

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
Томский государственный университет
систем управления и радиоэлектроники

М.М. Славникова

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОЭЛЕКТРОНИКИ

Учебное пособие
(для автоматизированной
технологии обучения)

Издание третье, исправленное и дополненное

Томск
ТУСУР
2014

УДК 621.753.1/.2
ББК 34.417.2я75
Н55

Славникова М.М.
Н55 **Физические основы микро- и нанoeлектроники : учеб. пособие / М.М. Славникова** – Томск : Томск. гос. ун-т систем упр. и радиоэлектроники, 2014. – 276 с.

Приводятся основные положения физики твердого тела, рассматривается статистика электронов и дырок, электропроводность, генерация и рекомбинация, диффузия и дрейф носителей заряда в полупроводниках. Описываются контактные и поверхностные явления в полупроводниках, основные процессы в электронно-дырочном переходе. В сокращенной форме рассматриваются гальвано-, термомагнитные, термоэлектрические и фотоэлектрические явления в полупроводниках. Широко используются упрощенные модели. Каждая глава заканчивается контрольными вопросами и задачами.

Для студентов, обучающихся по направлению 211000.62 «Конструирование и технология электронных средств», а также других специальностей, изучающих дисциплины «Физика полупроводников» и «Твердотельная электроника».

УДК 621.753.1/.2
ББК 34.417.2я75

Славникова Мария Михайловна

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО И НАНОЭЛЕКТРОНИКИ
Учебное пособие

Томский государственный университет
систем управления и радиоэлектроники.
634050, Томск, пр. Ленина, 40. Тел. (3822) 533018.

© Томск. гос. ун-т систем упр.
и радиоэлектроники, 2014
© Славникова М.М., 2014

Оглавление

Список условных обозначений и сокращений	7
Предисловие	11
Глава 1. Элементы зонной теории твердых тел	
§ 1.1. Энергия абсолютно свободного электрона и электрона в потенциальной яме	13
§ 1.2. Энергия электрона в свободном атоме	15
§ 1.3. Модель периодического потенциального поля в кристалле	18
§ 1.4. Зонный характер энергетического спектра электронов в кристалле	20
§ 1.5. Модель Кронига — Пенни	22
§ 1.6. Проводник, полупроводник, диэлектрик в свете зонной теории твердого тела	27
§ 1.7. Эффективная масса носителей заряда	30
§ 1.8. Физический смысл эффективной массы	32
§ 1.9. Понятие о дырках как носителях заряда	34
§ 1.10. Собственные полупроводники	34
§ 1.11. Донорный полупроводник	37
§ 1.12. Акцепторный полупроводник	39
Контрольные вопросы и задачи	41
Глава 2. Статистика электронов и дырок в полупроводниках	
§ 2.1. Распределение электронов в твердых телах по энергетическим уровням	45
§ 2.2. Плотность квантовых состояний	47
§ 2.3. Концентрация электронов и дырок в полупроводнике	50
§ 2.4. Положение уровня Ферми и концентрация свободных носителей в собственных полупроводниках	52
§ 2.5. Положение уровня Ферми и концентрация свободных носителей в примесных полупроводниках	54
§ 2.6. Основные и неосновные носители заряда. Закон действующих масс	57
§ 2.7. Влияние уровня легирования на температурную зависимость концентрации основных носителей заряда	59
§ 2.8. Компенсированные полупроводники	60
Контрольные вопросы и задачи	61
Глава 3. Электропроводность твердых тел	
§ 3.1. Влияние электрического поля на функцию распределения носителей заряда в полупроводниках	66
§ 3.2. Электропроводность полупроводников и подвижность носителей заряда	68
§ 3.3. Зависимость подвижности носителей заряда от температуры	70

§ 3.4. Зависимость электропроводности полупроводников от температуры	75
§ 3.5. Электропроводность металлов и сплавов	77
Контрольные вопросы и задачи	78
Глава 4. Электропроводность полупроводников в сильном электрическом поле	
§ 4.1. Влияние электрического поля на электропроводность полупроводников	83
§ 4.2. Термоэлектронная ионизация Френкеля	84
§ 4.3. Ударная ионизация	84
§ 4.4. Электростатическая ионизация	86
§ 4.5. Подвижность носителей заряда в сильном электрическом поле	88
§ 4.6. Эффект междолинного рассеяния электронов	89
§ 4.7. Электропроводность двухдолинных полупроводников в сильном электрическом поле	90
§ 4.8. Эффект Ганна	91
Контрольные вопросы и задачи	93
Глава 5. Генерация и рекомбинация носителей заряда	
§ 5.1. Тепловая генерация	97
§ 5.2. Неравновесные процессы в полупроводниках	98
§ 5.3. Рекомбинация свободных носителей заряда	100
§ 5.4. Скорость рекомбинации и время жизни неравновесных носителей заряда	101
§ 5.5. Рекомбинация через локальные уровни	102
§ 5.6. Зависимость времени жизни неравновесных носителей от концентрации основных носителей заряда при рекомбинации через локальные уровни	103
§ 5.7. Температурная зависимость времени жизни носителей заряда	106
§ 5.8. Поверхностная рекомбинация	107
Контрольные вопросы и задачи	108
Глава 6. Диффузия и дрейф носителей заряда в полупроводниках	
§ 6.1. Диффузионные потоки носителей заряда	113
§ 6.2. Диффузионный и дрейфовый токи носителей заряда	114
§ 6.3. Соотношение Эйнштейна	115
§ 6.4. Диффузионная длина носителей заряда	116
§ 6.5. Уравнение непрерывности	118
Контрольные вопросы и задачи	120
Глава 7. Контактные явления	
§ 7.1. Работа выхода электрона	123
§ 7.2. Контактная разность потенциалов	125
§ 7.3. Контакт металл-полупроводник	127

§ 7.4. Влияние внешнего электрического поля на высоту потенциального барьера и толщину запирающего слоя	131
§ 7.5. Вольт-амперная характеристика контакта металл-полупроводник	133
Контрольные вопросы и задачи	135
Глава 8. Физические явления в электронно-дырочном переходе	
§ 8.1. Электронно-дырочный переход в равновесном состоянии	139
§ 8.2. Вольт-амперная характеристика <i>p-n</i> -перехода	146
§ 8.3. Определение тока насыщения	149
§ 8.4. Пробой <i>p-n</i> -перехода	150
§ 8.5. Энергетические диаграммы <i>p-n</i> -перехода при наличии внешнего электрического поля	152
§ 8.6. Емкость <i>p-n</i> -перехода	153
§ 8.7. Структура биполярного транзистора	155
§ 8.8. Принцип работы биполярного транзистора	156
§ 8.9. Энергетические зонные диаграммы биполярного транзистора	159
§ 8.10. Коэффициент усиления биполярного транзистора по току	160
§ 8.11. Выходные вольт-амперные характеристики биполярного транзистора	163
Контрольные вопросы и задачи	165
Глава 9. Поверхностные явления в полупроводниках	
§ 9.1. Поверхностные состояния	176
§ 9.2. Приповерхностный слой объемного заряда	178
§ 9.3. Заряд в приповерхностной области полупроводника	180
§ 9.4. Емкость структуры металл-диэлектрик-полупроводник	182
§ 9.5. Эффект поля	185
§ 9.6. Полевые транзисторы с изолированным затвором	188
§ 9.7. Принцип работы и характеристики МДП-транзистора со встроенным каналом	190
§ 9.8. Принцип работы и характеристики МДП-транзистора с индуцированным каналом	191
Контрольные вопросы и задачи	193
Глава 10. Гальваномагнитные, термомагнитные и термоэлектрические явления в полупроводниках	
§ 10.1. Эффект Холла	199
§ 10.2. Использование эффекта Холла для определения основных параметров полупроводников	201
§ 10.3. Магниторезистивный эффект. Термомагнитные явления	203
§ 10.4. Эффект Зеебека	204
§ 10.5. Эффект Пельтье	206
§ 10.6. Эффект Томсона	208
Контрольные вопросы и задачи	209

Глава 11. Фотоэлектрические явления в полупроводниках	
§ 11.1. Спектры отражения и поглощения	212
§ 11.2. Фотопроводимость полупроводников	217
§ 11.3. Фотопроводимость при импульсном освещении	219
§ 11.4. Фотоэлектродвижущая сила	220
§ 11.5. Использование фотоэлектрических явлений в полупроводниковых приборах	222
Контрольные вопросы и задачи	227
Приложение. Некоторые сведения о строении твердых тел	230
Список рекомендуемой литературы	276

Список условных обозначений и сокращений

A	— постоянная Ричардсона, Холл-фактор
a^{\pm}	— ускорение носителя заряда (электрона, дырки)
B	— вектор магнитной индукции
c	— скорость света в вакууме, $c = 2,998 \cdot 10^8$ м/с
D_n	— коэффициент диффузии электронов
D_p	— коэффициент диффузии дырок
d_n, d_p	— толщина слоя объемного заряда в полупроводнике n - и p -типа соответственно
$d\omega(x, y, z)$	— вероятность обнаружения микрочастицы в момент времени t в элементе объемом $dx dy dz$
E	— полная энергия микрочастицы
E_k	— кинетическая энергия микрочастицы
E	— напряженность электрического поля
E_a	— энергия акцепторного уровня
E_c	— энергия, соответствующая нижнему краю зоны проводимости
E_d	— энергия донорного уровня
E_F	— энергия уровня Ферми
E_i	— энергия, соответствующая середине запрещенной зоны полупроводника
E_v	— энергия верхнего края (потолка) валентной зоны
F	— сила
$f(E)$	— распределение микрочастиц по энергии
$f_{Б-Э}(E, T)$	— функция распределения Бозе — Эйнштейна
$f_{М-Б}(E, T)$	— функция распределения Максвелла — Больцмана
$f_{Ф-Д}(E, T)$	— функция распределения Ферми — Дирака
G	— скорость генерации носителей заряда
h	— постоянная Планка, $h = 6,62 \cdot 10^{-34}$ Дж·с = $4,14 \cdot 10^{-15}$ эВ·с
I	— величина потока, тока
J	— интенсивность излучения
j	— плотность электрического тока
K	— кельвин, единица измерения температуры
k	— волновое число
k_0	— постоянная Больцмана, $k_0 = 1,38 \cdot 10^{-23}$ Дж/К = $8,625 \cdot 10^{-5}$ эВ/К
L_n	— диффузионная длина пробега электронов
L_p	— диффузионная длина пробега дырок
l	— средняя длина свободного пробега микрочастиц
l_n	— средняя длина свободного пробега микрочастиц при рассеянии на ионах
l_{Φ}	— средняя длина свободного пробега микрочастиц при рассеянии на фононах
m	— масса микрочастицы

m_0	— масса покоя электрона, $m_0 = 9,11 \cdot 10^{-31}$ кг
	— эффективная масса электрона
	— эффективная масса дырки
N	— концентрация атомов
N_a	— концентрация акцепторных уровней
N_c	— эффективная плотность состояний в зоне проводимости
N_d	— концентрация донорных уровней
$N(E)$	— плотность разрешенных состояний в интервале энергий от E до $E + dE$
N_v	— эффективная плотность состояний в валентной зоне
n_0	— равновесная концентрация электронов
n_i	— концентрация собственных носителей заряда
n_n	— концентрация основных носителей заряда — электронов в донорном полупроводнике
n_p	— концентрация неосновных носителей заряда — электронов в дырочном полупроводнике
n_Φ	— концентрация фононов
$P_{\text{вх}}$	— входная мощность
$P_{\text{вых}}$	— выходная мощность
p	— импульс микрочастицы
p_0	— равновесная концентрация дырок
p_p	— концентрация дырок в акцепторном полупроводнике
p_n	— концентрация дырок в донорном полупроводнике
Q_S	— плотность поверхностного заряда
q	— заряд электрона, $q = 1,6 \cdot 10^{-19}$ Кл
R	— коэффициент отражения
R_H	— постоянная Холла
R_n	— скорость рекомбинации электронов
R_p	— скорость рекомбинации дырок
r	— радиус-вектор
S	— площадь
s	— скорость поверхностной рекомбинации
T	— температура
T_i	— температура перехода полупроводника от примесной проводимости к собственной
T_s	— температура истощения примеси
U	— потенциальная энергия микрочастицы
V	— потенциал
V_H	— холловская разность потенциалов
V_k	— контактная разность потенциалов
v	— скорость микрочастицы
v_d	— скорость дрейфа микрочастицы
$v_{n \text{ диф}}$	— скорость диффузии электронов

m_p^*

v_p диф	— скорость диффузии дырок
W	— ширина базы биполярного транзистора
Z	— порядковый номер элемента в таблице Д.И. Менделеева
α	— коэффициент поглощения излучения; коэффициент ударной ионизации; коэффициент усиления биполярного транзистора по току
α'	— коэффициент пропорциональности
β	— коэффициент пропорциональности; коэффициент переноса
ΔE	— ширина запрещенной зоны
Δn	— концентрация неравновесных электронов
Δp	— концентрация неравновесных дырок
$\Delta \sigma$	— фотопроводимость
ϵ	— диэлектрическая проницаемость
ϵ_0	— диэлектрическая проницаемость вакуума, или электрическая постоянная, $\epsilon_0 = 8,85 \cdot 10^{-12}$ Ф/м
λ	— длина волны
μ	— подвижность
μ_n	— подвижность электронов
μ_p	— подвижность дырок
ν	— частота излучения
ρ	— удельное сопротивление
σ	— электропроводность
τ	— время релаксации, время жизни
τ_n	— время жизни неравновесных электронов
τ_p	— время жизни неравновесных дырок
ϕ	— потенциальная энергия электрона
ϕ_0	— высота потенциального барьера
ϕ_b	— энергетическое расстояние от уровня Ферми до дна зоны проводимости
χ	— термодинамическая работа выхода
ψ	— волновая функция
ω	— угловая частота
БПС	— быстрые поверхностные состояния
ВАХ	— вольт-амперная характеристика
ВФХ	— вольт-фарадная характеристика
КПД	— коэффициент полезного действия
КРП	— контактная разность потенциалов
МДП	— металл-диэлектрик-полупроводник (структура)
МПС	— медленные поверхностные состояния
ОПЗ	— область приповерхностного (пространственного) заряда
ПЭС	— прибор с зарядовой связью
СВЧ	— сверхвысокая частота
ФП	— фотопроводимость
ЭДС	— электродвижущая сила

Некоторые свойства германия и кремния

Параметр	Германий	Кремний
Ширина запрещенной зона при 300 К, эВ	0,67	1,11
Дрейфовая подвижность при 300 К, м ² /В·с: электронов дырок	0,39 0,18	0,15 0,06
Концентрация собственных носителей при 300 К, м ⁻³	$2,5 \cdot 10^{19}$	$1,5 \cdot 10^{16}$
Относительная диэлектрическая проницаемость	16	12

Предисловие

В настоящее время во всех областях промышленности, науки, техники и в быту используются электронные устройства вычислительной и измерительной техники, автоматики и видеоаппаратура. Это является результатом развития интегральной технологии, внедрение которой позволило наладить массовый выпуск микроэлектронных функциональных узлов различного назначения. В качестве таких узлов широко применяются полупроводниковые интегральные схемы. Функции, выполняемые интегральными схемами, позволяют собирать нужные электронные блоки без детального расчета отдельных каскадов.

Эффективное применение интегральных микросхем невозможно без знания физических основ и принципов работы отдельных элементов: полупроводниковых диодов различного назначения, транзисторов, фоторезисторов и других дискретных приборов. Теоретическим фундаментом твердотельной электроники для объяснения сложных и многообразных физических процессов, происходящих при протекании электрического тока в полупроводниковых структурах и в контактах различных полупроводников, являются современные представления о строении твердых тел, природе и физической сущности электричества, света, тепла и т.п.

Некоторые физические явления в твердых телах можно объяснить с позиций классической физики, в которой электрону приписываются только корпускулярные свойства, а свет рассматривается как поток электромагнитных волн. Однако для объяснения свойств полупроводников требуется еще привлечение квантовых представлений.

В предлагаемом учебном пособии рассматриваются энергетические спектры электронов в различных типах твердых тел. С позиций зонной теории описаны особенности электропроводности металлов, полупроводников и диэлектриков, поведение полупроводников в сильных электрических полях. Рассмотрены генерационно-рекомбинационные, диффузионные и поверхностные явления в полупроводниках, в структурах металл-полупроводник и металл-диэлектрик-полупроводник, изложена физика работы диодов на основе электронно-дырочных *p-n*-переходов, диодов Шоттки и Ганна, биполярных и полевых транзисторов и т.п.

Последние главы пособия посвящены описанию физики магнитно-термоэлектрических явлений в твердых телах, а также эффектов, возникающих при взаимодействии оптического излучения с полупроводниками. Кратко рассмотрены вопросы использования этих явлений и эффектов при создании дискретных элементов различного назначения (датчики Холла, термоэлементы, фоторезисторы, фотодиоды, солнечные элементы и пр.).

Контрольные вопросы и задачи, приведенные в конце каждой главы, начинаются с буквенно-цифрового кода. Использование кода дает возможность студентам с помощью микропроцессорного контролирующего

устройства СИМВОЛ-ВУЗ проверить понимание и усвоение изучаемого материала. Для этого каждый вопрос имеет готовые ответы, обозначенные номерами 1, 2, 3, ... Достаточно на подготовленном к работе устройстве после кнопки СБРОС нажать поочередно кнопки кода вопроса и ответа или ответов, после чего, нажав и отпустив кнопку КОНТРОЛЬ, можно по индикаторам ПРАВИЛЬНО и НЕПРАВИЛЬНО оценить результат.

Содержание учебного пособия соответствует требованиям к дисциплине «Физические основы микроэлектроники», предусмотренным Государственным образовательным стандартом высшего профессионального образования для специальностей 210201 «Проектирование и технология радиоэлектронных средств» и 210202 «Проектирование и технология электронно-вычислительных средств».

Курс «Физические основы микроэлектроники» дает студентам базовые знания для успешного освоения ряда специальных дисциплин, таких как «Элементы микроэлектроники», «Основы радиоэлектроники», «Конструирование и технология микросхем и микропроцессоров», «Устройства функциональной электроники», «Физика полупроводниковых приборов» и других.

Главы 1–5 написаны доцентом А.А. Широковым, главы 6–9 — доцентом М.М. Славниковой, главы 10, 11 — профессором Н.С. Несмеловым, введение — Н.С. Несмеловым и А.А. Широковым, приложение — Н.С. Несмеловым и М.М. Славниковой. Общая редакция осуществлена профессором Н.С. Несмеловым.

Первое издание учебного пособия осуществлено в 2000 году и очень быстро разошлось среди студентов. Опыт его использования в учебном процессе позволил выявить ряд недостатков. Например, установлено, что студенты, приступая к изучению курса «Физические основы микроэлектроники», довольно часто не имеют базовых представлений о кристаллической структуре твердого тела, дефектах кристаллической решетки и процессах диффузии атомов примеси в них. Во втором издании учебного пособия замеченные недостатки устранены. В списке рекомендуемой литературы указаны книги, которые можно использовать для более глубокого освоения изучаемой дисциплины.

Авторы

Глава 1. Элементы зонной теории твердых тел

§ 1.1. Энергия абсолютно свободного электрона и электрона в потенциальной яме

Электрический ток в твердых телах может быть обусловлен движением различных заряженных частиц, среди которых в первую очередь следует выделить электрон. Энергия электрона зависит от условий окружающей среды и может изменяться под воздействием температуры, излучений, электрического и магнитного полей и т.д.

Согласно гипотезе французского физика Луи де Бройля электрон, как все микрочастицы, обладает корпускулярно-волновыми свойствами (дуализмом). Электрону с энергией E и импульсом p должна соответствовать волна с частотой

$$(1.1)$$

где ω — угловая частота, рад/с; $\hbar = h/2\pi$ — приведенная постоянная Планка.

В то же время электрон как частица обладает импульсом

$$p = \hbar k = mv, \quad (1.2)$$

где

$$\psi(x, y, z, t) = C e^{i(\omega t - kr)}. \quad (1.3)$$

λ — длина волны; m — масса, v — скорость.

Исходя из приведенной гипотезы де Бройля, австрийский физик-теоретик Э. Шредингер в 1926 г. разработал теорию движения микрочастиц — квантовую механику, в основу которой положил волновое уравнение, играющее в атомных процессах фундаментальную роль и имеющее вид

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) + U(x, y, z) \psi = E \psi, \quad (1.4)$$

где $U(x, y, z)$ — потенциальная энергия микрочастицы; $\psi(x, y, z, t)$ — полная волновая функция.

Уравнение (1.4) удовлетворяется только комплексными волновыми функциями вида

$$\psi(x, y, z, t) = C e^{i(\omega t - kr)}, \quad (1.5)$$

где C — амплитуда волны; ω — угловая частота; \mathbf{k} — волновой вектор; \mathbf{r} — радиус-вектор; t — время.

Согласно интерпретации немецкого физика-теоретика М. Борна физический смысл имеет только произведение волновой функции $\psi(x, y, z, t)$ на сопряженную функцию $\psi^*(x, y, z, t)$:

$$(1.6)$$

Произведение характеризует плотность вероятности нахождения микрочастицы в момент времени t в единичном объеме пространства в окрестности точки, положение которой задано радиус-вектором \mathbf{r} .

Таким образом, вероятность обнаружить микрочастицу в момент времени t в выделенном объеме dV будет

$$d\omega(x, y, z, t) = \psi(x, y, z, t) \psi^*(x, y, z, t) dV. \quad (1.7)$$

Обычно изучение свойств электрона начинают с рассмотрения простейшей модели. Электрон перемещается абсолютно свободно вдоль одного направления, например x . В этом случае он не взаимодействует с другими частицами и не испытывает влияния внешних сил, то есть его потенциальная энергия равна нулю:

$$U(x) = 0. \quad (1.8)$$

Решение уравнения (1.4) при условии (1.8) приводит к выражению для энергии абсолютно свободного электрона:

$$(1.9)$$

Выражение (1.9) показывает, что энергия абсолютно свободного электрона, рассчитанная с позиций корпускулярно-волновой природы микрочастиц, совпадает с выражением из классической физики для кинетической энергии материального тела.

В общем случае, когда направление движения свободной микрочастицы не совпадает ни с одной из осей координат, волновое число представляет собой вектор

$$\mathbf{k}^2 = k_x^2 + k_y^2 + k_z^2, \quad (1.10)$$

и энергия абсолютно свободного электрона определяется выражением

$$E = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2). \quad (1.11)$$

Из формулы (1.11) с учетом (1.10) следует, что энергия абсолютно свободного электрона пропорциональна квадрату волнового вектора. Так как величина \mathbf{k} не имеет никаких ограничений, то свободный электрон

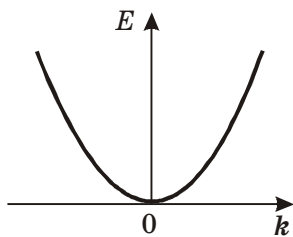


Рис. 1.1. Энергетический спектр электрона в вакууме

может обладать любой энергией. Энергетический спектр свободного электрона является сплошным (непрерывным), ему соответствует квадратичная парабола (рис. 1.1), т.е. дисперсионная кривая $E(\mathbf{k})$.

Если абсолютно свободное движение электрона в направлении, например, x ограничено бесконечно высокими потенциальными барьерами при $x = 0$ и $x = a$ (рис. 1.2), то решение уравнения Шредингера (1.4) дает сле-

$$\left(\frac{\partial \psi}{\partial x}\right)^* = \frac{\hbar^2 k_x^2}{2m}$$

дующее выражение для энергии электрона в бесконечно глубокой потенциальной яме:

$$E = \frac{\hbar^2 \pi^2}{2ma^2} n^2, \quad (1.12)$$

где $n = 1, 2, 3, \dots$; a — ширина потенциальной ямы.

Анализ выражения (1.12) показывает, что энергия электрона, «запертого» в бесконечно глубокой потенциальной яме, может принимать лишь вполне определенные дискретные значения, которые называются энергетическими уровнями. Электрон может иметь энергию, соответствующую одному из энергетических уровней. Целое число n , определяющее возможные значения энергии электрона, называется квантовым числом.

На рис. 1.3 показана схема расположения энергетических уровней электрона в бесконечно глубокой потенциальной яме. Как следует из формулы (1.12), энергетический зазор между разрешенными энергиями для электрона в яме будет увеличиваться при уменьшении длины a (см. рис. 1.2). Наоборот, увеличение размера a в пределе приведет к тому, что энергия электрона будет иметь сплошной (непрерывный) спектр, как показано на рис. 1.1.

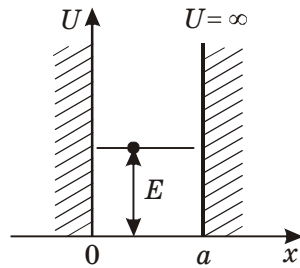


Рис. 1.2. Движение электрона в потенциальной яме

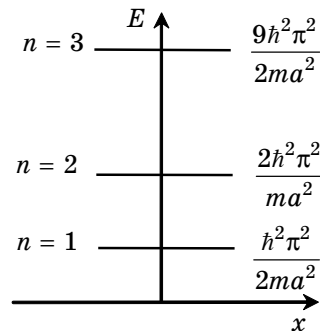


Рис. 1.3. Энергетический спектр электрона в потенциальной яме

§ 1.2. Энергия электрона в свободном атоме

Реальными примерами модели электрона в потенциальной яме являются электроны в свободных (индивидуальных) атомах. Действительно, из общих курсов химии и физики известно, что в свободных атомах электроны располагаются в потенциальной яме на энергетических уровнях, образующих дискретный спектр.

Согласно теории Бора в атоме водорода единственный электрон с отрицательным зарядом минус e вращается вокруг ядра с положительным зарядом eZ , где Z — порядковый номер элемента (в данном случае $Z = 1$).

Полная энергия электрона в атоме E равна сумме кинетической W и потенциальной U энергий:

$$E = W + U. \quad (1.13)$$

Потенциальная энергия представляет энергию взаимодействия электрона с ядром при вращении вокруг него по орбите с радиусом r . Сила электростатического (кулоновского) притяжения электрона к ядру равна

$$(1.14)$$

где $\epsilon_0 = 8,85 \cdot 10^{-12}$ Ф/м — абсолютная диэлектрическая проницаемость вакуума (электрическая постоянная).

