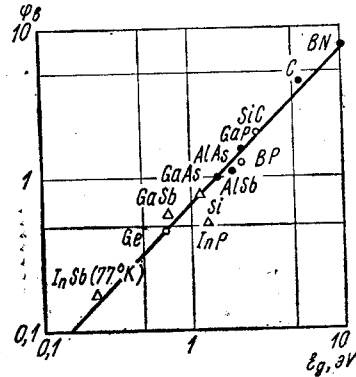


Рис. 1.17. Зависимость высоты потенциального барьера от ширины запрещенной зоны полупроводника: \circ — n -тип; \bullet — p -тип, Δ — оба типа проводимости.

Вследствие этого возникает изгиб зон. Если рассогласование решетки достигает нескольких процентов, то плотность таких состояний порядка 10^{14} см^{-2} . Если постоянные решетки отличаются менее чем на 1%, то плотность поверхностных состояний составляет 10^{13} см^{-2} и меньше, а их влияние на свойства гетероперехода будет более слабым.

В качестве примера гетеропереходов, для которых играют роль поверхностные состояния, могут быть взяты изотипные $n-n$ гетеропереходы Si—Ge. На рис. 1.18 для этих переходов представлена энергетическая диаграмма при разных концентрациях носителей заряда в Si и Ge. Как видно из рисунка, одновременно существует обедненный слой как в Si, так и в Ge, что определяется наличием поверхностных состояний.

В качестве гетероперехода с хорошо согласованными решетками можно привести систему GaAs — $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Все дело в том, что постоянная решетки AlAs, равная 5,66 Å, почти совпадает с постоянной решетки GaAs, равной 5,654 Å. В то же время ширина запрещенной зоны AlAs составляет 2,1 эВ, что больше, чем ширина запрещенной зоны GaAs, составляющей 1,41 эВ. Диаграмма энергетических зон гетеропереходов $\text{Al}_x\text{Ga}_{1-x}\text{As} - \text{GaAs}$ при разных уровнях легирования полупроводников, составляющих гетеропереход, показана на рис. 1.19. При этом получено: $\Delta E_c \sim 0,4 \text{ эВ}$; $\Delta E_v \approx 0$.

Большое количество исследований было посвящено изучению $p-n$ гомопереходов. В результате экспериментально установлено, что высота потенциального барьера зависит от концентрации примесей в p - и n -областях полупроводника, как это следует из приведенной выше теории без учета поверхностных состояний на границе раздела p - и n -областей. Однако отдельные результаты не удается объяснить, если не учитывать такие состояния.

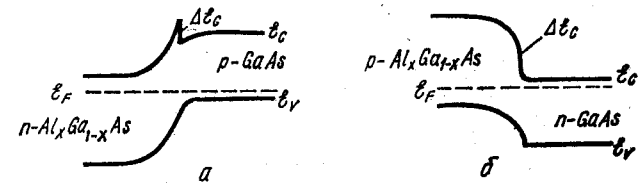


Рис. 1.19. Энергетическая диаграмма гетероперехода $\text{Al}_x\text{Ga}_{1-x}\text{As} - \text{GaAs}$ с согласованными постоянными решетками: a — переход $n\text{-Al}_x\text{Ga}_{1-x}\text{As} - p\text{-GaAs}$; b — переход $p\text{-Al}_x\text{Ga}_{1-x}\text{As} - n\text{-GaAs}$.

Глава 2. ОБЩИЙ АНАЛИЗ МОДЕЛИ ГЕТЕРО- И ГОМОПЕРЕХОДОВ ПРИ ПРИЛОЖЕНИИ НАПРЯЖЕНИЯ

2.1. ПОВЕДЕНИЕ ГЕТЕРО- И ГОМОПЕРЕХОДОВ ПРИ ПРИЛОЖЕНИИ НАПРЯЖЕНИЯ

Работа гетеро- и гомопереходов обычно связана с приложением к нему внешнего напряжения или возникновением напряжения вследствие внешних воздействий.

Гетеропереход, к которому приложено напряжение, показан на рис. 2.1, а. На рис. 2.1, б представлен случай, когда напряжение приводит к понижению высоты потенциальных барьеров как в правом, так и левом полупроводниках, а на рис. 2.1, в — случай, когда приложенное напряжение повышает высоту потенциального барьера в обоих полупроводниках.

На рис. 2.1 для определенности показан анизотипный гетеропереход, когда полупроводники 1 и 2 имеют разный тип проводимости (в данном случае $p-n$ переход). Однако картина не изменится, если взять изотипный гетеропереход, для которого тип проводимости полупроводников 1 и 2 одинаков (например, $n-n$ или $p-p$ переходы).

Изменение расстояний между уровнями Ферми в различных частях гетероперехода при пренебрежении падением напряжения в объеме полупроводников равно eV . Оно распределяется между различными частями контакта так, что

$$V = V_1 + V_3 + V_2,$$

где V_1 , V_3 и V_2 — части общего напряжения в первом полупроводнике, зазоре (промежуточном слое) и втором полупроводнике.

Приложение внешнего напряжения приводит к появлению в цепи тока, связанного с переносом в гетеропереходе носителей заряда: Этот ток обусловлен изменением условий для прохождения носителей заряда слева направо и справа налево. Например, для случая, представленного на рис. 2.1, б, после приложения напряжения электроны при прохождении из полупроводника 2 в полупроводник 1 преодолевают меньший потенциальный барьер, чем при отсутствии напряжения, а при прохождении из полупроводника 1 в 2 — больший. Это приводит к тому, что в цепи возникает электронный ток, направленный от полупроводника 2 к полупроводнику 1. Для того же случая дырки при приложенном напряжении легче переходят от полупроводника 1 к полупроводнику 2, чем при отсутствии напряжения. Это приводит к возникновению дырочного тока от полупроводника 1 к полупроводнику 2, который суммируется с электронным током.

Для случая, изображенного на рис. 2.1, в, ситуация обратная. Электронам при приложенном напряжении легче переходить от полупроводника 1 к полупроводнику 2, а дыркам труднее. Возникают электронный и дырочные токи обратного направления.

Все это условно показано на рисунке изменением длины стрелок, соответствующих токам электронов и дырок из полупроводника 1 и 2, при приложении напряжения.

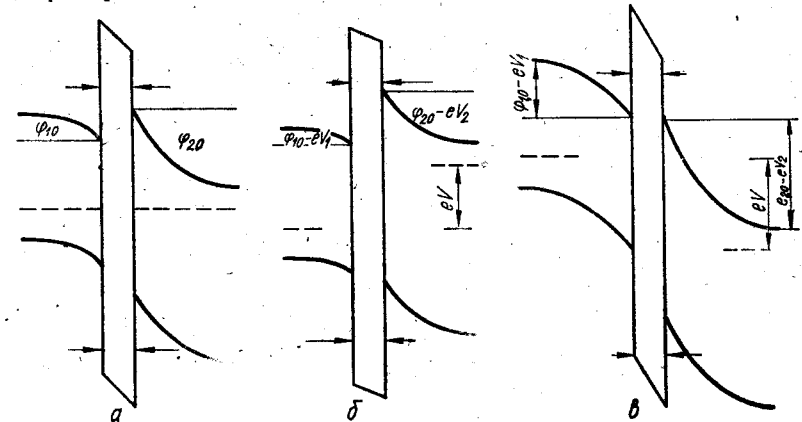


Рис. 2.1. Энергетическая диаграмма гетероперехода при разных приложенных напряжениях: а — $V = 0$; б — $V > 0$; в — $V < 0$.

Кроме того, важным является то, рассматриваются ли токи в анизотипном или изотипном гетеропереходе, а также то, идет ли перенос носителей заряда без изменения или с изменением вида разрешенных зон, т. е. идут ли носители заряда все время в зоне проводимости или валентной зоне или в процессе перехода от одного полупроводника к другому происходит переход носителей заряда из зоны проводимости в валентную зону и наоборот.

Если гетеропереход является изотипным (например, $n-n$ или $p-p$ переход) и межзонных переходов между зоной проводимости и валентной зоной нет, то движущиеся носители заряда везде являются основными. Если же гетеропереход анизотипный, т. е. $p-n$ или $n-p$, а межзонных переходов нет, то носители заряда, переходя из одного полупроводника в другой, становятся неосновными носителями заряда (инжекция). В случае $p-n$ гетероперехода электроны, которые переходят из полупроводника 2 в полупроводник 1, и дырки, которые переходят из полупроводника 2 в полупроводник 1, являются неосновными носителями заряда.

Если же в процессе перехода из одного полупроводника в другой происходит изменение вида разрешенной зоны, то неосновные носители заряда появляются только в случае, когда гетеропереход изотипный. Для анизотипного $p-n$ или $n-p$ гетероперехода при аналогичных переходах неосновные носители не возникают.

Все изложенное выше является важным для дальнейшего рассмотрения выпрямляющих свойств гетероперехода. Действительно,

если носители заряда во всех частях гетероперехода являются основными, то невозможно введение избыточной концентрации носителей в какой-либо части гетероперехода, так как на каждый введенный из одного электрода носитель заряда приходится носитель того же знака, выходящий из другого электрода. Если же носители заряда при переходе из одной части гетероперехода в другую становятся неосновными, то их заряд может быть компенсирован введением такого же количества основных носителей заряда. В этом случае возникает градиент концентрации введенных носителей заряда, так как их концентрация максимальна вблизи места их введения в полупроводник, т. е. вблизи собственно области контакта, что приводит к их диффузии от рассматриваемого контакта к омическим контактам.

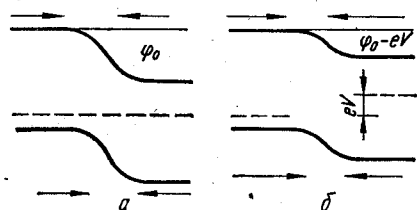


Рис. 2.2. Энергетическая диаграмма $p-n$ перехода.

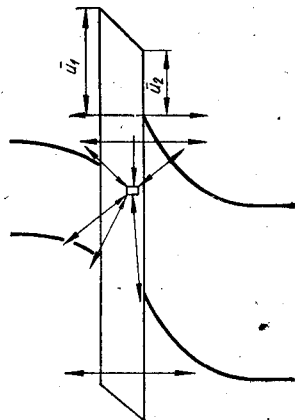


Рис. 2.3. Энергетическая диаграмма и токи в гетеропереходе.

Рассмотренная выше качественная картина прохождения носителей заряда через гетеропереход легко может быть распространена на контакт металл — полупроводник и гомо- $p-n$ переход.

Например, при отсутствии напряжения, приложенного к $p-n$ переходу, потоки электронов и дырок из n -типа полупроводника в p -тип и обратные потоки равны друг другу (рис. 2.2, а). При приложении напряжения эта ситуация изменяется. Когда напряжение приводит к понижению высоты барьера между n - и p -областями полупроводника (рис. 2.2, б), ток дырок из n -области в p -область и электронов из p -области в n -область остается неизменным, поскольку они не преодолевают никакого потенциального барьера. Ток же электронов из n -типа полупроводника в p -тип и дырок из p -типа полупроводника в n -область увеличивается, так как этим носителям заряда теперь нужно преодолевать меньший барьер по сравнению со случаем, когда напряжение равно нулю. В результате через $p-n$ переход пойдет ток.

Если бы напряжение повышало высоту потенциального барьера, то токи электронов и дырок, соответствующие переходам в направлении отсутствия барьера, не изменились бы, а в направлении повышения барьера уменьшились бы. Это привело бы к появлению тока, противоположного по направлению рассмотренному выше.

При приложении прямого напряжения, как следует из рассмотренного выше, изменяется и концентрация носителей заряда в n - и p -типе. Так, при понижении высоты потенциального барьера в n -области у $p-n$ перехода оказывается избыточная концентрация дырок Δp , а в p -области — избыточная концентрация электронов Δn и соответствующие им избыточные концентрации электронов в n -области и дырок в p -области, которые компенсируют заряд введенных неосновных носителей Δp и Δn .

С точки зрения общности рассмотрения прохождения носителей заряда в гетеропереходе целесообразно выделить несколько последовательных областей. Это квазинейтральные области 1 и 2 полупроводников, области пространственного заряда 1 и 2 полупроводников и область зазора (промежуточного слоя).

Прохождение носителей заряда через квазинейтральную область не отличается от обычной ситуации в объеме полупроводника. В областях пространственного заряда и зазоре возможны различные случаи.

В принципе возможны следующие пути прохождения носителей заряда через область пространственного заряда: прохождение носителя заряда в разрешенной зоне и туннелирование носителя заряда через область запрещенной зоны, прохождение носителя заряда с изменением вида разрешенной зоны.

При прохождении носителей заряда через область зазора возможно прямое туннелирование носителя заряда и прохождение носителей заряда с захватом на поверхностные уровни и последующим выбросом в одну из разрешенных зон. Все эти возможности проиллюстрированы на рис. 2.3.

При приложении напряжения носитель заряда в гетеропереходе движется через последовательно расположенные квазинейтральные области полупроводников, области пространственных зарядов и область зазора. Условия движения носителей в каждой из этих областей различны. Поэтому сквозное движение через гетеропереход при последовательном движении через различные области определяется преимущественно той областью, через которую движение носителей заряда наиболее затруднено. В этом случае можно говорить о том, что ток через гетеропереход контролируется той или иной выделенной областью.

С рассматриваемой точки зрения простыми системами являются контакт металл — полупроводник и гомо- $p-n$ переход, а более сложной системой является гетеропереход. В самом деле, из общего рассмотрения в случае контакта металл — полупроводник исключается область металла, а в $p-n$ гомопереходе — область зазора.

Дальнейший анализ рассмотренной здесь общей картины будет проводиться ниже.

2.2. ПРОХОЖДЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА ЧЕРЕЗ КВАЗИНЕЙТРАЛЬНУЮ ОБЛАСТЬ

В общем случае для описания движения носителей заряда в полупроводнике необходимо рассматривать совместно с записанным ранее уравнением Пуассона еще четыре уравнения:

$$\left. \begin{aligned} i_n &= eu_n n E + e D_n \text{grad } n; \\ i_p &= eu_p p E - e D_p \text{grad } p; \end{aligned} \right\} \quad (2.1)$$

$$\left. \begin{aligned} \frac{\partial n}{\partial t} &= g - \frac{n - n_0}{\tau_n} + \frac{1}{e} \text{div } i_n; \\ \frac{\partial p}{\partial t} &= g - \frac{p - p_0}{\tau_p} - \frac{1}{e} \text{div } i_p, \end{aligned} \right\} \quad (2.2)$$

где $u_n, u_p; D_n, D_p$ — подвижности и коэффициенты диффузии; n и p — концентрации; τ_n, τ_p — времена жизни электронов и дырок соответственно; g — концентрация носителей заряда, возникающих в единицу времени под действием внешних факторов.

Первые два уравнения — это диффузионно-дрейфовые уравнения для плотности тока электронов i_n и дырок i_p . Они отражают движение носителей заряда под действием поля или градиента концентраций в пределах одной из разрешенных зон.

Вторые два уравнения — это уравнения неразрывности, которые отражают изменение концентрации носителей заряда в разрешенных зонах.

В одномерном случае, когда источника внешних возбуждений нет и $g = 0$, эта система переписывается в следующем виде:

$$i_n = eu_n n E + e D_n \frac{dn}{dx}; \quad (2.3a)$$

$$i_p = eu_p p E - e D_p \frac{dp}{dx}; \quad (2.3b)$$

$$\frac{\partial n}{\partial t} = - \frac{n - n_0}{\tau_n} + \frac{1}{e} \cdot \frac{di_n}{dx}; \quad (2.4a)$$

$$\frac{\partial p}{\partial t} = - \frac{p - p_0}{\tau_p} - \frac{1}{e} \cdot \frac{di_p}{dx}. \quad (2.4b)$$

Система уравнений в общем случае не решается. Поэтому в дальнейшем рассмотрим наиболее интересные частные случаи, когда эта система имеет решение.

Прежде всего проанализируем прохождение носителей заряда через изотипный гетеропереход или $p-n$ переход с изменением вида разрешенной зоны, когда носители заряда остаются везде основными.

В этом случае не происходит изменения концентрации носителей заряда в разрешенных зонах за счет проходящего тока, следовательно:

$$p - p_0 = n - n_0 = 0 \quad \text{и} \quad \frac{di_n}{dx} = \frac{di_p}{dx} = 0.$$

Решение системы из четырех уравнений сводится к решению системы из двух уравнений (2.3a), (2.3b). При $\frac{dn}{dx} = \frac{dp}{dx} = 0$ имеем обычную связь плотности электронного и дырочного тока с проводимостью

$$i_n = eu_n n = \sigma_n E; \quad i_p = eu_p p = \sigma_p E.$$

Это означает, что в квазинейтральной области каждого полупроводника падает определенная часть общего напряжения.

$$V_b = I r, \quad (2.5)$$

где $r = \rho \frac{l}{S}$; I — общий ток.

Действительно, если домножить любое из выражений для i_p или i_n на площадь S и учесть, что $E = \frac{V_b}{l}$, где l — ширина квазинейтральной области, то легко получить записанное выше выражение.

При этом обычно V_b является малой частью приложенного общего напряжения и, следовательно, квазинейтральная область не контролирует общий ток через гетеропереход.

Если в процессе переноса тока в гетеро- или гомопереходе носители становятся неосновными, то $p - p_0$ и $n - n_0$ не равны нулю и следует рассматривать совместно систему уравнений (2.3a) — (2.4b).

Для определенности рассмотрим решение задачи, когда полупроводник является электронным с концентрацией электронов n_0 , а в квазинейтральную область вводятся дырки.

Пусть вначале уровень инжекции мал и $\Delta p \ll n_0$. Тогда в выражении (2.3b) для i_p можно пренебречь первым членом и

$$i_p = -e D_p \frac{dp}{dx}. \quad (2.6)$$

Подставляя это выражение в уравнение

$$- \frac{p - p_0}{\tau_p} - \frac{1}{e} \cdot \frac{di_p}{dx} = 0,$$

полученное из (2.4b) при $\frac{\partial p}{\partial t} = 0$, имеем

$$D_p \frac{d^2 p}{dx^2} = \frac{p - p_0}{\tau_p}. \quad (2.7)$$

Последнее выражение легко переписать в виде

$$\frac{d^2 p}{dx^2} = \frac{p - p_0}{l_p^2}, \quad (2.8)$$

где $l_p = \sqrt{D_p \tau_p}$.

Общее решение выражения (2.8)

$$p - p_0 = A e^{-\frac{x}{l_p}} + B e^{\frac{x}{l_p}}. \quad (2.9)$$

Постоянные A и B можно найти, если задать граничные условия. В случае широкой квазинейтральной области, когда $p - p_0 = \Delta p (\mathcal{L}_n)$ на границе области пространственного заряда при $x = \mathcal{L}_n$ и $p - p_0 = 0$ при $x = \infty$, получаем из второго граничного

условия $B = 0$, а из первого $A = \Delta p (\mathcal{L}_n) e^{\frac{\mathcal{L}_n}{l_p}}$. В результате

$$p - p_0 = \Delta p (\mathcal{L}_n) e^{-\frac{\mathcal{L}_n - x}{l_p}}. \quad (2.10)$$

В случае узкой квазинейтральной области, когда $p - p_0 = \Delta p (\mathcal{L}_n)$ при $x = \mathcal{L}_n$, а $p - p_0 = \Delta p (w_n)$ при $x = w_n$, для нахождения постоянных A и B следует решить два уравнения:

$$\Delta p (\mathcal{L}_n) = A e^{-\frac{\mathcal{L}_n}{l_p}} + B e^{\frac{\mathcal{L}_n}{l_p}}; \quad (2.11)$$

$$\Delta p (w_n) = A e^{-\frac{w_n}{l_p}} + B e^{\frac{w_n}{l_p}}.$$

Их совместное решение дает:

$$A = \frac{\Delta p (\mathcal{L}_n) e^{\frac{w_n}{l_p}} - \Delta p (w_n) e^{-\frac{\mathcal{L}_n}{l_p}}}{2 \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}};$$

$$B = \frac{\Delta p (w_n) e^{-\frac{\mathcal{L}_n}{l_p}} - \Delta p (\mathcal{L}_n) e^{-\frac{w_n}{l_p}}}{2 \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}}. \quad (2.12)$$

Окончательное выражение для $p - p_0$ можно записать в виде:

$$p - p_0 = \frac{\Delta p (\mathcal{L}_n) \operatorname{sh} \frac{w_n - x}{l_p}}{\operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}} + \frac{\Delta p (w_n) \operatorname{sh} \frac{x - \mathcal{L}_n}{l_p}}{\operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}}. \quad (2.13)$$

Теперь предположим, что уровень инжекции велик и $\Delta p \gg n_0$ во всем рассматриваемом пространстве. Тогда в уравнениях (2.3) нельзя пренебречь полевыми членами. Если электронный ток значительно меньше дырочного, из уравнения (2.3а) при $i_n = 0$ имеем

$$E = -\frac{kT}{e} \cdot \frac{1}{n} \cdot \frac{dn}{dx}. \quad (2.14)$$

Подставляя это значение поля в уравнение (2.3б), получаем

$$i_p = -eD_p \frac{p}{n} \cdot \frac{dn}{dx} - eD_p \frac{dp}{dx} = -2eD_p \frac{dp}{dx}, \quad (2.15)$$

поскольку при больших уровнях инжекции $p \simeq n$ и $\frac{dp}{dx} \simeq \frac{dn}{dx}$.

Уравнение неразрывности для дырок при больших уровнях инжекций для $\frac{\partial p}{\partial t} = 0$ приобретает вид

$$-\frac{p}{\tau_{\infty}} = \frac{1}{e} \cdot \frac{di_p}{dx}, \quad (2.16)$$

где учтено, что $\tau_p = \tau_{\infty}$ при $p \gg n_0$. После подстановки i_p (2.15) в это уравнение имеем

$$\frac{d^2 p}{dx^2} = \frac{p}{l_p^2}, \quad (2.17)$$

где $l_{p_0} = \sqrt{2D_p \tau_{\infty}}$.

Полученное уравнение для p является аналогичным уравнению (2.8), решение которого приводилось выше. Поэтому в случае большого уровня инжекции можно воспользоваться всеми результатами,

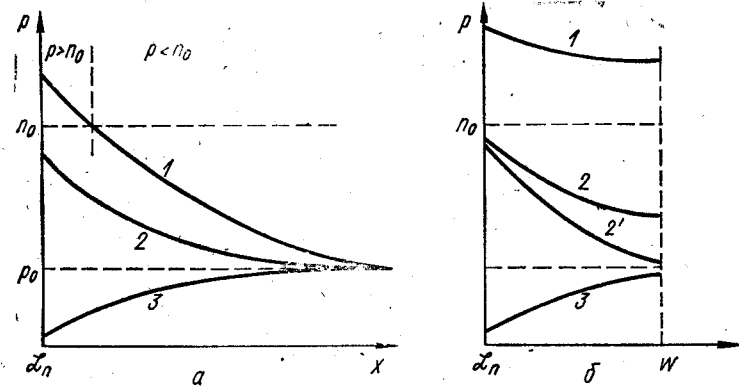


Рис. 2.4. Зависимость концентрации дырок p от расстояния для широкой (а) и узкой (б) квазинейтральной области:

1 — большой уровень инжекции; 2 и 2' — малый уровень инжекции (для 2 скорость рекомбинации меньше, чем для 2'); 3 — экстракция носителей.

полученными ранее при замене l_p на l_{p_0} . Однако необходимо помнить, что большой уровень инжекции во всем интервале x рассматриваемого пространства может быть получен только при $\Delta p (\mathcal{L}_n) \gg 0$ и конечной ширине квазинейтральной области. Это можно видеть из рис. 2.4, где показана зависимость $p - p_0$ от x , построенная по формулам (2.10) и (2.13) при разных значениях $p (\mathcal{L}_n)$.

Используя полученные значения $p - p_0$, можно найти i_p , если их подставить в формулы (2.6) и (2.15).

В результате имеем для широкой квазинейтральной области:

$$i_p = \frac{eD_p \Delta p (\mathcal{L}_n)}{l_p} e^{-\frac{\mathcal{L}_n - x}{l_p}}. \quad (2.18)$$

Для узкой квазинейтральной области получаем

$$i_p = \frac{eD_p}{l_p \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}} \left[\Delta p (\mathcal{L}_n) \operatorname{ch} \frac{w_n - x}{l_p} - \Delta p (w_n) \operatorname{ch} \frac{x - \mathcal{L}_n}{l_p} \right], \quad (2.19)$$

где для малого уровня инжекции следует брать l_p из выражения (2.8), а для большого l_{p_0} из выражения (2.17) с учетом (2.19), умноженного на 2.

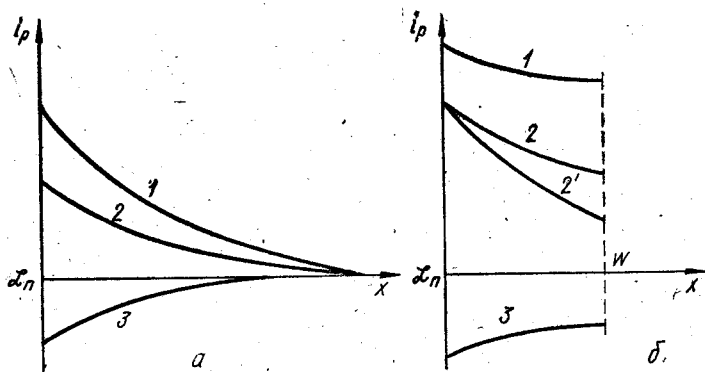


Рис. 2.5. Зависимость i_p от x для широкой (а) и узкой (б) квазинейтральной области: 1 — большой уровень инжекции; 2 и 2' — малый уровень инжекции ($V > 0$); 3 — экстракция ($V < 0$).

Как видно из последних выражений, дырочный ток максимален при $x = \mathcal{L}_n$ и убывает с ростом x . Уменьшение тока происходит с разной скоростью для широкой и узкой квазинейтральной области, как это показано на рис. 2.5. Причем для узкой квазинейтральной области можно изменять i_p в широких пределах путем изменения Δp (w_n).

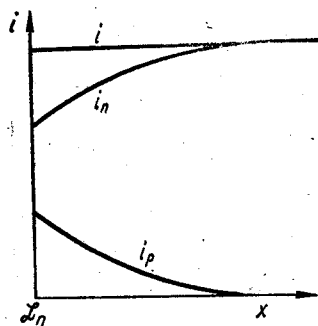


Рис. 2.6. Зависимость i , i_n и i_p от расстояния x для широкой квазинейтральной области.

Поскольку общий ток $i = i_n + i_p$ должен быть постоянным, то уменьшение i_p приводит к увеличению i_n , как это показано для одного из рассмотренных случаев на рис. 2.6.

Зависимость тока в квазинейтральной области от напряжения может быть получена при нахождении величины Δp (\mathcal{L}_n) из условий шивания решений в квазинейтральной области и области пространственного заряда, что будет сделано позже при рассмотрении конкретных вольт-амперных характеристик гетеро- и гомопереходов.

Как видно из приведенного выше анализа, движение неосновных носителей заряда через квазинейтральную область затруднено и последняя, в принципе, может контролировать ток неосновных носителей через гетеропереход.

И в заключение рассмотрим условия, при которых можно пренебречь полевыми членами в выражениях (2.3). Для этого подставим в (2.3б) решение (2.10). Тогда при $\Delta p(\mathcal{L}_n) \gg p_0$

$$i_p = eu_n \Delta p(\mathcal{L}_n) e^{-\frac{\mathcal{L}_n - x}{l_p}} E + \frac{eD_p \Delta p(\mathcal{L}_n)}{l_p} e^{-\frac{\mathcal{L}_n - x}{l_p}}.$$

Как видно из этого уравнения, первым членом можно пренебречь, если

$$u_p E \ll \frac{D_p}{l_p},$$

или используя соотношение Эйнштейна $\frac{u_p}{D_p} = \frac{e}{kT}$, если

$$el_p E \ll kT. \quad (2.20)$$

Полученное неравенство означает, что полевым членом в уравнениях (2.3б) можно пренебречь, если на диффузионной длине дырки приобретают энергию меньшую, чем их средняя тепловая энергия. При невыполнении этого неравенства первым членом в диффузионно-дрейфовом уравнении для дырок пренебрегать нельзя.

Аналогичные результаты могут быть получены и для электронов, вводимых в p -тип полупроводника. При этом в соответствующих формулах надо заменить величины и индексы со значком « p » на величины и индексы со значком « n ».

2.3. ПРОХОЖДЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА В ОБЛАСТЯХ ПРОСТРАНСТВЕННОГО ЗАРЯДА

В принципе, при отсутствии туннелирования система диффузионно-дрейфовых уравнений (2.1) должна быть справедливой и для области пространственного заряда, являющейся частью полупроводника. Однако, если в области пространственного заряда существует большое поле или она является достаточно тонкой, то это может оказаться и не так. В самом деле, система уравнений (2.1) и (2.3) остается справедливой, если в области пространственного заряда можно ввести понятие подвижности. Это означает, что дрейфовая скорость $v_{др}$ должна быть значительно меньше средней тепловой скорости $\bar{v}_{n,p}$, т. е. должно выполняться неравенство

$$\bar{v}_{др} = Eu_{n,p} \ll \bar{v}_{n,p}.$$

Домножая левую и правую части этого неравенства на $\bar{v}_{n,p}$ и учи-

тывая, что $\bar{v}_{n,p} = \left(\frac{8}{\pi} \cdot \frac{kT}{m_{n,p}^*}\right)^{1/2}$, а $u_{n,p} = \frac{e\tau_n}{m_{n,p}} = \frac{el_{n,p}}{m_{n,p}^* v_{n,p}}$, имеем

$$eEl_{n,p} \ll \frac{8}{\pi} kT. \quad (2.21)$$

Это означает, что уравнения (2.1) и (2.3) перестают быть справедливыми при полях, при которых на длине свободного пробега $l_{n,p}$ носитель заряда набирает энергию большую, чем средняя тепловая энергия. Кроме того, если при прохождении носителей заряда через область пространственного заряда существенно туннелирование, необходимо учитывать, что часть области пространственного заряда будет исключаться из области, где рассматривается движение

носителей заряда в разрешенных зонах полупроводника (см. рис. 2.3). Для характеристики туннелирования можно использовать коэффициент прозрачности, равный в ВКБ-приближении (приближении Венцеля — Крамера — Бриллюэна):

$$\mathcal{P}_n = \mathcal{P}_0 e^{-\frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m^*[\varphi(x) - \mathcal{E}]} dx} \quad (2.22)$$

где x_1 и x_2 — точки, в которых частица с энергией \mathcal{E} входит и выходит в барьер области пространственного заряда. Отметим, что ВКБ-приближение справедливо только тогда, когда выполняется соотношение $\frac{d\lambda}{dx} \ll 1$, где $\lambda = \frac{\hbar}{p(x)}$, а $p(x)$ — импульс частицы.

Для случая, когда область пространственного заряда является слоем Шоттки, $\varphi(x)$ задается выражением (1.16), а коэффициент прозрачности принимает вид

$$\mathcal{P}_n = \mathcal{P}_0 \exp \left\{ - \left[\frac{2m^* \varphi_0}{\hbar^2} \right]^{1/2} \times \right. \\ \left. \times \mathcal{L} \left(\sqrt{1 - x_E} - x_E \operatorname{arch} \frac{1}{\sqrt{x_E}} \right) \right\}, \quad (2.23)$$

Рис. 2.7. Движение носителей заряда в области пространственного заряда полупроводника: 1, 4 — диодная теория; 2, 3 — диффузионная теория.

где $x_E = \frac{\mathcal{E}}{\varphi_0}$, а \mathcal{E} отсчитывается от уровня $\varphi(\mathcal{L})$.

Как видно из последнего выражения, коэффициент прозрачности области пространственного заряда будет большим при $\mathcal{E} = \text{const}$ для сильно легированных полупроводников и больших высот потенциального барьера φ_0 . Для слабо легированных полупроводников туннелированием через область пространственного заряда обычно можно пренебречь.

В зависимости от условий движения носителей заряда в разрешенных зонах области пространственного заряда полупроводника нахождение вольт-амперных характеристик должно проводиться различным способом. Если носители заряда проходят область пространственного заряда без столкновений, то для нахождения вольт-амперной характеристики необходимо рассматривать поток носителей заряда из квазинейтральной области или в квазинейтральную область полупроводника над барьером области пространственного заряда. Для n -типа полупроводника это показано на рис. 2.7, где траектории движения электронов и дырок показаны горизонтальными стрелками 1 и 4 (отметим, что для 1 не соблюден масштаб вне ОПЗ).

Теория выпрямления, которая не учитывает столкновений носителей заряда с решеткой полупроводника в области пространственного заряда, называется диодной.

Если при прохождении носителя заряда через область пространственного заряда полупроводника он испытывает многократные столкновения, то для нахождения вольт-амперной характеристики необходимо детально анализировать условия прохождения носителя заряда в этой области. Соответствующие этому случаю траектории электронов также показаны схематически на рис. 2.7 ломаными линиями 2 и 3.

Теория выпрямления, в которой рассматривается взаимодействие электронов и дырок с полупроводником в области пространственного заряда, называется диффузионной. В диффузионной теории для области пространственного заряда в общем случае решаются диффузионно-дрейфовые уравнения (2.1) или в одномерном случае уравнения (2.3).

Когда область пространственного заряда полупроводника является слоем Шоттки, условия выполнения диодной или диффузионной теории могут быть определены из неравенств между длиной свободного пробега $l_{n,p}$ и шириной слоя Шоттки:

$$l_{n,p} \gg \mathcal{L}; \quad l_{n,p} \ll \mathcal{L},$$

где первое неравенство означает выполнение предположений диодной, а второе — диффузионной теорий.

Последние неравенства могут быть переписаны в ином виде, если учесть связь $E_0 = \frac{2\varphi_0}{e\mathcal{L}}$. Тогда

$$e l_n E_0 \gtrless 2\varphi_0. \quad (2.24)$$

Это неравенство означает, что диодная теория будет выполняться, если носитель заряда на длине свободного заряда приобретает энергию, значительно большую, чем удвоенная высота потенциального барьера области пространственного заряда. Диффузионная теория будет выполняться при обратных соотношениях между энергией, приобретаемой носителем заряда на длине свободного пробега, и высотой потенциального барьера.

Длина свободного пробега носителей заряда может быть выражена через подвижность $u_{n,p}$ следующим образом:

$$l_{n,p} = \frac{3\pi u_{n,p} m_{n,p}^* v_{n,p}}{8e}$$

Подставив постоянные величины, можно записать для $T = 300 \text{ K}$

$$l_{n,p} = 6,6 \cdot 10^{-9} u_{n,p} \left(\frac{m_{n,p}^*}{m_0} \right)^{1/2}$$

Теперь можно переписать неравенство (2.24) следующим образом:

$$E_0 \gtrless 3 \cdot 10^8 \frac{\varphi_0}{e u_{n,p}} \left(\frac{m_{n,p}^*}{m_0} \right)^{1/2}. \quad (2.25)$$

Как видно из последнего неравенства, для различных значений $u_{n,p}$, которая обычно изменяется в пределах $(10^2 \dots 10^4)$ см²/(В · с), Φ_0 и $m_{n,p}^*$ при постоянном E_0 в пределах $(10^2 \dots 10^5)$ В/см могут выполняться предположения либо диодной, либо диффузионной теорий.

Поскольку неравенство (2.25) определяется величинами высот барьеров и концентрации носителей заряда, от которой зависит величина напряженности поля, то даже для одного и того же полупроводника при разных высотах барьера или концентрациях может выполняться либо диодная, либо диффузионная теории. Возможно также выполнение предположений разных теорий для электронов и дырок из-за их разной подвижности и эффективной массы.

Следует особо отметить, что неравенства (2.24), (2.25) записаны по отношению к полю у границы полупроводника. Но в области пространственного заряда поле уменьшается от максимального значения у границы до нуля по мере удаления от этой границы. Поэтому неравенства, записанные в виде (2.24), (2.25), строго могут определить только условия применимости диффузионной теории.

В случае выполнения условия диодной теории у одной из границ слоя Шоттки, где E максимально, у другой границы, где $E \rightarrow 0$, эти условия выполняться не будут.

Для диффузионной теории, кроме того, следует учитывать более жесткие условия, определяемые неравенством (2.21), из которого следует, что энергия, приобретаемая на длине свободного пробега, не должна превышать $\frac{8}{\pi} kT$.

В том случае, когда справедливы уравнения (2.3) для области пространственного заряда, необходимо рассмотреть их решение. Пусть в области пространственного заряда i_n и $i_p = \text{const}$. Это означает, что в области пространственного заряда можно пренебречь изменением n и p за счет рекомбинации. Тогда диффузионно-дрейфовое уравнение для электронов (2.3а) может быть переписано в виде

$$\frac{dn}{dx} + \frac{n}{kT} \cdot \frac{d\Phi}{dx} = \frac{i_n}{eD_n}, \quad (2.26)$$

где использовано соотношение Эйнштейна $u_n = \frac{e}{kT} D_n$ и учтено,

$$\text{что } E(x) = \frac{1}{e} \frac{d\Phi(x)}{dx}.$$

При решении последнего уравнения для $i_n = \text{const}$ используем метод вариации постоянных. В этом случае решение уравнения (2.26) будем искать в виде

$$n(x) = C(x) e^{-\frac{\Phi(x)}{kT}}, \quad (2.27)$$

так как легко проверить, что $n(x) = C e^{-\frac{\Phi(x)}{kT}}$, где C — постоянная, является решением однородного уравнения

$$\frac{dn}{dx} + \frac{n}{kT} \frac{d\Phi}{dx} = 0.$$

Подставив (2.27) в (2.26), получим уравнение для нахождения

$$dC(x) = \frac{i_n}{eD_n} e^{\frac{\Phi(x)}{kT}} dx.$$

Интегрируя это выражение в пределах от x до x_1 , получим

$$C(x_1) - C(x) = \frac{i_n}{eD_n} \int_x^{x_1} e^{\frac{\Phi(x)}{kT}} dx, \quad (2.28)$$

где x_1 — точка, в которой задана концентрация $p(x)$ и $\Phi(x)$: Подставляя $C(x)$, найденное из выражения (2.28), в (2.27), можно получить $n(x)$ в следующем виде:

$$n(x) = \left[C(x_1) - \frac{i_n}{eD_n} \int_x^{x_1} e^{\frac{\Phi(x)}{kT}} dx \right] e^{-\frac{\Phi(x)}{kT}}. \quad (2.29)$$

Аналогичным образом может быть получено и решение диффузионно-дрейфового уравнения для дырочного тока. В результате

$$p(x) = \left[C_1(x_1) + \frac{i_p}{eD_p} \int_x^{x_1} e^{-\frac{\Phi(x)}{kT}} dx \right] e^{\frac{\Phi(x)}{kT}}, \quad (2.30)$$

где x_1 — также точка, в которой заданы концентрация $p(x)$ и $\Phi(x)$.

Полученные выражения для $n(x)$ и $p(x)$ могут быть использованы для нахождения токов в области пространственного заряда в предположениях диффузионной теории, если использовать граничные условия на границах с зазором и областью пространственного заряда.

Если в области пространственного заряда существенную роль играют межзонные переходы с участием локальных уровней, показанные на рис. 2.3, то уже нельзя считать, что i_n и i_p не зависят от координаты. В этом случае для нахождения токов через область пространственного заряда неправомерно применять полученные выше решения диффузионно-дрейфовых уравнений для $n(x)$ и $p(x)$, а необходимо решать полную систему уравнений (2.4). А эта задача, как уже отмечалось выше, в общем случае не решается.

Задача нахождения токов с участием локальных уровней может быть решена, если пренебречь сквозными токами в пределах разрешенных зон полупроводника. Тогда нахождение вольт-амперной характеристики сводится к рассмотрению перехода носителей заряда с зоны проводимости на уровень, с уровня в валентную зону и обратных переходов.

В этом случае для одного локального уровня с концентрацией N_t плотность тока между уровнем и зоной проводимости может быть записана так:

$$i_{nt} = e [k_{nt} (N_t - n_t) n - k_{tn} n_t],$$

а плотность тока между уровнем и валентной зоной так:

$$i_{vt} = e [k_{tv} n_t p - k_{vt} (N_t - n_t)],$$

где k_{nt} , k_{in} , k_{iv} , k_{vt} — кинетические коэффициенты, характеризующие переходы уровень — зона проводимости и уровень — валентная зона; n_i — концентрация электронов на уровне; n и p — концентрация электронов и дырок в зоне проводимости и валентной зоне в областях, с которыми идет электронный обмен.

За положительное направление тока здесь выбрано направление движения электронов от полупроводника 2 к полупроводнику 1.

Такая запись токов i_{nt} и i_{in} отражает следующие обстоятельства. Интенсивность переходов электронов зоны проводимости на уровень пропорциональна количеству свободных мест на уровне и концентрации электронов, участвующих в переходах, а интенсивность обратных переходов пропорциональна концентрации электронов на уровнях. Интенсивность электронных переходов с уровня в валентную зону пропорциональна концентрации электронов на уровне количеству свободных мест в валентной зоне, а интенсивность обратных переходов — пропорциональна концентрации свободных мест на уровне.

В заключение можно отметить, что движение носителей заряда через области пространственного заряда обычно и контролирует прохождение тока через гетеропереходы.

2.4. ПРОХОЖДЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА В ЗАЗОРЕ (ПРОМЕЖУТОЧНОМ СЛОЕ)

Как уже рассматривалось выше, возможны два параллельных механизма переноса носителей заряда через зазор (промежуточный слой), а именно, без участия и с участием пересадочных центров.

При рассмотрении туннелирования через зазор без участия пересадочных центров обычно предполагают, что зазор можно рассматривать как диэлектрик, для которого существует запрещенная зона. В этом случае, если потенциальная энергия меняется достаточно плавно, а прозрачность зазора достаточно мала (например, электрон проходит далеко под дном зоны проводимости зазора), можно пользоваться следующим выражением коэффициента прозрачности трапецидального барьера, полученным в ВКБ-приближении при интегрировании (2.22):

$$\mathcal{P} = \mathcal{P}_0 \exp \left\{ -\frac{4}{3} \left(\frac{2m^*}{\hbar^2} \right)^{1/2} \frac{d}{u_1 - u_2} [(u_1 - \mathcal{E})^{3/2} - (u_2 - \mathcal{E})^{3/2}] \right\}, \quad (2.31)$$

где \mathcal{E} отсчитывается от вершины потенциального барьера полупроводника 2 (рис. 2.3). Как видно из этого выражения, коэффициент прозрачности сильно изменяется с изменением высот потенциально-го барьера и толщины зазора d .

Если переходы идут с участием поверхностных уровней в качестве пересадочных центров, в стационарном случае можно записать следующие выражения для токов, характеризующих все возможные переходы между i -м уровнем и разрешенными зонами:

$$i_n^2 = e [K_{nt}^2 (n_i - m_i) n_k^2 - K_{in}^2 m_i]; \quad (2.32)$$

$$i_p^2 = e [K_{vt}^2 (n_i - m_i) - K_{iv}^2 m_i p_k^2]; \quad (2.33)$$

$$i_n^1 = e [K_{in}^1 m_i - K_{nt}^1 (n_i - m_i) n_k^1]; \quad (2.34)$$

$$i_p^1 = e [K_{iv}^1 m_i p_k^1 - K_{vt}^1 (n_i - m_i)]. \quad (2.35)$$

Здесь n_i — концентрация i -го поверхностного состояния; m_i — концентрация электронов на i -м уровне; $K_{nt}^{2,1}$, $K_{vt}^{2,1}$ — кинетические коэффициенты, характеризующие переход на i -е состояние; $K_{in}^{1,2}$, $K_{iv}^{1,2}$ — кинетические коэффициенты, характеризующие обратные переходы; $n_k^{1,2}$ и $p_k^{1,2}$ — концентрация электронов и дырок в тех областях полупроводников 1 и 2, с которыми идет электронный обмен. За положительное направление тока выбрано направление движения электронов от полупроводника 2 к полупроводнику 1. Такая запись токов через поверхностные уровни связана с тем, что токи с зоны проводимости полупроводников 1 и 2 на уровне пропорциональны концентрации свободных мест и концентрации электронов зоны проводимости, принимающих участие в электронном обмене, а обратные токи пропорциональны концентрации электронов на уровнях. Токи с уровня в валентные зоны полупроводников 1 и 2 пропорциональны концентрации электронов на уровне и количеству свободных мест в валентной зоне, а обратные токи пропорциональны количеству свободных мест на поверхностных уровнях.

Использование записанных в настоящем параграфе выражений позволяет проанализировать прохождение носителей заряда через зазор, который также может контролировать токи в гетеропереходе.

2.5. СООТНОШЕНИЕ МЕЖДУ НАПРЯЖЕНИЯМИ В ГЕТЕРОПЕРЕХОДЕ

Рассмотрим, как будет перераспределяться между полупроводниками и зазором приложенное к гетеропереходу напряжение V . Для нахождения соотношений между падениями напряжения в различных частях гетероперехода нужно использовать условия непрерывности вектора электрической индукции для случаев, когда к гетеропереходу не приложено и приложено напряжение. Эти соотношения для случая, когда на границах раздела полупроводников с зазором существуют поверхностные уровни, имеют вид:

$$\epsilon_0 \epsilon_1 E_1^0 = \epsilon_0 \epsilon_3 E_3^0 + \sigma_1^0; \quad (2.36)$$

$$\epsilon_0 \epsilon_3 E_3^0 = \epsilon_0 \epsilon_2 E_2^0 + \sigma_2^0; \quad (2.37)$$

$$\epsilon_0 \epsilon_1 E_1 = \epsilon_0 \epsilon_3 E_3 + \sigma_1; \quad (2.38)$$

$$\epsilon_0 \epsilon_3 E_3 = \epsilon_0 \epsilon_2 E_2 + \sigma_2, \quad (2.39)$$

где значки «0» у напряженности электрического поля E и плотности поверхностных зарядов означают, что к гетеропереходу напряжение не приложено. Вычитая уравнение (2.36) из (2.37) и (2.39) из (2.37),

можно получить систему уравнений

$$\epsilon_0 \epsilon_1 (E_1^0 - E_1) = \epsilon_0 \epsilon_3 (E_3^0 - E_3) + \sigma_1^0 - \sigma_1; \quad (2.40)$$

$$\epsilon_0 \epsilon_3 (E_3^0 - E_3) = \epsilon_0 \epsilon_2 (E_2^0 - E_2) + \sigma_2^0 - \sigma_2; \quad (2.41)$$

$$V = V_1 + V_3 + V_2, \quad (2.42)$$

для трех неизвестных V_1, V_3, V_2 .

Для решения системы уравнений (2.40) — (2.42) можно воспользоваться результатами п. 1.2. При этом следует учитывать, что высоты потенциального барьера в полупроводнике 1 и 2 $\phi_{1,2}$ изменятся на величину $eV_{1,2}$ и станут равными $\phi_{1,2} - eV_{1,2}$. Если ток, проходящий через область пространственного заряда, при этом не будет изменять $n(x)$ и нарушать условия реализации слоя Шоттки, то все рассмотренное выше при $V = 0$ для слоя Шоттки остается справедливым, но вместо граничного условия $\phi(0) = \phi_0$ нужно использовать условие $\phi_{1,2}(0) = \phi_{1,2} - eV$ и во всех полученных формулах вместо ϕ_0 подставлять $\phi_{1,2} - eV_{1,2}$. Например, ширина слоя Шоттки в полупроводнике 1 и 2 при равномерном распределении примесей запишется так:

$$\mathcal{L}_{1,2} = \left[\frac{2\epsilon_0 \epsilon (\phi_{1,2} - eV_{1,2})}{e^2 n_{1,2}} \right]^{1/2}.$$

Теперь ширина слоя Шоттки зависит от приложенного напряжения и уменьшается для напряжений, понижающих высоту потенциального барьера, прямых и увеличивается для обратных напряжений, увеличивающих высоту потенциального барьера.

Если к области пространственного заряда приложено напряжение, но проходящий ток не нарушает распределения подвижных зарядов в любой точке x , то остается справедливым проведенное выше рассмотрение общего случая при $V \neq 0$. Необходимо только в окончательных выражениях заменить величины ϕ_0 на $\phi_{1,2} - eV_{1,2}$.

Если ток нарушает распределение подвижных зарядов, то для нахождения $\phi(x)$ необходимо совместное решение уравнения Пуассона и диффузионно-дрейфовых уравнений. Такая задача в общем случае не решается аналитически.

Используя выражения (1.31) и (1.32), можно показать, что в случае диффузионной теории протекающий ток не нарушает распределение заряда электронов и дырок в области пространственного заряда, если выполняются неравенства:

$$|n(x_1)| \gg \frac{i_n}{eD_n} \int_x^{x_1} e^{-\frac{\phi(x)}{kT}} dx;$$

$$|p(x_1)| \gg \frac{i_p}{eD_p} \int_x^{x_1} e^{-\frac{\phi(x)}{kT}} dx.$$

Записанная система уравнений (2.40) — (2.42) может быть решена аналитически для случая, когда области пространственного

заряда в обоих полупроводниках являются слоем Шоттки. В этом случае:

$$E_1^0 - E_1 = - \left(\frac{2\rho_1}{\epsilon_0 \epsilon_1} \right)^{1/2} [(\phi_{10})^{1/2} - (\phi_1 - eV_1)^{1/2}];$$

$$E_2^0 - E_2 = - \left(\frac{2n_2}{\epsilon_0 \epsilon_2} \right)^{1/2} [(\phi_{20})^{1/2} - (\phi_2 - eV_2)^{1/2}].$$

После подстановки этих значений и величины $E_3^0 - E_3 = -\frac{V_3}{d}$ в уравнения (2.40) и (2.41), можно их решить относительно V_1 и V_2 , которые являются функцией V_3 . Затем, используя соотношение (2.42) и полученные решения, легко найти величину V_3 как функцию приложенного напряжения V .

В результате получаем

$$V_1 = \frac{\phi_{10}}{e} - \frac{1}{e} \left[\phi_{10}^{1/2} + \frac{V_1^i \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} - \frac{V_3 \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} \right]^2; \quad (2.43)$$

$$V_2 = \frac{\phi_{20}}{e} - \frac{1}{e} \left[\phi_{20}^{1/2} + \frac{V_2^i \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_2 n_2)^{1/2}} - \frac{V_3 \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_2 n_2)^{1/2}} \right]^2; \quad (2.44)$$

$$V_3 = -\frac{b}{2a} \pm \sqrt{\left(\frac{b}{2a}\right)^2 - \frac{c-V}{a}}, \quad (2.45)$$

где

$$V_1^i = \frac{(\sigma_1^0 - \sigma_1)}{\epsilon_0 \epsilon_3} d; \quad V_2^i = \frac{(\sigma_2^0 - \sigma_2)}{\epsilon_0 \epsilon_3} d;$$

$$a = \frac{1}{2e} \left(\frac{\epsilon_0 \epsilon_3}{d} \right)^2 \left(\frac{1}{\epsilon_0 \epsilon_1 \rho_1} + \frac{1}{\epsilon_0 \epsilon_2 n_2} \right);$$

$$b = 1 + \frac{2\epsilon_0 \epsilon_3}{ed} \left\{ \frac{\phi_{10}^{1/2}}{(2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} + \frac{\phi_{20}^{1/2}}{(2\epsilon_0 \epsilon_2 n_2)^{1/2}} + \frac{\epsilon_0 \epsilon_3}{2d} \left[\frac{V_1^i}{\epsilon_0 \epsilon_1 \rho_1} + \frac{V_2^i}{\epsilon_0 \epsilon_2 n_2} \right] \right\};$$

$$c = \frac{\phi_1 + \phi_2}{e} - \frac{1}{e} \left[\left(\phi_{10}^{1/2} + \frac{V_1^i \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} \right)^2 + \left(\phi_{20}^{1/2} + \frac{V_2^i \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_2 n_2)^{1/2}} \right)^2 \right].$$

Как видно из выражений (2.43) — (2.45), величины V_1, V_2 и V_3 при постоянном V зависят от толщины d и диэлектрической проницаемости зазора ϵ_3 , от концентраций носителей заряда в полупроводниках ρ_1 и n_2 , от их диэлектрических проницаемостей ϵ_1 и ϵ_2 , от высот барьеров ϕ_{10} и ϕ_{20} и параметров поверхностных уровней.

При отсутствии поверхностных уровней $V_1^i = V_2^i = 0$ и выражения (2.43) — (2.44) упрощаются:

$$V_1 = \frac{\phi_{10}}{e} - \frac{1}{e} \left[\phi_{10}^{1/2} - \frac{V_3 \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} \right]^2;$$

$$V_2 = \frac{\phi_{20}}{e} - \frac{1}{e} \left[\phi_{20}^{1/2} - \frac{V_3 \epsilon_0 \epsilon_3}{d (2\epsilon_0 \epsilon_2 n_2)^{1/2}} \right]^2;$$

$$b = 1 + \frac{2\epsilon_0 \epsilon_3}{ed} \left[\frac{\phi_{10}^{1/2}}{(2\epsilon_0 \epsilon_1 \rho_1)^{1/2}} + \frac{\phi_{20}^{1/2}}{(2\epsilon_0 \epsilon_2 n_2)^{1/2}} \right];$$

$c = 0.$

Как показывает анализ полученных выражений, падение напряжения в зазоре V_3 будет больше для больших p_1, n_2, d . Падение напряжения будет больше в полупроводнике с меньшими концентрациями примесей, высотами барьеров и диэлектрическими проницаемостями.

Поверхностные уровни могут как уменьшать, так и увеличивать значения V_1, V_2, V_3 при постоянных значениях остальных параметров. Это определяется тем, что поверхностные уровни могут как освобождаться, так и заполняться при приложении определенного V , что приводит к изменению знака V_1^i и V_2^i . В самом деле, заполнение поверхностных уровней определяется уровнем Ферми полупроводника, с которым идет наиболее интенсивный электронный обмен. Если обмен идет с ближайшим полупроводником, приложение прямого напряжения приводит к освобождению уровня, а обратного — к их заполнению. Это и может приводить к увеличению или уменьшению величин σ_1, σ_2 и $V_{1,2}^i$ в зависимости от свойств поверхностных уровней.

В частном случае, когда в гетеропереходе $p_1 \gg n_1$, что соответствует контакту металла с полупроводником, $\Phi_{10} \rightarrow 0, V_1 = 0$ и $V_1^i = 0$, имеем:

$$V_3 = -V_2^i + \frac{en_2\epsilon_0\epsilon_2d^2}{(\epsilon_0\epsilon_3)^2} + \frac{(2n_0\epsilon_0\epsilon_2\Phi_2)^{1/2}d}{\epsilon_0\epsilon_3} - \sqrt{\left[\frac{(2n_0\epsilon_0\epsilon_2\Phi_2)^{1/2}d}{\epsilon_0\epsilon_3} + \frac{en_2\epsilon_0\epsilon_2d^2}{(\epsilon_0\epsilon_3)^2} \right]^2 - \frac{2en_2\epsilon_0\epsilon_2d^2}{\epsilon_0\epsilon_3} (V + V_2^i)}; \quad (2.46)$$

$$V_2 = V - V_3.$$

При $d = 0$ $V_3 = 0, V_2 = V$ и все напряжение падает в области пространственного заряда полупроводника.

В другом частном случае, когда в гетеропереходе зазора нет и $\sigma_1 = \sigma_2 = 0, d = 0, V_3 = 0, V_1^i = 0$ и $V_2^i = 0$,

$$V_1 = \frac{V}{1 + \frac{\epsilon_1 p_1}{\epsilon_2 n_2}}; \quad (2.47)$$

$$V_2 = \frac{V}{1 + \frac{\epsilon_2 n_2}{\epsilon_1 p_1}}. \quad (2.48)$$

При $p \gg n_2$ $V_1 \rightarrow 0$, а $V_2 = V$, и при $n_2 \gg p_1$ $V_2 \rightarrow 0$, а $V_1 = V$ и все напряжение падает в полупроводнике с меньшей концентрацией примесей.

В заключение отметим, что хотя приведенные выше выражения для V_1, V_3 и V_2 получены для случая, когда в обеих частях гетероперехода реализуются слои Шоттки, качественные выводы о перераспределении напряжения между различными полупроводниками и зазором в зависимости от концентрации примесей в полупроводниках, толщины зазора и других параметров могут быть перенесены и на случай, когда области пространственного заряда в полупроводниках не являются слоями Шоттки.

Глава 3. ВЫПРЯМЛЕНИЕ В КОНТАКТЕ МЕТАЛЛ — ПОЛУПРОВОДНИК

3.1. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА КОНТАКТА С УЧЕТОМ ОДНОГО ТИПА НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим вольт-амперную характеристику контакта металл — полупроводник с зазором в предположениях диодной и диффузионной теорий для случая, когда ток переносится только основными носителями заряда без участия поверхностных уровней. Для n -типа полупроводника, изображенного на рис. 3.1, это будут электроны. Как уже рассмотрено выше, для диодной теории нахождение вольт-амперной характеристики сводится к вычислению разности плотности токов из полупроводника в металл i_2 и из металла в полупроводник i_1 , над барьером области пространственного заряда

$$i_n = i_2 - i_1, \quad (3.1)$$

где за положительное направление электронного тока выбрано направление, соответствующее току из полупроводника в металл. Для нахождения плотности токов i_2 и i_1 воспользуемся выражением для термоэлектронной эмиссии (1.1), на основании которого можно получить для равномерного распределения примесей в полупроводнике

$$i_{1,2} = e\mathcal{P}_n A T^2 e^{-\frac{\Phi_{1,2}^b}{kT}} = \mathcal{P}_n \frac{e\bar{v}_n n_0}{4} e^{-\frac{\Phi_{1,2}}{kT}}. \quad (3.2)$$

В этом выражении, справедливом только при $\Phi_{1,2}^b \gg kT$, дополнительно введено среднее значение коэффициента прозрачности электронов \mathcal{P}_n для учета того, что не все носители заряда проходят через зазор между металлом и полупроводником, и использованы соотношения:

$$\Phi_{1,2} = \Phi_{1,2}^b - \mu; \quad \bar{v}_n = \left(\frac{8kT}{\pi m^*} \right)^{1/2};$$

$$n_0 = 2 \left(\frac{2\pi m^* kT}{h^2} \right)^{3/2} e^{-\frac{\mu}{kT}}.$$

В ВКБ-приближении для трапециoidalного барьера, изображенного на рис. 3.1, средний коэффициент прозрачности электронов записан в п. 2.4 (формула 2.31).

Как видно из рис. 3.1, б, при приложении к контакту напряжения:

$$\begin{aligned} \Phi_1^b &= \Phi_0 + eV_1 + \mu, & \Phi_1 &= \Phi_0 + eV_1; \\ \Phi_2 &= \Phi_0 - eV_2 + \mu, & \Phi_2 &= \Phi_0 - eV_2, \end{aligned} \quad (3.3)$$

где V_1 и V_2 — части общего падения напряжения $V = V_1 + V_2$ в зазоре и области пространственного заряда полупроводника, связанные с V с помощью выражений, полученных в п. 2.5 предыдущей главы:

Используя выражения для $\phi_{1,2}$ и $\phi_{1,2}$, получаем, что вольт-амперная характеристика контакта металл — полупроводник в предположениях диодной теории имеет вид

$$i_n = eAT^2 \mathcal{P}_n e^{-\frac{\phi_b}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right) = \frac{en_0 \bar{v}_n \mathcal{P}_n}{4} e^{-\frac{\phi_0}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right). \quad (3.4)$$

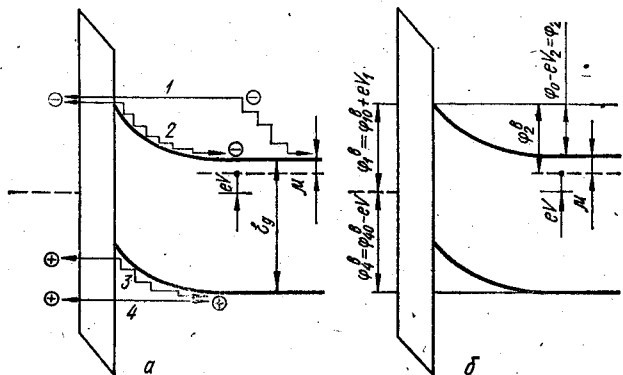


Рис. 3.1. Энергетическая диаграмма контакта металл — полупроводник при приложении напряжения:

а — токи через контакт (1, 4 — диодная теория, 2, 3 — диффузионная теория); б — принятые обозначения.

Диффузионная теория. В предположениях диффузионной теории для нахождения вольт-амперной характеристики в одномерном случае и учета только основных носителей заряда необходимо решать в области пространственного заряда диффузионно-дрейфовое уравнение (2.3а). Общее решение этого уравнения, полученное в предыдущей главе, теперь может быть переписано в виде

$$n(x) = \left[C(\mathcal{L}) + \frac{i_n}{eD_n} \int_x^{\mathcal{L}} e^{\frac{\phi(x)}{kT}} dx \right] e^{-\frac{\phi(x)}{kT}}, \quad (3.5)$$

где интегрирование проведено от x до $x = \mathcal{L}$, соответствующей концу области пространственного заряда, а знак плюс перед вторым членом связан с выбором положительного направления тока электронов от полупроводника к металлу.

Граничные условия, необходимые для решения задачи, можно записать для слоя Шоттки при равномерном распределении приме-

сей следующим образом:

$$\left. \begin{array}{l} \text{при} \quad x = 0, \quad \phi(0) = \phi_0 - eV_2 \\ \quad \quad n(0) = n_0 e^{-\frac{\phi_0 - eV_1}{kT}} - \frac{4i_n}{e\bar{v}_n \mathcal{P}_n}; \\ \text{при} \quad x = \mathcal{L}, \quad \phi(\mathcal{L}) = 0 \\ \quad \quad n(\mathcal{L}) = n_0. \end{array} \right\} \quad (3.6)$$

Если запись значений $\phi(0)$, $\phi(\mathcal{L})$ и $n(\mathcal{L})$ очевидна, то граничное условие для $n(0)$ требует специального пояснения.

Для нахождения $n(0)$ будем рассматривать разность плотности токов $i_n = i_2 - i_1$ в плоскости раздела полупроводника и зазора. Ток i_1 имеет то же значение, что и в случае диодной теории. Ток i_2 в случае диффузионной теории может быть записан через концентрацию $n(0)$ как плотность тока, проходящего через границу полупроводник — зазор:

$$i_2 = \frac{e\bar{v}_n n(0) \mathcal{P}_n}{4}. \quad (3.7)$$

В результате

$$i_n = \frac{e\bar{v}_n(0) \mathcal{P}_n}{4} - \frac{e\bar{v}_n n_0 \mathcal{P}_n}{4} e^{-\frac{\phi_0 + eV_1}{kT}}. \quad (3.8)$$

Из последнего равенства легко получить $n(0)$ в записанном выше виде.

Следует отметить, что, строго говоря, выражение (3.7) относится к симметричному распределению электронов по скоростям. При наличии тока симметрия нарушается. Если ток i_n мал по сравнению с тепловым током (3.7), то такое нарушение может быть учтено. В этом случае граничное условие для $n(0)$ принимает вид (3.6), если заменить \mathcal{P}_n на величину $\mathcal{P}_n \left(1 - \frac{1}{2} \mathcal{P}_n\right)$, которая при $\mathcal{P}_n \ll 1$ совпадает с \mathcal{P}_n , а при $\mathcal{P}_n = 1$ равна $1/2$.

Используя граничные условия при $x = \mathcal{L}$, находим, что $C(\mathcal{L}) = n_0$. Далее используем граничное условие для $x = 0$. Тогда

$$n_0 e^{\frac{\phi_0 - eV_1}{kT}} + \frac{i_n}{4e\bar{v}_n \mathcal{P}_n} = \left[n_0 + \frac{i_n}{eD_n} \int_0^{\mathcal{L}} e^{\frac{\phi(x)}{kT}} dx \right] e^{-\frac{\phi_0 - eV_2}{kT}}. \quad (3.9)$$

Решая это уравнение относительно тока, можно получить:

$$i_n = \frac{en_0 e^{-\frac{\phi_0}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right)}{4 - \frac{e\bar{v}_n \mathcal{P}_n}{e\bar{v}_n \mathcal{P}_n} - e^{-\frac{\phi_0 - eV_2}{kT}} |D_n \int_0^{\mathcal{L}} e^{\frac{\phi(x)}{kT}} dx}. \quad (3.10)$$

Для нахождения окончательного выражения вольт-амперной характеристики следует вычислить интеграл $\int_0^{\mathcal{L}} e^{\frac{\varphi(x)}{kT}} dx$. Но даже

для слоя Шоттки, когда $\varphi(x) \sim (\mathcal{L}^0 - x)^2$ и интеграл имеет довольно простой вид, точное значение интеграла может быть получено только численным способом.

Поэтому проведем приближенное вычисление этого интеграла, позволяющее получить аналитическое решение задачи, используя некоторые общие свойства подынтегральной функции, для чего представим интеграл в виде

$$\int_0^{\mathcal{L}} e^{\frac{\varphi(x)}{kT}} \left(\frac{d\varphi}{dx}\right)^{-1} d\varphi. \quad (3.11)$$

Поскольку функция $e^{\frac{\varphi(x)}{kT}}$ значительно сильнее изменяется с изменением φ по сравнению $\frac{d\varphi}{dx} \sim E(x)$ (см. предыдущую главу), то в интеграле (3.11) можно вынести $\frac{d\varphi}{dx}$ из-под знака интеграла в точке, где $e^{\frac{\varphi(x)}{kT}}$ максимально. Это соответствует вынесению $\left(\frac{d\varphi}{dx}\right)$ в точке вблизи $x = 0$. В результате можно получить

$$\begin{aligned} \int_0^{\mathcal{L}} e^{\frac{\varphi(x)}{kT}} \left(\frac{d\varphi}{dx}\right)^{-1} d\varphi &\simeq \left(\frac{d\varphi}{dx}\right)^{-1}_{x=0} \int_{\varphi_0 - eV_2}^0 e^{\frac{\varphi}{kT}} d\varphi = \\ &= kT \left(\frac{d\varphi}{dx}\right)^{-1}_{x=0} (1 - e^{-\frac{\varphi_0 - eV_2}{kT}}) \simeq -\frac{kT}{eE_0} e^{\frac{\varphi_0 - eV_2}{kT}}, \end{aligned} \quad (3.12)$$

где учтено, что в случае слоя Шоттки $e^{\frac{\varphi_0 - eV_2}{kT}} \gg 1$; $\frac{1}{e} \left(\frac{d\varphi}{dx}\right)_{x=0} = E_0$ — напряженность электрического поля при $x = 0$.

Подставляя (3.12) в (3.10), получаем окончательное выражение вольт-амперной характеристики в предположениях диффузионной теории

$$i_n = \frac{e\bar{v}_n \mathcal{P}_n n_0}{4} e^{-\frac{\varphi_0}{kT}} \left(\frac{E_0 u_n}{\bar{v}_n \mathcal{P}_n + E_0 u_n} \right) (e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}}), \quad (3.13)$$

где использовано также соотношение Эйнштейна $D_n = u_n \frac{kT}{e}$.