В этом случае потенциальная энергия электрона в атоме будет определяться формулой

$$U = \int F dr = \int \frac{q^2 Z}{4\pi\epsilon_0 r^2} dr = -\frac{q^2 Z}{4\pi\epsilon_0 r}. \quad (1.15)$$

Тогда на основании выражений (1.13) и (1.15) можно считать, что полная энергия электрона в атоме является отрицательной величиной.

Решение уравнения Шредингера (1.4) при подстановке в него (1.15) дает следующее выражение для энергии электрона в атоме водорода:

$$E = -\frac{q^4 Z^2 m}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} = -\frac{13,6}{n^2}, \quad (1.16)$$

где $n = 1, 2, 3, \dots$ — главное квантовое число.

В формуле (1.16) $E = -13,6/n^2$ энергия выражается в электронвольтах. В СИ $1 \text{ эВ} \cong 1,6 \cdot 10^{-19}$ Дж.

Состояние электрона в многоэлектронных атомах характеризуется четырьмя квантовыми числами:

- 1) главным $n = 1, 2, 3, \dots$;
- 2) орбитальным $l = 0, 1, 2, \dots, (n - 1)$;
- 3) магнитным $m = 0, \pm 1, \pm 2, \dots, \pm l$;
- 4) спиновым $s = \pm 1/2$.

Главное квантовое число n определяет порядковый номер разрешенной орбиты, или энергетического уровня, электронов.

Орбитальное квантовое число l связано с орбитальным квантовым моментом количества движения p_l электрона формулой

$$p_l = \hbar \sqrt{l(l+1)}. \quad (1.17)$$

Магнитное квантовое число m определяет проекцию момента p_l на некоторое направление, причем m принимает лишь целочисленные значения.

Спиновое квантовое число s в единицах \hbar выражает собственный механический момент количества движения электрона. Этот момент может

$$F = \frac{q^2 Z}{4\pi\epsilon_0 r^2}$$

быть направлен либо параллельно, либо антипараллельно орбитальному моменту количества движения.

Совокупность электронов, обладающих одним и тем же значением главного квантового числа, образует электронный слой атома. Слои, соответствующие значениям n от 1 до 4, обозначаются соответственно K , L , M , N . Внутри слоя электроны с различными орбитальными числами l образуют оболочки. Оболочки с $l = 0, 1, 2, 3$ обозначаются соответственно s, p, d, f .

Согласно принципу запрета Паули в одном квантовом состоянии, характеризуемом четырьмя квантовыми числами, может находиться не более одного электрона. Или в одном квантовом состоянии, характеризующемся тремя квантовыми числами n, l, m , может находиться не более двух электронов с антипараллельными спинами ($s = \pm 1/2$).

Число значений m , которое изменяется в пределах от $-l$ до $+l$, составляет $2l + 1$. Следовательно, с учетом спина максимальное число электронов в оболочке l будет равно

$$N_l = 2(2l + 1), \quad (1.18)$$

т.е. в оболочке s может находиться 2 электрона, в p -оболочке — 6, в d — 10 и в f -оболочке — 18 электронов.

Максимальное число электронов в слое с соответствующим главным квантовым числом n равно

$$N_n = \sum_{l=0}^{n-1} 2(2l + 1) = 2n^2. \quad (1.19)$$

Таким образом, в K -слое ($n = 1$) имеются только 2 электрона, в L -слое ($n = 2$) — 8 электронов, в M -слое ($n = 3$) — 18 электронов и в N -слое ($n = 4$) — 32 электрона.

В изолированном атоме s -уровни являются простыми, а все остальные — вырожденными. Последнее означает, что при заданных n и l , т.е. при заданной величине E , имеется $g = 2l + 1$ значений m и им может соответствовать g электронов. Следовательно, фактор вырождения g (без учета спина) для p -уровней равен 3, так что p -уровни имеют трехкратное вырождение, а d -уровни — пятикратное.

Рассмотрим, например, энергетическую схему атома натрия Na (рис. 1.4), который содержит 11 электронов ($Z = 11$). При заполнении слоев и оболочек первые 2 электрона стремятся занять более низкий энергетический уровень $1s^2$. Третий и четвертый электроны в соответствии с принципом Паули займут следующий самый близкий к ядру уровень второго слоя $2s^2$. В следующей $2p$ -оболочке находится 6 электронов, что можно рассматривать как совпадение трех энергетических уровней, учитывая трехкратное вырождение (см. выше). Далее должны заполняться уровни с $n = 3$ и т.д.

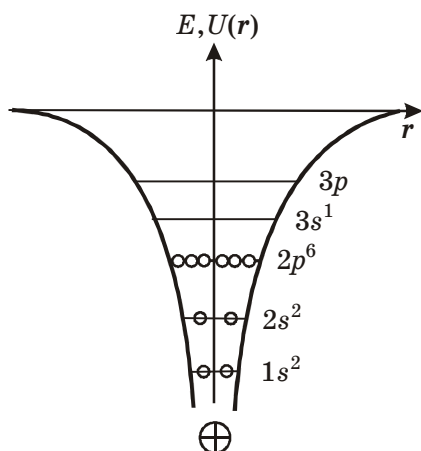
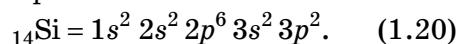
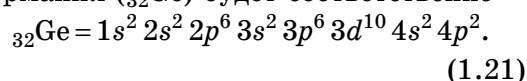


Рис. 1.4. Энергетический спектр электрона в изолированном атоме натрия

Обычно используют символическую форму записи состояния электронов в атомах. Например, атом кремния содержит 14 электронов (${}_{14}\text{Si}$), которые распределены по слоям и оболочкам следующим образом:



Распределение электронов в атоме германия (${}_{32}\text{Ge}$) будет соответственно



Нижние индексы 14 и 32 соответствуют порядковым номерам в таблице Менделеева, или числу электронов в атоме. Верхние индексы показывают число электронов в оболочках. Здесь необходимо отметить одинаковое заполнение электронами валентных оболочек кремния ($3s^2 3p^2$) и германия ($4s^2 4p^2$).

§ 1.3. Модель периодического потенциального поля в кристалле

Твердые тела, к которым относятся металлы и полупроводники, имеют обычно кристаллическое строение. Атомы в кристаллах располагаются упорядоченно и образуют кристаллическую решетку с одинаковым для каждого направления межатомным расстоянием a , величина которого менее 1 нм.

На рис. 1.5,а показана энергетическая диаграмма атомов натрия, отстоящих друг от друга на расстоянии значительно большем, чем межатомное расстояние в кристаллической решетке ($\Delta x \gg a$).

Потенциальные энергии электронов в изолированных атомах (рис. 1.5,а) и в кристалле (рис. 1.5,б) различаются. Это связано с тем, что в изолированном атоме существуют только две силы: сила кулоновского притяжения электронов к ядру и сила отталкивания между электронами.

В кристалле из-за близкого расстояния между атомами возникают дополнительные силы отталкивания и притяжения электронов и ядер всех атомов. Под влиянием этих сил происходит изменение системы энергетических уровней электронов в каждом из атомов кристалла. Энергия одних уровней уменьшается, других — возрастает. Это наблюдается при сближении атомов на расстояние, при котором возникает их сильное взаимодействие. В этом заключается первое следствие сближения атомов.

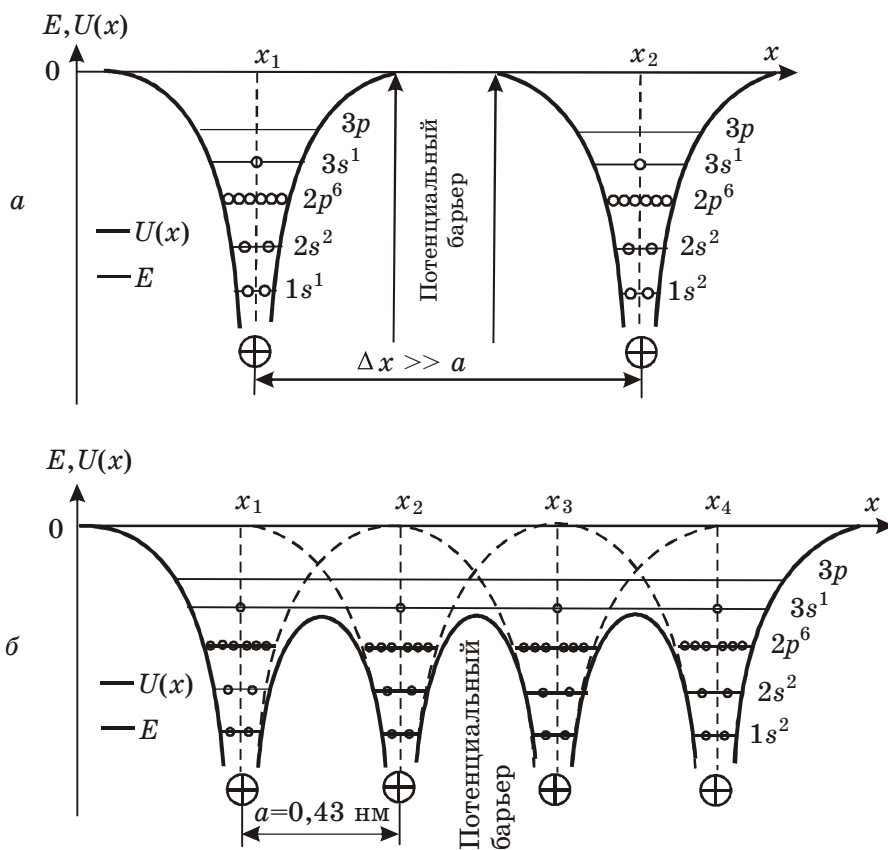


Рис. 1.5. Потенциальные барьеры в индивидуальных атомах (а) и в атомах кристалла с межатомным расстоянием $a = 0,43 \text{ нм}$ (б)

Второе следствие связано с тем, что валентные электронные оболочки соседних атомов могут не только соприкасаться, но и перекрываться. На рис. 1.5,б показана картина, соответствующая такому сближению атомов. Видно, что потенциальные штриховые кривые $U(x)$, разделяющие соседние атомы, частично перекрываются и дают результирующую (сплошную) потенциальную кривую, проходящую ниже уровня нулевой энергии. Это означает, что сближение атомов вызывает уменьшение как ширины, так и высоты потенциального барьера для электронов каждого атома в кристалле. Как видно из рис. 1.5,б, для валентных электронов потенциальный барьер в межатомном промежутке уменьшается и может исчезать. В результате валентные электроны получают возможность свободно перемещаться от одного атома к другому, чему соответствует состояние их полного обобществления. В то же время энергетическое состояние электронов на внутренних оболочках ($1s$, $2s$, $2p$) атомов остается практически

неизменным. В электрическом поле свободные электроны будут увлекаться в одном направлении, что приводит к возникновению электрического тока.

Таким образом, с помощью основных представлений физики можно создать простую картину возникновения в твердом теле свободных электронов, ответственных за электропроводность металлов. Однако рис. 1.5 является упрощенным. Валентные электроны всех атомов на рис. 1.5,б обладают одинаковой энергией. Все валентные электроны обобществлены и принадлежат кристаллу, который необходимо рассматривать как одну систему. Однако считать, что все электроны имеют одинаковую энергию в одной системе, является нарушением принципа Паули.

§ 1.4. Зонный характер энергетического спектра электронов в кристалле

Образование кристалла сопровождается не только снижением потенциального барьера для электронов, но и расщеплением их энергетических уровней на энергетические зоны (рис. 1.6).

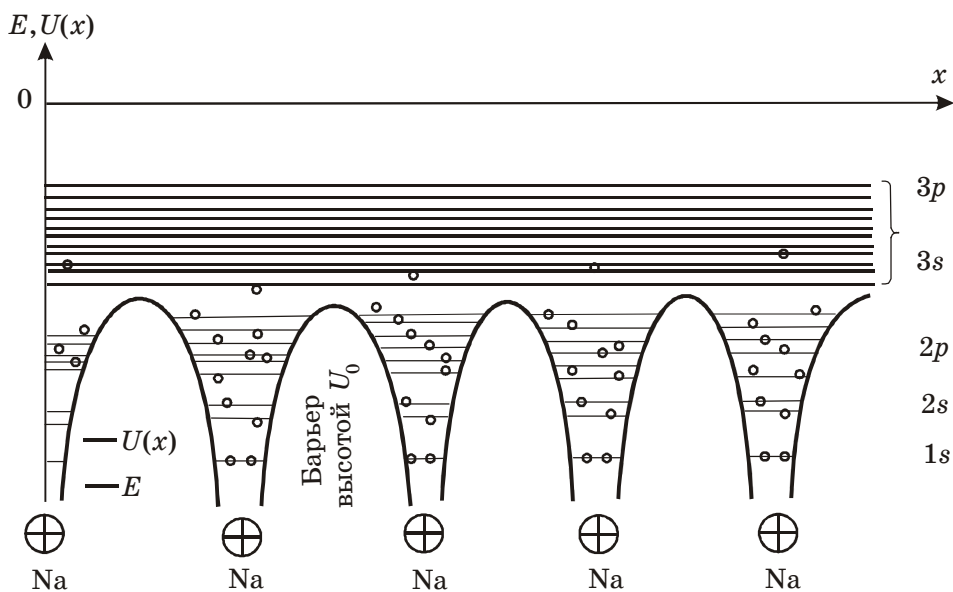


Рис. 1.6. Энергетический спектр электронов в атомах натрия, стоящих друг от друга на расстоянии, равном постоянной решетки

Каждая зона состоит из большого числа очень близко расположенных энергетических подуровней. В каждой зоне количество подуровней равно числу атомов в кристалле.

Каждый энергетический уровень электронов в отдельном (индивидуальном) атоме расщепляется в зону. Поэтому количество зон равно числу электронных уровней в атоме, что автоматически приводит к выполнению принципа Паули.

Все энергетические подуровни в зонах, расположенных ниже барьера высотой U_0 , заняты электронами, вследствие чего такие зоны называются заполненными. В кристалле Na ширина зон, образовавшихся из уровней оболочек $3s$ и $3p$ (см. рис. 1.6) такая, что самая верхняя частично заполненная зона $3s$ и пустая зона $3p$ перекрывают друг друга.

Полное объяснение механизма расщепления дискретных энергетических уровней электронов в атомах можно найти в квантовой физике. Простая модель Кронига — Пенни будет рассмотрена в § 1.5.

Реально с физической точки зрения в уравнении Шредингера (1.4) необходимо учесть взаимодействия между всеми электронами и ядрами в единичном объеме кристалла. Строгое решение этого уравнения, в котором примерно 10^{28} переменных, на современном техническом уровне ЭВМ получить невозможно.

Несколько допустимых приближений для электронов и ядер в кристалле упрощает решение уравнения Шредингера и показывает сложный характер расщепления энергетических уровней электронов в атомах при образовании кристалла. При сближении атомов начинает сказываться влияние на каждый атом электрических полей соседних атомов, что учитывается в уравнении Шредингера (1.4) поправкой потенциальной энергии $\Delta U(\mathbf{r})$. Это позволяет, не приводя здесь самого решения, указать качественные закономерности, вытекающие из него.

При сближении атомов происходит, как схематически показано на рис. 1.7, расщепление энергетических уровней электронов на $2N(2l + 1)$ близко расположенных подуровней, образующих разрешенную энергетическую зону. Здесь N — число атомов, l — орбитальное квантовое число (см. § 1.2). Уровню $1s$ в индивидуальном атоме ($x \gg a$) соответствует зона $1s$ в кристалле ($x = a_0$) (рис. 1.7, б), уровню $2s$ в атоме — зона $2s$ в кристалле и т.д. Расщепление (расстояние между подуровнями) тем больше, чем ближе сдвигаются атомы.

Самое сильное взаимодействие испытывают электроны наружной оболочки. Поэтому их энергетические уровни расщепляются более сильно, и ширина образующейся зоны в этом случае максимальна.

Взаимодействие между электронами внутренних орбит соседних атомов вследствие экранировки электронами внешних орбит значительно слабее. Следовательно, расщепление внутренних уровней будет слабым, а ширина энергетических разрешенных зон, соответствующих им, — меньше.

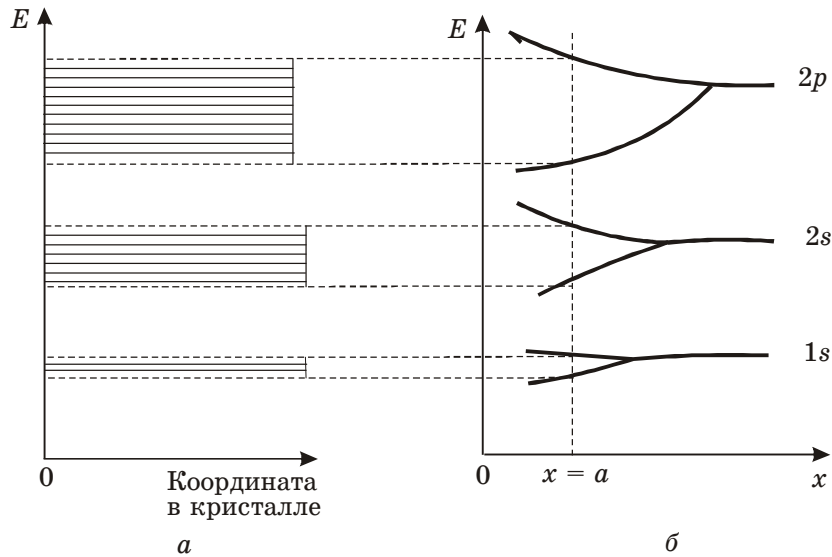


Рис. 1.7. Энергетические зоны в кристалле (а) и расщепление энергетических уровней электронов при сближении атомов (б)

На рис.1.7,б (справа) при больших расстояниях между атомами энергетические уровни дискретные, как в изолированных атомах. При сближении атомов верхние энергетические уровни начинают расщепляться в первую очередь, образуя зоны. При дальнейшем сближении расщепление верхнего $2p$ -уровня усиливается и начинается расщепление $2s$ -уровня. К тому моменту, когда атомы уплотняются до характерного для кристалла межатомного расстояния a , энергия электрона в кристалле может принимать только свободные разрешенные значения (в разрешенных зонах). Незаштрихованные зоны (рис. 1.7,а) соответствуют значениям энергии, которые электрон не может принимать. Таким образом, при объединении атомов в кристаллическую решетку твердого тела разрешенные уровни электронов изолированного атома превращаются в зоны разрешенных значений энергий, между которыми расположены зоны запрещенных значений энергий (запрещенные зоны).

§ 1.5. Модель Кронига — Пенни

В модели Кронига — Пенни используется основное свойство потенциального рельефа энергии электрона в кристалле — периодичность функции $U(x)$ (§ 1.3). Движение электрона рассматривается в цепочке прямоугольных потенциальных ям шириной a , отделенных друг от друга прямоугольными потенциальными барьерами высотой U_0 и шириной b

(рис. 1.8). Период повторения ям, равный $a + b$, представляет собой постоянную кристаллической решетки. Волновая функция, описывающая поведение электрона, является решением уравнения Шредингера (1.4). При движении электрона с энергией E вдоль оси x в цепочке потенциальных барьеров $U(x)$ (см. рис. 1.8) уравнение Шредингера будет иметь вид

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{8\pi^2 m}{h^2} [E - U(x)] \psi = 0. \quad (1.22)$$

Анализ модели Кронига — Пенни (см. рис. 1.8) следует провести для условий:

$$\begin{aligned} U(x) &= 0 \text{ при } 0 \leq x \leq a; \\ U(x) &= U_0 \text{ при } a \leq x \leq a + b. \end{aligned} \quad (1.23)$$

Решением уравнения (1.22) является функция Блоха

$$(1.24)$$

которая представляет плоскую волну с волновым числом k , промодулированную по амплитуде с помощью функции $U(x)$, период которой совпадает с периодом кристаллической решетки твердого тела.

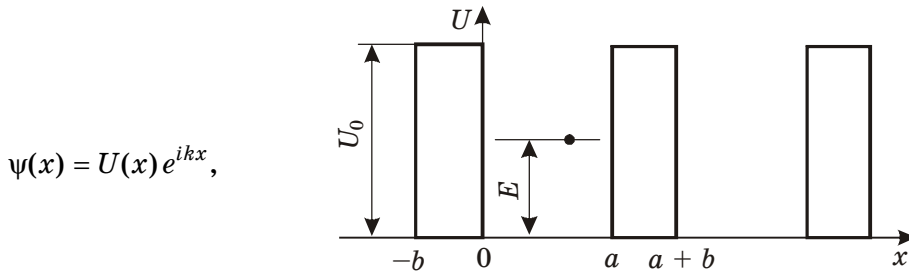


Рис. 1.8. Потенциальный рельеф в модели Кронига — Пенни

Дважды дифференцируя выражение (1.24) и подставляя в (1.22) с учетом условий (1.23), можно получить уравнение

$$\cos(\alpha a) + \frac{P}{\alpha a} \sin(\alpha a) = \cos(ka), \quad (1.25)$$

где

$$P = \lim_{\substack{b \rightarrow 0 \\ \beta \rightarrow \infty}} \frac{\beta^2 ab}{2}, \quad (1.26)$$

$$\alpha = \frac{2\pi}{h} \sqrt{2mE}, \quad (1.27)$$

$$\beta = \frac{2\pi}{h} \sqrt{2m(U_0 - E)}. \quad (1.28)$$

Трансцендентное уравнение (1.25) с учетом (1.26)–(1.28) связывает энергию электрона E с волновым числом k и постоянной кристаллической решетки a , т.е. определяет в неявном виде зависимость $E(k, a)$. Строгое решение этого уравнения возможно только в некоторых случаях: а) абсолютно свободный электрон; б) абсолютно связанный электрон; в) сильно связанный электрон.

Для случая «а» можно считать справедливым равенство $E = U_0$ и, подставляя $P = 0$ в (1.25) с учетом (1.27), показать, что энергия электрона описывается уже известной формулой (1.9). Условия $U = \infty$ и $P = \infty$ в соответствии с (1.26)–(1.28) характеризуют абсолютно связанный электрон, т.е. электрон в глубокой потенциальной яме (§ 1.1). Подставляя вышеуказанные условия в уравнение (1.25), нетрудно получить уже известную формулу (1.12).

Случай «в» является промежуточным между очень слабой ($P = 0$) и очень сильной ($P = \infty$) связями (случаи «а» и «б») электрона с атомом. Предполагается, что $P \gg 1$, т.е. величина P велика, но не обращается в бесконечность. Тогда уравнение (1.25) принимает вид выражения для энергии сильно связанного электрона:

$$E_n(k) = A_n + (-1)^n B_n \cos(ka), \quad (1.29)$$

где A_n и B_n — коэффициенты, зависящие от дискретного числа n .

Периодичности волновой функции для кристалла длиной L

$$\psi(x) = \psi(x + L) \quad (1.30)$$

соответствует периодичность волнового числа

$$k = \frac{2\pi}{L} j, \quad (1.31)$$

где $j = 0, \pm 1, \pm 2, \dots, \pm \frac{L}{2a}$.

В этом случае уравнение (1.29) принимает окончательный вид

$$E_n(k) = A_n + (-1)^n B_n \cos\left(\frac{2\pi}{L} ja\right), \quad (1.32)$$

из которого следует, что энергия электронов в периодическом поле кристалла дискретно зависит от двух чисел — n и j :

$$\begin{aligned} n &= 1, 2, 3, \dots; \\ j &= 0, \pm 1, \pm 2, \dots \end{aligned}$$

Анализ выражения (1.32) дает представление об энергетической структуре электронов в кристалле твердого тела.

При минимальных значениях $n = 1$ и $j = 0$

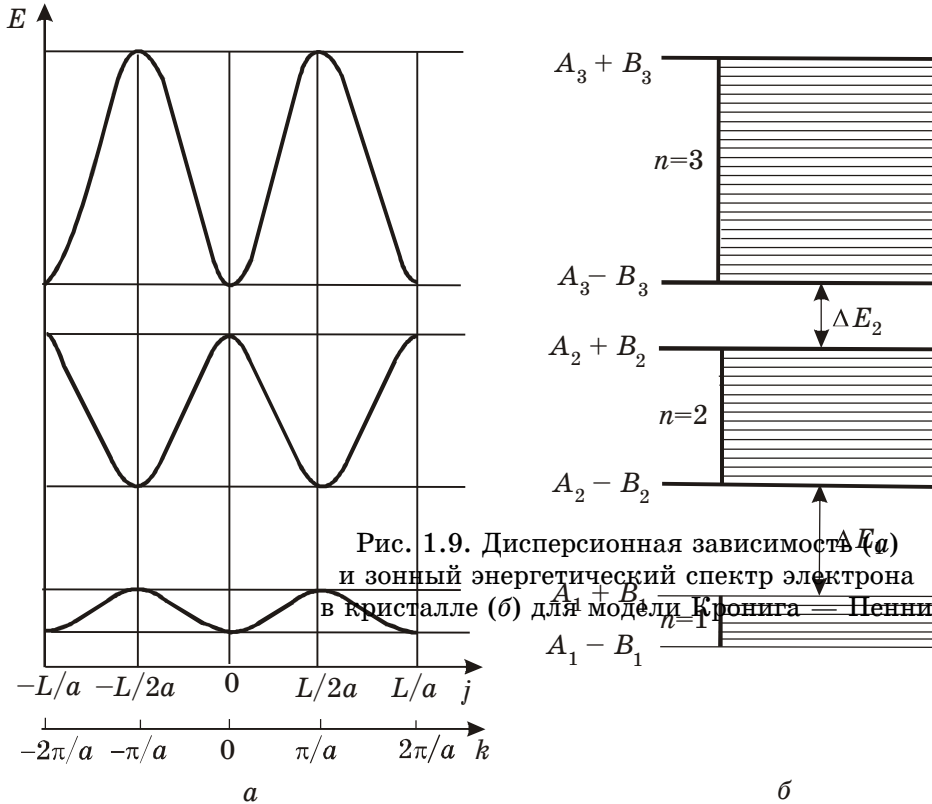
$$(1.33)$$

При $n = 1$ и $j_{\max} = L/2a$, что соответствует количеству периодов, т.е. атомов, в решетке вдоль оси x ,

$$\cos\left(\frac{2\pi}{L} ja\right) = -1, \quad (1.34)$$

$$E_1\left(\frac{L}{2a}\right) = A_1 + B_1. \quad (1.35)$$

Из (1.33) и (1.35) следует, что при $j_{\min} = 0$ $E_1 = E_{1\min}$, а при j_{\max} $E_1 = E_{1\max}$. Физически это означает, что при $n = 1$ (аналог главного квантового числа электрона в индивидуальном атоме) энергия электрона в кристалле может принимать полосу значений от $A_1 - B_1$ до $A_1 + B_1$, в которой дискретным значениям $j = 0, \pm 1, \pm 2, \dots, \pm L/2a$ соответствуют уровни энергии электрона, изменяющиеся пропорционально $\cos\left(\frac{2\pi}{L} ja\right)$ (рис. 1.9, а).



Для $n = 2$ энергия электронов в кристалле согласно уравнению (1.32) может быть выражена формулой

$$E_2(k) = A_2 + B_2 \cos\left(\frac{2\pi}{L} ja\right). \quad (1.36)$$

Если $j = j_{\min} = 0$,

$$E_2(0) = A_2 + B_2. \quad (1.37)$$

При $j = j_{\max} =$

$$(1.38)$$

Величина $E_2(0)$ соответствует максимальному значению энергии в полосе $n = 2$, величина $E_2(L/2a)$ — минимальному.

Для промежуточных значений j от 0 до $L/2a$ и периодически далее (см. рис. 1.9,а) энергия изменяется так же, как в полосе $n = 1$, пропорционально $\cos\left(\frac{2\pi}{L} ja\right)$.

В полосах $n = 3$ и $n = 4$ энергия электронов изменяется в зависимости от j (или k) аналогично полосам $n = 1$ и $n = 2$ соответственно.

Полосы разрешенных энергий электронов в кристалле называются разрешенными зонами (см. § 1.4). Количество разрешенных энергетических зон для электронов в кристалле равно числу оболочек в электронном слое с соответствующим главным квантовым числом n (см. § 1.2). Следовательно, энергетический спектр электронов в кристалле имеет зонный характер (§ 1.4), т.е. энергетические уровни атомов расщепляются в энергетические зоны (см. рис. 1.7). Ширина разрешенной зоны увеличивается с ростом n по квадратичному закону. Действительно, ранее, в § 1.5, можно было показать, что ширина разрешенной зоны (с индексом n) равна

$$(1.39)$$

и быстро увеличивается при удалении электронного слоя в индивидуальном атоме от ядра.

Разрешенные энергетические зоны для электронов в кристалле отделены одна от другой запрещенными зонами $\Delta E_1, \Delta E_2$ (рис. 1.9,б). Ширина запрещенной зоны уменьшается при удалении электронного слоя в атоме от ядра. Далее будет показано, что разрешенные и запрещенные зоны в твердых телах можно наблюдать и оценивать экспериментально.

$$\frac{L}{2a} \left(\frac{L}{2a} \right) = E_n^2$$

§ 1.6. Проводник, полупроводник, диэлектрик в свете зонной теории твердого тела

Энергетические уровни электронов в изолированном атоме расщепляются с учетом вырождения (см. § 1.2) при образовании кристалла на $2N(2l + 1)$ подуровней в каждой разрешенной зоне (§ 1.4, рис. 1.7,б). Если единичный объем кристалла (1 см^3) образуется из N атомов, то каждая разрешенная энергетическая зона будет содержать $2N(2l + 1)$ энергетических подуровней (см. § 1.4), которые отстоят друг от друга приблизительно на 10^{-22} эВ.

В природе выделяются три типа твердых тел: металлы, полупроводники, диэлектрики. Различия их электрических свойств в физике твердого тела объясняются разным заполнением электронами разрешенных энергетических зон.

Известно, что натрий обладает металлическими свойствами. Уровень $3s$ в атоме натрия содержит 1 электрон (см. рис. 1.4), вследствие чего зона $3s$ в кристалле натрия, содержащая $2N$ состояний (N — число атомов в кристалле натрия (см. рис. 1.5)), была бы заполнена только наполовину. Перекрывание зон $3s$ ($2N$ состояний) и $3p$ ($6N$ состояний) при образовании кристалла (см. рис. 1.6) увеличивает число разрешенных к заполнению состояний в верхней разрешенной зоне до $8N$ при наличии только N электронов, которые в соответствии с принципом Паули занимают нижние энергетические подуровни в этой зоне (рис. 1.10,а). Остальные энергетические подуровни верхней зоны остаются свободными (незаштрихованная область), т.е. способны принимать соответствующее количество электронов.

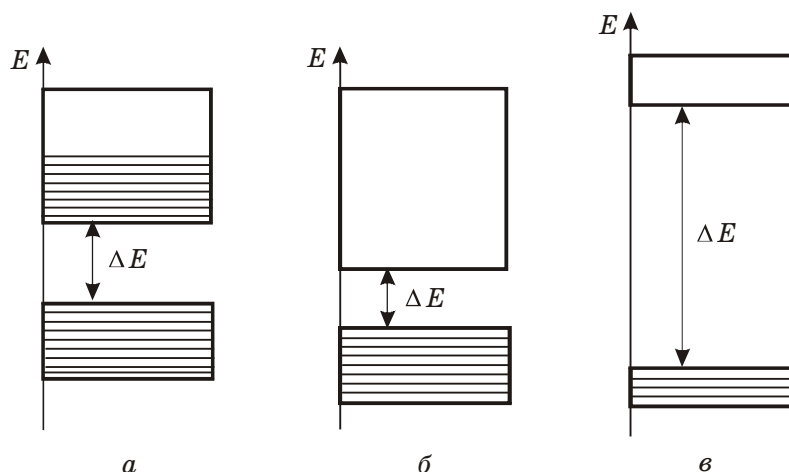


Рис. 1.10. Заполнение электронами верхних энергетических зон: в металлах (а), полупроводниках (б) и диэлектриках (в)

Таким образом, в металлах верхняя из зон, содержащих электроны, заполнена ими лишь частично. В ней имеются незанятые энергетические подуровни, и это создает возможность для изменения энергии электронов.

В электрическом поле с напряженностью E электрон, пройдя путь x , должен получить энергию

$$(1.40)$$

то есть совершить работу

$$(1.41)$$

Ускорение электрона в электрическом поле неизбежно связано с изменением (накоплением) его энергии и, следовательно, с переходом на более высокий свободный энергетический подуровень, который имеется в данной разрешенной зоне. При обычной разности потенциалов энергия, накопленная электроном в электрическом поле на длине свободного пробега, достигает значений выше 10^{-8} эВ, что достаточно для переходов по свободным энергетическим подуровням в верхней частично заполненной разрешенной зоне, которую принято называть зоной проводимости. Этот процесс связан с упорядоченным движением электронов по кристаллу, т.е. электрическим током. Отсюда следует, что твердые тела с частично заполненной электронами зоной проводимости обладают электропроводностью даже при температуре абсолютного нуля ($T = 0$ К), что характерно для металлов.

Иная ситуация в полупроводниках и диэлектриках. В этих материалах при $T = 0$ К валентные электроны полностью заполняют соответствующую разрешенную зону, которая называется валентной. Отсутствие свободных состояний в валентной зоне приводит к тому, что электроны этой зоны не могут менять свою энергию под действием электрического поля.

В разрешенной зоне, следующей за валентной и называемой зоной проводимости, при $T = 0$ К электронов нет, поскольку ширина запрещенной зоны ΔE обычно значительно больше, чем энергия, которую может получить электрон в электрическом поле на длине свободного пробега l :

$$(1.42)$$

Такое заполнение двух самых верхних разрешенных зон (зоны проводимости и валентной зоны) объясняет факт полного отсутствия электропроводности полупроводника и диэлектрика при температуре абсолютного нуля.

Условием возникновения электропроводности полупроводников является наличие свободных электронов в зоне проводимости или свободных энергетических уровней в валентной зоне. При нагревании полупроводника часть электронов валентной зоны получает тепловую энергию, величина которой больше запрещенной зоны. Электроны в этом случае перебрасываются на нижние уровни зоны проводимости и получают возможность для дальнейшего увеличения своей энергии в электрическом

~~$\Delta E = \hbar \omega$~~ $E \cdot l$.

поле, то есть теперь они могут перемещаться и участвовать в электрическом токе.

Таким образом, различие между металлами и полупроводниками по электрофизическим свойствам заключается в следующем:

1) металлы при $T = 0$ К хорошо проводят электрический ток, т.е. обладают высокой электропроводностью;

2) полупроводники электрический ток при температуре абсолютного нуля не пропускают, т.е. имеют бесконечно высокое электрическое сопротивление.

Энергетическая зонная диаграмма кристаллических диэлектриков (рис. 1.10,а) по виду подобна диаграмме для полупроводников (рис. 1.10,б). Энергетические зонные диаграммы электронов в диэлектриках отличаются большой шириной запрещенной зоны. В диэлектриках обычно $\Delta E \geq 3$ эВ, в распространенных (типичных) полупроводниках $\Delta E \approx 1$ эВ. К типичным кристаллическим диэлектрикам относятся углерод ($\Delta E = 5,3$ эВ), нитрид бора ($\Delta E = 4,6$ эВ) и окись алюминия ($\Delta E = 7$ эВ).

Вследствие широкой запрещенной зоны в диэлектриках при $T \approx 300$ К переход электронов из валентной зоны в зону проводимости затруднен, количество свободных электронов в зоне проводимости и свободных уровней в валентной зоне чрезвычайно мало. Поэтому твердые диэлектрики в обычных условиях плохо проводят электрический ток.

В дальнейшем можно рассматривать только две верхние разрешенные зоны (см. рис. 1.10). Более низкие зоны полностью заняты электронами. Отсутствие в них свободных квантовых состояний объясняет невозможность изменения энергии электронов.

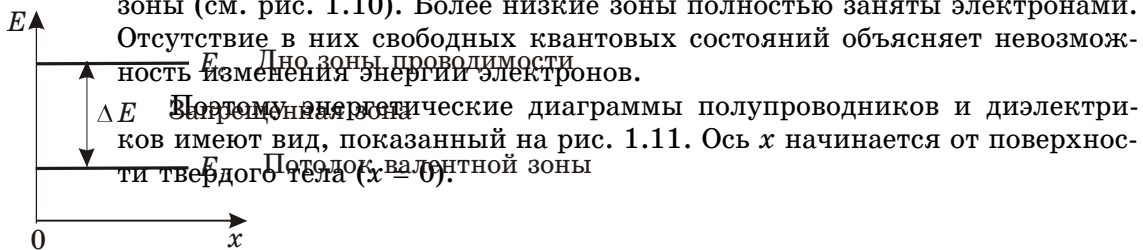


Рис. 1.11. Зонная энергетическая диаграмма для полупроводников и диэлектриков

§ 1.7. Эффективная масса носителей заряда

Перенос электрических зарядов в полупроводнике под действием внешнего электрического поля возможен только после переброса части электронов из валентной зоны в зону проводимости. При наличии свободных квантовых состояний (свободных энергетических уровней) в зоне проводимости появившиеся здесь электроны имеют возможность увеличивать свою энергию, т.е. перемещаться под действием внешнего электрического поля против вектора его напряженности \mathbf{E} со скоростью v_n (рис. 1.12). С другой стороны, в валентной зоне появляются свободные квантовые состояния и электроны этой зоны получают возможность увеличивать свою энергию, т.е. участвовать в переносе электрического заряда по валентной зоне.

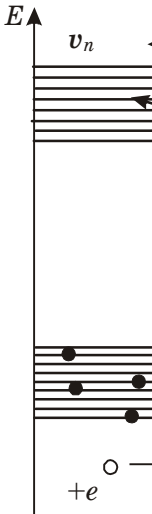
Таким образом, механизм переноса через полупроводник электрических зарядов обусловлен одновременным перемещением электронов как в зоне проводимости, так и в валентной зоне. Такое перемещение должно отличаться от движения абсолютно свободных электронов, так как в кристалле на электроны дей-

Рис 1.12. Возникновение носителей заряда в полупроводнике

ствует еще и периодическое потенциальное поле $E(k)$ (§§ 1.3–1.5). Вследствие этого, как показано в § 1.5, энергия электрона в кристалле является периодической функцией волнового числа k с периодом $2\pi/a$ (см. рис. 1.9,а). Это позволяет ограничить анализируемую область изменения числа k интервалом величиной $2\pi/a$ (обычно от $-\pi/a$ до $+\pi/a$), именуемым первой зоной Бриллюэна. Отсюда следует, что для качественного рассмотрения изменений энергии электронов в рамках модели Кронига — Пенни можно ограничиться интервалом $-\pi/a \leq k \leq +\pi/a$. В одномерном случае на краях и в центре разрешенных зон зависимость $E(k)$ проходит через экстремальные значения. На рис. 1.13,а изображена функция $E(k)$ в зоне Бриллюэна для зоны проводимости и валентной зоны и соответствующая зонная диаграмма (рис. 1.13,б).

В § 1.1 показано, что движение абсолютно свободного электрона может описываться с позиций как квантовой, так и классической физики (см. выражение (1.9)). Следовательно, движение электрона можно описать, пользуясь вторым законом Ньютона:

$$F = ma', \quad (1.43)$$



где a' — ускорение электрона под действием силы F , которая не зависит от координаты, т.е.

$$F(x) = \text{const.}$$

Потенциальное поле кристаллической решетки изменяет движение электрона в кристалле. В этом случае зависимости $E(k)$ для электрона в кристалле (рис. 1.13, *a*) и для абсолютно свободного электрона (см. рис. 1.1) существенно отличаются.

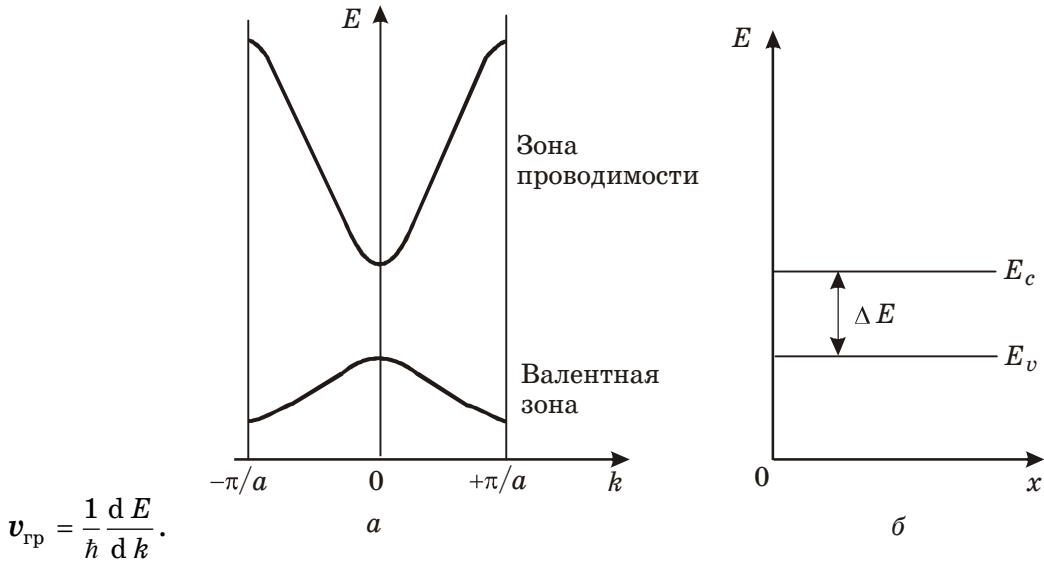


Рис. 1.13. Дисперсионные зависимости в первой зоне Бриллюэна для зоны проводимости и валентной зоны (*a*) и соответствующая зонная диаграмма для модели Кронига — Пенни (*б*)

Движение электрона в кристалле подчиняется законам квантовой физики и отождествляется с группой волн (волновым пакетом), распространяющихся во времени и в пространстве. Групповая скорость движения волнового пакета равна

$$v_{\text{гр}} = d\omega/dk. \quad (1.44)$$

С учетом выражения (1.1) нетрудно получить

$$(1.45)$$

Усредненное ускорение электрона в кристалле равно

$$a' = \frac{dv_{\text{гр}}}{dt} = \frac{1}{\hbar} \frac{d}{dt} \left(\frac{dE}{dk} \right) = \frac{1}{\hbar} \frac{d}{dk} \left(\frac{dE}{dt} \right). \quad (1.46)$$

Изменение энергии электрона dE , равное совершенной работе dA за время dt , будет

$$dE = dA = \mathbf{F} v_{\text{гр}} dt. \quad (1.47)$$

Из (1.47) следует, что

$$\frac{dE}{dt} = Fv_{\text{гр}}. \quad (1.48)$$

Объединяя равенства (1.46) и (1.48), получим выражение для ускорения электрона в направлении действия силы F :

$$(1.49)$$

Определив из (1.45)

$$\frac{dv_{\text{гр}}}{dk} = \frac{1}{\hbar} \frac{d^2 E}{dk^2} \quad (1.50)$$

и подставив в (1.49), получаем

$$a' = \frac{F}{\hbar} \frac{d^2 E}{dk^2}. \quad (1.51)$$

Сравнивая выражение (1.51) с уравнением Ньютона

$$a' = \frac{F}{m}, \quad (1.52)$$

видим, что роль массы электрона в кристалле играет величина

$$m^* = \frac{\hbar^2}{\frac{d^2 E}{dk^2}}, \quad (1.53)$$

которая называется эффективной массой носителя заряда.

$$\frac{d^2 E}{dk^2} \rightarrow F \cdot \frac{dv}{d}$$

§ 1.8. Физический смысл эффективной массы

Смысл введения эффективной массы для описания движения электрона в кристалле заключается в том, что с ее помощью учитывается совместное действие на электрон периодического потенциального поля атомов в кристалле и внешней силы. Эффективная масса электрона позволяет сложные законы движения электронов в кристалле свести к законам, которые по форме совпадают с известными законами классической механики.

В выражение для эффективной массы (1.53) входит вторая производная $d^2 E / dk^2$. Энергетический спектр электронов в кристалле в зоне Бриллюэна показан на рис. 1.13,а. Рассмотрим изменение энергии электрона в зоне проводимости для трех областей волнового числа:

1) при $k \approx 0$ наклон зависимости $E(k)$ увеличивается с ростом k . Зна-

чит, $\frac{dE}{dk} > 0$ и

Из формулы (1.53) следует, что эффективная

масса электрона у дна зоны проводимости ($E \approx E_c$) имеет положительный знак;

2) для электронов на энергетических уровнях в средней части зоны проводимости, т.е. при $E \approx E_c + \frac{1}{2} \Delta E$ а вторая производная

$\frac{d^2 E}{dk^2} = 0$. Из (1.53) для этого случая следует, что $m^* = \infty$;

3) рост энергии электронов на уровнях у потолка зоны проводимости ($k \approx \pm\pi/a$) замедляется, что свидетельствует об отрицательном знаке второй производной $d^2 E / dk^2$. Значит, эффективная масса электронов на верхних уровнях зоны проводимости отрицательна.

Из представленного выше анализа следует, что эффективная масса электрона не имеет физического смысла гравитационной массы, так как ее знак изменяется от положительного у дна до отрицательного у потолка разрешенной зоны. На средних уровнях разрешенной зоны

Проведенный выше анализ зависимости $E(k)$ для четной разрешенной зоны, т.е. зоны с четным квантовым числом n (см. рис. 1.9), приводит к тем же выводам об изменении знака эффективной массы и в других разрешенных зонах.

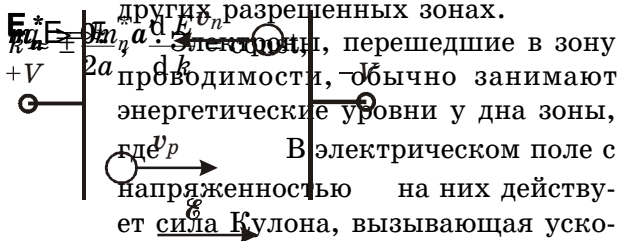


Рис. 1.14. Направления движения электронов и дырок во внешнем электрическом поле

Электрон, перешедший в зону проводимости, обычно занимает энергетические уровни у дна зоны, где $\frac{d^2 E}{dk^2} > 0$. В электрическом поле с напряженностью E на них действует сила Кулона, вызывающая ускоренное движение зарядов (электронов) против вектора напряженности электрического поля (рис. 1.14). В этом случае справедливо равенство

$$(1.54)$$

Таким образом, электроны на нижних уровнях зоны проводимости (у дна зоны) ускоряются в кристалле под действием внешней силы $F = -qE$. Их движение аналогично движению свободных частиц, имеющих эффективную массу m^* и скорость v_n . Однако сама эффективная масса не равна массе покоя электрона и может быть определена по формуле (1.53).

§ 1.9. Понятие о дырках как носителях заряда

Электроны поступают в зону проводимости с верхних уровней валентной зоны (см. рис. 1.12), расположенных у ее потолка при $k = 0$ (рис. 1.13,а). Зависимость $E(k)$ при $k = 0$ для валентной зоны по форме подобна зависимости $E(k)$ при $k = \pm\pi/a$ для зоны проводимости, где эффективная масса электрона $m_n^* < 0$. В этом случае для электронов валентной зоны выражение (1.54) будет иметь вид

$$(1.55)$$

что противоречит закону движения электронов под действием внешнего электрического поля.

Противоречие устраняется, если в валентной зоне рассматривать движение не электронов, а свободных квантовых состояний. При перемещении электрона вверх по оси E (рис. 1.12) освобождающееся квантовое состояние (энергетический уровень) передвигается вниз (вглубь валентной зоны). Это квантовое состояние отождествляется с частицей, имеющей равный по величине, но противоположный по знаку заряд $+q$ в сравнении с электроном, так как ее движение противоположно движению электрона. Тогда будет справедливым изменить знаки в уравнении (1.55) и считать, что эта частица, названная дыркой, имеет положительную эффективную массу m_p^* :

$$(1.56)$$

Под воздействием электрического поля в полупроводнике происходит перенос зарядов — электронов в зоне проводимости против вектора E (против поля) и дырок в валентной зоне по направлению поля. Дырки с меньшей энергией располагаются у потолка валентной зоны. Рост энергии дырок связан с их перемещением в глубину валентной зоны, то есть отсчет энергии дырок и электронов производится в противоположных направлениях. Эффективная масса дырки определяется формулой (1.53) и отличается величиной от эффективной массы электрона в зоне проводимости.

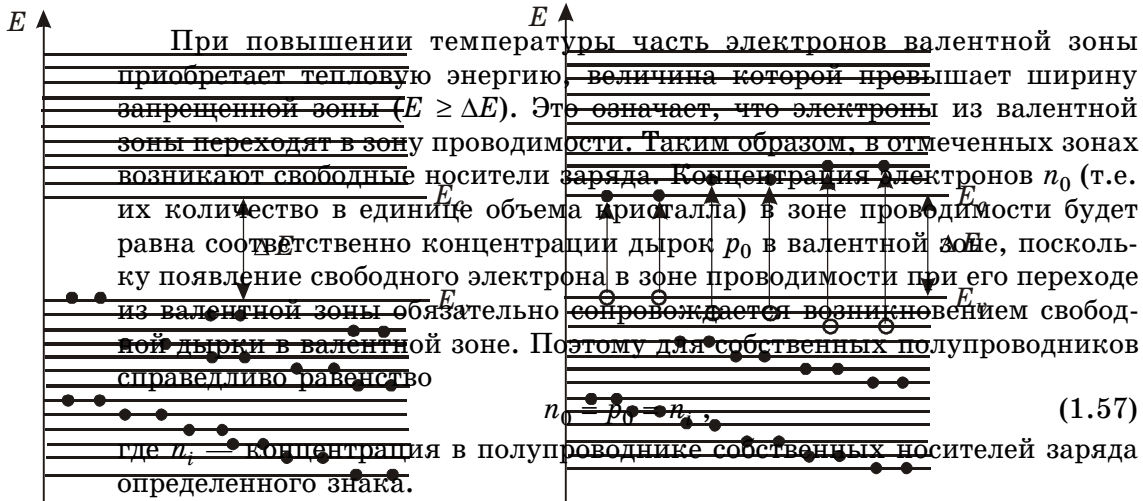
$$q\vec{E} = m_p^* \vec{a}_p^*$$

§ 1.10. Собственные полупроводники

Химически чистые полупроводники, свободные от дефектов кристаллической решетки, называются собственными. В запрещенной зоне собственных полупроводников отсутствуют какие-либо разрешенные энергетические уровни электронов (рис. 1.15,а). При температуре абсолютного нуля все энергетические уровни валентной зоны заняты электронами, а в зоне проводимости, в отличие от валентной, все энергетические уровни свободны. Свободные электроны в зоне проводимости и дырки в ва-

лентной зоне отсутствуют. Поэтому полупроводники при $T = 0 \text{ К}$ не проводят электрический ток (см. § 1.6).

Рис. 1.15. Зонные диаграммы полупроводника при $T = 0 \text{ К}$ (а) и при $T > 0 \text{ К}$ (б): • — электроны; о — дырки



Каждый полупроводник характеризуется основными физическими параметрами, в число которых входит ширина запрещенной зоны. При температуре 300 К (принята за комнатную, равную 27 °С) германий Ge имеет $\Delta E = 0,67 \text{ эВ}$, кремний Si — 1,11 эВ, арсенид галлия GaAs — 1,43 эВ. Интенсивность перехода электронов из валентной зоны в зону проводимости естественно снижается с ростом ширины запрещенной зоны

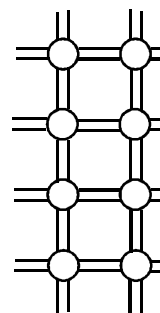
при постоянной температуре. Поэтому концентрация свободных носителей в вышеперечисленных собственных полупроводниках при $T = 300 \text{ K}$ будет равна:

- 1) в Ge $n_i = 2,5 \cdot 10^{13} \text{ см}^{-3}$;
- 2) в Si $n_i = 1,6 \cdot 10^{10} \text{ см}^{-3}$;
- 3) в GaAs $n_i = 1,1 \cdot 10^7 \text{ см}^{-3}$.