Анализ вольт-амперных характеристик. Выражение вольт-амперных характеристик как для диодной теории, так и для диффузионной может быть записано в общем виде:

$$i = i_s (e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}}), \quad (3.14)$$

где для диодной теории

$$i_s = \frac{e\bar{v}_n n_0 \mathcal{P}_n}{4} e^{-\frac{\varphi_0}{kT}}; \quad (3.15)$$

для диффузионной теории

$$i_s = \frac{e\bar{v}_n n_0 \mathcal{P}_n}{4} e^{-\frac{\varphi_0}{kT}} \frac{E_0 u_n}{\frac{\bar{v}_n \mathcal{P}_n}{4} + E_0 u_n}. \quad (3.16)$$

Используя выражения (3.14) — (3.16), можно детально проанализировать вольт-амперную характеристику.

Так, при $d \rightarrow 0$, когда $V \rightarrow 0$, $\mathcal{P}_n \rightarrow 1$ и $V_2 \rightarrow V$,

$$i = i_s (e^{\frac{eV}{kT}} - 1), \quad (3.17)$$

где для диодной теории

$$i_s = \frac{e\bar{v}_n n_0}{4} e^{\frac{\varphi_0}{kT}}; \quad (3.18)$$

для диффузионной теории

$$i_s = e u_n n_0 E_0 e^{-\frac{\varphi_0}{kT}} \quad (3.19)$$

(при записи значения i_s для диффузионной теории учтено, что средняя тепловая скорость значительно больше дрейфовой и $\bar{v}_n \gg 4E_0 u_n$).

Как видно из выражения (3.17), при толщине зазора $d = 0$ в случае диодной теории прямой ток при достаточно больших напряжениях пропорционален $e^{\frac{eV}{kT}}$, а обратный — постоянен. В случае диффузионной теории при тех же условиях прямой ток пропорционален $E_0 e^{\frac{eV}{kT}}$, а обратный E_0 .

Поскольку для прямого направления с ростом напряжения E_0 уменьшается, то это приведет к более слабому росту прямого тока для диффузионной теории по сравнению с диодной. При обратных напряжениях E_0 растет с увеличением напряжения. Это приводит к тому, что для диффузионной теории обратный ток не испытывает насыщения при росте обратного напряжения. Вольт-амперные характеристики для диодной и диффузионной теорий при $d = 0$ показаны на рис. 3.2. На этом рисунке учтено также то, что для диодной теории i_s больше, чем для диффузионной, так как

$$\bar{v}_n > 4u_n E_0.$$

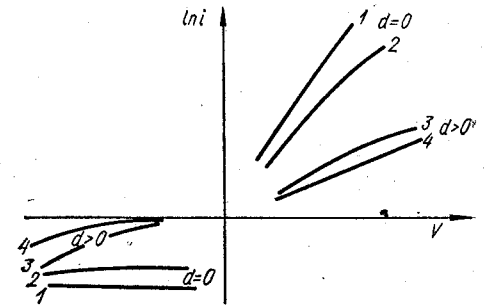


Рис. 3.2. Вольт-амперные характеристики контакта металла — полупроводник для одного типа носителей заряда (1, 3 — диодная теория; 2, 4 — диффузионная теория).

Если $d \neq 0$, то V_1 отлично от нуля, а V_2 от V . Это приводит к тому, что прямой ток растет с напряжением медленней и пропорционален $e \frac{eV_2}{kT}$ для диодной теории и диффузионной при достаточно малых коэффициентах прозрачности, когда $\frac{\bar{v}_n \mathcal{P}_n}{4} \ll E_0 \mu_n$. Для диффузион-

ной теории при $\frac{\bar{v}_n \mathcal{P}_n}{4} \gg E_0 \mu_n$ прямой ток пропорционален $E_0 e \frac{eV_2}{kT}$.

Обратный ток при $d \neq 0$ не испытывает насыщения и пропорционален $e \frac{eV_1}{kT}$ для диодной теории и диффузионной теории при $\frac{\bar{v}_n \mathcal{P}_n}{4} \ll E_0 \mu_n$. Для диффузионной теории при $\frac{\bar{v}_n \mathcal{P}_n}{4} \gg E_0 \mu_n$ обратный ток пропорционален $E_0 e \frac{eV_1}{kT}$.

Кроме того, с изменением толщины зазора изменяются величины i_s для диодной и диффузионной теории. Для диодной теории величина $i_s \sim \mathcal{P}_n$ и все время уменьшается с ростом d . Для диффузионной теории зависимость i_s от \mathcal{P}_n сложнее. Так, при малых d , когда $\frac{\bar{v}_n \mathcal{P}_n}{4} \ll E_0 \mu_n$, величина i_s не зависит от d . При достаточно больших толщинах d , когда $\frac{\bar{v}_n \mathcal{P}_n}{4} \ll E_0 \mu_n$, величина $i_s \sim \mathcal{P}_n$ и в точности совпадает с i_s , полученным в предположениях диодной теории.

Это показано на рис. 3.2, где также учтено, что i_s для диодной теории больше, чем для диффузионной теории в $1 + \frac{\bar{v}_n \mathcal{P}_n}{4E_0 \mu_n}$ раз, как это следует из сравнения выражения (3.4) и (3.13).

Наблюдаемое изменение зависимости тока от приложенного напряжения при изменении толщины зазора с физической точки зрения обуславливается тем, что прямой ток контролируется изменением высоты потенциального барьера для электронов, идущих из полупроводника в металл. В случае, когда $d=0$, эта высота изменяется на величину приложенного напряжения V . При $d \neq 0$ ее изменение определяется только частью приложенного напряжения, а именно напряжением V_1 .

Обратный ток как в случае диодной, так и диффузионной теории контролируется изменением высоты потенциального барьера для электронов, идущих из металла в полупроводник. Когда $d=0$, высота этого барьера не изменяется и обратный ток для диодной теории испытывает насыщение. Для диффузионной теории он пропорционален E_0 из-за того, что для прохождения тока также существенны условия в области пространственного заряда полупроводника. При $d \neq 0$ высота потенциального барьера для электронов, идущих из металла в полупроводник, изменяется на величину, определяемую напряжением V_1 . Поэтому обратный ток как для диодной, так и диффузионной теории не испытывает насыщения.

Наблюдаемые изменения i_s для диодной теории определяются тем, что во всем интервале изменения зазора ток контролируется прохождением через этот зазор.

Измерение i_s для диффузионной теории с физической точки зрения определяется тем, что при малых толщинах d ток контролируется прохождением носителей заряда через область пространственного заряда полупроводника, а при больших d прохождением через область зазора.

В заключение отметим, что формальное совпадение выражений вольт-амперных характеристик, полученных в приближениях диодной и диффузионной теорий в частном случае достаточно больших d , не означает, что диффузионная теория является более общей. Исходные предположения, закладываемые при решении задачи в предположениях диодной и диффузионной теорий, как уже указывалось, совершенно различны.

3.2. ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ КОНТАКТА С УЧЕТОМ ДВУХ ТИПОВ НОСИТЕЛЕЙ ЗАРЯДА (МАЛЫЙ УРОВЕНЬ ИНЖЕКЦИИ)

Пусть в переносе тока через контакт металл — n -полупроводник принимают участие электроны и дырки (см. рис. 3.1). Тогда общий ток равен сумме электронного и дырочного токов, т. е.

$$i = i_n + i_p.$$

Если уровень инжекции дырок в n -полупроводнике мал и $\Delta p \ll n_0$, то условия для прохождения электронов через контакт не изменяются и для тока i_n при равномерном распределении примесей можно воспользоваться выражениями, полученными в предыдущем параграфе для диодной и диффузионной теорий.

Для нахождения дырочного тока необходимо сшивать решения, полученные в предположении диодной или диффузионной теории для области пространственного заряда, с решением для квазинейтральной области.

Дырки, которые прошли через область пространственного заряда и попали в объем n -типа полупроводника, являются неосновными носителями заряда. Их движение описывается диффузионно-дрейфовым уравнением и уравнением неразрывности, решение которых для стационарного случая и одномерной задачи при малых уровнях инжекции получено в п. 2 гл. 2.

В точке $x = \mathcal{L}$ дырочный ток может быть получен из уравнения (2.18):

$$i_p = \frac{eD_p \Delta p_k}{l_p}. \quad (3.20)$$

Когда в области пространственного заряда можно пренебречь генерацией и рекомбинацией дырок, ток дырок постоянен во всей области пространственного заряда ($i_p(0) = i_p(\mathcal{L})$) и можно сшивать решения, полученные в любой плоскости области пространственного

заряда в предположениях диодной или диффузионной теорий, с решением при $x = \mathcal{L}$, полученным для квазинейтральной области.

Диодная теория. В случае диодной теории для нахождения вольт-амперной характеристики дырочного тока следует вычислить разность дырочных токов из полупроводника в металл i_3 и из металла в полупроводник i_4 , проходящих через область пространственного заряда

$$i_p = i_4 - i_3.$$

Здесь за положительное направление тока дырок выбрано направление, соответствующее току из металла в полупроводник, которое соответствует ранее выбранному положительному току электронов из полупроводника в металл.

Для плотности тока i_4 воспользуемся выражением термоэлектронной эмиссии, которое для дырок при $\varphi_4^b \gg kT$ записывается в виде

$$i_4 = e \mathcal{P}_p A T^2 e^{-\frac{\varphi_4^b}{kT}} = e \mathcal{P}_p \frac{\bar{v}_p p_0}{4} e^{-\frac{\varphi_4}{kT}}, \quad (3.21)$$

где

$$\varphi_4^b = \varphi_4 + \varepsilon - \mu = V_1 + V_2 + \varepsilon - \mu; \quad \bar{v}_p = \left(\frac{8kT}{\pi m_p^*} \right)^{1/2};$$

p_0 — равновесная концентрация дырок в n -полупроводнике. Ток i_3 может быть выражен следующим образом:

$$i_3 = \frac{e \bar{v}_p p(\mathcal{L})}{4} \mathcal{P}_p = \frac{e \bar{v}_p (p_0 + \Delta p_k)}{4} \mathcal{P}_p. \quad (3.22)$$

В выражениях (3.21) и (3.22) введено среднее значение коэффициента прозрачности дырок \mathcal{P}_p , учитывающего наличие зазора между металлом и полупроводником.

Теперь

$$i_p = \frac{e \bar{v}_p p_0 \mathcal{P}_p}{4} \left[e^{\frac{e(V_1+V_2)}{kT}} - \frac{p_0 + \Delta p_k}{p_0} \right]. \quad (3.23)$$

Подставляя в это выражение Δp_k , найденное из выражения (3.20) и равное $\Delta p_k = \frac{i_p l_p}{e D_p}$, получаем окончательное выражение для дырочного тока

$$i_p = \frac{e \bar{v}_p p_0 \mathcal{P}_p D_p}{4 D_p + \bar{v}_p \mathcal{P}_p l_p} \left[e^{\frac{e(V_1+V_2)}{kT}} - 1 \right]. \quad (3.24)$$

Суммарный ток через контакт в случае диодной теории может быть записан так:

$$i = (i_{sp} + i_{sn} e^{-\frac{eV_1}{kT}}) \left[e^{\frac{e(V_1+V_2)}{kT}} - 1 \right], \quad (3.25)$$

где

$$i_{sp} = \frac{e \bar{v}_p p_0 \mathcal{P}_p D_p}{4 D_p + \bar{v}_p \mathcal{P}_p l_p}; \quad (3.26)$$

$$i_{sn} = \frac{e \bar{v}_n n_0 \mathcal{P}_n}{4} e^{-\frac{\varphi_0}{kT}}. \quad (3.27)$$

Используя выражения для электронного и дырочного тока, можно найти коэффициент инжекции $\gamma = \frac{i_p}{i_n + i_p}$. В результате для диодной теории

$$\gamma = \left[1 + \frac{n_0}{p_0} \frac{\mathcal{P}_n}{\mathcal{P}_p} \frac{\bar{v}_n}{\bar{v}_p} e^{-\frac{\varphi_0 - eV_1}{kT}} \left(1 + \frac{\bar{v}_p \mathcal{P}_p l_p}{4 D_p} \right) \right]^{-1}. \quad (3.28)$$

Диффузионная теория. В случае диффузионной теории для нахождения дырочного тока в области пространственного заряда необходимо в одномерном случае решать диффузионно-дрейфовое уравнение (2.36). Как показано в предыдущей главе, его решение может быть представлено в виде

$$p(x) = \left[p(\mathcal{L}) + \frac{i_p}{e D_p} \int_x^{\mathcal{L}} e^{-\frac{\varphi(x)}{kT}} dx \right] e^{\frac{\varphi(x)}{kT}}. \quad (3.29)$$

Граничные условия для слоя Шоттки имеют вид: при $x = 0$

$$\varphi(0) = \varphi_0 - eV_2; \quad (3.30)$$

$$p(0) = p_0 e^{\frac{\varphi_0 + eV_1}{kT}} - \frac{4i_p}{e \bar{v}_p \mathcal{P}_p};$$

при $x = \mathcal{L}$, $\varphi(\mathcal{L}) = 0$

$$p(\mathcal{L}) = p_0 + \Delta p_k = p_0 + \frac{i_p l_p}{e D_p}. \quad (3.31)$$

Здесь второе граничное условие (3.30) получено при рассмотрении разности дырочных токов из металла в полупроводник и из полупроводника в металл аналогично тому, как это было сделано для электронного тока в предыдущем параграфе.

Используя записанные граничные условия, получаем, что

$$p_0 e^{\frac{\varphi_0 + eV_1}{kT}} - \frac{4i_p}{e \bar{v}_p \mathcal{P}_p} = \left[p_0 + \frac{i_p l_p}{e D_p} + \frac{i_p}{e D_p} \int_0^{\mathcal{L}} e^{-\frac{\varphi(x)}{kT}} dx \right] e^{\frac{\varphi_0 - eV_2}{kT}}. \quad (3.32)$$

Решая последнее уравнение относительно i_p , имеем

$$i_p = \frac{e p_0 e^{\frac{\varphi_0}{kT}} \left(e^{\frac{eV_1}{kT}} - e^{-\frac{eV_2}{kT}} \right)}{\frac{4}{\bar{v}_p D_p} + \frac{e}{\mu_p kT} e^{\frac{\varphi_0 - eV_2}{kT}} \left[\int_0^{\mathcal{L}} e^{-\frac{\varphi(x)}{kT}} dx + l_p \right]}. \quad (3.33)$$

Для нахождения интеграла $\int_0^{\infty} e^{-\frac{\varphi(x)}{kT}} dx$ при $\varphi(x) = \varphi(0) \left(1 - \frac{x}{\mathcal{L}}\right)^2$ можно перейти к переменным $y^2 = \frac{\varphi(0)}{kT} \left(1 - \frac{x}{\mathcal{L}}\right)^2$.

Тогда

$$\int_0^{\mathcal{L}} e^{-\frac{\varphi(x)}{kT}} dx = -\mathcal{L} \sqrt{\frac{kT}{\varphi_0 - eV_2}} \int_0^0 e^{-y^2} dy.$$

При достаточно больших значениях $\sqrt{\frac{\varphi_0 - eV_2}{kT}}$ в последнем интеграле можно перейти в пределе к ∞ .
В результате

$$\int_0^{\mathcal{L}} e^{-\frac{\varphi(x)}{kT}} dx = -\mathcal{L} \sqrt{\frac{kT}{\varphi_0 - eV_2}} \int_0^{\infty} e^{-y^2} dy = \frac{\mathcal{L}}{2} \sqrt{\frac{\pi kT}{\varphi_0 - eV_2}} = l_s.$$

Следует отметить, что всегда $l_s = \sqrt{\frac{kT}{8^2 e n_0}} < \mathcal{L} \ll l_p$, поскольку в противном случае следует учитывать генерацию или рекомбинацию неосновных носителей заряда в области пространственного заряда полупроводника.

Подставляя значение этого интеграла в выражение (3.33), можно найти окончательное выражение для дырочного тока в предположениях диффузионной теории:

$$i_p = \frac{e \rho_0 \bar{v}_p \mathcal{P}_p D_p [e^{\frac{e(V_1 + V_2)}{kT}} - 1]}{4 D_p e^{-\frac{\varphi_0 - eV_2}{kT}} + \bar{v}_p \mathcal{P}_p (l_p + l_s)}. \quad (3.34)$$

Суммарный ток в случае диффузионной теории:

$$i = (i_{sp} + i_{sn} e^{-\frac{eV_1}{kT}}) [e^{\frac{e(V_1 + V_2)}{kT}} - 1]; \quad (3.35)$$

где

$$i_{sn} = \frac{e \bar{v}_n \mathcal{P}_n n_0}{4} e^{-\frac{\varphi_0}{kT}} \left(\frac{E_0 u_n}{\bar{v}_n \mathcal{P}_n + E_0 u_n} \right);$$

$$i_{sp} = \frac{e \rho_0 \bar{v}_p \mathcal{P}_p D_p}{4 D_p e^{-\frac{\varphi_0 - eV_2}{kT}} + \bar{v}_p \mathcal{P}_p (l_p + l_s)}. \quad (3.36)$$

Используя полученные выражения, можно найти коэффициент инжекции для диффузионной теории в виде

$$\gamma = \left\{ 1 + \frac{n_0 \bar{v}_n \mathcal{P}_n u_n E_0 [4 D_p + \bar{v}_p \mathcal{P}_p (l_p + l_s) e^{-\frac{\varphi_0 - eV_2}{kT}}]}{\rho_0 \bar{v}_p \mathcal{P}_p D_p (4 u_n E_0 + \bar{v}_n \mathcal{P}_n)} \right\}^{-1} e^{-\frac{2\varphi_0 + e(V_1 - V_2)}{kT}}. \quad (3.37)$$

Анализ вольт-амперных характеристик и коэффициента инжекции. При отсутствии зазора, когда $d \rightarrow 0$, $V_1 \rightarrow 0$, $\mathcal{P}_n \rightarrow 1$, $\mathcal{P}_p \rightarrow 1$ и $V_2 = V$, для диодной и диффузионной теории вольт-амперная ха-

рактеристика записывается так:

$$i = (i_{sn} + i_{sp}) (e^{\frac{eV}{kT}} - 1). \quad (3.38)$$

Коэффициент инжекции

$$\gamma = \left(1 + \frac{i_{sn}}{i_{sp}} \right)^{-1},$$

где для диодной теории:

$$i_{sp} = \frac{e \bar{v}_p \rho_0 D_p}{4 D_p + \bar{v}_p l_p}; \quad i_{sn} = \frac{e \bar{v}_n n_0}{4} e^{-\frac{\varphi_0}{kT}}; \quad (3.39)$$

для диффузионной теории:

$$i_{sp} = \frac{e \bar{v}_p \rho_0 D_p}{4 D_p e^{-\frac{\varphi_0 - eV}{kT}} + \bar{v}_p (l_p + l_s)};$$

$$i_{sn} = e u_n E_0 n_0 e^{-\frac{\varphi_0}{kT}}. \quad (3.40)$$

Как видно, для диодной теории суммарный прямой ток при достаточно больших напряжениях пропорционален $e^{\frac{eV}{kT}}$, а обратный — постоянен. Коэффициент инжекции не зависит от напряжения.

Для диффузионной теории прямой ток может быть пропорциональным $E_0 e^{\frac{eV}{kT}}$, а обратный E_0 , если $i_{sn} \gg i_{sp}$. Если же $i_{sp} \gg i_{sn}$, то прямой ток пропорционален $e^{\frac{eV}{kT}}$ при $4 D_p e^{-\frac{\varphi_0 - eV}{kT}} \ll \bar{v}_p (l_p + l_s)$ или постоянен при $4 D_p e^{-\frac{\varphi_0 - eV}{kT}} \gg \bar{v}_p (l_p + l_s)$. Обратный ток постоянен при $\bar{v}_p (l_p + l_s) \gg 4 D_p e^{-\frac{\varphi_0 - eV}{kT}}$ и растёт пропорционально $e^{-\frac{eV}{kT}}$ при выполнении обратного неравенства.

Коэффициент инжекции примерно постоянен при выполнении неравенства

$$4 D_p e^{-\frac{\varphi_0 - eV}{kT}} \ll \bar{v}_p (l_p + l_s);$$

он несколько растёт в прямом направлении из-за изменения величины E_0 . При выполнении неравенства

$$4 D_p e^{-\frac{\varphi_0 - eV}{kT}} \gg \bar{v}_p (l_p + l_s)$$

коэффициент инжекции может расти с ростом увеличения прямого напряжения.

Для оценки максимального значения коэффициента инжекции при $d = 0$ в случае диодной и диффузионной теории оценим величину $\frac{i_{sn}}{i_{sp}}$. Поскольку эта величина пропорциональна $\frac{n_0}{\rho_0} e^{-\frac{\varphi_0}{kT}}$,

вначале найдем ее минимальное значение

$$\frac{n_0}{p_0} e^{-\frac{\varphi_0}{kT}} = \frac{N_c}{N_v} e^{-\frac{\mu}{kT}} e^{\frac{\varepsilon_g - 2\mu}{kT}}$$

Здесь для нахождения минимального значения искомой величины вместо φ_0 подставлено максимальное значение $\varphi_0 = \varepsilon_g - 2\mu$, при котором еще выполняются приближения слоя Шоттки.

Для диодной теории

$$\frac{i_{sn}}{i_{sp}} \geq \frac{4D_p + \bar{v}_p l_p}{4D_p} \frac{m_n^*}{m_p^*};$$

для диффузионной теории

$$\frac{i_{sn}}{i_{sp}} \geq \left(\frac{m_n^*}{m_p^*}\right)^{3/2} \frac{u_n E_0}{\bar{v}_p D_p} [4D_p e^{\frac{\varphi_0 - eV}{kT}} + \bar{v}_p (l_p + l_s)].$$

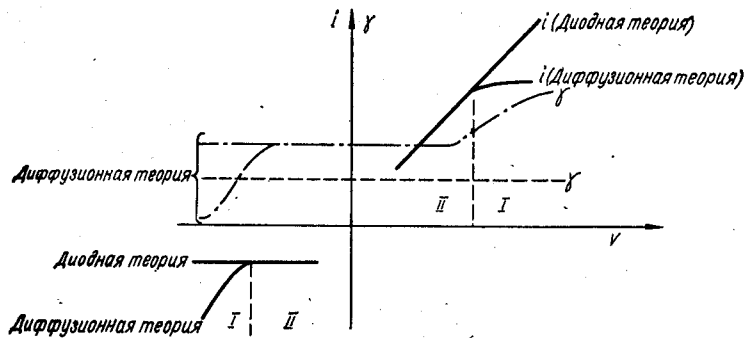


Рис. 3.3. Зависимость тока и коэффициента инжекции при $d = 0$ от напряжения для диодной и диффузионной теорий:

$$I - e^{-\frac{\varphi_0 - eV_1}{kT}} \gg \frac{v_p (l_p + l_s)}{4D_p}; \quad II - e^{-\frac{\varphi_0 - eV_2}{kT}} \ll \frac{\bar{v}_p (l_p + l_s)}{4D_p}.$$

Как видно из последних неравенств, γ может приближаться к единице только для случая, когда $m_p^* \gg m_n^*$. При $m_n^* = m_p^*$ для диодной теории $\gamma \leq 0,5$. Для диффузионной теории при $l_p \gg l_s$ коэффициент инжекции будет больше, чем для диодной, так как $\frac{u_n E_0}{v_p} \ll 1$.

Вольт-амперная характеристика и ход коэффициента инжекции при $d = 0$ для диодной и диффузионной теории показана на рис. 3.3.

В том случае, когда $d \neq 0$, а $V_2 \neq V$ и $V_1 \neq V$, вольт-амперная характеристика и коэффициент инжекции как для диодной, так и для диффузионной теории имеют более сложный вид.

Для диодной теории прямой ток пропорционален $e^{\frac{eV_2}{kT}}$, а обратный $e^{-\frac{eV_1}{kT}}$ при $i_{ns} e^{-\frac{eV_1}{kT}} \gg i_{ps}$. При $i_{ns} e^{-\frac{eV_1}{kT}} \ll i_{ps}$ прямой ток пропорционален $e^{\frac{eV}{kT}}$, а обратный постоянен. Коэффициент инжекции

растет пропорционально $e^{\frac{eV_1}{kT}}$, когда $\gamma \ll 1$, а затем стремится к единице, что видно из рис. 3.4.

Изображенное на этом рисунке увеличение степени роста тока с приложенным напряжением связано с переходом от электронного тока к дырочному. С физической точки зрения это обусловливается тем, что высота барьера для электронного тока из полупроводника в металл изменяется, как eV_2 , а для дырочного тока из металла в полупроводник, как $e(V_1 + V_2)$. Поэтому дырочный ток независимо от толщины барьера

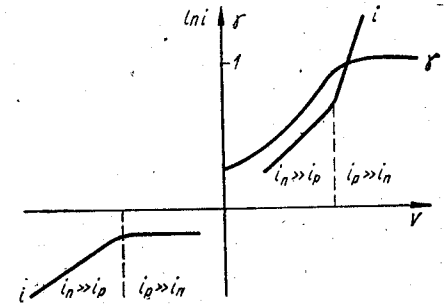


Рис. 3.4. Зависимость тока и коэффициента инжекции от напряжения для диодной теории при $d \neq 0$.

изменяется, как $e^{\frac{eV}{kT}}$.

Для диффузионной теории прямой ток в зависимости от соотношения электронной и дырочной составляющих может быть пропорциональным $e^{\frac{eV_2}{kT}}$, $E_0 e^{\frac{eV_2}{kT}}$, $e^{\frac{e(V_1 + V_2)}{kT}}$ или быть постоянным.

Обратный ток может быть пропорциональным E_0 , $e^{-\frac{eV_1}{kT}}$ или $e^{-\frac{e(V_1 - V_2)}{kT}}$. Коэффициент инжекции для диффузионной теории может изменяться от 0 до 1 в зависимости от параметров полупроводника, контакта и приложенного напряжения.

Зависимость γ от напряжения может иметь убывающий или возрастающий характер. При $V > V_2$ для прямого направления γ растет. При $V_1 < V_2$ γ растет, если $\frac{\bar{v}_p p l_p}{4D_p} e^{\frac{\varphi_0 - eV}{kT}} \gg 1$, и убывает, если $\frac{\bar{v}_p p l_p}{4D_p} e^{\frac{\varphi_0 - eV}{kT}} \ll 1$.

Вольт-амперные характеристики для диффузионной теории при $d \neq 0$ показаны на рис. 3.5, а зависимость коэффициента инжекции от напряжения — на рис. 3.6.

Более сложное изменение дырочного тока, по сравнению с электронным, обусловливается тем, что дырочный ток контролируется не только прохождением носителей заряда через зазор и область пространственного заряда, но также прохождением неос-

новых носителей заряда через квазинейтральную область объема полупроводника.

Для диодной теории при $4D_p \ll \bar{v}_p \mathcal{P}_p l_p$, которое выполняется для достаточно малых толщин зазора

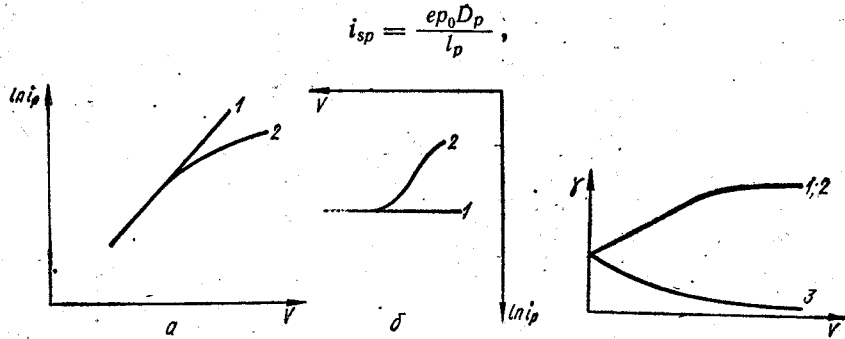


Рис. 3.5. Вольт-амперные характеристики тока неосновных носителей заряда для диффузионной теории; а — прямые токи; б — обратные токи;

$$1 - 4D_p e^{-\frac{\Phi_0 - eV_s}{kT}} \ll \bar{v}_p \mathcal{P}_p (l_p + l_s);$$

$$2 - 4D_p e^{-\frac{\Phi_0 - eV_s}{kT}} \gg \bar{v}_p \mathcal{P}_p (l_p + l_s).$$

а прохождение тока контролируется квазинейтральной областью полупроводника (в этом случае i_{sp} совпадает с аналогичным выражением, полученным в следующей главе для $p-n$ перехода).

При обратном неравенстве $4D_p \gg \bar{v}_p \mathcal{P}_p l_p$, выполняемом при достаточно больших толщинах зазора:

$$i_{sp} = \frac{e \bar{v}_p p_0 \mathcal{P}_p}{4}$$

и ток контролируется прохождением носителей заряда через область пространственного заряда и зазор.

Для диффузионной теории ток контролируется прохождением носителей заряда через квазинейтральную область полупроводника при выполнении неравенства

$4D_p \ll \bar{v}_p \mathcal{P}_p (l_p + l_s) e^{-\frac{\Phi_0 - eV_s}{kT}}$. В этом случае не очень больших толщин зазора

$$i_{sp} = \frac{e p_0 D_p}{(l_p + l_s)}$$

и также практически совпадает с i_{sp} для $p-n$ перехода.

При выполнении обратного неравенства

$$4D_p \gg \bar{v}_p \mathcal{P}_p (l_p + l_s) e^{-\frac{\Phi_0 - eV_s}{kT}},$$

которое справедливо при достаточно толстых зазорах, ток для диффузионной теории контролируется зазором и областью пространственного заряда. При этом

$$i_{sp} = \frac{e \bar{v}_p p_0 \mathcal{P}_p}{4} e^{-\frac{\Phi_0 - eV_s}{kT}}.$$

Как видно, последнее значение i_{sp} больше, чем соответствующее значение i_{sp} для диодной теории и зависит от высоты потенциального барьера и приложенного напряжения. Это обуславливается тем, что переход дырок для диффузионной теории идет вблизи верхушки потенциального барьера при $x = 0$, а для диодной теории у верха валентной зоны при $x = \mathcal{L}$.

3.3. РАСЧЕТ ВОЛЬТ-АМПЕРНОЙ ХАРАКТЕРИСТИКИ С УЧЕТОМ ТУННЕЛИРОВАНИЯ ЧЕРЕЗ ОБЛАСТЬ ПРОСТРАНСТВЕННОГО ЗАРЯДА

В предыдущих параграфах расчет вольт-амперной характеристики проводился в предположении, что носители заряда проходят над барьером области пространственного заряда полупроводника. Это справедливо для широких областей пространственного заряда. Если же область пространственного заряда достаточно узка, то возможно туннелирование носителей заряда также через нее.

Рассмотрение прохождения электронов и дырок через область пространственного заряда полупроводника показывает (см. рис. 3.1), что туннелирование может сказываться на прохождении основных носителей заряда и не сказывается на прохождении неосновных носителей заряда (туннельные токи электронов при разных энергиях показаны стрелками на рис. 3.7). Поэтому расчет вольт-амперных характеристик с учетом туннелирования сводится к нахождению вольт-амперной характеристики только основных носителей заряда — электронов.

При расчетах будем ограничиваться диодной теорией, так как для тонких областей пространственного заряда эта теория реализуется лучше.

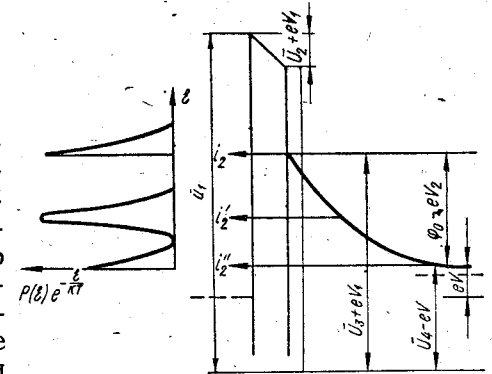


Рис. 3.7. Модель контакта с учетом туннелирования через область пространственного заряда.

Для диодной теории расчет вольт-амперной характеристики сводится к нахождению разности термоэлектронных токов из полупроводника в металл и из металла в полупроводник

$$i = i_2 - i_1.$$

Формула Ричардсона, описывающая термоэлектронную эмиссию, при этом должна быть переписана в более общем виде

$$i = \frac{AT^2}{kT} e^{-\frac{\mu}{kT}} \int_{\xi_0}^{\infty} e^{-\frac{\xi}{kT}} \mathcal{P}(\xi) d\xi, \quad (3.41)$$

где ξ_0 отсчитывается от дна зоны проводимости. В справедливости такой записи легко убедиться, если вынести коэффициент прозрачности $\mathcal{P}(\xi)$ из-под знака интеграла и проинтегрировать

$$\int_{\xi_0}^{\infty} e^{-\frac{\xi}{kT}} d\xi = kTe^{-\frac{\xi_0}{kT}}.$$

Необходимость более общей записи потока в виде (3.41) обусловлена тем, что теперь $\mathcal{P}(\xi)$ является сильной функцией энергии ξ и не может быть вынесена из-под знака интеграла.

Для рассматриваемого барьера при равномерном распределении примесей в ВКБ-приближении можно воспользоваться выражением

$$\mathcal{P}(\xi) = \mathcal{P}_n \exp \left\{ -\beta \left(\sqrt{1+x_\xi} - x_\xi \operatorname{arch} \frac{1}{\sqrt{x_\xi}} \right) \right\}, \quad (3.42)$$

где \mathcal{P}_n — коэффициент прозрачности зазора;

$$\beta = \left[\frac{2m_n^* (\varphi_0 - eV_2)}{\hbar^2} \right]^{1/2} \cdot \mathcal{L}, \quad x_\xi = \frac{\xi - u_4 - eV}{\varphi_0 - eV_2}.$$

Точное вычисление интеграла $\int_{\xi_0}^{\infty} e^{-\frac{\xi}{kT}} \mathcal{P}(\xi) d\xi$ провести нельзя. Поэтому проведем приближенное вычисление интеграла. Для этого проанализируем подынтегральную функцию выражения (3.41). Перепишем ее в следующем виде:

$$\mathcal{P}(\xi) e^{-\frac{\xi}{kT}} = e^{-F(\xi)}, \quad (3.43)$$

где

$$F(\xi) = \frac{\xi}{kT} - \ln \mathcal{P}_0 + \kappa_1 [(\bar{u}_1 - \xi_x)^{3/2} - (\bar{u}_2 - \xi_x)^{3/2}] + \beta \left(\sqrt{1-x_\xi} - x_\xi \operatorname{arch} \frac{1}{\sqrt{x_\xi}} \right). \quad (3.44)$$

Зависимость $F(\xi)$ от ξ при достаточно больших величинах \bar{u}_1 и \bar{u}_2 определяется главным образом первым и четвертым членами выражения (3.44). Анализируя эту зависимость, можно показать, что $F(\xi)$ имеет максимум (см. рис. 3.7). Величина энергии ξ_m , при которой наблюдается этот максимум, может быть получена из условия $\frac{dF(\xi)}{d\xi} = 0$ и определяется выражением

$$\frac{dF(\xi)}{d\xi} = \frac{1}{kT} - \frac{1}{\varphi_0 - eV_2} \operatorname{arch} \frac{1}{\sqrt{x_\xi}} = 0. \quad (3.45)$$

Решая это уравнение относительно ξ , найдем энергию, при которой $F(\xi)$ имеет максимум:

$$\xi_m = \bar{u}_4 + eV + (\varphi_0 - eV_2) x_\xi^m. \quad (3.46)$$

Здесь

$$x_\xi^m = \operatorname{ch}^{-2} \left[\frac{E_{00}}{kT} (1 + a_t) \right],$$

где

$$a_t = -\frac{3}{2} \kappa_1 kT [(\bar{u}_1 - \xi_m)^{1/2} - (\bar{u}_2 + eV_1 - \xi_m)^{1/2}] - kT \frac{d \ln \mathcal{P}_0}{d\xi} \Big|_{\xi_m};$$

$$E_{00} = e \left(\frac{\pi n_0 \hbar^2}{e_2 m_n^*} \right)^{1/2}.$$

Величина x_ξ^m определяет относительное положение максимума

$$x_\xi^m = \frac{\xi_m - u_4 - eV}{\varphi_0 - eV} \quad (3.47)$$

и зависит только от параметров полупроводника и зазора.

Как показывает анализ выражения (3.46), положение максимума заметно отличается от положения вершины барьера области пространственного заряда полупроводника (в этом случае и существенно туннелирование) при $\frac{kT}{E_{00}(1-a_t)}$ меньше нескольких единиц.

Величина E_{00} , которая играет главную роль в рассматриваемом случае, определяется только параметрами полупроводника. Она больше для полупроводников с большей концентрацией примесей n_0 и меньшей эффективной массой туннелирующих носителей заряда. Как показывают расчеты, величина E_{00} равна kT для комнатной температуры при $\frac{n_0}{\xi_2 m_n^*} = 2 \cdot 10^{15} \text{ см}^{-3}$. Это значение и может служить для оценки существенности туннелирования.

Учитывая наличие экстремума выражения (3.44), можно разложить функцию $F(\xi)$ в ряд

$$F(\xi) = F(\xi_m) + F''(\xi_m) (\xi - \xi_m)^2 + \dots$$

Ограничиваясь квадратичным членом, легко провести интегрирование выражения (3.41).

В результате получим вольт-амперную характеристику электронного тока с учетом туннелирования через область пространственного заряда полупроводника в следующем виде:

$$i = i_{se} e^{\frac{eV_2}{kT^*} (1 - \frac{T^*}{T})} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right). \quad (3.48)$$

Здесь

$$i_s = eAT^2 \mathcal{P}_n(V) e^{-\frac{\varphi_b}{kT^*}} = \frac{eV_n n_0}{4} \mathcal{P}_n(V) e^{-\frac{\varphi_0}{kT^*}};$$

$$\mathcal{P}_n(V) = \mathcal{P}_0 \left\{ 1 + \Phi \left[\frac{\operatorname{ch} \frac{E_{00}}{kT} (1 + a_t) \operatorname{cth}^{1/2} E_{00} (1 + a_t)}{E_{00} (\varphi_0 - eV_2)^{1/2}} \right] \right\} \times$$

$$\times \left(\frac{\varphi_0 - eV_2}{1 - \chi_g^m} \right)^{1/2} \frac{8}{e^{3/4} \hbar^{3/2}} \left(\frac{m^* e_2^3}{\pi n_0^3} \right)^{1/4} \left[\frac{E_{00}}{kT} (1 + a_t) \right]^{1/2} \times$$

$$\times \text{th}^{3/2} \left[\frac{E_{00}}{kT} (1 + a_t) \right] \exp \{ -\chi_1 [(\bar{u}_1 - \xi_m)^{3/2} - (\bar{u}_2 + eV_1 - \xi_m)^{3/2}] \};$$

$\Phi(\xi)$ — функция ошибок;

$$\frac{T^*}{T} = \frac{\frac{E_{00}}{kT} \text{cth} \left[\frac{E_{00}}{kT} (1 + a_t) \right]}{1 - 2 \frac{E_{00} a_t}{kT} \text{csh} \left[2 \frac{E_{00}}{kT} (1 + a_t) \right]} \quad (3.49)$$

В случае тесного контакта $d = 0$, $a_t = 0$ и выражение для T^* упрощается:

$$T^* = \frac{E_{00}}{k} \text{cth} \frac{E_{00}}{kT} \quad (3.50)$$

Для реального контакта с зазором величина T^* оказывается меньше, чем для случая $d = 0$ из-за величины a_t .

Когда отсутствует туннелирование через область пространственного заряда, $E_{00} \rightarrow 0$, $T^* \rightarrow T$ и выражение (3.48) переходит в выражение вольт-амперной характеристики (3.4), полученное ранее (следует отметить, что при этом предельном переходе величина i_{st} не совпадает с i_s (3.15) из-за неприменимости в этой точке квадратичного разложения).

В том случае, когда $E_{00} \gg kT$, туннелирование становится существенным. При этом T^* становится большим, чем T . Последнее приводит к более медленному росту прямого тока и более быстрому росту обратного тока с приложенным напряжением, чем без учета туннелирования. Действительно при этом прямой ток без учета

зависимости $\mathcal{P}_n(V)$ растет $i \sim e^{\frac{eV_2}{kT^*}}$, а обратный ток

$$i \sim e^{\frac{eV_2}{k} \left(\frac{1}{T^*} - \frac{1}{T} \right)} e^{-\frac{eV_1}{kT}}$$

Когда $E_{00} \gg kT$, выражения (3.49) и (3.50) переходят в выражение

$$T^* = \frac{E_{00}}{k} \quad (3.51)$$

Вольт-амперная характеристика контакта при этом описывается выражением

$$i = i_{st} e^{\frac{eV}{E_{00}}} \left(1 - e^{-\frac{eV}{kT}} \right), \quad (3.52)$$

где

$$i_{st} = \frac{e v_n n_0}{4} \mathcal{P}_n e^{-\frac{\varphi_0}{E_{00}}}$$

В этом случае обратный ток становится больше прямого и изменяется знак выпрямления.

Рассмотренный здесь переход от пренебрежимо малого туннелирования к очень сильному соответствует переходу от токов над вершиной потенциального барьера области пространственного заряда (ток i_2 на рис. 3.7) к токам у зоны проводимости (ток i_2' на рис. 3.7).

Используя формулы (3.48) и (3.14), можно найти отношение токов с учетом i_t и i без учета туннелирования через область пространственного заряда полупроводника:

$$\ln \frac{i_t}{i} = \ln \frac{i_{st}}{i_s} + \frac{\varphi_0 - eV_2}{kT} \left(1 - \frac{T}{T^*} \right). \quad (3.53)$$

Из последнего выражения видно, что ток с учетом туннелирования через область пространственного заряда полупроводника превышает ток без учета туннелирования в нем в тем большее количество раз, чем больше величина высоты потенциального барьера области пространственного заряда полупроводника при заданном напряжении, чем ниже температура T и больше величина T^* . В предельном случае, когда $T^* \gg T$, максимальное отношение

$$\ln \frac{i_t}{i} = \frac{\varphi_0 - eV_2}{kT} + \ln \frac{i_{st}}{i_s}.$$

На рис. 3.8 представлены расчетные зависимости $\ln \frac{i_{st}}{i_s}$ от $\frac{\varphi_0 - eV_2}{kT}$ при разных значениях E_{00} для значений $a_t = 0$ и $a_t = 0,5$ (для простоты полагалось, что i_{st} не зависит от напряжения). Как видно из этого рисунка, туннелирование при больших значениях $\frac{E_{00}}{kT}$ существенно изменяет вольт-амперную характеристику контакта.

Кроме того, влияние туннелирования уменьшает зависимость вольт-амперной характеристики контакта от температуры. Действительно, i_s при учете туннелирования пропорционально $e^{-\frac{\varphi_0}{kT^*}}$, а $T^* > T$.

3.4. ВЛИЯНИЕ СИЛ ИЗОБРАЖЕНИЯ НА ВОЛЬТ-АМПЕРНУЮ ХАРАКТЕРИСТИКУ КОНТАКТА

Как уже рассматривалось в п. 1.4., учет сил изображения приводит к тому, что ход потенциальной энергии вблизи границ раздела сред изменяется. Причем характер этого изменения зависит от соотношения диэлектрических проницаемостей рассматриваемых сред

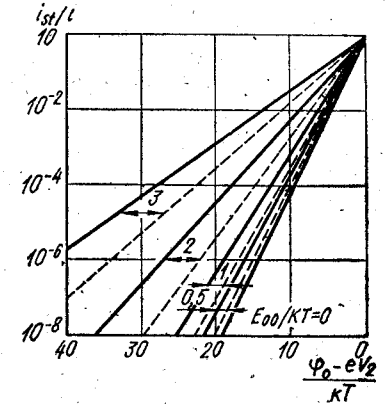


Рис. 3.8. Зависимость $\ln \frac{i_{st}}{i}$ от $\frac{\varphi_0 - eV_2}{kT}$ при $a_t = 0$ (—) и $a_t = 0,5$ (---).

и при $\epsilon_1 > \epsilon_2$ $\varphi^*(x) > 0$, а при $\epsilon_1 < \epsilon_2$ $\varphi^*(x) < 0$. Это показано для контакта металл — полупроводник с зазором на рис. 3.9. Как видно, силы изображения понижают высоту потенциального барьера зазора вблизи металла, могут понижать или повышать ее вблизи полупроводника, а также могут понижать или повышать высоту потенциального барьера области пространственного заряда полупроводника в зависимости от соотношения между величинами ϵ_1 и ϵ_2 .

Наиболее существенное влияние силы изображения оказывают на вольт-амперные характеристики в случае тонких зазоров и тонких областей пространственного заряда. Последнее означает, что наибольшее влияние силы изображения должны оказывать на вольт-амперные характеристики, полученные в предположениях диодной теории. Расчет вольт-амперной характеристики с учетом сил изображения в предположениях диодной теории может быть проведен таким же образом, как это рассмотрено п. 3.1, 3.2. В результате можно получить:

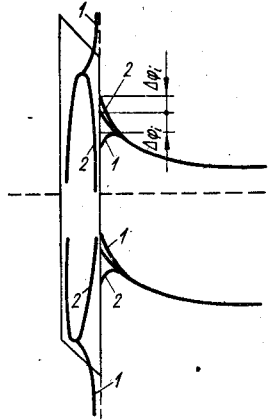


Рис. 3.9. Влияние сил изображения на ход потенциальной энергии:
1 — $\epsilon_2 < \epsilon_3$; 2 — $\epsilon_2 > \epsilon_3$.

$$i_n = \frac{en_0\bar{v}_n}{4} \mathcal{P}_n e^{-\frac{\varphi_0}{kT}} e^{-\frac{\Delta\varphi_i}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{\frac{eV_1}{kT}} \right); \quad (3.54)$$

$$i_p = \frac{ep_0\bar{v}_p D_p \mathcal{P}_p}{4D_p + \bar{v}_p \mathcal{P}_p l_p} \left[e^{\frac{e(V_1+V_2)}{kT}} - 1 \right]. \quad (3.55)$$

Как показывает сравнение этих выражений с полученными ранее без учета сил изображения, в первом выражении дополнитель-

но появляется член $e^{-\frac{\Delta\varphi_i}{kT}}$, а второе по форме совпадает с полученным ранее. Появление члена $e^{-\frac{\Delta\varphi_i}{kT}}$ понятно. Оно связано с изменением высоты потенциального барьера, над которой проходят носители заряда на величину $\Delta\varphi_i$ (см. рис. 3.9). Кроме того, силы изображения изменяют величину коэффициентов прозрачности зазора как для электронов \mathcal{P}_n , так и для дырок \mathcal{P}_p .

Расчет прозрачности зазора с учетом сил изображения может быть проведен только приближенно, если заменить реальный потенциальный барьер на прямоугольный той же площади, используя выражение

$$\bar{\varphi} = \frac{1}{d} \int_0^d \varphi(x) dx, \quad (3.56)$$

где $\varphi(x)$ относится к зазору.

Как показывает анализ при $\epsilon_3 < \epsilon_2$, $\bar{\varphi}$ с учетом сил изображения меньше, чем без учета сил изображения, что приводит к увеличению \mathcal{P}_n и \mathcal{P}_p за счет сил изображения.

Введение величины $\Delta\varphi_i$ приводит к уменьшению электронного тока через контакт при $\epsilon_3 < \epsilon_2$ и увеличению этого тока при $\epsilon_3 > \epsilon_2$.

Величина $\Delta\varphi_i$ может быть найдена аналитически при $\epsilon_3 > \epsilon_2$, если использовать выражение

$$\varphi^*(x) = \frac{(\epsilon_3 - \epsilon_2) e^2}{4\epsilon_2 \epsilon_0 (\epsilon_3 + \epsilon_2) x},$$

которое легко может быть получено из формулы (2.33) при пренебрежении влиянием третьей среды (здесь x отсчитывается от границы раздела полупроводник — зазор или полупроводник — металл). Для этого случая при равномерном распределении примесей можно записать суммарное φ_Σ , когда область пространственного заряда является слоем Шоттки, в виде:

$$\varphi_\Sigma(x) = \frac{e^2 n_0}{2\epsilon_2 \epsilon_0} (\mathcal{L} - x)^2 + \frac{e^2 (\epsilon_3 - \epsilon_2)}{4\epsilon_2 \epsilon_0 (\epsilon_3 + \epsilon_2) x}. \quad (3.57)$$

Такая запись справедлива при условии, что силы изображения не изменяют $\rho(x)$ в области пространственного заряда.

Предположим, что экстремум $\varphi_\Sigma(x)$ наблюдается при $x \ll \mathcal{L}$. Тогда последнее выражение может быть переписано в виде:

$$\varphi_\Sigma(x) = \varphi(0) \left(1 - \frac{2x}{\mathcal{L}} \right) + \frac{e^2 (\epsilon_3 - \epsilon_2)}{4\epsilon_2 \epsilon_0 (\epsilon_3 + \epsilon_2) x}, \quad (3.58)$$

где

$$\varphi(0) = \varphi_0 - eV_2.$$

Продифференцируем выражение (3.58) и приравняем полученный результат к нулю:

$$-\frac{e^2 n_0}{\epsilon_2 \epsilon_0} - \frac{e^2 (\epsilon_3 - \epsilon_2)}{4\epsilon_2 \epsilon_0 (\epsilon_3 + \epsilon_2) x_m^2} = 0. \quad (3.59)$$

Решение последнего уравнения относительно x_m дает

$$x_m = \left[\frac{\epsilon_2 - \epsilon_3}{(\epsilon_3 + \epsilon_2) \mathcal{L} n_0} \right]^{1/2}. \quad (3.60)$$

Оценим значение x_m . Например, при $\varphi_0 - eV_2 = 0,5$ эВ, $n_0 = 5 \cdot 10^{18} \text{ см}^{-3}$, $\epsilon_3 = \infty$ и $\epsilon_2 = 12\epsilon_0$, $\mathcal{L} = 1,2 \cdot 10^{-6} \text{ см}$, $x_m = 6 \times 10^{-8} \text{ см}$, что действительно мало по сравнению с \mathcal{L} . Для меньших концентраций это неравенство будет еще более усиливаться, так как

$$x_m \sim (n_0)^{1/2}, \quad \mathcal{L} \sim (n_0)^{-1/2}.$$

Теперь можно найти $\varphi_\Sigma(x_m)$, если подставить x_m , соответствующее максимуму φ_Σ , в выражение (3.58).

При этом

$$\varphi_\Sigma(x_m) = \varphi(0) \left[1 - \frac{2e^2}{\epsilon_0 \epsilon_2} \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_2} \mathcal{L} n_0 \right)^{1/2} \right]. \quad (3.61)$$

Значение $\Delta\varphi_i = \varphi(0) - \varphi_\Sigma(x_m)$, откуда

$$\Delta\varphi_i = \frac{2e^2}{\epsilon_0 \epsilon_2} \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_2} \mathcal{L} n_0 \right)^{1/2}. \quad (3.62)$$

Это выражение можно подставить в выражение (3.54) и рассчитать вольт-амперную характеристику. Анализ (3.62) показывает, что величина $\Delta\varphi_i$ растет для полупроводников с большими концентрациями и увеличивается с ростом неравенства между ϵ_3 и ϵ_2 . Максимальное значение $\Delta\varphi_i$ будет для тесного контакта металл — полупроводник, когда $\epsilon_3 = \epsilon_m = \infty$.

Величина $\Delta\varphi \sim (\varphi_0 - eV_2)^{1/4}$ и для обратного напряжения, когда $eV_2 \gg \varphi_0$ изменяется с напряжением, как $\Delta\varphi_i \sim V_2^{1/4}$. Оценка величины $\Delta\varphi_i$ для тех же параметров, что и x_m , дает $\Delta\varphi_i = 0,1$ эВ.

При обратном неравенстве между ϵ_3 и ϵ_2 , когда $\epsilon_3 < \epsilon_2$, аналитический расчет $\Delta\varphi_i$ провести нельзя. Для оценок в этом случае можно брать $\varphi_\Sigma(x)$ в выбранной точке x_0 , для которой начинается существенное туннелирование через область пространственного заряда. При этом для $x_0 \ll \mathcal{L}$ легко получить

$$\Delta\varphi_i = -\varphi(0) \frac{2x_0}{\mathcal{L}} + \frac{e^2(\epsilon_3 - \epsilon_2)}{4\epsilon_2\epsilon_0(\epsilon_3 + \epsilon_2)x_0}. \quad (3.63)$$

Как показывает анализ последнего выражения, величина $\Delta\varphi_i$ отрицательна при больших величинах x_0 , проходит через нуль при уменьшении x_0 , а затем становится положительной. Физический смысл имеют только положительные $\Delta\varphi_i$, так как при отрицательных $\Delta\varphi_i$ ток идет над барьером области пространственного заряда и силы изображения не влияют на прохождение носителей заряда через контакт.

Расчет показывает, что для $x_0 = 10^{-7}$ см; $\epsilon_2 = 12$; ϵ_0 и $\epsilon_3 = \infty$ второй член выражения (4.11) дает значение 0,03 эВ. Это значение, по-видимому, близко к максимально возможной величине $\Delta\varphi_i$, и таким образом повышение высоты потенциального барьера за счет сил изображения невелико.

3.5. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ДЛЯ ТОКОВ ЧЕРЕЗ ПОВЕРХНОСТНЫЕ СОСТОЯНИЯ

В общем случае токи через контакт могут идти как без участия, так и с участием поверхностных уровней, как это показано на рис. 3.10.

Рассчитаем вольт-амперную характеристику контакта с учетом токов через поверхностные уровни. Для простоты будем ограничиваться одним дискретным поверхностным уровнем, расположенным на границе раздела полупроводник — зазор, и рассмотрим обмен уровней только с зоной проводимости полупроводника, т. е. ограничимся токами, переносимыми основными носителями заряда.

В этом случае расчет общей вольт-амперной характеристики контакта как для диодной, так и диффузионной теории сводится к нахождению суммы

$$i_n = i_n^b + i_n^s. \quad (3.64)$$

где i_n^b — ток над барьером области пространственного заряда без учета уровней; i_n^s — ток через поверхностные уровни.

Если для диодной теории при подстановке в уравнение (3.64) можно воспользоваться ранее найденным решением для подбарьерного тока i_n^b , то для диффузионной теории нужно учесть, что ток через поверхностные уровни будет влиять на граничные условия, а это приведет к необходимости учета поверхностных уровней на более ранней стадии расчетов.

Рассматривая разность тока из металла в полупроводник и из полупроводника в металл в области вблизи границы раздела полупроводник — зазор так, как это было сделано в п. 3.1, можно получить, что для диффузионной теории с учетом токов через поверхностные уровни

$$n(0) = n_0 e^{\frac{\varphi_0 + eV_1}{kT}} - \frac{4(i_n - i_n^s)}{ev_n \mathcal{L}_n}. \quad (3.65)$$

Подставив $n(0)$ в выражение (3.9), можно показать для диффузионной теории

$$i_n = i_{nr}^b + \frac{4u_n E_2^0}{4u_n E_2^0 + v_n \mathcal{L}_n} i_n^s, \quad (3.66)$$

где i_{nr}^b — ранее найденный ток через контакт без учета поверхностных состояний (см. формулу (3.13)).

Таким образом, как в случае диодной, так и в случае диффузионной теории, расчет вольт-амперных характеристик контакта с учетом токов через поверхностные уровни сводится к дополнительному расчету токов через эти уровни.

Расчет токов через поверхностные состояния. Прежде всего рассмотрим, как собственно появляется ток через поверхностные уровни. Поверхностный уровень всегда обменивается электронами с полупроводником и металлом путем переходов уровень — зона, уровень — металл и обратных переходов.

При отсутствии приложенного к контакту напряжения каждый из переходов сбалансирован, так как в силу принципа детального равновесия противоположные переходы равны друг другу. Когда же к контакту прикладывается внешнее напряжение, этот баланс нарушается и через уровни идет ток.

Можно считать, что ток из зоны проводимости на уровень пропорционален числу свободных мест на уровне $n_i - m_i$ и числу электронов, которые могут попасть на уровень n_n , а обратный ток пропорционален числу электронов на уровне m_i . Ток из уровня в металл пропорционален числу электронов на уровне m_i и числу свободных мест в металле n_v , а обратный ток — числу свободных мест на уровне $n_i - m_i$ и числу электронов в металле n_m . В этом случае можно записать для токов зона проводимости — уровень и

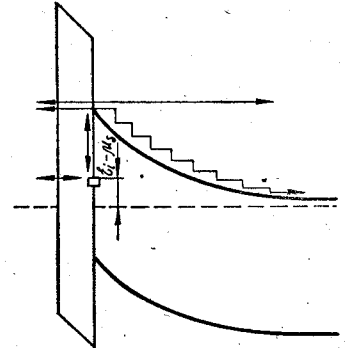


Рис. 3.10. Модель контакта с учетом токов через поверхностный уровень.

уровень — металл:

$$i_n^{ns} = e [K_{ni} (n_i - m_i) n_k - K_{in} m_i]; \quad (3.67)$$

$$i_n^{ms} = e [K_{mi} (n_i - m_i) n_m - K_{im} m_i n_v]. \quad (3.68)$$

Здесь n_i — концентрация поверхностного уровня; m_i — концентрация электронов на нем; K_{ni} и K_{mi} — кинетические коэффициенты, характеризующие переход на уровень из зоны проводимости и металла; K_{in} и K_{im} — обратные переходы; n_k — концентрация электронов в области полупроводника, с которой осуществляется электронный обмен; n_m и n_v — концентрация электронов и свободных мест в металле и областях, с которыми осуществляются электронные переходы.

В стационарном случае $i_n^{ns} = i_n^{ms} = i_n^s$ и можно найти

$$m_i = \frac{(K_{ni} n_k + K_{mi} n_m) n_i}{K_{ni} n_k + K_{in} + K_{mi} n_m + K_{im} n_v}. \quad (3.69)$$

Если подставить это значение m_i в формулу (3.67) или (3.68), легко получить следующее выражение для токов через поверхностные уровни:

$$i_n^s = \frac{(K_{im} K_{ni} n_k n_v - K_{mi} K_{in} n_m) n_i}{K_{mi} n_m + K_{im} n_v + K_{ni} n_k + K_{in}}. \quad (3.70)$$

В последнее выражение входят семь величин, требующих физической расшифровки: K_{im} , K_{mi} , K_{ni} , K_{in} , n_v , n_m , n_k .

Используя равенство нулю токов i_n^{ns} и i_n^{ms} при отсутствии приложенного напряжения, можно получить следующие соотношения между коэффициентами K_{ni} и K_{in} , а также K_{mi} и K_{im} :

$$\begin{aligned} K_{ni}^0 (n_i - m_i^0) n_k^0 &= K_{in}^0 m_i^0; \\ K_{mi}^0 (n_i - m_i^0) n_m^0 &= K_{im}^0 m_i^0 n_v^0, \end{aligned} \quad (3.71)$$

где значки «0» означают, что данные величины относятся к случаю $V = 0$.

Эти соотношения позволяют уменьшить число параметров задачи до пяти: K_{ni} , K_{mi} , n_m^0 , n_v^0 и n_k , если использовать соотношения:

$$\begin{aligned} K_{in}^0 &= K_{ni}^0 M_i n_k^0, \\ K_{im}^0 &= K_{mi}^0 M_i \frac{n_m^0}{n_v^0}, \end{aligned} \quad (3.72)$$

где

$$M_i = \left(\frac{n_i}{m_i^0} - 1 \right) = e^{\frac{\varepsilon_i - \mu_{s0}}{kT}},$$

и учесть, что K_{ni} и K_{mi} существенно не зависят от приложенного напряжения; $K_{in}^0 = K_{in}$, $K_{ni}^0 = K_{ni}$, $K_{mi}^0 = K_{mi}$, $K_{im}^0 = K_{im}$.

Параметр K_{ni} может быть найден, если рассмотреть поток электронов, падающих на поверхностные уровни и захватываемых ими, в виде

$$K_{ni} = \bar{v}_n \bar{c}_n, \quad (3.73)$$

где \bar{v}_n и \bar{c}_n — средняя тепловая скорость и сечение захвата электронов зоны проводимости на поверхностный уровень.

Параметр K_{mi} может быть рассмотрен аналогичным образом. Тогда

$$K_{mi} = \bar{v}_n^m \bar{c}_n^m, \quad (3.74)$$

где \bar{v}_n^m — скорость электронов металла, попадающих на уровень; \bar{c}_n^m — среднее сечение захвата электронов металла поверхностным уровнем, которое зависит от наличия и свойств зазора.

Концентрация электронов в металле, способных переходить на поверхностные уровни:

$$n_m = \int_{\xi} N(\xi) f(\xi) d\xi. \quad (3.75)$$

Концентрация свободных мест в металле, куда идут электронные переходы из уровня:

$$n_v = \int_{\xi} N(\xi) [1 - f(\xi)] d\xi, \quad (3.76)$$

где $N(\xi) = \frac{4\pi (2m_m^*)^{3/2}}{\hbar} \xi^{1/2}$ — плотность состояний в металле;

$f(\xi) = (1 + e^{\frac{\varepsilon - \mu}{kT}})^{-1}$ — функция их заполнения.

Интегрирование записанных выражений должно проводиться по всем энергиям, соответствующим возможным переходам.

Если электронные переходы из уровней в металл и обратные переходы идут без изменения энергии, то в соответствующие интегралы можно ввести пределы $\xi_{im} - kT$ и $\xi_{im} + kT$, где ξ_i — положения уровня, отсчитанные от дна зоны проводимости металла. Теперь можно заменить интегрирование по ξ умножением подынтегральной функции, взятой в точке ξ_{im} , на $2kT$. В результате получим:

$$n_m = \frac{4\pi (2m_m^*)^{3/2}}{\hbar^3} \int_{\xi} \frac{\xi^2 d\xi}{1 + e^{\frac{\varepsilon - \mu}{kT}}} = \frac{Q}{1 + M_i e^{\frac{eV_1}{kT}}}; \quad (3.77)$$

$$n_v = \frac{4\pi (2m_m^*)^{3/2}}{\hbar^3} \int_{\xi} \frac{\xi^{1/2} d\xi}{1 + e^{\frac{\mu - \varepsilon}{kT}}} = \frac{Q}{1 + M_i e^{-\frac{eV_1}{kT}}}. \quad (3.78)$$

Здесь $Q = \frac{8\pi kT (2m_m^*)^{3/2}}{\hbar^3} (\xi_{im})^{1/2}$; ξ_{im} отсчитывается от дна зоны проводимости металла.

Подставляя найденные параметры в выражение (3.70), легко получить как для диодной, так и для диффузионной теории, что

$$i_n^s = \frac{\bar{v}_n \bar{c}_n n_0 M_i n_i e^{-\frac{\varphi_0}{kT}} \left(\frac{n_k}{n_k^0} e^{\frac{eV}{kT}} - 1 \right)}{(1 + M_i e^{\frac{eV_1}{kT}}) \left[1 + K \left(\frac{n_k}{n_k^0} + M_i \right) \right]}, \quad (3.79)$$

где

$$K = \frac{\bar{v}_n \bar{c}_n n_0}{\bar{v}_n^m \bar{c}_n^m Q} e^{-\frac{\varphi_0}{kT}}.$$

Как видно из выражения (3.79), ток через поверхностные состояния пропорционален их концентрации n_i и зависит от их энергетического положения M_i , сечений захвата электронов зоны проводимости \bar{c}_n , металла \bar{c}_n^m , высоты потенциального барьера φ_0 , концентрации носителей заряда в полупроводнике n_0 , приложенного напряжения V , а также концентрации электронов в области полупроводника, с которой идет электронный обмен.

Для диодной теории

$$n_k = n_0 e^{\frac{\varphi_0 - eV_s}{kT}} = n_k^0 e^{\frac{eV_s}{kT}},$$

что соответствует концентрации электронов в точке $x = \mathcal{L}$, способных преодолеть потенциальный барьер области пространственного заряда. Эта концентрация равна $n_k^0 = n_0 e^{-\frac{\varphi_0}{kT}}$ при $V = 0$ и увеличивается при уменьшении потенциального барьера.

Для диффузионной теории, когда электронные переходы происходят в области вблизи $x = 0$, величину $n_k = n(0)$ можно получить, если подставить в выражение (3.65) суммарный ток $i_n = i_n^b + i_n^s$ и использовать рассчитанное ранее значение i_n^b (см. (3.13)).

В результате имеем

$$n(0) = \frac{n_0 e^{-\frac{\varphi_0}{kT}}}{(4E_2 \mu_n + \bar{v}_n \mathcal{P}_n)} \left(\bar{v}_n \mathcal{P}_n e^{-\frac{eV_1}{kT}} + 4E_0 \mu_n e^{-\frac{eV_2}{kT}} - \frac{4i_{ns}}{en_0 e^{-\frac{\varphi_0}{kT}}} \right). \quad (3.80)$$

Для диффузионной теории величина $n(0) = n_0 e^{-\frac{\varphi_0}{kT}}$ при $V = 0$. Но в отличие от диодной теории она может как увеличиваться, так и уменьшаться при уменьшении высоты потенциального барьера при $v \neq 0$. Это связано с тем, что граничная концентрация может определяться как электронным обменом области полупроводника вблизи $x = 0$ с объемом полупроводника, так и с металлом или поверхностными уровнями. В первом случае концентрация $n(0)$, как и для диодной теории, растет с уменьшением высоты потенциального барьера области пространственного заряда. Во втором

она уменьшается, поскольку при тех же условиях растет высота потенциального барьера для электронов, идущих в полупроводник из металла.

Записанные выше выражения для n_k и $n(0)$ позволяют построить вольт-амперные характеристики с учетом токов через поверхностный уровень как в предположениях диодной, так и диффузионной теорий. Однако для общего анализа полученные выражения слишком сложны. Поэтому проведем анализ полученных выражений для ряда частных случаев.

Анализ токов через поверхностные состояния. Для диодной теории после подстановки n_k в выражение (3.79) имеем

$$i_n^s = \frac{\bar{v}_n \bar{c}_n n_i M_i n_0 e^{-\frac{\varphi_0}{kT}} \left(e^{\frac{eV}{kT}} - 1 \right)}{(1 + M_i e^{\frac{eV_1}{kT}}) \left[1 + K \left(e^{\frac{eV_2}{kT}} + M_i \right) \right]}. \quad (3.81)$$

Пусть главную роль играет электронный обмен уровней с металлом. Это означает, что интенсивность электронных переходов между поверхностным уровнем и металлом значительно больше, чем между уровнем и полупроводником. Тогда, учитывая, что $c_n^m \gg c_n$, можно получить $K \rightarrow 0$. Выражение (3.81) при этом переходит в следующее:

$$i_n^s = \frac{\bar{v}_n \bar{c}_n n_i M_i n_0 e^{-\frac{\varphi_0}{kT}} \left(e^{\frac{eV}{kT}} - 1 \right)}{1 + M_i e^{\frac{eV_1}{kT}}}. \quad (3.82)$$

В рассматриваемом случае ток через поверхностные состояния пропорционален сечению захвата электрона зоны проводимости поверхностным уровнем \bar{c}_n и не зависит от сечений захвата электронов металла. Последнее легко понять, если учесть, что при последовательном прохождении электронов из полупроводника в металл и наоборот с участием пересадочного уровня при преимущественном обмене с металлом, лимитирующей является стадия переноса электрона на этапе полупроводник — поверхностный уровень. Обмен уровней с металлом в рассматриваемом случае определяет заполнение поверхностного уровня.