Энергетическая диаграмма рис. 1.15 соответствует идеальной периодичности кристаллического поля, что обусловлено отсутствием дефектов в кристаллической решетке собственного полупроводника. Схема двумерной кристаллической решетки, характерной для Ge и Si, показана на рис. 1.16,а.

Рис 1.16. Двумерная кристаллическая решетка полупроводника (а) и образование собственных носителей заряда при разрыве одной из ковалентных связей (б)

Каждый атом в кристаллической решетке связан ковалентными химическими связями с четырьмя ближайшими такими же атомами. В каждой ковалентной связи участвует пара электронов. Разрыв одной ковалентной связи в результате вышеупомянутого теплового воздействия приводит к освобождению одного электрона от связи (рис. 1.16,б). Процесс превращения связанного электрона в свободный электрон носит название генерации. На месте ушедшего электрона образуется незавершенная связь, которая имеет избыточный положительный заряд. Электрон, вырванный из ковалентной связи, соответствует свободному электрону в зоне проводимости, а вакантное место в ковалентной связи — дырке в валентной зоне. В целом кристалл остается электронейтральным, так как каждой образовавшейся положительно заряженной дырке соответствует свободный отрицательно заряженный электрон.



§ 1.11. Донорный полупроводник

Рассмотренная в § 1.10 зонная структура собственного полупроводника соответствует идеально правильному периодическому потенциальному полю в кристалле. Отклонения от идеальности обуславливают нарушения энергетической зонной структуры. В реальных кристаллах практически всегда имеются нарушения кристаллической решетки, называемые дефектами.

Рассмотрим случай, когда в кристалле четырехвалентного германия или кремния часть атомов в узлах кристаллической решетки замещена атомами пятивалентного мышьяка, фосфора или сурьмы. На установление химической связи с четырьмя соседними атомами германия атом мышьяка, у которого валентная оболочка содержит пять электронов, должен затратить четыре валентных электрона (рис. 1.17, *a*).

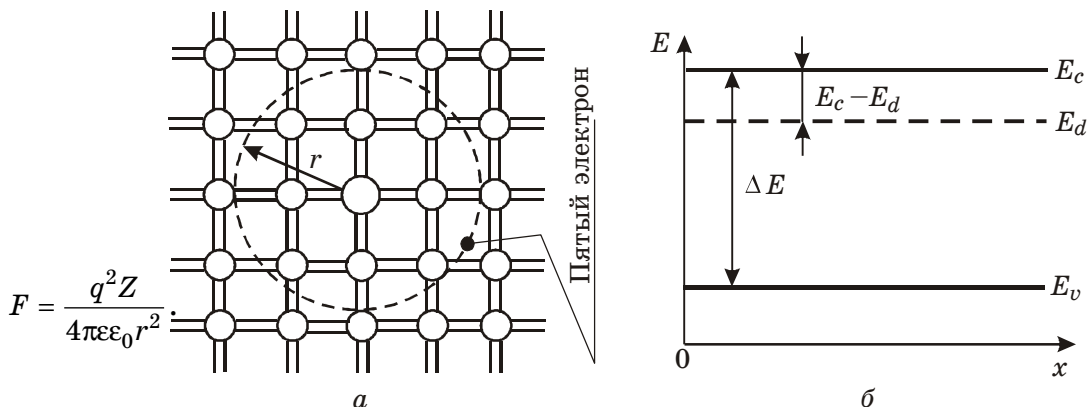


Рис 1.17. Двумерная кристаллическая решетка полупроводника с примесными атомами As (*a*) и соответствующая зонная структура с донорным уровнем E_d (*б*)

Пятый электрон, не принимая участия в образовании ковалентных связей между атомом мышьяка и атомами германия, будет взаимодействовать с большим числом атомов кристалла и в результате окажется слабо связанным с атомом мышьяка силой кулоновского взаимодействия, величина которой в ϵ раз ниже, чем при их взаимодействии в вакууме (см. выражение (1.14)):

$$(1.58)$$

В выражении (1.58) величина ϵ представляет собой диэлектрическую проницаемость германия. Поэтому пятый электрон вращается по орбите большего радиуса, чем в индивидуальном атоме мышьяка. Действительно,

радиус орбиты валентных электронов в индивидуальном атоме мышьяка $r=0,053$ нм увеличивается до $r=7$ нм (рис. 1.17,а). При таком возрастании радиуса орбиты пятый валентный электрон можно рассматривать как электрон в атоме водорода (см. § 1.2). Тогда энергия связи этого электрона с ядром атома мышьяка может быть описана формулой (1.16), как для электрона в атоме водорода, с тем отличием, что вместо диэлектрической проницаемости ϵ_0 в ней должна быть величина $\epsilon\epsilon_0$ и вместо реальной массы электрона m — его эффективная масса в зоне проводимости m_n^* .

В таком случае энергия связи E_d электрона с ионом мышьяка выражается формулой, аналогичной (1.16):

$$(1.59)$$

Если учесть, что в германии $m_n^* = 0,22m$ и $\epsilon = 16$, то получаем $E_d \approx 0,01$ эВ. Если электрону сообщить такую энергию, то он оторвется от атома мышьяка или другого элемента V группы, например фосфора, сурьмы, и приобретет способность свободно перемещаться в решетке германия.

Используя зонную теорию, этот процесс можно описать следующим образом. Энергетические уровни пятых электронов примесных атомов мышьяка (или других элементов V группы) размещаются в запрещенной зоне вблизи дна зоны проводимости, отстоя от нее на энергетическом расстоянии $E_c - E_d \approx 0,01$ эВ (рис. 1.17,б).

При сообщении электрону энергии $E_c - E_d$ он отрывается от пятиявалентного (примесного) атома и приобретает способность перемещаться в кристалле, превращаясь в свободный электрон (рис. 1.18,а), т.е. переходит в зону проводимости (рис. 1.18,б). В узлах кристаллической решетки после ухода пятых валентных электронов остаются положительно заряженные донорные ионы, которые в электрическом токе участвовать не могут вследствие химических связей с окружающими атомами германия.

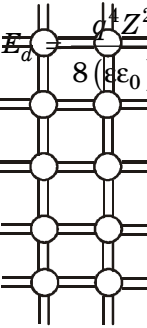


Рис 1.18. Ионизация донорного атома в двумерной кристаллической решетке полупроводника (а) и соответствующая зонная диаграмма (б)

Так как энергия ионизации примесных электронов значительно меньше энергии ионизации собственных электронов ($E_c - E_d \ll \Delta E$), то в первую очередь при нагревании кристалла в зону проводимости будут переходить электроны с донорных уровней, вследствие чего их концентрация будет во много раз превышать концентрацию собственных электронов. Таким образом, примеси, которые являются дополнительным источником электронов, называются донорами, а полупроводник с донорной примесью — донорным. Его также называют электронным полупроводником или полупроводником n -типа. Отсюда следует, что концентрация свободных носителей заряда (электронов) в полупроводнике n -типа

$$n_{n0} \gg n_i. \quad (1.60)$$

§ 1.12. Акцепторный полупроводник

Рассмотрим случай, когда атом германия в узле кристаллической решетки замещен трехвалентным атомом бора В или каким-либо другим элементом третьей группы таблицы Менделеева (In, Ga, Al). Три валентных электрона примесного атома бора идут на образование трех ковалентных связей с соседними атомами германия. Однако на укомплектование четвертой связи у атома бора нет еще одного валентного электрона (рис. 1.19,а).

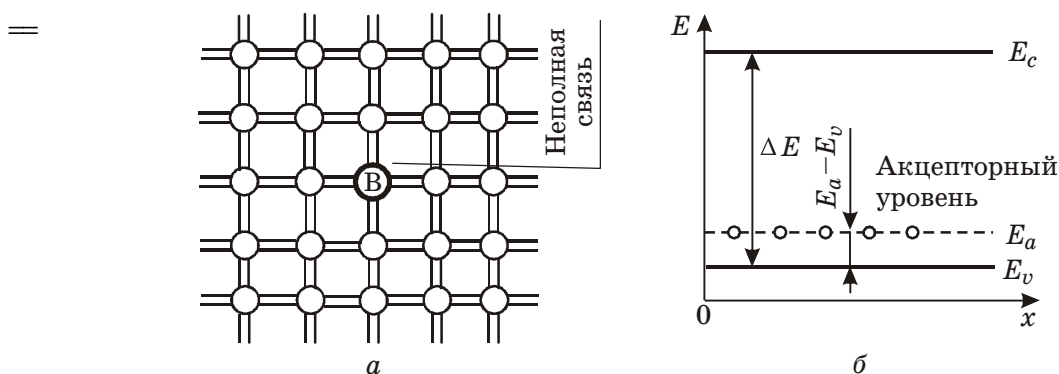


Рис 1.19. Двумерная кристаллическая решетка полупроводника, содержащего примеси атомов бора (а), и соответствующая этому случаю зонная диаграмма (б)

Создание полной ковалентной связи с четвертым атомом германия обеспечивается переходом к атому бора электрона от одной из соседних связей между атомами германия. При разрыве ковалентной связи Ge—Ge освобождается один электрон, который уходит на достройку неполной связи В—Ge. Естественно, для разрыва химической ковалентной связи между собственными атомами полупроводника требуется энергия. В собственном

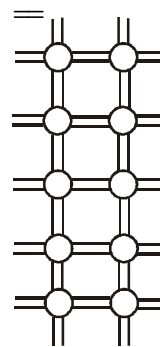
полупроводнике эта энергия равна ширине запрещенной зоны, и тогда электрон из валентной зоны переходит в зону проводимости. При наличии в кристаллической решетке германия примесных атомов III группы, в данном случае бора, для образования полной связи, то есть для перехода электрона от связи Ge—Ge к связи B—Ge, требуется энергия, величина которой значительно меньше ширины запрещенной зоны. На зонной энергетической диаграмме (рис. 1.19,б) эта энергия $E_a - E_v$ соответствует энергетическому уровню E_a , который называется акцепторным уровнем.

Для образования полной связи примесных акцепторных атомов (элементов третьей группы) с атомами германия обычно требуется энергия около 0,01 эВ. Тогда электрон из валентной зоны уходит на создание полной химической связи атомов бора и германия. При этом в валентной зоне появляется свободная дырка с положительным зарядом, а атом бора становится отрицательно заряженным ионом (рис. 1.20,а). В этом случае величина $E_a - E_v$ представляет собой энергию ионизации акцепторного атома.

Рис. 1.20. Ионизация атома акцепторной примеси в двумерной кристаллической решетке полупроводника (а) и соответствующая этому случаю зонная диаграмма (б)

Близкое расположение акцепторного уровня к валентной зоне ($E_a - E_v \approx 0,01$ эВ) приводит к тому, что при повышении температуры от 0 К электрон из валентной зоны в первую очередь переходит на акцепторный уровень. Связываясь с атомом бора (или других элементов третьей группы), он теряет возможность свободного перемещения в решетке германия. В результате образуется только свободная дырка в валентной зоне (рис. 1.20,б). Эти дырки как подвижные положительные носители заряда принимают участие в электрическом токе через полупроводник, который называется акцепторным. В акцепторном полупроводнике концентрация дырок p_{p0} значительно выше, чем в собственном, где $p_i = n_i$:

$$p_{p0} \gg p_i \cdot \quad (1.61)$$



Учитывая неравенство (1.61), рассмотренный акцепторный полупроводник часто называют дырочным полупроводником или полупроводником p -типа.

Контрольные вопросы и задачи

1. В физике твердого тела используются выражения для энергии электрона:

$$1) E = \frac{\hbar^2 \pi^2}{2ma^2} n^2;$$

$$2) E = -\frac{q^2 Z^2 m}{8\epsilon_0^2 \hbar^2} \frac{1}{n^2};$$

$$3) E = \frac{\hbar^2 k^2}{2m},$$

где $\hbar = h/2\pi$ — приведенная постоянная Планка; m — масса микрочастицы; n — квантовое число; a — ширина потенциальной ямы; q — заряд электрона; Z — заряд ядра; ϵ_0 — электрическая постоянная; h — постоянная Планка; k — волновое число.

Укажите номер формулы, справедливой для описания:

(01ЭЭ) энергии абсолютно свободного электрона;

(ПЭЭЯ) энергии электрона в прямоугольной бесконечно глубокой потенциальной яме;

(КСЭА) энергии электрона в индивидуальном атоме.

2. (В4КГ). Укажите в виде общей последовательности цифр распределение валентных электронов по оболочкам в индивидуальных атомах кремния и германия.

3. (7БЭК). Какие силы определяют потенциальную энергию электронов в кристалле?

Ответы:

1) силы притяжения электронов к ядрам;

2) силы взаимного отталкивания ядер;

3) силы притяжения электронов к ядрам своих атомов, силы взаимного отталкивания всех ядер и силы взаимного отталкивания электронов всех атомов кристалла;

4) силы взаимного отталкивания между всеми электронами.

4. (Б4УК). Что происходит с энергетическими уровнями электронов при сближении изолированных атомов и образовании кристалла? Укажите правильные ответы, для чего после кода вопроса последовательно нажмите кнопки с номерами ответов:

1) энергетические уровни электронов в изолированном атоме изменяют первоначальное положение;

2) энергетические уровни электронов всех атомов сливаются в один общий уровень;

3) валентные электронные оболочки соседних атомов перекрываются;

4) энергетические уровни валентных электронов исчезают;

5) при объединении индивидуальных атомов в кристалл дискретные энергетические уровни электронов расщепляются на энергетические зоны.

5. (МАЦД). Укажите основные выводы зонной теории об энергетическом распределении электронов в твердых телах. Энергия электронов в твердых телах может принимать значения, соответствующие:

1) дискретным энергетическим уровням электронов в атомах этого тела;

2) энергетическим уровням в разрешенных зонах;

3) энергетическим уровням в запрещенных зонах.

6. (6ДЭР). Как распределяются электроны по энергетическим зонам в полупроводнике при температуре абсолютного нуля?

Ответы:

1) энергетические уровни верхней разрешенной зоны частично заняты электронами, а в нижних зонах электроны отсутствуют;

2) все энергетические уровни разрешенных зон заняты электронами;

3) все энергетические уровни нижних разрешенных зон заполнены электронами, тогда как в последней наивысшей зоне электроны отсутствуют.

7. (2УМД). Укажите особенности металлов, полупроводников и диэлектриков при температуре абсолютного нуля с точки зрения зонной теории твердого тела.

Ответы:

1) энергетическая диаграмма металлов представляет собой одну зону проводимости, в которой есть электроны и свободные энергетические уровни;

2) на энергетических диаграммах полупроводников и диэлектриков изображаются дно зоны проводимости E_c и потолок валентной зоны E_v , между которыми находится запрещенная зона шириной ΔE ;

3) в зонах проводимости полупроводников и диэлектриков при $T = 0 \text{ К}$ свободные электроны отсутствуют;

4) в валентных зонах полупроводников и диэлектриков при $T = 0 \text{ К}$ все энергетические уровни заняты электронами;

5) энергетические диаграммы полупроводников отличаются от диаграмм диэлектриков меньшей шириной запрещенной зоны.

8. (ЗЯМП). Укажите особенности электропроводности металлов и полупроводников при температуре абсолютного нуля.

Ответы:

1) металлы и полупроводники при $T = 0 \text{ К}$ имеют бесконечно низкую электропроводность;

2) металлы обладают высокой электропроводностью, так как при $T = 0$ К в зоне проводимости есть свободные электроны и свободные энергетические уровни;

3) полупроводники при $T = 0$ К имеют бесконечно высокую электропроводность;

4) полупроводники при $T = 0$ К имеют бесконечно малую электропроводность, так как в зоне проводимости отсутствуют свободные электроны, а в валентной зоне — свободные энергетические уровни.

9. (У5ЭФ). Найдите точные определения для эффективной массы электрона.

Ответы:

1) эффективная масса электрона m_n^* равна его весу при движении в зоне проводимости;

2) эффективная масса электрона равна его весу в состоянии покоя в валентной зоне;

3) эффективная масса электрона не имеет физического смысла и представляет собой коэффициент пропорциональности в формуле $F = m^* a'$ для описания движения электрона в рамках классической механики Ньютона;

4) общее математическое выражение для эффективной массы имеет

вид

$m^* = \hbar^2 / d^2$ **10. (ЭНДР).** Найдите правильные определения дырки как носителя заряда в полупроводнике.

Ответы:

1) отсутствие электрона в зоне проводимости дает свободное квантовое состояние, которое является дыркой;

2) отсутствие атома в кристаллической решетке полупроводника представляет собой дырку;

3) свободное квантовое состояние в валентной зоне представляет собой дырку;

4) дырка — это освобождающийся энергетический уровень в валентной зоне, перемещение которого направлено встречно движению электронов, занимающих последовательно освобождающиеся квантовые состояния.

11. (И5СП). Найдите правильные определения собственного полупроводника.

Ответы:

1) это полупроводник с идеальной кристаллической структурой;

2) это полупроводник, в котором есть специально введенные примеси;

3) это полупроводник, в котором концентрации свободных электронов и дырок всегда равны;

- 4) это полупроводник, в котором нет каких-либо дефектов;
- 5) это полупроводник, в котором концентрация собственных носителей заряда преобладает над концентрациями электронов или дырок.

12. (БПДП). Выберите правильные определения. Донорный полупроводник представляет собой следующие виды материалов:

- 1) полупроводник электронного типа проводимости;
- 2) полупроводник *n*-типа;
- 3) полупроводник с примесным уровнем в верхней половине запрещенной зоны (вблизи дна зоны проводимости);
- 4) четырехвалентный полупроводник с примесными атомами V группы (Sb, As, ...);
- 5) четырехвалентный полупроводник (Ge, Si) с примесными атомами III группы (B, In, ...);
- 6) полупроводник, в котором концентрация электронов выше концентрации дырок.

13. (Л5АП). Выберите правильные определения акцепторного полупроводника. Акцепторный полупроводник представляет собой:

- 1) дырочный полупроводник;
- 2) полупроводник *p*-типа;
- 3) четырехвалентный полупроводник с примесными атомами V группы;
- 4) четырехвалентный полупроводник с примесными атомами III группы;
- 5) четырехвалентный полупроводник с примесным уровнем в запрещенной зоне вблизи потолка валентной зоны E_v ;
- 6) полупроводник, в котором концентрация дырок выше концентрации электронов.

Глава 2. Статистика электронов и дырок в полупроводниках

§ 2.1. Распределение электронов в твердых телах по энергетическим уровням

Распределение электронов по энергетическим уровням в разрешенных зонах подчиняется требованиям квантовой статистики для фермионов — электронов, протонов или других микрочастиц с полуцелым спином $1/2\hbar, 3/2\hbar, \dots$, т.е. принципу Паули. В коллективе фермионы могут занимать только свободное квантовое состояние. В этом случае встает вопрос о соотношении между числом квантовых состояний G_c и количеством микрочастиц N , которые могли бы их занять. Если число микрочастиц в коллективе сравнимо с числом квантовых состояний, по которым они могут распределяться,

$$N/G_c \approx 1, \quad (2.1)$$

то такие системы называются вырожденными.

Распределение фермионов по квантовым состояниям в вырожденном коллективе описывается функцией распределения Ферми — Дирака $f_{\text{Ф-Д}}$ (отсюда и название класса микрочастиц):

$$f_{\text{Ф-Д}}(E, T) = \frac{1}{e^{(E-E_F)/kT} + 1}. \quad (2.2)$$

Выражение (2.2) в случае электронов в твердых телах показывает вероятность $f_{\text{Ф-Д}}(E, T)$ того, что при данной температуре T электрон займет энергетический уровень E . Здесь E_F — параметр этого распределения, называемый уровнем Ферми или химическим потенциалом.

Анализируя (2.2), нетрудно видеть, что при $T = 0$ К функция $f_{\text{Ф-Д}}(E, T)$ принимает только два значения:

1) для уровней, расположенных ниже уровня Ферми ($0 < E < E_F$), $f_{\text{Ф-Д}}(E, 0) = 1$. Это означает, что все энергетические уровни, расположенные ниже уровня Ферми, заняты электронами;

2) уровни с энергией $E > E_F$, т.е. $E - E_F > 0$, характеризуются вероятностью заполнения $f_{\text{Ф-Д}}(E, 0) = 0$, что соответствует уровням, свободным от электронов.

Таким образом, при температуре абсолютного нуля вероятность заполнения электронами уровней при $E = E_F$ изменяется скачком от нуля до единицы. Это показывает, что при $T = 0$ К самый верхний энергетический уровень, занятый электронами, является уровнем Ферми (рис. 2.1). Такое распределение электронов по энергетическим уровням при $T = 0$ К соответствует заполнению электронами верхней зоны (см. рис. 1.9,а) в металлах (зоны проводимости).

Рис. 2.1. График функции распределения Ферми — Дирака

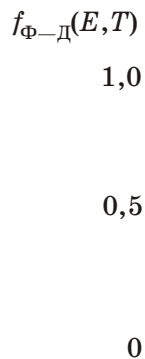
При повышении температуры наибольшее изменение в заполнении энергетических уровней наблюдается вблизи уровня Ферми. Часть электронов за счет тепловой энергии может перейти из состояний, расположенных ниже уровня Ферми ($E < E_F$), в свободные состояния, которые находятся выше него ($E > E_F$).

Расчет вероятности заполнения энергетических уровней по формуле (2.2) показывает, что при изменении энергии уровня в интервале $E_F \pm 3k_0T$, где $k_0 = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана, величина $f_{\text{Ф-Д}}(E, T)$ изменяется от 0,048 до 0,954. Из зависимости $f_{\text{Ф-Д}}(E, T) = f((E - E_F)/k_0T)$, показанной на рис. 2.1, видно, что при $T > 0$ К вероятность заполнения электронами уровней, расположенных выше уровня Ферми в области до $+3k_0T$, понижается с ростом энергии от 0,5 до 0, тогда как вероятность заполнения уровней, которые находятся ниже уровня Ферми в диапазоне до $-3k_0T$, с уменьшением энергии уровня повышается от 0,5 до 1. При этом число освобождающихся энергетических уровней, расположенных ниже E_F , равно числу занятых уровней, расположенных выше уровня Ферми, что подтверждается равенством заштрихованных площадок, ограниченных зависимостями $f_{\text{Ф-Д}}(E, T)$ для $T = 0$ К и $T > 0$ К (см. рис. 2.1).

Таким образом, вероятность заполнения энергетических уровней заметно отличается от 1 или 0 только в интервале $\pm 3k_0T$ вблизи уровня Ферми. Сам уровень Ферми можно представить как уровень, вероятность заполнения которого при температуре выше абсолютного нуля равна 0,5.

Энергетические уровни $E - E_F > 3k_0T$ имеют малую вероятность заполнения электронами. В этом случае анализ выражения (2.2) показыва-

ет, что для $E > E_F + 3k_0T$ величина $e^{\frac{E-E_F}{k_0T}} \gg 1$ и



(2.3)

где $f_{\text{М-Б}}(E, T)$ — классическая функция распределения Максвелла — Больцмана.

При $E - E_F > 3k_0T$ квантовая функция Ферми — Дирака и классическая функция Максвелла — Больцмана отличаются менее чем на 1 %.

Таким образом, для энергий $E > E_F + 3k_0T$ к электронной системе с достаточной степенью точности можно применить классическую статистику Максвелла — Больцмана. Это объясняется тем, что вследствие большого количества свободных квантовых состояний принцип Паули для электронов уже не является определяющим. Если количество частиц N значительно меньше числа квантовых состояний G_c :

$$N / G_c \ll 1, \quad (2.4)$$

то в распоряжении каждого электрона имеется множество различных свободных квантовых состояний. В этом случае электронный коллектив можно назвать невырожденным электронным газом, распределение которого по энергетическим уровням описывается функцией распределения Максвелла — Больцмана (2.3). Таким образом, невырожденный электронный газ подчиняется законам классической статистики.

§ 2.2. Плотность квантовых состояний

$$f_{\text{Ф-Д}}(E, T) = \frac{1}{e^{E-E_F}} = f_{\text{М-Б}}(E, T),$$

Важнейшая задача статистики состоит в определении числа частиц, энергия которых лежит в заданном интервале. Для ее решения необходимо знать число квантовых состояний в рассматриваемом энергетическом интервале и вероятность нахождения частиц в этих состояниях.

Пусть в кристалле единичного объема в интервале энергий от E до $E + dE$ имеется dZ квантовых состояний (с учетом спина). Обозначим через $N(E)$ плотность состояний, т.е. число состояний в единичном интервале энергии для единичного объема кристалла. Тогда согласно определению

$$N(E) = \frac{dZ}{dE}. \quad (2.5)$$

В классической механике состояние материальной частицы определяется в шестимерном пространстве координат x, y, z и импульсов p_x, p_y, p_z . Такое пространство называется фазовым, а точки x, y, z, p_x, p_y, p_z , определяющие состояние частицы, — фазовыми точками. Величина, равная

$$\Delta\Gamma = dx dy dz dp_x dp_y dp_z = \Delta\Gamma_V \Delta\Gamma_p, \quad (2.6)$$

называется элементом фазового пространства, который представляет собой произведение элементов пространства координат

$$\Delta\Gamma_V = dx dy dz \quad (2.7)$$

и пространства импульсов

$$\Delta\Gamma_p = dp_x dp_y dp_z. \quad (2.8)$$

Координаты и составляющие импульса классической частицы могут изменяться непрерывно, поэтому не существует каких-либо ограничений на минимальный размер элементов $\Delta\Gamma_V$, $\Delta\Gamma_p$ и, следовательно, $\Delta\Gamma$.