В зависимости от энергетического положения уровня, которое определяет величину M_i , изменяется величина тока через уровень и характер зависимости тока от напряжения.

При преимущественном обмене уровня с металлом величина тока вначале, когда $1 \gg M_i e^{\frac{eV_1}{kT}}$, растет с ростом M_i , а затем, когда $1 \ll M_i e^{\frac{eV_1}{kT}}$, выходит на насыщение.

Если выполняется неравенство $1 \gg M_i e^{\frac{eV_1}{kT}}$, прямой ток через поверхностный уровень пропорционален $e^{\frac{eV}{kT}}$, а обратный испытывает насыщение. В этом случае, когда уровни расположены значи-

тельно ниже уровня Ферми, зазор не проявляется и вольт-амперная характеристика по общему виду совпадает с вольт-амперной характеристикой для токов без участия поверхностных состояний при $d = 0$ (см. п. 3.1).

Если выполняется неравенство $1 \ll M_i e^{\frac{eV_1}{kT}}$, прямой ток через поверхностные уровни пропорционален $e^{\frac{eV_2}{kT}}$, а обратный пропорционален $e^{-\frac{eV_1}{kT}}$. В этом случае вольт-амперная характеристика по виду

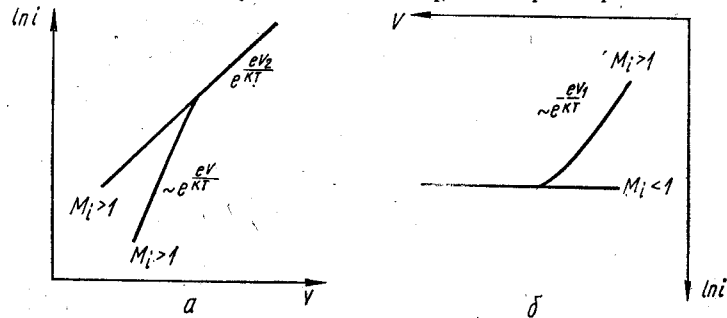


Рис. 3.11. Вольт-амперные характеристики для токов с участием поверхностных уровней при преимущественном обмене уровней с металлом: а — прямой ток; б — обратный ток.

совпадает с вольт-амперной характеристикой для токов без участия поверхностных уровней с учетом зазора (см. п. 3.1).

Возможный характер прямой и обратной ветвей вольт-амперной характеристики при преимущественном обмене с металлом показан на рис. 3.11.

Пусть главную роль играет электронный обмен поверхностного уровня с полупроводником. Это означает, что интенсивность электронных переходов между поверхностным уровнем и полупроводником значительно больше, чем между уровнем и металлом. Тогда, учитывая, что $c_n^m \gg c_n^m$, можно положить $K \rightarrow \infty$. Выражение (3.81) при этом переходит в следующее:

$$i_n^s = \frac{e \bar{v}_n^m c_n^m n_i M_i Q (e^{\frac{eV}{kT}} - 1)}{(1 + M_i e^{\frac{eV_1}{kT}}) (e^{\frac{eV_2}{kT}} + M_i)} \quad (3.83)$$

В рассматриваемом случае ток через поверхностные состояния пропорционален сечению захвата поверхностным уровнем электронов металла и не зависит от сечений захвата электронов зоны проводимости. Последнее определяется тем, что при последовательном прохождении электронов из полупроводника в металл и наоборот с участием пересадочного уровня при преимущественном обмене с полупроводником лимитирующей является стадия переноса электрона на этапе металл — поверхностный уровень. Обмен уровней с полупроводником определяет заполнение поверхностного уровня.

В зависимости от величины M_i , зависящей от энергетического положения уровня, ток через поверхностный уровень может вначале при $1 \gg M_i e^{\frac{eV_1}{kT}}$ и $1 \gg M_i e^{-\frac{eV_2}{kT}}$ расти, а затем при $1 \ll M_i e^{\frac{eV_1}{kT}}$ и $1 \ll e^{-\frac{eV_2}{kT}}$ уменьшаться.

Для разных величин M_i наблюдается изменение характера зависимости тока от напряжения.

Так, при $1 \gg M_i e^{\frac{eV_1}{kT}}$ и $1 \gg M_i e^{-\frac{eV_2}{kT}}$ прямой ток пропорционален $e^{\frac{eV_2}{kT}}$, а обратный пропорционален $e^{-\frac{eV_1}{kT}}$.

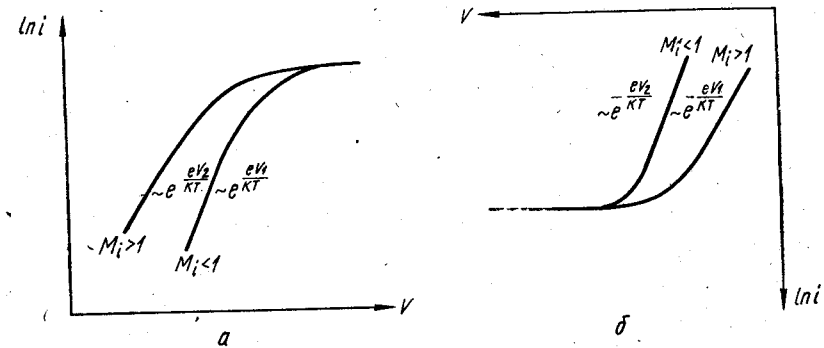


Рис. 3.12. Вольт-амперные характеристики для токов с участием поверхностного уровня при преимущественном электронном обмене с полупроводником: а — прямой ток; б — обратный ток.

Вольт-амперная характеристика при этом совпадает по виду с характеристикой контакта без учета токов через поверхностные уровни при $d \neq 0$ (см. п. 3.1), но с взаимной заменой V_1 на V_2 и другим значением i_s (особо следует обратить внимание на то, что величина i_s для токов через поверхностные уровни при преимущественном обмене с полупроводником не зависит от величины Φ_0).

При обратных неравенствах $1 \ll M_i e^{\frac{eV_1}{kT}}$ и $1 \ll M_i e^{-\frac{eV_2}{kT}}$ прямой ток пропорционален $e^{\frac{eV_2}{kT}}$, а обратный пропорционален $e^{-\frac{eV_1}{kT}}$.

При этом вольт-амперная характеристика по виду совпадает с вольт-амперной характеристикой без учета токов через поверхностные уровни при $d \neq 0$ (см. п. 3.1), но с другим значением величины i_s .

Возможный характер прямой и обратной ветви вольт-амперной характеристики при преимущественном обмене с полупроводником показан на рис. 3.12.

Пусть для диффузионной теории величина $n(0)$ не зависит от i_n^s и последним членом в выражении (3.80) можно пренебречь. Это означает, что ток через поверхностные уровни достаточно мал, чтобы влиять на величину $n(0)$.

Кроме того, будем предполагать, что выполняются два неравенства $4E_0\mu_n \gg \bar{v}_n \mathcal{P}_n$ и $4E_0\mu_n \gg \frac{\bar{v}_n \mathcal{P}_n e^{-\frac{eV}{kT}}}{\frac{\Phi_0 - eV}{kT}}$. Тогда, как видно из выражения (3.80), $n(0) = n_0 e^{-\frac{\Phi_0 - eV}{kT}}$ и совпадает с n_k в предположениях диодной теории. Поэтому совпадает и выражение для i_n^s в предположениях диффузионной теории с выражением для i_n^s в предположениях диодной теории.

В связи с этим для рассматриваемого случая диффузионной теории можно воспользоваться проведенным выше анализом тока через поверхностный уровень, полученным в предположениях диодной теории.

При выполнении обратных неравенств $4E_0\mu_n \ll \bar{v}_n \mathcal{P}_p$ и $4E_0\mu_n \ll \frac{\bar{v}_n \mathcal{P}_p e^{-\frac{eV}{kT}}}{\frac{\Phi_0 + eV_1}{kT}}$ выражение (3.80) дает $n(0) = n_0 e^{-\frac{\Phi_0 + eV_1}{kT}}$. Подставляя полученное значение $n(0)$ в (3.79), имеем:

$$i_n^s = \frac{\bar{v}_n c_n n_i M_i n_0 e^{-\frac{\Phi_0}{kT}} \left(e^{-\frac{eV_1}{kT}} e^{\frac{eV_1}{kT}} - 1 \right)}{(1 + M_i e^{\frac{eV_1}{kT}}) [1 + K \left(e^{-\frac{eV_1}{kT}} + M_i \right)]} = 0.$$

В рассматриваемом случае ток через поверхностные уровни не идет.

Сравнение двух последних рассмотренных случаев показывает, что ток через поверхностные уровни в предположениях диффузионной теории чувствителен к тому, что определяет электронный обмен области пространственного заряда в окрестности $x = 0$ полупроводник или металл.

Теперь пусть для диффузионной теории можно пренебречь влиянием на $n(0)$ тока без участия поверхностных уровней и $i_n = i_n^s$.

Тогда из (3.65) видно, что $n(0) = n_0 e^{-\frac{\Phi_0 - eV}{kT}}$; как показано выше, такое значение $n(0)$ не позволяет получить ток через поверхностные уровни (см. (3.84)).

Различие в поведении токов через поверхностные уровни для диодной и диффузионной теории определяется тем, что для диодной теории при приложении напряжения всегда можно «разбалансировать» равенства (3.67) или (3.68), а для диффузионной теории изменения с напряжением различных величин, входящих в (3.67) или (3.68), могут компенсировать друг друга и такой «разбаланс» невозможен.

В заключение проведем сравнение токов без участия и с участием поверхностных уровней для тех случаев, когда они существенны.

Для этого рассмотрим ситуацию, когда $V \rightarrow 0$ и $e^{-\frac{eV_1}{kT}}$, $e^{-\frac{eV_2}{kT}}$, $e^{-\frac{eV}{kT}}$ можно разложить в ряд и ограничиться первым или вторым членом разложения.

Тогда выражение (3.81) переходит в следующее:

$$i_{ns} = \frac{e}{kT} \frac{e v_n c_n n_i M_i n_0 e^{-\frac{\Phi_0}{kT}} V}{(1 + M_i) [1 + K(1 + M_i)]}. \quad (3.85)$$

Выражение (3.23) в том же приближении имеет вид

$$i_n^b = \frac{e \bar{v}_n n_0 \mathcal{P}_n e^{-\frac{\Phi_0}{kT}}}{4} \frac{e}{kT} V. \quad (3.86)$$

Сравнивая выражения (3.85) и (3.86), видим, что токи через поверхностные уровни будут преобладать при

$$\mathcal{P}_n > \frac{4 c_n n_i M_i}{(1 + M_i) [1 + K(1 + M_i)]}. \quad (3.87)$$

При обратном неравенстве преобладает ток без участия поверхностных уровней.

Для оценок можно взять $n_i = 10^{12} \text{ см}^{-2}$ и $\bar{c}_n = 10^{-15} \text{ см}^{-2}$. Тогда при $K = 0$ и $M_i \gg 1$ ток через уровни будет преобладать при $\mathcal{P}_n < 4 \cdot 10^{-3}$. При обратном неравенстве током через уровни можно пренебречь.

3.6. ПАРАМЕТРЫ ВОЛЬТ-АМПЕРНОЙ ХАРАКТЕРИСТИКИ И ИХ АНАЛИЗ

В общем виде вольт-амперная характеристика контакта с учетом всех рассмотренных выше эффектов описывается выражением

$$i = i_s \left(e^{\frac{eV_s}{kT}} - e^{-\frac{eV_1}{kT}} \right), \quad (3.88)$$

где величина i_s , являющаяся обобщенным параметром вольт-амперной характеристики, для различных частных случаев может быть записана в явном виде (см. п. 3.1—3.5) и может являться слабой или сильной функцией приложенного напряжения.

Так, для диодной и диффузионной теории для токов основных носителей заряда без участия поверхностных состояний, а также для токов через поверхностные уровни, расположенные достаточно высоко над уровнем Ферми, параметр i_s , как будет подробнее проанализировано дальше, слабо зависит от приложенного напряжения. Для случая, когда существенно туннелирование основных носителей заряда через область пространственного заряда полупроводника, а также для токов через поверхностные уровни, расположенные ниже уровня Ферми, i_s может сильно зависеть от напряжения.

Далее будем учитывать, что для практического использования контактов металл-полупроводник с барьером Шоттки чрезвычайно важным является степень роста тока с приложенным прямым напряжением, которая может в полулогарифмическом масштабе

характеризоваться величиной

$$\alpha = \frac{d \ln i}{dV} = \frac{e}{nkT}. \quad (3.89)$$

В такой записи отражены обозначения, применяемые в различных литературных источниках (величина $n = \frac{e}{kT\alpha}$, применяемая обычно в зарубежной литературе, может быть легко найдена при известном параметре α). Поскольку в общем случае для обратного направления также наблюдается существенный рост тока с приложенным напряжением, можно аналогичным образом ввести величину α^* , характеризующую обратную ветвь вольт-амперной характеристики контакта.

В результате, используя выражение (3.88), для прямого направления при $e \frac{eV}{kT} \gg 1$ можно записать

$$\alpha = \frac{e}{kT} \cdot \frac{dV_2}{dV} + \frac{d \ln i_s}{dV}, \quad (3.90a)$$

для обратного при $e \frac{eV}{kT} \ll 1$

$$\alpha^* = \frac{e}{kT} \cdot \frac{dV_1}{dV} - \frac{d \ln i_s^*}{dV}. \quad (3.90b)$$

Таким образом, нахождение величин α и α^* , в первую очередь, сводится к нахождению производных $\frac{dV_2}{dV}$ и $\frac{dV_1}{dV}$. Из выражений (3.90a) и (3.90b) следует, что при $\frac{d \ln i_s}{dV} = -\frac{d \ln i_s^*}{dV}$ между параметром α и α^* существует следующая связь:

$$\alpha^* = \frac{e}{kT} - \alpha. \quad (3.91)$$

Используя выражение (2.41), можно показать, что

$$\frac{dE_3}{dV} = \frac{\epsilon_3}{\epsilon_3} \frac{dE_2}{dV} + \frac{1}{\epsilon_0 \epsilon_3}, \quad \frac{d\sigma_2}{dV},$$

где для контакта металл — полупроводник $V = V_1 + V_2$.

Учитывая соотношения:

$$E_3 = -\frac{V_1^0 - V_1}{d}; \quad \frac{dE_2}{dV} = -\frac{eE_2}{2(\varphi_0 - eV_2)} \cdot \frac{dV_2}{dV} = \frac{1}{\mathcal{L}} \cdot \frac{dV_2}{dV};$$

$$\sigma_2 = em_i; \quad 1 = \frac{dV_1}{dV} + \frac{dV_2}{dV},$$

легко получить:

$$\frac{dV_2}{dV} = \left(1 - \frac{de}{\epsilon_0 \epsilon_1} \cdot \frac{dm_i}{dV}\right) \left(1 + \frac{\epsilon_2}{\epsilon_1} \cdot \frac{d}{\mathcal{L}}\right)^{-1}; \quad (3.92)$$

$$\frac{dV_1}{dV} = \left(\frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} + \frac{ed}{\epsilon_0 \epsilon_1} \cdot \frac{dm_i}{dV}\right) \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}}\right)^{-1}. \quad (3.93)$$

Параметры α и α^* при пренебрежении зависимостью i_s от напряжения. В том случае, когда i_s не зависит от приложенного к контакту напряжения, анализ параметров α и α^* сводится к анализу $\frac{dV_2}{dV}$ и $\frac{dV_1}{dV}$, так как при этом:

$$\alpha \equiv \alpha_0 = \frac{e}{kT} \left(1 - \frac{ed}{\epsilon_0 \epsilon_1} \cdot \frac{dm_i}{dV}\right) \left(1 + \frac{\epsilon_2}{\epsilon_1} \frac{d}{\mathcal{L}}\right)^{-1}; \quad (3.94)$$

$$\alpha^* \equiv \alpha_0^* = \frac{e}{kT} \left(\frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} + \frac{ed}{\epsilon_0 \epsilon_1} \cdot \frac{dm_i}{dV}\right) \left(1 + \frac{\epsilon_2}{\epsilon_1} \frac{d}{\mathcal{L}}\right)^{-1}. \quad (3.95)$$

В частном случае, когда $\frac{dm_i}{dV} = 0$:

$$\alpha_0 \equiv \alpha_0^0 = \frac{e}{kT} \left(1 + \frac{\epsilon_2}{\epsilon_1} \cdot \frac{d}{\mathcal{L}}\right)^{-1}; \quad (3.96a)$$

$$\alpha_0^* \equiv \alpha_0^{0*} = \frac{e}{kT} \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}}\right)^{-1}. \quad (3.96b)$$

Как видно из формулы (3.96a), для контакта металл — полупроводник с зазором α_0^0 всегда меньше $\frac{e}{kT}$ и зависит от параметров полупроводника, зазора и приложенного напряжения. Анализ выражений (3.96a) и (3.96b) показывает, что параметр α_0^0 уменьшается с ростом d , отношения $\frac{\epsilon_2}{\epsilon_1}$ и концентрации носителей заряда в полупроводнике n_0 . Кроме того, параметр α_0^0 больше для больших значений φ_0 и уменьшается с ростом прямого напряжения. Величина α_0^{0*} также зависит от перечисленных выше параметров, но ведет себя с их изменением противоположным образом ($\alpha_0^{0*} = \frac{e}{kT} - \alpha_0^0$).

Влияние на параметры α и α^* величины $\frac{dm_i}{dV}$ может быть проанализировано при использовании результатов предыдущего параграфа. Так, используя результаты предыдущего параграфа, можно показать, что при преимущественном обмене поверхностных уровней с полупроводником, выражение (3.69) переходит в следующее:

$$m_i = \frac{n_i}{1 + \frac{K_{in}}{K_{ni} n_k}}. \quad (3.97)$$

Подставляя полученные выше значения K_{in} , K_{ni} и n_k , можно показать, что для поверхностного уровня при преимущественном обмене с полупроводником в предположениях диодной теории, а также диффузионной при $i_n^s \rightarrow 0$, $4u_n E_2^0 \gg \bar{v}_n \mathcal{P}_n$ и $4u_n E_2^0 \gg \bar{v}_n \mathcal{P}_n e^{-\frac{eV_2}{kT}}$

$$m_i = \frac{n_i}{1 + M_i e^{-\frac{eV_2}{kT}}}. \quad (3.98)$$

Для поверхностной полузаполненной зоны в этом случае можно показать, что

$$m_i = N_i (\mu_{s0} + eV_2). \quad (3.99)$$

При преимущественном обмене поверхностных уровней с металлом на основании результатов предыдущего параграфа имеем, что выражение (3.69) сводится к такому:

$$m_i = \frac{n_i}{1 + \frac{K_{im} n_v}{K_{mi} n_m}}. \quad (3.100)$$

После подстановки в это выражение K_{im} , K_{mi} , n_v и n_m получаем для дискретного уровня

$$m_i = \frac{n_i}{1 + M_i e^{\frac{eV_1}{kT}}}. \quad (3.101)$$

Аналогичный результат также можно получить для преимущественного обмена поверхностного уровня с полупроводником в предположении диффузионной теории, если $i_n^s = i_n$ или $i_n^s \rightarrow 0$, а $4u_n E_2^0 \ll \bar{v}_n \mathcal{P}_n$ и $4u_n E_2^0 \ll \bar{v}_n \mathcal{P}_n e^{\frac{eV}{kT}}$. В дальнейшем для простоты этот случай выделяться отдельно не будет.

Для поверхностной полузаполненной зоны с постоянной плотностью концентрации N_i при преимущественном обмене с металлом можно получить

$$m_i = N_i (\mu_{s0} - eV_1). \quad (3.102)$$

Используя полученные значения m_i , можно показать, что в случае преимущественного обмена поверхностных состояний с полупроводником:

для дискретного уровня

$$\frac{dm_i}{dV} = \frac{e}{kT} \frac{dV_2}{dV} M_i e^{\frac{eV_2}{kT}} n_i \left(1 + M_i e^{\frac{eV_2}{kT}}\right)^{-2}; \quad (3.103)$$

для полузаполненной зоны поверхностных состояний с постоянной плотностью концентрации N_i

$$\frac{dm_i}{dV} = e N_i^0 \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} N_i^0\right)^{-1}. \quad (3.104)$$

В случае преимущественного обмена поверхностных состояний с металлом:

для дискретного уровня

$$\frac{dm_i}{dV} = -\frac{e}{kT} \cdot \frac{dV_1}{dV} M_i e^{\frac{eV_1}{kT}} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-2}; \quad (3.105)$$

для полузаполненной зоны поверхностных состояний с постоянной плотностью концентрации

$$\frac{dm_i}{dV} = -e N_i^0 \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} N_i^0\right)^{-1}; \quad (3.106)$$

Подставляя соответствующие значения $\frac{dm_i}{dV}$ в выражение для α_0 , можно получить в случае преимущественного обмена поверхностных уровней с полупроводником: для дискретного уровня

$$\alpha_0 = \frac{e}{kT} \left[1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} \cdot \frac{n_i}{kT} M_i e^{-\frac{eV_2}{kT}} \left(1 + M_i e^{-\frac{eV_2}{kT}}\right)^{-2}\right]^{-1}; \quad (3.107)$$

для полузаполненной зоны

$$\alpha_0 = \frac{e}{kT} \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} N_i^0\right)^{-1}. \quad (3.108)$$

В случае преимущественного обмена с металлом выражения имеют вид:

для дискретного уровня

$$\alpha_0 = \frac{e}{kT} \left[1 + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} \frac{n_i}{kT} M_i e^{\frac{eV_1}{kT}} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-2}\right] \left[1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{e^2 d}{\varepsilon_0 \varepsilon_1} \cdot \frac{n_i}{kT} \times \right. \\ \left. \times M_i e^{\frac{eV_1}{kT}} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-2}\right]^{-1}; \quad (3.109)$$

для полузаполненной зоны поверхностных состояний

$$\alpha_0 = \frac{e}{kT} \left(1 + \frac{de^2}{\varepsilon_0 \varepsilon_1} N_i^0\right) \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} + \frac{de^2}{\varepsilon_0 \varepsilon_1} N_i^0\right)^{-1}. \quad (3.110)$$

Анализ полученных выражений показывает, что для преимущественного обмена с полупроводником наличие поверхностных состояний всегда приводит к уменьшению величины α_0 , а для преимущественного обмена с металлом — к увеличению α_0 . В первом случае при устремлении n_i или N_i^0 к бесконечности $\alpha_0 \rightarrow 0$, а во втором $\alpha_0 \rightarrow \frac{e}{kT}$. Это указывает путь повышения величины α_0 до значения $\frac{e}{kT}$ в случае реального контакта с зазором между металлом и полупроводником.

Следует также отметить, что для дискретного уровня в случае преимущественного обмена с полупроводником максимальное влияние поверхностных состояний на величину α будет наблюдаться при $\exp \frac{E_i - \mu_{s0} - eV_2}{kT} = 1$, что соответствует совпадению уровня Ферми с положением дискретного уровня (при этом член выражения (3.107), зависящий от n_i , равен $\frac{de^2 n_i}{\varepsilon_0 \varepsilon_1 kT}$). При больших и меньших напряжениях влияние дискретного уровня на вольт-

амперную характеристику ослабевает и, в принципе, может оказаться сколь угодно малым. Это может приводить к немонотонному изменению параметров α и α^* с напряжением.

Рассмотрение параметров α и α^* с учетом зависимости i_s от напряжения. Для учета зависимости i_s от напряжения дополнительно проанализируем величину $\frac{d \ln i_s}{dV}$ и для различных частных случаев.

В случае, когда токи через поверхностные уровни не играют роли, а также можно пренебречь туннелированием через область пространственного заряда полупроводника для одного типа носителей заряда

$$\frac{d}{dV} \ln i_s = \frac{d}{dV} \ln \mathcal{P}_n^* + \frac{e}{kT} \frac{d}{dV} \Delta\Phi_i, \quad (3.111)$$

где для диодной теории $\mathcal{P}_n^* = \mathcal{P}_n$, а для диффузионной

$$\mathcal{P}_n^* = \frac{4E_2^0 u_n \mathcal{P}_n}{\mathcal{P}_n v_n + 4E_2^0 u_n}.$$

Легко показать, что для диодной и диффузионной теории при выполнении неравенства $\mathcal{P}_n \ll \frac{4E_2^0 u_n}{v_n}$

$$\frac{d}{dV} \ln \mathcal{P}_n^* \approx \left(\frac{2m^*}{\Phi} \right)^{1/2} \frac{ed^*}{2\hbar} \frac{dV_i}{dV} = a_t \frac{dV_i}{dV}. \quad (3.112)$$

Здесь $\frac{dV_i}{dV} = 1 - \alpha_0 \left(\frac{e}{kT} \right)^{-1}$; $d^* \leq d$ из-за влияния сил изображения. Величина $\frac{d}{dV} \Delta\Phi_i$ может быть представлена в виде

$$\frac{d}{dV} \Delta\Phi_i = - \frac{\Delta\Phi_i e}{4(\Phi_0 - eV_2)} \cdot \frac{dV_2}{dV}, \quad (3.113)$$

где $\frac{dV_2}{dV} = \frac{\alpha_0}{e/kT}$, а для $\Delta\Phi_i$ использовано выражение (3.63).

Все это позволяет выразить суммарное значение величин α и α^* для рассматриваемого случая в следующем общем виде:

$$\alpha = \alpha_0 \left[1 - a_t \frac{kT}{e} - \frac{\Delta\Phi_i}{n_i(\Phi_0 - eV_2)} \right] + a_i; \quad (3.114)$$

$$\alpha^* = \frac{e}{kT} - \alpha_0 \left[1 - a_t \frac{kT}{e} - \frac{\Delta\Phi_i}{n_1(\Phi_0 - eV_2)} \right] - a_i, \quad (3.115)$$

где α_0 задается формулами (3.107) — (3.110).

Оценки показывают, величина $a_t \frac{kT}{e}$ обычно не превышает значения 0,2 при толщинах 30...60 Å. Учет сил изображения также не приводит к заметным поправкам к построенным величинам.

При достаточно больших значениях $(\Phi_0 - eV_2)$

$$\frac{\Delta\Phi_i}{4(\Phi_0 - eV_2)} \ll 1.$$

В случае, когда можно пренебречь токами через поверхностные уровни, в приближении диффузионной теории при выполнении неравенства $\mathcal{P}_n \gg \frac{4E_2^0 u_n}{v_n}$

$$\frac{d}{dV} \ln \mathcal{P}_n^* = - \frac{1}{e\mathcal{L}E_2^0} \cdot \frac{dV_2}{dV}. \quad (3.116)$$

Тогда

$$\alpha = \alpha_0 \left[1 - \frac{kT}{e\mathcal{L}E_2^0} - \frac{\Delta\Phi_i}{4(\Phi_0 - eV_2)} \right]. \quad (3.117)$$

Для равномерного распределения примесей

$$\alpha = \alpha_0 \left[1 - \frac{kT}{2(\Phi_0 - eV_2)} - \frac{\Delta\Phi_i}{4(\Phi_0 - eV_2)} \right]. \quad (3.118)$$

Оценки дополнительного члена $\frac{kT}{2(\Phi_0 - eV_2)}$ для равномерного распределения примесей дают величины меньше 0,25 для высот потенциального барьера $(\Phi_0 - eV_2) \geq 2kT$, которые согласуются с критериями получения слоя Шоттки.

Проведенный подробный анализ выражения (3.111) показывает, что учет зависимости от напряжения коэффициента прозрачности \mathcal{P}_n и поля E_2^0 дает поправки в величину α , которые не превышают 20%, поэтому в первом приближении можно считать, что для рассматриваемых случаев $\alpha \simeq \alpha_0$.

Однако зависимость коэффициента прозрачности \mathcal{P}_n или поля E_2^0 от напряжения существенно может изменять величину α^* .

Например, в случае тесного контакта металла с полупроводником, когда $d = 0$,

$$\alpha = \frac{e}{kT} \left[1 - \frac{\Delta\Phi_i}{4(\Phi_0 - eV)} \right] \simeq \frac{e}{kT}. \quad (3.119)$$

Но в этом же случае $\alpha^* \neq 0$, а задается выражением

$$\alpha^* = \frac{e}{kT} \cdot \frac{\Delta\Phi_i}{4(\Phi_0 - eV)}, \quad (3.120)$$

которое, как проанализировано выше, может достичь достаточно большой величины $\alpha^* = 0,2 \frac{e}{kT}$.

В том случае, когда можно пренебречь токами через поверхностные уровни и существенную роль играют туннельные токи через область пространственного заряда полупроводника, при пренебрежении зависимостью от напряжения коэффициента прозрачности $\mathcal{P}_n(V)$

$$\frac{d \ln i_s}{dV} = \frac{e}{kT^*} \left(1 - \frac{T^*}{T} \right) \frac{dV_2}{dV}. \quad (3.121)$$

Это позволяет записать:

$$\alpha \simeq \alpha_0 \frac{T}{T^*}; \quad (3.122)$$

$$\alpha^* \simeq \frac{e}{kT} - \alpha_0 \frac{T}{T^*}. \quad (3.123)$$

Поскольку, как показано выше, величина T^* может существенно превышать величину T , то учет туннелирования может привести к существенным изменениям величин α и α^* .

Например, для контакта без зазора при $E_{00} \gg kT$ $\alpha = \frac{e}{E_{00}}$, $\alpha^* = \frac{e}{kT}$, что существенно отличается от величин $\alpha \simeq \frac{e}{kT}$, $\alpha^* \rightarrow 0$, которые соответствуют тесному контакту без учета туннелирования.

Когда главную роль играют токи через локальный поверхностный уровень, для нахождения величины $\frac{d}{dV} \ln i_s$ в случае диодной теории следует пользоваться выражением (3.81). В результате при предположении, что c_n , c_n^m и Q_m постоянны, получим

$$\frac{d}{dV} \ln i_s = \frac{e}{kT} \frac{dV_1}{dV} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1} - \frac{e}{kT} K \frac{dV_2}{dV} e^{-\frac{\varphi_0}{kT}} e^{\frac{eV_2}{kT}} \left[1 + K e^{-\frac{\varphi_0}{kT}} \left(e^{\frac{eV_2}{kT}} + M_i\right)\right]^{-1}. \quad (3.124)$$

В случае преимущественного обмена с полупроводником $K \rightarrow \infty$ и

$$\frac{d}{dV} \ln i_s = \frac{e}{kT} \left[\left(1 - \frac{dV_2}{dV}\right) \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1} - \frac{dV_2}{dV} \left(1 + M_i e^{-\frac{eV_2}{kT}}\right)^{-1} \right]. \quad (3.125)$$

Здесь $\frac{dV_2}{dV} = \alpha_0 \left(\frac{e}{kT}\right)^{-1}$, где α_0 задается выражением (3.107).

В случае преимущественного обмена с металлом $K \rightarrow 0$ и

$$\frac{d}{dV} \ln i_s = \frac{e}{kT} \left[\left(1 - \frac{dV_2}{dV}\right) \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1} \right], \quad (3.126)$$

где $\frac{dV_2}{dV}$ может быть получено из выражения (3.92).

Это позволяет найти выражение для α в случае, когда главную роль играют токи через поверхностные уровни. Для преимущественного обмена поверхностных уровней с полупроводником

$$\alpha = \alpha_0 \left[1 - \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1} - \left(1 + M_i e^{-\frac{eV_2}{kT}}\right)^{-1} \right] + \frac{e}{kT} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1}, \quad (3.127)$$

где α_0 определяется соотношением (3.107).

Для преимущественного обмена поверхностных уравнений с металлом

$$\alpha = \alpha_0 \left[1 - \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1} \right] + \frac{e}{kT} \left(1 + M_i e^{\frac{eV_1}{kT}}\right)^{-1}, \quad (3.128)$$

где α_0 определяется соотношением (3.109). Величина α^* легко определяется из соотношения (3.95).

В рассматриваемом случае величины α могут существенно образом отличаться от величины α_0 . Например, при $M_i e^{\frac{eV_1}{kT}} \ll 1$ из (3.127) видно, что $\alpha = \frac{e}{kT}$, в то время как α_0 может быть значительно меньше $\frac{e}{kT}$.

Проведенный выше анализ величин α и α^* относился к токам основных носителей заряда. Для токов неосновных носителей заряда, рассмотренных в п. 3.2, анализ величин α и α^* может быть проведен аналогичным образом. Например, для неосновных носителей заряда в предположениях диодной теории можно получить

$$\alpha \simeq \frac{e}{kT}; \quad \alpha^* \rightarrow 0.$$

О связи параметров α и α^* с параметрами $\bar{\alpha}$ и $\bar{\alpha}^*$. При теоретическом анализе вольт-амперных характеристик удобно пользоваться аппроксимацией, обычно хорошо описывающей экспериментальные характеристики:

$$i = i_s e^{\bar{\alpha}V} - i_s^* e^{-\bar{\alpha}^*V} = i_s \left(\frac{i_s}{i_s^*} e^{\bar{\alpha}V} - e^{-\bar{\alpha}^*V} \right), \quad (3.129a)$$

или

$$i = i_s (e^{\bar{\alpha}V} - 1) - i_s^* (e^{-\bar{\alpha}^*V} - 1). \quad (3.129b)$$

(Эти выражения практически идентичны и лишь несколько отличаются только при малых токах).

Параметры $\bar{\alpha}$ и $\bar{\alpha}^*$ в формулах (3.129a) и (3.129b) в общем случае могут отличаться от значений α и α^* , проанализированных выше. Как легко показать:

$$\alpha = V \frac{d\bar{\alpha}}{dV} + \bar{\alpha}; \quad (3.130)$$

$$\alpha^* = V \frac{d\bar{\alpha}^*}{dV} + \bar{\alpha}^*. \quad (3.131)$$

Из этих выражений видно, что $\alpha = \bar{\alpha}$ и $\alpha^* = \bar{\alpha}^*$ только в том случае, если $\frac{d \ln \alpha}{d \ln V}$, $\frac{d \ln \alpha^*}{d \ln V} \ll 1$ или $\bar{\alpha}$ и $\bar{\alpha}^*$ постоянны.

Строгие выражения для $\bar{\alpha}$ и $\bar{\alpha}^*$ могут быть получены из сравнения выражений (3.88) и (3.129a).

Для случая, когда i_s и i_s^* постоянны:

$$\bar{\alpha} = \frac{e}{kT} \cdot \frac{V_s}{V}; \quad (3.132)$$

$$\bar{\alpha}^* = \frac{e}{kT} \cdot \frac{V_s}{V}, \quad (3.133)$$

где V_2 и V_1 получены в гл. 2. Из (3.132) и (3.133) видно, что параметры α и α^* равны $\bar{\alpha}$ и $\bar{\alpha}^*$, если $\frac{dV_2}{dV} = \frac{V_2}{V}$ и $\frac{dV_1}{dV} = \frac{V_1}{V}$ ($V_2 \sim V$ и $V_1 \sim V$).

В том случае, если V_1 и V_2 не являются удобными аналитическими функциями напряжения, для нахождения выражений $\bar{\alpha}$ и $\bar{\alpha}^*$ через известные значения α и α^* можно воспользоваться очевидными соотношениями:

$$\bar{\alpha} = \frac{\int_0^V \alpha(V) dV}{V}; \quad (3.134)$$

$$\bar{\alpha}^* = \frac{\int_0^V \alpha^*(V) dV}{V}. \quad (3.135)$$

Это позволяет найти и проанализировать точные или приближенные выражения для $\bar{\alpha}$ и $\bar{\alpha}^*$.

Следует особо отметить, что рассмотренные значения α и α^* , а также $\bar{\alpha}$ и $\bar{\alpha}^*$ относятся к статической вольт-амперной характеристике. Это означает, что поверхностные состояния полностью успевают «следить» за приложенным напряжением. Если это не так, то значения α , α^* и $\bar{\alpha}$, $\bar{\alpha}^*$ будут изменяться с частотой.

3.7. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ВОЛЬТ-АМПЕРНЫХ ХАРАКТЕРИСТИК

Подробные исследования выпрямляющих свойств контакта металл — полупроводник показали, что они зависят от тех же факторов, от которых зависят параметры энергетической диаграммы контакта, а именно: удельного сопротивления и типа проводимости полупроводника, обработки поверхности полупроводника, материала металлического электрода и метода его нанесения.

Наиболее полно исследовались контакты металл — полупроводник, в которых реализовался слой Шоттки (так называемые диоды с барьером Шоттки). Для них, в частности, показано, что на параметры вольт-амперной характеристики α , α^* , i_s , i_s^* существенным образом влияет обработка поверхности полупроводника перед нанесением металла. Это иллюстрируется рис. 3.13, где представлены типичные вольт-амперные характеристики контактов, полученных для основных носителей заряда на поверхностях кремния, для которых принимались (I) или не принимались (II) специальные меры по удалению слоя окисла.

Как видно, для поверхности с минимальным слоем окисла прямая ветвь вольт-амперной характеристики контакта в полулогарифмическом масштабе линейна, а наклон близок к $\alpha = \frac{e}{kT}$. Для поверхности со слоем окисла прямая ветвь вольт-амперной характеристики в полулогарифмическом масштабе не является прямой, а ее наклон при любых напряжениях $\alpha < \frac{e}{kT}$. Обратный ток для кон-

тактов с $\alpha < \frac{e}{kT}$ сильнее изменяется с приложенным напряжением, чем для контактов с $\alpha \approx \frac{e}{kT}$.

Такое поведение прямой и обратной ветви вольт-амперной характеристики может быть объяснено наличием диэлектрического зазора в контактах, где $\alpha < \frac{e}{kT}$ (см. п. 3.6).

Действительно, прямыми опытами было установлено, что при окислении скототой в вакууме поверхности кремния наблюдается переход от характеристики с $\alpha \approx \frac{e}{kT}$ к ха-

рактеристикам с $\alpha < \frac{e}{kT}$.

Вольт-амперные характеристики с $\alpha \approx \frac{e}{kT}$ достаточно

хорошо описываются теоретическими формулами при $d = 0$, учитывающими силы изображения, туннельный эффект и другие явления (смотри кривые 1—3 на рис. 3.13, а). Однако детальное сравнение теории и эксперимента показывает, что все же имеются факты, которые не могут быть объяснены без учета диэлектрического зазора и поверх-

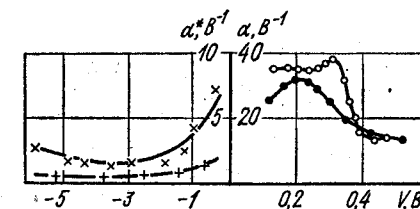


Рис. 3.14. Зависимости параметров α и α^* от напряжения для контактов никель-*p*-кремний.

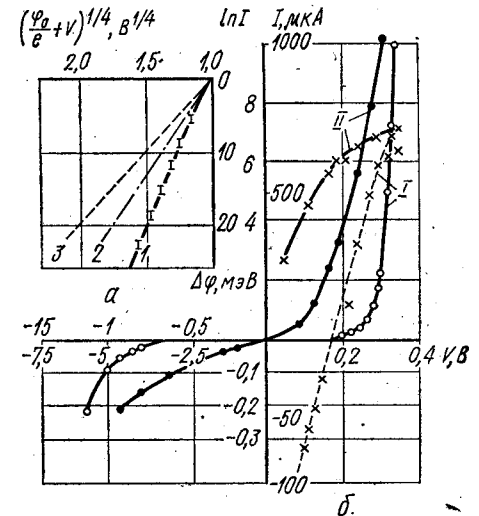


Рис. 3.13. Вольт-амперные характеристики контакта металл — *p*-кремний:

1 — экспериментальные данные; 2 — расчет с учетом сил изображения и поверхностных состояний; 3 — расчет с учетом только сил изображения.

ностных состояний. К ним относятся детальный ход параметров α и α^* с изменением напряжения, представленный на рис. 3.14. Как видно из этого рисунка, зависимость параметра α от V имеет сложный вид, что не может быть объяснено без учета диэлектрического зазора и поверхностных состояний (см. например, выражения (3.107) или (3.109)). Без

учета этих факторов не может быть строго объяснена и зависимость α^* от напряжения.

Сравнение теории и эксперимента для случая, когда $\alpha < \frac{e}{kT}$, показывает, что наблюдаемые экспериментально зависимости не могут быть объяснены без учета зазора и поверхностных состояний (штриховая кривая на рис. 3.13, б построена по формуле, учитывающей зазор с $d = 5 \text{ \AA}$). Особенно большие различия теории и эксперимента наблюдаются для обратной ветви вольт-амперной характеристики. Об этом же свидетельствует детальный ход зависимости α и α^* от напряжения.

Наличие зазора и поверхностных состояний как для контактов с $\alpha \approx \frac{e}{kT}$, так и для контактов с $\alpha < \frac{e}{kT}$ подтверждают и другие наблюдаемые закономерности.

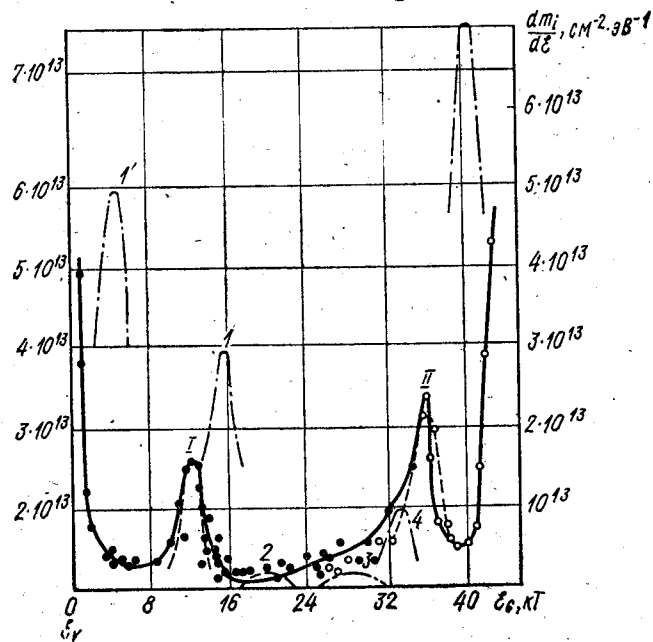


Рис. 3.15. Энергетическое распределение поверхностных состояний для контактов вольфрам-кремний при $d = 30 \dots 40 \text{ \AA}$ (● — n-кремний; ○ — p-кремний).

Величины i_s и i_s^* зависят от толщины диэлектрического зазора. Причем, даже для диодов с параметром $\alpha \approx \frac{e}{kT} i_s$, рассчитанное по известным значениям ϕ_0 и других параметров, не совпадает с измеренным значением, если не учитывать коэффициента прозрачности \mathcal{P}_n зазора и других факторов.

Таким образом, проведенный выше анализ результатов экспериментального исследования вольт-амперных характеристик диодов с барьером Шоттки показывает, что наблюдаемые закономерности подтверждают основные выводы описанной теории при учете зазора и поверхностных состояний даже для диодов с $\alpha \approx \frac{e}{kT}$.

Анализируя зависимости емкости, измеренной в СВЧ-диапазоне, от приложенного напряжения удалось непосредственно определить спектр поверхностных состояний в контакте кремния с различными металлами. На рис. 3.15 приведена типичная для поверхности (III) кремния зависимость плотности поверхностных состояний от энергии. Как видно из рисунка, для изучаемых прижимных контактов вольфрам-кремний реализуется случай непрерывного спектра в нижней и верхней части запрещенной зоны полупроводника, на который наложен спектр дискретных уровней I и II (теоретические кривые для уровней изображены пунктиром). На том же рисунке для сравнения представлены данные для свободной поверхности (максимумы 1—4) и результаты, полученные для структур металл-окисел-полупроводник. Как видно, данные для контакта металл-полупроводник больше соответствуют результатам на свободной поверхности, хотя и несколько отличаются от них. Это может быть связано с влиянием металла на энергетический спектр поверхностных состояний.

В подтверждение этого свидетельствует изменение наблюдаемого спектра с изменением толщины диэлектрического зазора.

Знание энергетического спектра поверхностных состояний позволило провести более детальное сравнение теоретических и экспериментальных вольт-амперных характеристик для контактов с

$\alpha < \frac{e}{kT}$. Для этого использовалась найденная из СВЧ-измерений зависимость V_1 и V_2 от напряжения V . Для расчета вольт-амперных характеристик без учета токов через поверхностные состояния применялась формула (3.4), а с учетом токов через поверхностные состояния — формула (3.83). Результаты расчета зависимости $\ln i$ от V путем параллельного переноса вдоль оси $\ln i$ наилучшим образом

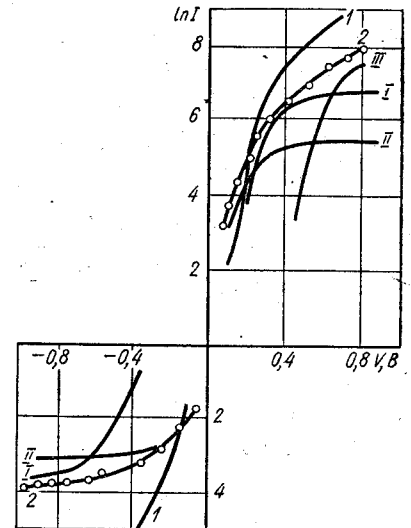


Рис. 3.16. Экспериментальные и теоретические вольт-амперные характеристики контактов вольфрам-кремний: 1 — без учета токов через поверхностные уровни; 2 — для токов через поверхностные уровни; 3 — расчетные токи через поверхностные уровни.

согласовывались с экспериментальными точками (по существу, расчет проводился с точностью до величины i_s , которая определялась из совпадения теоретических* и экспериментальных результатов).

Типичные результаты сравнения теории и эксперимента для контактов вольфрам — p -кремний представлены на рис. 3.16. Как видно из этого рисунка, расчет для токов без участия поверхностных уравнений не дает совпадения с экспериментом при любых постоянных i_s . Расчет же для токов с участием поверхностных состояний хорошо совпадает с экспериментом как для прямого, так и для обратного напряжения.

Таким образом, из изложенного следует, что для объяснения вольт-амперных характеристик контакта в общем случае следует учитывать зазор и поверхностные состояния, через которые может проходить ток.

Глава 4. ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ $p-n$ ПЕРЕХОДА

4.1. ОБЩЕЕ РАССМОТРЕНИЕ НЕВЫРОЖДЕННОГО $p-n$ ПЕРЕХОДА

На рис. 4.1 представлен $p-n$ переход, у которого и p - и n -области являются невырожденными полупроводниками и пренебрежимо мала концентрация глубоких уровней. В таком переходе при прямых и не очень больших обратных напряжениях нет ни туннельных переходов носителей, ни междузонных переходов заряда между зоной проводимости и валентной зоной с участием глубоких уровней и движущиеся через $p-n$ переход электроны находятся в зоне проводимости, а движущиеся дырки в валентной зоне и в n - и p -области $p-n$ перехода.

Как показано в гл. 2, при этом для $p-n$ перехода, к которому приложено напряжение, происходит изменение концентрации неосновных носителей заряда в n - и p -области. В связи с этим при рассмотрении токов через $p-n$ переход в квазинейтральных областях полупроводника необходимо анализировать диффузионно-дрейфовые процессы движения неосновных носителей заряда.

При рассмотрении прохождения тока через собственно $p-n$ переход, как и для контакта металл — полупроводник, важно то, испытывают или не испытывают носители заряда столкновения при прохождении области пространственного заряда. Если носители заряда не испытывают столкновений, то справедливы предположения диодной теории и носители движутся без изменения энергии в области пространственного заряда. Если же носители заряда испытывают многократные столкновения, то справедливы предположения диффузионной теории и носители заряда изменяют свою энергию, проходя через область пространственного заряда. Это показано стрелками и ломаными линиями на рис. 4.1.

При последовательном прохождении носителей заряда в пределах каждой из разрешенных зон ток через $p-n$ переход может контролироваться как квазинейтральными областями, так и областью пространственного заряда.

В зависимости от величины проходящего тока можно различать случаи малых и больших уровней инжекции. При малых уровнях инжекции концентрация инжектированных неосновных носителей заряда всегда значительно меньше концентрации основных носителей заряда. Для n -области $p-n$ перехода это означает, что $\Delta p \ll \ll n_n$, а для p -области, что $\Delta n \ll p_p$. В этом случае, когда хотя бы одно из записанных выше неравенств нарушается, наблюдаются большие уровни инжекции.

Для нахождения вольт-амперной характеристики $p-n$ перехода необходимо найти сумму плотностей токов электронов и дырок в любом сечении $p-n$ перехода

$$i = i_n + i_p. \quad (4.1)$$

В том случае, когда ширина $p-n$ перехода достаточно мала по сравнению с диффузионной длиной для электронов и дырок ($\mathcal{L}_p \ll \ll l_p$ и $\mathcal{L}_n \ll l_n$), рекомбинацией электронов и дырок в области пространственного заряда можно пренебречь. Тогда токи электро-

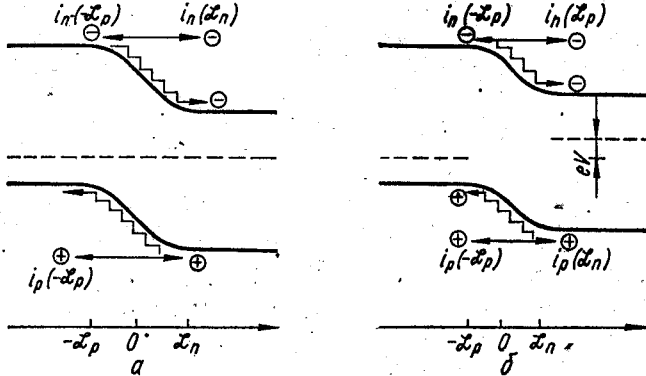


Рис. 4.1. Модель $p-n$ перехода:
а - $V = 0$; б - $V > 0$.

нов и дырок не изменяются в пределах области пространственного заряда:

$$\begin{aligned} i_n(-\mathcal{L}_p) &= i_n(\mathcal{L}_n); \\ i_p(-\mathcal{L}_p) &= i_p(\mathcal{L}_n). \end{aligned} \quad (4.2)$$

Это позволяет при расчете вольт-амперной характеристики рассмотреть сумму плотностей токов электронов и дырок, отнесенных к разным плоскостям, и записать

$$i = i_n(-\mathcal{L}_p) + i_p(\mathcal{L}_n). \quad (4.3)$$

При этой записи электронный ток берется на границе квазинейтральной p -области с областью пространственного заряда, а дырочный на границе квазинейтральной n -области с областью пространственного заряда. Такое разнесение плоскостей для тока электронов и дырок позволяет упростить задачу нахождения вольт-амперной характеристики и свести ее к решению уравнений, относящихся к квазинейтральным областям, а не к области пространственного заряда, как это делалось для контакта металл — полупроводник. В этом случае можно сразу получить общее решение как для диодной, так и диффузионной теории.

Так, для $p-n$ перехода с широкой базой, используя выражение вида (2.18), получаем, что для малых уровней инжекции после подстановки $x = \mathcal{L}_n$ и $x = -\mathcal{L}_p$

$$i = e \left(\frac{D_n \Delta n(-\mathcal{L}_p)}{l_n} + \frac{D_p \Delta p(\mathcal{L}_n)}{l_p} \right). \quad (4.4)$$

Для $p-n$ перехода с узкой базой на основании выражения (2.19) имеем для малых уровней инжекции после подстановки $x = \mathcal{L}_n$ и $x = -\mathcal{L}_p$

$$\begin{aligned} i = e \left\{ \frac{D_p \left[\Delta p(\mathcal{L}_n) \operatorname{ch} \frac{w_n - \mathcal{L}_n}{l_p} - \Delta p(w_n) \right]}{l_p \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}} + \right. \\ \left. + \frac{D_n \left[\Delta n(-\mathcal{L}_p) \operatorname{ch} \frac{\mathcal{L}_p - w_p}{l_n} - \Delta p(-w_p) \right]}{l_n \operatorname{sh} \frac{\mathcal{L}_p - w_p}{l_n}} \right\}. \end{aligned} \quad (4.5)$$

Для больших уровней инжекции, используя выражение (2.19), можно получить, что если весь ток через $p-n$ переход дырочный при $x = \mathcal{L}_n$, то

$$i = i_p(\mathcal{L}_n) = \frac{2e \left[\Delta p(\mathcal{L}_n) \operatorname{ch} \frac{w_n - \mathcal{L}_n}{l_{p3}} - \Delta p(w_n) \right] D_p}{l_{p3} \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_{p3}}}. \quad (4.6)$$

Как видно из выражений (4.4) — (4.6), окончательное решение задачи сводится к нахождению величин $\Delta p(\mathcal{L}_n)$ и $\Delta p(w_n)$, $\Delta n(-\mathcal{L}_p)$ и $\Delta p(-w_p)$, или величин $p(\mathcal{L}_n) = \Delta p(\mathcal{L}_n) + p_n$ и $n(-\mathcal{L}_p) = \Delta n(-\mathcal{L}_p) + n_p$.

Будем искать величины $p(\mathcal{L}_n)$ и $n(-\mathcal{L}_p)$ отдельно для случаев, когда для области пространственного заряда справедлива диодная или диффузионная теория. Для этого рассмотрим токи собственно через область пространственного заряда.

По диодной теории дырочный ток через $p-n$ переход можно записать как разность тока над барьером области пространственного заряда из p -области в n -область и тока из n -области в p -область и получить аналогично тому, как и для контакта металл — полупроводник в п. 3.1, что

$$i_p(\mathcal{L}_n) = \frac{ep_p \bar{v}_p}{4} e^{\frac{\varphi_0 - eV}{kT}} - \frac{ep(\mathcal{L}_n) \bar{v}_p}{4}. \quad (4.7)$$

Если дырочный ток через $p-n$ переход контролируется прохождением дырок через квазинейтральную n -область $p-n$ перехода, то $i_p(\mathcal{L}_n)$ значительно меньше, чем каждый из членов правой части равенства (4.7). В этом случае можно положить $i_p(\mathcal{L}_n) = 0$ и, решая (4.7), получить следующее значение:

$$p(\mathcal{L}_n) = p_p e^{-\frac{\varphi_0 - eV}{kT}} = p_p e^{\frac{eV}{kT}}, \quad (4.8)$$

где использовано очевидное соотношение $p_n = p_p e^{-\frac{\varphi_0}{kT}}$.

Если же прохождение дырок через $p-n$ переход контролируется как квазинейтральной областью, так и областью пространственного заряда, то для нахождения $p(\mathcal{L}_n)$ необходимо решить уравнение (4.7), в котором сохранен ток $i_p(\mathcal{L}_n)$. В результате имеем

$$p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}} - \frac{4i_p(\mathcal{L}_n)}{e\nu_p}. \quad (4.8a)$$

Для нахождения $p(\mathcal{L}_n)$ в предположениях диффузионной теории воспользуемся понятиями квазиуровней Ферми для электронов и дырок μ_n и μ_p , с помощью которых можно ввести суммарную концентрацию электронов и дырок в зоне проводимости и валентной зоне при отсутствии термодинамического равновесия следующим образом:

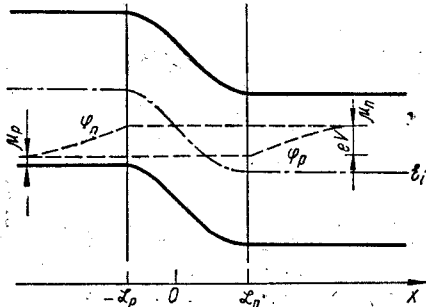


Рис. 4.2. Уровень и квазиуровни Ферми в $p-n$ переходе.

$$n = N_c e^{-\frac{\mu_n}{kT}} = n_i e^{\frac{e(\psi - \varphi_n)}{kT}}; \quad (4.9)$$

$$p = N_v e^{\frac{\mu_p - \psi}{kT}} = n_i e^{\frac{e(\varphi_p - \psi)}{kT}},$$

где $\varphi_{n,p} = -\frac{\mu_{n,p}}{e}$ и $\psi = -\frac{e_i}{e}$ показаны на рис. 4.2.

Подставляя n и p в диффузионно-дрейфовые уравнения, легко получить:

$$i_n = -eu_n n \frac{d\psi}{dx} + D_n \left[n_i \frac{e}{kT} e^{\frac{e(\psi - \varphi_n)}{kT}} \left(\frac{d\psi}{dx} - \frac{d\varphi_n}{dx} \right) \right] = -eu_n n \frac{d\varphi_n}{dx}; \quad (4.10)$$

$$i_p = -eu_p p \frac{d\psi}{dx} - D_p \left[n_i \frac{e}{kT} e^{\frac{e(\varphi_p - \psi)}{kT}} \left(\frac{d\varphi_p}{dx} - \frac{d\psi}{dx} \right) \right] = -eu_p p \frac{d\varphi_p}{dx}. \quad (4.11)$$

Из этих выражений видно, что электронный и дырочный токи пропорциональны градиентам квазиуровней Ферми для электронов и дырок.

Рассмотрим рис. 4.2, где показаны уровень Ферми и квазиуровни Ферми для электронов и дырок. В n -области и p -области при малых уровнях инжекции квазиуровни Ферми для основных носителей заряда электронов и дырок совпадают с уровнем Ферми. Квазиуровни Ферми для неосновных носителей заряда — электронов в дырочной области и дырок в электронной — отличаются от уровня Ферми вблизи $p-n$ перехода и совпадают с уровнем Ферми достаточно далеко от него.

Если использовать соотношение (4.11) для дырочного тока, то можно найти отношение токов $i_p(-\mathcal{L}_p)$ к току $i_p(\mathcal{L}_n)$ в виде

$$\frac{i_p(-\mathcal{L}_p)}{i_p(\mathcal{L}_n)} = \frac{p_p}{p(\mathcal{L}_n)} \cdot \frac{\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p}}{\frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}}. \quad (4.12)$$

Теперь учтем, что при принятых выше предположениях в пределах области пространственного заряда нет междузонных электронных и дырочных переходов и $\frac{i_p(-\mathcal{L}_n)}{i_p(\mathcal{L}_n)} = 1$. Тогда (4.12) можно переписать в виде

$$\frac{\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p}}{\frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}} = \frac{p(\mathcal{L}_n)}{p_p}. \quad (4.13)$$

Поскольку при малых уровнях инжекции $p(\mathcal{L}_n) \ll p_p$, то $\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p} \ll \frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}$, следовательно, φ_p практически изменяется только вне области $p-n$ перехода. Как видно из рис. 4.2, при этом $\varphi_p(\mathcal{L}_n) - \varphi_p(\infty) = V$.

Теперь, используя выражение (4.9), запишем

$$\frac{p(\mathcal{L}_n)}{p_n} = e^{\frac{\varphi_p - \varphi_p(\infty)}{kT}} = e^{\frac{\varphi_p(\mathcal{L}_n) - \varphi_p(\infty)}{kT}}, \quad (4.14)$$

или

$$p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}}. \quad (4.15)$$

Это выражение совпадает с выражением (4.8), полученным ранее в предположениях диодной теории для случая, когда дырочный ток контролируется квазинейтральной n -областью. Последнее не удивительно, поскольку равенство (4.13), по существу, также означает, что дырочный ток для диффузионной теории контролируется квазинейтральной n -областью, а не областью пространственного заряда.

Полученные выражения для $p(\mathcal{L}_n)$ в предположениях диодной и диффузионной теорий связывают плотность инжектированных дырок на границе $p-n$ перехода с n -областью с приложенным напряжением и плотностью неосновных носителей в n -области.

В случае, когда ток через $p-n$ переход контролируется квазинейтральной n -областью, величина $p(\mathcal{L}_n)$ зависит только от равновесной концентрации неосновных носителей заряда в n -области и приложенного напряжения (см. (4.8) и (4.15)). Особо отметим, что в величину $p(\mathcal{L}_n)$ не входят в явном виде параметры области пространственного заряда $p-n$ перехода. Аналогичным образом можно найти граничные условия для концентрации электронов $n(-\mathcal{L}_p)$.

Если же прохождение дырок через $p-n$ переход контролируется как квазинейтральной областью, так и областью пространственного заряда, то для нахождения $p(\mathcal{L}_n)$ необходимо решить уравнение (4.7), в котором сохранен ток $i_p(\mathcal{L}_n)$. В результате имеем

$$p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}} - \frac{4i_p(\mathcal{L}_n)}{e\nu_p}. \quad (4.8a)$$

Для нахождения $p(\mathcal{L}_n)$ в предположениях диффузионной теории воспользуемся понятиями квазиуровней Ферми для электронов и дырок μ_n и μ_p , с помощью которых можно ввести суммарную концентрацию электронов и дырок в зоне проводимости и валентной зоне при отсутствии термодинамического равновесия следующим образом:

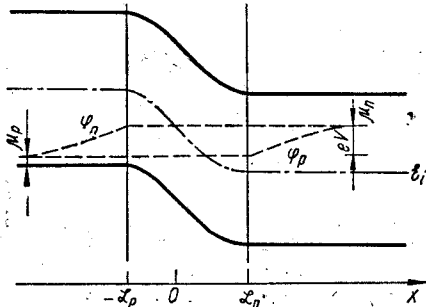


Рис. 4.2. Уровень и квазиуровни Ферми в $p-n$ переходе.

$$n = N_c e^{-\frac{\mu_n}{kT}} = n_i e^{-\frac{e(\psi - \varphi_n)}{kT}}; \quad (4.9)$$

$$p = N_v e^{-\frac{\mu_p - \psi}{kT}} = n_i e^{-\frac{e(\varphi_p - \psi)}{kT}},$$

где $\varphi_{n,p} = -\frac{\mu_{n,p}}{e}$ и $\psi = -\frac{\xi_i}{e}$ показаны на рис. 4.2.

Подставляя n и p в диффузионно-дрейфовые уравнения, легко получить:

$$i_n = -eu_n n \frac{d\psi}{dx} + D_n \left[n_i \frac{e}{kT} e^{-\frac{e(\psi - \varphi_n)}{kT}} \left(\frac{d\psi}{dx} - \frac{d\varphi_n}{dx} \right) \right] = -eu_n n \frac{d\varphi_n}{dx}; \quad (4.10)$$

$$i_p = -eu_p p \frac{d\psi}{dx} - D_p \left[n_i \frac{e}{kT} e^{-\frac{e(\varphi_p - \psi)}{kT}} \left(\frac{d\varphi_p}{dx} - \frac{d\psi}{dx} \right) \right] = -eu_p p \frac{d\varphi_p}{dx}. \quad (4.11)$$

Из этих выражений видно, что электронный и дырочный токи пропорциональны градиентам квазиуровней Ферми для электронов и дырок.

Рассмотрим рис. 4.2, где показаны уровень Ферми и квазиуровни Ферми для электронов и дырок. В n -области и p -области при малых уровнях инжекции квазиуровни Ферми для основных носителей заряда электронов и дырок совпадают с уровнем Ферми. Квазиуровни Ферми для неосновных носителей заряда — электронов в дырочной области и дырок в электронной — отличаются от уровня Ферми вблизи $p-n$ перехода и совпадают с уровнем Ферми достаточно далеко от него.

Если использовать соотношение (4.11) для дырочного тока, то можно найти отношение токов $i_p(-\mathcal{L}_p)$ к току $i_p(\mathcal{L}_n)$ в виде

$$\frac{i_p(-\mathcal{L}_p)}{i_p(\mathcal{L}_n)} = \frac{p_p}{p(\mathcal{L}_n)} \cdot \frac{\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p}}{\frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}}. \quad (4.12)$$

Теперь учтем, что при принятых выше предположениях в пределах области пространственного заряда нет междузонных электронных и дырочных переходов и $\frac{i_p(-\mathcal{L}_n)}{i_p(\mathcal{L}_n)} = 1$. Тогда (4.12) можно переписать в виде

$$\frac{\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p}}{\frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}} = \frac{p(\mathcal{L}_n)}{p_p}. \quad (4.13)$$

Поскольку при малых уровнях инжекции $p(\mathcal{L}_n) \ll p_p$, то $\frac{d\varphi_p}{dx} \Big|_{-\mathcal{L}_p} \ll \frac{d\varphi_p}{dx} \Big|_{\mathcal{L}_n}$, следовательно, φ_p практически изменяется только вне области $p-n$ перехода. Как видно из рис. 4.2, при этом $\varphi_p(\mathcal{L}_n) - \varphi_p(\infty) = V$.

Теперь, используя выражение (4.9), запишем

$$\frac{p(\mathcal{L}_n)}{p_n} = e^{-\frac{\varphi_p - \varphi_e[\varphi_p(\mathcal{L}_n) - \varphi_p(\infty)]}{kT}}, \quad (4.14)$$

или

$$p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}}. \quad (4.15)$$

Это выражение совпадает с выражением (4.8), полученным ранее в предположениях диодной теории для случая, когда дырочный ток контролируется квазинейтральной n -областью. Последнее не удивительно, поскольку равенство (4.13), по существу, также означает, что дырочный ток для диффузионной теории контролируется квазинейтральной n -областью, а не областью пространственного заряда.

Полученные выражения для $p(\mathcal{L}_n)$ в предположениях диодной и диффузионной теорий связывают плотность инжектированных дырок на границе $p-n$ перехода с n -областью с приложенным напряжением и плотностью неосновных носителей в n -области.

В случае, когда ток через $p-n$ переход контролируется квазинейтральной n -областью, величина $p(\mathcal{L}_n)$ зависит только от равновесной концентрации неосновных носителей заряда в n -области и приложенного напряжения (см. (4.8) и (4.15)). Особо отметим, что в величину $p(\mathcal{L}_n)$ не входят в явном виде параметры области пространственного заряда $p-n$ перехода. Аналогичным образом можно найти граничные условия для концентрации электронов $n(-\mathcal{L}_p)$.

В результате в предположениях диодной теории, когда электронный ток контролируется квазинейтральной p -областью,

$$n(-\mathcal{L}_p) = n_p e^{\frac{eV}{kT}}. \quad (4.16)$$

Если электронный ток контролируется как областью пространственного заряда, так и квазинейтральной областью в предположениях диодной теории, имеем

$$n(-\mathcal{L}_p) = n_p e^{\frac{eV}{kT}} - \frac{4i_n(-\mathcal{L}_p)}{ev_n}. \quad (4.16a)$$

В предположениях диффузионной теории граничная концентрация

$$n(-\mathcal{L}_n) = n_p e^{\frac{eV}{kT}}, \quad (4.17)$$

что совпадает с концентрацией, задаваемой выражением (4.16).

Отметим, что, как будет показано ниже, полученные граничные условия для $p(\mathcal{L}_n)$ и $n(-\mathcal{L}_p)$ играют определяющую роль для получения зависимости тока от напряжения в $p-n$ переходе.

В связи с этим хотелось бы еще раз подчеркнуть, что записанные выше общепринятые граничные условия в виде (4.8), (4.15), (4.16) и (4.17) получены при определенных приближениях, а именно:

предполагается отсутствие междузонных переходов в области пространственного заряда $p-n$ перехода;

предполагается, что ток через $p-n$ переход контролируется квазинейтральными областями;

пренебрегается нарушением функции распределения носителей заряда проходящим током.

Найденные граничные условия для $p(\mathcal{L}_n)$ и $n(-\mathcal{L}_p)$ пригодны как в случае малых уровней инжекции, так и больших, если выполняются предположения, принятые при их получении.

Для нахождения неизвестных величин $\Delta p(\omega_n)$ и $\Delta n(-\omega_p)$ можно воспользоваться соотношениями:

$$i_p(\omega_n) = eS_p \Delta p(\omega_n) = eS_p [p(\omega_n) - p_n]; \quad (4.18)$$

$$i_n(-\omega_p) = eS_n \Delta n(-\omega_p) = eS_n [n(-\omega_p) - n_p], \quad (4.19)$$

где S_p и S_n — скорости поверхностной рекомбинации дырок и электронов на омических контактах к n - и p -областям $p-n$ перехода. Эти соотношения выражают тот факт, что дырочный и электронный токи через омические контакты пропорциональны избыточной концентрации неосновных носителей заряда у этих контактов.

4.2. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА $p-n$ ПЕРЕХОДА С ШИРОКОЙ БАЗОЙ ПРИ МАЛЫХ УРОВНЯХ ИНЖЕКЦИИ

Рассмотрим вольт-амперную характеристику невырожденного $p-n$ перехода с широкой базой, когда размеры p - и n -областей существенно превышают длину диффузионного смещения неосновных носителей заряда и $l_p \ll \omega_n$, а $l_n \ll \omega_p$.

Вначале получим и проанализируем вольт-амперную характеристику для случая, когда дырочный и электронный токи контролируются квазинейтральными областями. При этом граничные значения $p(\mathcal{L}_n)$ и $p(-\mathcal{L}_p)$ одинаковы как для диодной, так и диффузионной теории.

Подставляя выражения (4.8) или (4.15), а также (4.16) или (4.17) в выражение (4.4), можно получить полную плотность тока через $p-n$ переход с широкой базой в виде:

$$i = i_s \left(e^{\frac{eV}{kT}} - 1 \right), \quad (4.20)$$

где $i_s = i_{sp} + i_{sn}$;

$$i_{sp} = \frac{ep_n D_p}{l_p}; \quad (4.21)$$

$$i_{sn} = \frac{en_p D_n}{l_n}. \quad (4.22)$$

Полученное выражение вольт-амперной характеристики $p-n$ перехода по внешнему виду совпадает с выражением вольт-амперной характеристики контакта металл — полупроводник при отсутствии зазора для токов основных носителей заряда без участия поверхностных уровней в переносе тока, вольт-амперной характеристики с участием поверхностных уровней в переносе тока при преимущественном обмене уровней с металлом (см. п. 3.1, 3.5), а также с вольт-амперной характеристикой контакта металл — полупроводник для неосновных носителей заряда, полученной в предположениях диодной теории при любых зазорах (см. п. 3.2).

Прямой ток для $p-n$ перехода, как и для рассмотренных выше случаев, растет с напряжением пропорционально $e^{\frac{eV}{kT}}$, а обратный достигает насыщения.

Однако плотность тока насыщения i_s для $p-n$ перехода может существенно отличаться от соответствующих плотностей токов для контакта металл — полупроводник как по величине, так и по физической природе.

Например, для контакта металл — полупроводник без зазора в случае диодной теории при переносе тока основными носителями заряда величина $i_s = \frac{ev_n n_0}{4} e^{-\frac{\Phi_0}{kT}}$, что равно количеству носителей заряда, падающих в единицу времени на границу раздела металл — полупроводник.

Для $p-n$ перехода плотность тока i_s определяется числом дырок, генерированных или рекомбинирующих в n -области на длине l_p и электронов, генерированных или рекомбинирующих в p -области на длине l_n в единицу времени.

Величины i_{sn} и i_{sp} , задаваемые выражениями (4.21) и (4.22), могут быть переписаны в виде:

$$i_{sn} = \frac{eD_n n_p}{l_n} = \frac{eD_n l_n n_p}{i_n^2} = \frac{e l_n n_p}{\tau_n};$$

$$i_{sp} = \frac{eD_p p_n}{l_p} = \frac{eD_p l_p p_n}{i_p^2} = \frac{e l_p p_n}{\tau_p},$$

откуда очевидна правильность сделанного выше утверждения.

Прямой ток p — n перехода при этом связан с инжекцией неосновных носителей заряда, которая существенно зависит от приложенного напряжения и определяется рекомбинацией введенных носителей в областях l_n и l_p , которая от напряжения не зависит. Обратный ток определяется темпом генерации носителей заряда в областях l_n и l_p и последующим их выведением через p — n переход, которые не зависят от приложенного напряжения.

В связи со сказанным и наблюдается столь различное поведение прямого и обратного токов с приложенным напряжением.

Величина i_s для p — n перехода при принятых предположениях о том, что ток контролируется прохождением через квазинейтральные области полупроводника, всегда меньше, чем для рассмотренного выше случая контакта металл—полупроводник, где ток контролируется прохождением носителей заряда через область пространственного заряда.

Математически это сводится к выполнению неравенства для тока электронов

$$\frac{D_n}{l_n} n_p \ll \frac{\bar{v}_n n_0}{4} e^{-\frac{\varphi_0}{kT}}.$$

Если учесть, что $n_p = n_n e^{-\frac{\varphi_0}{kT}}$, и принять $n_n = n_0$, последнее неравенство принимает следующий вид:

$$\frac{D_n}{l_n} \ll \frac{\bar{v}_n}{4}.$$

Например, для слабо легированного германия при $D_n \approx 100 \text{ см}^2 \times \text{с}^{-1}$ и $\bar{v}_n \approx 2 \cdot 10^7 \text{ см} \cdot \text{с}^{-1}$ последнее неравенство выполняется вплоть до $l_n > 2 \cdot 10^{-4} \text{ см}$, а обычные значения $l_n = 10^{-1} \dots 10^{-2} \text{ см}$ и записанное выше неравенство выполняется очень хорошо.

Для слабо легированного кремния при $D_n = 40 \text{ см}^2 \cdot \text{с}^{-1}$ и тех же значениях \bar{v}_n последнее неравенство выполняется лучше и справедливо вплоть до $l_n > 8 \cdot 10^{-5} \text{ см}$. Хотя для кремния обычно l_n меньше, чем для германия, и составляет $10^{-1} \dots 10^{-3} \text{ см}$, этого достаточно для хорошего выполнения рассматриваемого неравенства.

Важным для p — n перехода является отношение электронного тока к дырочному, а также коэффициент инжекции. Используя выражения (4.21) и (4.22), можно получить следующее отношение элек-

тронного тока к дырочному:

$$\frac{i_n}{i_p} = \frac{D_n n_p}{l_n} \frac{l_p}{D_p p_n} = \frac{u_n n_n}{u_p p_p} \frac{l_p}{l_n} = \frac{\sigma_n l_p}{\sigma_p l_n},$$

где использованы соотношения

$$\frac{u_n}{D_n} = \frac{e}{kT}; \quad n_n p_n = p_p n_p = n_i^2.$$

Как видно, это отношение определяется главным образом отношением проводимостей n - и p -областей полупроводника:

Коэффициент инжекции p — n перехода нужно отдельно определять для n -области и p -области, так как для них отличаются неосновные носители заряда: для n -области это дырки, а для p -области — электроны.

Для p -области

$$\gamma_p = \frac{i_n}{i_n + i_p} = \frac{1}{1 + \frac{\sigma_p l_n}{\sigma_n l_p}}; \quad (4.23)$$

для n -области

$$\gamma_n = \frac{i_p}{i_n + i_p} = \frac{1}{1 + \frac{\sigma_n l_p}{\sigma_p l_n}}. \quad (4.24)$$

Как видно из полученных выражений, в p — n переходе больше ток того знака, для которого больше проводимость соответствующей области. Так, если проводимость p -области больше, чем n -области, то ток будет дырочным, а если n -области больше, чем p -области, то ток в основном будет электронным.

Соответственно будет изменяться коэффициент инжекции. В первом случае он будет стремиться к единице для n -области и к нулю для p -области, а во втором к нулю для n -области и к единице для p -области.

Если ток через p — n переход контролируется кроме квазинейтральных областей также областью пространственного заряда, для нахождения вольт-амперной характеристики необходимо использовать выражения (4.8а) и (4.16а). Подставляя эти выражения в (4.4), можно решить полученное выражение относительно i_p (\mathcal{L}_n) и i_n ($-\mathcal{L}_p$) и получить окончательную вольт-амперную характеристику p — n перехода с широкой базой в виде (4.20), где

$$i_s = i_{sp} + i_{sn};$$

$$i_{sp} = \frac{e p_n D_p}{l_p \left(1 + \frac{4D_p}{\bar{v}_p l_p}\right)}; \quad (4.25)$$

$$i_{sn} = \frac{e n_p D_n}{l_n \left(1 + \frac{4D_n}{\bar{v}_n l_n}\right)}. \quad (4.26)$$

Как видно из последних выражений, при $\frac{4D_p}{v_p l_p} \ll 1$ и $\frac{4D_n}{v_n l_n} \ll 1$ они переходят в полученные ранее выражения (4.21) и (4.22) для случая, когда токи контролируются квазинейтральными областями.

При обратных неравенствах $\frac{4D_p}{v_p l_p} \gg 1$ и $\frac{4D_n}{v_n l_n} \gg 1$, выражения (4.25) и (4.26) переходят в следующие:

$$i_{sp} = \frac{e\bar{v}_p p_n}{4} = \frac{e\bar{v}_p p_p}{4} e^{-\frac{\varphi_0}{kT}};$$

$$i_{sn} = \frac{e\bar{v}_n n_p}{4} = \frac{e\bar{v}_n n_n}{4} e^{-\frac{\varphi_0}{kT}}.$$

В этом случае, когда ток через $p-n$ переход контролируется областью пространственного заряда, они полностью совпадают с соответствующими выражениями для контакта металл — полупроводник без зазора, полученными в п. 3.1, 3.2.

В заключение следует обратить внимание на то, что для таких широко распространенных полупроводников, как германий и кремний, ток через $p-n$ переход обычно контролируется квазинейтральными областями.

4.3. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА $p-n$ ПЕРЕХОДА С УЗКОЙ БАЗОЙ

Рассчитаем вольт-амперную характеристику невырожденного $p-n$ перехода при малых уровнях инжекции для случая, когда размеры p - и n -областей сравнимы с диффузионной длиной смещения неосновных носителей заряда. Для этого можно воспользоваться выражением (4.5), куда необходимо подставить величины $\Delta p(\mathcal{L}_n)$, $\Delta n(-\mathcal{L}_p)$, $\Delta p(-w_p)$ и $\Delta p(w_n)$. Величины $\Delta p(\mathcal{L}_n)$ и $\Delta n(-\mathcal{L}_p)$ можно найти, используя полученные ранее граничные условия для $p(\mathcal{L}_n)$ и $n(-\mathcal{L}_p)$. Для нахождения величины $\Delta p(w_n)$ подставим в выражение (2.19) $x = w_n$. В результате имеем для дырочного тока

$$i_p(w_n) = \frac{eD_p}{l_p \operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p}} \left[\Delta p(\mathcal{L}_n) - \Delta p(w_n) \operatorname{ch} \frac{w_n - \mathcal{L}_n}{l_p} \right]. \quad (4.27)$$

Теперь, приравняв выражения (4.27) и (4.18), можно получить уравнение относительно $\Delta p(w_n)$, решение которого дает

$$\Delta p(w_n) = \frac{\Delta p(\mathcal{L}_n)}{\operatorname{ch} \frac{\mathcal{L}_n - w_n}{l_p} + \frac{S_p l_p}{D_p} \left[\operatorname{sh} \frac{w_n - \mathcal{L}_n}{l_p} \right]}. \quad (4.28)$$

Аналогичным образом можно получить

$$\Delta n(-w_p) = \frac{\Delta n(-\mathcal{L}_p)}{\operatorname{ch} \frac{\mathcal{L}_p - w_p}{l_n} + \frac{S_n l_n}{D_n} \operatorname{sh} \frac{\mathcal{L}_p - w_p}{l_p}}. \quad (4.29)$$

Подставляя теперь значения $\Delta p(\mathcal{L}_n)$ и $\Delta n(-\mathcal{L}_p)$, для которых использованы выражения (4.8) и (4.17), а также значения $\Delta p(w_n)$ и $\Delta n(-w_p)$, полученные выше (см. (4.28) и (4.29)), в выражение (4.5), можно найти вольт-амперную характеристику $p-n$ перехода с узкой базой, по виду совпадающую с (4.20), где $i_s = i_{sp} + i_{sn}$:

$$i_{sp} = \frac{eD_p p_n}{l_p} \cdot \frac{1 + \frac{D_p}{S_p l_p} \operatorname{th} \frac{w_n - \mathcal{L}_n}{l_p}}{\frac{D_p}{S_p l_p} + \operatorname{th} \frac{w_n + \mathcal{L}_n}{l_p}}; \quad (4.30)$$

$$i_{sn} = \frac{eD_n n_p}{l_n} \cdot \frac{\left(1 + \frac{D_n}{S_n l_n} \operatorname{th} \frac{-w_p + \mathcal{L}_p}{l_n} \right)}{\left(\frac{D_n}{S_n l_n} + \operatorname{th} \frac{-w_p + \mathcal{L}_p}{l_n} \right)}. \quad (4.31)$$

Таким образом, общий вид вольт-амперной характеристики для $p-n$ перехода с узкой и широкой базой совпадает. Однако сами величины токов для таких переходов могут существенно отличаться. Это определяется тем, что для $p-n$ перехода величина i_s зависит дополнительно от таких параметров, как S_p , S_n , w_p , w_n , \mathcal{L}_p и \mathcal{L}_n .

Отношение токов для $p-n$ переходов с узкой и широкой базой можно найти, если использовать выражения (4.21), (4.22) и (4.30), (4.31). Для простоты дальнейшего анализа при этом можно ограничиться случаем, когда $i_{sp} \gg i_{sn}$.

Тогда $i_{sp} = i_s$;

$$\frac{i_s}{i_{s\infty}} = \frac{1 + \frac{D_p}{S_p l_p} \operatorname{th} \frac{w_n - \mathcal{L}_n}{l_p}}{\frac{D_p}{S_p l_p} + \operatorname{th} \frac{w_n - \mathcal{L}_n}{l_p}}, \quad (4.32)$$

где i_s относится к $p-n$ переходу с узкой, а $i_{s\infty}$ с широкой базой.

При изменении величины $\frac{D_p}{S_p l_p}$ последнее отношение принимает вид:

$$\frac{i_s}{i_{s\infty}} = \operatorname{th} \frac{w_n - \mathcal{L}_n}{l_p} \quad \text{для} \quad \frac{D_p}{S_p l_p} \gg 1; \quad (4.33a)$$

$$\frac{i_s}{i_{s\infty}} = 1 \quad \text{для} \quad \frac{D_p}{S_p l_p} = 1; \quad (4.33b)$$

$$\frac{i_s}{i_{s\infty}} = \operatorname{cth} \frac{w_n - \mathcal{L}_n}{l_p} \quad \text{для} \quad \frac{D_p}{S_p l_p} \ll 1. \quad (4.33в)$$

Зависимость $\frac{i_s}{i_{s\infty}}$ от $\frac{w_n - \mathcal{L}_n}{l_p}$ при различных значениях скорости рекомбинации показана на рис. 4.3. Из выражения (4.32) и рис. 4.3 видно, что при $S_p = 0$ дырочный ток уменьшается, а при $S_p = \infty$ увеличивается с ростом $\frac{w_n - \mathcal{L}_n}{l_p}$. При $S_p = \frac{D_p}{l_p}$ ток $p-n$ перехода

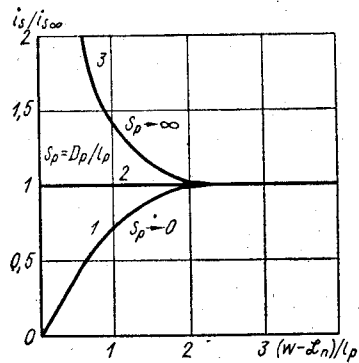


Рис. 4.3. Зависимость отношения $i_s/i_{s\infty}$ от $\frac{w_n - L_n}{l_p}$ при разных значениях скорости рекомбинации на невыпрямляющем контакте.

с узкой базой не зависит от $\frac{w_n - L_n}{l_p}$ и равен току $p-n$ перехода с широкой базой.

При $\frac{w_n - L_n}{l_p} \ll 1$, когда наблюдается максимальное отличие в $\frac{i_s}{i_{s\infty}}$, можно гиперболический тангенс разложить в ряд $\text{th} \frac{w_n - L_n}{l_p} \approx \frac{w_n - L_n}{l_p}$. Тогда

$$\frac{i_s}{i_{s\infty}} = \frac{D_p (w_n - L_n) + S_p l_p^2}{D_p l_p + (w_n - L_n) S_p} \quad (4.34)$$

Из выражения (4.32) видно, что при изменении скорости поверхностной рекомбинации от 0 до ∞ величина тока насыщения для $p-n$ перехода с узкой базой может изменяться в $\left(\frac{l_p}{w_n - L_n}\right)^2$ раз.

Выражение для плотности дырочного тока насыщения $p-n$ перехода с узкой базой для случая $\frac{w_n - L_n}{l_p} \ll 1$ можно переписать также в виде

$$i_s = \frac{e D_p p_n}{l_{\text{эф}}}, \quad (4.35)$$

где величину

$$l_{\text{эф}} = l_p \frac{\frac{D_p}{l_p} + \frac{w_n - L_n}{l_p} S_p}{\frac{D_p}{l_p} \frac{w_n - L_n}{l_p} + S_p} \quad (4.36)$$

можно рассматривать как эффективную диффузионную длину для дырок.

При $S_p = \infty$ эффективная диффузионная длина $l_{\text{эф}} = w_n - L_n$. Это означает, что вся рекомбинация происходит на невыпрямляющем электроде на расстоянии $w_n - L_n$ от $p-n$ перехода. При $S_p = \frac{D_p}{l_p}$ величина $l_{\text{эф}} = l_p$, т. е. эффективная диффузионная длина совпадает с диффузионной длиной для дырок в объеме n -области полупроводника. В этом случае рекомбинация на электроде происходит с такой скоростью, которая имела бы место в $p-n$ переходе с толстой базой в пространстве, расположенном за плоскостью, отстоящей от $p-n$ перехода на расстоянии $w_n - L_n$. При $S_p = 0$ ве-

личина $l_{\text{эф}} = \frac{l_p^2}{w_n - L_n}$, т. е. она больше диффузионной длины дырок в базе во столько раз, во сколько диффузионная длина больше толщины базы. Это вызвано тем, что вследствие малой скорости рекомбинации на контакте возрастает концентрация дырок в базе.

С увеличением толщины базы влияние скорости рекомбинации на контакт на ток насыщения быстро уменьшается. Уже при $w_n - L_n = 2l_p$ этим влиянием можно пренебречь.

Аналогичное рассмотрение может быть проведено и для электронного тока.

Используя полученные выше выражения для i_{sn} и i_{sp} , легко показать, что для $p-n$ перехода с узкой базой изменяется также коэффициент инжекции неосновных носителей заряда как в n -, так и в p -области. Действительно:

$$\gamma_n = \left[1 + \frac{\sigma_n}{\sigma_p} \frac{l_p}{l_n} \frac{\left(1 + \frac{D_n}{S_n l_n} \text{th} \frac{L_p - w_p}{l_n}\right)}{\left(\frac{D_n}{S_n l_n} + \text{th} \frac{L_p - w_p}{l_n}\right)} \right] \times \left[\frac{\left(\frac{D_p}{S_p l_p} + \text{th} \frac{w_n - L_n}{l_p}\right)}{\left(1 + \frac{D_p}{S_p l_p} \text{th} \frac{w_n - L_n}{l_p}\right)} \right]^{-1}; \quad (4.37)$$

$$\gamma_p = \left[1 + \frac{\sigma_p}{\sigma_n} \frac{l_n}{l_p} \frac{\left(\frac{D_n}{S_n l_n} + \text{th} \frac{L_p - w_p}{l_n}\right)}{\left(1 + \frac{D_n}{S_n l_n} \text{th} \frac{L_p - w_p}{l_n}\right)} \right] \times \left[\frac{\left(1 + \frac{D_p}{S_p l_p} \text{th} \frac{w_n - L_n}{l_p}\right)}{\left(\frac{D_p}{S_p l_p} + \text{th} \frac{w_n - L_n}{l_p}\right)} \right]^{-1}. \quad (4.38)$$

Как видно, коэффициент инжекции может изменяться при изменении не только величин $\sigma_n, \sigma_p, l_n, l_p$, а также при изменении величин w_n, w_p или S_n, S_p . В $p-n$ переходе идет больший ток того знака, для которого больше величина скорости поверхностной рекомбинации S и меньше величина w .

С физической точки зрения изменение величин i_{sp} и i_{sn} для $p-n$ перехода с узкой базой с изменением w_p, w_n, S_p и S_n связано с тем, что изменяются величины градиентов концентраций $\frac{dp}{dx}$ и $\frac{dn}{dx}$, которым пропорциональны токи.

4.4. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА $p-n$ ПЕРЕХОДА ПРИ БОЛЬШИХ УРОВНЯХ ИНЖЕКЦИИ

Рассмотрим вольт-амперную характеристику $p-n$ перехода при больших уровнях инжекции. При этом нужно, как уже рассматривалось выше, решать полные диффузионно-дрейфовые уравнения для электронов и дырок совместно с уравнениями неразрывности. Такая задача в общем виде не решается. Поэтому ограничимся рассмотрением частного случая $p-n$ перехода с тонкой n -областью, когда $w - \mathcal{L}_n \ll l_p$ при $S_p \rightarrow \infty$. Будем также предполагать, что коэффициент инжекции $\gamma_n \rightarrow 1$, т. е. весь ток в n -области является дырочным. При этих предположениях для нахождения вольт-амперной характеристики можно воспользоваться выражением (4.6).

В результате при $p(\mathcal{L}_n) = p_n e^{\frac{eV_n}{kT}}$, $p(w_n) = p_n$ и $l_{p2} = w_n - \mathcal{L}_n$, когда $\text{cth} \frac{w_n - \mathcal{L}_n}{l_{p2}} \approx \frac{l_{p2}}{w_n - \mathcal{L}_n}$, имеем

$$i = \frac{2eD_p p_n}{w_n - \mathcal{L}_n} e^{\frac{eV_n}{kT}}. \quad (4.39)$$

В этом выражении учтено, что $e^{\frac{eV_n}{kT}} \gg 1$, так как большие уровни инжекции реализуются только для больших прямых напряжений на $p-n$ переходе V_n .

Падение напряжения собственно на $p-n$ переходе

$$V_n = V - V_b, \quad (4.40)$$

где V_b — падение напряжения на квазинейтральной области. Используя выражение (2.14), можно найти

$$V_b = \int_{\mathcal{L}_n}^{w_n} E dx = \frac{kT}{e} \int_{\mathcal{L}_n}^{w_n} d \ln n = \frac{kT}{e} \ln \frac{n(\mathcal{L}_n)}{n(w_n)}. \quad (4.41)$$

Если учесть, что $n(\mathcal{L}_n) = p(\mathcal{L}_n) = p_n e^{\frac{eV_n}{kT}}$, а $n(w_n) = n_n$, то окончательно

$$V_b = \frac{kT}{e} \ln \frac{p_n}{n_n} + V_n. \quad (4.42)$$

Подставляя это выражение в (4.40), можно получить

$$V_n = \frac{V}{2} - \frac{1}{2} \frac{kT}{e} \ln \frac{p_n}{n_n} = \frac{V}{2} - \frac{1}{2} \frac{kT}{e} \ln \frac{p_n^2}{n_n^2}. \quad (4.43)$$

Теперь, подставляя это выражение в выражение для тока (4.39), имеем окончательное выражение для вольт-амперной характеристики:

$$i = \frac{2eD_p p_n}{w_n - \mathcal{L}_n} e^{\frac{eV}{2kT}}. \quad (4.44)$$

Для большого уровня инжекции ток $p-n$ перехода с тонкой базой пропорционален величинам n_i и $2D_p$ и изменяется с приложенным напряжением пропорционально $e^{\frac{eV}{2kT}}$, т. е. более слабо, чем для малых уровней инжекции.

Как видно из приведенного выше расчета, более медленное изменение тока с приложенным напряжением при больших уровнях инжекции обуславливается тем, что при больших уровнях инжекции на $p-n$ переходе падает только часть приложенного напряжения. Остальная часть напряжения приложена к квазинейтральной области полупроводника (см. выражение (4.43)).

При еще больших уровнях инжекции уже нельзя считать, что $i_n(\mathcal{L}_n) = 0$. В самом деле при $p = n$ устремить i к бесконечности при $\gamma = 1$ нельзя. Используя диффузионно-дрейфовые уравнения для дырочного и электронного токов, можно показать, что при достаточно больших токах и напряженностях поля, когда $i_p \rightarrow eu_p p E$ и $i_n \rightarrow eu_n p E$, коэффициент инжекции

$$\gamma \rightarrow \frac{1}{1 + \frac{u_p}{u_n}}.$$

В этом случае решение задачи при больших уровнях инжекции усложняется.

Его можно найти, предположив, что падением напряжения в $p-n$ переходе можно пренебречь по сравнению с падением напряжения в квазинейтральной области.

Тогда, используя диффузионно-дрейфовые выражения для n -области базы, можно записать

$$E = \frac{i}{e(u_{pp} + u_n(p + n_n))}. \quad (4.45)$$

Подставим это значение E в выражение для V_b , получим при $p \gg n_n$

$$V_b = - \int_0^{w_n} E dx = - \frac{i}{eu_p \left(1 + \frac{u_n}{u_p}\right)} \int_0^{w_n} \frac{dx}{p}. \quad (4.46)$$

Для нахождения связи между dx и dp используем уравнение неразрывности для дырочного тока

$$\frac{1}{e} \cdot \frac{di_p}{dx} = - \frac{p}{\tau_\infty},$$

куда подставим

$$i_p = eu_p p E = \frac{ip}{p + \frac{u_n}{u_p}(p + n_n)}.$$

В результате

$$\frac{i}{e} \cdot \frac{\frac{u_n}{u_p} n_n \frac{dp}{dx}}{\left[p + \frac{u_n}{u_p}(p + n_n)\right]^2} = - \frac{p}{\tau_\infty},$$

откуда можно найти при $p \gg n_n$.

$$dx = - \frac{in_n \tau_\infty \frac{u_n}{u_p}}{e \left(\frac{u_n}{u_p} + 1 \right)^2} \cdot \frac{dp}{p^3}. \quad (4.47)$$

Интегрируя последнее выражение от $x = 0$ до $x = w_n$, имеем

$$w_n = \frac{in_n \tau_\infty \frac{u_n}{u_p}}{2e \left(\frac{u_n}{u_p} + 1 \right)^2} \left[\frac{1}{p^2(0)} - \frac{1}{p^2(w_n)} \right]. \quad (4.48)$$

Подставляя dx в выражение (4.46), после интегрирования получаем

$$V_b = \frac{i^2 \frac{u_n}{u_p} n_n \tau_\infty}{3e^2 u_p \left(\frac{u_n}{u_p} + 1 \right)^3} \left[\frac{1}{p^3(w_n)} - \frac{1}{p^3(0)} \right]. \quad (4.49)$$

Если $p(0) \gg p(w_n)$, из выражений (4.49) и (4.48) легко получить окончательное выражение вольт-амперной характеристики n -области для достаточно больших уровней инжекции в виде

$$i = \frac{9}{8} \frac{en_n u_n u_p \tau_\infty}{w_n^3} V_b^2. \quad (4.50)$$

Как видно из последнего выражения, при принятых предположениях плотность тока пропорциональна квадрату напряжения, падающему в n -области, и обратно пропорциональна кубу ее толщины.

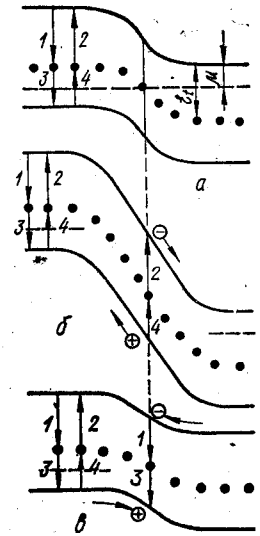
Зависимость тока от напряжения в рассмотренном случае ослабляется по сравнению со случаем, когда играет роль падение напряжения на p - n переходе.

4.5. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА p - n ПЕРЕХОДА С УЧЕТОМ ГЕНЕРАЦИИ — РЕКОМБИНАЦИИ В ОБЛАСТИ ПРОСТРАНСТВЕННОГО ЗАРЯДА

В предыдущих параграфах вывод вольт-амперных характеристик проводился в предположениях, что в области пространственного заряда нет генерации или рекомбинации носителей заряда. При этом считалось, что достаточным условием для выполнения этих предположений являются неравенства $\mathcal{L}_n \ll l_p$ и $\mathcal{L}_p \ll l_n$. В настоящем параграфе будет показано, что выполнение записанных условий еще не означает того, что можно пренебречь генерацией или рекомбинацией в области пространственного заряда, если темп генерации или рекомбинации зависит от поля, и будут рассчитаны точки рекомбинации и генерации в области пространственного заряда.

Рис. 4.4. Генерационно-рекомбинационные процессы в p - n переходе:

$a - V = 0, np = n_i^2$; $b - V < 0, np \ll n_i^2$; $c - V > 0, np > n_i^2$.



Наиболее распространенным механизмом генерации — рекомбинации является механизм с участием локальных центров, как это показано на рис. 4.4.

Для этого механизма вероятность переходов 1 пропорциональна числу электронов в зоне проводимости n и числу свободных уровней с концентрацией N_t , т. е. $N_t(1-f)$, где f — функция, характеризующая степень заполнения уровней. Вероятность переходов 2 пропорциональна числу электронов на уровнях, т. е. $N_t f$. В результате изменение электронов в зоне проводимости в единицу времени определяется разностью переходов 1 и 2 и выражается как

$$-\frac{dn}{dt} = \gamma_n N_t n (1-f) - \beta_n N_t f, \quad (4.51)$$

где γ_n и β_n — коэффициенты пропорциональности.

В условиях термодинамического равновесия $\frac{dn}{dt} = 0$, а $n = n_0$ и $f = f_0$. Это позволяет из (4.51) получить

$$\beta_n = \gamma_n n_0 \frac{1-f_0}{f_0}.$$

Подставляя сюда в качестве f_0 функцию распределения Ферми

$$f_0 = \frac{1}{1 + e^{\frac{\mathcal{E}_t - \mu}{kT}}}$$

и концентрацию электронов в зоне проводимости при отсутствии вырождения

$$n_0 = N_c e^{-\frac{\mu}{kT}},$$

можно найти, что

$$\beta_n = \gamma_n n_i,$$

где

$$n_i = N_c e^{-\frac{\mathcal{E}_t}{kT}} \quad (4.52)$$

совпадает с концентрацией электронов в полупроводнике, в котором уровень Ферми проходит через положение локального рекомбинационного центра.

Аналогичным образом можно рассмотреть изменение числа дырок в валентной зоне. Так, число дырочных переходов 4 пропорционально числу дырок p и числу свободных от дырок уровней, т. е. $N_t f$, а число обратных переходов 3 пропорционально числу дырок на уровнях $N_t(1-f)$.

В результате

$$-\frac{dp}{dt} = \gamma_p N_t p f - \beta_p N_t (1-f). \quad (4.53)$$

Используя условия равновесия, когда $\frac{dp}{dt} = 0$, $p = p_0$, $f = f_0$, имеем $\beta_p = \gamma_p p_1$, где

$$p_1 = N_D e^{-\frac{E_f - \mu}{kT}} \quad (4.54)$$

есть концентрация дырок в полупроводнике при совпадении уровня Ферми с уровнем рекомбинационного центра.

В стационарных условиях изменение концентрации электронов в зоне проводимости равно изменению концентрации дырок в валентной зоне, т. е. $\frac{dn}{dt} = \frac{dp}{dt}$. Приравнявая (4.51) и (4.53), найдем

$$f = \frac{\gamma_n n + \gamma_p p_1}{\gamma_n (n + n_1) + \gamma_p (p + p_1)} \quad (4.55)$$

Подставляя (4.55) в выражение (4.51) или (4.55), можно получить для стационарного случая

$$\frac{dn}{dt} = \frac{dp}{dt} = \frac{\gamma_n \gamma_p N_t (pn - n_1 p_1)}{\gamma_n (n + n_1) + \gamma_p (p + p_1)} \quad (4.56)$$

Затем полученное выражение можно разделить на $\gamma_n \gamma_p N_t$.

В результате при рекомбинации носителей заряда в полупроводнике с участием локальных уровней темп изменения концентрации дырок в валентной зоне со временем определяется следующим выражением:

$$\frac{dp}{dt} = \frac{pn - n_i^2}{\tau_{n0}(p + p_1) + \tau_{p0}(n + n_1)} \quad (4.57)$$

где p_1 , n_1 — концентрация носителей заряда в валентной зоне и зоне проводимости, когда уровень Ферми совпадает с энергетическим положением рекомбинационного центра; $\tau_{n0} \approx \frac{1}{\gamma_n N_t}$, $\tau_{p0} = \frac{1}{\gamma_p N_t}$ времена жизни электронов и дырок в сильно легированных полупроводниках и учитывается, что $n_1 p_1 = n_i^2$.

Из выражения (4.57) видно, что темп генерации — рекомбинации зависит от величин концентрации электронов n и дырок p в разрешенных зонах. Последние в свою очередь зависят от напряженности электрического поля. Так, при больших отрицательных напряжениях свободные электроны и дырки быстро удаляются из области пространственного заряда (см. рис. 4.4, $V < 0$) и можно считать, что $p \rightarrow 0$ и $n \rightarrow 0$. При этом наблюдается максимальный темп генерации, который постоянен и равен

$$\frac{dp}{dt} = \frac{n_i^2}{\tau_{n0} p_1 + \tau_{p0} n_1} \quad (4.58)$$

При положительных напряжениях удаление носителей из области пространственного заряда идет менее интенсивно, чем при $V = 0$ (см. рис. 4.4 $V > 0$) и наблюдается рекомбинация

$$\frac{dp}{dt} \approx - \frac{np}{\tau_{n0}(p_1 + p) + \tau_{p0}(n_1 + n)} \quad (4.59)$$

Общий ток генерации — рекомбинации, который, как ясно из сказанного выше, не приводит к инжекции неосновных носителей заряда в квазинейтральные области, можно получить, если найти интегралы

$$i = e \left[\int_{-\mathcal{L}_p}^0 \left(\frac{dp}{dt} \right)_p dx + \int_0^{\mathcal{L}_n} \left(\frac{dp}{dt} \right)_n dx \right] \quad (4.60)$$

куда подставить выражение (4.57) для p - и n -областей. Однако этот интеграл в общем виде не берется.

С достаточной степенью точности решение задачи может быть проведено для случаев больших отрицательных и положительных напряжений, когда преобладает либо генерация, либо рекомбинация носителей заряда. При обратных напряжениях, если использовать выражение (4.58) и подставить его в (4.60), имеем в случае несимметричного p - n перехода при $\mathcal{L}_n \gg \mathcal{L}_p$

$$i_r = e \frac{n_i^2}{\tau_{n0} p_1 + \tau_{p0} n_1} \int_0^{\mathcal{L}} dx = e \frac{n_i^2 \mathcal{L}}{\tau_{n0} p_1 + \tau_{p0} n_1} \quad (4.61)$$

где $\mathcal{L} \approx \mathcal{L}_n$.

Как видно из этого выражения, генерационный ток тем больше, чем больше величины n_i и \mathcal{L} и меньше величины τ_{n0} , τ_{p0} , p_1 и n_1 . Наибольший вклад в генерационный ток при $\tau_{n0} \sim \tau_{p0}$ дают уровни, расположенные вблизи середины запрещенной зоны, когда $p_1 = n_1 = n_i$. В этом случае

$$i_r = \frac{en_i \mathcal{L}}{\tau_{n0} + \tau_{p0}} \quad (4.62)$$

и генерационный ток пропорционален концентрации носителей заряда в собственном полупроводнике n_i , ширине области пространственного заряда \mathcal{L} и обратно пропорционален сумме времен жизни $\tau_{n0} + \tau_{p0}$.

Как видно из выражений (4.61) и (4.62), генерационный ток изменяется с приложенным напряжением, поскольку с напряжением изменяется ширина области пространственного заряда \mathcal{L} . В случае равномерного распределения примесей в n -области

$$i_r \sim (\Phi_0 - eV)^{1/2}$$

Обычно эта зависимость служит для доказательства того, что генерационный ток в p - n переходе играет определяющую роль.

Для численной оценки, чем контролируется прохождение тока через p - n переход — генерацией носителей заряда через локальные уровни в области пространственного заряда или диффузией в квазинейтральной области, найдем отношение i_r к току i_s , полученному в п. 4.1.

Если воспользоваться выражениями (4.21) и (4.62), можно получить

$$\frac{i_r}{i_d} = \frac{en_i \mathcal{L}}{\tau_{n0} + \tau_{p0}} \frac{l_p}{eD_p p n} = \frac{n_n}{n_i} \cdot \frac{\mathcal{L}}{l_p} \left(\frac{l_p}{l_{p3}} \right)^2, \quad (4.63)$$

где $l_{p3} = \sqrt{D_p(\tau_{p0} + \tau_{n0})}$.

Как видно из выражения (4.63), действительно, неравенство $\mathcal{L} \ll \ll l_p$ еще не означает, что генерационным током можно пренебречь, так как величины отношений $\frac{n_n}{n_i}$ и $\frac{l_p}{l_{p3}}$ могут быть значительно больше единицы. Доля генерационного тока пропорциональна уровню легирования полупроводника и обратно пропорциональна собственной концентрации. Поэтому такой ток будет играть более существенную роль для сильно легированных широкозонных полупроводников. Например, при $n_n = 10^{18} \text{ см}^{-2}$, $\mathcal{L} \approx 2 \cdot 10^{-5} \text{ см}$, $l_p = 10^{-1} \text{ см}$ и $l_p = l_{p3}$ для Ge $n_i = 2 \cdot 10^{13} \text{ см}^{-3}$ и $\frac{i_r}{i_d} = 0,1$, а для Si $n_i = 10^{10} \text{ см}^{-3}$ и $\frac{i_r}{i_d} \approx 2 \cdot 10^2$.

При прямых напряжениях расчет рекомбинационного тока более сложен, так как необходимо знать величины n и p в любом сечении области пространственного заряда. Для упрощения расчетов предположим, что p - n переход строго симметричен. Тогда в центре перехода $p = n = n_i$ при отсутствии напряжения $V = 0$, а при положительном напряжении $V > 0$

$$p = n = n_i e^{\frac{eV}{2kT}}. \quad (4.64)$$

Подставив это значение p и n в формулу (4.59) для уровня, лежащего в центре запрещенной зоны, где $n_1 = p_1 = n_i$, имеем

$$\frac{dp}{dt} = - \frac{n_i}{\tau_{p0} + \tau_{n0}} e^{\frac{eV}{2kT}}. \quad (4.65)$$

Используя то, что выражение (4.59) для $\tau_{n0} \approx \tau_{p0}$ имеет резкий максимум при $n = p$ в интервале изменения φ от $\varphi_{\text{max}} - kT$ до $\varphi_{\text{max}} + kT$, введем соответствующий интервал изменения

$$\Delta x = \frac{\Delta \varphi}{\left(\frac{d\varphi}{dx} \right)_{\varphi_{\text{max}}}} \approx \frac{2kT \mathcal{L}}{\varphi_0 - eV}. \quad (4.66)$$

Теперь можно приближенно вычислить интеграл (4.60), если подставить в него (4.59) при $n \approx p$ и заменить интегрирование умножением на Δx , соответствующему выражению (4.66). В результате

$$i_{\text{рек}} = \frac{2e \mathcal{L} n_i}{\tau_{p0} + \tau_{n0}} \frac{kT}{\varphi_0 - eV} e^{\frac{eV}{2kT}}. \quad (4.67)$$

Как видно из этого выражения, для генерационного тока, как и для рекомбинационного, ток имеет большую величину при боль-

ших \mathcal{L} , n_i и меньших $\tau_{p0} + \tau_{n0}$. Рекомбинационный ток существенно растет с напряжением и пропорционален $e^{\frac{eV}{2kT}}$.

Полученное значение рекомбинационного тока можно сравнить со значением диффузионного тока. При $\frac{eV}{2kT} \gg 1$ их отношение

$$\begin{aligned} \frac{i_{\text{рек}}}{i_d} &= \frac{2n_i \mathcal{L}}{\tau_{n0} + \tau_{p0}} \frac{l_p}{D_p p n} \cdot \frac{kT}{(\varphi_0 - eV)} e^{-\frac{eV}{2kT}} = \\ &= 2 \frac{n_n}{n_i} \frac{\mathcal{L}}{l_p} \left(\frac{l_p}{l_{p3}} \right)^2 e^{-\frac{eV}{2kT}} \frac{kT}{(\varphi_0 - eV)}. \end{aligned} \quad (4.68)$$

Как видно, это отношение меньше, чем аналогичное отношение для обратных напряжений как за счет множителя $\frac{kT}{\varphi_0 - eV}$, так и главным образом за счет множителя $e^{-\frac{eV}{2kT}}$. С ростом прямого напряжения отношение $\frac{i_{\text{рек}}}{i_d}$ убывает и при достаточно больших напряжениях всегда будет преобладать диффузионный ток. Так, например, для параметров, при которых проводились выше оценки отношения генерационного тока к диффузионному, рекомбинационным током можно будет пренебречь и для кремния, если $V \gg 0,1 \text{ В}$.

4.6. ИЗБЫТОЧНЫЕ ТОКИ В ТОНКИХ p - n ПЕРЕХОДАХ

В достаточно тонком p - n переходе при наличии локальных уровней в области пространственного заряда могут проходить избыточные токи, которые определяются процессами туннелирования, связанными с этими уровнями.

Возможные пути прохождения избыточного тока с участием уровней показаны на рис. 4.5. Как видно из этого рисунка, при достаточно узкой ширине области пространственного заряда электрон зоны проводимости может сначала протуннелировать на локальный уровень, а затем, разменяв энергию, попасть в валентную зону (путь А). Этим же путем могут происходить и обратные переходы электронов из валентной зоны в зону проводимости. При этом электрон сначала приобретает энергию, а затем происходит туннелирование.

Другой возможностью для попадания электрона из зоны проводимости в валентную зону является переход электрона из зоны проводимости на уровень с последующим туннелированием через область пространственного заряда в валентную зону (путь В). Обратные переходы в этом случае идут сначала путем туннелирования, а затем с приобретением энергии.

И, наконец, возможен переход электрона из зоны проводимости в валентную зону и обратно по пути С, когда электрон рассеивает или приобретает свою энергию в промежутках между последовательными туннельными переходами через локальные уровни.

Рассмотренная здесь картина отличается от проанализированной в предыдущем параграфе в случае генерационно-рекомбинационных токов тем, что в процессе межзонных переходов существенны не только вертикальные переходы с изменением энергии, но и туннельные переходы при постоянной энергии. При этом, как это отмечалось и в предыдущем параграфе, не происходит инъекции неосновных носителей заряда и в p и n области полупроводника.

Общее выражение для туннельного тока с зоны проводимости на локальные уровни по пути A может быть получено из следующих

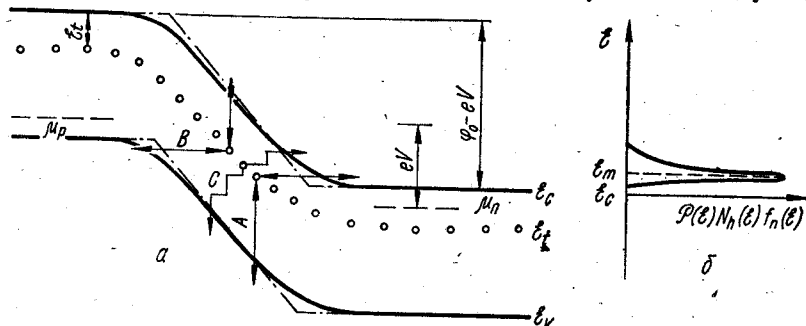


Рис. 4.5. Пути перехода носителей заряда в тонком p - n переходе с участием локальных уровней.

соображений. Поток электронов di_{nt} , туннелирующих из n области p - n перехода на локальные уровни в интервале энергий от ξ до $\xi + d\xi$, пропорционален числу электронов в n области $N_n(\xi) f_n(\xi)$, числу незанятых локальных уровней $N_t(\xi) [1 - f_t(\xi)]$ и вероятности туннелирования $\mathcal{P}_{nt}(\xi)$. Это позволяет записать

$$di_{nt}(\xi) = C \mathcal{P}_{nt}(\xi) N_n(\xi) f_n(\xi) N_t(\xi) [1 - f_t(\xi)] d\xi, \quad (4.69)$$

где C — коэффициент пропорциональности, $N_n(\xi)$ и $N_t(\xi)$ — плотности состояний; f_n и f_t — функции заполнения.

Аналогичным образом можно рассмотреть обратный ток из локальных уровней в зону проводимости и записать

$$di_{in}(\xi) = C \mathcal{P}_{in}(\xi) N_n(\xi) [1 - f_n(\xi)] N_t(\xi) f_t(\xi) d\xi. \quad (4.70)$$

При такой записи учтено, что ток с уровней в зону проводимости пропорционален количеству свободных мест в зоне проводимости $N_n(\xi) [1 - f_n(\xi)]$, количеству занятых локальных уровней $N_t(\xi) f_n(\xi)$ и вероятности обратного туннелирования $\mathcal{P}_{in}(\xi)$.

Результирующий ток в интервале энергий от ξ до $\xi + d\xi$ по пути A равен разности записанных потоков:

$$di = di_{nt} - di_{in}. \quad (4.71)$$

Полный туннельный ток зона проводимости — локальные уровни (путь A) можно получить, интегрируя последнее выражение, куда подставлены (4.69) и (4.70), по всем энергиям, для которых воз-

можны электронные переходы

$$i = C \int_{\xi} \mathcal{P}(\xi) [f_n(\xi) - f_t(\xi)] N_n(\xi) N_t(\xi) d\xi. \quad (4.72)$$

При записи последнего выражения учтено, что $\mathcal{P}_{n \rightarrow t} = \mathcal{P}_{t \rightarrow n}$.

Выражение (4.72) позволяет проанализировать общий характер зависимости тока от приложенного напряжения, когда туннельные токи зона проводимости — локальные уровни играют контролируемую роль для последовательных переходов электронов между зоной проводимости и валентной зоной по пути A . Кстати, если бы лимитирующей была стадия перехода электронов на этапе локальные уровни — валентная зона, задача нахождения токов была бы аналогична, рассмотренной в предыдущем параграфе.

Анализ выражения (4.72) показывает, что при отсутствии напряжения $f_n(\xi) = f_t(\xi)$ и $i = 0$.

При прямых напряжениях $f_n(\xi) > f_t(\xi)$ и прямой ток вначале растет с приложенным напряжением, что определяется ростом величины $\mathcal{P}(\xi)$. Затем происходит резкое уменьшение тока нуля, поскольку при достаточно больших прямых напряжениях, когда локальные уровни смещаются ниже уровня, соответствующего дну зоны проводимости n -типа полупроводника, в интервале энергий интегрирования $N_t(\xi) \rightarrow 0$.

При обратных напряжениях $f_t(\xi) > f_n(\xi)$ и туннельный ток изменяет направление. Характер изменения тока с приложенным напряжением при этом определяется изменением произведения величин $\mathcal{P}(\xi)$ и $f_t(\xi)$, где величина $f_t(\xi)$ уменьшается с ростом напряжения, а $f_t(\xi)$ растет. Для определения точного характера изменения $f_t(\xi)$ с напряжением необходимо учитывать обмен локальных уровней с валентной зоной, как это, например, сделано в предыдущем параграфе, что существенно усложняет задачу.

Аналогичным образом может быть рассмотрен туннельный ток валентная зона — локальные уровни по пути B . В результате можно показать, что

$$i = C \int_{\xi} \mathcal{P}(\xi) (f_t - f_p) N_p(\xi) N_t(\xi) d\xi, \quad (4.73)$$

где C — коэффициент пропорциональности; $\mathcal{P}(\xi)$ — вероятность туннелирования электрона на участке валентная зона — локальный уровень; f_p — функция заполнения состояний валентной зоны; $N_p(\xi)$ — плотность состояний в валентной зоне. Интегрирование выражения проводится по всем энергиям, где возможны туннельные переходы.

Анализ этого выражения приводит к результатам, аналогичным полученным выше при анализе выражения (4.72).

Так, прямой ток может с изменением напряжения иметь максимум, а обратный ведет себя сложным образом в зависимости от поведения величин $f_t(\xi)$ и $\mathcal{P}(\xi)$.

Вольт-амперная характеристика для прямого направления тока как для пути A , так и для пути B , когда $f_n \gg f_t$ и $f_p \gg f_t$, может быть найдена и в аналитическом виде.

Продолжаем это для туннельного тока по пути A . Для простоты заменим реальный потенциальный барьер треугольным барьером, как это показано пунктиром на рис. 4.5. Тогда в ВКБ-приближении, используя (2.22), можно получить

$$\mathcal{P} = \mathcal{P}_0 \exp\left(-\frac{4(2m^*)^{1/2} \xi_t^{3/2}}{3e\hbar E}\right), \quad (4.74)$$

где E — постоянное поле, равное среднему полю в области пространственного заряда $E = \frac{\varphi_0 - eV}{\mathcal{L}}$ (E — половина максимального поля).

Если подставить в выражение (4.74) значение E , то для \mathcal{L} , задаваемого выражением (1.49), можно получить, что

$$\mathcal{P} = \mathcal{P}_0 e^{-\sigma_\kappa (\varphi_0 - eV)^{1/2}}, \quad (4.75)$$

где

$$\sigma_\kappa = \frac{4}{3} \cdot \frac{(2m^*)^{1/2} \xi_t^{3/2}}{3e\hbar} \left[\frac{\varepsilon \varepsilon_0 (n_n + p_p)}{2e^2 n_n p_p} \right]^{1/2}. \quad (4.76)$$

Используя полученное значение $\mathcal{P}(\xi)$ и функцию Ферми для f_n , легко видеть, что основной переход носителей заряда осуществляется на уровни, расположенные в окрестности энергий вблизи дна зоны проводимости ξ_m (см. рис. 4.5). Это связано с тем, что \mathcal{P} при принятых приближениях не зависит от ξ , а $f_n(\xi)$ резко убывает. Поэтому можно переписать выражение (4.72) для прямого направления тока, когда $f_n \gg f_t$, в виде

$$i \sim \int_{\xi} \mathcal{P}(\xi) N_n(\xi) f_n(\xi_m) N_t(\xi) \delta(\xi + \xi_m) d\xi, \quad (4.77)$$

где $\delta(\xi + \xi_m)$ — дельта-функция.

Теперь после интегрирования можно получить, что

$$i \sim n_n N_t \mathcal{P}(\xi_m) \sim n_n N_t e^{-\sigma_\kappa (\varphi_0 - eV)^{1/2}} \quad (4.78)$$

при $\varphi_0 - eV - \xi_t > 0$; $i = 0$ при $\varphi_0 - eV - \xi_t < 0$.

При записи последнего выражения учтено, что $N(\xi_m) f(\xi_m) \sim n_n$ и $N_t(\xi_m) = N_t$ для равномерного распределения примесей.

Как видно из выражения (4.78), прямой ток вначале растет с приложенным напряжением и пропорционален $\mathcal{P}(\xi)$. Это изменение происходит более медленно, чем для диффузионных токов (см. п. 3.1). При дальнейшем увеличении напряжения, когда туннелирование на уровни становится невозможным, прямой ток уменьшается до нуля.

Особо отметим, что избыточный ток существенно растет с увеличением концентрации основных носителей заряда в полупроводнике. Это связано главным образом с уменьшением σ_κ и ростом величины \mathcal{P} с увеличением n_n . Поэтому особенно существенную роль избыточ-

ные токи могут играть для $p-n$ перехода, рассмотренного в следующем параграфе, когда n - и p -области становятся вырожденными.

Аналогичным способом можно рассмотреть и избыточный ток по пути B . В этом случае также ток пропорционален N_t и \mathcal{P} и зависит от напряжения и концентрации основных носителей заряда, как это рассмотрено выше.

Рассмотрение тока для переходов по пути C значительно сложнее. Однако и здесь можно предположить, что ток пропорционален N_t и суммарному коэффициенту прозрачности $\mathcal{P} = \mathcal{P}_1 \cdot \mathcal{P}_2 \cdot \mathcal{P}_3 \dots \approx \mathcal{P}_n$, где $\mathcal{P}_1, \mathcal{P}_2, \dots$ — коэффициенты прозрачности при туннельных переходах между двумя соседними уровнями. Это приводит к таким же качественным результатам, которые рассмотрены выше.

4.7. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА $p-n$ ПЕРЕХОДА МЕЖДУ ВЫРОЖДЕННЫМИ ПОЛУПРОВОДНИКАМИ

До сих пор рассматривались $p-n$ переходы для не очень больших концентраций доноров и акцепторов, когда n - и p -области были не вырожденными полупроводниками и уровень Ферми располагался в запрещенной зоне. Если увеличивать концентрацию как в p -, так и в n -области полупроводника, можно достичь вырождения, когда уровень Ферми будет располагаться в разрешенных зонах полупроводника. Схема $p-n$ перехода для вырожденных полупроводников изображена на рис. 4.6. Как видно, для этого $p-n$ перехода характерно наличие заполненного электронами интервала в зоне проводимости и заполненного дырками интервала в валентной зоне.

Кроме того, поскольку ширина области заряда зависит от концентрации примесей в полупроводнике, то для вырожденного полупроводника она становится значительно меньше, чем для невырожденного. Так, например, если принять за условия вырождения $n = N_c$ и $p = N_v$, то при комнатной температуре для вырожденного полупроводника при $m_n^* = m^* = m_0$ и $p \geq 3 \cdot 10^{19} \text{ см}^{-3}$. Используя выражение (1.18) для этого случая, можно показать, что $\mathcal{L} \leq 10^{-6} \text{ см}$. А для таких узких областей пространственного заряда возможно туннелирование носителей заряда из заполненных электронами областей n -типа или p -типа полупроводника в свободные места p -типа или n -типа полупроводника.

В то же время для вырожденного $p-n$ перехода резко уменьшаются токи над барьером области пространственного заряда вследствие того, что высота барьера велика ($\varphi_0 > \xi_2$) и соответственно p_n и p_p в формулах (4.21), (4.22) малы.

Все это приводит к тому, что вольт-амперная характеристика $p-n$ перехода с вырожденными полупроводниками по крайней мере

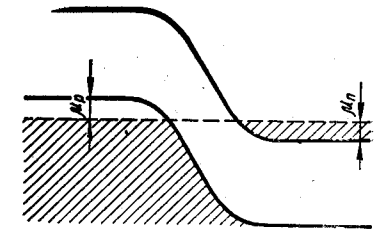


Рис. 4.6. Схема $p-n$ перехода между вырожденными полупроводниками.

при не очень больших прямых напряжениях определяется туннельным механизмом прохождения носителей заряда, а не рассмотренным выше диффузионным механизмом, связанным с прохождением носителей заряда над барьером области пространственного заряда.

Общее выражение для туннельного тока через $p-n$ переход может быть получено из следующих соображений. Поток электронов $di_{n,p}$ туннелирующих из n -области $p-n$ перехода в p -область и обладающих энергией в интервале от \mathcal{E} до $\mathcal{E} + d\mathcal{E}$, пропорционален числу электронов n -области, числу незапятных состояний в p -области и вероятности туннелирования. Это позволяет записать

$$di_{n \rightarrow p}(\mathcal{E}) = C \mathcal{P}_{n \rightarrow p} f_n(\mathcal{E}) N_n(\mathcal{E}) [1 - f_p(\mathcal{E})] N_p(\mathcal{E}) d\mathcal{E},$$

где C — коэффициент пропорциональности; $\mathcal{P}_{n,p}$ — вероятность туннелирования; $f_{p,n}$ — функции Ферми; $N(\mathcal{E})_{n,p}$ — плотность состояний в p - и n -областях.

Аналогичным образом можно рассмотреть поток электронов, туннелирующих из p - в n -область полупроводника и получить, что

$$di_{p \rightarrow n}(\mathcal{E}) = C \mathcal{P}_{p \rightarrow n} f_p(\mathcal{E}) N_p(\mathcal{E}) [1 - f_n(\mathcal{E})] N_n(\mathcal{E}) d\mathcal{E}.$$

Результирующий ток в интервале энергий от \mathcal{E} до $\mathcal{E} + d\mathcal{E}$ равен разности записанных потоков

$$di = di_{n \rightarrow p} - di_{p \rightarrow n}.$$

Полный туннельный ток через $p-n$ переход можно получить интегрируя последнее выражение по всем энергиям, для которых возможны электронные переходы:

$$i = eC \int_{\mathcal{E}} [f_n(\mathcal{E}) - f_p(\mathcal{E})] N_n(\mathcal{E}) N_p(\mathcal{E}) d\mathcal{E}. \quad (4.79)$$

При записи последнего выражения учтено, что $\mathcal{P}_{n \rightarrow p} = \mathcal{P}_{p \rightarrow n} = \mathcal{P}$.

Записанное выражение позволяет найти общий вид вольт-амперной характеристики $p-n$ перехода между вырожденными полупроводниками, не прибегая к детальным расчетам.

Для этого рассмотрим достаточно низкие температуры, когда функцию Ферми можно аппроксимировать ступенчатой функцией и $f_n = f_p = 1$ при $\mathcal{E} < \mu_n, \mu_p$, а $f_n = f_p = 0$ при $\mathcal{E} > \mu_n, \mu_p$.

Для напряжения $V = 0$ всегда $f_n = f_p$ и ток $i = 0$. При приложении напряжения увеличивающего высоту барьера уровни энергии в зонах смещаются так, что уровень Ферми в n -области понижается по отношению к уровню Ферми p -области и против заполненных мест валентной зоны оказываются свободные места зоны проводимости. Это приводит к $f_p > f_n$ и появлению электронного тока от p -области к n -области. Такой ток растет с увеличением приложенного напряжения из-за роста разности $f_n - f_p$.

При приложении напряжения, уменьшающего высоту потенциального барьера, уровень Ферми в n -типе полупроводника поднимается по отношению уровня Ферми в p -типе полупроводника. Когда напряжения малы, это приводит к тому, что против заполненных уровней n -типа оказывается некоторое количество свободных

мест в p -типе и $f_n > f_p$. При этом появляется электронный ток, направленный от n - к p -области полупроводника. Этот ток будет расти с приложенным напряжением до тех пор, пока уровень Ферми в n -типе не достигнет верха валентной зоны. При дальнейшем увеличении напряжения начнется уменьшение тока из-за того, что для части электронов зоны проводимости уже нет мест, куда можно туннелировать, т. е. уменьшаются величины f_n и $N_p(\mathcal{E})$.

Когда приложенное напряжение приведет к тому, что верх валентной зоны окажется ниже дна зоны проводимости, туннельный

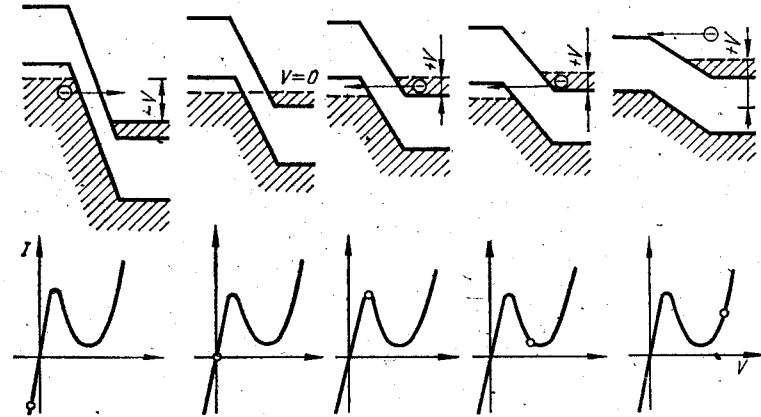


Рис. 4.7. Энергетическая диаграмма и вольт-амперная характеристика туннельно прозрачного $p-n$ перехода при различных напряжениях.

ток вообще прекращается, поскольку $N_p(\mathcal{E}) = 0$. При дальнейшем увеличении напряжения начнут проявляться надбарьерные переходы носителей заряда и ток снова будет расти. Все это показано на рис. 4.7; на котором представлены токи для различных случаев, рассмотренных выше. Как видно из этого рисунка, вольт-амперная характеристика $p-n$ перехода между вырожденными полупроводниками характеризуется резким ростом тока для обратного напряжения и экстремумом тока для прямого напряжения. Соответственно дифференциальное сопротивление рассмотренного $p-n$ перехода положительно для обратного напряжения и начального участка прямого напряжения и отрицательно при достаточно больших прямых напряжениях. Эта особенность вольт-амперной характеристики $p-n$ перехода между вырожденными полупроводниками является основой применения для так называемых туннельных диодов.

Для нахождения аналитического выражения вольт-амперной характеристики необходимо иметь выражение для \mathcal{P} , $f_{n,p}$ и $N_{n,p}$. Точное выражение для $\mathcal{P}(\mathcal{E})$ неизвестно. Однако при интегрировании можно воспользоваться тем обстоятельством, что для небольших интервалов энергий значение $\mathcal{P}(\mathcal{E})$ приблизительно постоянно и его можно вынести из-под знака интеграла.

Сама же величина \mathcal{P} может быть найдена для случая, когда потенциальный барьер для туннелирующих носителей заряда может быть аппроксимирован треугольным барьером. Тогда в ВКБ-приближении:

$$\mathcal{P} = \mathcal{P}_0 \exp\left(-\frac{4\sqrt{2m^*} \xi_0^{3/2}}{3e\hbar E}\right), \quad (4.80)$$

где $E = \frac{\Phi_0 - eV}{\mathcal{P}}$ — напряженность постоянного поля (это соответствует аппроксимации, рассмотренной в предыдущем параграфе). После подстановки E величина \mathcal{P} может быть переписана в виде

$$\mathcal{P} = \mathcal{P}_0 e^{-\sigma_{\kappa}(\Phi_0 - eV)^{1/2}}, \quad (4.81)$$

где

$$\sigma_{\kappa} = \frac{4\sqrt{2m^*} \xi_0^{3/2}}{3e\hbar} \left[\frac{e_0 E (n_n + p_p)}{2e^2 n_n p_p} \right]^{1/2}.$$

Как видно из последнего выражения, величина \mathcal{P} существенно зависит от концентрации примесей в p - и n -областях полупроводника, а также приложенного напряжения.

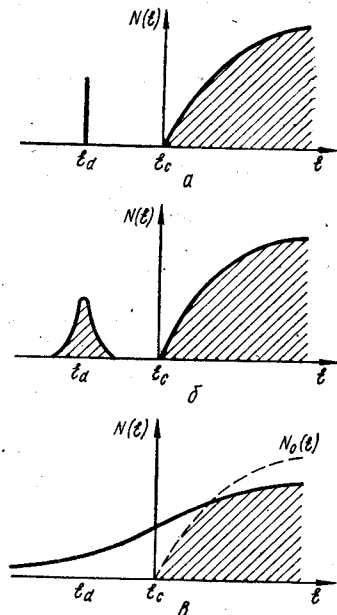


Рис. 4.8. Плотность состояний в слабо легированном (а), промежуточно легированном (б) и сильно легированном (в) полупроводнике (----- параболическая аппроксимация).

Сложнее положение с величинами N_n и N_p . Они неизвестны для вырожденных полупроводников. Это связано с двумя причинами. Во-первых, в вырожденных полупроводниках примесные уровни образуют примесную зону, во-вторых, в сильно легированных полупроводниках образуются так называемые хвосты плотности состояний. Это приводит к изменению хода плотности состояний, так как это показано на рис. 4.8.

Кроме того, подстановка полных значений f_n и f_p в интеграл (4.79) дает неберущиеся в аналитическом виде интегралы.

В силу изложенного точный аналитический вид вольт-амперной характеристики получить нельзя.

Приближенное аналитическое выражение вольт-амперной характеристики может быть получено, если аппроксимировать N_n и N_p выражениями, справедливыми для невырожденного полупроводника (пунктирная линия на рис. 4.8, в):

$$N_n \sim \begin{cases} (\xi - \xi_c)^{1/2} & \text{для } \xi > \xi_c; \\ 0 & \text{для } \xi < \xi_c; \end{cases} \quad (4.82)$$

$$N_p \sim \begin{cases} (\xi_v - \xi)^{1/2} & \text{для } \xi < \xi_v; \\ 0 & \text{для } \xi > \xi_v. \end{cases} \quad (4.83)$$

и предположить, что вырождение не очень велико так, что $\mu_n \leq 2kT$ и $\mu_p \leq 2kT$.

Кроме того, предположим, что функции Ферми $f_n = \frac{1}{1 + e^{\frac{\xi - \mu_n - eV}{kT}}}$ и $f_p = \frac{1}{1 + e^{\frac{\mu_p - \xi}{kT}}}$ могут быть разложены по малому параметру $\frac{\xi - \mu_n - eV}{kT}$ и $\frac{\mu_p - \xi}{kT}$ (здесь для n области учитывается сдвиг уровня Ферми на величину eV по сравнению с уровнем Ферми p -области).

При этом функции Ферми приобретают вид:

$$f_n = \frac{1}{2} - \frac{\xi - (\mu_n + eV)}{4kT}; \quad (4.84)$$

$$f_p = \frac{1}{2} + \frac{\mu_p - \xi}{4kT}. \quad (4.85)$$

Подставляя выражение (4.82) — (4.85) в интеграл (4.79), получим

$$i = C_1 \int_{\xi_c + eV}^{\xi_v} \frac{eV}{4kT} (\xi - \xi_c + eV)^{1/2} (\xi_v - \xi)^{1/2} d\xi. \quad (4.86)$$

Последнее выражение можно проинтегрировать.

В результате проведенных приближенных расчетов для случая, когда вырождение не очень велико, а $N_n(\xi)$ и $N_p(\xi)$ можно аппроксимировать выражениями, справедливыми для невырожденного полупроводника, для туннельных токов получаем следующее выражение вольт-амперной характеристики:

$$i = C_2 V (\mu_n + \mu_p - eV)^2. \quad (4.87)$$

Дифференцируя выражение по напряжению (4.87), можно получить

$$\frac{di}{dV} = C_2 (\mu_n + \mu_p - eV) (\mu_n + \mu_p - 3eV). \quad (4.88)$$

Как видно из последнего выражения, дифференциальная проводимость с ростом напряжения изменяет знак.