Микрочастицы (в частности, электрон) обладают волновыми свойствами. В квантовой механике известны соотношения неопределенностей Гейзенберга, которые связывают достижимые точности в определении координат микрочастицы и проекций импульсов:

$$dx dp_x > h; \quad dy dp_y > h; \quad dz dp_z > h. \quad (2.9)$$

Учитывая (2.9), можно сделать вывод, что два состояния микрочастицы (x, y, z, p_x, p_y, p_z) и $(x + dx, y + dy, z + dz, p_x + dp_x, p_y + dp_y, p_z + dp_z)$ можно различить, если выполняется принцип неопределенностей

$$dx dy dz dp_x dp_y dp_z > h^3. \quad (2.10)$$

Поэтому в квантовых статистиках существует минимальный размер элемента фазового пространства, связываемый с отдельным квантовым состоянием, который на основе формул (2.6) и (2.10) будет равен

$$\Delta\Gamma = \Delta\Gamma_V \Delta\Gamma_p = h^3. \quad (2.11)$$

Для свободных микрочастиц удобнее рассматривать не шестимерное фазовое пространство, а трехмерное пространство импульсов. В этом случае элемент пространства координат равен просто единичному объему полупроводника

$$\Delta\Gamma_V = dx dy dz = 1. \quad (2.12)$$

В этом объеме находятся квантовые микрочастицы (электроны, дырки и др.). Тогда минимальный элемент трехмерного пространства импульсов для квантовой микрочастицы будет определяться выражением

$$\Delta\Gamma_p = h^3. \quad (2.13)$$

Каждому такому элементу в пространстве импульсов соответствует одно квантовое состояние.

Определим плотность состояний у дна зоны проводимости. В этом случае энергия свободного электрона у дна зоны с учетом (1.9) может быть записана в виде

$$E = E_c + \frac{p^2}{2m_n^*}. \quad (2.14)$$

Выделим шаровой слой в пространстве импульсов между сферами с радиусами p и $p + dp$ (рис. 2.2). Объем этого слоя составляет величину

$$\Delta V_p = 4\pi p^2 dp. \quad (2.15)$$

В соответствии с выражениями (2.13) и (2.15), учитывая, что в одном элементе ΔV_p могут находиться два электрона с противоположно направленными спинами, получим, что число различных квантовых состояний в объеме ΔV_p равно

$$dZ = 2 \frac{\Delta V_p}{h^3} = \frac{8\pi}{h^3} p^2 dp. \quad (2.16)$$

Из равенства (2.14) следует:

$$p^2 = 2m_n^*(E - E_c), \quad (2.17)$$

$$p = (2m_n^*)^{1/2} (E - E_c)^{1/2}, \quad (2.18)$$

$$dp = \frac{1}{2} (2m_n^*)^{1/2} (E - E_c)^{-1/2} dE. \quad (2.19)$$

Подставив равенства (2.16)–(2.19) в соотношение (2.5), получим выражение для плотности квантовых состояний у дна зоны проводимости:

$$N(E) = \frac{4\pi}{h^3} (2m_n^*)^{3/2} (E - E_c)^{1/2}. \quad (2.20)$$

Отсчет энергии дырок в валентной зоне начинается от ее потолка E_v . Поэтому формула для энергии свободной дырки будет иметь вид (с учетом, что энергия дырок увеличивается начиная с E_v по направлению в валентную зону, т.е. в обратную сторону от направления роста энергии электрона):

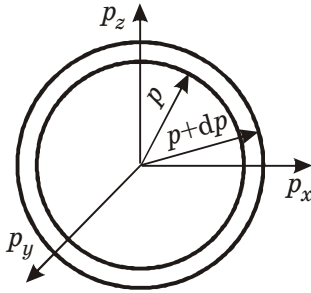


Рис. 2.2. Шаровой слой в пространстве импульсов

$$E = E_v - \frac{p^2}{2m_p^*}. \quad (2.21)$$

Расчеты, аналогичные проделанным выше, дают следующее выражение для плотности квантовых состояний в валентной зоне:

$$N(E) = \frac{4\pi}{h^3} (2m_p^*)^{3/2} (E_v - E)^{1/2}. \quad (2.22)$$

Из формул (2.20) и (2.22) следует, что если энергия носителей заряда в полупроводниках является квадратичной функцией импульса, то плотность состояний для электронов у дна зоны проводимости и для дырок у потолка валентной зоны $N(E)$ имеет зависимость от энергии вида $(E - E_c)^{1/2}$ и $(E_v - E)^{1/2}$ (рис. 2.3).

Рис. 2.3. Энергетическая зависимость функции плотности состояний в разрешенных зонах

§ 2.3. Концентрация электронов и дырок в полупроводнике

На основе вышеописанных представлений об энергетическом распределении квантовых состояний в разрешенных зонах (см. § 2.2) и о распределении электронов по энергетическим уровням (см. § 2.1) можно вычислить концентрацию электронов $n(E)$ в энергетическом интервале от E до $E + dE$:

$$n(E) dE = N(E) f_{\Phi-\text{д}}(E) dE. \quad (2.23)$$

В § 1.10 показано, что за счет тепловой энергии электроны из валентной зоны переходят на свободные уровни зоны проводимости. В результате этих переходов в единичном объеме полупроводника на энергетических уровнях у дна зоны проводимости появляются свободные электроны, концентрация которых равна n_0 , а в валентной зоне вблизи ее потолка — свободные дырки с концентрацией p_0 .

Здесь возникает вопрос: какова вероятность того, что тот или иной уровень в зоне проводимости будет занят электроном, а тот или иной уровень валентной зоны лишится своего электрона и превратится в дырку?

Согласно статистике Ферми — Дирака (см. § 2.1) вероятность того, что состояние с энергией E при данной температуре T занято электроном $f_n(E, T)$, выражается функцией, аналогичной выражению (2.2):

$$f_n(E, T) = \frac{1}{\exp\left(\frac{E - E_F}{k_0 T}\right) + 1}. \quad (2.24)$$

Очевидно, вероятность того, что уровень в валентной зоне не занят электроном, т.е. вероятность нахождения на этом уровне дырки $f_p(E, T)$, будет определяться выражением

$$f_p(E, T) = 1 - f_n(E, T) \quad (2.25)$$

или, учитывая (2.24),

$$f_p(E, T) = \frac{1}{\exp\left(\frac{E_F - E}{k_0 T}\right) + 1}. \quad (2.26)$$

Концентрацию электронов в зоне проводимости полупроводника можно получить из равенства (2.23) после подстановки в него выражений (2.20) и (2.24):

$$n_0 = \int_{E_c}^{\infty} \frac{4\pi}{h^3} (2m_n^*)^{3/2} (E - E_c)^{1/2} \frac{1}{e^{\frac{E - E_F}{k_0 T}} + 1} dE. \quad (2.27)$$

При расчете концентрации электронов необходимо охватить весь интервал энергетических уровней, начиная от дна зоны проводимости E_c и до ∞ , так как степень заполнения этих уровней, например, с ростом температуры изменяется только у дна этой зоны.

Подставляя в (2.23) формулы (2.22) для плотности квантовых состояний в валентной зоне и (2.25) для вероятности их заполнения дырками, получаем выражение для равновесной концентрации дырок в валентной зоне:

$$p_0 = \int_{-\infty}^{E_v} \frac{4\pi}{h^3} (2m_p^*)^{3/2} (E_v - E)^{1/2} \frac{1}{e^{\frac{E_F - E}{k_0 T}} + 1} dE. \quad (2.28)$$

Нижний предел интегрирования в выражении (2.28) выбран произвольно, так как он не должен влиять на величину p_0 , поскольку дырки располагаются только на уровнях вблизи потолка валентной зоны.

Вычисление интегралов (2.27) и (2.28) приводит к следующему результату для концентраций электронов и дырок в невырожденных полупроводниках:

$$n_0 = 2 \left(\frac{2\pi m_n^* k_0 T}{h^2} \right)^{3/2} \exp\left(-\frac{E_c - E_F}{k_0 T}\right) = N_c \exp\left(-\frac{E_c - E_F}{k_0 T}\right); \quad (2.29)$$

$$p_0 = 2 \left(\frac{2\pi m_p^* k_0 T}{h^2} \right)^{3/2} \exp\left(-\frac{E_F - E_v}{k_0 T}\right) = N_v \exp\left(-\frac{E_F - E_v}{k_0 T}\right), \quad (2.30)$$

где

$$N_c = 2 \left(\frac{2\pi m_n^* k_0 T}{h^2} \right)^{3/2}, \quad (2.31)$$

$$N_v = 2 \left(\frac{2\pi m_p^* k_0 T}{h^2} \right)^{3/2}. \quad (2.32)$$

Величины N_c и N_v имеют название эффективной плотности состояний соответственно в зоне проводимости и в валентной зоне. В германии при $T = 300 \text{ К}$ $N_c \approx 1 \cdot 10^{19} \text{ см}^{-3}$, $N_v = 6 \cdot 10^{18} \text{ см}^{-3}$.

Из сравнения формулы (2.23) с (2.29) и (2.30) с учетом (2.3) вытекает физический смысл понятия эффективной плотности состояний для зоны проводимости и валентной зоны полупроводника. Вводя понятие эффективной плотности состояний, мы заменяем реальные функции (см. выражения (2.20) и (2.22)), зависящие от энергии, некоторыми эффективными плотностями состояний, расположенными у дна зоны проводимости или потолка валентной зоны. Использование эффективных плотностей состояний позволяет получить те же результаты, что и использование реальных функций плотности состояний, но значительно упрощает расчеты.

§ 2.4. Положение уровня Ферми и концентрация свободных носителей в собственных полупроводниках

В соответствии с условием электрической нейтральности полупроводника сумма всех зарядов в нем равна нулю:

$$(+q) p_0 + (-q) n_0 = 0. \quad (2.33)$$

Для собственного полупроводника (§1.10) известно равенство концентраций электронов и дырок. Согласно выражению (1.57)

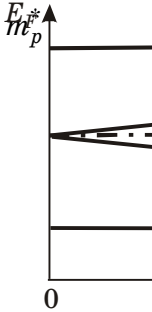
$$n_0 = p_0 = n_i.$$

Приравнивая (2.29) и (2.30), получаем точное выражение для положения уровня Ферми в собственном полупроводнике:

$$E_F = \frac{E_c + E_v}{2} + \frac{3}{4} k_0 T \ln \frac{m_p^*}{m_n^*}. \quad (2.34)$$

Из (2.34) видно, что уровень Ферми при $T = 0 \text{ К}$ лежит в середине запрещенной зоны $(E_c + E_v)/2 = E_i$ и его положение линейно зависит от температуры (рис. 2.4). С ростом температуры уровень Ферми приближается к той зоне, в которой носители заряда имеют меньшую эффективную массу. Поскольку $k_0 T$ — величина малая, а m_n^* сравнима с m_p^* , то уровень Ферми в собственном полупроводнике обычно незначительно смещен относительно середины запрещенной зоны.

Рис. 2.4. Зависимость энергии Ферми от температуры для собственного полупроводника



Используя равенство (1.57), можно определить концентрацию носителей заряда в собственном полупроводнике:

$$(2.35)$$

Таким образом, из (2.35) следует, что концентрация носителей заряда в собственном полупроводнике экспоненциально возрастает при увеличении температуры и уменьшении ширины запрещенной зоны:

$$n_i \sim \exp(-\Delta E/2k_0T). \quad (2.36)$$

В § 1.10 приводились значения ΔE и соответствующие им величины n_i для собственных полупроводников германия, кремния и арсенида галлия, что согласуется с вышеизложенным. Можно заметить, что увеличение температуры германия от 100 до 600 К повышает величину n_i на 17 порядков.

Логарифмируя выражение (2.35), получаем

$$(2.37)$$

Первый член в правой части (2.37) зависит от температуры значительно слабее, чем второй. Поэтому температурная зависимость логарифма

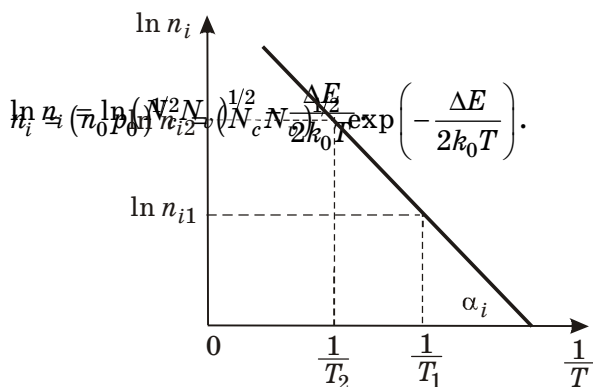


Рис. 2.5. Зависимость концентрации носителей от температуры для собственного полупроводника

концентрации собственных носителей заряда $\ln n_i$ будет представлять прямую линию (рис. 2.5), тангенс угла наклона которой связан с шириной запрещенной зоны выражением

$$\operatorname{tg} \alpha_i = -\frac{\Delta E}{2k_0}. \quad (2.38)$$

Таким образом, если для полупроводника экспериментально измерена зависимость концентрации собственных носителей заряда от температуры, то из нее легко определить ширину запрещенной зоны материала по формуле

$$\Delta E = 2k_0 \frac{\ln n_{i2} - \ln n_{i1}}{1/T_1 - 1/T_2}, \quad (2.39)$$

где T_1 и T_2 — два произвольно выбранных значения температуры; n_{i1} и n_{i2} — концентрации собственных носителей заряда в полупроводнике при соответствующих температурах T_1 и T_2 (см. рис. 2.5).

§ 2.5. Положение уровня Ферми и концентрация свободных носителей в примесных полупроводниках

Рассмотрим полупроводник, в который введена донорная примесь с концентрацией N_d . Уравнение электронейтральности в этом случае будет иметь вид

$$(2.40)$$

что имеет следующий физический смысл. В донорном полупроводнике свободные носители заряда в зоне проводимости возникают вследствие

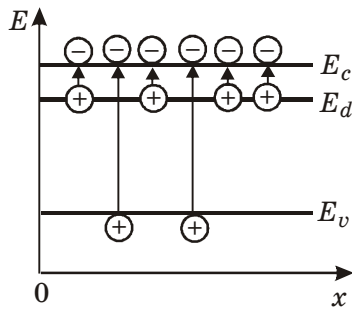


Рис. 2.6. Схема возникновения подвижных и неподвижных зарядов в донорном полупроводнике

переходов электронов из валентной зоны и примесного уровня E_d (рис. 2.6), что сопровождается появлением положительно заряженных дырок с концентрацией p_0 и донорных ионов с концентрацией N_d^+ . Таким образом, отрицательный заряд электронов в полупроводнике полностью компенсируется положительным зарядом донорных ионов и дырок.

Из уравнения электронейтральности (2.40) можно определить положение уровня Ферми в донорном полупроводнике. Для перевода электрона из валентной зоны в зону проводимости необходима энергия, равная ширине запрещенной зоны ($\Delta E = E_c - E_v$), тогда как для перевода электрона с донорного уровня требуется энергия значительно меньшая, равная энергии ионизации примеси $E_c - E_d$. Поэтому при низкой температуре основную роль играют электроны, перешедшие в зону проводимости с донорного уровня. В этом случае

$$(2.41)$$

и уравнение электронейтральности (2.40) можно записать в виде

$$n_0 = N_d^+ \quad (2.42)$$

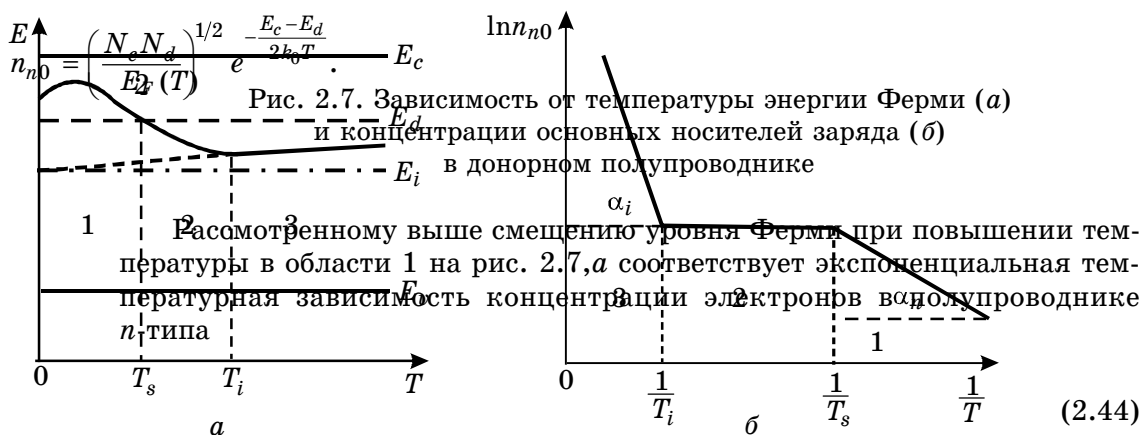
Таким образом, при низких температурах в полупроводнике имеются только электроны в зоне проводимости и донорные ионы в узлах кристаллической решетки. В этом случае решение уравнения (2.42) приводит к следующему выражению для положения уровня Ферми:

$$E_F = \frac{E_c + E_d}{2} + \frac{k_0 T}{2} \ln \frac{N_d}{2N_c} \quad (2.43)$$

Из уравнения (2.43) видно, что при $T = 0$ уровень Ферми в донорном полупроводнике лежит посередине между дном зоны проводимости и примесным уровнем. Температурная зависимость E_F обусловлена температур-

$$p_0 \ll p_0^+ N_d^+$$

ной зависимостью N_c и множителем k_0T . С увеличением T в области низких температур, когда $2N_c < N_d$, уровень Ферми сначала приближается к E_c , а затем начинает опускаться в запрещенную зону (рис. 2.7,а). При $2N_c = N_d$ положение уровня Ферми снова будет $E_F = (E_c + E_d)/2$, и при достижении температуры T_s уровень Ферми пересекает донорный уровень. Температура T_s называется температурой истощения примеси. Дальнейшее повышение температуры сопровождается ростом N_c (см. формулу (2.31)), и уровень Ферми продолжает снижаться до температуры T_i , начиная с которой донорный полупроводник проявляет свойства собственного полупроводника, так как при $T = T_i$ уровень Ферми должен приблизиться к середине запрещенной зоны E_i или пересечь ее.



Тангенс угла наклона зависимости $\ln n_{n0} = f(1/T)$ в области 1 (рис. 2.7,б) определяет глубину залегания донорного уровня $E_c - E_d$ относительно дна зоны проводимости:

$$-\operatorname{tg} \alpha_n = \frac{E_c - E_d}{2k_0} \quad (2.45)$$

Эта область смещения уровня Ферми при повышении температуры описывается выражением (2.43) и называется областью слабой ионизации

примеси (область 1 на рис. 2.7,а). Экспоненциальный рост концентрации электронов в полупроводнике n -типа показан в полулогарифмических координатах $\ln n_{n0} = f(1/T)$ в виде прямой линии на рис. 2.7,б (область 1). Значения температуры истощения

$$T_s = \frac{E_c - E_d}{k_0 \ln \frac{2N_c}{N_d}} \quad (2.46)$$

обычно невелики. Например, для донорной примеси в германии, когда $E_c = E_d \approx 0,01$ эВ и $N_d = 10^{16}$ см⁻³, $T_s = 32$ К.

При повышении температуры от 0 К до T_s концентрация электронов в зоне проводимости возрастает и сравнивается с концентрацией атомов донорной примеси ($n_0 = N_d$). Это означает, что практически все донорные атомы ионизированы и при отсутствии других источников носителей заряда концентрация электронов в зоне проводимости не будет зависеть от температуры (область 2 на рис. 2.7,б). Эта область температур, при которой имеет место полная ионизация примесных атомов, называется областью истощения примеси (область 2, рис. 2.7,б). Из условия $n_0 = N_d$ с учетом (2.29) получаем $E_F = E_c + k_0 T \ln N_d / N_c$. При увеличении температуры, начиная с T_i , рост концентрации электронов в зоне проводимости будет связан с переходом электронов из валентной зоны. В этом случае положение уровня Ферми определяется уравнением (2.34), и его температурная зависимость становится аналогичной случаю собственного полупроводника (см. рис. 2.4). На рис. 2.7,а показан ход зависимости $E_F(T)$ в области температур выше T_i в предположении, что эффективная масса у электронов меньше, чем у дырок.

Рост концентрации электронов, начиная с температуры T_i , показан на рис. 2.7,б (область 3). При температуре

$$T_i = \frac{\Delta E}{k_0 \ln \left(\frac{N_c N_v}{N_d^2} \right)} \quad (2.47)$$

свойства донорного полупроводника перестают отличаться от аналогичных свойств собственного полупроводника.

В области 3 на рис. 2.7,б при $T > T_i$ зависимость $\ln n_{n0}(1/T)$ не отличается от такой же зависимости для собственного полупроводника (см. рис. 2.5), и по ее наклону относительно оси $1/T$ можно определить ширину запрещенной зоны (см. § 2.4, формула (2.39)).

В акцепторном полупроводнике при $T = 0$ К уровень Ферми расположен в середине интервала $E_a - E_v$ (рис. 2.8). При $T = T_s$ $E_F = E_a$, что соответствует полной ионизации акцепторных атомов. Концентрация дырок при $T < T_s$ определяется по формуле

$$(2.48)$$

При $T = T_s$ $p_{p0} = N_a$ и не изменяется при дальнейшем увеличении температуры до T_i . При $T > T_i$ дырочный полупроводник проявляет свойства собственного полупроводника. Значения температур T_s и T_i определяются по формулам:

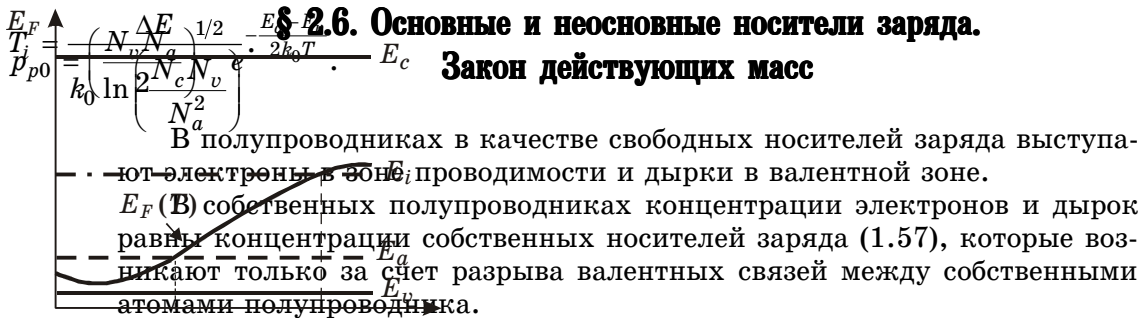
$$T_s = \frac{E_a - E_v}{k_0 \ln \left(\frac{2N_v}{N_a} \right)}, \quad (2.49)$$

Рис. 2.8. Зависимость энергии Ферми от температуры для акцепторного полупроводника

где N_a — концентрация акцепторной примеси,

$$(2.50)$$

Зависимость $\ln p_{p0}(1/T)$ подобна показанной на рис. 2.7,б.



§ 2.6. Основные и неосновные носители заряда. Закон действующих масс

В полупроводниках в качестве свободных носителей заряда выступают электроны в зоне проводимости и дырки в валентной зоне. E_F (В) собственных полупроводниках концентрации электронов и дырок равны концентрации собственных носителей заряда (1.57), которые возникают только за счет разрыва валентных связей между собственными атомами полупроводника.

В примесных полупроводниках наибольшее количество свободных носителей заряда (при $T > 0$ К) поставляют примесные атомы — доноры в полупроводнике n -типа и акцепторы в полупроводнике p -типа. Донорные атомы в полупроводнике n -типа поставляют дополнительные электроны в зону проводимости, вследствие чего выполняется неравенство (1.60). Акцепторные атомы в дырочном полупроводнике поставляют в валентную зону дополнительные дырки, вследствие чего концентрация дырок значительно превышает концентрацию собственных дырок (1.61).

В соответствии с вышеизложенным полупроводник называется электронным, если выполняется неравенство (1.60), и дырочным, если справедливо неравенство (1.61). Тогда электроны в электронном полупроводнике, а дырки в дырочном являются основными носителями заряда.

Однако в названных полупроводниках при $T > 0$ К всегда есть неосновные носители заряда — дырки в электронном полупроводнике и электроны в дырочном полупроводнике.

Энергия ионизации примесных атомов, т.е. глубина залегания донорного или акцепторного уровня, всегда меньше ширины запрещенной зоны полупроводника. Поэтому в примесных полупроводниках концентрация основных носителей заряда всегда больше концентрации неосновных носителей:

$$n_{n0} \gg p_{n0}, \quad (2.51)$$

$$p_{p0} \gg n_{p0}, \quad (2.52)$$

где p_{n0} — равновесная концентрация дырок в полупроводнике n -типа; n_{p0} — равновесная концентрация электронов в полупроводнике p -типа.

При любой температуре в любом полупроводнике устанавливается равновесное состояние, при котором концентрации электронов и дырок выражаются соответственно формулами (2.29) и (2.30). При этом должно выполняться соотношение (2.35), которое можно представить в виде

$$(2.53)$$

Это равенство носит название закона действующих масс и показывает, что в невырожденном полупроводнике произведение равновесных концентраций электронов и дырок равно квадрату концентрации собственных носителей заряда. Соотношение (2.53) справедливо только для носителей заряда, которые созданы за счет теплового воздействия на невырожденный полупроводник и находятся с ним в тепловом равновесии.

Для электронного полупроводника закон действующих масс (выражение (2.53)) будет иметь вид

$$n_{n0} p_{n0} = n_i^2 \quad (2.54)$$

и для акцепторного —

$$p_{p0} n_{p0} = n_i^2, \quad (2.55)$$

т.е. произведение равновесных концентраций основных и неосновных носителей заряда в данном невырожденном полупроводнике равно квадрату концентрации собственных носителей заряда (при соответствующей температуре).