Напряжение, соответствующее максимуму тока, может быть получено при $\frac{di}{dV} = 0$:

$$V_{\max} = \frac{1}{3e} (\mu_n + \mu_p). \quad (4.89)$$

Как видно, это напряжение растет с ростом μ_n и μ_p , т. е. с ростом легирования n - и p -области.

На падающем участке вольт-амперной характеристики дифференциальное сопротивление имеет минимум. Напряжение, соответствующее минимуму отрицательного сопротивления равно

$$V_{\min} = \frac{2}{3e} (\mu_n + \mu_p) = 2V_{\max} \quad (4.90)$$

и также растет при легировании p - и n -области полупроводника.

Используя полученные выше результаты настоящего и предыдущего параграфов, можно найти полную вольт-амперную характеристику p - n перехода между вырожденными полупроводниками (рис. 4.9).

Как видно из рисунка, в общем случае прямой ток после прохождения максимума не устремляется к нулю. Это связано с наличием избыточного тока через локальные уровни в запрещенной зоне, рассмотренного в предыдущем параграфе. При этом может наблюдаться еще один максимум тока. Затем ток начинает расти с приложенным напряжением. Этот рост определяется обычным

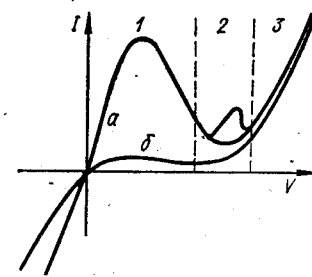


Рис. 4.9. Суммарная вольт-амперная характеристика для p - n перехода между вырожденными полупроводниками (а — туннельный диод, б — обращенный диод):

1 — туннельный ток; 2 — избыточный ток; 3 — диффузионный ток.

надбарьерным механизмом прохождения носителей заряда, рассмотренным ранее. Обратный ток растет согласно рассмотренному выше туннельному механизму. Следует особое внимание обратить на то, что на различных участках вольт-амперной характеристики проявляются существенно различные механизмы прохождения тока. И, если на участках 1 и 2 ток идет за счет межзонных переходов, то на участке 3 за счет переходов в пределах одной и той же разрешенной зоны.

Особым случаем рассмотренной здесь картины является случай, когда вырождение в одном из полупроводников слабое. Тогда, пик тока в прямом направлении уменьшается и может становиться несущественным. При этом как бы меняется прямое и обратное направление напряжения для $p-n$ перехода и соответствующий диод называется обращенным. Вольт-амперная характеристика обращенного диода также представлена на рис. 4.9.

4.8. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ $p-n$ ПЕРЕХОДОВ

Экспериментальные исследования выпрямляющих свойств $p-n$ переходов проводились на большом количестве полупроводниковых материалов как для слабо легированных, так и вырожденных p - и n -областей. Исследовались $p-n$ переходы, изготовленные самыми различными способами, в частности, $p-n$ переходы, полученные путем диффузии доноров в p -полупроводник или акцепторов в n -полупроводник, сплавлением полупроводника с металлом, содержащим примеси донорного или акцепторного типа, путем ионной имплантации примесей, создающих $p-n$ переход и т. д.

Наибольшее количество экспериментальных исследований посвящено германиевым и кремниевым $p-n$ переходам, поскольку в настоящее время они находят самое широкое практическое применение. Исследования германиевых и кремниевых $p-n$ переходов проводились как при малых, так и больших уровнях инжекции.

Условием малого уровня инжекции является для дырочных токов выполнение неравенства

$$p(\mathcal{L}_n) \ll n_n.$$

После подстановки в это неравенство $p(\mathcal{L}_n)$, найденного из уравнения (4.4), оно может быть переписано в виде

$$\frac{i_p l_p}{e D_p} \ll n_n. \quad (4.91)$$

Аналогичное неравенство для случая, когда ток через $p-n$ переход является электронным, имеет вид

$$\frac{i_n l_n}{e D_n} \ll p_p. \quad (4.92)$$

При выполнении обратных неравенств уровень инжекции будет большим.

Для резких германиевых $p-n$ переходов при малых уровнях инжекции экспериментальные данные обычно хорошо описываются

простой теорией, учитывающей прохождение электронов и дырок над барьером области пространственного заряда. Это показано на рис. 4.10, где показана типичная вольт-амперная характеристика германиевого $p-n$ перехода для токов, удовлетворяющих условиям малых уровней инжекции. Как видно из этого рисунка, прямая ветвь вольт-амперной характеристики хорошо описывается экспоненциальным законом и прямой ток пропорционален $e^{\frac{eV}{kT}}$, как это

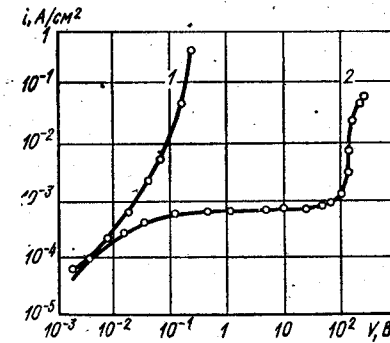


Рис. 4.10. Вольт-амперная характеристика германиевого $p-n$ перехода (сплошные линии рассчитаны теоретически по формуле (4.20)): 1 — прямой ток; 2 — обратный ток.

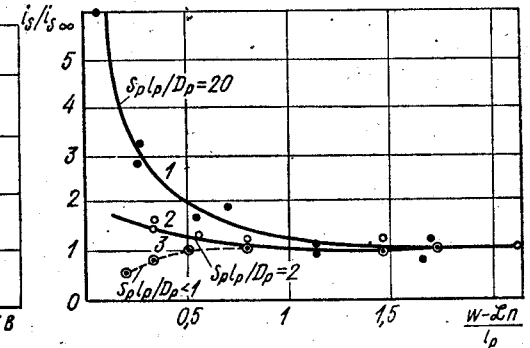


Рис. 4.11. Экспериментальные зависимости $\frac{i_s}{i_{s\infty}}$ германиевых $p-n$ переходов от $\frac{w-L_n}{l_p}$ для разных значений $\frac{S_p l_p}{D_p}$.

и следует из формулы (4.20). Обратный ток вплоть до пробоя испытывает насыщение, что также соответствует формуле (4.22).

Величина i_s изменяется с изменением уровня легирования p - и n -областей пропорционально p_n или n_p и зависит от диффузионной длины, как это и предсказывается теорией.

Для $p-n$ переходов с узкой базой проявляется зависимость величины i_s от ширины базы и скорости поверхностной рекомбинации. Это иллюстрируется рис. 4.11, на котором представлена зависимость величины i_s от ширины базы при разных значениях величины $\frac{S_p l_p}{D_p}$. Как видно из этого рисунка, для $\frac{S_p l_p}{D_p} > 1$ величина i_s растет, а для $\frac{S_p l_p}{D_p} < 1$ убывает с уменьшением ширины

базы, как это и следует из рассмотренной выше теории, для токов без учета генерации — рекомбинации и других эффектов.

Для кремниевых $p-n$ переходов между невырожденными полупроводниками вольт-амперная характеристика для малых уровней инжекции, как правило, не описывается выражением (4.20).

На рис. 4.12 показана типичная прямая ветвь вольт-амперной характеристики, а на рис. 4.13 — обратная ветвь вольт-амперной характеристики кремниевых $p-n$ перехода при малых уровнях инжекции. Как можно видеть, для прямой ветви характерно наличие

двух участков. При малых напряжениях прямой ток пропорционален $e^{eV/2kT}$, а при больших становится пропорциональным $e^{eV/kT}$. Обратный ток во всем интервале вплоть до начала пробоя пропорционален $\left(\frac{\Phi_0 - eV}{\Phi_0}\right)^{1/2} \sim \mathcal{L}$. Такой вид вольт-амперной характеристики может быть объяснен тем, что для обратных и небольших прямых напряжений для кремниевых $p-n$ переходов существенную роль играют генерационно-рекомбинационные токи. Вольт-амперная характеристика при этом описывается выражениями (4.61) и (4.67). При больших прямых напряжениях начинает идти

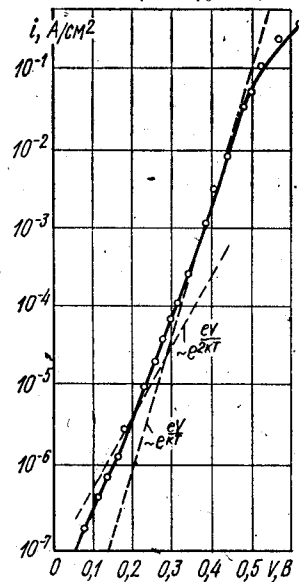


Рис. 4.12. Прямая ветвь вольт-амперной характеристики кремниевое $p-n$ перехода.

ток над барьером области пространственного заряда, который описывается выражением (4.20). Убедительным доказательством такого объяснения обратных и небольших прямых токов являются опыты, в которых специально вводились глубокие примеси в область $p-n$ перехода. При этом показано, что обратные токи существенно увеличиваются с ростом концентрации вводимой примеси.

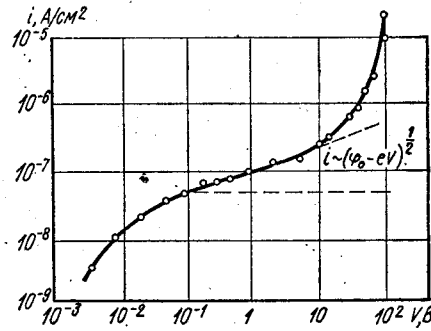


Рис. 4.13. Обратная ветвь вольт-амперной характеристики кремниевое $p-n$ перехода.

Наблюдаемое экспериментально отличие кремниевых и германиевых $p-n$ переходов определяется большей шириной запрещенной зоны кремния по сравнению с шириной запрещенной зоны германия. Это приводит к тому, что собственная концентрация n_i в кремнии меньше, чем в германии.

Отношение

$$\frac{(n_i)_{\text{Ge}}}{(n_i)_{\text{Si}}} = \left[\frac{(m_n^* m_p^*)_{\text{Ge}}}{(m_n^* m_p^*)_{\text{Si}}} \right]^{1/2} e^{\frac{\mathcal{E}_{\text{Si}} - \mathcal{E}_{\text{Ge}}}{2kT}}$$

Если сюда подставить соответствующие значения, то можно получить, что $\frac{(n_i)_{\text{Ge}}}{(n_i)_{\text{Si}}} = 1,7 \cdot 10^3$. В результате надбарьерные

токи для кремния уменьшаются по сравнению с токами для германия. Это изменение может быть рассчитано для дырочного тока из соотношения

$$\frac{i_{\text{Ge}}}{i_{\text{Si}}} = \frac{(D_p)_{\text{Ge}}}{(D_p)_{\text{Si}}} \cdot \frac{(l_p)_{\text{Si}}}{(l_p)_{\text{Ge}}} \left(\frac{n_{i\text{Ge}}}{n_{i\text{Si}}} \right)^2 \left(\frac{n_{n\text{Si}}}{n_{n\text{Ge}}} \right),$$

которое легко получается на основании выражения (4.20). Подставляя в последнюю формулу соответствующие значения, получаем, что при

$$\frac{(D_p)_{\text{Ge}}}{(D_p)_{\text{Si}}} \cdot \frac{(l_p)_{\text{Si}}}{(l_p)_{\text{Ge}}} \cdot \frac{(n_{n\text{Si}})}{(n_{n\text{Ge}})} \approx 1$$

$$\frac{i_{\text{Ge}}}{i_{\text{Si}}} = 3 \cdot 10^6.$$

Отношение генерационно-рекомбинационных токов для германия и кремния определяется выражением

$$\frac{(i_{r-p})_{\text{Ge}}}{(i_{r-p})_{\text{Si}}} = \frac{\mathcal{L}_{\text{Ge}} (\tau_{p0} + \tau_{n0})_{\text{Si}}}{\mathcal{L}_{\text{Si}} (\tau_{p0} + \tau_{n0})_{\text{Ge}}} \cdot \frac{n_{i\text{Ge}}}{n_{i\text{Si}}}.$$

При $\frac{\mathcal{L}_{\text{Ge}}}{\mathcal{L}_{\text{Si}}} \frac{(\tau_{p0} + \tau_{n0})_{\text{Si}}}{(\tau_{p0} + \tau_{n0})_{\text{Ge}}} \approx 1$ это дает $\frac{(i_{r-p})_{\text{Ge}}}{(i_{r-p})_{\text{Si}}} = 1,7 \cdot 10^3$.

Таким образом, для кремниевых диодов роль генерационно-рекомбинационных токов резко возрастает, что и наблюдается экспериментально.

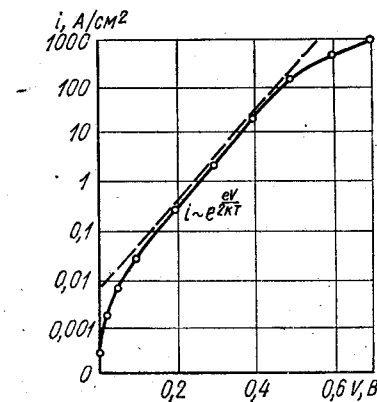


Рис. 4.14. Прямая ветвь вольт-амперной характеристики германиевого диода при больших уровнях инжекций.

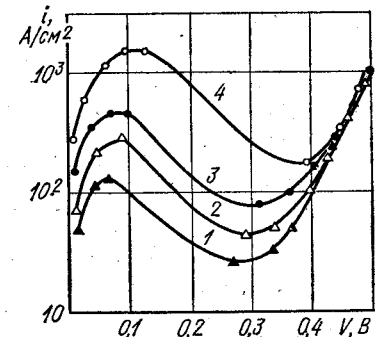


Рис. 4.15. Вольт-амперная характеристика германиевого $p-n$ перехода между вырожденными полупроводниками при разных уровнях легирования 1 — $p = 1,5 \cdot 10^{19} \text{ см}^{-3}$; 2 — $p = 2,5 \cdot 10^{19} \text{ см}^{-3}$; 3 — $p = 3,5 \cdot 10^{19} \text{ см}^{-3}$; 4 — $p = 5,3 \cdot 10^{19} \text{ см}^{-3}$.

Аналогичное возрастание генерационно-рекомбинационных токов наблюдается и для других полупроводников с широкой шириной запрещенной зоны, например GaAs, GaP и др.

При больших уровнях инжекции теория предсказывает рост прямого тока пропорциональный $e \frac{eV}{2kT}$, а затем замедление роста тока с приложенным напряжением. Это и наблюдается экспериментально. Типичная экспериментальная вольт-амперная характеристика, подтверждающая сказанное, показана на рис. 4.14.

Особое внимание было уделено исследованию $p-n$ перехода между вырожденными полупроводниками. Типичные вольт-амперные характеристики германиевых $p-n$ переходов между вырожденными n - и p -областями показаны на рис. 4.15. Как видно из приведенного рисунка, ток при росте прямого напряжения вначале растет, затем падает и снова растет, как это и предсказано теорией, рассмотренной в предыдущих параграфах. Напряжение, при котором наблюдается максимум, смещается в сторону больших напряжений при увеличении степени легирования p -области полупроводника, что также предсказано теорией (см. формулу (4.89)).

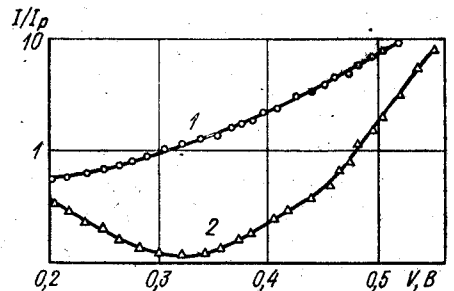


Рис. 4.16. Избыточный ток I/I_p германиевых $p-n$ переходов между вырожденными полупроводниками с примесью меди (1) и никеля (2).

Экспериментальное значение минимального тока не равно нулю, что можно объяснить привлечением избыточных токов. Эти токи должны зависеть от концентрации глубоких уровней в области пространственного заряда. И, действительно, специальное введение глубоких уровней в область пространственного заряда показало существенное изменение токов в минимуме, как это показано на рис. 4.16.

Таким образом, разработанная к настоящему времени теория $p-n$ переходов между невырожденными и вырожденными p - и n -областями хорошо объясняет все наблюдаемые экспериментально закономерности.

Глава 5. ВЫПРЯМЛЯЮЩИЕ СВОЙСТВА ГЕТЕРОПЕРЕХОДОВ

5.1. ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ ИЗОТИПНЫХ ГЕТЕРОПЕРЕХОДОВ

Рассмотрим вольт-амперные характеристики изотипных гетеропереходов на примере $n-n$ гетероперехода. Вначале не будем учитывать тока через поверхностные уровни, туннелирование через область пространственного заряда и силы изображения. Это даст возможность рассмотреть только токи основных носителей заряда в пределах одного вида разрешенных зон над соответствующими потенциальными барьерами.

Примем также для определенности здесь и далее, что в полупроводнике 2 (рис. 5.1) всегда осуществляется заборный изгиб зон. Тогда все возможные конфигурации потенциальных барьеров в рассматриваемом гетеропереходе ограничиваются случаями, представленными на рис. 5.1.

На рис. 5.1, а, б показаны барьеры, образованные за счет контактной разности потенциалов. При этом, как уже обсуждалось в гл. 1, если в полупроводнике 2 задан заборный изгиб зон, то в полупроводнике 1 реализуется антизаборный. Такие гетеропереходы в дальнейшем будут именоваться гетеропереходами вида антизаборный — заборный слой.

На рис. 5, в, г показаны барьеры, которые образованы за счет отрицательного заряда поверхностных состояний. В этом случае реализуется одновременно заборный слой в полупроводниках 2 и 1. Такие гетеропереходы в дальнейшем будут называться гетеропереходами вида заборный — заборный слой.

Модели гетеропереходов рис. 5.1, а, в отличаются от моделей рис. 5.1, б, г еще и тем, что в первом случае после прохождения через потенциальный барьер полупроводника 2 носитель заряда преодолевает дополнительный барьер, связанный с разрывом в зонах проводимостей полупроводников 1 и 2, а во втором не преодолевает такого барьера. Если в гетеропереходе носитель заряда при движении от полупроводника 2 к полупроводнику 1 преодолевает дополнительный потенциальный барьер, он будет в дальнейшем называться гетеропереходом с положительным скачком (рис. 5.1, а, в). Если в гетеропереходе носитель заряда при движении от полупроводника 2 к полупроводнику 1 не преодолевает дополнительного барьера, он будет называться гетеропереходом с отрицательным скачком (рис. 5.1, б, г).

Обратим особое внимание на то, что на рис. 5.1 изображены гетеропереходы при отсутствии приложенного напряжения. При

приложении внешнего напряжения высоты потенциальных барьеров в полупроводниках 1 и 2 будут изменяться, и в гетеропереходе может изменяться вид скачка. Поэтому понятия «отрицательный» или «положительный» скачок относятся только к определенному интервалу приложенных напряжений. Об этом необходимо всегда помнить при рассмотрении конкретных вольт-амперных характеристик гетероперехода в предположениях диодной или диффузионной теорий.

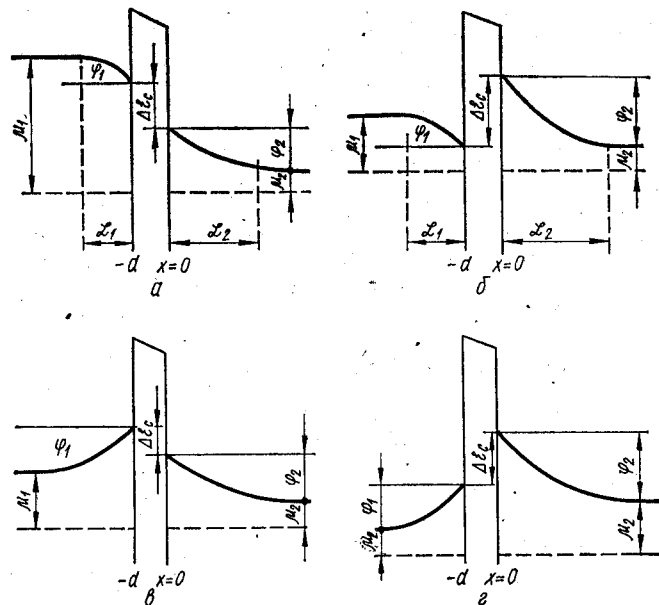


Рис. 5.1. Энергетические диаграммы изотипных и анизотипных гетеропереходов.

Диодная теория. Задача нахождения вольт-амперной характеристики в предположениях диодной теории, как известно, сводится к нахождению токов термоэлектронной эмиссии над потенциальными барьерами полупроводников 1 и 2, а затем и нахождению разности этих токов: $i = i_2 - i_1$ (здесь и далее за положительное напряжение тока по аналогии с контактом металл-полупроводник будет принято направление тока электронов справа налево, т. е. через заборный слой полупроводника 2, который понижается приложенным напряжением, к полупроводнику 1).

В связи с изложенным модель изотипного гетероперехода, рассматриваемая в предположениях диодной теории, в литературе называется эмиссионной моделью.

Используя ту же методику, что и при расчете вольт-амперной характеристики контакта металл-полупроводник для случая, когда уровень Ферми не пересекает разрешенных зон и нет вырождения, легко показать, что для гетероперехода вида антизаборный — за-

борный слой при положительном скачке

$$i = AT^2 \mathcal{P}_n e^{-\frac{\mu_1}{kT}} \left(e^{\frac{eV}{kT}} - 1 \right) = \frac{eV_1 n_1}{4} \mathcal{P}_n \left(e^{\frac{eV}{kT}} - 1 \right). \quad (5.1)$$

Для гетероперехода вида антизаборный — заборный слой и вида заборный — заборный слой при отрицательном скачке легко найти

$$i = AT^2 \mathcal{P}_n e^{-\frac{\Phi_2 + \mu_2}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{e(V_1 + V_3)}{kT}} \right) = \frac{eV_2 n_2 \mathcal{P}_n}{4} e^{-\frac{\Phi_2}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{e(V_1 + V_3)}{kT}} \right). \quad (5.2)$$

Для гетероперехода вида заборный — заборный слой при положительном скачке расчет приводит к вольт-амперной характеристике вида

$$i = AT^2 \mathcal{P}_n e^{-\frac{\Phi_1 + \mu_1}{kT}} \left(e^{\frac{V_2 + V_3}{kT}} - e^{-\frac{eV_1}{kT}} \right) = \frac{eV_1 n_1 \mathcal{P}_n}{4} e^{-\frac{\Phi_1}{kT}} \left(e^{\frac{V_2 + V_3}{kT}} - e^{-\frac{eV_1}{kT}} \right). \quad (5.3)$$

В выражениях (5.1) — (5.3) и далее в этой главе V_1 , V_3 и V_2 — части напряжения, падающие на полупроводнике 1, промежуточном слое и полупроводнике 2, рассмотренные в гл. 2. Положительный знак этих напряжений соответствует понижению общей высоты потенциального барьера заборного слоя полупроводника 2 и повышению или понижению потенциального барьера для полупроводника 1 в зависимости от направления первоначального изгиба зон.

Диффузионная теория. В предположениях диффузионной теории необходимо решать диффузионно-дрейфовое уравнение в полупроводниках 1 и 2, а затем сшивать полученные выражения. В связи с этим в литературе при рассмотрении диффузионной теории изотипного гетероперехода обычно говорят о диффузионной модели гетероперехода. Используя полученные ранее результаты при рассмотрении контакта металл — полупроводник и пренебрегая генерацией и рекомбинацией носителей в области пространственного заряда, получим

$$n_1(-d) = \left[n_1(-\mathcal{L}_1 - d) - \frac{i}{eD_1} \int_{-d}^{-\mathcal{L}_1 - d} e^{\frac{\Phi}{kT}} dx \right] e^{\pm \frac{\Phi_1 + eV_1}{kT}}, \quad (5.4)$$

где верхний знак в показателе последней экспоненты относится к антизаборному, а нижний к заборному изгибу зон:

$$n_2(0) = \left[n_2(\mathcal{L}_2) + \frac{i}{eD_2} \int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx \right] e^{-\frac{\Phi_2 - eV_2}{kT}}. \quad (5.5)$$

Рассматривая разность потоков носителей заряда через промежуточный слой гетероперехода вида антизаборный — заборный

и заборный — заборный слой с отрицательным скачком, когда уровень Ферми не пересекает разрешенных зон, можно записать

$$i = \frac{e\mathcal{P}n}{4} (v_2 n_2(0) - v_1 n_1(-d)) e^{-\frac{\Delta\delta_c}{kT}} e^{-\frac{eV_2}{kT}}. \quad (5.6)$$

Аналогичным образом для гетеропереходов вида антизаборный — заборный и заборный — заборный слой при положительном скачке

$$i = \frac{e\mathcal{P}n}{4} (v_2 n_2(0) e^{-\frac{\Delta\delta_c}{kT}} e^{\frac{eV_2}{kT}} - v_1 n_1(-d)). \quad (5.7)$$

Используя записанную систему уравнений, легко найти общее выражение вольт-амперной характеристики. Для этого можно подставить $n_2(0)$ и $n_1(-d)$ из выражений (5.4) и (5.5) в выражения (5.6) и (5.7) и решить полученные уравнения относительно тока.

В результате для гетеропереходов с отрицательным скачком

$$i = \frac{e\mathcal{P}n}{4} \cdot \frac{v_2 n_2 e^{-\frac{\Phi_2 - eV_2}{kT}} - v_1 n_1 e^{-\frac{\Delta\delta_c}{kT}} e^{\pm \frac{\Phi_1 \mp eV_1}{kT}} e^{-\frac{eV_2}{kT}}}{1 - \frac{e^2 v_2 \mathcal{P}n}{4u_2 kT} e^{-\frac{\Phi_2 - eV_2}{kT}} \int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx + \frac{e^2 v_1 \mathcal{P}n}{4u_1 kT} e^{-\frac{\Delta\delta_c + eV_2}{kT}} e^{\pm \frac{\Phi_1 \mp eV_1}{kT}} \int_{-d}^{-\mathcal{L}_1 - d} e^{\frac{\Phi}{kT}} dx}, \quad (5.8)$$

а для гетеропереходов с положительным скачком

$$i = \frac{e\mathcal{P}n v_2 n_2 e^{-\frac{\Delta\delta_c}{kT}} e^{-\frac{\Phi_2 - eV_2}{kT}} e^{\frac{eV_2}{kT}} - v_1 n_1 e^{\pm \frac{\Phi_1 - eV_1}{kT}}}{4 \left(1 - \frac{e^2 v_2 \mathcal{P}n}{4u_2 kT} e^{-\frac{\Delta\delta_c}{kT}} e^{\frac{eV_2}{kT}} e^{-\frac{\Phi_2 - eV_2}{kT}} \int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx + \frac{e^2 v_1 \mathcal{P}n}{4u_1 kT} e^{\pm \frac{\Phi_1 \mp eV_1}{kT}} \int_{-d}^{-\mathcal{L}_1 - d} e^{\frac{\Phi}{kT}} dx \right)}. \quad (5.9)$$

Поскольку при отсутствии напряжения ток через гетеропереход равен нулю, то из выражений (5.8) и (5.9) можно получить

$$v_2 n_2 e^{-\frac{\Phi_2}{kT}} = v_1 n_1 e^{\pm \frac{\Phi_1}{kT}} e^{-\frac{\Delta\delta_c}{kT}}, \quad (5.10a)$$

или

$$v_2 n_2 e^{-\frac{\Phi_2}{kT}} e^{-\frac{\Delta\delta_c}{kT}} = v_1 n_1 e^{\pm \frac{\Phi_1}{kT}}. \quad (5.10b)$$

Это позволяет переписать числитель выражений (5.8) и (5.9) через параметры одного из полупроводников, например полупроводника 2. В результате полу-

чаем вместо выражения (5.8)

$$i = \frac{e v_2 n_2 \mathcal{P}n}{4} e^{-\frac{\Phi_2}{kT}} e^{-\frac{\Delta\delta_c}{kT}} \times \frac{e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} e^{-\frac{eV_2}{kT}}}{1 - \frac{e^2 v_2 \mathcal{P}n}{4u_2 kT} e^{-\frac{\Phi_2 - eV_2}{kT}} \int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx + \frac{e^2 v_1 \mathcal{P}n}{4u_1 kT} e^{-\frac{\Delta\delta_c + eV_2}{kT}} e^{\pm \frac{\Phi_1 \mp eV_1}{kT}} \int_{-d}^{-\mathcal{L}_1 - d} e^{\frac{\Phi}{kT}} dx} \quad (5.11)$$

и вместо выражения (5.9)

$$i = \frac{e v_2 n_2 \mathcal{P}n}{4} e^{-\frac{\Phi_2}{kT}} e^{\frac{\Delta\delta_c}{kT}} \times \frac{(e^{\frac{eV_2}{kT}} e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}})}{1 - \frac{e^2 v_2 \mathcal{P}n}{4u_2 kT} e^{-\frac{\Delta\delta_c}{kT}} e^{\frac{eV_2}{kT}} e^{-\frac{\Phi_2 - eV_2}{kT}} \int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx + \frac{e^2 v_1 \mathcal{P}n}{4u_1 kT} e^{\pm \frac{\Phi_1 \mp eV_1}{kT}} \int_{-d}^{-\mathcal{L}_1 - d} e^{\frac{\Phi}{kT}} dx}. \quad (5.12)$$

Окончательные выражения вольт-амперной характеристики могут быть записаны после нахождения интегралов в знаменателе выражений (5.11) и (5.12).

Для простоты предположим, что последними членами знаменателя выражений (5.11) и (5.12) можно пренебречь. Это означает, что выполняется соотношение

$$n_1(-\mathcal{L}_1 - d) \gg \frac{i}{eD_1} \int_0^{-\mathcal{L}_1 - d} e^{\frac{\Phi(x)}{kT}} dx.$$

Такое предположение, как видно из выражения (5.4), эквивалентно предположению о том, что распределение концентрации в области пространственного заряда полу-

проводника 1 описывается законом Больцмана и $n_1(-\mathcal{L}_1 - d) = n(-d) e^{\pm \frac{\Phi_1 \mp eV_1}{kT}}$. Использование закона Больцмана для $n_1(-\mathcal{L}_1 - d)$ позволяет упростить окончательное выражение вольт-амперной характеристики без ущерба для общности полученных результатов. Действительно, с одинаковым правом можно было бы предположить, что можно пренебречь предпоследними членами выражений (5.11) и (5.12) (это означало бы выполнение закона Больцмана в полупроводнике 2).

При принятых предположениях нахождение интеграла для слоя Шоттки в полупроводнике аналогично проведенному выше для контакта металл — полупроводник и

$$\int_0^{\mathcal{L}_2} e^{\frac{\Phi}{kT}} dx \simeq -\frac{kT}{eE_2^0} e^{\frac{\Phi_2 - eV_2}{kT}}.$$

Подставляя этот интеграл в записанные выражения вольт-амперной характеристики, получаем вместо выражения (5.11) для гетеропереходов с отрицательным скачком

$$i = \frac{en_2v_2\mathcal{P}_n}{4} e^{-\frac{\varphi_2}{kT}} e^{-\frac{\Delta\mathcal{E}_c}{kT}} \frac{E_2^0u_2}{\left(E_2^0u_2 + \frac{v_2\mathcal{P}_n}{4}\right)} \times \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} e^{-\frac{eV_3}{kT}}\right), \quad (5.13)$$

а вместо выражения (5.12) для гетеропереходов с положительным скачком

$$i = \frac{e^2n_2v_2\mathcal{P}_n}{4} e^{-\frac{\varphi_2}{kT}} e^{-\frac{\Delta\mathcal{E}_c}{kT}} \frac{E_2^0u_2}{E_2^0u_2 + \frac{v_2\mathcal{P}_n}{4}} e^{-\frac{\Delta\mathcal{E}_c}{kT}} e^{\frac{eV_3}{kT}} \times \left(e^{\frac{eV_2}{kT}} e^{\frac{eV_3}{kT}} - e^{-\frac{eV_1}{kT}}\right). \quad (5.14)$$

Анализ выражений вольт-амперной характеристики. Как видно из записанных выше выражений, вольт-амперные характеристики имеют различный вид в зависимости от высот потенциальных барьеров полупроводников 1 и 2, а также от принятых при их выводе предположений.

В том случае, когда промежуточный слой стремится к нулю, $d \rightarrow 0$, $V_3 \rightarrow 0$ и $\mathcal{P}_n \rightarrow 1$, а выражение (5.1), полученное в предположениях диодной теории для гетероперехода вида антизапорный — запорный слой с положительным скачком, может быть записано в виде

$$i = i_s \left(e^{\frac{eV}{kT}} - 1\right). \quad (5.15)$$

При этом выражения (5.2) — (5.4), (5.13) и (5.14) как в предположениях диодной, так и диффузионной теории принимают вид

$$i = i_s \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}}\right). \quad (5.16)$$

В общих выражениях (5.15) и (5.16) i_s имеет разное значение для различных частных случаев.

Как видно, выражение (5.15) по форме совпадает с выражением вольт-амперной характеристики контакта металл — полупроводник без зазора или p — n перехода. Выражение (5.16) даже при отсутствии промежуточного слоя совпадает по виду с выражением вольт-амперной характеристики для контакта с промежуточным слоем, где роль падения напряжения в этом слое играет падение напряжения в полупроводнике 1.

В случае, описываемом выражением (5.15), прямой ток растет с приложенным напряжением как $e^{\frac{eV}{kT}}$, а обратный испытывает

насыщение. В случае, описываемом выражением (5.16), при постоянном i_s прямой ток растет медленней с приложенным напряжением, так как он пропорционален $e^{\frac{eV_2}{kT}}$, а обратный не испытывает насыщения и растет пропорционально $e^{-\frac{eV_1}{kT}}$.

В общем случае, когда промежуточный слой сохраняется; падение напряжения в нем приводит к изменению вольт-амперных характеристик гетероперехода. Так, при наличии промежуточного слоя ослабляется с напряжением рост прямого тока и усиливается рост обратного для случаев, описываемых выражениями (5.2), (5.13), и усиливается рост прямого тока и ослабляется обратного для случаев, описываемых выражениями (5.3) и (5.14). В случае диодной

теории прямой ток может быть пропорциональным $e^{\frac{eV}{kT}}$, $e^{\frac{eV_2}{kT}}$ или $e^{\frac{e(V_2+V_3)}{kT}}$, а обратный постоянен или пропорционален $e^{-\frac{e(V_1+V_3)}{kT}}$, $e^{-\frac{eV_1}{kT}}$. В случае диффузионной теории прямой ток

может быть пропорциональным $e^{\frac{eV_2}{kT}}$, $e^{\frac{e(V_2+V_3)}{kT}}$, а обратный пропорциональным $e^{-\frac{e(V_1+V_3)}{kT}}$ или $e^{-\frac{eV_1}{kT}}$.

С физической точки зрения характеристика вида (5.15) обуславливается тем, что для диодной теории при положительном скачке барьеров полупроводников 1 и 2, а обратный ток идет в направлении от полупроводника 1 к полупроводнику 2 и не зависит от изменения потенциальных барьеров.

Характеристика вида (5.16) обуславливается тем, что во всех остальных случаях для диодной и диффузионной теории прямой ток контролируется изменением с напряжением высоты потенциального барьера 2 полупроводника, а обратный изменением потенциального барьера полупроводника 1.

Разная величина i_s для вольт-амперной характеристики вида (5.16) в случае диодной теории обуславливается тем, что ток выражается через параметры разных полупроводников. Для диффузионной теории различие величин i_s в формуле (5.16) для различных случаев определяется еще и различным знаком потенциального скачка.

Изменение вида вольт-амперной характеристики с изменением ширины промежуточного слоя d с физической точки зрения обуславливается тем, что изменяются высоты барьеров для носителей заряда при их движении от полупроводника 2 к полупроводнику 1 и обратно. Теперь для движущихся носителей заряда нужно учитывать также изменение высоты потенциального барьера промежуточного слоя. Это может ослаблять или усиливать зависимость прямых и обратных токов от напряжения.

При изменении толщины промежуточного слоя изменяется также величина i_s . Причем при достаточно больших толщинах ток для гетероперехода, как и для контакта металл — полупроводник в предположениях диффузионной теории не зависит от условий прохождения носителей заряда в области пространственного заряда и контролируется изменением высот потенциальных барьеров полупроводников 1 и 2.

5.2. ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ АНИЗОТИПНЫХ ГЕТЕРОПЕРЕХОДОВ

Рассмотрим вольт-амперные характеристики анизотипных гетеропереходов на примере анизотипного $p-n$ гетероперехода. Вначале также пренебрегаем токами через поверхностные уровни, туннелированием через область пространственного заряда и силами изображения.

При принятых предположениях носители заряда, которые прошли через гетеропереход от полупроводника 2 к полупроводнику 1, становятся неосновными. Это позволяет так же, как и для обычного гомо- $p-n$ перехода, сводить решение задачи для анизотипного $p-n$ гетероперехода к решению задачи в квазинейтральной области, а собственно гетеропереход будет определять граничные условия. Принципиально возможные конфигурации высот барьеров, которые можно рассматривать для анизотипных $p-n$ гетеропереходов при предположении, что в полупроводнике 2 реализуется запиорный слой, совпадают с рассмотренными ранее для случая изотипных гетеропереходов (см. рис. 5.1). При этом, как и раньше, в случаях 5.1, а, б изгибы зон определяются главным образом контактной разностью потенциалов, а в случае 5.1, в, г — поверхностными электронными состояниями.

Для нахождения граничных условий в предположениях диодной теории можно воспользоваться тем, что общий ток можно найти как разность надбарьерных токов между полупроводниками 1 и 2. При этом можно воспользоваться результатами п. 5.1, если учесть, что переход из полупроводника, в который введены неосновные носители заряда определяется не равновесной концентрацией n_1 , а суммарной концентрацией носителей заряда $n_1(-\mathcal{L}_1 - d)$. Тогда выражения (5.1) — (5.3) можно переписать в виде, учитывающем изменение концентрации. В результате

$$i = \frac{ev_1 \mathcal{F} n_1}{4} \left(e^{\frac{eV}{kT}} - \frac{n_1(-\mathcal{L}_1 - d)}{n_1} \right) \quad (5.17)$$

для гетероперехода вида антизапорный — запиорный слой с положительным скачком.

Для гетероперехода с отрицательным скачком

$$i = \frac{ev_1 \mathcal{F} n_2}{4} e^{-\frac{\Delta \mathcal{E}_c \pm \varphi_1}{kT}} \left[e^{\frac{eV_2}{kT}} - \frac{n_1(-\mathcal{L}_1 - d)}{n_1} e^{-\frac{e(V_1 + V_2)}{kT}} \right], \quad (5.18)$$

где знак «+» перед φ_1 относится к гетеропереходу вида запиорный — запиорный слой, а знак «-» к гетеропереходу вида антизапорный — запиорный слой.

Для гетероперехода вида запиорный — запиорный слой с положительным скачком выражение (5.3) переписывается в виде

$$i = \frac{ev_1 n_1 \mathcal{F} n}{4} e^{-\frac{\varphi_1}{kT}} \left[e^{\frac{V_2 + V_3}{kT}} - \frac{n_1(-\mathcal{L}_1 - d)}{n_1} e^{-\frac{eV_1}{kT}} \right]. \quad (5.19)$$

При условии, что суммарный ток значительно меньше, чем каждый из членов, стоящих в правой части уравнений (5.17) — (5.19), для всех случаев, изображенных на рис. 5.1, можно получить

$$n_1(-\mathcal{L}_1 - d) = n_1 e^{\frac{eV}{kT}} \quad (5.20)$$

В предположениях диффузионной теории граничные условия легко найти, если пренебречь генерацией и рекомбинацией носителей заряда в каждом из полупроводников. Тогда можно воспользоваться выражениями (5.4) — (5.7), записанными в предыдущем параграфе. Нашей задачей теперь является нахождение величины $n_1(-\mathcal{L}_1 - d)$ через приложенное к гетеропереходу напряжение. Учитывая, что при отсутствии напряжения суммарный ток равен нулю, можно из (5.6) найти

$$v_2 n_2^0(0) = v_1 n_1^0(-d) e^{-\frac{\Delta \mathcal{E}_c}{kT}}, \quad (5.21)$$

из (5.8)

$$v_2 n_2^0(0) e^{-\frac{\Delta \mathcal{E}_c}{kT}} = v_1 n_1^0(-d), \quad (5.22)$$

где дополнительные значки «0» означают $V = 0$.

Если предположить, что в выражениях (5.6) и (5.7) каждый из членов правой части значительно больше, чем суммарный ток i , то как для случая, описываемого уравнением (5.6), так и для случая, описываемого уравнением (5.7), можно найти после подстановки (5.21) и (5.22) в (5.6) и (5.7) соотношение

$$\frac{n_1^0(-d)}{n_2^0(0)} e^{\frac{eV_2}{kT}} = \frac{n_1(-d)}{n_2(0)}. \quad (5.23)$$

После подстановки в это выражение $n_1(-d)$ и $n_2(0)$ из выражений (5.4) и (5.5), в которых пренебрегаются члены пропорциональные токам, можно получить в предположениях диффузионной теории

$$n_1(-\mathcal{L}_1 - d) = n_1 e^{\frac{eV}{kT}} \quad (5.24)$$

Как видно, при принятых приближениях (5.24) совпадает с выражением (5.20), полученным в предположениях диодной теории.

Граничные условия (5.20) и (5.24) для $p-n$ гетероперехода в свою очередь совпадают с граничными условиями для обычного $p-n$ перехода. Поэтому совпадут и окончательные выражения вольт-амперной характеристики. Например, для $p-n$ гетероперехода с широкой базой для электронного тока

$$i_n = \frac{en_1 D_1}{l_1} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (5.25)$$

Здесь значки «1» при разных параметрах, как и ранее, относятся к первому полупроводнику.

Анализ последнего выражения не отличается от анализа аналогичного выражения для обычного $p-n$ перехода и проведен выше. Следует при этом только особо отметить, что выражение (5.25) не зависит от толщины промежуточного слоя.

Полученные выше результаты можно объяснить, если учесть, что для анизотипного гетероперехода рассматривается ток в квазинейтральной области, а области пространственного заряда определяют только граничные условия задачи.

В принятых приближениях, когда для диодной теории общий ток контролируется квазинейтральной областью, а для диффузионной в области пространственного заряда выполняются соотношения Больцмана, граничные условия для гетероперехода совпадают с граничными условиями для $p-n$ перехода. В результате совпадают и конечные результаты для токов через анизотипный гетеропереход и гомо- $p-n$ переход.

Рассмотренная выше модель анизотипного гетероперехода, когда ток определяется диффузией неосновных носителей заряда в квазинейтральной области полупроводника, в литературе называется диффузионной моделью.

Если в предположениях диодной теории в выражениях (5.17) — (5.19) нельзя считать ток малым, то необходимо решать полные выражения относительно n_1 ($-\mathcal{L}_1 - d$). Затем n_1 ($-\mathcal{L} - d$) можно подставить в выражение вида (4.4) и после решения полученного уравнения относительно i получаем для анизотипного гетероперехода

$$i = \frac{i_{sn} \left(e^{\frac{eV}{kT}} - 1 \right)}{1 + \frac{i_{sn}}{i_{sd}}}, \quad (5.26)$$

где

$$i_{sn} = \frac{eD_1 n_1}{l_1}$$

совпадает с электронным током насыщения для $p-n$ гомоперехода, а i_{sd} совпадает с обратным током гетероперехода, в котором нет инжекции;

$$i_{sd} = \frac{ev_1 n_1 \mathcal{P}_n}{4}$$

для гетероперехода вида антизапорный — запорный слой с положительным скачком;

$$i_{sd} = \frac{ev_1 n_1 \mathcal{P}_n}{4} e^{-\frac{\Delta\mathcal{E}_c \pm \Phi_1}{kT}} e^{-\frac{e(V_1 + V_d)}{kT}}$$

для гетеропереходов с отрицательным скачком;

$$i_{sd} = \frac{ev_1 n_1 \mathcal{P}_n}{4} e^{-\frac{\Phi_1}{kT}} e^{-\frac{eV_1}{kT}}$$

для гетероперехода вида запорный — запорный слой с положительным скачком.

Как видно из выражения (5.26), при $i_{sn} \ll i_{sd}$, оно переходит в полученное ранее, где ток контролируется квазинейтральной областью. При $i_{sn} \gg i_{sd}$

$$i = i_{sd} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (5.27)$$

В этом случае ток контролируется областями пространственного заряда полупроводников 1 и 2, а также промежуточным слоем.

Как видно, здесь качественно наблюдается такое же изменение вида вольт-амперной характеристики, как и для $p-n$ гомоперехода в п. 4.2. Однако, если для германиевых и кремниевых гомо- $p-n$ переходов обычно реализуется случай, когда ток контролируется квазинейтральными областями полупроводника, то для анизотипных гетеропереходов как из-за малых значений диффузионных длин, так из-за влияния промежуточного слоя ток обычно контролируется областями пространственного заряда и промежуточным слоем.

Модель анизотипного гетероперехода, для которой ток контролируется областями пространственного заряда и зазором, называется эмиссионной моделью.

Если в выражениях (5.4) или (5.5) членами, пропорциональными току, пренебречь нельзя, для диффузионной теории можно получить вольт-амперную характеристику, отличающуюся от рассмотренной выше.

Таким же методом можно получить и проанализировать вольт-амперную характеристику для дырочных токов в анизотипном гетеропереходе. Однако еще раз напомним, что в силу большого различия высот барьеров для дырок и электронов в гетеропереходе обычно существует только один из токов.

5.3. ВЛИЯНИЕ ТУННЕЛИРОВАНИЯ ЧЕРЕЗ ОБЛАСТЬ ПРОСТРАНСТВЕННОГО ЗАРЯДА И СИЛ ИЗОБРАЖЕНИЯ НА ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ ГЕТЕРОПЕРЕХОДОВ

Если область пространственного заряда хотя бы одного из полупроводников достаточно узкая, возможно туннелирование носителей через эту область. Для определенности будем рассматривать

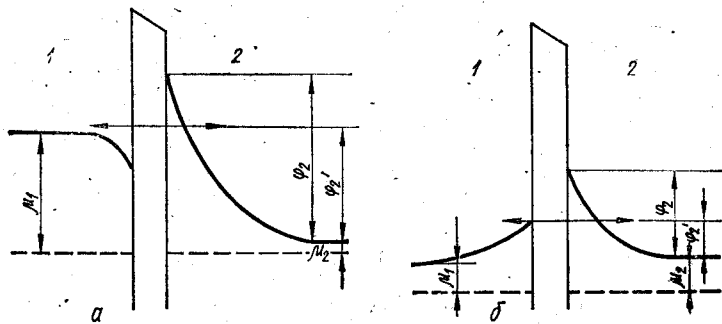


Рис. 5.2. Туннельные токи в гетеропереходе без изменения вида разрешенных зон.

туннелирование в полупроводнике 2. Когда при туннельном переходе носитель заряда оказывается в разрешенной зоне полупроводника, происходит изменение суммарного тока через гетеропереход. Туннелирование через область пространственного заряда может сказываться как при прямых, так и обратных напряжениях, при прохождении носителей заряда в пределах одной зоны над положительным зубцом, как это показано на рис. 5.2, а также только при обратных напряжениях при туннелировании между различными зонами, как это, например, показано на рис. 5.3.

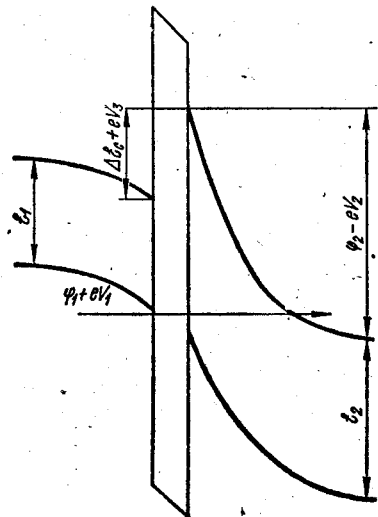


Рис. 5.3. Туннельные токи в гетеропереходе при обратных напряжениях с изменением вида разрешенных зон.

При расчете вольт-амперной характеристики при туннельном прохождении носителей заряда над положительным зубцом можно воспользоваться результатами, полученными ранее для диодной теории, которая более справедлива для тонких областей пространственного заряда.

Если туннелирование идет при энергиях выше, чем дно зоны проводимости полупроводника 1, то расчет вольт-амперной характеристики может быть проведен так же, как и для контакта металл-полупроводник в п. 3.3.

В результате для гетероперехода, используя выражение (3.48), получим

$$i = i_{st} e^{\frac{eV_2}{kT^*}} \left(1 - \frac{T^*}{T} \right) \left(e^{\frac{eV_2}{kT}} - e^{-\frac{e(V_1+V_2)}{kT}} \right), \quad (5.28)$$

где величина T^* задается (3.49).

Записанное выше выражение отличается от (3.48) только тем, что в экспоненте, описывающей обратную ветвь вольт-амперной характеристики, вместо напряжения V_1 стоит напряжение $V_1 + V_2$.

Сравнивая это выражение с выражением (5.2), полученным для гетеропереходов с отрицательным скачком, можно видеть, что туннелирование приводит к более медленному росту прямого тока и более быстрому росту обратного с напряжением.

Для гетероперехода, как и для контакта металл-полупроводник при увеличении туннелирования может происходить изменение знака выпрямления, когда обратный ток при одинаковых напряжениях становится больше прямого.

При достаточно больших $E_{00} \gg kT$ и $T^* = \frac{E_{00}}{k}$, выражение вольт-амперной характеристики может быть переписано в виде

$$i = i_{st} e^{\frac{eV_2}{E_{00}}} \left(1 - e^{-\frac{eV}{kT}} \right). \quad (5.29)$$

Теперь предположим, что главную роль играет туннелирование при энергии, соответствующей дну зоны проводимости (рис. 5.2, а) или вершине потенциального барьера (рис. 5.2, б) полупроводника 1.

Это означает, что при этой энергии коэффициент прозрачности

$$\mathcal{P}_n \gg e^{-\frac{\Delta E_c - eV_1}{kT}},$$

так как только при выполнении последнего неравенства туннельный ток значительно больше надбарьерного.

Тогда в предположениях диодной теории для случая, изображенного на рис. 5.2, а, для $n-n$ гетероперехода

$$i = \frac{e v_2 n_2}{4} \mathcal{P}_n e^{-\frac{\Phi_2}{kT}} \left(e^{\frac{eV_2}{kT}} - 1 \right), \quad (5.30)$$

где Φ_2 и V_2' — пониженные значения высоты потенциального барьера полупроводника 2 и части падающего на нем напряжения. Используя рис. 5.2, легко показать:

$$\Phi_2 = \mu_1 - \mu_2; \\ \Phi_2 - eV_2' = \mu_1 - \mu_2 - eV.$$

В результате выражение вольт-амперной характеристики (5.30) переписывается в виде

$$i = \frac{e v_2 n_2}{4} \mathcal{P}_n e^{-\frac{\mu_1 - \mu_2}{kT}} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (5.31)$$

В предположениях диодной теории для случая, изображенного на рис. 5.2, б, для $n-n$ гетероперехода

$$i = \frac{e v_2 n_2}{4} \mathcal{P}_n e^{-\frac{\Phi_2}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{e(V_1+V_2)}{kT}} \right). \quad (5.32)$$

Используя полученные выше выражения для φ_2 и $\varphi_2 - eV_2$, можно (5.32) переписать в виде

$$i = \frac{eV_2 n_2}{4} \mathcal{P}_n e^{-\frac{\mu_1 - \mu_2}{kT}} \left(e^{\frac{eV}{kT}} - e^{-\frac{e(V_1 + V_3)}{kT}} \right). \quad (5.33)$$

Коэффициент прозрачности в выражениях (5.31) и (5.33) отражает туннелирование через область зазора и пространственного заряда полупроводника 2.

Если пренебречь изменением с напряжением прозрачности зазора, то, используя те же приближения, что и при получении выражения (4.74), можно показать, что

$$\mathcal{P}_n = \mathcal{P}_0 e^{-\sigma_k (\Delta \mathcal{E}_c + eV_1)^{3/2} (\varphi_2 - eV_2)^{-1/2}}$$

где $\sigma_k = \frac{4}{3e\hbar} \left(\frac{8m^* \epsilon_0 \epsilon_2}{n_2} \right)^{1/2}$.

Как видно, учет туннелирования может существенно изменять зависимость тока от приложенного напряжения. Например, выражение (5.30) существенно отличается от аналогичного выражения (5.2), а выражение (5.31) от соответствующего ему выражения (5.3). Туннелирование приводит к изменению как величины тока, так и зависимости прямого и обратного тока от напряжения.

При расчете обратной ветви вольт-амперной характеристики с учетом туннелирования можно воспользоваться теми же принципами, что и при расчете тока в туннельных $p-n$ переходах. При этом в предположении, что туннелирование лимитирует процесс переноса заряда, ток пропорционален коэффициентам прозрачности промежуточного слоя и области пространственного заряда, т. е.

$$i = A \mathcal{P}_n.$$

Если предположить, что туннелирование идет вблизи верха валентной зоны, то, используя те же приближения, что и для $p-n$ перехода при получении выражения (4.74), а также пренебрегая изменением с напряжением коэффициента прозрачности зазора, можно получить

$$i \sim e \sigma_k (\mathcal{E}_1 + \Delta \mathcal{E}_c + eV_3)^{3/2} (\varphi_2 - eV_2)^{-1/2}, \quad (5.34)$$

где $\sigma_k = \frac{8 \left(\frac{m^* \epsilon_0 \epsilon_2}{n_2} \right)^{1/2}}{3e\hbar}$.

Как видно из выражения (5.34), зависимость туннельного тока от напряжения здесь имеет более сложный вид, чем для гомо- $p-n$ перехода. Это связано с влиянием падения напряжения на зазоре и полупроводнике 1. Только при $V_1 = V_3 = 0$ и $V_2 = V$ выражение (5.34) дает ту же зависимость тока от напряжения, что и для гомо- $p-n$ перехода.

Как и для контакта металл — полупроводник, для гетероперехода существенными могут быть силы изображения. Эти силы могут повышать или понижать высоты потенциальных барьеров в зависимости от соотношения диэлектрических проницаемостей полупроводников и промежуточного слоя. Силы изображения могут увеличивать и уменьшать ток через гетеропереход, что необходимо учитывать при строгом анализе экспериментальных результатов.

5.4. ВОЛЬТ-АМПЕРНЫЕ ХАРАКТЕРИСТИКИ ГЕТЕРОПЕРЕХОДОВ ДЛЯ ТОКОВ ЧЕРЕЗ ПОВЕРХНОСТНЫЕ УРОВНИ ПРИ ДВИЖЕНИИ НОСИТЕЛЕЙ В ПРЕДЕЛАХ ОДНОГО ТИПА РАЗРЕШЕННЫХ ЗОН

Ниже будут рассмотрены токи через поверхностные уровни в изотипных и анизотипных гетеропереходах, когда носители заряда движутся в разрешенных зонах одного вида (зонах проводимости или валентных зонах полупроводности 1 и 2). Физические модели рассчитываемых гетеропереходов показаны на рис. 5.4, где на рис. 5.4, а представлен гетеропереход с противоположным, а на рис. 5.4, б — с одинаковым изгибом зон. Используя полученные ранее результаты (см. гл. 2), можно записать для токов полупроводник 1 — уровень при ранее выбранном направлении прямого тока от полупроводника 2 к полупроводнику 1 в виде

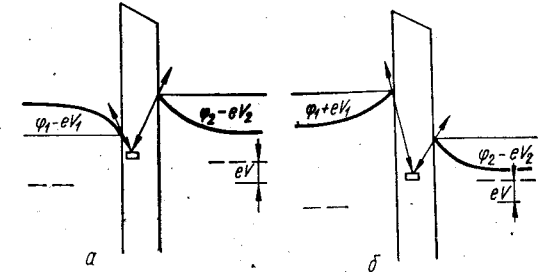


Рис. 5.4. Токи через поверхностные уровни без изменения вида разрешенных зон.

$$i_n^{(1)} = e [K_{ni}^{(1)} m_i - K_{ni}^{(1)} (n_i - m_i) n_k^{(1)}], \quad (5.35)$$

для токов полупроводник 2 — уровень в виде

$$i_n^{(2)} = e [K_{ni}^{(2)} (n_i - m_i) n_k^{(2)} - K_{in}^{(2)} m_i]. \quad (5.36)$$

Здесь, как и раньше, K_{ni} и K_{in} — кинетические коэффициенты, характеризующие электронный обмен между разрешенными зонами и поверхностным уровнем; n_k — концентрации электронов в областях, с которыми идет электронный обмен уровня и которые различаются для диодной и диффузионной теории.

В стационарном случае $i_n^{(1)} = i_n^{(2)}$. Поэтому, приравняв (5.35) и (5.36), легко найти

$$m_i = \frac{(K_{ni}^{(1)} n_k^{(1)} + K_{ni}^{(2)} n_k^{(2)}) n_i}{K_{in}^{(1)} + K_{in}^{(2)} + K_{ni}^{(1)} n_k^{(1)} + K_{ni}^{(2)} n_k^{(2)}}. \quad (5.37)$$

Если подставить это значение в формулу (5.35) или (5.36) можно получить следующее общее выражение для тока через поверхностные уровни в случае гетероперехода:

$$i = \frac{e n_i (K_{in}^{(1)} K_{ni}^{(2)} n_k^{(2)} - K_{ni}^{(1)} K_{in}^{(2)} n_k^{(1)})}{(K_{in}^{(1)} + K_{in}^{(2)} + K_{ni}^{(1)} n_k^{(1)} + K_{ni}^{(2)} n_k^{(2)})}. \quad (5.38)$$

Таким образом, нахождение окончательной вольт-амперной характеристики сводится к физической расшивке параметров

$$K_{ni}^{(1)}, K_{in}^{(1)}, K_{ni}^{(2)}, K_{in}^{(2)}, n_k^{(1)} \text{ и } n_k^{(2)}.$$

Используя равенства (5.35) и (5.36) при напряжении, равном нулю, можно получить:

$$\begin{aligned} K_{in}^{(10)} m_i^{(0)} &= K_{ni}^{(10)} (n_i - m_i) n_k^{(10)}; \\ K_{ni}^{(20)} (n_i - m_i) n_k^{(20)} &= K_{in}^{(20)} m_i^{(0)}, \end{aligned} \quad (5.39)$$

где значки «0» означают, что данная величина относится к случаю, когда напряжение, приложенное к гетеропереходу, равно нулю.

Параметры $K_{ni}^{(1)}$ и $K_{ni}^{(2)}$ по аналогии с рассмотренным ранее случаем для токов через провизи в контакте металл — полупроводник можно записать как произведение тепловых скоростей на сечение захвата:

$$K_{ni}^{(1)} = v_1 c_1^t; \quad K_{ni}^{(2)} = v_2 c_2^t. \quad (5.40)$$

Если предположить, что $K_{ni}^{(1)}$, $K_{ni}^{(2)}$, $K_{in}^{(1)}$ и $K_{in}^{(2)}$ не зависят от напряжения, используя соотношения (5.40), можно получить вольт-амперную характеристику в виде

$$i = \frac{en_i M_i v_1 c_1^t v_2 c_2^t n_k^{(10)} n_k^{(20)} \left(\frac{n_k^{(20)}}{n_k^{(20)}} - \frac{n_k^{(10)}}{n_k^{(10)}} \right)}{v_1 c_1^t (M_i n_k^{(10)} + n_k^1) + v_2 c_2^t (M_i n_k^{(20)} + n_k^2)}, \quad (5.41)$$

где $M_i = e^{\frac{\delta_i - \mu}{kT}}$.

Теперь необходимо конкретизировать зависимость $n_k^{(1)}$ и $n_k^{(2)}$ от напряжения. Это может быть сделано отдельно для изотипного и анизотипного гетероперехода.

Изотипный гетеропереход. Для диодной теории, когда электронный обмен идет через область пространственного заряда полупроводников, для изотипного $n-n$ гетероперехода вида антизапорный — зазорный слой

$$n_k^1 = n_k^{(10)} = n_1. \quad (5.42)$$

Для диодной теории в случае гетероперехода вида зазорный — зазорный слой, а также для диффузионной теории, когда можно пренебречь нарушением функции распределения носителей заряда проходящим током,

$$n_k^{(1)} = n_1 e^{\mp \frac{\varphi_1 \pm eV_1}{kT}}, \quad (5.43)$$

где верхний знак относится к зазорному, а нижний к антизазорному слою.

В предположении диодной теории, а также диффузионной при пренебрежении нарушением функции распределения носителей заряда проходящим током

$$n_k^{(2)} = n_2 e^{-\frac{\varphi_2 - eV_2}{kT}}. \quad (5.44)$$

Используя выражения (5.42) — (5.44), можно записать вольт-амперную характеристику изотипного гетероперехода вида антизапорный — зазорный слой в предположениях диодной теории так:

$$i = \frac{en_i M_i (v_1 c_1^t) (v_2 c_2^t) n_1 n_2 e^{-\frac{\varphi_2}{kT}} (e^{\frac{eV_2}{kT}} - 1)}{(v_1 c_1^t) n_1 (M_i + 1) + v_2 c_2^t n_2 e^{-\frac{\varphi_2}{kT}} (M_i + e^{\frac{eV_2}{kT}})}. \quad (5.45)$$

Используя выражения (5.43), (5.44), легко получить вольт-амперную характеристику изотипного гетероперехода вида зазорный — зазорный слой в предположениях диодной теории, а также в предположениях диффузионной теории

$$i = \frac{en_i M_i (v_1 c_1^t) (v_2 c_2^t) n_1 n_2 e^{\mp \frac{\varphi_1 \pm \varphi_2}{kT}} (e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}})}{(v_1 c_1^t) n_1 e^{\mp \frac{\varphi_1}{kT}} (M_i + e^{-\frac{eV_1}{kT}}) + (v_2 c_2^t) n_2 e^{-\frac{\varphi_2}{kT}} (M_i + e^{\frac{eV_2}{kT}})}, \quad (5.46)$$

где, как и раньше, верхний знак в показателях экспонент относится к зазорному, а нижний — к антизазорному изгибу зон.

Как показывает анализ выражений (5.45) и (5.46), величины токов через поверхностные уровни пропорциональны их концентрации n_i и зависят от их положения M_i , а также сечений захвата электронов зоны проводимости полупроводников 1 и 2 c_1^t и c_2^t . Кроме того, токи через поверхностные уровни определяются концентрациями носителей заряда n_1 и n_2 , высотами потенциальных барьеров φ_1 и φ_2 в обоих полупроводниках. Зависимость тока от напряжения

определяется комбинацией экспонент вида $e^{\frac{eV_2}{kT}}$ и $e^{-\frac{eV_1}{kT}}$. Полный анализ полученных выражений достаточно сложен. Поэтому проведем анализ полученных выражений для частных случаев, когда главную роль играет электронный обмен с полупроводником 1 и полупроводником 2.

Если главную роль играет электронный обмен с полупроводником, то $c_1^t \gg c_2^t$ и можно пренебречь вторым членом знаменателя выражений (5.45) — (5.46). В результате:

$$i = \frac{en_i M_i v_2 c_2^t n_2 e^{-\frac{\varphi_2}{kT}} (e^{\frac{eV_2}{kT}} - 1)}{(M_i + 1)}; \quad (5.47)$$

$$i = \frac{en_i M_i v_1 c_1^t n_2 e^{-\frac{\varphi_2}{kT}} (e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}})}{(M_i + e^{-\frac{eV_1}{kT}})}. \quad (5.48)$$

Как видно из последних выражений, при преимущественном обмене поверхностных уровней с полупроводником 1 токи пропорциональны сечениям захвата электронов зоны проводимости полупроводника 2. Это, как и для контакта металл — полупроводник, опре-

деляется тем, что при последовательном прохождении электронов через уровни контролирующим является этап менее интенсивного обмена. По этой же причине токи также пропорциональны концентрации носителей заряда в полупроводнике 2 и зависят от высоты потенциального барьера для него. В зависимости от положения уровня ток при прочих одинаковых параметрах может увеличиваться с ростом величины M_i , а затем выходить на насыщение.

Анализ выражений (5.47) и (5.48) показывает, что прямой ток через гетеропереход может быть пропорциональным $e \frac{eV_2}{kT}$ или $e \frac{e(V_2+V_1)}{kT}$, а обратный пропорциональным $e^{-\frac{eV_1}{kT}}$ или не зависеть от напряжения.

Если главную роль играет электронный обмен уровня с полупроводником 2 и $c_2^i \gg c_1^i$, то можно пренебречь первым членом знаменателя выражений (5.45) — (5.46) и переписать их в виде:

$$i = \frac{en_i M_i v_1 c_1^i n_1 \left(e^{\frac{eV_2}{kT}} - 1 \right)}{\left(M_i + e^{\frac{eV_2}{kT}} \right)} \quad (5.49)$$

$$i = \frac{en_i M_i v_1 c_1^i n_1 e^{\mp \frac{\Phi_1}{kT}} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right)}{\left(M_i + e^{-\frac{eV_2}{kT}} \right)} \quad (5.50)$$

Как показывает анализ последних двух выражений, при преимущественном обмене поверхностных уровней с полупроводником 2 токи через поверхностные уровни пропорциональны их концентрации, и определяются параметрами, характеризующими полупроводник 1.

В зависимости от величины M_i , определяемой энергетическим положением поверхностных уровней, ток через уровни может расти, а затем выходить на насыщение.

Анализ выражений (5.49), (5.50) показывает, что прямой ток через гетеропереход может быть пропорционален $e \frac{eV_2}{kT}$ или не зависеть от напряжения, обратный ток пропорционален $e^{-\frac{eV_2}{kT}}$, $e^{-\frac{eV_1}{kT}}$ или изменяется более сложно.

Таким образом, проведенное рассмотрение токов через поверхностные уровни при движении носителей заряда в пределах одного типа разрешенных зон показывает, что в зависимости от реализации конкретных условий вольт-амперная характеристика гетероперехода изменяется самым существенным образом.

Анизотипный гетеропереход. В случае анизотипного $p-n$ гетероперехода в p -области электроны становятся неосновными носителями заряда. При этом плотность тока на границе квазинейтральной области и области пространственного заряда полупроводника

1 для гетероперехода с широкой базой, как и для $p-n$ перехода, задается выражением

$$i = \frac{eD_1}{l_1} \Delta n_1 (-\mathcal{L}_1 - d) = \frac{eD_1}{l_1} [n_1 (-\mathcal{L}_1 - d) - n_1]. \quad (5.51)$$

В случае диодной теории для анизотипного гетероперехода вида антизапорный — запорный слой

$$n_1 (-\mathcal{L}_1 - d) = n_k^{(1)}.$$

Для диодной теории гетероперехода вида запорный — запорный слой, а также диффузионной теории при пренебрежении нарушения функции распределения носителей заряда проходящим током

$$n_1 (-\mathcal{L}_1 - d) = n_k^{\pm} e^{\pm \frac{\Phi_1 + eV_1}{kT}},$$

где верхний знак в показателе степени относится к антизапорному, а нижний — к запорному слою.

Теперь расчет вольт-амперной характеристики может быть сведен к совместному решению уравнений (5.41) и (5.51) с учетом существующей связи между $n_1 (-\mathcal{L}_1 - d)$ и n_k^1 .

Для простоты будем рассматривать два частных случая. Пусть ток через гетеропереход контролируется движением носителей заряда в квазинейтральной области полупроводника 1. Тогда в выражении (5.41) можно положить ток $i = 0$. С физической точки зрения это означает, что ток является малой разностью двух больших величин. В результате легко найти

$$n_k^{(1)} = n_k^{(1c)} \frac{n_k^{(2)}}{n_k^{(20)}} = n_k^{(10)} e^{\frac{eV_2}{kT}}, \quad (5.52)$$

где в гетеропереходах вида запорный — запорный слой для диодной теории $n_k^{(10)} = n_1 (-\mathcal{L}_1 - d)$, а в гетеропереходах вида антизапорный — запорный слой для диодной теории, а также для диффузионной теории при принятых приближениях $n_k^{(10)} = n_1 e^{\mp \frac{\Phi_1}{kT}}$.

Все это позволяет записать для диодной теории и гетероперехода вида запорный слой

$$n_1 (-\mathcal{L}_1 - d) = n_1 e^{\frac{eV}{kT}}. \quad (5.53a)$$

Для гетероперехода вида антизапорный — запорный слой в предположениях диодной теории, а также в предположениях диффузионной теории при принятых выше приближениях

$$n_1 (-\mathcal{L}_1 - d) = n_1 e^{\frac{e(V_2+V_1)}{kT}}. \quad (5.53b)$$

Подставляя найденные выражения $n_1 (-\mathcal{L}_1 - d)$ в выражение (5.56), получаем для диодной теории и гетероперехода вида

запорный — запорный-слой

$$i = i_s \left(e^{\frac{eV_2}{kT}} - 1 \right). \quad (5.54a)$$

Для диодной теории и гетероперехода вида антизапорный — запорный слой, а также для диффузионной теории

$$i = i_s \left(e^{\frac{e(V_2+V_1)}{kT}} - 1 \right). \quad (5.54b)$$

В выражениях (5.54a) и (5.54b)

$$i_s = \frac{eD_1 n_1}{l_1}$$

Как видно из выражений (5.54), прямой ток анизотипного гетероперехода пропорционален $e^{\frac{eV_2}{kT}}$ или $e^{\frac{e(V_1+V_2)}{kT}}$, а обратный испытывает насыщение. Величина тока насыщения i_s определяется параметрами квазинейтральной области полупроводника 1.

Пусть ток контролируется прохождением носителей заряда через поверхностные уровни. Тогда можно приравнять нулю выражение (5.51). Это также означает то, что ток является малой разностью двух больших величин.

Тогда легко найти

$$n_1 (-\mathcal{L}_1 - d) = n_1,$$

следовательно, уровень инжекции очень мал. Это позволяет воспользоваться для n_k выражениями (5.42) и (5.43).

В результате можно получить вольт-амперную характеристику, совпадающую по виду с вольт-амперной характеристикой для изотипного гетероперехода и воспользоваться проведенным выше анализом. При этом следует специально отметить, что n_1 для анизотипного гетероперехода — это равновесная концентрация неосновных носителей, которая значительно меньше равновесной концентрации основных носителей заряда n_1 в изотипном гетеропереходе.

5.5. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ГЕТЕРОПЕРЕХОДОВ ДЛЯ ТОКОВ ЧЕРЕЗ ПОВЕРХНОСТНЫЕ УРОВНИ ПРИ ДВИЖЕНИИ НОСИТЕЛЕЙ С ИЗМЕНЕНИЕМ ВИДА РАЗРЕШЕННЫХ ЗОН

Пусть носители заряда при прохождении через поверхностные уровни изменяют вид разрешенной зоны и переходят из зоны проводимости в валентную зону и наоборот. В изотипном $n-n$ гетеропереходе при таком переходе носитель заряда становится неосновным носителем и происходит инжекция, в анизотипном $p-n$ гетеропереходе носитель заряда остается основным и инжекции нет.

В случае анизотипного гетероперехода рассматриваемая в настоящем параграфе модель обычно в литературе называется эмиссионно-рекомбинационной.

Как для анизотипного, так и изотипного гетероперехода, используя физическую модель, изображенную на рис. 5.5, где представлены гетеропереходы с разным и одинаковым изгибом зон, можно записать для токов из полупроводника 1 на уровень

$$i_{11} = e [K_{iv}^{(1)} m_i p_k^{(1)} - K_{vi}^{(1)} (n_i - m_i)], \quad (5.55)$$

а для токов из полупроводника 2 на уровень

$$i_{21} = e [K_{ni}^{(2)} (n_i - m_i) n_k^{(2)} - K_{in}^{(2)} m_i], \quad (5.56)$$

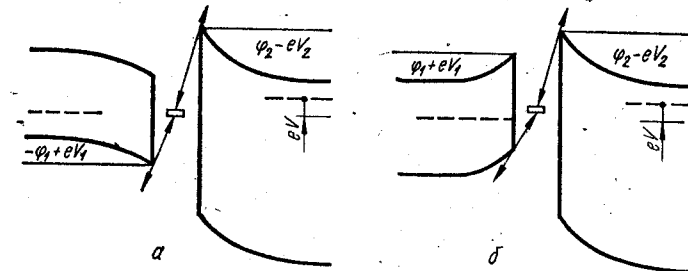


Рис. 5.5. Токи через поверхностные уровни при изменении вида разрешенных зон.

где $K_{iv}^{(1)}$ и $K_{vi}^{(1)}$ — кинетические коэффициенты, характеризующие электронный обмен уровня с валентной зоной полупроводника 1; p_k — концентрация дырок в области полупроводника 1, с которой идет электронный обмен, имеющая различное значение для диодной и диффузионной теории.

Для стационарного случая, когда $i_{11} = i_{21}$, легко найти

$$m_i = \frac{(K_{vi}^{(1)} + K_{ni}^{(2)} n_k^{(2)}) n_i}{K_{vi}^{(1)} + K_{iv}^{(1)} p_k^{(1)} + K_{ni}^{(2)} n_k^{(2)} + K_{in}^{(2)}}. \quad (5.57)$$

Подставляя это значение m_i в формулы (5.55) или (5.56), можно получить

$$i = \frac{e n_i (K_{iv}^{(1)} K_{ni}^{(2)} n_k^{(2)} - K_{vi}^{(1)} K_{in}^{(2)})}{(K_{vi}^{(1)} + K_{iv}^{(1)} p_k^{(1)} + K_{ni}^{(2)} n_k^{(2)} + K_{in}^{(2)})}. \quad (5.58)$$

Таким образом, для рассматриваемого случая вольт-амперная характеристика гетероперехода может быть найдена после определения коэффициентов $K_{iv}^{(1)}$, $K_{vi}^{(1)}$, $K_{ni}^{(2)}$, $K_{in}^{(2)}$ и величин $p_k^{(1)}$ и $n_k^{(2)}$.

Используя равенство (5.55) при $V = 0$, можно найти

$$K_{iv}^{(10)} m_i^0 p_k^0 = K_{vi}^{(10)} (n_i - m_i^0). \quad (5.59)$$

Кроме того, по аналогии с контактом металл—полупроводник запишем

$$K_{iv}^{(1)} = v_1 c_1^i, \quad (5.60)$$

где v_1 — средняя тепловая скорость дырок в полупроводнике 1; c_1^i — сечение захвата дырок из полупроводника 1 локальным уровнем.

Если пренебречь зависимостью $K_{cv}^{(1)}$, $K_{iv}^{(1)}$ и $K_{in}^{(2)}$, $K_{iv}^{(2)}$ от напряжения и использовать для определения $K_{in}^{(2)}$ и $K_{iv}^{(2)}$ результат предыдущего параграфа, после подстановки полученных кинетических коэффициентов в (5.58) можно записать следующее общее выражение вольт-амперной характеристики:

$$i = \frac{en_i v_2 c_2 v_1 c_1 (n_k^{(2)} p_k^{(1)} - n_k^{(20)} p_k^{(10)})}{v_1 c_1 (p_k^{(1)} + M_i^{-1} p_k^{(10)}) + v_2 c_2 (n_k^{(2)} + n_k^{(20)})}. \quad (5.61)$$

Величина $n_k^{(2)}$ в этом выражении совпадает с полученной в предыдущем параграфе. Величина $p_k^{(1)}$ будет рассмотрена отдельно для анизотипного и изотипного гетеропереходов.

Анизотипный гетеропереход. В анизотипном переходе при переходах с изменением вида разрешенных зон носители заряда в обоих полупроводниках являются основными.

Для диодной теории и случая, изображенного на рис. 5.5, а:

$$p_k^{(1)} = p_1 e^{-\frac{\varphi_1 - eV_1}{kT}}. \quad (5.62)$$

Для той же теории и случая, изображенного на рис. 5.5, б:

$$p_k^{(1)} = p_1. \quad (5.63)$$

В предположениях диффузионной теории при достаточно малых токах, когда не нарушается функция распределения носителей заряда,

$$p_k^{(1)} = p_1 e^{\mp \frac{\varphi_1 \pm eV_1}{kT}}, \quad (5.64)$$

где верхний знак в показателе степени относится к заборному, а нижний — к антизаборному слою.

Используя выражения (5.62) — (5.64), можно получить для диодной теории и гетероперехода вида заборный — заборный слой, а также для диффузионной теории при принятых предположениях о малости тока

$$i = \frac{en_i v_2 c_2 v_1 c_1 n_2 e^{-\frac{\varphi_2}{kT}} p_1 e^{\mp \frac{\varphi_1}{kT}} (e^{\frac{e(V_2+V_1)}{kT}} - 1)}{v_1 c_1 p_1 e^{\mp \frac{\varphi_1}{kT}} (e^{\frac{eV_1}{kT}} + M_i) + v_2 c_2 n_2 e^{-\frac{\varphi_2}{kT}} (e^{\frac{eV_2}{kT}} + 1)}. \quad (5.65)$$

Для диодной теории и гетероперехода вида антизаборный — заборный слой имеем

$$i = \frac{en_i v_2 c_2 v_1 c_1 n_2 e^{-\frac{\varphi_2}{kT}} p_1 (e^{\frac{eV_2}{kT}} - 1)}{v_1 c_1 p_1 (1 + M_i^{-1}) + v_2 c_2 n_2 e^{-\frac{\varphi_2}{kT}} (e^{\frac{eV_2}{kT}} + 1)}. \quad (5.66)$$

Как показывает анализ полученных выражений, величины токов через поверхностные уровни зависят от параметров обоих полупроводников, параметров поверхностных состояний и приложенного

напряжения. Их подробный анализ может быть проведен при преимущественном обмене поверхностного уровня с полупроводником 1 или полупроводником 2, как это было проведено в предыдущем параграфе.

При этом остаются справедливыми основные выводы о том, что ток пропорционален концентрации поверхностных уровней и определяется параметрами того полупроводника, обмен с которым контролирует ток. Прямой ток для рассматриваемого случая может быть пропорциональным $e^{\frac{eV_2}{kT}}$, $e^{\frac{eV_1}{kT}}$, $e^{\frac{e(V_2+V_1)}{kT}}$ или постоянным. Обратный ток может изменяться как $e^{-\frac{eV_2}{kT}}$, $e^{-\frac{eV_1}{kT}}$ или быть постоянным.

Изотипный переход. В изотипном гетеропереходе при переходах с изменением вида разрешенных зон носители заряда в одном из полупроводников становятся неосновными. Пусть это полупроводник 1. Тогда плотность тока на границе квазинейтральной области и области пространственного заряда полупроводника 1 для гетероперехода с широкой базой

$$i = \frac{eD_1}{l_1} \Delta p_1 (-\mathcal{L}_1 - d) = \frac{eD_1}{l_1} [p_1 (-\mathcal{L}_1 - d) - p_1]. \quad (5.67)$$

Для диодной теории и модели гетероперехода, показанной на рис. 5.5, б:

$$p_1 (-\mathcal{L}_1 - d) = p_k^1. \quad (5.68)$$

Для диодной теории и модели, представленной на рис. 5.5, б, а также для диффузионной теории при пренебрежении нарушения функции распределения проходящим током

$$p_1 (-\mathcal{L}_1 - d) = p_k^1 e^{\pm \frac{\varphi_1 \mp eV_1}{kT}}. \quad (5.69)$$

Теперь для нахождения вольт-амперной характеристики изотипного гетероперехода необходимо, как и в предыдущем параграфе, решить совместно уравнения (5.61) и (5.67). Для простоты снова рассмотрим два частных случая.

Пусть ток контролируется движением носителей заряда в квазинейтральной области полупроводника 1. Тогда можно приравнять нулю выражение (5.61).

В результате

$$p_k^1 = p_k^{(10)} \frac{n_k^{(20)}}{n_k^{(2)}} = p_k^{(10)} e^{-\frac{eV_2}{kT}}, \quad (5.70)$$

где для диодной теории и случая, изображенного на рис. 5.5, б $p_k^{(10)} = p_1$, а для диодной теории и случая, изображенного на

рис. 5.5, а, а также для диффузионной теории $p_k^{(10)} = p_1 e^{\mp \frac{\varphi_1}{kT}}$.

После подстановки p_k^1 в выражения (5.68), (5.69) и подстановки последних в (5.67)

$$i = i_s \left(e^{-\frac{eV_2}{kT}} - 1 \right) \quad (5.71. a)$$

для диодной теории и гетероперехода вида запертый — запертый слой;

$$i = i_s \left(e^{-\frac{e(V_2+V_1)}{kT}} - 1 \right) \quad (5.71. б)$$

для диодной теории и гетероперехода вида антизапертый — запертый слой, а также для диффузионной теории, где

$$i_s = \frac{eD_1 p_1}{l_1}$$

Как видно из полученных выражений, прямой ток постоянен для рассматриваемого случая, а обратный пропорционален $e^{-\frac{eV_2}{kT}}$ или $e^{-\frac{e(V_2+V_1)}{kT}}$. При этом изменяется знак выпрямления по сравнению со случаем, рассмотренным в предыдущем параграфе.

Пусть теперь ток контролируется прохождением носителей заряда через поверхностные уровни. Тогда можно положить равным нулю выражение (5.67) и получить, что $p_1 (-\mathcal{L}_1 - d) = p_1$.

Это позволяет воспользоваться для p_k^1 выражениями (5.62) — (5.64). В результате задача нахождения вольт-амперной характеристики для изотипного гетероперехода и ее анализа сводится к рассмотренной раньше задаче для анизотипного гетероперехода.

5.6. ГЕНЕРАЦИОННО-РЕКОМБИНАЦИОННЫЕ И ТУННЕЛЬНО-РЕКОМБИНАЦИОННЫЕ ТОКИ В ГЕТЕРОПЕРЕХОДЕ

В анизотипных гетеропереходах могут наблюдаться такие же генерационно-рекомбинационные токи, как и для обычных гомо- $p-n$ переходов. Возникновение таких токов иллюстрируется рис. 5.6.

Генерационно-рекомбинационные токи могут реализоваться как в узкозонном, так и широкозонном полупроводнике. Затем носители заряда, определяющие сквозной ток, должны пройти промежуточный слой и оказаться в разрешенных зонах другого полупроводника. Если при этом сквозной ток контролируется собственно генерационно-рекомбинационными процессами, то для нахождения генерационно-рекомбинационных токов в гетеропереходах можно воспользоваться той же методикой, что и в п. 4.5 при рассмотрении обычного $p-n$ перехода. При этом, например, обратные токи могут быть записаны с помощью формулы (4.62), где все величины должны относиться к полупроводнику 1 и 2.

В результате можно получить для локальных уровней, расположенных вблизи середины запрещенной зоны полупроводников

1 и 2:

$$i_r = e \left(\frac{n_{i1} \mathcal{L}_1}{\tau_{n0}^{(1)} + \tau_{p0}^{(1)}} + \frac{n_{i2} \mathcal{L}_2}{\tau_{n0}^{(2)} + \tau_{p0}^{(2)}} \right) \quad (5.72)$$

В зависимости от параметров полупроводников 1 и 2, определяющую роль может играть первый или второй член последнего выражения.

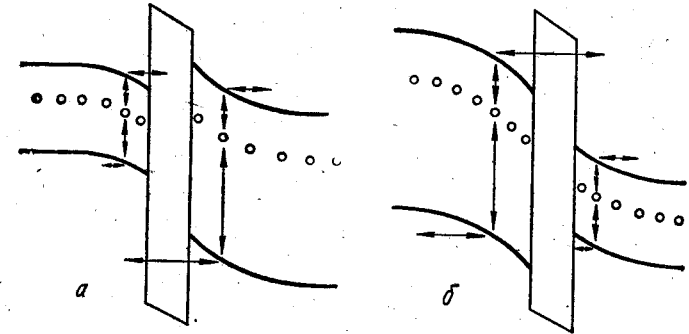


Рис. 5.6. Генерационно-рекомбинационные токи в гетеропереходах.

При этом обратный ток может быть пропорциональным $(\Phi_1 + eV_1)^{1/2}$ или $(\Phi_2 - eV_2)^{1/2}$.

Кроме того, в гетеропереходах существенную роль могут играть так называемые туннельно-рекомбинационные токи (рис. 5.7). Соответствующая модель гетероперехода называется туннельно-

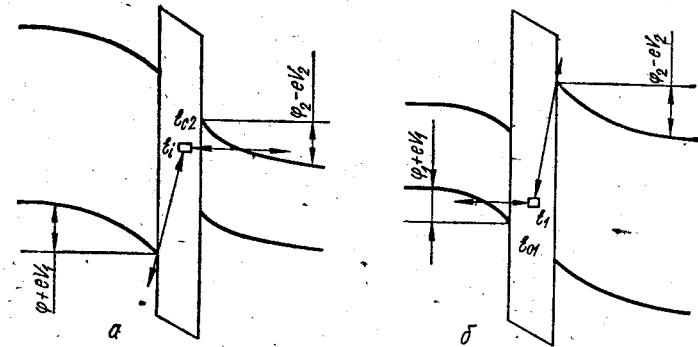


Рис. 5.7. Туннельно-рекомбинационные токи в гетеропереходах.

рекомбинационной. Как видно из рис. 5.7, в случае туннельно-рекомбинационных токов один из электронных переходов на поверхностные уровни осуществляется с помощью туннельного эффекта.

Расчет вольт-амперных характеристик для туннельно-рекомбинационных токов может быть проведен аналогично тому, как проводился для избыточных токов через гомо- $p-n$ переход.

Если рассмотреть разность потока электронов зона проводимости — поверхностный уровень и поверхностный уровень — зона

проводимости (см. рис. 5.7, а), можно получить

$$i = C \int_{\xi} \mathcal{P}_n(\xi) [f_n(\xi) - f_i(\xi)] N_n(\xi) N_i(\xi) d\xi, \quad (5.73)$$

где $f_n(\xi)$ и $f_i(\xi)$ — функции распределения электронов в зоне проводимости и на поверхностных уровнях; $N_n(\xi)$ и $N_i(\xi)$ — плотности состояний для зоны проводимости и поверхностных уровней;

$\mathcal{P}(\xi)$ — коэффициент прозрачности для туннелирующих электронов. Интегрирование проводится по всем энергиям ξ , при которых осуществляются электронные переходы.

Если туннельные переходы происходят без изменения энергии ξ_i , то последнее выражение может быть переписано в виде

$$i \sim \int_{\xi} \mathcal{P}_n(\xi) [f_n(\xi) - f_i(\xi)] N_n(\xi) N_i(\xi) \delta(\xi + \xi_i) d\xi. \quad (5.74)$$

где $\delta(\xi + \xi_i)$ — дельта-функция.

После интегрирования этого выражения для локального уровня, расположенного в точке ξ_i , имеем

$$i \sim n_i \mathcal{P}_n(\xi_i) (\varphi_2 - eV_2 + \xi_i - \xi_{c2})^{1/2} [f_n(\xi_i) - f_i(\xi_i)], \quad (5.75)$$

где учтено, что $N_n(\xi_i) \sim (\varphi_2 - eV_2 + \xi_i - \xi_{c2})^{1/2}$.

Величина коэффициента прозрачности $\mathcal{P}(\xi_i)$ может быть получена аналогично тому, как это было сделано в п. 5.3. В результате, пренебрегая зависимостью положения уровня от падения напряжения в зазоре, получим

$$\mathcal{P}_n = \mathcal{P}_0 e^{-\sigma_k(\varphi_2 - eV_2)^{-1/2}}, \quad (5.76)$$

где

$$\sigma_k = \frac{8}{3e\hbar} \left(\frac{m^* e_0 e_2}{n_2} \right)^{1/2} (\xi_{c2} - \xi_i)^{1/2}.$$

Функция распределения электронов в полупроводнике — это функция распределения Ферми. При отсутствии вырождения ее можно записать так:

$$f_n(\xi_i) = e^{-\frac{\xi_i - \mu_s - eV_2}{kT}} = M_i e^{\frac{eV_2}{kT}}. \quad (5.77)$$

Функция распределения для уровней $f_i(\xi_i) = \frac{m_i}{n_i}$ может быть получена при рассмотрении электронных переходов поверхностных уровней с полупроводником 1 и 2. Пусть ток контролируется туннельными переходами зона проводимости — поверхностный уровень. Тогда заполнение уровня определяется обменом с валентной зоной полупроводника 1.

В этом случае на основании (5.57)

$$f_i = \frac{1}{1 + \frac{K_{iv} p_n^1}{K_{vi}}}. \quad (5.78)$$

Для анизотипного гетероперехода, используя (5.63), можно получить

$$f_i = \frac{1}{1 + M_i} \simeq M_i^{-1}. \quad (5.79)$$

Подставляя в (5.78) выражения (5.62) или (5.64), имеем для анизотипного гетероперехода

$$f_i = \frac{1}{1 + M_i e^{\frac{eV_1}{kT}}}. \quad (5.80)$$

Для анизотипного гетероперехода после подстановки в (5.78) выражения (5.70) имеем

$$f_i = \frac{1}{1 + M_i e^{-\frac{eV_2}{kT}}} \simeq M_i^{-1} e^{\frac{eV_2}{kT}}. \quad (5.81)$$

Теперь можно записать выражение вольт-амперной характеристики для различных частных случаев.

Если f_i задается выражением (5.79), то

$$i \sim n_i \mathcal{P}_n M_i^{-1} (e^{\frac{eV_2}{kT}} - 1). \quad (5.82)$$

Если f_i задается выражением (5.80), при $M_i e^{\frac{eV_1}{kT}} \gg 1$ получаем

$$i \sim n_i \mathcal{P}_n M_i^{-1} (e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}}). \quad (5.83)$$

Если f_i задается выражением (5.81), то $i = 0$ и туннельно-рекомбинационный ток отсутствует.

Как видно из полученных выражений, прямой ток через гетеропереход может быть пропорциональным $\mathcal{P}_n e^{\frac{eV_2}{kT}}$, а обратный — \mathcal{P}_n или $\mathcal{P}_n e^{-\frac{eV_1}{kT}}$.

Изложенное справедливо, пока выполняется неравенство $\xi_{c2} - \xi_i < \varphi_2 - eV_2$. При обратном неравенстве туннельно-рекомбинационные токи равны нулю, поскольку для поверхностного уровня нет разрешенных состояний в полупроводнике 2, куда возможны туннельные переходы без изменения энергии.

Аналогичным образом можно проанализировать и туннельно-рекомбинационный ток для модели, показанной на рис. 5.7, б.

5.7. ПАРАМЕТРЫ α И α^* ВОЛЬТ-АМПЕРНОЙ ХАРАКТЕРИСТИКИ ГЕТЕРОПЕРЕХОДА

Для гетероперехода, как и для контакта металл-полупроводник, большую роль играет степень роста тока с приложенным напряжением.

Для прямого и обратного направлений тока, как и для контакта металл-полупроводник:

$$\alpha = \frac{d \ln i_{np}}{dV}; \quad \alpha^* = \frac{d \ln i_{обр}}{dV}.$$

Как показал проведенный выше анализ, для различных частных случаев прямая ветвь вольт-амперной характеристики гетероперевода может быть представлена в виде

$$i_{\text{пр}} = i_s e^{f_1(V_1, V_3, V_2)}, \quad (5.84)$$

а обратная в виде

$$i_{\text{обр}} = i_s e^{-f_2(V_1, V_3, V_2)}. \quad (5.85)$$

Здесь i_s является более слабой функцией напряжения, чем экспоненциальная функция вида $e^{f_1(V_1, V_3, V_2)}$ и $e^{-f_2(V_1, V_3, V_2)}$.

Учитывая выражения (5.84) и (5.85), можно записать параметры α и α^* в следующем общем виде:

$$\alpha = \frac{df_1}{dV_1} \frac{dV_1}{dV} + \frac{df_1}{dV_3} \frac{dV_3}{dV} + \frac{df_1}{dV_2} \frac{dV_2}{dV} + \frac{d \ln i_s}{dV}. \quad (5.86)$$

$$\alpha^* = \frac{df_2}{dV_1} \frac{dV_1}{dV} + \frac{df_2}{dV_3} \frac{dV_3}{dV} + \frac{df_2}{dV_2} \frac{dV_2}{dV} + \frac{d \ln i_s}{dV}. \quad (5.87)$$

Таким образом, задача нахождения α и α^* , в первую очередь, сводится к нахождению производных $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$.

Расчет и анализ величин $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$.

Для нахождения производных $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ можно воспользоваться системой уравнений (2.42), где для общности учтен заряд, связанный с поверхностными состояниями на границах раздела полупроводник 1 — промежуточный слой и промежуточный слой — полупроводник 2.

Дифференцируя выражения (2.42) по напряжению V , можно получить систему трех уравнений относительно $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$. После решения этой системы уравнений имеем:

$$\frac{dV_1}{dV} = \frac{\frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} \frac{d\sigma_2}{dV} + \epsilon_0 \epsilon_3 \frac{dE_3}{dV_3} + \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2}\right) \frac{d\sigma_1}{dV}}{\epsilon_0 \epsilon_1 \frac{dE_1}{dV_1} \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} + \frac{\epsilon_3}{\epsilon_1} \frac{dE_3}{dV_3} \frac{dV_1}{dE_1}\right)}; \quad (5.88)$$

$$\frac{dV_3}{dV} = \frac{1 + \frac{1}{\epsilon_0 \epsilon_2} \frac{dV_2}{dE_2} \frac{d\sigma_2}{dV} - \frac{1}{\epsilon_0 \epsilon_1} \frac{dV_1}{dE_1} \frac{d\sigma_1}{dV}}{1 + \frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} + \frac{\epsilon_3}{\epsilon_1} \frac{dE_3}{dV_3} \frac{dV_1}{dE_1}}; \quad (5.89)$$

$$\frac{dV_2}{dV} = \frac{\epsilon_0 \epsilon_3 \frac{\epsilon_1}{\epsilon_2} \frac{dE_1}{dV_1} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} - \frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} \frac{d\sigma_1}{dV} - \left(\frac{dE_1}{dV_1} \frac{dV_2}{dE_2} \frac{\epsilon_1}{\epsilon_2} + \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} \frac{\epsilon_3}{\epsilon_2}\right) \frac{d\sigma_2}{dV}}{\epsilon_0 \epsilon_1 \frac{dE_1}{dV_1} \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{dE_3}{dV_3} \frac{dV_2}{dE_2} + \frac{\epsilon_3}{\epsilon_1} \frac{dE_3}{dV_3} \frac{dV_1}{dE_1}\right)}. \quad (5.90)$$

Подставим в эти выражения величины производных $\frac{dE_1}{dV_1}$, $\frac{dE_3}{dV_3}$ и $\frac{dE_2}{dV_2}$.

В том случае, когда область пространственного заряда является слоем Шоттки, как уже рассматривалось в п. 3.6, $\frac{dE_2}{dV_2} = \frac{1}{\mathcal{L}_2}$. Аналогично можно получить для слоя Шоттки в полупроводнике 1, что $\frac{dE_1}{dV_1} = \frac{1}{\mathcal{L}_1}$. Для промежуточного слоя, в котором можно пренебречь объемным зарядом, $\frac{dE_3}{dV_3} = \frac{1}{d}$.

Если в полупроводнике 1 реализуется антизапорный слой, то выражение для E_1 может быть получено путем интегрирования уравнения Пуассона (1.7), в которое необходимо подставить $\rho(x)$ в виде $\rho(x) = en_1 e^{\frac{\phi(x)}{kT}}$ (см. (1.10)). В результате после интегрирования уравнения Пуассона от $\phi(x) = 0$ до $\phi(x) = \phi_1 - eV_1$ получаем

$$E_1 = - \left(\frac{2kT}{\epsilon_0 \epsilon_1} n_1 e^{\frac{\phi_1 - eV_1}{kT}} \right)^{1/2}. \quad (5.91)$$

После дифференцирования E_1 по V_1

$$\frac{dE_1}{dV_1} = \frac{1}{\mathcal{L}_1} \left(\frac{4(\phi_1 - eV_1)}{kT} \right)^{1/2} e^{\frac{\phi_1 - eV_1}{kT}} = \frac{f}{\mathcal{L}_1}, \quad (5.92)$$

где \mathcal{L}_1 — ширина слоя Шоттки с высотой барьера, равной по величине высоте барьера антизапорного слоя; $f = \left[\frac{4(\phi_1 - eV_1)}{kT} \right]^{1/2} e^{\frac{\phi_1 - eV_1}{kT}}$. Как видно из последнего выражения, для антизапорного слоя $f \gg 1$.

Теперь можно подставить полученные выражения для $\frac{dE_1}{dV}$, $\frac{dE_3}{dV_3}$ и $\frac{dE_2}{dV_2}$ в уравнения (5.88) — (5.90) и получить:

$$\frac{dV_1}{dV} = \frac{\frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} \frac{d\sigma_2}{dV} + \frac{\epsilon_0 \epsilon_3}{d} + \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d}\right) \frac{d\sigma_1}{dV}}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}\right)}; \quad (5.93)$$

$$\frac{dV_3}{dV} = \frac{1 + \frac{\mathcal{L}_2}{\epsilon_0 \epsilon_2} \frac{d\sigma_2}{dV} - \frac{\mathcal{L}_1}{f \epsilon_0 \epsilon_1} \frac{d\sigma_1}{dV}}{1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}}; \quad (5.94)$$

$$\frac{dV_2}{dV} = \frac{\left[\frac{\epsilon_0 \epsilon_3}{d} \frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} - \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \frac{d\sigma_1}{dV} - \left(\frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \right) \frac{d\sigma_2}{dV} \right]}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}\right)}, \quad (5.95)$$

где для слоя Шоттки $f = 1$; для антизапорного слоя $f \gg 1$.

Полученные выражения для $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ могут быть проанализированы для различных частных случаев.

Пусть поверхностных состояний нет и $\frac{d\sigma_1}{dV} = \frac{d\sigma_2}{dV} = 0$.
Тогда:

$$\frac{dV_1}{dV} = \frac{\frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}}{\left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}\right)}; \quad (5.96)$$

$$\frac{dV_3}{dV} = \frac{1}{\left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}\right)}; \quad (5.97)$$

$$\frac{dV_2}{dV} = \frac{\frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d}}{\left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d}\right)}. \quad (5.98)$$

Как показывает анализ этих выражений, с увеличением толщины промежуточного слоя d $\frac{dV_1}{dV}$ и $\frac{dV_2}{dV}$ уменьшаются, а $\frac{dV_3}{dV}$ растет. При уменьшении концентрации n_1 растет $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и убывает $\frac{dV_2}{dV}$. При уменьшении концентрации n_2 растет $\frac{dV_2}{dV}$, $\frac{dV_3}{dV}$ и уменьшается $\frac{dV_1}{dV}$. Для хорошо развитого антизапорного слоя, когда $f \gg 1$ $\frac{dV_1}{dV} \rightarrow 0$, а выражения для $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ совпадают с аналогичными выражениями для контакта металл — полупроводник, полученными в п. 3.6.

Пусть ширина промежуточного слоя стремится к нулю и $d = 0$. Тогда выражения (5.96)–(5.98) принимают вид:

$$\frac{dV_1}{dV} = \frac{\frac{\mathcal{L}_1}{f \epsilon_1}}{\left(\frac{\mathcal{L}_2}{\epsilon_2} + \frac{\mathcal{L}_1}{f \epsilon_1}\right)}; \quad (5.99)$$

$$\frac{dV_3}{dV} = 0; \quad (5.100)$$

$$\frac{dV_2}{dV} = \frac{\frac{\mathcal{L}_2}{\epsilon_2}}{\frac{\mathcal{L}_2}{\epsilon_2} + \frac{\mathcal{L}_1}{f \epsilon_1}}. \quad (5.101)$$

В этом случае величина $\frac{dV_1}{dV}$ убывает, а величина $\frac{dV_2}{dV}$ растет при увеличении n_1 и уменьшении n_2 . Для антизапорного слоя, когда $f \gg 1$ $\frac{dV_1}{dV} = 0$, а $\frac{dV_2}{dV} = 1$.

Если $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ не равны нулю, то для подробного анализа $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ необходимо найти зависимости $m_1(V)$ и $m_2(V)$.

Для этого используем результаты предыдущего параграфа, где найдены величины f_i для различных частных случаев при учете обмена уровней с валентной зоной полупроводника 1 (аналогичный анализ показывает, что такие же выражения для f_i получаются при учете обмена уровня с зоной проводимости полупроводников 1 и 2). Учтем, что $\frac{dm_i}{dV} = n_i \frac{df_i}{dV}$.

Используя выражение (5.79), имеем $\frac{dm_i}{dV} = 0$ и $\frac{d\sigma_1}{dV} = \frac{d\sigma_2}{dV} = 0$.

Если f_i задается выражением (5.80), то можно получить, что $\frac{dm_i}{dV} < 0$, следовательно, $\frac{d\sigma_1}{dV} < 0$ и $\frac{d\sigma_2}{dV} > 0$.

Если f_i задается выражением (5.81), то $\frac{dm_i}{dV} > 0$, следовательно, $\frac{d\sigma_1}{dV} > 0$ и $\frac{d\sigma_2}{dV} > 0$.

Таким образом, в зависимости от механизма преимущественного обмена с полупроводником 1 или 2 и механизма прохождения тока $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ могут быть положительными или отрицательными.

Теперь проанализируем, как $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ влияют на величины производных $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$.

Если $\frac{d\sigma_1}{dV} > 0$, то $\frac{dV_1}{dV}$ растет, а $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ уменьшаются. При $\frac{d\sigma_1}{dV} > 0$ $\frac{dV_1}{dV}$ уменьшается, а $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$ растут.

Если $\frac{d\sigma_2}{dV} > 0$, то $\frac{dV_1}{dV}$ и $\frac{dV_3}{dV}$ растут, а $\frac{dV_2}{dV}$ уменьшается.

При обратном неравенстве $\frac{d\sigma_3}{dV} < 0$, $\frac{dV_1}{dV}$ и $\frac{dV_3}{dV}$ становятся меньше, а $\frac{dV_2}{dV}$ больше.

Параметры α и α^* без учета зависимости i_s от напряжения. Проведем анализ параметров α и α^* для различных частных случаев. Возможный вид функций f_1 и f_2 , полученный из анализа приведенных выше результатов, представлен в табл. 5.1. В этой же таблице даны номера использованных формул. При этом учитывалось, что в зависимости от параметров полупроводников 1 и 2, промежуточного слоя и приложенного напряжения для конкретного выражения вольт-амперной характеристики возможны участки с различными $f_1(V_1, V_3, V_2)$ и $f_2(V_1, V_3, V_2)$.

Таблица 5.1. Параметры f_1 и f_2 для различных частных случаев

$f_1(V_1, V_2, V_3)$		$f_2(V_1, V_2, V_3)$	
$\frac{e}{kT} V$	(5.1), (5.25), (5.26), (5.31), (5.33)	0	(5.1), (5.25), (5.26), (5.31), (5.45), (5.46), (5.54), (5.65), (5.66), (5.82)
$\frac{e}{kT} V_1$	(5.2), (5.13), (5.14), (5.26), (5.45), (5.46), (5.54), (5.65), (5.66), (5.82), (5.83)	$\frac{e}{kT} (V_1 + V_3)$	(5.2), (5.13), (5.14), (5.26), (5.33), (5.46)
$\frac{e}{kT} (V_2 + V_3)$	(5.3), (5.14), (5.26)	$\frac{e}{kT} V_1$	(5.3), (5.14), (5.26), (5.46), (5.65), (5.83)
$\frac{e}{kT} (V_2 + V_1)$	(5.46), (5.54), (5.65)	$\frac{e}{kT} (V_2 + V_1)$	(5.71)
$\frac{eV_2}{kT^*}$	(5.28)	$\frac{e}{kT} V - \frac{eV_2}{kT^*}$	(5.28)
0	(5.45), (4.46), (5.54), (5.66), (5.71)	$\frac{eV_2}{kT}$	(5.45), (5.65), (5.66), (5.71)

Как показывает анализ табл. 5.1, наиболее часто встречаются случаи, когда $f_1 = \frac{e}{kT} V_2$; $f_2 = 0$. С физической точки зрения это означает, что чаще всего прямой ток контролируется изменением высоты потенциального барьера полупроводника 2, а обратный ток не зависит от изменения потенциальных барьеров ни в полупроводнике 1, ни в полупроводнике 2.

Рассмотрение табл. 5.1 показывает также, что примерно для половины проанализированных случаев вольт-амперные характеристики описываются одной и той же формулой, где выполняется соотношение $f_1 + f_2 = \frac{e}{kT} V$. При этом $\alpha + \alpha^* = \frac{e}{kT}$.

Используя данные табл. 5.1 и ранее полученные значения производных $\frac{dV_1}{dV}$, $\frac{dV_3}{dV}$ и $\frac{dV_2}{dV}$, можно получить α и α^* для различных частных случаев.

Если $f_1 = \frac{e}{kT} V$ и $f_2 = 0$, то $\alpha = \frac{e}{kT}$; $\alpha^* = 0$.

Здесь параметры α и α^* не зависят от параметров гетероперехода.

Значения α и α^* для случая $f_1 = \frac{e}{k} V_2$, $f_2 = \frac{e}{kT} (V_1 + V_3)$:

$$\alpha = \frac{e}{kT} \frac{\left[\frac{\epsilon_0 \epsilon_3}{d} \frac{\mathcal{L}_2 f}{\epsilon_2 \mathcal{L}_1} - \frac{\epsilon_2 \mathcal{L}_2}{\epsilon_2 d} \frac{d\sigma_1}{dV} - \left(\frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \right) \frac{d\sigma_2}{dV} \right]}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d} \right)}; \quad (5.102)$$

$\alpha^* =$

$$= \frac{e}{kT} \frac{\frac{\epsilon_0 \epsilon_3}{d} + \frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} + \left(\frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \right) \frac{d\sigma_2}{dV} + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} \frac{d\sigma_1}{dV}}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d} \right)}$$

Анализ выражения для α здесь сводится к проведенному выше анализу выражения для $\frac{dV_2}{dV}$. Параметр $\alpha^* = \frac{e}{kT} - \alpha$ может быть проанализирован на основании анализа параметра α . Отметим, что $\alpha \sim \frac{1}{T}$ и $\alpha^* \sim \frac{1}{T}$.

Оценим величины α и α^* при $\frac{d\sigma_1}{dV} = \frac{d\sigma_2}{dV} = 0$. Примем: $\mathcal{L}_1 = \mathcal{L}_2 = 10^{-5}$ см; $d = 2 \cdot 10^{-7}$ см; $\epsilon_1 = \epsilon_3 = \epsilon_2 = 10$; $f = 1$. Тогда $\alpha = 0,49 e/kT$; $\alpha^* = 0,51 e/kT$. Если $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ отличны от нуля, α и α^* могут уменьшаться или увеличиваться по сравнению с приведенными значениями.

Если $f_1 = \frac{e}{kT} (V_2 + V_3)$; $f_2 = \frac{e}{kT} V_1$:

$$\alpha = \frac{e}{kT} \frac{\frac{\epsilon_0 \epsilon_2 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \right) - \left(1 + \frac{\epsilon_3 \mathcal{L}_1}{\epsilon_2 d} \right) \frac{d\sigma_1}{dV} + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \frac{d\sigma_2}{dV}}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d} \right)}; \quad (5.103)$$

$$\alpha^* = \frac{e}{kT} \frac{\frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} \frac{d\sigma_2}{dV} + \frac{\epsilon_0 \epsilon_3}{d} + \left(1 + \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} \right) \frac{d\sigma_1}{dV}}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d} \right)}$$

В этом случае $\alpha \sim \frac{1}{T}$, а его анализ сводится к проведенному выше анализу $\frac{dV_1}{dV}$. Параметр $\alpha = \frac{e}{kT} - \alpha^*$ не требует отдельного анализа. Оценка величин α и α^* при тех же параметрах, что и в предыдущем случае дает $\alpha = 0,51 \frac{e}{kT}$ и $\alpha^* = 0,49 \frac{e}{kT}$.

Величины $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ также могут изменить эти значения α и α^* .

Для случая $f_1 = f_2 = \frac{e}{kT} (V_2 + V_1)$:

$$\alpha = \frac{e}{kT} \frac{\epsilon_0 \epsilon_3 \left(1 + \frac{\epsilon_1 \mathcal{L}_2 f}{\epsilon_2 \mathcal{L}_1} \right) + \frac{d\sigma_1}{dV} - \frac{\epsilon_1 \mathcal{L}_2 f}{\epsilon_2 \mathcal{L}_1} \frac{d\sigma_2}{dV}}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_1 d} \right)}; \quad (5.104)$$

$\alpha^* = \alpha.$

Как показывает анализ выражения (5.104), с ростом ширины промежуточного слоя $\alpha = \alpha^*$ уменьшается. При уменьшении концентрации n_1 и n_2 величина α растет. В зависимости от знака $\frac{d\sigma_1}{dV}$ и $\frac{d\sigma_2}{dV}$ $\alpha = \alpha^*$ может увеличиваться и уменьшаться. Как и ранее, параметры α и α^* пропорциональны $\frac{1}{T}$.

Следует особо подчеркнуть, что для последнего случая α и α^* относятся к разным вольт-амперным характеристикам. Аналогичные оценки показывают, что при принятых выше параметрах $\alpha = \alpha^* = 0,98 e/kT$.

Если $f_1 = \frac{eV_2}{kT}$; $f_2 = \frac{e}{kT} V - \frac{eV_2}{kT}$, то:

$$\alpha = \frac{e}{kT^*} \frac{\left[\frac{\epsilon_0 \epsilon_3}{d} \frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} - \frac{\epsilon_3}{\epsilon_2} \frac{\mathcal{L}_2}{d} \frac{d\sigma_1}{dV} - \left(\frac{\epsilon_1}{\epsilon_2} \frac{\mathcal{L}_2 f}{\mathcal{L}_1} + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} \right) \frac{d\sigma_2}{dV} \right]}{\frac{\epsilon_0 \epsilon_1 f}{\mathcal{L}_1} \left(1 + \frac{\epsilon_3 \mathcal{L}_2}{\epsilon_2 d} + \frac{\epsilon_3 \mathcal{L}_1}{f \epsilon_2 d} \right)}$$

$$\alpha^* = \frac{e}{kT} - \alpha. \quad (5.105)$$

Анализ зависимости α от параметров полупроводников 1 и 2 и промежуточного слоя, связанный с величиной $\frac{dV_2}{dV}$, был проведен выше. К этому теперь надо добавить анализ зависимости T^* от этих параметров. Такой анализ был проведен в гл. 3 при рассмотрении работы контакта металл — полупроводник. В частности, было показано, что при $E_{00} \gg kT$ величина T^* и туннелирование существенным образом уменьшают величину α и увеличивают величину α^* .

Отметим, что принципиальным отличием последнего случая, когда существенно туннелирование, от всех предыдущих случаев является малая зависимость параметра α от температуры. При этом α^* от температуры может зависеть сильно и быть пропорциональной $\frac{1}{T}$.

Влияние зависимости i_s от напряжения на α и α^* . Поправки к величинам α и α^* за счет члена $\frac{d \ln i_s}{dV}$ могут быть найдены, как и для контакта металл — полупроводник в гл. 3.

Например, для случая, когда $i_s \sim \mathcal{P}_n$ (см. (5.31) — (5.34), (5.76), (5.82), (5.83)), а величина \mathcal{P}_n задается выражением (5.76),

$$\frac{d \ln i_s}{dV} = \frac{e\sigma_k}{2(\varphi_2 - eV_2)^{3/2}} \frac{dV_2}{dV}.$$

Оценка последнего выражения показывает, что при $\mathcal{P}_n = 10^{-5}$ и $(\varphi_2 - eV_2) = 1\text{В}$ $\frac{d \ln i_s}{dV} \simeq 11,5 \frac{dV_2}{dV} \text{В}^{-1}$. Это существенная поправка к величинам α и α^* . Аналогичные поправки могут быть найдены и для других выражений \mathcal{P}_n , использованных в настоящей

главе. Отметим, что учет туннелирования и величины \mathcal{P}_n приводит к более слабой зависимости параметров α и α^* от температуры. В общем виде теперь можно записать:

$$\alpha = A + \frac{B}{T}; \quad \alpha^* = C + \frac{D}{T}.$$

5.8. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ВОЛЬТ-АМПЕРНЫХ ХАРАКТЕРИСТИК ГЕТЕРОПЕРЕХОДОВ

В настоящее время экспериментально исследовано большое количество различных изотипных и анизотипных гетеропереходов, причем основное внимание уделялось гетеропереходам Ge—Si, Ge—GaAs, GaAs—AlAs и ряду других.

Для выяснения возможных механизмов прохождения тока в гетеропереходах обычно исследовались вольт-амперные характеристики и их зависимость от температуры.

Оказалось, что прямые ветви гетеропереходов можно описать с помощью одного или нескольких участков, подчиняющихся закону,

$$i = i_s e^{\alpha V} = i_s e^{\frac{eV}{nkT}}, \quad (5.106)$$

а обратные ветви описывались с помощью либо аналогичного выражения с заменой α на α^* , либо подчинялись степенному закону

$$i = BV^m, \quad (5.107)$$

как это в общем и следует из рассмотренных выше теоретических выражений.

Конкретную интерпретацию экспериментальных вольт-амперных характеристик рассмотрим на примере Ge—Si гетеропереходов. Это связано с тем, что, во-первых, такие гетеропереходы относятся к одним из наиболее изученных, во-вторых, в таких гетеропереходах существенны поверхностные состояния и модель токопереноса должна быть более общей.

Анизотипные гетеропереходы. Типичные прямые ветви вольт-амперных характеристик сплавного гетероперехода $p\text{-Ge—}n\text{-Si}$, снятые при различных температурах, показаны на рис. 5.8. Как видно из этого рисунка, прямые ветви вольт-амперных характеристик имеют три участка при низких температурах, а при высоких один, которые описываются выражением (5.106) с разными i_s и α .

По мнению некоторых исследователей, участки при малых токах связаны с поверхностной утечкой, по мнению иных, эти участки связаны с избыточными туннельными токами, поскольку эти участки описываются выражением (5.106) с α слабо зависящим от температуры. Однако, поскольку такие токи наблюдаются и в образцах с малой концентрацией ($\leq 10^{16} \text{ см}^{-3}$), то для существования туннельных токов следует предполагать наличие областей $p\text{—}n$ перехода, где локальное легирование значительно больше среднего.

Это серьезное, но пока мало обоснованное дополнительное предположение.

Вторые и третьи участки прямых ветвей вольт-амперных характеристик могут быть описаны формулой (5.106), где $\alpha = \left(A + \frac{B}{T}\right) V$.

Исследование кинетики установления токов показали, что инжекция в области токов, соответствующих второму и третьему участкам, практически отсутствует. Все это позволяет привлечь для объяснения указанных участков вольт-амперной характеристики туннельно-рекомбинационные токи. Относительно места их протекания и их локальности реализуется изложенное выше.

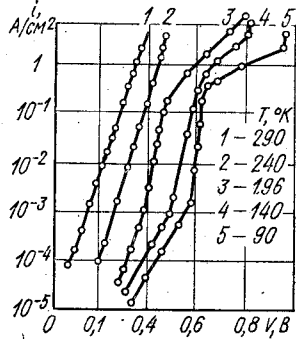


Рис. 5.8. Прямые ветви вольт-амперных характеристик сплавных гетеропереходов $p\text{-Ge}-n\text{-Si}$ при разных температурах.

Используя экспериментальные результаты и теоретические формулы, можно оценить некоторые параметры туннельно-рекомбинационной модели. В частности, если $n_i = 10^{13} \text{ см}^{-3}$, сечение захвата уровней ($10^{-15} \dots 10^{-14}$) см^2 .

Типичные обратные ветви сплавных гетеропереходов $p\text{-Ge}-n\text{-Si}$ показаны на рис. 5.9. Как видно из этого рисунка, обратные вольт-амперные характеристики удовлетворительно описываются формулой (5.107), где m имеет разное значение для различных участков.

При значениях $m \approx 1/2$ обратная ветвь вольт-амперной характеристики может быть объяснена генерационно-рекомбинационными процессами в области пространственного заряда. При больших m следует привлекать другие механизмы прохождения токов. Например, $m \geq 3$ может быть получено, если предположить, что существенную роль играет туннелирование электронов из валентной зоны широкозонного материала $n-p$ гетероперехода в зону проводимости узкозонного, как это рассмотрено в п. 5.3 (строго говоря, при этом должна наблюдаться экспоненциальная зависимость тока от напряжения).

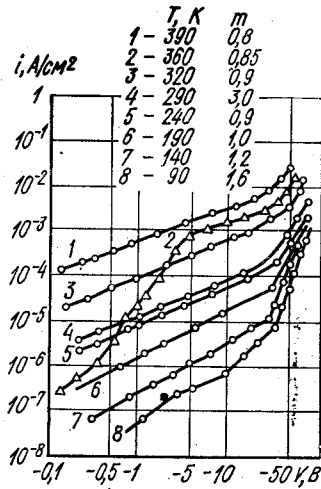


Рис. 5.9. Обратные ветви вольт-амперных характеристик сплавных гетеропереходов $p\text{-Ge}-n\text{-Si}$ при разных температурах.

Изотипные гетеропереходы. Наиболее подробные исследования изотипных гетеропереходов проводились для структур $n\text{-Ge}-n\text{-Si}$. Полученные вольт-амперные характеристики можно было разделить на три группы:

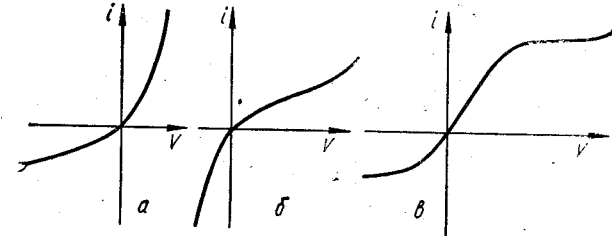


Рис. 5.10. Вольт-амперные характеристики изотипного гетероперехода $\text{Ge}-\text{Si}$.

с хорошим выпрямлением и прямым током, соответствующим энергетической модели п. 5.1 (рис. 5.10, а); с хорошим выпрямлением, но противоположной полярности (рис. 5.10, б); с двойным насыщением в прямом и обратном направлениях (рис. 5.10, в).

Первые два типа вольт-амперных характеристик могут быть объяснены изменением падения напряжения в полупроводниках 1 и 2 как без привлечения, так и с привлечением поверхностных уровней в качестве пересадочных центров для носителей заряда, участвующих в переносе тока через гетеропереход.

Например, изменяя степень легирования полупроводников 1 и 2, можно получить $V_1 \ll V_2$ или $V_1 \gg V_2$. В первом случае прямой ток контролируется напряжением V_2 , а обратный изменяется мало, во втором прямой ток изменяется мало, а обратный контролируется напряжением V_1 (см. табл. 5.1). Это подтверждено экспериментально и иллюстрируется рис. 5.11, а, б, где показаны равновесные диаграммы энергетических зон соответствующих $n-n$ гетеропереходов $\text{Ge}-\text{Si}$ при разных уровнях легирования, соответствующие вольт-амперным характеристикам рис. 5.10, а, б.

Вольт-амперная характеристика изотипного гетероперехода с двойным насыщением (рис. 5.10, в) никак не может быть объяснена простым перераспределением напряжения между полупроводниками. Как уже анализировалось в п. 5.4, для ее объяснения необходимо обязательно привлекать участие поверхностных уровней в переносе тока (см. (5.46)).

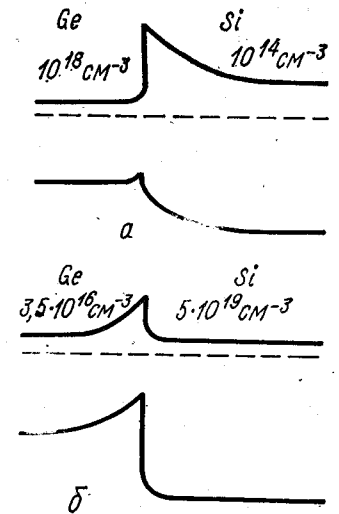


Рис. 5.11. Равновесные диаграммы энергетических зон, соответствующие рис. 5.10.

Поскольку в предыдущих двух случаях исследованные гетеропереходы получались по той же технологии, т. е. имеют тот же порядок концентрации поверхностных уровней, то нет оснований не учитывать их участия в переносе тока.

Таким образом, все многообразие вольт-амперных характеристик изотипного гетероперехода Ge—Si вида запиорный — запиорный слой может быть объяснено только с привлечением участия поверхностных уровней в переносе тока.

Аналогичные результаты при исследовании вольт-амперных характеристик получены и на других анизотипных и изотипных гетеропереходах. Характерным для этих результатов, как и для рассмотренного примера гетероперехода Ge—Si, является то, что для объяснения сложного характера вольт-амперной характеристики привлекают обычно несколько механизмов прохождения тока. И, несмотря на то, что уже накоплен обширный экспериментальный материал, удовлетворительное объяснение вольт-амперных характеристик удалось привести только для отдельных случаев, причем полученные параметры модели часто лежат за пределами физически разумных. Все это связано со следующим. Во-первых, теория гетеропереходов создана для упрощенных моделей. В реальном гетеропереходе переходной слой может иметь сложную структуру, поверхностные уровни могут быть распределены сложным образом по пространству промежуточного слоя и по энергиям. Во-вторых, к настоящему времени не разработан полный комплекс методик, который бы позволял независимым образом находить все параметры, определяющие вольт-амперную характеристику гетероперехода. В результате в настоящее время невозможна независимая проверка принимаемой для объяснения экспериментальных данных модели гетероперехода.

Глава 6. ПРОБОЙНЫЕ ЯВЛЕНИЯ

6.1. ВИДЫ ПРОБОЯ

При больших обратных напряжениях в гомо- и гетеропереходах экспериментально наблюдается резкий рост тока с приложенным напряжением, связанный с появлением дополнительных носителей заряда, принимающих участие в переносе тока (рис. 6.1). Рассмотрим возможные физические причины, которые вызывают появление таких добавочных носителей заряда. При этом ограничимся рассмотрением процессов в области пространственного заряда и не будем рассматривать процессы в промежуточных слоях, процессы, связанные с локальными уровнями и т. п. Это позволит рассмотреть пробойные явления в контакте металл — полупроводник, $p-n$ переходе и гетеропереходе на общем примере $p-n$ перехода.

В настоящее время различают три основные разновидности пробоя $p-n$ перехода: лавинный, туннельный (зинеровский), тепловой.

Суть явления лавинного пробоя заключается в том, что носители заряда в сильном поле области пространственного заряда приобретают энергию, достаточную для ионизации атомов полупроводника, т. е. для образования новых носителей заряда. Эти носители заряда в свою очередь приводят к возникновению новых электронов и дырок и т. д. (рис. 6.2). Количество носителей заряда, участвующих в переносе тока, резко возрастает и наступает пробой.

При туннельном пробое рост носителей заряда, принимающих участие в переносе тока, происходит вследствие прямого туннелирования электронов из валентной зоны в зону проводимости, как показано на рис. 6.3. Такой процесс происходит без изменения энергии и, следовательно, столкновения носителей заряда с атомами решетки полупроводника.

Как ясно из изложенного выше, лавинный пробой будет происходить в достаточно широких областях пространственного заряда (слабо легированные полупроводники), когда наблюдаются многократные столкновения носителей заряда с решеткой в области пространственного заряда, т. е. при справедливости предположений, характерных для диффузионной теории. Туннельный пробой наблюдается для узких областей пространственного заряда (сильно легированные полупроводники), в которых нет столкновений с решеткой и осуществляются критерии диодной теории. Оба рассмотренные вида пробоя являются электрическими и зависят от величины поля в области пространственного заряда. Наблюдаемые при лавинном

и туннельном пробоях вольт-амперные характеристики показаны на рис. 6.1.

В зависимости от места локализации максимальных полей они могут осуществляться в объеме или на поверхности $p-n$ перехода. Связано это с тем, что заряд на поверхностных уровнях в месте выхода $p-n$ перехода на поверхность может изменять ширину области пространственного заряда и, следовательно, величину напряженности электрического поля. Это видно из рис. 6.4. Если знак поверхностного заряда совпадает со знаком ионизированных примесей в более высокоомной области полупроводника, то напряженность поля у поверхности будет выше, чем в объеме. Вероятность пробоя на поверхности при этом больше, чем в объеме $p-n$ перехода. Если же заряд на поверхностных уровнях противоположен заряду ионизированных примесей в слабо легированной области, то напряженность поля у поверхности уменьшается и пробой будет возникать в объеме полупроводника. Физически это определяется тем, что заряд более сильно легированной области компенсируется не только зарядом слабо легированной области, но и поверхностным зарядом.

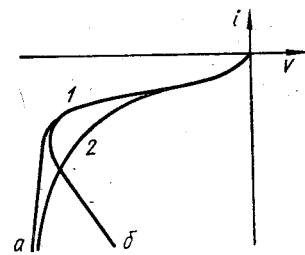


Рис. 6.1. Типичные вольт-амперные характеристики при больших обратных напряжениях: а — лавинный и туннельный пробой; б — тепловой пробой.

Физически это определяется тем, что заряд более сильно легированной области компенсируется не только зарядом слабо легированной области, но и поверхностным зарядом.

В реальных $p-n$ переходах обычно существует некоторое количество дефектов: нарушения кристаллической решетки, неоднородность в распределении примесей и т. п. Все это приводит к тому,

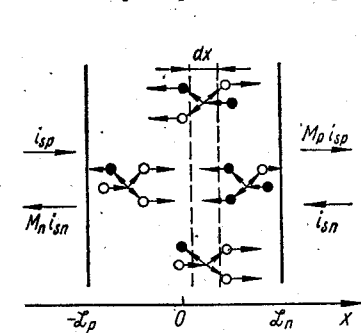


Рис. 6.2. Схема лавинного пробоя.

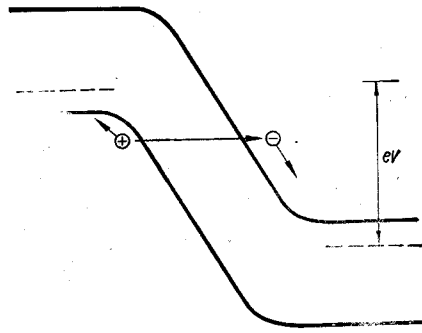


Рис. 6.3. Электронные переходы при туннельном пробое.

что в области таких дефектов возникают неоднородности электрического поля. В местах максимального поля и начинает развиваться электрический пробой, который в этом случае является локальным. Такой локальный пробой также может быть объемным и поверхностным. Если локальный пробой развивается в области очень малого сечения, то вследствие большого последовательного сопротивления для области пробоя его влияние на вольт-амперную харак-

теристику несущественно и наблюдается просто более плавный рост тока с обратным напряжением по сравнению с нелокальным пробоем (см. рис. 6.1, кривые 1, 2).

Особым видом локального электрического пробоя является спорадический пробой, когда образуются импульсы тока и возникают микроплазмы.

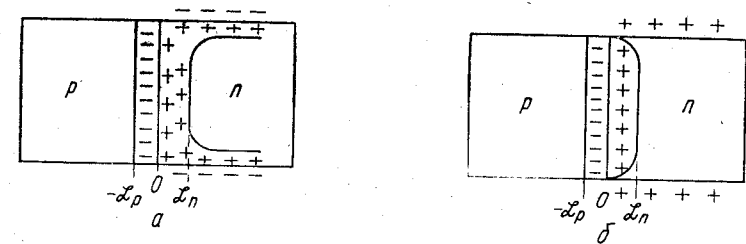


Рис. 6.4. К влиянию поверхности на электрический пробой: а — поверхность ослабляет пробой; б — поверхность усиливает пробой.

Последний, тепловой вид пробоя связан с тем, что при прохождении тока через $p-n$ переход может происходить его нагрев. Этот нагрев приводит к увеличению числа носителей заряда и, следовательно, к росту тока. Такой вид пробоя существенным образом зависит от теплоотвода из области $p-n$ перехода. Он может наблюдаться или не наблюдаться в тех же $p-n$ переходах в зависимости от условий теплоотвода. При этом виде пробоя наблюдается вольт-амперная характеристика типа показанной на рис. 6.1, б.

6.2. ЛАВИННЫЙ ПРОБОЙ

Выведем основное уравнение ударной ионизации, определяющее условие лавинного пробоя. Будем считать, что в область $p-n$ перехода из p -области входит дырочный ток $i_{sp}(-L_p)$. Если поле в области $p-n$ перехода имеет достаточную величину, то дырочный ток с расстоянием увеличивается и при выходе из области пространственного заряда имеет величину $i_p(L_n) = M_p i_{sp}(-L_p)$, где M_p — коэффициент лавинного умножения дырок. Аналогично ток электронов i_{sn} будет увеличиваться при прохождении $p-n$ перехода справа налево и через границу $p-n$ перехода выходит ток $i_n(L_p) = M_n i_{sn}(L_n)$ (см. рис. 6.2). Суммарный ток $i = i_p(L_n) + i_n(-L_p)$ в стационарном состоянии является постоянным.

Приращение дырочного потока в точке x внутри области пространственного заряда $p-n$ перехода равно числу электронно-дырочных пар, генерируемых в секунду на расстоянии dx :

$$d \frac{i_p}{e} = \frac{i_p}{e} \alpha_p dx + \frac{i_n}{e} \alpha_n dx, \quad (6.1)$$

где α_p и α_n — коэффициенты ударной ионизации для дырок и электронов. Последнее выражение после добавления и вычитания

$\frac{i_p}{e} \alpha_n dx$ может быть переписано в виде

$$\frac{di_p}{dx} - (\alpha_p - \alpha_n) i_p = \alpha_n i. \quad (6.2)$$

Уравнение (6.2) имеет вид $y' + \mathcal{P}y = Q$, где $y = i_p$. Решение его

$$y = \left[\int_{-\mathcal{L}_p}^x Q e^{-\int_{-\mathcal{L}_p}^x \mathcal{P} dx'} dx' + C \right] / e^{-\int_{-\mathcal{L}_p}^x \mathcal{P} dx'}$$

В результате при граничных условиях:

$$x = -\mathcal{L}_p, \quad i_p = i_{sp}(-\mathcal{L}_p), \quad i_n = i_{sp}(-\mathcal{L}_p);$$

$$x = \mathcal{L}_n, \quad i_p = i - i_{sn}(\mathcal{L}_n), \quad i_n = i_{n0}(\mathcal{L}_n)$$

имеем

$$i_p(x) = \exp \left[\int_{-\mathcal{L}_p}^x (\alpha_p - \alpha_n) dx' \right] \left\{ i \int_{-\mathcal{L}_p}^x \alpha_n \exp \left[- \int_{-\mathcal{L}_p}^{x'} (\alpha_p - \alpha_n) dx'' \right] + i_{sp} \right\}. \quad (6.3)$$

В случае, если лавинный процесс определяется дырками, из (6.3), используя тождество

$$\int_{-\mathcal{L}_p}^{\mathcal{L}_n} \alpha_n \exp \left[- \int_{-\mathcal{L}_p}^x (\alpha_p - \alpha_n) dx' \right] dx = \exp \left[- \int_{-\mathcal{L}_p}^{\mathcal{L}_n} (\alpha_p - \alpha_n) dx \right] -$$

$$- 1 + \int_{-\mathcal{L}_p}^{\mathcal{L}_n} \alpha_p \exp \left[- \int_{-\mathcal{L}_p}^x (\alpha_p - \alpha_n) dx' \right] dx,$$

можно получить

$$1 - \frac{1}{M_p} = \int_{-\mathcal{L}_p}^{\mathcal{L}_n} \alpha_p \exp \left[- \int_{-\mathcal{L}_p}^x (\alpha_p - \alpha_n) dx' \right] dx. \quad (6.4)$$

Напряжение лавинного пробоя V_{np} можно определить как напряжение, при котором величина M_p стремится к бесконечности. При этом выражение (6.4) переписывается в виде

$$\int_{-\mathcal{L}_p}^{\mathcal{L}_n} \alpha_p \exp \left[- \int_{-\mathcal{L}_p}^x (\alpha_p - \alpha_n) dx' \right] dx = 1. \quad (6.5)$$

Если лавинный процесс определяется не дырками, а электронами, то аналогичные рассуждения приводят к условию

$$\int_{-\mathcal{L}_p}^{\mathcal{L}_n} \alpha_n \exp \left[- \int_{-\mathcal{L}_p}^x (\alpha_n - \alpha_p) dx' \right] dx = 1. \quad (6.6)$$

Полученные выражения в общем виде еще сложны для аналитического расчета, поскольку α_n и α_p являются функциями напряженности электрического поля. В общем виде они решаются только с помощью ЭВМ.

Для аналитического расчета можно использовать следующие приближения для связи α_p и α_n , а также зависимости α_p от напряженности поля:

$$\alpha_p = K \alpha_n; \quad \alpha_p = A E^m, \quad (6.7)$$

где K и m — постоянные величины.

Тогда выражение (6.5) принимает вид

$$\frac{K \ln K}{K-1} = A \int_{-\mathcal{L}_p}^{\mathcal{L}_n} E^m dx. \quad (6.8)$$

Подставляя в (6.8) $E(x)$ для слоя Шоттки при равномерном распределении примесей для несимметричного $p-n$

перехода, когда $n_n \ll p_p$ и $\mathcal{L} = \mathcal{L}_n$, в виде $E(x) = -\frac{en_n}{e\epsilon_0} (\mathcal{L} - x)$ получаем

$$V_{np} = \frac{1}{2} \left[\frac{\ln K}{K-1} \frac{m+1}{A} \right]^{\frac{2}{m+1}} \left(\frac{e\epsilon_0}{e} \right)^{\frac{m-1}{m+1}} n_n^{\frac{1-m}{m+1}}. \quad (6.9)$$

Таким образом, пробивное напряжение резкого несимметричного

$p-n$ перехода падает с ростом концентрации как $n_n^{\frac{1-m}{m+1}}$. На рис. 6.5 приведены рассчитанные по (6.9) зависимости пробивного напряжения для ряда полупроводников. Можно показать, что они для Ge и Si довольно хорошо подчиняются закону

$$V_{np} = a \rho^{0.8}, \quad (6.10)$$

где ρ — удельное сопротивление полупроводника.