Закон действующих масс позволяет по известной концентрации основных носителей рассчитать концентрацию неосновных и наоборот.

$$n_0 p_0 = n_i^2.$$

§ 2.7. Влияние уровня легирования на температурную зависимость концентрации основных носителей заряда

Температурная зависимость концентрации электронов в донорном полупроводнике с низким уровнем легирования (невысокая концентрация примесей N_{d1}) показана на рис. 2.9,а. В этом случае зависимость $\ln n(1/T)$ не отличается от рассмотренной ранее (см. рис. 2.7,б). Увеличение концентрации доноров ($N_{d2} > N_{d1}$) вызывает смещение зависимости $\ln n(1/T)$, рост температуры истощения примеси ($T_{s2} > T_{s1}$) и температуры перехода к собственной проводимости ($T_{i2} > T_{i1}$), а также уменьшение угла наклона ($\alpha_{n2} < \alpha_{n1}$). Аналогичное изменение зависимости $\ln p(1/T)$ наблюдается и для акцепторных полупроводников.

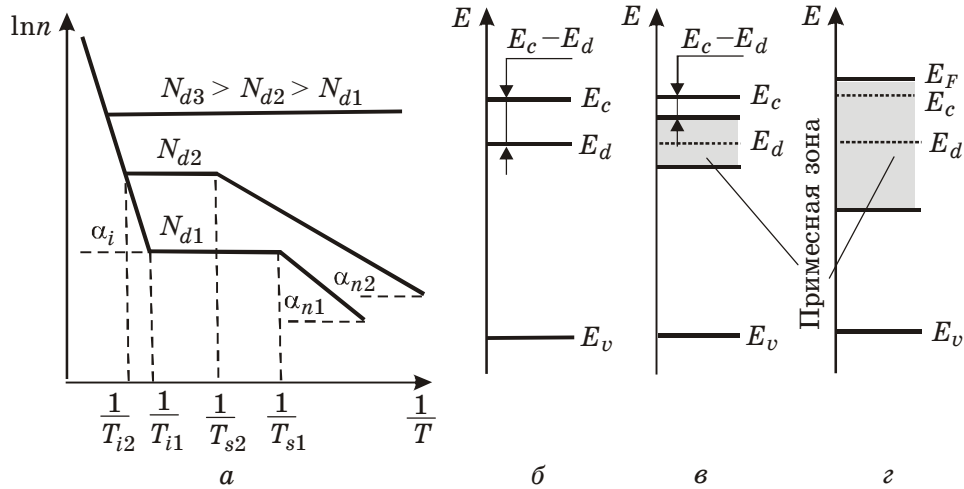


Рис. 2.9. Температурные зависимости концентрации электронов в полупроводнике (а) и соответствующие зонные диаграммы полупроводника (б, в, г) при различных концентрациях донорной примеси

Уменьшение угла наклона α_{n2} с ростом концентрации донорных атомов (повышение степени легирования) до N_2 говорит об уменьшении энергии ионизации доноров, на что указывает формула (2.45). Это связано с тем, что увеличение количества донорных атомов в единичном объеме полупроводника приводит к уменьшению расстояния между ними и, следовательно, к повышению их взаимодействия. При достаточно большой концентрации доноров N_{d2} это взаимодействие становится настолько сильным, что донорный уровень E_d (рис. 2.9,б) расщепляется в примесную зону (рис. 2.9,в), ширина которой возрастает с увеличением взаимодействия

донорных атомов, т.е. с ростом их концентрации. Расщепление донорного уровня в примесную зону объясняется перекрытием волновых функций валентных электронов взаимодействующих донорных атомов.

Увеличение концентрации доноров, например, в германии до $N_{d3} \approx 3 \cdot 10^{18} \text{ см}^{-3}$ приводит к настолько большому расщеплению донорного уровня, что примесная зона перекрывается с зоной проводимости, образуя единую зону разрешенных энергий (рис. 2.9,з). Вследствие этого энергия ионизации примесных атомов обращается в нуль, а концентрация электронов в зоне проводимости перестает зависеть от температуры (рис. 2.9,а, зависимость $\ln n(1/T)$ при N_{d3}). Уровень Ферми у таких полупроводников находится в зоне проводимости, и состояния ниже E_F у дна зоны заняты практически полностью, как у металлов. Однако с повышением температуры в связи с перемещением уровня Ферми в запрещенную зону и его смещением к середине запрещенной зоны, как было показано в § 2.5, примесные электроны перераспределяются по большому числу состояний у дна зоны, степень заселенности этих состояний уменьшается и при снижении уровня Ферми ниже дна примесной подзоны на $\sim 3k_0T$ вырождение снимается.

§ 2.8. Компенсированные полупроводники

Рассмотрим случай, когда полупроводник одновременно содержит донорную и акцепторную примеси, т.е.

$$N_d \neq 0; N_a \neq 0. \quad (2.56)$$

На рис. 2.10,а показана зонная энергетическая диаграмма такого полупроводника.

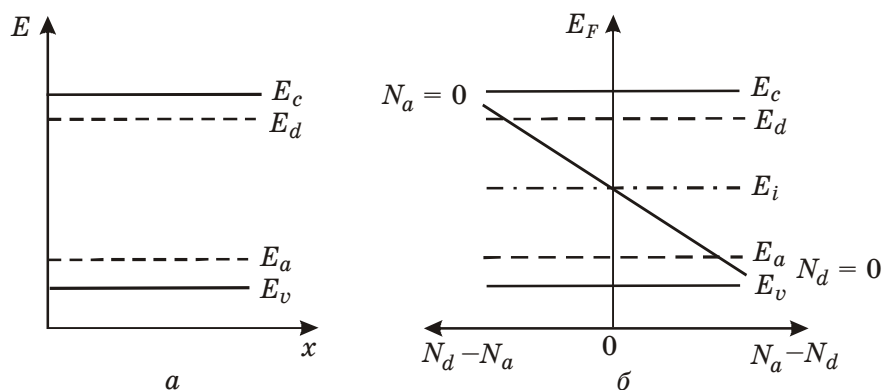


Рис. 2.10. Зонная диаграмма полупроводника, содержащего донорную и акцепторную примеси (а), и зависимость положения уровня Ферми в запрещенной зоне такого полупроводника от степени компенсации этих примесей (б)

Согласно статистике Ферми — Дирака электроны занимают наименьшие энергетические состояния. Поэтому валентные электроны донорных атомов при $T > 0 \text{ К}$ в первую очередь будут переходить на акцепторные уровни. Если

$$N_d = N_a, \quad (2.57)$$

то наблюдается компенсация положительных донорных ионов отрицательно заряженными ионами акцепторов. В результате в валентную зону и в зону проводимости не поставляются дополнительные дырки и электроны. Происходит полная взаимная компенсация примесей, и свободные носители в разрешенных зонах возникают только за счет переброса электронов из валентной зоны в зону проводимости. Следовательно, компенсированный полупроводник в случае равенства (2.57) будет вести себя аналогично собственному.

Если

$$N_d > N_a, \quad (2.58)$$

то полупроводник будет донорным в результате частичной компенсации акцепторов. Результирующая, так называемая эффективная, концентрация в этом случае будет

$$N'_d = N_d - N_a. \quad (2.59)$$

Когда в полупроводнике

$$N_a > N_d, \quad (2.60)$$

то получают частично скомпенсированный акцепторный полупроводник, близкий по свойствам к полупроводнику, содержащему только акцепторы с эффективной концентрацией

$$N'_a = N_a - N_d. \quad (2.61)$$

На рис. 2.10,б показана зависимость положения уровня Ферми от соотношения между величинами N_d и N_a при низких температурах ($T < T_i$). При равенстве (2.57) уровень Ферми находится в середине запрещенной зоны. Преобладание доноров приводит к смещению уровня Ферми в верхнюю половину запрещенной зоны в зависимости от величины N_d . Уровень E_F смещается в нижнюю половину запрещенной зоны при увеличении эффективной концентрации акцепторов (2.61).

Контрольные вопросы и задачи

1. Распределение микрочастиц по квантовым состояниям описывается следующими формулами:

$$1) f_{\Phi-д}(E, T) = \frac{1}{e^{\frac{E-E_F}{k_0T}} + 1}; \quad 2)$$

$$3) f_{M-B}(E, T) = \frac{1}{e^{\frac{E-E_F}{k_0T}}}; \quad 4) f_{B-\Delta}(E, T) = \frac{1}{e^{\frac{E-E_F}{k_0T}} - 1};$$

5)

Укажите номера правильных формул для распределения микрочастиц в нижеуказанных коллективах:

(ЛБНК) невырожденные коллективы;

(МПВК) вырожденные коллективы.

2. (ЛТПС). Выберите правильные с физической точки зрения определения плотности и эффективной плотности квантовых состояний в зоне проводимости и в валентной зоне полупроводника:

1) понятие плотности квантовых состояний справедливо в целом для полос энергетических уровней в разрешенных зонах и представляет собой число состояний в единичном интервале энергии для единичного объема кристалла;

2) плотность квантовых состояний в зоне проводимости равна числу электронов, которые при заданной температуре могут перейти в зону проводимости;

3) плотность квантовых состояний в валентной зоне равна числу электронов, покинувших ее при заданной температуре;

4) плотность квантовых состояний в зоне проводимости равна числу энергетических уровней во всех атомах, объединенных в кристалл;

5) эффективная плотность квантовых состояний в зоне проводимости N_c и в валентной зоне N_v — это плотность состояний в указанных зонах, приведенная к дну зоны проводимости E_c и к потолку валентной зоны E_v .

3. (ЮРУФ). Что происходит с уровнем Ферми в собственном полупроводнике при повышении температуры от абсолютного нуля?

Ответы:

1) положение уровня Ферми смещается от середины запрещенной зоны к той разрешенной зоне, где носители заряда имеют меньшую эффективную массу;

2) уровень Ферми смещается к дну зоны проводимости при $m_p^* > m_n^*$;

3) уровень Ферми смещается к потолку валентной зоны, если эффективная масса электрона больше эффективной массы дырки;

4) в собственном полупроводнике положение уровня Ферми не зависит от температуры.

4. (ФУТЗ). Выберите правильную формулу, которая описывает температурную зависимость концентрации носителей заряда в собственном полупроводнике:

$$f(E, T) = \frac{1}{e^{\frac{E-E_F}{k_0T}}}$$

$$1) n_i = (N_c N_d)^{1/2} e^{-\frac{\Delta E}{2k_0 T}}; \quad 2) n_i = (N_c N_v)^{1/2} e^{-\frac{\Delta E}{2k_0 T}};$$

$$3) n_i = (N_v N_a)^{1/2} e^{-\frac{\Delta E}{2k_0 T}}; \quad 4) n_i = (N_v N_a)^{1/2} e^{-\frac{\Delta E}{2k_0 T}},$$

где N_c, N_v — эффективные плотности квантовых состояний в разрешенных зонах; N_a, N_d — концентрации донорных и акцепторных ионов.

5. (ЮОУФ). Опишите температурную зависимость положения уровня Ферми в донорном полупроводнике начиная от $T = 0$ К и до $T \approx 380$ К.

Ответы:

1) при повышении температуры от 0 К уровень Ферми смещается от донорного уровня и оказывается в зоне проводимости;

2) уровень Ферми опускается в валентную зону;

3) зависимость $E_F(T)$ имеет сложный характер. С ростом температуры от $T = 0$ К уровень Ферми сначала поднимается к E_c , потом опускается почти до середины запрещенной зоны и далее ход зависимости $E_F(T)$ аналогичен зависимости в собственном полупроводнике.

6. (5АПС). При какой температуре донорный и акцепторный полупроводники приобретают свойства собственного полупроводника?

Ответы:

1) при любых температурах примесные полупроводники сохраняют свои свойства, так как примесные атомы остаются в объемах кристаллов;

2) при $T = T_s$, когда уровень Ферми пересекает донорный или акцепторный уровень;

3) при $T \geq T_i$, когда уровень Ферми должен оказаться в середине запрещенной зоны;

4) при температурах в области $T_s < T < T_i$, когда все примесные уровни ионизированы.

7. (ВППП). Укажите основные и неосновные носители заряда в донорном и акцепторном полупроводниках:

1) в донорном полупроводнике основные носители заряда — дырки в валентной зоне, образующиеся при переходе из нее электронов в зону проводимости;

2) в донорном (электронном) полупроводнике концентрация электронов в зоне проводимости за счет их поступления с донорного уровня при $0 < T < T_i$ выше, чем концентрация дырок в валентной зоне. Поэтому электроны в донорном полупроводнике называются основными носителями заряда, а дырки — неосновными;

3) в акцепторном полупроводнике основные носители заряда — это электроны, перешедшие при $T > 0$ в валентную зону с акцепторного уровня, а неосновные носители — дырки, оставшиеся на акцепторном уровне;

4) в акцепторном (дырочном) полупроводнике основными носителями заряда являются дырки в валентной зоне, образующиеся при переходе

из нее электронов при $0 < T < T_i$ на акцепторный уровень, а электроны, количество которых в зоне проводимости меньше, чем дырок в валентной зоне, представляют неосновные носители.

8. (И1ТЗ). Как изменяется концентрация основных носителей заряда при увеличении температуры полупроводника?

Ответы:

1) концентрации электронов в электронном полупроводнике и дырок в акцепторном растут экспоненциально при увеличении температуры, начиная с $T > T_i$, вследствие тепловой ионизации примесных уровней;

2) концентрация основных носителей заряда увеличивается экспоненциально в области температур $T > T_i$ вследствие ионизации собственных атомов полупроводника;

3) в области $T < T_s$ концентрация основных носителей заряда экспоненциально возрастает вследствие ускорения процесса ионизации примесных атомов;

4) при увеличении температуры в области $T_s \leq T \leq T_i$ концентрация основных носителей заряда остается постоянной вследствие полной ионизации примесных атомов;

5) в области $T < T_s$ наблюдается преобладающий процесс ионизации собственных атомов полупроводника.

9. (БКЭФ). Укажите формулу, описывающую закон действующих масс:

1) $n_{n0}n_{p0} = p_n^2$;

4) $p_{p0}n_{p0} = n_i^2$;

2) $p_{p0}n_{n0} = n_i^2$;

5) $p_{n0}n_i = n_{p0}^2$;

3) $n_{n0}p_{n0} = n_i^2$;

6) $n_{n0}p_{n0} = p_i^2$.

10. (ЯГНН). Как зависит концентрация неосновных носителей заряда в полупроводниках от ширины запрещенной зоны?

Ответы:

1) концентрация неосновных носителей заряда экспоненциально увеличивается или уменьшается при соответствующем уменьшении или увеличении ширины запрещенной зоны, так как переход электронов в зону проводимости через запрещенную зону обязательно приводит к возникновению в валентной зоне дырки и какие-то носители из двух указанных выше видов окажутся неосновными;

2) зависит в соответствии с выражениями $n_{n0}p_{n0} = n_i^2$, $p_{p0}n_{p0} = n_i^2$.

3) концентрация неосновных носителей заряда зависит только от положения донорного или акцепторного уровня в запрещенной зоне и не зависит от ширины запрещенной зоны;

4) зависит только концентрация основных носителей заряда, а концентрация неосновных носителей всегда остается постоянной.

11. (КФПЗ). Найдите концентрацию неосновных носителей заряда в германии, легированном атомами фосфора в концентрации 10^{21} м^{-3} . Считать, что все примесные атомы при $T = 300 \text{ К}$ ионизированы и концентрация собственных носителей заряда равна $2,1 \cdot 10^{19} \text{ м}^{-3}$.

Ответ набрать в виде ряда из 5 цифр без разделительных знаков: целое число — одна цифра, дробная часть — две цифры, степень — две цифры. Единицу концентрации (м^{-3}) не указывать.

12. (Л2НН). Найдите концентрацию неосновных носителей заряда в германии с примесью атомов In. Условия аналогичны задаче 11.

13. (Б5ВК). Каким образом примеси в полупроводнике могут влиять на ширину его запрещенной зоны ΔE ?

Ответы:

1) увеличение концентрации донорных или акцепторных атомов приводит только к росту концентрации основных носителей заряда соответствующего типа без изменения величины ΔE ;

2) увеличение концентрации доноров или акцепторов в германии до $N_d \approx 3 \cdot 10^{24} \text{ м}^{-3}$ приводит к расщеплению примесного уровня так, что образуется примесная зона разрешенных энергетических уровней, которая перекрывается с зоной проводимости или валентной зоной, и ширина запрещенной зоны уменьшается;

3) увеличение концентрации примесных атомов в полупроводнике может только увеличить ширину запрещенной зоны вследствие дополнительного взаимодействия.

Глава 3. Электропроводность твердых тел

§ 3.1. Влияние электрического поля на функцию распределения носителей заряда в полупроводниках

В отсутствие внешнего электрического поля свободные носители заряда в полупроводниковом кристалле совершают беспорядочное, т.е. хаотическое, тепловое движение, что объясняется взаимодействием электронов с нарушениями периодического поля кристаллической решетки. Это приводит к изменению вектора скорости носителей заряда, вследствие чего изменяется как направление движения, так и абсолютная величина их скорости. В результате взаимодействия носителей заряда электронный газ в кристалле находится в равновесном состоянии, которое описывается равновесными функциями распределения Максвелла — Больцмана, если газ не вырожден (см. формулу (2.3)), или Ферми — Дирака в случае его вырождения (см. выражение (2.2)).

Классическая формула (1.9) для кинетической энергии электрона с эффективной массой (1.53) принимает вид

$$E = \frac{m^* v^2}{2}. \quad (3.1)$$

В этом случае функции распределения Максвелла — Больцмана и Ферми — Дирака в зависимости от компоненты скорости v_x по направлению x будут иметь вид, показанный на рис. 3.1.

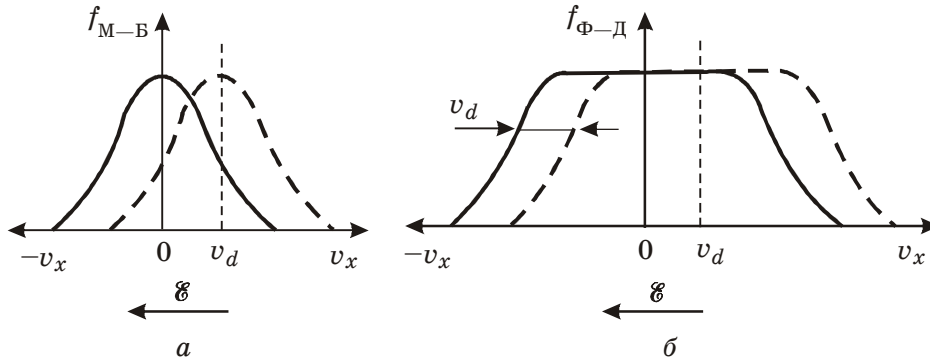


Рис. 3.1. Функции распределения электронов по скоростям для невырожденного (а) и вырожденного (б) электронного газа

Так как равновесные функции распределения (сплошные кривые) симметричны относительно $v_x = 0$, т.е. количество электронов, движущихся в противоположных направлениях, одинаково, и поэтому в отсутствие

внешнего электрического поля нет направленного потока электронов, представляющего электрический ток. Таким образом, в результате взаимодействия с дефектами электроны хаотически распределяются по кристаллу и их средняя скорость равна нулю.

Во внешнем электрическом поле, напряженность которого равна \mathcal{E} , на электрон действует сила

$$F = -q \mathcal{E}, \quad (3.2)$$

вследствие чего происходит накопление энергии и перераспределение электронов по энергетическому спектру. Тепловое равновесие нарушается, и функции распределения электронов по скоростям (энергиям) становятся неравновесными (рис. 3.1, штриховые кривые). Функции распределения смещаются по оси скоростей в направлении действующей силы на некоторую величину, называемую скоростью дрейфа v_d . Под скоростью дрейфа следует понимать результирующую скорость движения коллектива электронов в направлении, заданном внешней силой (электрическим полем). При этом траектория движения каждого (отдельного) электрона вследствие взаимодействия с дефектами кристаллической решетки остается очень сложной (рис. 3.2). Одним из существенных видов нарушения периодичности кристаллической решетки являются тепловые колебания атомов.

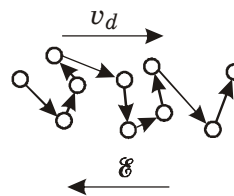


Рис. 3.2. Траектория движения отдельного электрона в твердом теле

Траектория движения электронов, попадающих в область воздействия колеблющихся атомов (фононов), искривляется.

В реальных кристаллах существует множество дефектов, например примесные атомы, дислокации, вакансии и другие, которые приводят к искажению траектории прямолинейного движения электронов, т.е. к рассеянию. Процесс рассеяния электронов можно объяснить и с позиций взаимодействия материальных частиц, т.е. учитывая корпускулярную точку зрения, и в представлении волновой природы электрона.

Влияние кристаллической решетки на электрон, движущийся в ней под действием внешней силы F (см. (3.2)), формально можно свести к противодействию силы сопротивления F_c , как показано на рис. 3.3.

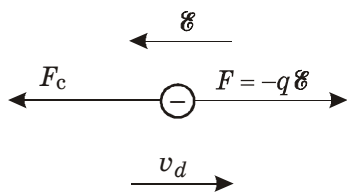


Рис. 3.3. Схематическое изображение сил, действующих на свободный электрон твердого тела

Уравнение направленного движения электрона с учетом сил (рис. 3.3) и формулы (1.43) можно записать следующим образом:

$$(3.3)$$

Из выражения (3.3) видно, что после приложения электрического поля дрейфовая

скорость электронов v_d будет возрастать, т.е. они будут двигаться ускоренно до тех пор, пока сила сопротивления

$$F_c = \frac{m_n^* v_d(t)}{\tau} \quad (3.4)$$

не достигнет одинаковой величины с силой (3.2), действующей со стороны электрического поля. Когда эти силы сравниваются, результирующая сила, действующая на электрон, и ускорение его движения будут равны нулю. Начиная с этого момента, направленное движение электронов будет совершаться с постоянной скоростью дрейфа

$$v_d = -\frac{q \mathbf{E} \tau}{m_n^*} \quad (3.5)$$

в направлении, противоположном вектору \mathbf{E} , вследствие отрицательного заряда электрона.

Выключение электрического поля сопровождается уменьшением дрейфовой скорости электронов вследствие рассеяния на дефектах решетки, и электронный газ будет переходить в равновесное состояние. Такие процессы установления равновесия в системе, ранее выведенной из этого состояния, называются релаксацией. Полагая в формуле (3.5) $\mathbf{E} = \frac{d\mathbf{E}}{dt} \tau$ получаем уравнение, описывающее переход электронного газа в равновесное состояние, т.е. процесс его релаксации:

$$\frac{dv_d(t)}{dt} + \frac{v_d(t)}{\tau} = -\frac{q \mathbf{E}(t)}{m_n^*} \quad (3.6)$$

Интегрируя (3.6), получаем

$$v_d(t) = v_{d0} e^{-t/\tau} - \frac{q \mathbf{E}(t) \tau}{m_n^*} \quad (3.7)$$

где v_{d0} — дрейфовая скорость электронов в момент выключения электрического поля, т.е. в момент $t = 0$.

Из уравнения (3.7) следует, что время релаксации τ характеризует скорость установления в системе равновесного состояния, т.е., фактически, определяет время затухания направленного потока электронов после выключения внешнего электрического поля. За время $t = \tau$ скорость направленного (дрейфового) движения электронов уменьшается в e раз. Время τ называют временем релаксации.

§ 3.2. Электропроводность полупроводников и подвижность носителей заряда

В полупроводниках при $T > 0$ К в электрическом поле с напряженностью \mathbf{E} возникает дрейф электронов и дырок (см. рис. 3.2), которые являются носителями отрицательных и положительных зарядов. Таким обра-

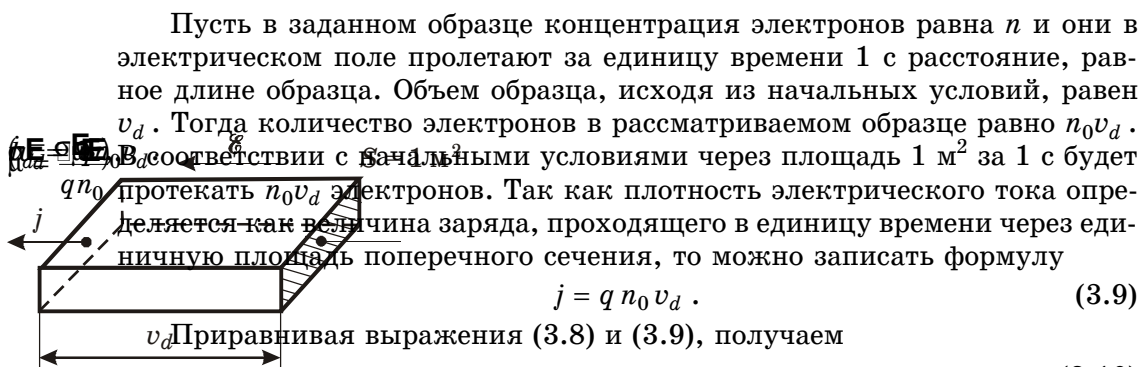
зом, в полупроводнике возникает электрический ток, плотность которого j согласно закону Ома равна

$$(3.8)$$

где коэффициент σ — удельная электрическая проводимость (электропроводность), $\text{Ом}^{-1} \cdot \text{м}^{-1}$ или $\text{См} \cdot \text{м}^{-1}$.

Рассчитаем электропроводность (в дальнейшем будем использовать термин «проводимость») полупроводникового образца с единичной площадью поперечного сечения $S = 1 \text{ м}^2$ и с длиной, численно равной скорости дрейфа v_d (рис. 3.4).

Рис. 3.4. К расчету плотности тока в проводнике



Из равенства (3.10) следует, что

$$v_d = \frac{\sigma}{q n_0} \epsilon, \quad (3.11)$$

т.е. скорость дрейфа носителей заряда пропорциональна напряженности электрического поля. Выражение (3.11) можно записать в виде равенства с коэффициентом пропорциональности μ :

$$(3.12)$$

где

$$(3.13)$$

представляет собой подвижность носителей заряда. Из равенства (3.12) следует, что подвижность равна дрейфовой скорости носителей заряда

в электрическом поле единичной напряженности. Единица измерения подвижности в СИ — $1 \text{ м}^2/(\text{В} \cdot \text{с})$.

Из (3.13) проводимость равна

$$\sigma = q \mu n_0 . \quad (3.14)$$

В полупроводниках в электрическом токе участвуют электроны и дырки. Поэтому общее выражение для проводимости полупроводника имеет вид

$$\sigma = q \mu_n n_0 + q \mu_p p_0 , \quad (3.15)$$

где μ_n — подвижность электронов; μ_p — подвижность дырок.

В электронном полупроводнике обычно $n_{n0} \gg p_{n0}$. Поэтому проводимость полупроводника n -типа выражается формулой

$$\sigma_n \approx q \mu_n n_{n0} . \quad (3.16)$$

В дырочном полупроводнике $p_{p0} \gg n_{p0}$, и проводимость будет

$$\sigma_p \approx q \mu_p p_{p0} . \quad (3.17)$$

С учетом неосновных носителей заряда уточненные формулы имеют вид:

для донорного полупроводника

$$\sigma \approx q \mu_n n_{n0} + q \mu_p p_{n0} ; \quad (3.18)$$

для акцепторного полупроводника

$$\sigma \approx q \mu_p p_{p0} + q \mu_n n_{p0} . \quad (3.19)$$

В собственном полупроводнике концентрации электронов и дырок равны собственной концентрации (1.57), и

$$\sigma_i \approx q (\mu_n + \mu_p) n_i . \quad (3.20)$$

§ 3.3. Зависимость подвижности носителей заряда от температуры

Длина свободного пробега носителей заряда

Направленное движение носителей заряда в полупроводнике в заданном электрическом поле совершается с некоторой постоянной скоростью дрейфа, что обусловлено рассеянием носителей на дефектах кристаллической решетки. Параметром, характеризующим рассеяние носителей заряда, является подвижность, определяемая по формуле (3.13). Из соотношений (3.5) и (3.12) можно получить формальное выражение для подвижности носителей заряда:

$$\mu = \frac{q\tau}{m^*} . \quad (3.21)$$

Иногда удобно пользоваться понятием средней длины свободного пробега носителей заряда l , равной среднему отрезку пути, который прохо-

дит носитель заряда между двумя последовательными актами рассеяния. Если бы носитель уже в единичном акте рассеяния полностью терял скорость в направлении первоначального движения, то средняя длина свободного пробега l равнялась бы произведению средней скорости движения v на время релаксации τ , которое в данном случае представляло бы собой среднее время движения носителя заряда между актами рассеяния:

$$l = v\tau. \quad (3.22)$$

В этом случае выражение (3.21) с учетом (3.22) принимает вид

$$(3.23)$$

Рассеяние носителей заряда на фононах

При высоких температурах в полупроводниках основное влияние на подвижность носителей заряда оказывает рассеяние на тепловых колебаниях кристаллической решетки — фононах. Концентрация квазичастиц — фононов — пропорционально увеличивается с ростом температуры:

$$n_{\text{ф}} \sim T. \quad (3.24)$$

Если носитель заряда прекращает движение в данном направлении уже при единичном столкновении с фононом, то средняя длина свободного пробега

$$\mu = \frac{q}{m^*} \frac{l}{v} = \frac{q}{m^*} \frac{l}{v} \left(\frac{3k_0T}{m^*} \right)^{1/2}, \quad l \propto \frac{1}{T}. \quad (3.25)$$

Средняя энергия теплового движения частицы, как известно из об-

щей физики, равна $\frac{3}{2} k_0T$. Приравнивая эту энергию к кинетической энергии частицы $mv^2/2$, получаем, что средняя скорость хаотического (теплового) движения равняется

$$(3.26)$$

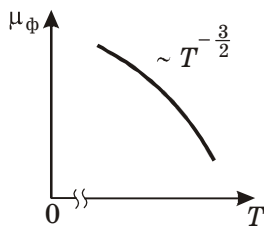


Рис. 3.5. Зависимость подвижности носителей заряда от температуры для полупроводника (рассеяние на фононах)

т.е. $v \propto T^{1/2}$.

Тогда, принимая во внимание температурные зависимости средней длины свободного пробега (3.25) и средней скорости теплового движения (3.26), из (3.23) получаем, что подвижность носителей заряда при рассеянии на фононах (на тепловых колебаниях атомов кристаллической решетки) уменьшается с ростом температуры пропорционально $T^{-3/2}$. Графический вид зависимости $\mu_{\text{ф}}(T)$ в области повышенных температур показан на рис. 3.5.

Рассеяние на ионах примесей

Рассеяние на ионизированных атомах примесей в полупроводниках заметно проявляется в области низких температур, когда концентрация фононов еще незначительна. Рассматриваемый процесс представляет собой искривление траектории движения носителей заряда под влиянием кулоновской силы, действующей на электрон или дырку со стороны рассеивающего центра. Если таким центром является положительный донорный ион, то рассеивающая сила, описываемая формулой (1.14), вызывает отклонение электрона и дырки от исходного направления движения в электрическом поле ионизированного донора (рис. 3.6). Предположим, что электрон начинает свое движение из бесконечности со скоростью v по прямолинейной траектории, проходящей на расстоянии b от кулоновского центра рассеяния. Расстояние b называется прицельным. На рисунке также показана траектория дырки, начавшей движение с той же скоростью и на том же прицельном расстоянии. Носители заряда в общем случае пролетают на различных расстояниях от заряженных центров и отклоняются на разные углы. Поэтому рассеяние по рассматриваемому механизму столь же хаотично, как и рассеяние на тепловых колебаниях атомов решетки.

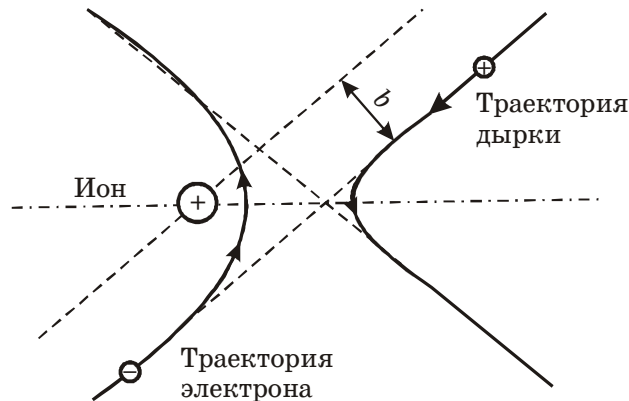


Рис. 3.6. Траектория движения электрона и дырки в электрическом поле ионизированного донорного атома

Рассеяние на ионах примеси возникает в результате взаимодействия зарядов. Отклонение носителей заряда увеличивается с уменьшением скорости движения и прицельного расстояния. Математически отклонение рассчитывают путем определения траектории движения заряженной частицы в центральном поле иона, подобно задаче о рассеянии α -частиц на ионах, которая решена Э. Резерфордом. Формула для длины свободного пробега при рассеянии электронов на ионах примеси имеет следующий вид:

$$l \approx v^4 m_n^* \left(\frac{\epsilon}{Zq} \right). \quad (3.27)$$

Все величины в выражении (3.27) использовались ранее. Из выражения (3.27) следует, что длина свободного пробега увеличивается и, следовательно, вероятность рассеяния уменьшается, если скорость движения, эффективная масса носителей заряда и диэлектрическая проницаемость кристалла возрастают. Физический смысл выражения (3.27) становится

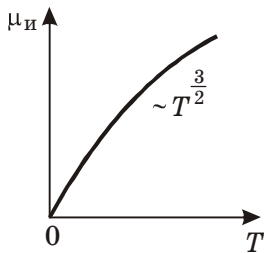


Рис. 3.7. Зависимость подвижности носителей заряда в полупроводнике от температуры (рассеяние на ионах примесей)

очевидным с позиций динамики материальной точки в классической механике. Математический анализ кинетических явлений в твердых телах показывает, что при рассеянии на ионах примесей подвижность носителей возрастает за счет увеличения скорости v в выражении (3.27). Для невырожденного полупроводника

$$\mu_{и} \sim T^{3/2}, \quad (3.28)$$

т.е. подвижность увеличивается с ростом температуры (рис. 3.7).

Рассеяние на ионах примесей преобладает над другими механизмами рассеяния в области температур около 30–70 К, т.е.

в условиях, когда тепловые колебания атомов в кристаллической решетке не ограничивают длину свободного пробега носителей заряда, но тепловая энергия достаточна для ионизации примесных атомов.

Смешанное рассеяние

В реальных кристаллах обычно наиболее эффективны два вида рассеяния: рассеяние на тепловых колебаниях и рассеяние на примесных ионах. Соответственно температурную зависимость подвижности в области низких температур часто можно объяснить рассеянием носителей заряда на ионах примеси, когда выражение для температурной зависимости имеет вид, подобный (3.28):

$$(3.29)$$

где α' — коэффициент пропорциональности.

В области высоких температур зависимость $\mu(T)$ определяется рассеянием носителей заряда на фононах, и ее можно описать выражением

$$\mu_{ф} = \beta' T^{-3/2}, \quad (3.30)$$

где β' — коэффициент пропорциональности.

При одновременном действии различных механизмов рассеяния можно считать, что числа столкновений носителей заряда, вызванных различными видами рассеяния, суммируются:

$$\frac{1}{l} = \frac{1}{l_{\text{ф}}} + \frac{1}{l_{\text{и}}} + \dots, \quad (3.31)$$

где $l_{\text{ф}}$ и $l_{\text{и}}$ — средние длины свободного пробега при рассеянии на фононах и ионах примеси.

В этом случае результирующая подвижность определяется как

$$(3.32)$$

или, подставляя (3.29) и (3.30),

$$(3.33)$$

Графический вид результирующей зависимости $\mu(T)$ можно представить, объединяя рис. 3.5, 3.7 или, что будет точнее, используя выражение (3.33). На рис. 3.8 представлены зависимости подвижности от температуры при смешанном механизме рассеяния для различных концентраций примесей доноров или акцепторов N . При малых температурах (меньше T_1, T_2, T_3), когда рассеяние на тепловых колебаниях решетки еще слабое, подвижность определяется рассеянием носителей на ионах примесей. С увеличением температуры роль рассеяния на фононах возрастает и, начиная со значений T_1, T_2 и T_3 , становится преобладающей. Подвижность при дальнейшем росте температуры уменьшается пропорционально $T^{-3/2}$.

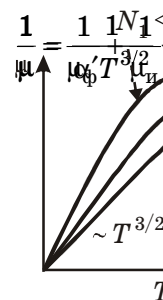


Рис. 3.8. Температурные зависимости подвижности носителей заряда в невырожденном полупроводнике при различных концентрациях примесных атомов донорного или акцепторного типа

Увеличение концентрации ионизированной примеси ($N_3 > N_2 > N_1$) приводит к тому, что рассеяние на тепловых колебаниях становится преобладающим при более высоких температурах ($T_3 > T_2 > T_1$). Увеличение концентрации примесей закономерно приводит к уменьшению подвиж-

ности. Как показано на рис. 3.8, рост величины N сопровождается снижением зависимости $\mu(T)$.

Реально в германии подвижность электронов в зоне проводимости при рассеянии на колебаниях решетки выражается формулой

$$\mu_n = 3800 \left(\frac{T}{300} \right)^{-1,66}, \text{ см}^2/(\text{В} \cdot \text{с}), \quad (3.34)$$

для дырок в валентной зоне справедливо выражение

$$\mu_p = 1800 \left(\frac{T}{300} \right)^{-2,33}, \text{ см}^2/(\text{В} \cdot \text{с}). \quad (3.35)$$

Из соотношений (3.34) и (3.35) следует, что при 300 К подвижности носителей заряда будут:

$$\begin{aligned} \mu_n &= 0,38 \text{ м}^2/(\text{В} \cdot \text{с}), \\ \mu_p &= 0,18 \text{ м}^2/(\text{В} \cdot \text{с}). \end{aligned}$$

Различие подвижностей электронов и дырок согласно формуле (3.23) можно объяснить их разной эффективной массой.

§ 3.4. Зависимость электропроводности полупроводников от температуры

В § 3.2 показано, что электропроводность (проводимость) полупроводника можно выразить формулой (3.15). Для рассмотрения ее температурной зависимости удобнее ограничиться собственным и донорным полупроводниками.

В полупроводнике любого типа проводимости зависимость $\sigma_0(T)$ определяется температурными зависимостями подвижности $\mu(T)$ и концентраций носителей заряда $n_i(T)$, $n(T)$ и $p(T)$. Ранее было показано, что в широком диапазоне температур подвижность определяется смешанным рассеянием носителей заряда на ионах примесей и на тепловых колебаниях решетки (фононах) и возрастает или уменьшается по степенному закону $\mu \propto T^{\pm 3/2}$.

Температурная зависимость концентрации носителей заряда в основном обусловлена интенсивностью перехода электронов из валентной зоны и с примесных уровней в зону проводимости. Как показано на рис. 2.5 и 2.7,б, зависимость $n(T)$ определяется шириной запрещенной зоны и глубиной залегания в ней примесного уровня, т.е. энергией активации, вследствие чего зависимости концентраций носителей заряда от температуры (2.29) и (2.30) являются экспоненциальными. Поэтому зависимости $N_c(T)$ и $N_v(T)$ — степенные функции — не проявляются в зависимостях $n_i(T)$ (рис. 2.5) и $n(T)$ (рис. 2.7,б).

Тогда, принимая во внимание вышеизложенное, выражение для проводимости собственного полупроводника будет иметь вид

$$\sigma_i = \sigma_{0i} e^{-\frac{\Delta E}{2k_0 T}}, \quad (3.36)$$

где σ_{0i} — величина, слабо зависящая от температуры, т.е.

Логарифмируя (3.36), получаем

$$\ln \sigma_i = \ln \sigma_{0i} - \frac{\Delta E}{2k_0 T}. \quad (3.37)$$

Отсюда следует, что зависимость $\ln \sigma_i (1/T)$ представляет прямую линию (рис. 3.9), что совпадает с зависимостью $\ln n_i (1/T)$, показанной на рис. 2.5. С помощью измеренной и построенной зависимостей $\ln \sigma_i (1/T)$ нетрудно рассчитать ширину запрещенной зоны полупроводника.

Действительно, как показано на рис. 3.9 и в уравнении (3.37),

$$\ln \sigma_{i1} = \ln \sigma_0 - \frac{\Delta E}{2k_0 T_1}, \quad (3.38)$$

$$\ln \sigma_{i2} = \ln \sigma_0 - \frac{\Delta E}{2k_0 T_2}. \quad (3.39)$$

Вычитая (3.38) из (3.39), получим

$$\ln \sigma_{i2} - \ln \sigma_{i1} = \frac{\Delta E}{2k_0} \left(\frac{1}{T_1} - \frac{1}{T_2} \right), \quad (3.40)$$

откуда следует важная с практической точки зрения формула для определения ΔE :

$$(3.41)$$

Здесь удобно использовать постоянную Больцмана

$$k_0 = 8,625 \cdot 10^{-5} \text{ эВ/К.}$$

Зависимость проводимости донорного полупроводника от температуры, исходя из (2.44), можно выразить формулой

$$\sigma_n(T) = \sigma_{0n} e^{-\frac{E_c - E_d}{k_0 T}}, \quad (3.42)$$

где $\sigma_{0n}(T) \approx \text{const.}$

Различие в величинах σ_{0i} и σ_{0n} можно установить, анализируя формулы (3.14) и (3.20) и соответственно подставляя в них выражения (2.44) и (2.35). При этом можно убедиться, что вид температурной зависимости

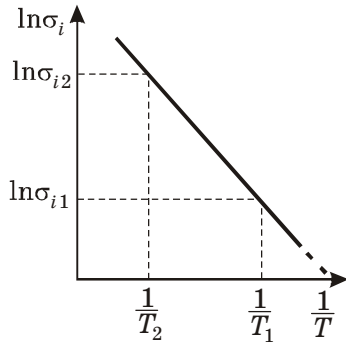


Рис. 3.9. Зависимость проводимости собственного полупроводника от температуры

$$\frac{\sigma_{0i}(T)}{\Delta E} \approx \frac{dn}{2k_0}$$

проводимости будет соответствовать виду температурной зависимости концентрации электронов (см. рис. 2.7,б). Однако необходимо заметить, что в области 2 (рис. 2.7,б), где $n_{n0}(T) = \text{const}$, на зависимости $\ln \sigma_n(1/T)$, представленной на рис. 3.10 (область 2), будет проявляться температурная зависимость подвижности при смешанном рассеянии (рис. 3.8). Объяснять зависимость $\ln \sigma(1/T)$ следует с использованием обсуждения кривой $\ln n(1/T)$, приведенного в § 2.5.

В области 1 (рис. 3.10) проводимость полупроводника n -типа обусловлена дрейфом электронов, поступающих в зону проводимости с донорного уровня. В связи с этим область 1 называется областью примесной проводимости. Согласно формуле (2.45) угол наклона α_n соответствует глубине залегания донорного уровня.

Начиная с температуры T_s и до T_i проявляется область 2, в которой все донорные атомы ионизированы, и электроны с донорных уровней в зону проводимости больше не поступают. Проводимость полупроводника в области 2 обусловлена участием в электрическом токе всех электронов с донорных атомов. Подвижность этих электронов и, следовательно, проводимость определяются рассеянием на донорных ионах (вблизи T_s), а затем на фононах (вблизи T_i). Этим объясняется, что в области 2, которая называется областью истощения, примесная зависимость $\ln \sigma(1/T)$ проходит через максимум.

В области 3, начиная с температуры T_i , концентрация собственных электронов в зоне проводимости, поступивших из валентной зоны, становится преобладающей по сравнению с примесными электронами, и область 3 поэтому называется областью собственной проводимости. Угол наклона α_i позволяет в соответствии с выражениями (2.39) и (3.41) определить ширину запрещенной зоны данного полупроводника n -типа. Из вышеизложенного следует, что для области 3 в проводимости как донорного, так и акцепторного полупроводников участвуют электронно-дырочные пары.

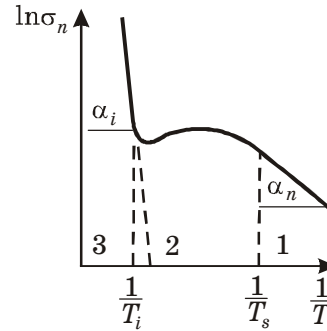


Рис. 3.10. Зависимость проводимости примесного полупроводника от температуры

§ 3.5. Электропроводность металлов и сплавов

В металлах в электрическом токе участвуют только свободные электроны, и представления о дырке для этого класса твердых тел не требуется. В связи с этим электропроводность металлов и полупроводников сильно выраженного n -типа описывается одной формулой (3.14). Так как электронный газ в металлах вырожден, то концентрация электронов

не зависит от температуры. Тогда зависимость $\sigma(T)$ определяется только влиянием температуры на подвижность носителей заряда, которая связана с рассеянием на тепловых колебаниях атомов кристаллической решетки, т.е. на фононах. Концентрация фононов увеличивается, как отмечалось выше, с ростом температуры пропорционально ее величине, и соответственно длина свободного пробега l будет обратно пропорциональна температуре. В электрическом токе в металлах участвуют электроны, располагающиеся вблизи уровня Ферми. Поэтому их скорость v_d также не зависит от температуры. Тогда в соответствии с выражением (3.23) можно получить

$$\sigma \propto T^{-1}. \quad (3.43)$$

и, следовательно,

$$\sigma \propto T^{-1}. \quad (3.44)$$

Удельное сопротивление, которое определяется формулой

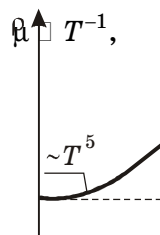
$$\rho = \frac{1}{\sigma}, \quad (3.45)$$

следовательно, в области высоких температур будет увеличиваться пропорционально температуре (рис. 3.11).

При низких температурах вследствие малой концентрации фононов удельное сопротивление пропорционально температуре в пятой степени, так как в этой области в процесс рассеяния включаются многократные взаимодействия электронов с фононами.

Рис. 3.11. Зависимость удельного сопротивления металла от температуры

При температурах, близких к абсолютному нулю, концентрация фононов мала и уже не определяет процесс рассеяния. Основное значение в области низких температур имеет рассеяние на примесных атомах, которые практически всегда остаются в металлах высокой степени чистоты, а также на точечных дефектах кристаллической решетки. В этом случае, как отмечалось выше, $v_d(T) = \text{const}$, удельное сопротивление не зависит от температуры и равно $\rho_{\text{ост}}$. Величина $\rho_{\text{ост}}$ называется остаточным сопротивлением.



Контрольные вопросы и задачи

1. (ВБЭЯ). Какую траекторию движения имеют свободные носители заряда в полупроводнике в отсутствие и при наличии внешнего электрического поля? Выберите правильные ответы:

1) в отсутствие внешнего электрического поля носители заряда совершают хаотическое тепловое движение вследствие их столкновений друг с другом, с атомами и с дефектами в кристалле;

2) во внешнем электрическом поле хаотическое движение электронов превращается в их прямолинейное ускоренное перемещение против вектора напряженности E ;

3) электроны при включенном и при выключенном внешнем электрическом поле совершают хаотическое движение. Однако в электрическом поле в беспорядочном движении электронов можно выделить результирующую компоненту скорости, направленную строго против вектора E , что соответствует понятию дрейфа электронов против электрического поля, а дырок по направлению поля;

4) в электрическом поле дрейфовая скорость движения электронов против поля возрастает до бесконечности под действием внешней силы $F = -qE$;

5) в электрическом поле скорость электронов увеличивается до тех пор, пока сила, действующая со стороны поля, не сравняется с силой сопротивления со стороны атомов, совершающих тепловые колебания, других электронов и дефектов кристаллической решетки, после чего электроны приобретают направленное движение с постоянной скоростью дрейфа.

2. (КБЭТ). Укажите выражения, которые описывают закон Ома в раз-

ных полупроводниках:

- 1) $\mu_n = \frac{qE}{m_n^* v}$ 2) $\mu_p = \frac{qE}{m_p^* v}$
 3) $\mu_n = \frac{qE}{m_n^* v}$ 4) $\mu_p = \frac{qE}{m_p^* v}$
 5) $\mu_n = \frac{qE}{m_n^* v}$

где σ — удельная электрическая проводимость (электропроводность); j — плотность электрического тока; μ_n и μ_p — подвижности электронов и дырок; n_0 , p_0 — равновесные концентрации электронов и дырок.

3. (ПШМЮ). Укажите правильные выражения для подвижности свободных носителей заряда:

- 1) $\mu_n = \frac{\sigma_n}{qn}$ 2) $\mu_n = \frac{\sigma_n}{qn}$ 3) $\mu_p = \frac{q}{m_p^* v}$
 4) $\mu_p = \frac{q\tau}{m_p^*}$ 5) $\mu_p = \frac{\sigma_p}{qp}$ 6) $\mu_n = \frac{q\tau}{m_n^*}$

4. (КБСП). Укажите правильный физический смысл величин в формуле

$$l = v\tau.$$

Ответы:

- 1) l — длина кристалла;

- 2) l — средняя длина свободного пробега носителя заряда;
 - 3) v — средняя скорость движения носителя заряда;
 - 4) τ — среднее время движения носителя заряда между двумя последовательными актами рассеяния;
 - 5) τ — время пролета носителя заряда через кристалл.
5. (ЮФДС). Укажите физический смысл величин в формуле

$$v_d = \mu E.$$

Ответы:

- 1) v_d — скорость направленного движения носителя заряда при условии, что кристалл находится в электрическом поле;
- 2) v_d — скорость дрейфа носителя заряда;
- 3) — энергия носителя заряда;
- 4) — напряженность электрического поля;
- 5) μ — дрейфовая скорость носителя заряда в электрическом поле единичной напряженности;
- 6) μ — подвижность носителя заряда.

6. (АЮПД). Укажите физический смысл величин в формуле

Ответы:

- 1) μ — подвижность носителей заряда;
- 2) m^* — эффективная масса носителя заряда;
- 3) m^* — гравитационная масса (вес) носителя заряда;
- 4) q — электрический заряд носителя ($q = 1,6 \cdot 10^{-19}$ Кл);
- 5) τ — время релаксации носителей заряда;
- 6) τ — время затухания направленного потока носителей заряда после выключения внешнего электрического поля;
- 7) τ — время, за которое скорость направленного движения носителей заряда уменьшается в $e \approx 2,7$ раз.

7. (КНТЗ). Как подвижность носителей заряда в полупроводнике зависит от его температуры?

Ответы:

1) в области низких температур, когда концентрация фононов мала, а примесные атомы уже ионизированы, электроны и дырки рассеиваются на примесных ионах, т.е. траектория движения носителей заряда искривляется за счет кулоновского взаимодействия с донорными или акцепторными ионами. При повышении температуры полупроводника возрастает тепловая скорость носителей заряда и уменьшается время их взаимодействия с ионами и отклонение от направленного движения. Вследствие вышеизложенного подвижность свободных носителей заряда в области низких температур возрастает при повышении температуры: $\mu \propto T^{3/2}$;

2) при малой концентрации фононов в области низких температур подвижность остается постоянной при увеличении температуры;

$$\mu = \frac{q\tau}{m^*}.$$

3) в реальных полупроводниках в области низких температур зависимость $\mu(T)$ объясняется рассеянием носителей заряда на ионах примесей и имеет вид $\mu_{и} = \alpha' T^{3/2}$, где α' — коэффициент пропорциональности;

4) в области высоких температур основное влияние на подвижность носителей заряда оказывает рассеяние на фононах. Температурные зависимости средней длины свободного пробега и средней скорости теплового движения носителей заряда являются причиной уменьшения их подвижности с ростом температуры пропорционально $T^{-3/2}$.

8. (Д2МР). Укажите правильное описание процесса рассеяния носителей заряда в реальных полупроводниках в областях низких и высоких температур.

Ответы:

1) в реальных полупроводниках наблюдается смешанный механизм рассеяния на ионах примесей и на фононах. Результирующая подвижность определяется выражением

$$\frac{1}{\mu} = \frac{1}{\alpha' T^{3/2}} + \frac{1}{\beta' T^{-3/2}};$$

2) в реальных полупроводниках рассеяние носителей заряда на ионах при повышении температуры приводит к уменьшению подвижности, тогда как при рассеянии на фононах подвижность только увеличивается;

3) результирующая зависимость $\mu(T)$ в реальных полупроводниках имеет вид колоколообразной кривой, что говорит о конкурирующем характере двух видов рассеяния — на ионах в области низких и на фононах в области высоких температур;

4) увеличение концентрации примесей сопровождается снижением максимальной подвижности на зависимости $\mu(T)$ и смещением максимума в область более высоких температур, что говорит о преобладании ионного рассеяния над фононным в более широком температурном диапазоне.