В случае лавинного пробоя V_{np} может зависеть также от температуры. Это определяется тем, что с изменением температуры изменяется длина свободного пробега и энергия, которую может набрать носитель заряда между столкновениями. Поскольку при рассеянии на колебаниях решетки длина свободного пробега с ростом температуры уменьшается, то энергия, приобретаемая носителем заряда на длине свободного пробега, также уменьшается. Это приводит к тому, что с ростом температуры при лавинном пробое напряжение пробоя увеличивается.

Вольт-амперная характеристика при лавинном пробое может быть получена, если использовать эмпирическую зависимость

$$M = \frac{1}{1 - \left(\frac{V}{V_{np}} \right)^b}, \quad (6.11)$$

где для Ge и Si $b = 3, \dots 5.5$.

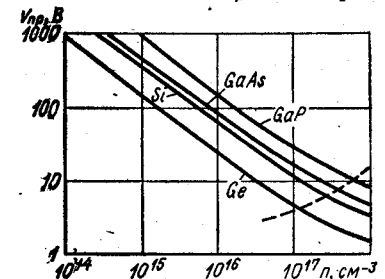


Рис. 6.5. Зависимость напряжения лавинного пробоя от концентрации примесей n для несимметричного резкого $p-n$ перехода.

Подставляя (6.11) в

$$i = i_p(\mathcal{L}_n) + i_n(-\mathcal{L}_p) = M_p i_{p0}(-\mathcal{L}_p) + M_n i_{n0}(\mathcal{L}_n),$$

получаем

$$i = \frac{i_s}{1 - \left(\frac{V}{V_{пр}}\right)^b}, \quad (6.12)$$

где $i_s = i_{sn} + i_{sp}$ — ток насыщения $p-n$ перехода.

При увеличении обратного тока следует также учитывать падение напряжения на объемном сопротивлении p - и n -областей. Это приводит к сглаживанию вольт-амперной характеристики в области пробоя по сравнению с тем, когда падением ее в объеме полупроводника можно пренебречь.

6.3. ТУННЕЛЬНЫЙ ПРОБОЙ

Для нахождения вольт-амперной характеристики при туннельном пробое можно воспользоваться результатами, полученными в п. 4.7 для туннельных диодов. При этом, подставляя (4.80) в (4.79), для обратного тока легко получить

$$i = \frac{\sqrt{2m^*} e^2 EV}{4\pi\hbar^2 \mathcal{E}_g^{1/2}} \exp\left(-\frac{4\sqrt{m^*} \mathcal{E}_g^{3/2}}{3eE\hbar}\right). \quad (6.13)$$

Как видно из последней формулы, туннельный ток i сильно зависит от величины поля E .

Можно ввести критическое значение напряженности поля $E_{кр}$, когда плотность туннельного тока становится преобладающей над плотностью обычного надбарьерного тока и начинается пробой. Поскольку максимальное поле в $p-n$ переходе реализуется в точке $x = 0$, то можно использовать это значение в качестве выражения для $E_{кр}$.

Для случая, когда и в n - и p -полупроводниках область пространственного заряда является слоем Шоттки, при равномерном распределении примесей:

$$E_{кр} = -\frac{en_n \mathcal{L}_n}{\epsilon\epsilon_0} = -\frac{ep_p \mathcal{L}_p}{\epsilon\epsilon_0}. \quad (6.14)$$

Здесь использовано равенство величины E на границе раздела n - и p -областей.

Подставляя в это выражение \mathcal{L}_n или \mathcal{L}_p , легко получить

$$V_{пр} = \frac{E_{кр}^2 \epsilon\epsilon_0}{2e^2} \left(\frac{1}{p_p} + \frac{1}{n_n}\right) + \frac{\Phi_0}{e}. \quad (6.15)$$

Как видно из последнего выражения, при туннельном пробое и $E_{кр} = \text{const}$ пробивное напряжение обратно пропорционально первой степени концентрации примесей в n - и p -областях.

При пренебрежении зависимостью подвижности от концентрации и $\frac{\Phi_0}{e} \ll V_{пр}$ выражение (6.15) можно переписать в виде

$$V_{пр} = a\rho_n + b\rho_p, \quad (6.16)$$

где ρ_n и ρ_p — удельные сопротивления n - и p -областей; a и b — постоянные. Величина $V_{пр}$ уменьшается прямо пропорционально удельному сопротивлению и туннельный пробой более существенен для сильно легированных полупроводников.

Зависимость $V_{пр}$ от температуры при туннельном пробое определяется, главным образом, зависимостью от температуры величины ширины запрещенной зоны \mathcal{E}_g . Для Ge, Si, GaAs ширина запрещенной зоны уменьшается с ростом температуры, а это означает, что напряжение туннельного пробоя в этих полупроводниках, в отличие от лавинного пробоя, с температурой уменьшается. Однако эти изменения невелики и для туннельного пробоя обычно предполагают, что напряжение пробоя на зависит от температуры.

6.4. ТЕПЛОВОЙ ПРОБОЙ

Вид вольт-амперной характеристики при тепловом пробое может быть определен из условий теплового баланса в $p-n$ переходе. При стационарном режиме количество теплоты, выделяемой в $p-n$ переходе, должно равняться количеству теплоты, отводимой от него.

Выделяемая на $p-n$ переходе в виде теплоты мощность

$$\mathcal{P}_{\text{выд}} = VI, \quad (6.17)$$

а отводимая от $p-n$ перехода мощность

$$\mathcal{P}_{\text{отв}} = \frac{T - T_0}{R_T}, \quad (6.18)$$

где T и T_0 — температура $p-n$ перехода и окружающей среды соответственно, R_T — тепловое сопротивление $p-n$ перехода.

Приравнявая $\mathcal{P}_{\text{выд}} = \mathcal{P}_{\text{отв}}$, имеем

$$VI = \frac{T - T_0}{R_T}. \quad (6.19)$$

В этом выражении надо учесть, что обратный ток I является функцией температуры.

В самом деле, как показывает анализ результатов гл. 4, величина i_s для $p-n$ перехода с широкой и узкой базой может изменяться от температуры из-за изменения с температурой величин D_p , l_p , D_n , l_n , а также величин p_n и n_p . Если пренебречь более слабой зависимостью от напряжения первых четырех величин, то i_s будет

определяться только величинами $p_n = \frac{n_i^2}{n_n}$ и $n_p = \frac{n_i^2}{p_p}$. Когда в интересующем нас температурном интервале концентрация основных

носителей заряда не изменяется, то $i_s \sim n_i^2 \sim e^{-\frac{\mathcal{E}_g}{kT}}$. Это позволяет

записать обратный ток через $p-n$ переход с широкой и узкой базой в виде

$$I = I_0 e^{-\frac{\xi_g}{kT}}, \quad (6.20a)$$

где I_0 является слабой функцией напряжения.

В том случае, когда обратный ток $p-n$ перехода определяется генерацией носителей заряда в области пространственного заряда аналогичный анализ показывает, что $i_s \sim n_i$, следовательно:

$$I = I_0 e^{-\frac{\xi_g}{2kT}}, \quad (6.20b)$$

где I_0 также слабо зависит от температуры.

Подставляя величину T , найденную из выражения (6.20a) или (6.20b), в (6.19), легко найти выражение обратной ветви вольт-амперной характеристики при тепловом пробое:

$$V = \frac{\frac{\xi_g}{mK \ln \frac{I_0}{I}} - T_0}{IR_T}, \quad (6.21)$$

где $m = 1$ для диффузионных токов $p-n$ перехода и $m = 2$ для генерационных токов.

Полученное выражение задает вольт-амперную характеристику вида рис. 6.1, б. Это связано с тем, что обратный ток увеличивается с увеличением температуры. Таким образом, рост тока ведет к росту выделяемой мощности и, следовательно, температуры, а рост температуры — к росту тока. При этом для малых нагревов ток $I = I_{\text{const}} + \Delta I$, где $\Delta I \ll I_{\text{const}}$ и рост V определяется увеличением числителя выражения (6.21). Для достаточно больших нагревов $\Delta I \gg I_{\text{const}}$ и величина V уменьшается, так как главную роль начинает играть рост знаменателя выражения (6.21).

Используя выражение (6.21), можно получить I_m , T_m и V_m , соответствующие точке перегиба на вольт-амперной характеристике. При этом для простоты учтем, что температура $p-n$ перехода T не сильно отличается от T_0 и $T - T_0 \ll T_0$. Тогда выражение (6.20) может быть переписано в виде

$$I = I_0 e^{-\frac{\xi_g}{mKT}} e^{\frac{\xi_g}{mKT_0}} e^{-\frac{\xi_g}{mKT_0}} = I_{\infty} e^{-\frac{\xi_g}{mKT_0}} (T=T_0), \quad (6.22)$$

где $I_{\infty} = I_0 e^{-\frac{\xi_g}{mKT_0}}$.

Из выражения (6.22) можно найти величину $(T - T_0)$. Подставляя эту величину в выражение (6.19), можно получить связь напряжения V с током I в виде

$$V = \frac{mKT_0^2 \ln \frac{I}{I_{\infty}} - kT_0^2}{IR_T \xi_g}, \quad (6.23)$$

Дифференцируя последнее выражение по току и приравнявая производную нулю, можно получить, что

$$1 - \ln \frac{I}{I_{\infty}} = 0.$$

Отсюда ток, соответствующий точке, когда $\frac{dV}{dI} = 0$:

$$I_m = I_{\infty} e = I_0 e^{-\frac{\xi_g}{mKT_0} + 1}. \quad (6.24)$$

Как видно из (6.24), температура, соответствующая точке, когда $\frac{dV}{dI} = 0$:

$$T_m = T_0 \left(1 + \frac{mKT_0}{\xi_g} \right). \quad (6.25)$$

Значение $V_m = V_{\text{пр}}$ может быть получено после подстановки I_m в (6.23):

$$V_{\text{пр}} = \frac{mKT_0^2}{R_T \xi_g I_0 e^{-\frac{\xi_g}{mKT_0} + 1}}. \quad (6.26)$$

Как показывает анализ выражения (6.28), для теплового механизма пробоя напряжение пробоя зависит от теплового сопротивления R_T , величины m , характеризующей механизм прохождения обратного тока, ширины запрещенной зоны ξ_g , а также температуры окружающей среды T_0 . Напряжение теплового пробоя больше для малых значений R_T , т. е. при хорошем теплоотводе от $p-n$ перехода. Кроме того, напряжение теплового пробоя будет большим для полупроводников с большей шириной запрещенной зоны и меньшей температурой окружающей среды. Это определяется меньшими значениями начальных токов через $p-n$ переход при малых значениях напряжения, когда нагрев еще не существен. Этим же определяется зависимость $V_{\text{пр}}$ от механизма прохождения тока.

6.5. МИКРОПЛАЗМА

Как уже отмечалось, локализация тока в отдельных участках может приводить к возникновению микроплазм, в результате чего при обратных напряжениях возникают импульсы тока. Амплитуда таких импульсов растет с увеличением обратного напряжения. Это сопровождается увеличением длительности импульсов и уменьшением их скважности вплоть до достижения непрерывного режима работы.

Для объяснения возникновения микроплазм наиболее широко используется модель, основанная на аналогии с явлениями в газовом разряде. В случае $p-n$ перехода эта модель сводится к введению локальных участков с $p-i-n$ структурой. Распределение поля в этой структуре имеет вид, представленный на рис. 6.6. Когда микроплазма не включена, количество подвижных носителей заряда

в канале микроплазмы мало и ток через нее практически равен нулю даже при напряжении, превышающем напряжение пробоя. Поле в *i*-слое при этом приблизительно постоянно. Как только какой-либо носитель заряда в канале микроплазмы «запустит» лавину, микроплазма включается. При этом в средней части *i*-области высокая плотность электронов и дырок взаимно компенсируется и поле мало. Образование максимума электрического поля на краях *i*-области приводит к росту в этих областях коэффициента ударной ионизации $\alpha_{n,p}$, который не может быть скомпенсирован уменьшением величины $\alpha_{n,p}$ в центре, так как $\alpha_{n,p} \sim E^m$, где $m \gg 1$. Таким образом, пространственный заряд, связанный с протеканием тока в *p-i-n*-структуре, приводит к увеличению полного умножения $M_{n,p}$. Увеличение полного умножения приведет к дополнительному

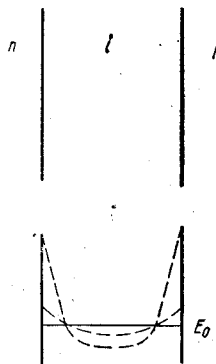


Рис. 6.6. К возникновению микроплазм в *p-i-n* структурах.

росту тока, искажению поля и т. д. Нарастание тока будет продолжаться до тех пор, пока поле на краях запирающего слоя не достигает таких значений, когда зависимость $\alpha_{n,p}$ от напряженности поля *E* не ослабится. При этом дополнительное увеличение тока не будет приводить к росту умножения носителей. К стабилизации тока ведут также такие эффекты, как разогрев *p-n*-перехода, диффузия носителей заряда из плазмы и другие.

Основным механизмом, приводящим к выключению микроплазмы, считается вылет всех носителей из области микроплазмы без единого акта ионизации. На процесс выключения может также влиять диффузия носителей из канала микроплазмы в радиальном направлении. В результате выключения снова будет устанавливаться примерно постоянное поле (рис. 6.6) и т. д.

При увеличении тока растет амплитуда импульсов. Это приводит к росту длительности импульсов вследствие того, что уменьшается вероятность достаточно больших флуктуаций тока в канале микроплазмы, прекращающих ток. Для больших напряжений достигается стабильная область, когда разряд в микроплазме осуществляется непрерывно. В стабильной области микроплазма обладает положительным сопротивлением.

В общем случае для возникновения микроплазм требуются структурные несовершенства и дефекты, приводящие либо к локальному увеличению электрического поля, либо к локальному увеличению коэффициентов ионизации.

В качестве таких несовершенств в кремнии могут служить, например, включения SiO_2 , карбида кремния, дислокации и т. д. Большую роль в образовании микроплазм могут также играть поверхностные состояния. Возникновение микроплазм сопровождается свечением отдельных точек *p-n*-перехода, которое связано с высокой температурой электронов и дырок.

6.6. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ПРОБОЯ

При экспериментальных исследованиях электрического пробоя особое внимание обращалось на два обстоятельства. Во-первых, осуществлялся хороший тепловой контакт исследуемых *p-n* или гетеропереходов с внешней средой, чтобы не было нагрева. Во-вторых, принимались меры по устранению поверхностного и локального пробоев. Для этого обычно применялись соответствующие

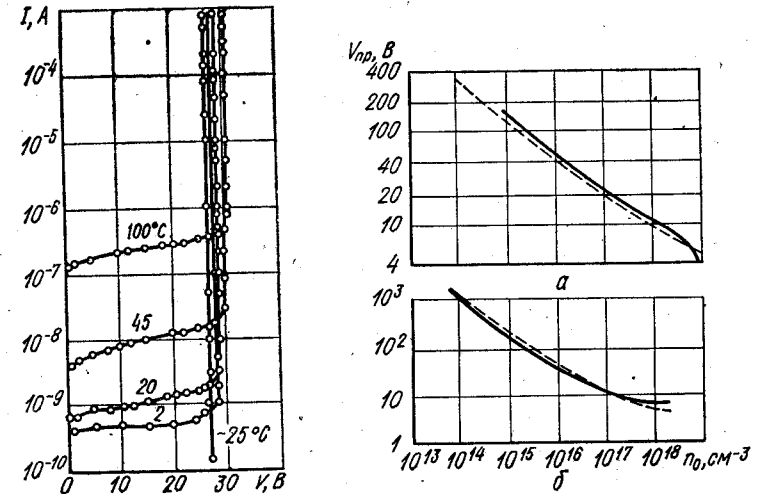


Рис. 6.7. Обратная ветвь вольт-амперной характеристики при различных температурах ($p_0 = 2,5 \cdot 10^{15} \text{ см}^{-3}$).

Рис. 6.8. Зависимость напряжения пробоя *p-n* перехода от концентрации основных носителей в менее легированной области для Ge (а) и Si (б). Пунктир — теоретические кривые.

поверхностные обработки или охранные кольца, а также однородные образцы с малым количеством дефектов.

Исследовались зависимости обратной ветви вольт-амперной характеристики *p-n* переходов от температуры, что, в частности, позволяет определить вид электрического пробоя при различных концентрациях примесей в полупроводнике.

Типичные обратные вольт-амперные характеристики кремниевго $n^+ - p$ перехода при небольших концентрациях легирующей примеси для *p*-области показаны на рис. 6.7. Как видно, из этого рисунка, при определенных напряжениях начинается резкий рост тока с приложенным напряжением, т. е. наблюдается пробой. Величина $V_{пр}$ увеличивается с ростом температуры. Это соответствует лавинному пробую, как и следует из теоретических расчетов для концентрации $p_0 = 2,5 \cdot 10^{15} \text{ см}^{-3}$.

Зависимость экспериментально найденной величины напряжения пробоя *p-n* перехода от концентрации показана на рис. 6.8. Как видно, напряжение пробоя уменьшается с ростом концентрации основных носителей заряда в менее легированной области.

Пунктирные кривые, построенные для лавинного пробоя по формулам п. 6.2, хорошо совпадают с экспериментальными значениями.

При больших концентрациях, чем 10^{17} см^{-3} для Ge и $3 \cdot 10^{17} \text{ см}^{-3}$ для Si, пробой оказывается туннельным. При этом температурный ход зависимости $V_{\text{пр}}$ значительно ослабевает, а $V_{\text{пр}}$ уменьшается с ростом температуры.

Аналогичные результаты получены для контактов металл — полупроводник и гетеропереходов.

В том случае, когда теплоотвод от $p-n$ перехода или гетероперехода был недостаточным, наблюдался тепловой пробой, характеристика которого соответствовала виду рис. 6.1, б. Величина пробойного напряжения зависела от теплового сопротивления и температуры окружающей среды, как это предсказано теорией теплового пробоя.

Особое внимание исследователей было уделено изучению влияния неоднородностей на пробой. В частности, исследовались микроплазмы. Комбинируя методы диффузии и сплавления, удалось создавать искусственные модельные микроплазмы с размерами порядка $5 \cdot 10^{-6} \text{ см}$. Вольт-амперная характеристика таких микроплазм была ступенчатой и подчинялась закономерностям, вытекающим из развитых выше для микроплазм представлений. Экспериментально также обнаружено, что микроплазмы часто наблюдаются у выхода $p-n$ перехода на поверхность, что соответствует теоретическим представлениям о влиянии поверхностных состояний на возникновение микроплазм. При этом на кремнии появление микроплазм у поверхности сопровождалось излучением видимого света.

Таким образом, экспериментальные результаты подтверждают основные модели пробоя, рассмотренные выше.

Глава 7. ЭКВИВАЛЕНТНЫЕ СХЕМЫ ГЕТЕРО- И ГОМОПЕРЕХОДОВ

7.1. ОБЩИЕ ПОЛОЖЕНИЯ

Пусть на гетеро- или гомопереход кроме постоянного напряжения V подается переменный сигнал малой амплитуды ΔV с частотой ω . Тогда суммарное напряжение

$$\tilde{V} = V + \Delta V e^{j\omega t}. \quad (7.1)$$

Это приведет к тому, что через систему пойдет ток, который при достаточно малых величинах ΔV может быть представлен в аналогичном виде

$$\tilde{i} = i + \Delta i e^{j\omega t}. \quad (7.2)$$

Здесь Δi включает амплитуду токов, связанных с переносом свободных носителей заряда $\Delta i_{n,p}$, и максвелловские токи смещения Δi_m , т. е. в общем случае

$$\Delta i = \Delta i_n + \Delta i_p + \Delta i_m, \quad (7.3)$$

где Δi_n , Δi_p и Δi_m нужно относить к одной и той же плоскости сечения гетеро- или гомоперехода.

В результате можно выразить полную проводимость гомо- или гетероперехода так:

$$\frac{1}{Z} = \frac{\Delta i_n}{\Delta V} + \frac{\Delta i_p}{\Delta V} + \frac{\Delta i_m}{\Delta V}. \quad (7.4)$$

Таким образом, задача нахождения полного сопротивления гетеро- и гомопереходов сводится к нахождению величин Δi_n , Δi_p и Δi_m .

Если на рассматриваемой частоте ω справедливы статические вольт-амперные характеристики, полученные ранее, то они могут быть использованы для нахождения Δi_n и Δi_p .

Такая ситуация наблюдается для контакта металл — полупроводник, гетероперехода и $p-n$ перехода между вырожденными полупроводниками, если нет инжекции неосновных носителей заряда и в переносе тока не участвуют локальные уровни вплоть до очень высоких частот.

Частоты, ниже которых справедливы статические вольт-амперные характеристики, определяются инерционностью основных носителей заряда в объеме полупроводников, а также временами их прохождения через области пространственного заряда.

Инерционность основных носителей заряда в объеме полупроводника определяется постоянной времени Максвелла

$$\tau_m = \frac{\epsilon_0 \epsilon}{\sigma},$$

где σ — удельная проводимость полупроводника. При $\epsilon = 10$ и $\sigma = 10^2 \dots 10^{-1}$ Омсм, которые характерны для полупроводников, наиболее широко применяемых для полупроводниковых приборов, $\tau_m = 10^{-14} \dots 10^{-11}$ с.

По порядку величины время пролета через область пространственного заряда

$$\tau_{пр} \approx \frac{v}{\mathcal{L}}.$$

Для ширины области пространственного заряда $\mathcal{L} = 10^{-4} \dots 10^{-6}$ см и $v = 10^7$ см/с время пролета изменяется и составляет $10^{-11} \dots 10^{-13}$ с.

Таким образом, по крайней мере, до частоты 10^{10} Гц⁻¹, соответствующей миллиметровому диапазону длин волн, для контакта металл — полупроводник, p — n перехода и гетероперехода при отсутствии инжекции и локальных уровней, через которые идет ток, для нахождения Δi_n можно использовать статические вольт-амперные характеристики.

Если в контакте металл — полупроводник, p — n переходе или в гетеропереходе токи идут с участием пересадочных центров, например через поверхностные уровни, то статические вольт-амперные характеристики, полученные выше, справедливы только для не очень высоких частот. Это определяется инерционностью перезаполнения пересадочных центров. Времена, которые характеризуют этот процесс, обычно составляют $10^{-8} \dots 10^{-3}$ с.

Если в гомо- или гетеропереходе происходит введение неосновных носителей заряда, то статические вольт-амперные характеристики, найденные в предыдущих главах, становятся несправедливыми на частотах, когда сказывается инерционность неосновных носителей заряда, определяемая временами их рекомбинации. Времена рекомбинации неосновных носителей заряда обычно составляют $10^{-2} \dots 10^{-8}$ с.

В двух последних случаях, когда нужно рассмотреть Δi_n и Δi_p в широкой области частот, уже нельзя использовать статические вольт-амперные характеристики, а необходимо анализировать исходные выражения для токов.

При нахождении Δi_n и Δi_p будем предполагать, что величина приложенного напряжения достаточно мала, так что все величины, являющиеся функцией приложенного напряжения, можно разложить в ряд по малому параметру и ограничиться первым членом разложения.

Например, для контакта металл — полупроводник при учете основных носителей заряда такими величинами являются напряжения V_1 , V_2 , кинетические коэффициенты K_{im} и K_{mi} , concentra-

ция электронов на поверхностных уровнях m_i , граничная напряженность поля E_2 и граничная концентрация электронов n_k . Все эти величины можно записать в следующем виде:

$$\begin{aligned} \tilde{V}_1 &= V_1 + \Delta V_1 e^{j\omega t}; & \tilde{V}_2 &= V_2 + \Delta V_2 e^{j\omega t}; & \tilde{m}_i &= m_i + \Delta m_i e^{j\omega t}; \\ \tilde{E}_2 &= E_2 + \Delta E_2 e^{j\omega t}; & \tilde{K}_{mi} &= K_{mi} + \Delta K_{mi} e^{j\omega t}; & & \\ \tilde{K}_{im} &= K_{im} + \Delta K_{im} e^{j\omega t}; & \tilde{n}_k &= n_k + \Delta n_k e^{j\omega t}. & & \end{aligned} \quad (7.5)$$

Для p — n перехода между невырожденными полупроводниками, кроме того, необходимо учитывать изменение с переменным напряжением концентрации неосновных носителей заряда:

$$\tilde{p} = p + \Delta p e^{j\omega t}; \quad \tilde{n} = n + \Delta n e^{j\omega t}. \quad (7.6)$$

Для гетероперехода аналогичным образом можно представить величины \tilde{V}_1 , \tilde{V}_2 , \tilde{m}_i , \tilde{E}_2 , \tilde{K}_{in}^1 , \tilde{K}_{in}^2 , \tilde{K}_{ni}^1 , \tilde{K}_{ni}^2 , \tilde{K}_{iv}^1 , \tilde{K}_{iv}^2 , \tilde{K}_{iv}^1 , \tilde{K}_{iv}^2 , \tilde{n}_1 , \tilde{n}_2 , \tilde{p}_1 , \tilde{p}_2 и др.

Используя выражения статических вольт-амперных характеристик или исходные выражения для тока с учетом выражений вида (7.5), (7.6), можно найти величины Δi_n и Δi_p . При этом в качестве плоскости сечения, через которую определяется ток, удобно выбрать плоскость, соответствующую границе раздела полупроводника 2 с зазором или полупроводником 1.

Величина Δi_m , связанная с максвелловскими токами смещения, при выбранной плоскости сечения на границе полупроводника 2 может быть найдена в виде

$$\Delta i_m e^{j\omega t} = \epsilon_0 \epsilon_2 \frac{\partial E_2}{\partial t}. \quad (7.7)$$

Ток Δi_m при этом может быть найден в общем виде для контакта металл — полупроводник, для p — n и для гетероперехода. Подставим E_2 из системы (7.5) в (7.4):

$$\Delta i_m = j\omega \epsilon_0 \epsilon_2 \Delta E_2. \quad (7.8)$$

Величина ΔE_2 в случае слоя Шоттки может быть получена из выражения (1.17) через величину ΔV_2

$$\Delta E_2 = \frac{dE_2}{dV_2} \Delta V_2 = - \frac{eE_2}{2(\varphi_0 - eV_2)} \Delta V_2 = \frac{\Delta V_2}{\mathcal{L}}. \quad (7.9)$$

Подставляя это выражение в (7.8), имеем

$$\Delta i_m = j\omega \frac{\epsilon_0 \epsilon_2}{\mathcal{L}} \Delta V_2 = j\omega C_2 \Delta V_2, \quad (7.10)$$

где

$$C_2 = \frac{\epsilon_0 \epsilon_2}{\mathcal{L}} = \left[\frac{\epsilon_0 \epsilon_2 e^2 n_0}{2(\varphi_0 - eV_2)} \right]^{1/2} \quad (7.11)$$

величина, соответствующая емкости плоского конденсатора с расстоянием между пластинами, равным \mathcal{L} . Физическая природа этой емкости определяется тем, что при приложении к области пространственного заряда переменного напряжения происходит изменение

ее ширины и, следовательно, заряда этой области (рис. 7.1). Поэтому такая емкость может быть получена и так:

$$C_2 = \frac{dQ}{dV_2} = \frac{d(en_0\mathcal{L})}{dV_2} = \frac{\epsilon_0\epsilon_2}{\mathcal{L}}. \quad (7.12)$$

Из приведенных выше рассуждений видно, что, хотя полученная емкость и совпадает с емкостью плоского конденсатора, она по своему происхождению отличается от нее, так как в нашем случае происходит изменение ширины области пространственного заряда, эквивалентной расстоянию между пластинами полупроводника.

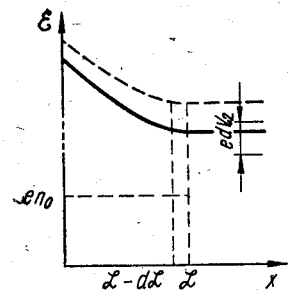


Рис. 7.1. Изменение заряда в области пространственного заряда с приложенным напряжением:

$V = 0$ — сплошная линия;
 $V + dV$ — пунктирная линия.

Емкость, связанная с изменением ширины области пространственного заряда, называется барьерной. Барьерная емкость не зависит от частоты приложенного переменного сигнала вплоть до очень высоких частот, когда сказывается инерционность, связанная с движением основных носителей заряда в объеме полупроводника или в области пространственного заряда. С увеличением прямого напряжения барьерная емкость растет, а обратного — уменьшается

пропорционально $(\phi_0 - eV_2)^{-1/2}$, что определяется изменением с напряжением ширины области пространственного заряда \mathcal{L} . Детализация Δi_m для различных частных случаев сводится к нахождению величины ΔV_2 .

Дальнейшее рассмотрение величин Δi_n , Δi_p и Δi_m будет проведено ниже для различных конкретных случаев. Основное внимание будет уделяться тем случаям, когда величины Δi_n , Δi_p определяются качественно различными физическими явлениями. Исходя из этого, для контакта металл — полупроводник будут рассмотрены токи основных носителей заряда, когда поверхностных уровней нет, поверхностные уровни есть и влияют на перераспределение напряжения между зазором и полупроводником, поверхностные уровни определяют ток. Для $p-n$ перехода будут рассмотрены токи неосновных носителей заряда, а также явления в области пробоя. Для гетероперехода будет проведен общий анализ всех возможных случаев на основе проведенного раньше рассмотрения контакта металл — полупроводник и $p-n$ перехода.

7.2. ЭКВИВАЛЕНТНАЯ СХЕМА КОНТАКТА МЕТАЛЛ — ПОЛУПРОВОДНИК ДЛЯ ОДНОГО ТИПА НОСИТЕЛЕЙ ЗАРЯДА БЕЗ УЧЕТА ПОВЕРХНОСТНЫХ УРОВНЕЙ

Рассмотрим эквивалентную схему контакта металл — полупроводник для случая, когда ток переносится только основными носителями заряда и можно пренебречь туннелированием, силами изображения и т. п. Тогда вольт-амперная характеристика как для ди-

одной, так и для диффузионной теории при пренебрежении токами через поверхностные уровни может быть представлена в следующем общем виде:

$$i = i_s \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right). \quad (7.13)$$

Эта характеристика может быть использована для нахождения величины Δi_n на частотах, значительно меньших частот, при которых возникает сдвиг фаз между током и напряжением за счет инерционности основных носителей заряда.

При этом можно считать, что выражение (7.13) справедливо и для суммы величин \tilde{i} , \tilde{V}_2 , \tilde{V}_1 , \tilde{i}_s . Тогда при пренебрежении зависимостью $i_s(V)$ как более слабой по сравнению с экспоненциальными членами после подстановки \tilde{i} , \tilde{V}_2 и \tilde{V}_1 из (7.5) в (7.13) имеем

$$i + \Delta i_n e^{j\omega t} = i_s \left(e^{\frac{e}{kT}(V_2 + \Delta V_2 e^{j\omega t})} - e^{-\frac{e}{kT}(V_1 + \Delta V_1 e^{j\omega t})} \right).$$

Раскладывая в ряд по малым величинам $\frac{e}{kT} \Delta V_2$ и $\frac{e}{kT} \Delta V_1$ экспоненты и ограничиваясь первым членом разложения, получаем

$$\Delta i_n = i_s \frac{e}{kT} \left(e^{\frac{eV_2}{kT}} \Delta V_2 + e^{-\frac{eV_1}{kT}} \Delta V_1 \right).$$

Подставляя полученные значения Δi_n , а также Δi_m в выражение (7.4), можно записать

$$\frac{1}{Z} = i_s \frac{e}{kT} \left(e^{\frac{eV_2}{kT}} \frac{\Delta V_2}{\Delta V} + e^{-\frac{eV_1}{kT}} \frac{\Delta V_1}{\Delta V} \right) + j\omega C_2 \frac{\Delta V_2}{\Delta V}. \quad (7.14)$$

Для нахождения $\frac{\Delta V_2}{\Delta V}$ и $\frac{\Delta V_1}{\Delta V}$ в случае контакта металл — полупроводник воспользуемся выражением (2.6) при $\sigma_2 = 0$. После подстановки в это выражение \tilde{V}_2 и \tilde{V}_1 и разложения по малым величинам ΔV_2 и ΔV_1 имеем:

$$\frac{\Delta V_1}{\Delta V} = \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right)^{-1} = (1 - \bar{\alpha}_0); \quad (7.15)$$

$$\frac{\Delta V_2}{\Delta V} = \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right)^{-1} = \bar{\alpha}_0, \quad (7.16)$$

где $\bar{\alpha}_0$ связано с α_0^0 , найденным в п. 3.6, следующим соотношением:

$$\bar{\alpha}_0 = \frac{\alpha_0^0}{e/kT}.$$

Теперь можно записать окончательное выражение для $\frac{1}{Z}$ в виде

$$\frac{1}{Z} = i_s \frac{e}{kT} \left[\left(e^{\frac{eV_2}{kT}} + e^{-\frac{eV_1}{kT}} \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right) \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right)^{-1} \right] + j\omega C_2 \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right)^{-1} = \frac{1}{R} + j\omega C, \quad (7.17)$$

где R и C здесь и далее в этой главе относятся к площади 1 см^2 . Как видно из последнего выражения, проводимость собственно контакта металл — полупроводник может быть представлена в виде параллельно включенных сопротивлений и емкости контакта, не зависящих от частоты. Общая эквивалентная схема может быть представлена в виде, изображенном на рис. 7.2, где для общности также учтено последовательное сопротивление объема полупроводника r .

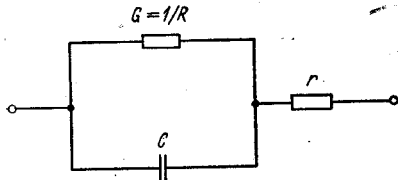


Рис. 7.2. Эквивалентная схема гетеро- и гомоперехода.

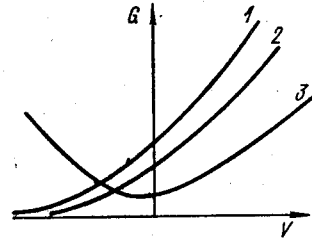


Рис. 7.3. Зависимость проводимости контакта металл — полупроводник от напряжения: 1 — контакт при $d = 0$, диодная теория; 2 — контакт при $d = 0$, диффузионная теория; 3 — контакт при $d \neq 0$.

Проанализируем выражения для $\frac{1}{R}$ и C для различных частных случаев. При уменьшении толщины зазора до нуля $V_1 \rightarrow 0$, $V_2 \rightarrow V$ и полученные выражения для проводимости и емкости переходят в следующие:

$$\frac{1}{R} = i_s \frac{e}{kT} e^{\frac{eV}{kT}}; \quad (7.18)$$

$$C = \frac{\epsilon_0 \epsilon_2}{\mathcal{L}} = \left[\frac{\epsilon_0 \epsilon_2 e^2 n_0}{2(\phi_0 - eV)} \right]^{1/2}. \quad (7.19)$$

В этом случае проводимость контакта изменяется для диодной теории пропорционально $e^{\frac{eV}{kT}}$, а для диффузионной пропорционально $(\phi_0 - eV)^{1/2} e^{\frac{eV}{kT}}$. Емкость контакта является барьерной и растет для прямого направления и убывает для обратного пропорционально $(\phi_0 - eV)^{1/2}$. Все это показано на рис. 7.3 и 7.4.

Если $d \neq 0$, то тогда для анализа зависимости $\frac{1}{R}$ и C от напряжения следует использовать полное выражение (7.17). При этом проводимость в прямом направлении при пренебрежении слабой зависимостью \mathcal{L} от V пропорциональна $e^{\frac{eV_2}{kT}}$, а в обратном пропорциональна $e^{-\frac{eV_1}{kT}}$. Для случая с зазором проводимость для прямого смещения меньше, чем проводимость тесного контакта, так как меньше величина i_s и $e^{\frac{eV_2}{kT}} < e^{\frac{eV}{kT}}$. Для обратного смещения

проводимость тесного контакта может превышать проводимость контакта с зазором из-за экспоненциального члена $e^{\frac{eV_1}{kT}}$ (рис. 7.3).

Следует особо отметить, что активная составляющая проводимости контакта в виде (7.17) и (7.18) может быть получена простым дифференцированием статической вольт-амперной характеристики по напряжению. Это и не удивительно, так как в рассматриваемом интервале частот статическая характеристика не изменяется с частотой.

В случае контакта с зазором можно показать, что суммарная емкость контакта металл — полупроводник может быть преобразована следующим образом:

$$C = C_2 \left(1 + \frac{\epsilon_2 d}{\epsilon_1 \mathcal{L}} \right)^{-1} = C_2 \left(1 + \frac{C_2}{C_1} \right)^{-1} = \frac{C_1 \cdot C_2}{C_1 + C_2}, \quad (7.20)$$

где $C_1 = \frac{\epsilon_0 \epsilon_1}{d}$ — емкость зазора, а $C_2 = \frac{\epsilon_0 \epsilon_2}{\mathcal{L}}$ — барьерная емкость области пространственного заряда.

Выражение (7.20) может быть переписано и в другом виде:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}, \quad (7.21)$$

из которого ясно, что суммарная емкость есть результат последовательного включения емкости зазора и барьерной емкости. Зависимость суммарной емкости C от напряжения для контакта с зазором будет более слабой, чем зависимость барьерной емкости от напряжения для тесного контакта как в силу того, что \mathcal{L} зависит от V_2 , а не от V , так и в силу того, что C_1 не зависит от напряжения. При больших обратных напряжениях, когда $C_1 \gg C_2$, $C \simeq C_2$, а при больших прямых, когда $C_2 \gg C_1$, $C \simeq C_1$. Кроме того, поскольку рассматривается последовательное включение емкости зазора и области пространственного заряда, суммарная емкость контакта с зазором при прочих одинаковых параметрах меньше, чем емкость контакта без зазора (рис. 7.4).

Учет туннелирования через область пространственного заряда и сил изображения приведет к изменению величины активного сопротивления контакта, но не изменит его емкости.

7.3. ВЛИЯНИЕ ПОВЕРХНОСТНЫХ УРОВНЕЙ НА ЭКВИВАЛЕНТНУЮ СХЕМУ КОНТАКТА

Пусть в контакте металл — полупроводник, ток в котором переносится основными носителями заряда, существуют поверхностные уровни. Предположим, что эти уровни не участвуют в переносе тока.

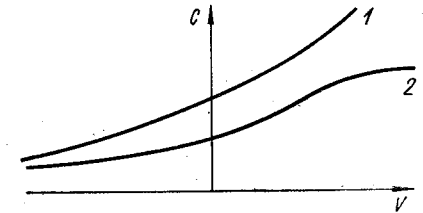


Рис. 7.4. Зависимость емкости контакта металл — полупроводник от напряжения: 1 — контакт без зазора; 2 — контакт с зазором.

Тогда расчет величины Δi_n и последующее нахождение $\frac{1}{Z}$ может проводиться с использованием статической вольт-амперной характеристики вида (7.13), как это проделано в предыдущем параграфе. Но теперь следует заново найти величины ΔV_1 и ΔV_2 , которые подставляются в (7.14), с учетом влияния на эти величины поверхностных уровней. Возьмем для определенности один локальный уровень, расположенный на границе раздела зазор — полупроводник.

Подставив в выражение (2.41), которое справедливо и для суммы постоянной и переменной составляющих, \tilde{V}_1 и \tilde{V}_2 , легко получить:

$$\Delta V_1 = -\frac{\varepsilon_2 d}{\varepsilon_1} \Delta E_2 + \frac{ed}{\varepsilon_0 \varepsilon_1} \Delta m_i; \quad (7.22)$$

$$\Delta V_2 = \Delta V - \frac{\varepsilon_2 d}{\varepsilon_1} \Delta E_2 - \frac{ed}{\varepsilon_0 \varepsilon_1} \Delta m_i.$$

Используя (7.9), выражения (7.22) можно переписать в виде:

$$\begin{aligned} \Delta V_1 &= (1 - \tilde{\alpha}_0) \Delta V + \frac{ed}{\varepsilon_0 \varepsilon_1} \Delta m_i; \\ \Delta V_2 &= \tilde{\alpha}_0 \left(\Delta V - \frac{ed}{\varepsilon_0 \varepsilon_1} \Delta m_i \right). \end{aligned} \quad (7.23)$$

Таким образом, задача нахождения $\frac{1}{Z}$ сводится к нахождению величины Δm_i . Для нахождения неизвестного значения Δm_i запишем уравнение баланса носителей заряда на уровне

$$\frac{dm_i}{dt} = K_{ni} (n_i - m_i) n_k - K_{in} m_i + K_{mi} (n_i - m_i) - K_{im} m_i, \quad (7.24)$$

где учтены обмены с зоной проводимости и металлом. Предположим, что это выражение справедливо и для величин \tilde{m}_i , \tilde{V}_2 , \tilde{V}_1 и т. д. Подставляя в это выражение \tilde{m}_i , \tilde{n}_k , \tilde{K}_{im} и \tilde{K}_{mi} , можно получить

$$\Delta m_i = \frac{K_{ni} (n_i - m_i) \Delta n_k - \Delta K_{mi} (n_i - m_i) - \Delta K_{im} m_i}{i\omega + K_{ni} n_k + K_{in} + K_{mi} + K_{im}}. \quad (7.25)$$

Для нахождения Δn_k подставим в $n_k = n_k^0 e^{\frac{eV_2}{kT}}$ справедливое для диодной теории выражение \tilde{n}_k и \tilde{V}_2 и проведем разложение по малым величинам:

$$\Delta n_k = n_k^0 e^{\frac{eV_2}{kT}} \frac{e \Delta V_2}{kT}. \quad (7.26)$$

Аналогично можно найти ΔK_{mi} и ΔK_{im} , если воспользоваться выражениями (3.72) и (3.74):

$$\Delta K_{mi} = \frac{dK_{mi}}{dV_1} \Delta V_1 = -\frac{e}{kT} \frac{K_{im} K_{mi}}{K_{im} + K_{mi}} \Delta V_1; \quad (7.27)$$

$$\Delta K_{im} = \frac{dK_{im}}{dV_1} \Delta V_1 = \frac{e}{kT} \frac{K_{im} K_{mi}}{K_{im} + K_{mi}} \Delta V_1. \quad (7.28)$$

Подставляя значения Δn_k , ΔK_{mi} и ΔK_{im} в (7.25), имеем

$$\Delta m_i = \frac{a_1 \Delta V}{i\omega + b_1} = \frac{a_1 b_1 \Delta V}{\omega^2 + b_1^2} - \frac{j\omega a_1 \Delta V}{\omega^2 + b_1^2}, \quad (7.29)$$

где

$$a_1 = -\frac{e}{kT} \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} \right)^{-1} \left[\frac{K_{im} K_{mi} n_i}{K_{im} + K_{mi}} \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} - K_{ni} (n_i - m_i) n_k \right]; \quad (7.30)$$

$$\begin{aligned} b_1 &= K_{mi} + K_{im} + K_{ni} n_k + K_{in} + \left\{ \frac{K_{im} K_{mi} n_i}{K_{im} + K_{mi}} + K_{ni} (n_i - m_i) n_k \right\} \times \\ &\times \frac{ed}{\varepsilon_0 \varepsilon_1} \frac{e}{kT} \left(1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} \right)^{-1} \end{aligned} \quad (7.31)$$

Как видно из выражения (7.29), Δm_i представляется в виде суммы действительной и мнимой составляющих. С физической точки зрения это определяется тем, что величина Δm_i может иметь фазовый сдвиг по отношению к приложенному напряжению. С изменением частоты Δm_i изменяется от максимальной величины

$$\Delta m_i = \frac{a_1}{b_1} \Delta V$$

при $\omega = 0$, когда поверхностные состояния полностью следят за приложенным напряжением, до минимальной

$$\Delta m_i = 0$$

при $\omega \rightarrow \infty$, когда поверхностные состояния уже не успевают изменяться при изменении напряжения.

Анализ полученных выражений для a_1 и b_1 показывает, что величина b_1 всегда положительна и имеет смысл $\frac{1}{\tau_i}$, где τ_i — время перезаплатнения поверхностного уровня, а величина a_1 может быть как положительной, так и отрицательной. Например, при преимущественном обмене поверхностного уровня с металлом $a_1 < 0$, а при преимущественном обмене с полупроводником $a_1 > 0$. В первом случае при положительных ΔV $\Delta m_i < 0$ и уровень освобождается, а во втором $\Delta m_i > 0$ и уровень заполняется.

Подставляя ΔE_2 и Δm_i в выражение (7.22) и (7.23), можно получить:

$$\Delta V_1 = \left[(1 - \tilde{\alpha}_0) + \frac{a_2 b_1}{\omega^2 + b_1^2} - \frac{j\omega a_2}{\omega^2 + b_1^2} \right] \Delta V; \quad (7.32)$$

$$\Delta V_2 = \left(\tilde{\alpha}_0 - \frac{a_2 b_1}{\omega^2 + b_1^2} + \frac{j\omega a_2}{\omega^2 + b_1^2} \right) \Delta V, \quad (7.33)$$

где $a_2 = \frac{e d a_1 \alpha_0}{\varepsilon_0 \varepsilon_1}$.

Как видно из этих выражений, величины ΔV_1 и ΔV_2 также могут иметь сдвиг фаз по отношению к приложенному напряжению

и изменяются с частотой. Влияние поверхностных уровней приводит к увеличению величины ΔV_1 и уменьшению величины ΔV_2 при преимущественном их обмене с полупроводником и уменьшению ΔV_1 и увеличению ΔV_2 при преимущественном их обмене с металлом. Это влияние максимально при $\omega = 0$, а затем ослабевает с ростом частоты, а когда $\omega \rightarrow \infty$, добавки к ΔV_1 и ΔV_2 , определяемые поверхностными уровнями, стремятся к нулю.

Подставляя выражения (7.32) и (7.33) в (7.14) и (7.10), имеем

$$\frac{1}{Z} = \frac{1}{R} + j\omega C,$$

где

$$\frac{1}{R} = i_s \frac{e}{kT} \left\{ \left(\bar{\alpha}_0 - \frac{a_2 b_1}{\omega^2 + b_1^2} \right) e^{\frac{eV_2}{kT}} + \left[(1 - \bar{\alpha}_0) + \frac{a_2 b_1}{\omega^2 + b_1^2} \right] e^{-\frac{eV_1}{kT}} \right\} - \frac{\omega^2 a_2}{\omega^2 + b_1^2} C_2; \quad (7.34)$$

$$C = i_s \frac{e}{kT} \frac{a_2}{\omega^2 + b_1^2} \left(e^{\frac{eV_2}{kT}} - e^{-\frac{eV_1}{kT}} \right) + C_2 \left(\bar{\alpha}_0 - \frac{a_2 b_1}{\omega^2 + b_1^2} \right). \quad (7.35)$$

Как видно, и в рассматриваемом случае эквивалентную схему контакта также можно представить в виде параллельно включенных активной и реактивной составляющих, как это показано на рис. 7.2. Однако здесь в общем случае как активная, так и реактивная составляющие зависят от частоты.

Отметим также, что в величину активной составляющей входит член, пропорциональный C_2 , а в выражение для реактивной составляющей — член, пропорциональный проходящему над барьером области пространственного заряда току i_n . Это связано с комплексным характером величин ΔV_1 и ΔV_2 .

Активную составляющую проводимости в общем виде теперь нельзя получить простым дифференцированием статической вольт-амперной характеристики, несмотря на то, что она остается справедливой во всем интервале рассматриваемых частот, поскольку теперь величины V_1 и V_2 зависят от частоты.

Легко показать, что при $d = 0$, когда $a_2 = 0$, $V_1 = 0$, $V_2 = V$, $\bar{\alpha}_0 = 1$, выражения для активной и реактивной составляющих сводятся к ранее полученным выражениям (7.18) и (7.19) для контакта без зазора. При этом уровни не проявляются в эквивалентной схеме. Физическая причина такого явления заключается в том, что уровни влияют на полное сопротивление только через изменение соотношения между ΔV_1 и ΔV_2 . Если же $d = 0$, то всегда $\Delta V_1 = 0$, а $\Delta V_2 = \Delta V$.

В том случае, когда $d \neq 0$, поверхностные уровни изменяют как активную, так и реактивную составляющую проводимости собственно контакта.

При изменении частоты от $\omega = 0$ до $\omega = \infty$ активная составляющая контакта изменяется от величины

$$G = i_s \frac{e}{kT} \left\{ \left(\bar{\alpha}_0 - \frac{a_2}{b_1} \right) e^{\frac{eV_2}{kT}} + \left[1 - \bar{\alpha}_0 + \frac{a_2}{b_1} \right] e^{-\frac{eV_1}{kT}} \right\}$$

до величины

$$G = i_s \frac{e}{kT} \left\{ \left(\bar{\alpha}_0 e^{\frac{eV_2}{kT}} + (1 - \bar{\alpha}_0) e^{-\frac{eV_1}{kT}} \right) - a_2 C_2 \right\}.$$

Можно показать, что активная составляющая проводимости при $\omega = 0$ совпадает с активной составляющей, полученной при дифференцировании статической вольт-амперной характеристики с учетом зависимости ΔV_1 и ΔV_2 от Δm_i . Активная составляющая при $\omega \rightarrow \infty$ не может быть получена в результате такого дифференцирования. Это связано с тем, что при $\omega \rightarrow \infty$ в активную составляющую вносит вклад ток Δi_m из-за двойного сдвига фаз, получаемого в результате перемножения мнимых составляющих в токе смещения Δi_m . При этом вклад поверхностных уровней в проводимость не исчезает на высоких частотах.

Можно показать, что в зависимости от преимущественного механизма обмена уровней с полупроводником или металлом проводимость, связанная с надбарьерными токами, с частотой может убывать или расти. Так, при преимущественном обмене уровней с металлом $a_1 < 0$ и эта проводимость с частотой для прямого смещения уменьшается, а для обратного растет. Если же главную роль играет обмен с полупроводником $a_1 > 0$ и проводимость на высоких частотах при прямых напряжениях больше, а при обратных меньше, чем на высоких частотах.

Реактивная составляющая, связанная с надбарьерными токами, убывает с частотой. На низких частотах знак этой составляющей зависит от механизма обмена и направления приложенного напряжения. Так, при обмене с полупроводником ($a_1 > 0$) эта составляющая положительна и несет емкостной характер для прямого смещения и отрицательна и носит индуктивный характер для отрицательных смещений. Если же главную роль играет обмен с металлом ($a_1 < 0$), то реактивная составляющая, связанная с уровнями, отрицательна для прямого направления и положительна для обратного.

7.4. ЭКВИВАЛЕНТНАЯ СХЕМА КОНТАКТА ДЛЯ ТОКОВ ЧЕРЕЗ ПОВЕРХНОСТНЫЕ УРОВНИ

Рассмотрим теперь эквивалентную схему контакта для случая, когда главную роль играют токи через поверхностные уровни, которые обмениваются электронами с зоной проводимости полупроводника и металлом. В этом случае для нахождения Δi_n нельзя использовать статическую вольт-амперную характеристику, полученную

в результате подстановки в (3.67) и (3.68) статических значений m_i .

Однако для расчета Δi_n можно использовать сами исходные выражения вида (3.67) и (3.68), которые справедливы при любых частотах, когда не проявляется инерционность основных носителей заряда. Учитывая то, что выбранная нами плоскость, через которую проходит ток Δi_n , расположена на границе раздела полупроводник — зазор, для нахождения Δi_n нужно использовать выражение для тока из зоны проводимости на поверхностный уровень. В результате

$$\Delta i_n = e [K_{ni}(n_i - m_i) \Delta n_k - K_{ni} \Delta m_i - K_{im} \Delta m_i], \quad (7.36)$$

где стационарное значение m_i задается выражением (3.69).

Теперь можно воспользоваться найденными в п. 7.3 значениями Δn_k и Δm_i и получить выражение для Z , состоящее из суммы активной и реактивной составляющих:

$$\frac{1}{Z} = \frac{1}{R} + j\omega C = G + j\omega C,$$

где

$$\frac{1}{R} = \frac{ea_1}{\omega^2 + b_1^2} \left[\omega^2 \left(\bar{\alpha}_0 - \frac{a_m}{a_1} \right) + b_1^2 \left(\frac{b_m}{b_1} - \frac{a_m}{a_1} \right) \right]; \quad (7.37)$$

$$C = \left[e \left(\bar{\alpha}_0 - \frac{b_m}{b_1} \right) \frac{a_1 b_1}{\omega^2 + b_1^2} + C_2 \bar{\alpha}_0 \right]. \quad (7.38)$$

Здесь

$$a_m = - \frac{e}{kT} \bar{\alpha}_0 \frac{K_{im} K_{mi} n_i}{K_{im} + K_{mi}} \cdot \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}}; \quad (7.39)$$

$$b_m = K_{mi} + K_{im} + \frac{K_{im} K_{mi}}{K_{im} + K_{mi}} \left(\frac{1}{\bar{\alpha}_0} - \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}} \right) \frac{ed}{\varepsilon_0 \varepsilon_1} \cdot \frac{e}{kT} \bar{\alpha}_0. \quad (7.40)$$

Как видно, и в этом случае эквивалентная схема контакта может быть представлена в обычном виде (см. рис. 7.2), где активная и реактивные составляющие в общем случае зависят от частоты переменного сигнала.

Рассматривая (7.37), можно показать, что G и C всегда положительны. В самом деле, при преимущественном обмене с металлом

$$\frac{a_m}{a_1} \simeq 1; \quad \frac{b_m}{b_1} = 1; \quad G \simeq \frac{ea_1 \omega^2}{\omega^2 + b_1^2} (\bar{\alpha}_0 - 1);$$

при $a_1 < 0$ и $\alpha_0 - 1 > 0$ $G > 0$.

При преимущественном обмене с полупроводником

$$\frac{a_m}{a_1} = 0, \quad \frac{b_m}{b_1} = 0;$$

$$G \simeq \frac{ea_1 \omega^2}{\omega^2 + b_1^2} \bar{\alpha}_0$$

при $a_1 > 0$ $G > 0$.

Величина C также всегда положительна, при преимущественном обмене с металлом

$$C = e (\bar{\alpha}_0 - 1) \frac{a_1 b_1}{\omega^2 + b_1^2} + C_2 \bar{\alpha}_0 > 0,$$

при преимущественном обмене с полупроводником

$$C = \frac{ea_1 b_1}{\omega^2 + b_1^2} \bar{\alpha}_0 + C_2 \bar{\alpha}_0 > 0.$$

С изменением частоты проводимость и емкость изменяются между двумя крайними значениями. Максимальное изменение с частотой проводимости определяется отношением

$$\frac{G|_{\omega=0}}{G|_{\omega=\infty}} = \frac{\frac{b_m}{b_1} - \frac{a_m}{a_1}}{\bar{\alpha}_0 - \frac{a_m}{a_1}}.$$

Рассматривая последнее выражение, можно видеть, что изменения сопротивления с частотой могут быть значительными. Аналогичное отношение для емкости равно при преимущественном обмене поверхностных уровней с металлом

$$\frac{C|_{\omega=0}}{C|_{\omega=\infty}} \leq 1 + \frac{\varepsilon_2 d}{\varepsilon_1 \mathcal{L}};$$

при преимущественном обмене с полупроводником

$$\frac{C|_{\omega=0}}{C|_{\omega=\infty}} \leq 1 + \frac{\varepsilon_1 \mathcal{L}}{\varepsilon_2 d}.$$

Из последних двух выражений видно, что изменение емкости с частотой для тонких зазоров будет больше для случая обмена с полупроводником по сравнению со случаем обмена с металлом.

7.5. ЭКВИВАЛЕНТНАЯ СХЕМА $p-n$ ПЕРЕХОДА

Рассмотрим эквивалентную схему $p-n$ перехода для малых уровней инжекции при пренебрежении генерацией и рекомбинацией в области пространственного заряда. Для простоты ограничимся случаем широкой базы.

Для $p-n$ перехода уже нельзя найти Δi_n или Δi_p из готового выражения статической вольт-амперной характеристики, а необходимо решать задачу с самого начала для случая суммы постоянных и переменных составляющих концентраций электронов и дырок. В случае, когда в квазинейтральной области ток является диффузионным, для нахождения Δi_p необходимо решать уравнения

$$i_p = -e D_p \frac{dp}{dx}; \quad (7.41)$$

$$\frac{\partial p}{\partial t} = - \frac{p - p_0}{\tau_p} - \frac{1}{e} \cdot \frac{\partial i_p}{\partial x} = - \frac{p - p_0}{\tau_p} - D_p \frac{d^2 p_p}{dx^2}. \quad (7.42)$$

Подставляя в выражение (7.41) \tilde{p} и \tilde{i}_p и группируя члены, зависящие и не зависящие от времени, можно получить следующих два уравнения:

$$i_p = -eD_p \frac{dp}{dx}; \quad \Delta i_p = -eD_p \frac{d\Delta p}{dx}. \quad (7.43)$$

Первое из этих уравнений совпадает с использованным ранее при расчете вольт-амперной характеристики и позволяет находить постоянный ток через $p-n$ переход, а второе дает связь между амплитудными значениями Δi_p и Δp . Подставляя в выражение (7.42) \tilde{p} и группируя члены, зависящие и не зависящие от времени, можно также получить два выражения:

$$\frac{d^2 p}{dx^2} = \frac{p - p_0}{l_p^2}; \quad \frac{d^2 \Delta p}{dx^2} = \frac{\Delta p}{D_p \tau_p} + \frac{j\omega \Delta p}{D_p}. \quad (7.44)$$

Первое из этих выражений уже использовалось. Оно относится к стационарным значениям концентрации p . Второе выражение относится к амплитудным значениям Δp . Его легко переписать в виде, аналогичном первому выражению:

$$\frac{d^2 \Delta p}{dx^2} = \frac{\Delta p}{(l_p^*)^2}, \quad (7.45)$$

если обозначить

$$l_p^* = \frac{l_p}{(1 + j\omega\tau_p)^{1/2}}. \quad (7.46)$$

Решение этого уравнения по форме совпадает с решением уравнения для стационарного случая и может быть записано в виде

$$\Delta p = Ae^{-\frac{x}{l_p^*}} + Be^{\frac{x}{l_p^*}}. \quad (7.47)$$

Из граничного условия $\Delta p = 0$ при $x = \infty$ $B = 0$. Из граничного условия $\Delta p = \Delta p(\mathcal{L}_n)$ при $x = \mathcal{L}_n$ можно найти

$$A = \Delta p(\mathcal{L}_n) e^{\frac{\mathcal{L}_n}{l_p^*}}. \quad (7.48)$$

Для нахождения $\Delta p(\mathcal{L}_n)$ можно воспользоваться граничным условием $p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}}$, пригодным также для \tilde{p} и \tilde{V} . В результате после подстановки в это условие p и \tilde{V} и разделения членов, зависящих и не зависящих от времени, можно получить

$$\Delta p(\mathcal{L}_n) = p_n e^{\frac{eV}{kT}} \frac{e\Delta V}{kT}. \quad (7.49)$$

Окончательно

$$\Delta p = p_n \frac{e\Delta V}{kT} e^{\frac{eV}{kT}} e^{\frac{\mathcal{L}_n - x}{l_p^*}}. \quad (7.50)$$

Подставляя это выражение в формулу (7.43), имеем

$$\Delta i_p = \frac{e^2 D_p p_n}{l_p^*} \cdot \frac{\Delta V}{kT} e^{\frac{eV}{kT}} e^{\frac{\mathcal{L}_n - x}{l_p^*}}. \quad (7.51)$$

Если, как и для постоянного тока, находить $\Delta i_p(\mathcal{L}_n)$, то

$$\Delta i_p(\mathcal{L}_n) = \frac{e^2 D_p p_n}{l_p^*} \cdot \frac{\Delta V}{kT} e^{\frac{eV}{kT}}. \quad (7.52)$$

Аналогичным образом можно получить величину

$$\Delta i_n(-\mathcal{L}_p) = \frac{e^2 D_n n_p}{l_n^*} \cdot \frac{\Delta V}{kT} e^{\frac{eV}{kT}}, \quad (7.53)$$

где

$$l_n^* = \frac{l_n}{(1 + j\omega\tau_n)^{1/2}}.$$

Теперь нужно найти величину Δi_m для $p-n$ перехода. Для этого подставим в выражение (7.10) величину ΔV_2 , найденную с помощью формулы (2.48). В результате

$$\begin{aligned} \Delta i_m &= \frac{\varepsilon_0 \varepsilon_2}{\mathcal{L}_n} \cdot \frac{p_p}{p_p + n_n} \Delta V = \\ &= \frac{\varepsilon_0 \varepsilon_2}{\left[\frac{2\varepsilon_2 \varepsilon_0 \varphi_0 p_p}{e^2 (n_n + p_p) n_n} - \frac{2\varepsilon_2 \varepsilon_0 eV p_p}{e^2 (n_n + p_p) n_n} \right]^{1/2}} \frac{p_p \Delta V}{n_n + p_p} = \frac{\varepsilon_0 \varepsilon_2}{\mathcal{L}} \Delta V, \end{aligned}$$

где

$$\mathcal{L} = \left[\frac{2\varepsilon_0 \varepsilon_2 (\varphi_0 - eV) (p_p + n_n)}{e^2 p_p n_n} \right]^{1/2}$$

общая ширина области пространственного заряда $p-n$ перехода. Подставляя Δi_n , Δi_p , Δi_m в выражение (7.4), получаем для $p-n$ перехода

$$\frac{1}{Z} = \frac{e^2}{kT} e^{\frac{eV}{kT}} \left[\frac{D_n n_p}{l_n} (1 + j\omega\tau_n)^{1/2} + \frac{D_p p_n}{l_p} (1 + j\omega\tau_p)^{1/2} \right] + j\omega \frac{\varepsilon_0 \varepsilon_2}{\mathcal{L}}. \quad (7.54)$$

Как показывает анализ этого выражения, полная проводимость $p-n$ перехода состоит из активной и реактивной составляющей.

Действительно, если использовать формулу Муавра, то

$$\sqrt{1 + j\omega\tau_p} = \frac{\sqrt{1 + \sqrt{1 + (\omega\tau_p)^2}}}{\sqrt{2}} + j \frac{\omega\tau_p}{\sqrt{2}\sqrt{1 + (\omega\tau_p)^2}}, \quad (7.55)$$

$$\sqrt{1 + j\omega\tau_n} = \frac{\sqrt{1 + \sqrt{1 + (\omega\tau_n)^2}}}{\sqrt{2}} + j \frac{\omega\tau_n}{\sqrt{2}\sqrt{1 + (\omega\tau_n)^2}}.$$

Тогда

$$\frac{1}{Z} = G + j\omega C,$$

где

$$G = G_{p0} \frac{\sqrt{1 + \sqrt{1 + (\omega\tau_p)^2}}}{\sqrt{2}} + G_{n0} \frac{\sqrt{1 + \sqrt{1 + (\omega\tau_n)^2}}}{\sqrt{2}}; \quad (7.56)$$

$$C = G_{p0} \frac{\tau_p}{\sqrt{2} \sqrt{1 + \sqrt{1 + (\omega\tau_p)^2}}} + G_{n0} \frac{\tau_n}{\sqrt{2} \sqrt{1 + \sqrt{1 + (\omega\tau_n)^2}}} + \frac{\varepsilon_1 \varepsilon_2}{\mathcal{L}} = C_d + C_b. \quad (7.57)$$

Здесь

$$G_{p0} = \frac{eD_p p_n}{l_p} \cdot \frac{e}{kT} e^{\frac{eV}{kT}};$$

$$G_{n0} = \frac{eD_n n_p}{l_n} \cdot \frac{e}{kT} e^{\frac{eV}{kT}},$$

где C_d — диффузионная емкость, связанная с инжекцией неосновных носителей заряда; C_b — барьерная емкость, связанная с максвелловскими токами смещения (эквивалентную схему см. на рис. 7.2).

Как видно из выражения (7.56), проводимость p — n перехода зависит от частоты. При $\omega\tau_p, \omega\tau_n \ll 1$ проводимость постоянна, а при $\omega\tau_p, \omega\tau_n \gg 1$ она пропорциональна $\sqrt{\omega}$. Диффузионная емкость также зависит от частоты. При $\omega\tau_p, \omega\tau_n \ll 1$ она постоянная, а при обратном неравенстве она обратно пропорциональна $\sqrt{\omega}$.

Природа такого изменения проводимости и диффузионной емкости с частотой заключается в том, что они связаны с движением неосновных носителей заряда в квазинейтральной области. Этот процесс является инерционным и приводит к частотной зависимости проводимости и диффузионной емкости.

Проводимость p — n перехода при $\omega \rightarrow 0$, равная $G = G_{p0} + G_{n0}$, может быть получена также дифференцированием статической вольт-амперной характеристики. Это определяется тем, что при достаточно низких частотах инжектированные неосновные носители заряда успевают «следить» за приложенным напряжением.

Диффузионная емкость при $\omega = 0$, когда инжектированные носители «следят» за приложенным напряжением, может быть найдена как производная $C_d = \frac{dQ_{p,n}}{dV}$, где $Q_{p,n}$ — накопленный в квазинейтральной области заряд неосновных носителей. Например, для заряда дырок

$$Q_p = e \int_{\mathcal{L}_n}^{\infty} p_n e^{\frac{eV}{kT}} e^{-\frac{\mathcal{L}_n - x}{l_p}} dx = -e l_p p_n e^{\frac{eV}{kT}}.$$

Используя это значение Q_p , можно получить

$$C_d = \frac{dQ_p}{dV} = \frac{eD_p p_n}{l_p} \cdot \frac{e}{kT} \cdot \frac{\tau_p}{2},$$

которая равна диффузионной емкости при $\omega\tau_p \ll 1$, связанной с дырками.

С изменением напряжения проводимость и диффузионная емкость экспоненциально растут с напряжением для прямого направления и уменьшаются для обратного. Барьерная емкость изменяется с напряжением как $(\varphi_0 - eV)^{1/2}$ и не зависит от частоты.

В связи с этим в зависимости от напряжения и частоты для емкости p — n перехода могут реализоваться разные случаи. В частности, суммарная емкость на низких частотах при прямых смещениях, как правило, является диффузионной, а при обратных всегда преобладает барьерная емкость. При больших частотах барьерная емкость всегда больше, чем диффузионная.

Аналогичное выражение для эквивалентной схемы может быть получено и для p — n перехода с узкой базой. В этом случае будут отличаться величины G_{p0} и G_{n0} , а также изменена зависимость активной и реактивной составляющей от частоты переменного сигнала при сохранении ее общего характера.

Если в p — n переходе между вырожденными полупроводниками нет инжекций носителей заряда, то величина Δi_n может быть найдена с использованием вольт-амперной характеристики, проанализированной в п. 4.7. При этом для нахождения Δi_n можно использовать, например, выражение (4.87). В результате легко получить значение активного сопротивления такого p — n перехода, совпадающее со значением, полученным при дифференцировании вольт-амперной характеристики (см. выражение (4.88)). Величина Δi_n для p — n перехода между вырожденными полупроводниками действительна и не дает вклада в реактивную составляющую сопротивления. Поэтому емкость туннельного p — n перехода является чисто барьерной и совпадает с полученной выше (см. выражение (7.57)).

7.6. ЭКВИВАЛЕНТНАЯ СХЕМА ГЕТЕРОПЕРЕХОДА

Гетеропереход является более сложной системой, чем контакт металл — полупроводник и p — n гетеропереход. В гетеропереходе могут играть роль зазор, поверхностные состояния, участвующие в переносе тока без изменения и с изменением вида разрешенных зон, а также инжекция носителей заряда в полупроводники 1 и 2. Поэтому нахождение в аналитическом виде выражений для эквивалентной схемы гетероперехода значительно усложняется. Однако в результате можно получить новые аналитические выражения, но не новые физические результаты.

В самом деле, если проводить в случае изотипного гетероперехода расчет для надбарьерных токов при пренебрежении поверхностными уровнями можно для нахождения Δi_n воспользоваться выражением статических вольт-амперных характеристик, а для нахождения Δi_m — выражением (7.10).

Поскольку при этом вольт-амперная характеристика в общем случае может быть записана в виде $i_n = i(V_1, V_3, V_2)$, то в общем

виде

$$\Delta i_n = \frac{di}{dV_1} \Delta V_1 + \frac{di}{dV_3} \Delta V_3 + \frac{di}{dV_2} \Delta V_2.$$

Таким образом, нахождение Δi_n сводится к нахождению величин ΔV_1 , ΔV_3 и ΔV_2 . При отсутствии поверхностных состояний эти величины являются действительными.

Нахождение Δi_m сводится к подстановке действительной величины ΔV_2 в выражение (7.10). В результате получается мнимое значение Δi_m .

Все это позволяет получить эквивалентную схему гетероперехода, изображенную на рис. 7.2, где R и C — сопротивление и емкость собственно гетероперехода, а r сумма последовательных сопротивлений объемов полупроводников 1 и 2. При этом R и C зависят от параметров зазора, области пространственного заряда и параметров объема полупроводников 1 и 2.

В частности, емкость изотипного гетероперехода вида антизапорный — запорный слой, когда $\frac{dV_1}{dV} \ll \frac{dV_3}{dV}$ и $\frac{dV_1}{dV} \ll \frac{dV_2}{dV}$ совпадает с емкостью контакта металл — полупроводник, а для гетероперехода вида запорный — запорный слой можно, используя результаты п. 5.7, получить

$$C = \frac{C_3}{\left[1 + \frac{C_3}{C_2} + \frac{C_3}{C_1}\right]},$$

где

$$C_1 = \frac{\epsilon_0 \epsilon_1}{\mathcal{L}_1}; \quad C_3 = \frac{\epsilon_0 \epsilon_3}{d}; \quad C_2 = \frac{\epsilon_0 \epsilon_2}{\mathcal{L}_2}.$$

Эта емкость может быть представлена так:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_3} + \frac{1}{C_2},$$

что обуславливается последовательным включением областей пространственного заряда полупроводников 1 и 2 и области зазора. Особенности поведения суммарной емкости гетероперехода C с приложенным напряжением определяются тем, что C_1 и C_2 ведут себя по-разному с приложенным напряжением. Так, для изотипного гетероперехода вида запорный — запорный слой при приложении прямого напряжения емкость C_2 увеличивается, а емкость C_1 уменьшается. Это может приводить к тому, что суммарная емкость с изменением напряжения может иметь максимум.

Если в гетеропереходе есть поверхностные уровни, которые не участвуют в переносе тока, а только изменяют соотношение между V_1 , V_3 и V_2 , то для нахождения Δi_n в случае изотипного гетероперехода можно воспользоваться теми же вольт-амперными характеристиками, что и выше. Учет поверхностных уровней при этом приводит к другим значениям ΔV_1 , ΔV_3 и ΔV_2 , которые теперь являются комплексными величинами. Все это приводит к изменению величин

R и C , их зависимости от частоты, как это рассмотрено выше для контакта металл — полупроводник.

В том случае, когда в анизотипном гетеропереходе происходит инжекция неосновных носителей заряда, нахождение его эквивалентной схемы следует производить таким же образом, как и для $p-n$ перехода. При этом следует учитывать, что для гетероперехода в силу различия высот потенциальных барьеров для электронов и дырок нужно учитывать только одну составляющую тока — либо дырочную, либо электронную.

В результате полученная эквивалентная схема имеет вид, представленный на рис. 7.2, и те же основные свойства, что и эквивалентная схема $p-n$ перехода.

Сопротивление гетероперехода и его емкость при этом зависят от частоты. Эта зависимость определяется временем жизни неосновных носителей заряда в полупроводниках 1 или 2.

Емкость в этом случае представляет собой сумму барьерной и диффузионной емкостей, которые по-разному зависят от частоты и приложенного напряжения.

Когда ток в гетеропереходе идет с участием поверхностных уровней, в качестве пересадочных центров, то для нахождения величин Δi_n и Δi_p уже нельзя использовать окончательные статические вольтамперные характеристики, а нужно использовать исходные выражения вида (5.29), (5.30) или (5.51), (5.52), которые справедливы для любых частот, ограниченных инерционностью основных носителей заряда.

В результате также можно получить эквивалентную схему вида рис. 7.2. Элементы эквивалентной схемы R и C при этом зависят от частоты, что определяется инерционностью заполнения поверхностных состояний, если не происходит инжекции неосновных носителей заряда, и определяется инерционностью заполнения поверхностных уровней и временем жизни неосновных носителей заряда, если такая инжекция есть.

7.7. ЭКВИВАЛЕНТНАЯ СХЕМА $p-n$ ПЕРЕХОДА НА УЧАСТКЕ ЛАВИННОГО ПРОБОЯ

Эквивалентную схему $p-n$ перехода на участке лавинного пробоя будем рассматривать при следующих предположениях:

1) частота электрического сигнала много меньше частоты соударений носителей заряда с решеткой, так что можно не учитывать реальную траекторию носителя заряда, а рассматривать их усредненное движение в сильном электрическом поле;

2) умножение квазистационарно, т. е. вероятность ионизации является однозначной функцией напряженности поля и не зависит явно от координаты и времени;

3) напряженность поля изменяется достаточно плавно, чтобы коэффициенты ионизации носителей заряда можно считать однозначными функциями величины напряженности поля.

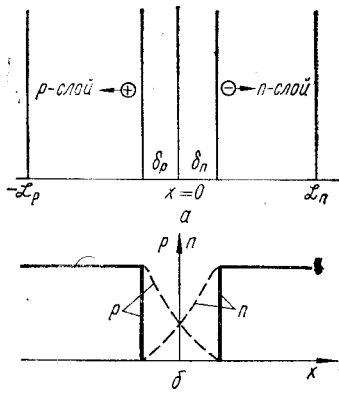


Рис. 7.5. Модель $p-n$ перехода в режиме лавинного умножения.

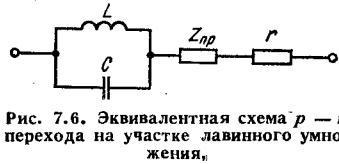


Рис. 7.6. Эквивалентная схема $p-n$ перехода на участке лавинного умножения.

Будем считать, что для $p-n$ перехода в режиме лавинного умножения можно выделить тонкий эквивалентный слой умножения, в котором сосредоточено все умножение, и две области дрейфа носителей заряда, как это показано на рис. 7.5, а. В эквивалентном слое умножения суммарная плотность заряда подвижных носителей равна нулю. Из этого слоя в области дрейфа выходит полный ток умножения, который в p -полупроводнике целиком переносится дырками, а в n -полупроводнике электронами. Ширина эквивалентного слоя умножения, в котором истинное распределение электронов и дырок может быть заменено прямоугольным, как это показано на рис. 7.5, б, из физических соображений равна эффективной длине пробега между ионизирующими столкновениями $\delta = \delta_n + \delta_p$.

Полное сопротивление $p-n$ перехода в режиме лавинного умножения

$$Z = \frac{\Delta V_n}{\Delta i_n + \Delta i_{cm}} + \frac{\Delta V_{np}}{\Delta i_{np} + \Delta i_{cm}}, \quad (7.58)$$

где ΔV_n и ΔV_{np} , а также Δi_n , Δi_{np} и Δi_{cm} — части амплитуд напряжения и плотности тока проводимости и смещения на слое умножения и пролетных участках.

При выполнении принятых предположений решение поставленной задачи в общем случае сводится к решению в различных областях $p-n$ перехода следующей системы уравнений:

$$\frac{\partial p}{\partial t} = \frac{1}{e} \cdot \frac{\partial i_p}{\partial x} + (\alpha_n i_n + \alpha_p i_p); \quad (7.59)$$

$$\frac{\partial n}{\partial t} = -\frac{1}{e} \cdot \frac{\partial i_n}{\partial x} + (\alpha_n i_n + \alpha_p i_p); \quad (7.60)$$

$$\frac{dE}{dx} = \frac{1}{\epsilon_0 \epsilon} \left(\rho(x) + \frac{i_p(x, t)}{v_p} - \frac{i_n(x, t)}{v_n} \right); \quad (7.61)$$

$$i(t) = i_n(x, t) + i_p(x, t) + \epsilon_0 \epsilon \frac{\partial E(x, t)}{\partial t}, \quad (7.62)$$

где

$$i_n = env_n, \quad i_p = epv_p.$$

Уравнения (7.59), (7.60) являются уравнениями неразрывности с учетом ударной ионизации носителей заряда, уравнение (7.61) — уравнение Пуассона, а уравнение (7.62) описывает суммарный ток через $p-n$ переход.

Поскольку в общем случае коэффициенты ударной ионизации для электронов и дырок α_n и α_p зависят от величины напряженности электрического поля E , система уравнений (7.59) — (7.62) в частных производных нелинейна и не может

быть решена аналитически даже для малых амплитуд переменного сигнала. Поэтому для аналитического решения задачи нужны дальнейшие упрощения.

Дополнительно будем предполагать, что

$$v_n = v_p = v_{др} = \text{const}; \quad \alpha_n = \alpha_p = \alpha_{n,p}.$$

Это означает, что скорости электронов и дырок, а также коэффициенты ударной ионизации равны.

Теперь сложим уравнения (7.59) и (7.60) и проинтегрируем обе части полученного выражения в пределах слоя умножения. В результате получим

$$\begin{aligned} \frac{1}{v_{др}} \cdot \frac{\partial}{\partial t} \int_{-\delta_p}^{\delta_n} [i_n(x, t) + i_p(x, t)] dx + (i_p - i_n) \int_{-\delta_p}^{\delta_n} &= \\ = 2 \int_{-\delta_p}^{\delta_n} \alpha_{n,p}(x, t) [i_n(x, t) + i_p(x, t)] dx. \end{aligned} \quad (7.63)$$

Если полагать, что суммарная плотность тока в слое умножения в каждый момент времени одинакова, равна $i_n(t) = i_n(x, t) + i_p(x, t)$ и не зависит от координаты, то последнее равенство запишется в виде

$$\tau_6 \frac{di_n}{dt} = 2i_n \int_{-\delta_p}^{\delta_n} \alpha_{n,p} dx - (i_p - i_n) \int_{-\delta_p}^{\delta_n} & \quad (7.64)$$

где $\tau_6 = \frac{\delta}{v_{др}}$.

За пределами слоя умножения в сечении $x = -\delta_n$ ток дырок состоит целиком из тока i_{sp} , генерированного вне слоя умножения, так что

$$(i_p - i_n) |_{-\delta_n} = (2i_p - i_n) |_{-\delta_n} = 2i_{sp} - i_n.$$

Ток электронов в сечении $x = -\delta_p$ состоит из обратного тока насыщения i_{sn} . Поэтому

$$(i_p - i_n) |_{-\delta_p} = (i_n - 2i_n) |_{-\delta_p} = (i_n - 2i_{sn}) |_{-\delta_p}.$$

Используя записанные граничные условия, получим

$$2i_n \int_{-\delta_n}^{\delta_p} \alpha_{n,p} dx - (i_p - i_n) \int_{-\delta_n}^{\delta_p} = 2i_n \int_{-\delta_n}^{\delta_p} \alpha_{n,p} dx - 2i_n + 2i_s, \quad (7.65)$$

где $i_s = i_{sn} + i_{sp}$.

Подставив (7.65) в (7.64), приходим к уравнению

$$\frac{di_n}{dt} = i_n \frac{2}{\tau_6} \left[\int_{-\delta_p}^{\delta_n} \alpha_{n,p} dx - 1 \right] + \frac{2i_s}{\tau_6}. \quad (7.66)$$

Упростим уравнение (7.66), подставив вместо $\alpha_{n,p}$ среднюю величину $\bar{\alpha}$, полученную путем интегрирования $\alpha_{n,p}$ по всему слою умножения.

Тогда выражение (7.66) при пренебрежении малым членом $\frac{2i_s}{\tau_6}$ переписется в виде

$$\frac{di_n}{dt} = \frac{2i_n}{\tau_6} [\bar{\alpha}\delta - 1]. \quad (7.67)$$

В предположениях малости сигнала:

$$\begin{aligned}\bar{\alpha} &= \bar{\alpha}_0 + \Delta\bar{\alpha}e^{j\omega t} = \bar{\alpha}_0 + \frac{\partial\bar{\alpha}}{\partial E} \Delta E e^{j\omega t}; \\ \bar{\alpha}\delta &= \left(\bar{\alpha}_0\delta + \frac{\partial\bar{\alpha}}{\partial E} \delta\Delta E e^{j\omega t}\right) = 1 + \delta \frac{\partial\bar{\alpha}}{\partial E} \Delta E e^{j\omega t}; \\ i_{л} &= i_{л0} + \Delta i_{л} e^{j\omega t}; \\ E &= E_0 + \Delta E e^{j\omega t}.\end{aligned}$$

Подставляя записанные выше выражения в (7.67) и пренебрегая членами высшего порядка малости, можно получить

$$\Delta i_{л} = \frac{2 \frac{\partial\bar{\alpha}}{\partial E} i_{л0} \Delta E_{л} \delta}{j\omega\tau_6}. \quad (7.68)$$

Ток смещения в слое умножения

$$\Delta i_{см} = j\omega\epsilon_0\epsilon\Delta E_{л}. \quad (7.69)$$

Учитывая, что для слоя умножения при принятых предположениях $\Delta E_{л} = \frac{\Delta V_{л}}{\delta}$, можно записать импеданс слоя умножения

$$Z_{л} = \frac{\delta}{j\omega\epsilon_0\epsilon} \left[1 - \frac{\omega_r^2}{\omega^2}\right]^{-1}, \quad (7.70)$$

где

$$\omega_r = \sqrt{\frac{2 \frac{\partial\bar{\alpha}}{\partial E} v_{др} i_{лс}}{\epsilon_0\epsilon}}.$$

Этот слой может быть представлен, как параллельная LC-цепочка. Действительно, величина $\frac{1}{Z} = \frac{\Delta i_{л}}{\Delta V_{л}} + \frac{\Delta i_{см}}{\Delta V_{л}}$ с использованием (7.68) и (7.69) может быть представлена в виде

$$\frac{1}{Z} = j\omega C + \frac{1}{j\omega L},$$

где индуктивность и емкость цепочки соответственно равны:

$$L = \frac{\tau_6}{2i_{л0} \frac{\partial\alpha}{\partial E}}; \quad (7.71)$$

$$C = \frac{\epsilon_0\epsilon}{\mathcal{L}}. \quad (7.72)$$

Таким образом, тонкий слой умножения ведет себя как параллельная резонансная цепочка с собственной резонансной частотой, равной ω_r . При частотах $\omega < \omega_r$ сопротивление слоя умножения носит емкостной, а при $\omega > \omega_r$ — индуктивный характер.

Рассмотрим эквивалентную схему пролетных участков. В предположении о насыщении дрейфовой скорости носителей заряда переменная составляющая плотности тока проводимости образует незатухающую волну, в которой имеются только фазовые изменения и которая распространяется с этой дрейфовой скоростью.

Таким образом, для n -области

$$\Delta i_{пр} = \Delta i_{л} e^{-\frac{j\omega x}{v_{др}}} = \gamma_{л} \Delta i_{л} e^{-\frac{j\omega x}{v_{др}}}, \quad (7.73)$$

где $\gamma_{л} = \frac{\Delta i_{л}}{\Delta i}$ — комплексное число, равное отношению плотности тока лавины к плотности полного тока.

Напомним, что при принятых предположениях в любом сечении пролетного пространства плотность полного тока равна сумме плотности тока проводимости и плотности тока смещения и не зависит от координаты:

$$\Delta i = \Delta i_{пр} + \Delta i_{см} \neq f(x).$$

Плотность тока смещения может быть определена для произвольной точки x пролетного слоя так:

$$\Delta i_{см} = j\omega\epsilon_0\epsilon\Delta E(x).$$

Это позволяет получить выражение для переменной составляющей электрического поля в пространстве дрейфа n -области как функции от x в виде

$$\Delta E(x) = \Delta i \frac{1 - \gamma_{л} e^{-\frac{j\omega x}{v_{др}}}}{j\omega\epsilon_0\epsilon}. \quad (7.74)$$

Интегрирование уравнения (7.74) по длине дрейфа $\mathcal{L}_n - \delta_n$ дает выражение для переменной составляющей напряжения на пролетном участке n -области

$$\Delta V(\mathcal{L}_n - \delta_n) = \frac{(\mathcal{L}_n - \delta_n) \Delta i}{j\omega\epsilon_0\epsilon} \left[1 - \gamma_{л} \left(\frac{1 - e^{-j\theta_{др}^n}}{j\theta_{др}^n}\right)\right], \quad (7.75)$$

где $\theta_{др}^n = \frac{\omega(\mathcal{L}_n - \delta_n)}{v_{др}}$ — пролетный угол в пространстве дрейфа n -области.

Теперь легко получить выражение для импеданса пролетного участка n -области в виде

$$\begin{aligned}Z_{др}^n &= \frac{\Delta V_{др}}{\Delta i} = \frac{1}{\omega C_{др}^n} \left[\frac{1}{\gamma} \left(\frac{1 - \cos \theta_{др}^n}{\theta_{др}^n}\right) + \right. \\ &\quad \left. + \frac{j}{\omega C_{др}^n} \left(-1 + \gamma_{л} \frac{\sin \theta_{др}^n}{\theta_{др}^n}\right) \right], \quad (7.76)\end{aligned}$$

где $C_{др}^n = \frac{\epsilon_0\epsilon}{(\mathcal{L}_n - \delta_n)}$ — емкость пролетного участка n -области.

Теперь воспользуемся выражениями (7.68) и (7.69) и найдем $\gamma_{л} = \frac{\Delta i_{л}}{\Delta i_{л} + \Delta i_{см}}$. Легко показать, что

$$\gamma_{л} = \left[1 - \frac{\omega^2}{\omega_r^2}\right]^{-1}. \quad (7.77)$$

В результате полное сопротивление пролетного слоя n -области перепишется в виде

$$Z_{др}^n = \frac{1}{\omega C_{др}^n} \left[\left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \left(\frac{1 - \cos \theta_{др}^n}{\theta_{др}^n}\right) + \frac{j}{\omega C_{др}^n} \left[-1 + \left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \frac{\sin \theta_{др}^n}{\theta_{др}^n} \right] \right]. \quad (7.78)$$

Эквивалентная схема пролетного слоя может быть представлена в виде параллельно включенных активной и реактивной составляющих, зависящих от частоты, причем активное сопротивление становится отрицательным для всех частот, лежащих выше ω_r , кроме точек, где оно равно нулю.

Аналогичным образом можно найти полное сопротивление пролетного слоя p -области

$$Z_{др}^p = \frac{1}{\omega C_{др}^p} \left[\left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \left(\frac{1 - \cos \theta_{др}^p}{\theta_{др}^p}\right) + \frac{j}{\omega C_{др}^p} \left[-1 + \left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \frac{\sin \theta_{др}^p}{\theta_{др}^p} \right] \right], \quad (7.79)$$

где

$$C_{др}^p = \frac{\varepsilon_0 \varepsilon}{-(\mathcal{L}_p - \delta_p)}; \quad \theta_{др}^p = -\omega \frac{(\mathcal{L}_p - \delta_p)}{v_{др}}.$$

Полная эквивалентная схема p — n перехода на участке лавинного пробоя может быть представлена в виде суммы импедансов слоя умножения, областей дрейфа, а также сопротивления квазинейтральных областей и показана на рис. 7.6. Величина полного сопротивления

$$Z = \frac{\delta}{j\omega \varepsilon_0 \varepsilon} \left[1 - \frac{\omega_r^2}{\omega^2}\right]^{-1} + \frac{1}{\omega} \sum_{i=n,p} \left\{ \frac{1}{C_{др}^i} \left[\left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \times \left(\frac{1 - \cos \theta_{др}^i}{\theta_{др}^i} \right) + \frac{j}{C_{др}^i} \left[-1 + \left(1 - \frac{\omega^2}{\omega_r^2}\right)^{-1} \frac{\sin \theta_{др}^i}{\theta_{др}^i} \right] \right] \right\} + r. \quad (7.80)$$

Как показывает анализ полученных результатов, главными особенностями полного сопротивления p — n перехода на участке лавинного пробоя являются зависимость реактивной и активной составляющей от частоты и существование отрицательного динамического активного сопротивления. Последнее используется в лавинно-пролетных диодах для генерации и усиления колебаний.

Расчетные зависимости полного сопротивления p — n перехода от частоты и тока имеют ряд характерных черт. В первую очередь к ним относятся резкое изменение полного сопротивления вблизи характеристических значений частоты ω , и зависимость ω , от плотности постоянного тока i_n , когда реактивное X и активное R сопротивления p — n перехода изменяют знак. На частотах, ниже некоторого значения ω_r , активное сопротивление p — n перехода положительно при всех токах. Причем зависимость этого сопротивления от тока имеет вид резонансных кривых, которые смещаются при увеличении частоты в сторону больших токов. На больших частотах сопротив-

ление p — n перехода становится отрицательным с очень малых токов. При увеличении тока модуль этого сопротивления резко возрастает, затем уменьшается и уходит в область положительных значений. Реактивное сопротивление при токах меньших определенного значения является емкостным, а при больших — индуктивным. В зависимости от величины тока этот переход наблюдается при разных значениях частоты.

В заключение следует подчеркнуть, что сложный характер зависимостей параметров рассмотренной здесь эквивалентной схемы p — n перехода на участке лавинного пробоя определяется, помимо прочих факторов, конечным временем движения носителей через область пространственного заряда. Этим рассмотренный здесь случай принципиально отличается от всех ранее рассмотренных.

7.8. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ГОМО- И ГЕТЕРОПЕРЕХОДОВ ПРИ МАЛОМ ПЕРЕМЕННОМ НАПРЯЖЕНИИ

Для исследования эквивалентной схемы гомо- и гетеропереходов применяются методы, основанные на сравнении измеряемого полного сопротивления с известным, либо на измерении амплитуд ΔV и Δi и сдвига фаз между ними. При этом необходимо, чтобы выполнялось условие малого сигнала $\Delta V \ll \frac{kT}{e}$.

В результате проведенных исследований установлено, что эквивалентная схема контакта металл — полупроводник, p — n перехода и гетероперехода в общем случае может быть представлена в виде, изображенном на рис. 7.2.

Подробные исследования параметров R и C эквивалентной схемы контакта показали, что они зависят от удельного сопротивления и типа проводимости полупроводника, обработки поверхности полупроводника, материала металла и метода его нанесения, т. е. зависят от тех же факторов, что и параметры вольт-амперной характеристики. Для установления физической модели контактов проводилось исследование зависимости R и C от приложенного напряжения и частоты.

На рис. 7.7 приведены типичные зависимости параметров R и C контакта металл — полупроводник с различной толщиной промежуточного слоя, определенной из величины C_1 при больших прямых напряжениях, от напряжения. Как видно из рисунка, сопротивление диода с $d = 25 \text{ \AA}$ в одинаковом интервале напряжений слабее зависит от напряжения в прямом направлении и сильнее в обратном, чем для диода с $d \leq 10 \text{ \AA}$. Емкость диода с $d = 10 \text{ \AA}$ меньше емкости диода с $d \leq 10 \text{ \AA}$ и при достаточно больших напряжениях практически не зависит от напряжения. Описанные экспериментально закономерности качественно соответствуют теоретическим, рассмотренным в предыдущих параграфах.

Проведенные исследования зависимости R и C от частоты при фиксированном значении постоянного напряжения показывают, что

для контактов с $d \leq 10 \text{ \AA}$ в диапазоне частот $10^3 \dots 10^7 \text{ Гц}$ R и C практически не зависят от частоты. Для контактов с $d \geq 25 \text{ \AA}$, как правило, R и C зависят от частоты. Примеры таких зависимостей показаны на рис. 7.8. Как видно, с повышением частоты как сопротивление R , так и емкость диода C уменьшаются. Анализ этих зависимостей показывает, что в этом частном случае они качественно согласуются с моделью контакта с учетом зазора и поверхностных состояний при предположении об участии поверхностных состояний

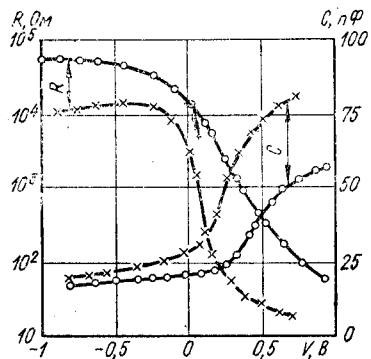


Рис. 7.7. Зависимость R и C от напряжения для контакта $\text{Pd}-p\text{-Si}$ ($\rho = 0,4 \text{ Ом}\cdot\text{см}$): $\times\times\times$ — $d \leq 10 \text{ \AA}$; $\circ\circ\circ$ — $d = 25 \text{ \AA}$.

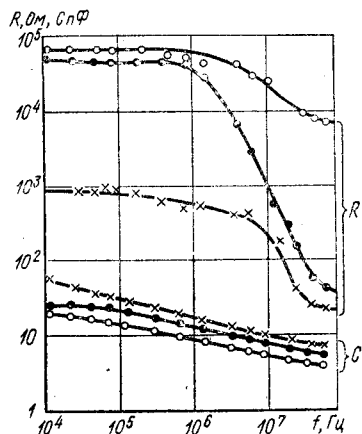


Рис. 7.8. Зависимость R и C от частоты переменного сигнала для контакта $\text{Au}-\text{Si}$ ($\rho = 0,7 \text{ Ом}\cdot\text{см}$): \dots — $V = 0$; $\times\times\times$ — $V = 0,5 \text{ В}$; $\circ\circ\circ$ — $V = -0,5 \text{ В}$.

в переносе тока и предположении, что главную роль играет обмен этих состояний с полупроводником.

Таким образом, полученные при исследовании элементов эквивалентной схемы экспериментальные результаты подтверждают общую модель, в которой учитываются зазор и поверхностные состояния. Однако для контакта металл — полупроводник в настоящее время еще не получено полного количественного согласия между разработанными теоретическими моделями и экспериментом.

Подробно исследовалась также эквивалентная схема $p-n$ переходов. Зависимость активной и реактивной составляющих германиевых $p-n$ переходов, для которых можно пренебречь генерационно-рекомбинационными токами от частоты и напряжения, находится в хорошем согласии с теорией. Это иллюстрируется рис. 7.9, где представлены экспериментальные точки и теоретические кривые проводимости, а также рис. 7.10, где построены зависимости C от R при разных I . Последние зависимости должны укладываться на прямую, поскольку

$$C = C_d + C_6 = C_6 + \left(\frac{e}{kT}\right)^2 (I + I_s)^2 \frac{\tau_p R}{2}$$

при любых частотах (см. (7.56) и (7.57)), что и наблюдается экспериментально.

Из наклона прямых может быть найдено время жизни, определяющее частотную зависимость R и C , а из отсечки — величина барьерной емкости. Сравнение значений времени жизни, полученных из частотных зависимостей R и C , а также из других измерений, обнаруживает неплохое совпадение.

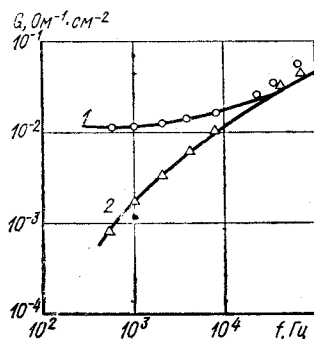


Рис. 7.9. Зависимость проводимости германиевого $p-n$ перехода от частоты: 1 — активная составляющая; 2 — реактивная составляющая (сплошные кривые — теория).

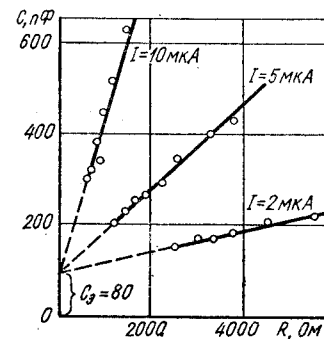


Рис. 7.10. Зависимость C от R для $p-n$ перехода при разных токах положительного смещения.

Таким образом, разработанная теория для $p-n$ перехода хорошо описывает экспериментальные результаты, если можно пренебречь генерационно-рекомбинационными токами.

Эквивалентные схемы различных гетеропереходов широко исследовались с целью установления их физической модели. При этом установлено, что в общем случае R и C зависят от напряжения и частоты. Много работ было посвящено исследованию вольт-емкостных характеристик, позволяющих с более простых позиций интерпретировать физическую модель исследуемого гетероперехода.

Это можно иллюстрировать на примере резких изотипных гетеропереходов германий — кремний.

Типичная зависимость емкости $n-n$ гетероперехода $\text{Ge}-\text{Si}$ от напряжения при разных частотах представлена на рис. 7.11. Как видно из этого рисунка, емкость при фиксированной частоте убывает с ростом положительных и отрицательных напряжений. С ростом частоты при постоянных значениях напряженная емкость уменьшается. Наблюдаемая зависимость емкости от напряжения может быть объяснена, если предположить, что в исследованных гетеропереходах существуют два, включенных навстречу друг другу, барьера Шоттки, как это показано на рис. 5.1, в, г. Зависимость емкости от частоты можно объяснить влиянием поверхностных состояний. Это все качественно согласуется с данными о рассогласовании постоянных решеток Ge и Si , приводящих к существованию на границе раздела полупроводников большой концентрации по-

верхностных состояний. Однако количественного описания экспериментальных результатов достигнуто не было.

И вообще, несмотря на большое количество экспериментальных исследований гетеропереходов, лишь в очень небольшом количестве случаев удалось дать полуколичественное объяснение наблюдаемым зависимостям с точки зрения рассмотренных выше физических моделей.

Для сопоставления малосигнальной теории $p-n$ перехода в режиме лавинного пробоя исследовалось полное сопротивление $p-n$ переходов в широком диапазоне сверхвысоких частот. Полученные

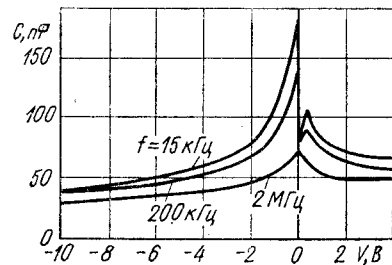


Рис. 7.11. Зависимость емкости изотипного $p-n$ гетероперехода Ge-Si от приложенного напряжения на различных частотах

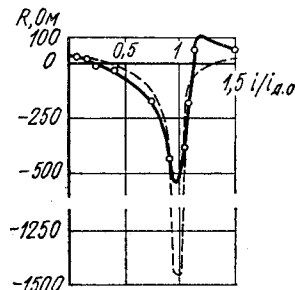


Рис. 7.12. Сравнение экспериментальных (—) зависимостей $R(I)$ с расчетными (---).

экспериментальные результаты подтвердили качественные выводы рассмотренной теории. Так, на частотах выше некоторого значения активное сопротивление $p-n$ перехода становилось отрицательным. Для таких частот при увеличении тока модуль активного сопротивления возрастает, а затем уменьшается, проходит через нуль и уходит в область положительных значений. Все это показано на рис. 7.12. На этом же рисунке даны расчетные кривые. Как видно, наблюдается хорошее не только качественное, но и количественное совпадение экспериментальных и теоретических результатов (некоторые различия могут объясняться неоднородным пробоем).

Аналогичные результаты, показывающие справедливость описанной выше малосигнальной теории работы $p-n$ перехода на участке лавинного пробоя, были получены также при изучении реактивной составляющей полного сопротивления $p-n$ перехода.

Глава 8. ОМИЧЕСКИЕ КОНТАКТЫ

8.1. СВОЙСТВА И МОДЕЛЬ ОМИЧЕСКИХ КОНТАКТОВ

Омический контакт металл — полупроводник служит для подведения электрического тока или напряжения к полупроводникам и полупроводниковым приборам. Основное требование к омическим контактам заключается в том, что они не должны вносить заметного дополнительного сопротивления в рабочем интервале напряжений и токов. Последнее означает, что сопротивление омических контактов должно быть достаточно малым, а его вольт-амперная характеристика должна быть линейной, т. е. для таких контактов должен выполняться закон Ома.

Кроме того, к омическим контактам иногда предъявляются и другие требования. Например, как правило, омические контакты должны быть неинжектирующими, иметь хорошую теплопроводность, т. е. к ним предъявляются различные конкретные требования.

Создать идеальный омический контакт, полностью отвечающий различным требованиям, практически невозможно. Поэтому реальные омические контакты представляют собой большее или меньшее приближение к идеальному.

Если ограничиться рассмотрением основных электрических свойств омического контакта, прежде всего необходимо рассмотреть условия в области пространственного заряда, при которых эта область имеет малое и близкое к линейному сопротивление. При этом можно выделить три случая, когда свойства контакта металл — полупроводник приближаются к свойствам омического контакта. Для n -типа полупроводника при учете тока основных носителей заряда, когда инжекция незначительна, они показаны на рис. 8.1. В случае, изображенном на рис. 8.1, а, в контакте изгиб зон антизапорный. Для случая рис. 8.1, б изгиб зон запорный, но ширина области пространственного заряда очень мала. На рис. 8.1, в представлен запорный слой с достаточно малой высотой барьера. Основные носители заряда в рассмотренных случаях либо идут в разрешенной зоне, либо туннелируют через область пространственного заряда, как показано стрелками.

При рассмотрении прохождения носителей заряда через промежуточный слой между металлом и полупроводником, как и в гл. 3, можно рассматривать токи с участием и без участия поверхностных уровней.

Используя результаты, полученные в предыдущих главах, можно найти сопротивление омического контакта при напряжениях $V_1 + V_2 \ll V_b$, где V_1, V_2 и V_b — падение напряжения в промежу-

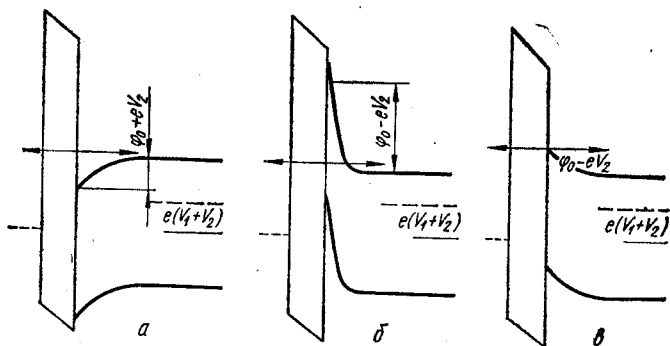


Рис. 8.1. Ход потенциала в области пространственного заряда полупроводника для омических контактов.

точном слое области пространственного заряда и объеме полупроводника. При этом можно ограничиться приближениями диодной теории, так как при этом (см. гл. 3) токи больше, а сопротивление меньше, чем в случае выполнения предположений диффузионной теории, что и требуется для омических контактов.

8.2. ОМИЧЕСКИЕ КОНТАКТЫ ДЛЯ ТОКОВ БЕЗ УЧАСТИЯ ПОВЕРХНОСТНЫХ УРОВНЕЙ

Антизапорный слой. Для антизапорного слоя в случае диодной теории, по аналогии с результатами п. 3.1, можно записать вольт-амперную характеристику для тока электронов в виде

$$i = \frac{en_0\bar{v}_n\mathcal{P}_n}{4} \left(1 - e^{-\frac{e(V_1+V_2)}{kT}}\right), \quad (8.1)$$

где учтено, что ток электронов из полупроводника в металл идет без преодоления потенциального барьера, а для обратного тока такой барьер существует и изменяется с приложенным напряжением.

Для нахождения сопротивления омического контакта предположим, что $\frac{e(V_1+V_2)}{kT} \ll 1$. Тогда, раскладывая экспоненту в выражении (8.1) в ряд и ограничиваясь первым членом, имеем

$$i = \frac{en_0\bar{v}_n\mathcal{P}_n}{4} \frac{e}{kT} (V_1 + V_2). \quad (8.2)$$

Сопротивление омического контакта

$$R = \frac{V_1 + V_2}{i}, \text{ т. е. } R = \frac{4kT}{e^2 n_0 \bar{v}_n \mathcal{P}_n}. \quad (8.3)$$

Как видно из последнего выражения, величина R убывает с ростом концентрации носителей заряда n_0 и коэффициента прозрачности \mathcal{P}_n . Оценки показывают, что при комнатной температуре для $\mathcal{P}_n = 1$ и $n_0 = 10^{15} \text{ см}^{-3}$ $R = 5 \cdot 10^{-5} \text{ Ом} \cdot \text{см}^{-2}$, а для $n_0 = 10^{18} \text{ см}^{-3}$ $R = 5 \cdot 10^{-8} \text{ Ом} \cdot \text{см}^{-2}$. Эти значения сопротивления омического контакта удовлетворяют большинство реальных случаев.

Туннелирование через область пространственного заряда. При учете сильного туннелирования для нахождения сопротивления омического контакта можно воспользоваться выражением (3.52). Раскладывая экспоненту этого выражения в ряд при $\frac{e(V_1+V_2)}{kT} \ll 1$, можно получить

$$i = \frac{en_0\bar{v}_n\mathcal{P}_n}{4} \frac{e}{kT} e^{-\frac{\phi_0}{E_{00}}} (V_1 + V_2), \quad (8.4)$$

где \mathcal{P}_n учитывает также туннелирование через область пространственного заряда.

Отсюда можно получить сопротивление омического контакта

$$R = \frac{4kT}{e^2 n_0 \bar{v}_n \mathcal{P}_n} e^{\frac{\phi_0}{E_{00}}}. \quad (8.5)$$

Сопротивление омического контакта в случае туннелирования также зависит от концентрации n_0 и коэффициента прозрачности \mathcal{P}_n . Кроме того, сопротивление зависит от высоты потенциального барьера ϕ_0 и параметра E_{00} . Как показывает сравнение этого сопротивления с сопротивлением для антизапорного слоя, оно имеет большую величину, чем сопротивление для антизапорного слоя, поскольку

для рассмотренного случая $\mathcal{P}_n < 1$, а $e^{\frac{\phi_0}{E_{00}}} > 1$.

Слабый запорный изгиб зон. В случае диодной теории для тока электронов при этом можно воспользоваться выражением вольт-амперной характеристики вида (3.4). При $\frac{eV_2}{kT} \ll 1$ и $\frac{eV_1}{kT} \ll 1$ оно переписывается в виде

$$i = \frac{\bar{v}_n n_0 \mathcal{P}_n}{4} e^{-\frac{\phi_0}{kT}} \frac{e}{kT} (V_1 + V_2), \quad (8.6)$$

откуда

$$R = \frac{4kT}{e^2 \bar{v}_n n_0 \mathcal{P}_n} e^{\frac{\phi_0}{kT}}. \quad (8.7)$$

Полученное сопротивление, как и в предыдущих случаях, зависит от величин n_0 и \mathcal{P}_n и уменьшается с их ростом. Кроме того, как и в случае сильного туннелирования, сопротивление омического контакта уменьшается с уменьшением величины высоты барьера. Оценки показывают, что для комнатной температуры при $\mathcal{P}_n = 1$ $\phi_0 = 2,3 kT$ и $n_0 = 10^{15} \text{ см}^{-3}$, $R = 5 \cdot 10^{-4} \text{ Ом} \cdot \text{см}^{-2}$, а при тех же параметрах и $n_0 = 10^{18} \text{ см}^{-3}$, $R = 5 \cdot 10^{-7} \text{ Ом} \cdot \text{см}^{-2}$. Это все еще достаточно малые значения сопротивления для многих реальных случаев. Кроме того, при расчете сопротивления запорного слоя следует учитывать, что для тесного контакта высота потенциального барьера дополнительно уменьшается за счет сил изображения, как это показано в гл. 3.

Антизапорный слой. Для антизапорного слоя, используя формулу (3.79), при $\frac{n_k}{n_0} = 1$, можно получить

$$i_n^s = \frac{\bar{e}v_n \bar{c}_n n_0 M_i n_i (e^{\frac{e(V_1+V_2)}{kT}} - 1)}{(1 + M_i e^{\frac{eV_1}{kT}}) [1 + K(1 + M_i)]} \quad (8.8)$$

При $\frac{e(V_1 + V_2)}{kT} \ll 1$ после разложения экспонент последнего уравнения в ряд получаем

$$i_n^s = \frac{\bar{e}v_n \bar{c}_n n_0 M_i n_i \bar{c}_n}{(1 + M_i) [1 + K(1 + M_i)]} \frac{e}{kT} (V_1 + V_2). \quad (8.9)$$

Отсюда сопротивление омического контакта

$$R = \frac{(1 + M_i) [1 + K(1 + M_i)] kT}{e^2 \bar{v}_n \bar{c}_n n_0 M_i n_i \bar{c}_n}. \quad (8.10)$$

Как видно, сопротивление омического контакта для антизапорного слоя в этом случае зависит не только от концентрации носителей заряда в полупроводнике n_0 , но и от параметров M_i , K , c_n и n_i , определяющих свойства поверхностных состояний. Сопротивление уменьшается с уменьшением K , т. е. с переходом к преимущественному обмену уровней с металлом, с ростом величин n_i и c_n . Зависимость R от величины M_i разная для случаев, когда уровни преимущественно обмениваются с полупроводником или металлом. Так, при преимущественном обмене с полупроводником, когда $K \rightarrow \infty$, величина R вначале убывает, а затем увеличивается при росте M_i . В этом случае имеется экстремальное значение $M_i = 1$, соответствующее минимальному сопротивлению. При преимущественном обмене уровней с металлом, когда $K \rightarrow 0$, величина R вначале убывает с ростом M_i , а затем перестает зависеть от M_i .

Оценки величины R для комнатной температуры при обмене с металлом для $n_0 = 10^{15} \text{ см}^{-3}$, $M_i \gg 1$, $n_i = 10^{13} \text{ см}^{-2}$ и $c_n = 10^{-15} \text{ см}^{-2}$ дают $R = 5 \cdot 10^{-3} \text{ Ом} \cdot \text{см}^{-2}$. При тех же параметрах, но $n_0 = 10^{18} \text{ см}^{-2}$, $R = 5 \cdot 10^{-6} \text{ Ом} \cdot \text{см}^{-2}$. В этом случае R достаточно мало значения только для сильно легированных полупроводников.

Слабый запорный изгиб зон. В этом случае при справедливости диодной теории для вольт-амперной характеристики можно воспользоваться выражением (3.82). При $\frac{e(V_1 + V_2)}{kT} \ll 1$ это позволяет получить:

$$i = \frac{\bar{e}v_n \bar{c}_n n_i M_i n_0 e^{-\frac{\phi_0}{kT}}}{(1 + M_i) [1 + K(1 + M_i)]} \cdot \frac{e}{kT} (V_1 + V_2); \quad (8.11)$$

$$R = \frac{(1 + M_i) [1 + K(1 + M_i)] kT}{e^2 \bar{v}_n \bar{c}_n n_i M_i n_0} e^{\frac{\phi_0}{kT}}. \quad (8.12)$$

Сравнение последнего выражения с выражением (8.10) показывает, что они различаются только на величину множителя $e^{\frac{\phi_0}{kT}}$. Поэтому о зависимости (8.12) от n_0 , M_i , c_n , K , n_i можно утверждать то же, что и выше. Однако из-за множителя $e^{\frac{\phi_0}{kT}}$ сопротивление омического контакта для запорных слоев при прочих равных условиях будет большим, чем для антизапорного слоя.

8.4. ПОЛУЧЕНИЕ И ИССЛЕДОВАНИЕ ОМИЧЕСКИХ КОНТАКТОВ

Сопротивление омического контакта

$$R = \frac{4kT}{en_0 \bar{v}_n B}, \quad (8.13)$$

где $B \leq 1$ и различно для различных частных случаев, рассмотренных в предыдущих параграфах. Причем $B = 1$ реализуется только для контакта без зазора при антизапорном слое, для всех остальных случаев $B < 1$.

Большое количество экспериментальных работ было посвящено созданию омических контактов на различных полупроводниках, характеризующихся минимальным значением R . Такие работы в первую очередь были ориентированы на проверку теоретической зависимости R от концентрации примесей n_0 ($R \sim \frac{1}{n_0}$).

Большинство экспериментальных исследований подтверждают эту зависимость. На рис. 8.2 в качестве примера приведены зависимости R от n_0 для омических контактов Al—Si и In—GaAs. Сплошная линия на этом рисунке построена по формуле (4.1) при $B = 10^{-2}$. Последнее означает, что в исследованных омических контактах либо существенен промежуточный слой или запорный изгиб зон, если токи идут без участия поверхностных уровней, либо токи идут через поверхностные уровни.

Для проверки последнего утверждения проводились подробные исследования влияния на сопротивление контактов обработки поверхности полупроводника, метода нанесения металла, влияния температурной обработки созданных контактов и т. п.

Оказалось, что действительно обработка поверхности полупроводников в травителях, приводящих к уменьшению толщины окисных пленок, дает меньшие сопротивления омических контактов. Например, для контактов Al—n—Si после травления в CP-8 и

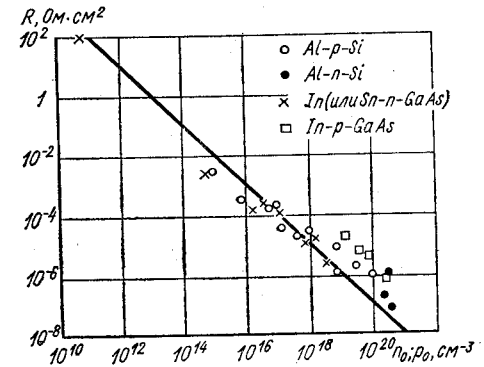


Рис. 8.2. Зависимость сопротивления омического контакта от концентрации носителей заряда.

окисления в HNO_3 , сопротивление составляет $1,6 \cdot 10^{-4} \text{ Ом} \cdot \text{см}^{-2}$, а после удаления окисного слоя в HF — $8,4 \cdot 10^{-5} \text{ Ом} \cdot \text{см}^{-2}$.

Вид металла и способ его нанесения также сказываются на сопротивлении омических контактов. Так, омические контакты $\text{Ti}-n\text{-Si}$ имели существенно меньшее сопротивление, чем контакты $\text{Al}-n\text{-Si}$.

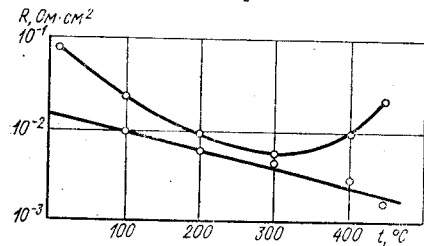


Рис. 8.3. Зависимость R от температуры обработки контакта: 1 — $\text{Al}-n\text{-Si}$ (верхняя кривая), 2 — $\text{Ni}-n\text{-Si}$ (нижняя кривая).

Проводились также экспериментальные исследования зависимости сопротивления омических контактов от режимов прогрева изготовленных контактов. В качестве примера на рис. 8.3 изображена зависимость R от температуры прогрева контакта для контактов $\text{Al}-n\text{-Si}$ и $\text{Ni}-n\text{-Si}$. Из рисунка видно, что для контактов $\text{Al}-n\text{-Si}$ существует оптимальная обработка контактов, а для $\text{Ni}-n\text{-Si}$ сопротивление уменьшается с ростом температуры обработки. Все приведенные данные свидетельствуют о сложности модели омического контакта, в которой в общем случае для объяснения экспериментальных результатов и, в частности, того факта, что всегда $B < 1$, следует учитывать промежуточный слой и поверхностные электронные состояния.

При повышенных температурах подложки при напылении металла иногда образуются соединения металл — полупроводник, характеризующиеся малыми сопротивлениями. Например, контакты $\text{Pt}-\text{Si}$, в которых образуются силициды платины Pt_3Si_2 , имеют высокое сопротивление.

При повышенных температурах подложки при напылении металла иногда образуются соединения металл — полупроводник, характеризующиеся малыми сопротивлениями. Например, контакты $\text{Pt}-\text{Si}$, в которых образуются силициды платины Pt_3Si_2 , имеют высокое сопротивление.

ЗАКЛЮЧЕНИЕ

Существует большое количество одноконтактных и многоконтактных приборов на основе контакта металл — полупроводник, гомо- и гетероперехода. Они преобразуют световые сигналы в электрические, электрические — в световые, тепловую энергию в электрическую и наоборот, давление в электрический сигнал, энергию потока частиц в электрический сигнал и т. д. Конкретные типы приборов могут создаваться преимущественно на основе контактов металл — полупроводник (обычно диоды с барьером Шоттки) и на основе гомо- или гетеропереходов.

Например, приборы для сверхвысоких частот обычно создаются на основе контакта металл — полупроводник, биполярные транзисторы — на основе $p-n$ перехода, а полупроводниковые лазеры на основе гомо- и гетеро- $p-n$ переходов.

В работе полупроводниковых приборов на основе контакта металл — полупроводник, гомо- и гетероперехода используются следующие контактные свойства: существование потенциальных барьеров для подвижных носителей заряда; зарядные свойства областей пространственного заряда (изменение заряда ионизированных примесей и подвижных носителей в области пространственного заряда с напряжением); существование электрического поля областей пространственного заряда; свойства промежуточного слоя и поверхностных состояний в контакте.

Хотя выделение перечисленных основных свойств контактов является в определенной мере условным, поскольку существование потенциального барьера (см. гл. 1) приводит и к наличию области пространственного заряда и электрического поля в этой области, оно позволяет выявить основные особенности явлений, происходящих в приборах. Например (см. гл. 2—5), именно изменение высоты потенциальных барьеров под действием приложенного напряжения приводит к протеканию тока основных и неосновных носителей заряда через контакт металл — полупроводник, гомо- и гетеропереход, изменение заряда в области пространственного заряда с приложенным напряжением приводит к появлению барьерной емкости (см. гл. 7).

Применение того или иного вида контактов для решения различных практических задач определяется конкретными достоинствами контакта металл — полупроводник, $p-n$ гомоперехода и гетероперехода.

Преимущества контакта металл — полупроводник таковы: большое быстродействие, определяемое отсутствием инжекции неосновных носителей заряда; принципиальная простота технологии; способность выполнять различные функции; возможность создания приборов на различных полупроводниках. Поэтому основная область применения контакта металл — полупроводник — преобразователи, усилители и генераторы сверхвысоких частот, импульсная техника, приборы на различных полупроводниках.

Приборы на основе $p-n$ гомопереходов в настоящее время являются наиболее распространенными приборами полупроводниковой электроники. Это связано со следующими их достоинствами: свойством инжекции неосновных носителей заряда; универсальностью приборов; отработанностью технологии изготовления приборов для таких полупроводников, как кремний и германий.

Однако быстродействие таких приборов меньше, чем у приборов на основе контакта металл — полупроводник, и определяется оно временем жизни или переноса неосновных носителей заряда. Поэтому приборы на основе $p-n$ перехода широко применяются в области не очень высоких частот.

Из рассмотренных в настоящем пособии приборов очевидным преимуществом обладают приборы на основе гетеропереходов. Они фактически объединяют достоинства контакта металл — полупроводник и $p-n$ гомоперехода. Однако эти приборы еще не нашли широкого применения главным образом из-за технологических трудностей создания приборов с наперед заданными параметрами.

КОНТРОЛЬНЫЕ ВОПРОСЫ

К главе 1

Что такое гетеро- и гомопереход? Почему гомопереход и контакт металл — полупроводник могут рассматриваться как частный случай гетероперехода? Какими свойствами обладают запорный или обедненный слой, антизапорный или обогащенный? В каком случае образуются запорные и антизапорные слои для n - и p -типа полупроводника? Чем определяется величина разрывов в валентной зоне и зоне проводимости гетероперехода? Как отличаются области пространственного заряда для анизотипного и изотипного гетеропереходов? Как влияют поверхностные состояния на образование гетероперехода? Что такое слой Шоттки и каковы условия его существования? От чего зависит ширина слоя Шоттки? Чем определяется соотношение между высотами барьеров в гетеропереходе? Чем определяется общая ширина области пространственного заряда гетероперехода? Как влияют силы изображения на ход потенциала в гетеропереходе при разных соотношениях диэлектрической постоянной двух полупроводников? Какие основные результаты экспериментального исследования модели контакта металл — полупроводник, гомо- и гетероперехода?

К главе 2

Какие основные области можно выделить в гетеропереходе при рассмотрении движения носителей заряда? Чем отличается прохождение носителей заряда через изотипный и анизотипный гетеропереход? Как проходят основные и неосновные носители квазинейтральную область? Определить возможные механизмы прохождения носителей заряда через область пространственного заряда? Какие критерии диодной и диффузионной теории выпрямления? Чем определяются соотношения между напряжениями в различных частях гетероперехода?

К главе 3

Какие приближения допущены при рассмотрении диодной и диффузионной теории выпрямления контакта металл — полупроводник? Чем отличаются выражения вольт-амперной характеристики в предположениях диодной и диффузионной теории для одного типа носителей заряда? Когда и почему они совпадают? Как отличаются вольт-амперные характеристики для случая, когда есть промежуточный слой и когда его нет? Какие особенности вольт-амперной характеристики и коэффициента инжекции в предположениях диодной и диффузионной теории при учете неосновных носителей заряда? Как изменяет вольт-амперную характеристику туннелирование через область пространственного заряда основных и неосновных носителей заряда? Почему учет туннелирования более важен для диодной теории, чем для диффузионной? Как влияют силы изображения на вольт-амперную характеристику контакта с промежуточным слоем и без него? Какова природа токов через поверхностные уровни? Как понимать, что главную роль играет обмен поверхностных уровней с полупроводником или металлом? Когда совпадают и различаются вольт-амперные характеристики для токов через поверхностный уровень в предположениях диодной и диффузионной теории? Когда токи через поверхностные уровни будут больше, чем токи без участия поверхностных уровней? Чем определяются параметры α и α^* вольт-амперной характеристики? Какая связь между параметрами вольт-амперной характеристики α и $\bar{\alpha}$, α^* и $\bar{\alpha}^*$? Какие основные закономерности при исследовании вольт-амперной характеристики контакта металл — полупроводник получены экспериментальным путем?

К главе 4

Чем отличается подход при рассмотрении вольт-амперной характеристики контакта металл — полупроводник и $p-n$ перехода? Какие предположения допущены при рассмотрении концентрации неосновных носителей заряда на границе области пространственного заряда и квазинейтральной области в диодной и диффузионной теории? Какие граничные условия в квазинейтральной области для концентрации неосновных носителей справедливы при широкой и узкой квазинейтральной области? Какой общий вид имеют вольт-амперные характеристики

p — n переходов с широкой и узкой базой при малых уровнях инжекции? Как коэффициент инжекции зависит от параметров n - и p -областей, скорости поверхностной рекомбинации на невыпрямляющих контактах? Какие особенности имеет вольт-амперная характеристика и коэффициент инжекции при больших уровнях инжекции? Как проходит ток в условиях генерации — рекомбинации в p — n переходе? Почему при прямых напряжениях существенна рекомбинация, а при обратных — генерация? Какое соотношение между током генерации и рекомбинации и током над барьером области пространственного заряда? Как в p — n переходе идут избыточные токи? Как идут токи в p — n переходе между вырожденными полупроводниками? Какой вид имеет вольт-амперная характеристика для туннельного диода? Как подтверждаются выводы теории выпрямления p — n перехода экспериментально?

К главе 5

Как идут носители заряда в изотипных гетеропереходах? Как идут носители заряда в анизотипных переходах? В каких предположениях вольт-амперные характеристики изотипных и анизотипных гетеропереходов совпадают с вольт-амперными характеристиками контакта металл — полупроводник и p — n перехода? Как влияют туннелирование через область пространственного заряда и силы изображения на вольт-амперные характеристики гетеропереходов? Какие особенности вольт-амперных характеристик гетероперехода для токов через поверхностные уровни? Что означает преимущественный обмен поверхностных уровней с одним или другим полупроводником? Как сказывается на рассмотрении вольт-амперных характеристик сохранение или изменение вида разрешенной зоны для изотипного и анизотипного полупроводника? Какие особенности рассмотрения генерационно-рекомбинационных и избыточных токов в гетеропереходе по сравнению с p — n переходом? Какие трудности возникают при сравнении теории и эксперимента в случае гетероперехода?

К главе 6

В чем заключается суть пробоя? Какие механизмы могут приводить к возникновению пробоя? Какая зависимость пробойного напряжения от параметров полупроводника для лавинного пробоя? Чем определяется пробойное напряжение для туннельного пробоя? Каковы особенности теплового пробоя? Какая физическая природа микроплазм? Как экспериментально можно различить механизмы пробоя?

К главе 7

Какие физические причины могут вызывать сдвиг фаз между амплитудой тока и амплитудой переменного напряжения? Что такое барьерная емкость и какова ее физическая природа? Как изменяется сопротивление и емкость контакта металл — полупроводник для одного типа носителей заряда от напряжения для разных толщин промежуточного слоя? Как влияют поверхностные уровни на эквивалентную схему контакта металл — полупроводник? Почему и как сопротивление и емкость контакта при учете поверхностных состояний зависят от частоты? Почему сопротивление и емкость p — n перехода зависят от частоты? Что такое диффузионная емкость и какова ее физическая природа? Как различаются барьерная и диффузионная емкости? Какие особенности имеет эквивалентная схема гетероперехода? Чем отличается эквивалентная схема p — n перехода на участке лавинного пробоя? Какие основные результаты экспериментальных исследований эквивалентной схемы гомо- и гетеропереходов?

К главе 8

Какие основные требования предъявляются к омическим контактам? При каких условиях в области пространственного заряда может осуществляться омический контакт? Как зависит сопротивление омического контакта от концентрации основных носителей заряда? Как зависит сопротивление омического контакта от концентрации поверхностных состояний? Как проверены экспериментально основные теоретические представления об омических контактах?

СПИСОК ЛИТЕРАТУРЫ

Губанов А. И. Теория выпрямляющего действия полупроводников.— М.: ГИТТЛ, 1956.

Пикус Г. Е. Основы теории полупроводниковых приборов.— М.: Наука, 1965.

Федотов Я. А. Основы физики полупроводниковых приборов.— М.: Сов. радио, 1969.

Зи С. М. Физика полупроводниковых приборов.— М.: Энергия, 1973.

Алферов Ж. И.— В кн.: Физика сегодня и завтра.— М.: Наука, 1973.

Стриха В. И. Теоретические основы работы контакта металл — полупроводник.— К.: Наук. думка, 1974.

Стриха В. И., Бузанева Е. В., Радзневский И. А. Полупроводниковые приборы с барьером Шоттки.— М.: Сов. радио, 1974.

Милнс А., Фойхт Д. Гетеропереходы и переходы металл — полупроводник.— М.: Мир, 1975.

Климов Б. Н., Цукерман Н. М. Гетеропереходы в полупроводниках.— Саратов: Изд-во Саратовского университета, 1976.

Шарма Б. Л., Пурохит Р. К. Полупроводниковые гетеропереходы.— М.: Сов. радио, 1979.

Викулин И. М., Стафеев В. И. Физика полупроводниковых приборов.— М.: Сов. радио, 1980.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Барьер потенциальный
— влияние разности работ выхода 13—17
образование поверхностных состояний 13, 19, 26
— приложенное напряжение 36—38
— силы изображения 30—32
— соотношение между высотами барьеров 25—27
— Шоттки 22, 23
- Быстродействие диодов 184
- Вольт-амперная характеристика гетеропереходов
— анализ 138—140, 142, 143, 152, 155, 159
— в диффузионной теории 135—138, 141, 144, 152, 154, 156
— влияние туннелирования через область пространственного заряда 143, 145
— сил изображения 146
— генерационно-рекомбинационные токи 156, 157
— токи через поверхностные уровни 147—150, 152, 154, 156
— при движении носителей в пределах одного типа разрешенных зон 147—150, 152
— с изменением вида разрешенных зон 154, 156
— анизотипных 142, 143, 152, 154
— изотипных 135, 138, 149, 156
— параметры и их анализ 159—167
— экспериментальные исследования 167—170
- Вольт-амперная характеристика контакта металл — полупроводник
— анализ 58—61, 64, 69, 72, 73, 81—85
— в диодной теории 58—69
— диффузионной теории 53—61, 64—69
— влияние сил изображения 73—76
— параметры и их анализ 85—94
— с учетом одного типа носителей заряда 55—61
— двух типов носителей заряда 61—69
— туннелирования через область пространственного заряда 69—73
— токов через поверхностные состояния 76—85
— экспериментальные исследования 94—98
- Вольт-амперная характеристика $p-n$ -перехода
— анализ 105, 106, 109—111, 117—119, 127, 128
— в диодной теории 105
— в диффузионной теории 105
— для невырожденных полупроводников при малых уровнях инжекции 105, 109
— с широкой базой 115
— с узкой базой 109
— для вырожденных полупроводников 127
— зависимость от скорости поверхностной рекомбинации 109, 110
— избыточные токи в тонких $p-n$ -переходах 122
— при больших уровнях инжекции 112, 114
— с учетом генерации — рекомбинации 117, 118
— экспериментальные исследования 128—132
- Гетероопереход
анизотипный 16
изотипный 16
— зонная модель 14—19
— образование 14—19
— при приложении напряжения 36—39
— параметры эквивалентной схемы 199—201
— зависимость от напряжения 200
— от частоты 201
— применение 218
— соотношение между напряжениями 51—54
— эквивалентная схема 188
- Диоды с барьером Шоттки (см. также Контакт металл — полупроводник с барьером Шоттки)
Емкость гетероперехода 200
Емкость контакта металл — полупроводник
— области пространственного заряда (барьерная) 185, 188, 189
— зазора (промежуточного слоя) 192
— с учетом поверхностных уровней 192
- Емкость $p-n$ -перехода
— барьерная 198
— диффузионная 198
- Зазор (промежуточный слой)
— коэффициент прозрачности 50
— для трапециoidalного барьера 50
— для электронов 55, 135

— для дырок 62
— влияние сил изображения 74
Заряд пространственный 13, 14, 18, 21, 23, 24

Изгиб энергетических зон (см. также Барьер потенциальный) 13—19, 21, 23—24

Инжекция
в контакте металл — полупроводник 63, 64, 67, 68
— $p-n$ -переходе 37, 106, 107, 111, 118,
— гетеропереходе 37, 38

Контакт металл — полупроводник с барьером Шоттки
— зонная модель 23

— параметры эквивалентной схемы
— без учета поверхностных состояний 187, 188

— — зависимость от напряжения 188
— с учетом поверхностных состояний 192

— — зависимость от частоты 192, 193
— для токов через поверхностные уровни 194

— — зависимость от частоты 194, 196
— применение 217, 218
— эквивалентная схема 188

Контакт омический
— свойства и модель 211, 219
— сопротивление 212—215

— для токов без участия поверхностных состояний 212, 213
— для токов через поверхностные состояния 214

— при учете туннелирования через область пространственного заряда 213
— влияние обработки поверхности полупроводника 215

— метод нанесения металла 216
— температурная обработка контакта 216
— зависимость от концентрации носителей заряда 215

Коэффициент инжекции
— контакт металл — полупроводник 63, 65
— анализ 64, 67
— для диодной теории 63
— для диффузионной теории 64
— $p-n$ -перехода 107, 111
— с узкой базой 111
— с широкой базой 107

Коэффициенты кинетические 50, 51, 78, 147, 153

Область пространственного заряда
— антизапорный слой 13, 14, 16
— запорный слой 13, 14
— коэффициент прозрачности 46, 70, 71, 122, 126
— ширина 12, 24, 29, 30

Переход $p-n$
— зонная модель 17
— параметры эквивалентной схемы 197, 198, 206

— — зависимость от частоты 198, 209
— — от напряжения 198, 209
— применение 217

— эквивалентная схема 188, 202
— на участке лавинного пробоя 202

Поверхностные электронные состояния 17, 18, 25, 38
— параметры 25, 33—35, 78—80, 147, 148, 153

— перезаполнение 78, 147, 148, 153
— спектр энергетический 96, 97
ток через поверхностный уровень 77, 147, 153

Потенциальная энергия электрона (см. энергия потенциальная)
Потенциальный барьер (см. барьер потенциальный)

Пробой
— вольт-амперные характеристики 172, 176, 178
— коэффициент лавинного умножения 173

— ударной ионизации 173
— лавинный 171, 173, 176
— микроплазма 173, 179, 180
— пробивное напряжение 175—177, 179

— тепловой 173, 177, 179
— туннельный 171, 176
— экспериментальные исследования 180—182

Пуассона уравнение 20, 23, 24, 202

Разность потенциалов контактная 15, 18, 27, 28, 32

Силы изображения 30—32, 73—76, 146
Сопротивление дифференциальное
— контакта металл — полупроводник 187, 188, 192
— $p-n$ -перехода 198
— гетероперехода 200

Шоттки барьер (см. барьер Шоттки)
— слой Шоттки 22—25
— ширина 23, 25, 29, 30

Энергия электрона потенциальная 20, 21, 30

Виталий Илларионович Стриха

**КОНТАКТНЫЕ
ЯВЛЕНИЯ
В ПОЛУПРОВОДНИКАХ**

Редактор *И. В. Мисюренко*
Переплет художника *Г. М. Балюна*
Художественный редактор *С. П. Духленко*
Технический редактор *Т. И. Трофимова*
Корректор *С. Я. Куроченко*

Информ. бланк № 6702

Сдано в набор 31.10.80. Подп. в печать 08.12.81.
БФ 10209. Формат 60×90^{1/16}. Бумага типогр. № 3.
Лит. гарн. Выс. печать 14 печ. л. 14,12 кр.-отт. 14,82
уч.-изд.л. Тираж 2000. Изд. № 4905. Зак. № 1—687.
Цена 65 к.

Головное издательство издательского объединения
«Вища школа», 252054, Киев-54, ул. Гоголевская, 7.
Отпечатано с матриц Головного предприятия рес-
публиканского производственного объединения
«Полиграфкнига» Госкомиздата УССР, в Харь-
ковской городской типографии № 16, Харьков-3,
ул. Университетская, 16. Зак. 1716.