9. (ЛЯЭП). Проанализируйте и объясните причины температурной зависимости электропроводности полупроводников. Укажите правильные ответы.

Ответы:

1) зависимость $\sigma(T)$ в общем случае объясняется температурными зависимостями концентраций свободных носителей заряда и их подвижностей;

2) зависимость концентрации носителей заряда от температуры в области примесной проводимости связана с тепловыми перебросами носителей с примесных уровней в соответствующие разрешенные зоны и имеет экспоненциальный характер:

$$\sigma_n(T) = \sigma_{0n} e^{-\frac{E_c - E_d}{k_0 T}};$$

3) в области собственной проводимости зависимость $\sigma(T)$ имеет вид

$$\sigma_i(T) = \sigma_{0i} e^{-\frac{\Delta E}{k_0 T}}$$

и объясняется температурной зависимостью концентрации собственных носителей заряда, поступающих в результате их тепловых перебросов из валентной зоны в зону проводимости;

4) в области истощения примеси концентрация свободных носителей заряда не зависит от температуры. Зависимость $\sigma(T)$ в этой области связана только с рассеянием носителей заряда на ионах и на фононах;

5) зависимость электропроводности полупроводников и металлов от температуры обусловлена только влиянием температуры на подвижность носителей заряда;

6) температурная зависимость электропроводности металлов связана только с влиянием температуры на подвижность электронов, так как электронный газ в металлах вырожден и концентрация носителей заряда $n(T) = \text{const}$.

10. (ХЛЭП). Какие параметры полупроводников можно определить из температурной зависимости их электропроводности?

Ответы:

- 1) подвижность носителей заряда;
- 2) концентрацию дырок и электронов;
- 3) удельное сопротивление при заданной температуре;
- 4) ширину запрещенной зоны из зависимости $\ln \sigma(1/T)$ в области собственной проводимости.

Глава 4. Электропроводность полупроводников в сильном электрическом поле

§ 4.1. Влияние электрического поля на электропроводность полупроводников

Электропроводность полупроводников в общем случае определяется концентрацией основных и неосновных носителей заряда и их подвижностью (выражения (3.18) и (3.19)).

Температурные зависимости концентрации, подвижности, электропроводности в слабых электрических полях рассмотрены в гл. 2 и 3.

Экспериментальная зависимость электропроводности от напряженности электрического поля показана на рис. 4.1. В области слабых полей она представляет горизонтальную прямую, что говорит о постоянстве коэффициента σ в формуле (3.8). Это означает, что закон Ома выполняется только в области слабых полей, то есть до некоторой критической величины $E_{кр}$. Рост электропроводности в области сильных полей объясняется:

- 1) возникновением дополнительных носителей, что обусловлено их генерацией в сильных электрических полях;
- 2) изменением подвижности носителей заряда в электрических полях.

Величину Δv можно оценить по изменению скорости носителей заряда в сильном электрическом поле. Если Δv — некоторая добавка к скорости за счет влияния поля, то условием слабого электрического поля будет

$$\Delta v \ll v_T. \quad (4.1)$$

Так как величина тепловой скорости $v_T \approx 10^5$ м/с, то условие (4.1) в соответствии с формулой (3.12) выполняется до критического поля $\approx 10^4$ В/см.

Сильное электрическое поле вызывает повышение концентрации носителей заряда. Рассмотрим три механизма этого явления: термоэлектронную ионизацию Френкеля, электростатическую ионизацию Зиннера (туннельный эффект) и ударную ионизацию.

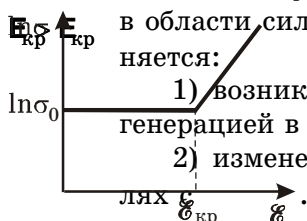


Рис. 4.1. Зависимость проводимости полупроводника от напряженности электрического поля

§ 4.2. Термоэлектронная ионизация Френкеля

В сильном электрическом поле изменяется форма зависимости потенциальной энергии электронов в атомах (см. рис. 1.4) от расстояния между ними x , как показано на рис. 4.2 для одного атома. В электрическом поле с напряженностью \mathcal{E} на периодический потенциал $U(x)$ накладывается дополнительный потенциал

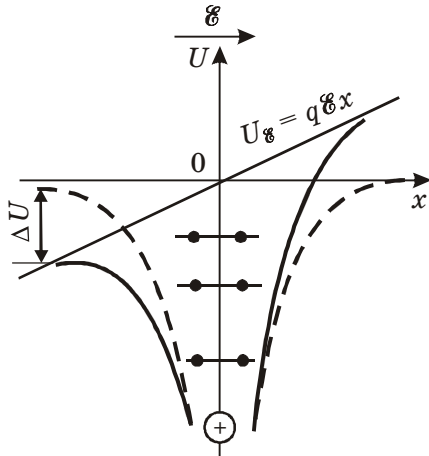


Рис. 4.2. Изменение формы потенциального барьера для электрона в примесном атоме при наложении внешнего электрического поля

и потенциальный барьер принимает вид, показанный на рис. 4.2 сплошной линией. В соответствии с направлением вектора напряженности электрического поля левый потенциальный барьер снизится на величину ΔU .

Если потенциальная яма на рис. 4.2 относится к донорному атому в полупроводнике n -типа, то концентрация электронов согласно формуле (2.44) будет определяться выражением

$$(4.2)$$

т.е. должна возрасти за счет снижения энергии ионизации донорного атома на величину ΔU . Это очевидно, если выражение (4.2) переписать с учетом (2.44) в виде

$$n = n_{n0} e^{-\frac{\Delta E}{2k_0 T}}. \quad (4.3)$$

Таким образом, термоэлектронная ионизация заключается в снижении потенциального барьера, что облегчает тепловую ионизацию примесных атомов. Эффект наблюдается в области температур ниже температуры истощения полупроводника $T < T_s$ (§ 2.5). Заметный рост концентрации носителей заряда за счет термоэлектронной ионизации происходит при $E = 10^4 - 10^5$ В/см.

§ 4.3. Ударная ионизация

Свободные носители заряда в разрешенных зонах полупроводника при движении в кристалле под влиянием электрического поля приобретают энергию, величина которой определяется формулой (1.40). Свободные носители, разгоняясь при движении в кристалле, сталкиваются с атомами решетки и теряют часть энергии, накопленной на длине сво-

$$n = \left(\frac{qEx}{N_c N_d} \right)^2$$

бодного пробега. Если напряженность электрического поля настолько велика, что электроны за счет его влияния, ускоряясь, приобретают энергию, равную ширине запрещенной зоны, то при соударении с основными атомами кристаллической решетки происходит выбивание дополнительных электронов (рис. 4.3). На энергетической диаграмме этому соответствует переброс электрона из валентной зоны в зону проводимости.

Описанный процесс называется ударной ионизацией атомов полупроводника. Ионизацию могут вызвать и дырки, так как движение дырок со скоростью v_p является лишь способом описания движения совокупности электронов в валентной зоне полупроводника.

Таким образом, при ударной ионизации электрон ионизирует атом, с которым он сталкивается в достаточно сильном электрическом поле, в результате чего в зону проводимости переходит еще один электрон и далее будут ускоряться уже два электрона. Эти электроны при соответствующих условиях ионизируют следующие два атома, что приводит к ускорению четырех электронов, и, таким образом, количество свободных электронов будет увеличиваться по геометрической прогрессии, т.е. лавинообразно.

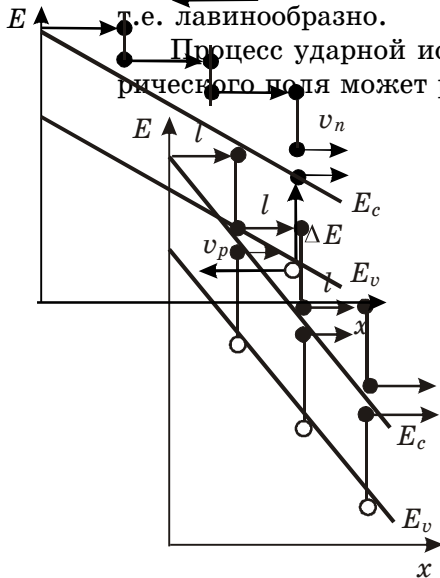


Рис. 4.4. Диаграмма высоковольтной ионизации

Рис. 4.3. Зонная диаграмма полупроводника в сильном электрическом поле

Процесс ударной ионизации в зависимости от напряженности электрического поля может развиваться разными путями. В полях низкой напряженности электрон на длине свободного пробега приобретает энергию, которая меньше энергии ионизации атома ΔE . Однако если при столкновении электрон теряет меньше энергии, чем приобретает на длине свободного пробега, то скорость его будет непрерывно увеличиваться и энергия электрона становится равной ширине запрещенной зоны и достаточной для ионизации собственных атомов полупроводника, как показано на рис. 4.3. Такую ударную ионизацию называют низковольтной.

В полях высокой напряженности электрон может приобрести энергию, необходимую для ионизации атома, уже на длине свободного пробега (рис. 4.4). Такую ударную ионизацию называют высоковольтной.

Количественно процесс ударной ионизации характеризуется коэффициентом ударной ионизации α , который численно равен количеству пар носителей заряда, образуемых первичным носителем на единице пути, и сильно зависит от напряженности электрического поля.

Если электрон набирает энергию, равную ΔE , на одной длине свободного пробега, то зависимость подчиняется экспоненциальному закону:

$$(4.4)$$

На практике часто пользуются эмпирической аппроксимацией

$$\alpha = AE^m, \quad (4.5)$$

где m — показатель степени, который в различных материалах изменяется от 5 до 8; A — константа.

Явление ударной ионизации используется на практике в p - n -переходах, в частности в стабилитронах, где можно получить электрические поля величиной 10^5 – 10^6 В/см без электрического пробоя полупроводника. Ударная ионизация донорных и акцепторных атомов в германии наблюдается в полях с В/см, что связано с незначительной глубиной залегания соответствующих уровней в запрещенной зоне (см. § 1.11 и § 1.12).

§ 4.4. Электростатическая ионизация

Образование свободных электронно-дырочных пар в полупроводнике возможно только при переходе электронов из валентной зоны в зону проводимости, на что требуется энергия не менее ΔE .

Ранее при объяснении явлений в полупроводниках предполагалось, что края энергетических зон занимают горизонтальное положение (рис. 4.5,а), что можно считать справедливым только для случая слабого электрического поля. В кристалле, который находится в сильном электрическом поле, края энергетических зон будут проходить под углом α к горизонтали (рис. 4.5,б). Природа наклона зависимости $E(x)$ рассматривалась выше (§ 4.2).

При $\alpha = 0$ (рис. 4.5,а) возможны только вертикальные переходы электронов из валентной зоны в зону проводимости. Когда $\alpha > 0$, появляется возможность горизонтального перехода AB , как показано на рис. 4.5,б. При таких межзонных переходах энергия электронов не изменяется, т.е. переход происходит без энергетических затрат. Электрон как бы проходит по траектории AB сквозь энергетический барьер треугольной формы ABC . Это явление в физике называется туннельным эффектом.

Вероятность туннелирования сквозь барьер характеризуется коэффициентом прозрачности D .

Из квантовой механики известно, что коэффициент прозрачности определяется формулой

$$\alpha(E) \approx e^{-\frac{\Delta E}{qEL}}$$

(4.6)

Из (4.6) следует, что вероятность межзонных туннельных переходов электронов определяется шириной запрещенной зоны. Однако, ввиду того что концентрация электронов в валентной зоне больше, направление туннелирования устанавливается из валентной зоны в зону проводимости. В полупроводнике, помещенном в электрическое поле, угол наклона энергетических зон α_b увеличивается с ростом напряженности электрического поля до α'_b , вследствие чего ширина барьера (AB , рис. 4.5,б) уменьшается до AB' и вероятность туннелирования, как следует из (4.6), возрастает.



§ 4.5. Подвижность носителей заряда в сильном электрическом поле

Подвижность носителей заряда определяется формулами (3.21) и (3.23), которые справедливы для случаев слабых электрических полей, когда выполняется неравенство

$$v_T \gg v_d. \quad (4.7)$$

Действительно, средняя тепловая скорость ($v_T = 10^5$ м/с) значительно превышает среднюю скорость дрейфа носителей заряда, которая в слабых полях достигает 10^4 м/с. В общем случае время релаксации в формуле (3.21) зависит от скоростей теплового и дрейфового движения:

$$(4.8)$$

Тогда в слабых полях скорость электрона может быть представлена только тепловой составляющей скорости, и согласно (3.23) $\mu(E) = \text{const}$. Подобное объяснение справедливо и для дырки.

В сильных электрических полях, когда напряженность выше критической величины, дрейфовая скорость увеличивается в соответствии с формулой (3.12), и при анализе процессов рассеяния пренебрежение вторым членом в знаменателе (4.8) оказывается недопустимым. Следовательно, становится очевидным, что подвижность носителей заряда должна изменяться в сильных электрических полях.

Эксперименты показывают, что при рассеянии носителей на тепловых колебаниях атомной решетки подвижность будет уменьшаться:

$$(4.9)$$

При рассеянии на примесных ионах

$$\mu \propto E^{3/2}, \quad (4.10)$$

т.е. наблюдается рост подвижности, обусловленный сокращением времени взаимодействия носителя с кулоновским полем примесного иона.

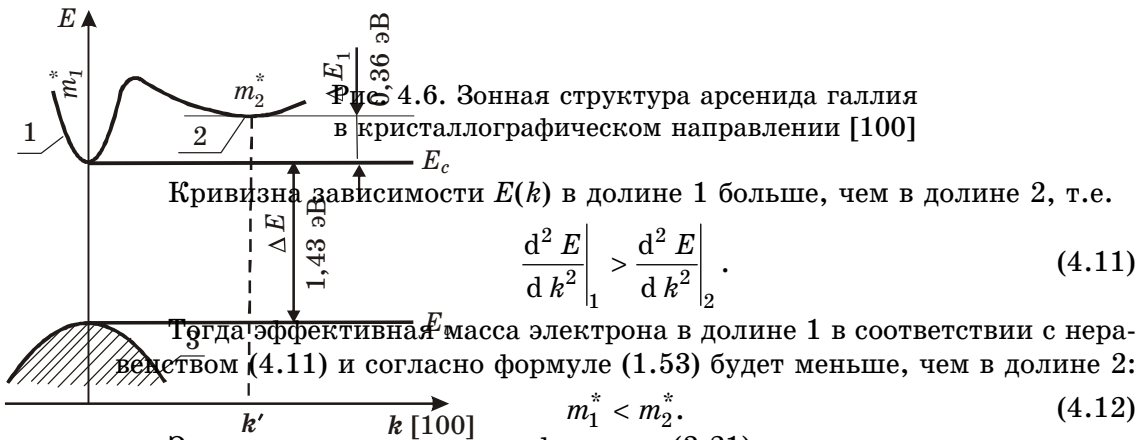
Изменения подвижности при $E > E_{кр}$ сильнее проявляются в полупроводниках с большой подвижностью носителей. Так, для Ge *n*-типа, в котором подвижность электронов $\mu_n \approx 3800$ см²/(В · с), отклонение от закона Ома наблюдается при $E = 400$ В/см, тогда как в Ge *p*-типа, в котором подвижность дырок ниже ($\mu_p = 1800$ см²/(В · с)), величина $E_{кр}$ больше ($E_{кр} = 1900$ В/см).

$$\frac{E}{E_{кр}} \gg \frac{E_{кр}^{-1/2}}{v_T + v_d}$$

§ 4.6. Эффект междолинного рассеяния электронов

Рассмотрим эффект междолинного рассеяния электронов, который связан с их междолинным переходом.

В некоторых полупроводниках, например GaAs, InP и др., зона проводимости, если рассматривать зависимость $E(k)$, имеет более одного минимума, называемого долиной. На рис. 4.6 показана структура энергетических зон арсенида галлия n -типа в зависимости от волнового вектора k в кристаллографическом направлении [100]. Цифрами 1 и 2 обозначены соответственно первая и вторая долины зоны проводимости, 3 — валентная зона.



Это неравенство с учетом формулы (3.21) приводит к различию подвижностей электронов в рассматриваемых долинах. Действительно, экспериментальные данные показывают, что эффективной массе электрона в долине 1 $m_1^* = 0,072m_0$, где m_0 — масса свободного электрона, соответствует подвижность $\mu_1 = (5-8) \cdot 10^3 \text{ см}^2/(\text{В} \cdot \text{с})$. В долине 2 более «тяжелому» электрону ($m_2^* = 1,2m_0$) соответствует значительно меньшая подвижность $\mu_2 = 100 \text{ см}^2/(\text{В} \cdot \text{с})$, т.е.

$$\mu_2 \ll \mu_1. \quad (4.13)$$

Различие между подвижностями электронов в долинах 1 и 2 лежит в основе механизма междолинного рассеяния электронов, которое отражается на электропроводности вышеуказанных полупроводников в сильных электрических полях.

§ 4.7. Электропроводность двухдолинных полупроводников в сильном электрическом поле

В слабых электрических полях ($E < E_{кр}$) электроны находятся в термодинамическом равновесии с решеткой полупроводника и занимают энергетические уровни вблизи дна зоны проводимости, т.е. находятся в долине 1. При этом условии можно считать, что в долине 2 свободные электроны отсутствуют (рис. 4.7,а). Тогда зависимость тока через образец *n*-GaAs подчиняется закону Ома (3.8) и описывается выражением

$$(4.14)$$

где $n_0 = n_1$, т.е. плотность тока определяется концентрацией электронов n_1 в долине 1, их подвижностью μ_1 и линейно возрастает с увеличением напряженности поля (рис. 4.8, область 1).

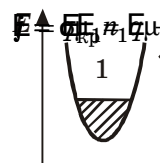


Рис. 4.7. Заполнение электронами первой и второй долин зоны проводимости арсенида галлия в слабом (а), среднем (б), сильном (в) электрическом поле

При напряженности электрического поля некоторые свободные электроны, получив дополнительную энергию больше ΔE_1 (см. рис. 4.6), оказываются в долине 2, где их концентрация становится равной n_2 (рис. 4.7,б).

Переход части электронов из долины 1 в долину 2 сопровождается уменьшением их подвижности, в то время как скорость тепловой генерации остается неизменной, т.е. выполняется равенство

$$n_1 + n_2 = n_0 . \quad (4.15)$$

В электрическом токе участвуют n_1 «легких» электронов и n_2 — «тяжелых», и плотность тока определяется выражением

$$(4.16)$$

Скорость роста тока снижается (рис. 4.8, область 2а). По мере увеличения напряженности электрического поля концентрация «тяжелых» электронов в долине 2 растет, а «легких» в долине 1 на такую же величину уменьшается.

Перечисленные процессы приводят к уменьшению плотности электрического тока с ростом напряженности электрического поля, т.е. наблюдается участок отрицательного дифференциального сопротивления (рис. 4.8, область 2б).

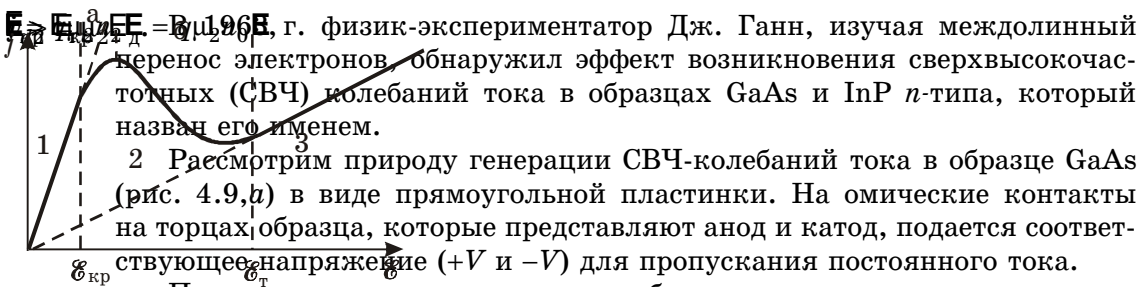
Рис. 4.8. Вольт-амперная характеристика арсенида галлия

При напряженности $E_{кр}$ все электроны находятся в долине 2 (рис. 4.7, в), и плотность тока равна

$$(4.17)$$

т.е. на участке 3 (рис. 4.8), как и на участке 1, снова выполняется закон Ома, но угол наклона зависимости $j(E)$ вследствие неравенства (4.13) будет меньше.

§ 4.8. Эффект Ганна



В 1966 г. физик-экспериментатор Дж. Ганн, изучая междолинный перенос электронов, обнаружил эффект возникновения сверхвысокочастотных (СВЧ) колебаний тока в образцах GaAs и InP *n*-типа, который назвал его именем.

Рассмотрим природу генерации СВЧ-колебаний тока в образце GaAs (рис. 4.9, а) в виде прямоугольной пластинки. На омические контакты на торцах образца, которые представляют анод и катод, подается соответствующее напряжение $(+V$ и $-V)$ для пропускания постоянного тока.

Полупроводниковые материалы обычно неоднородны по удельному сопротивлению. Пусть в образце на расстоянии x_1 от катода существует микрообласть повышенного сопротивления шириной Δx . Напряженность электрического поля в этой области будет выше, чем в других частях образца, и при повышении внешнего напряжения электрическое поле именно здесь в первую очередь достигнет значения $E_{кр}$ (рис. 4.9, б), при котором начинается междолинный переход электронов (§ 4.7).

Вследствие перехода электронов из долины 1 в долину 2 при уменьшения подвижности электронов в области Δx скорость их перемещения к аноду будет

$$(4.18)$$

Справа и слева от области Δx совершают дрейф электроны с более высокой скоростью

$$(4.19)$$

Таким образом, в зоне Δx двигаются по направлению к аноду «тяжелые» электроны со скоростью v_{d2} , величина которой будет меньше, чем

скорость перемещения «легких» электронов v_{d1} в частях образца справа и слева от области Δx . В результате «легкие» электроны слева догоняют «тяжелые» электроны, а электроны справа уходят вперед от области «тяжелых» электронов, вследствие чего на участке от x_1 до x_2 создается двойной слой противоположных зарядов (рис. 4.9, в). Левая часть слоя имеет повышенную концентрацию электронов и заряжена отрицательно. К нему справа примыкает положительно заряженный слой, обедненный электронами. Эта область двойного заряда называется электростатическим доменом.

Рис. 4.9. Образование доменов в диодах Ганна: *a* — конструкция диода; *б* — распределение напряженности электрического поля в диоде; *в* — распределение концентрации носителей заряда

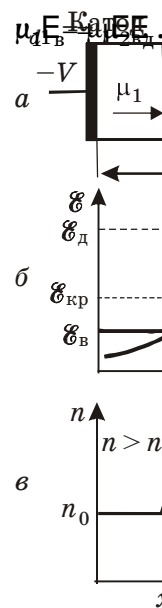
Очень быстрое возникновение домена ($t \approx 10^{-12}$ с) сопровождается увеличением сопротивления в его области. Тогда в этой части напряженность поля возрастает, а за областью домена — снижается. Происходит перераспределение электрического поля, которое за-

канчивается выравниванием скоростей дрейфа электронов внутри и вне домена, т.е. установлением равновесия

$$(4.20)$$

Установившаяся скорость дрейфа электронов в домене будет ниже максимальной скорости дрейфа равновесных электронов и фактически представляет скорость движения домена к аноду. Ширина домена $x_2 - x_1$ не превышает нескольких десятков микрон, тогда как на рис. 4.9, в она соизмерима с длиной образца L .

Таким образом, в момент включения внешнего напряжения ($t = t_0$, рис. 4.10) через образец должен протекать ток j_{\max} , величина которого определяется формулой (4.14). Однако у катода, где может находиться высокоомный приконтактный слой, возникает домен, и ток практически



мгновенно падает до j_{\min} . В течение времени домен дрейфует к аноду, и ток через образец не изменяется, т.е. остается равным j_{\min} .

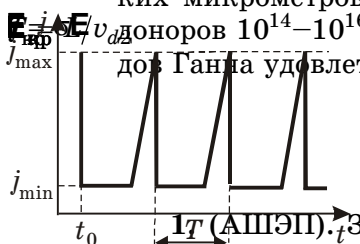
Напряженность электрического поля вне домена E_v меньше критической величины ($E_v < E_{cr}$, рис. 4.9,б), вследствие чего два домена не могут образоваться одновременно. На аноде домен разрушается, ток в цепи возрастает до величины j_{\max} , пока вследствие сопутствующего роста напряженности электрического поля до величины E_{cr} не возникнет новый домен и далее повторится процесс его дрейфа через весь образец длиной L до разрушения на аноде.

Таким образом, частота колебаний тока $f = 1/T$ зависит от длины образца, подвижности «тяжелых» электронов и может достигать десятков гигагерц.

Приборы на основе эффекта Ганна применяются в СВЧ-генераторах и усилителях, быстродействующих цифровых и аналоговых устройствах и называются диодами Ганна.

Для изготовления диодов Ганна необходимы чрезвычайно чистые и однородные полупроводниковые кристаллы с малой длиной (от нескольких микрометров до нескольких сотен микрометров), с концентрацией до $10^{14} - 10^{16} \text{ см}^{-3}$. Таким условиям в технологии производства диодов Ганна удовлетворяют полупроводниковые эпитаксиальные пленки.

Рис. 4.10. Зависимость тока от времени для диода Ганна



Контрольные вопросы и задачи

Зависит ли электропроводность полупроводников от напряженности электрического поля? Укажите правильные ответы.

Ответы:

- 1) электропроводность полупроводников σ зависит только от температуры и не зависит от напряженности электрического поля;
- 2) электропроводность полупроводников не зависит от напряженности электрического поля только в слабых полях;
- 3) до критической напряженности электрического поля $\sigma(E) = \text{const}$ и выполняется закон Ома $j = \sigma E$;
- 4) при $E = E_{cr}$ происходит электрический пробой, и полупроводник приобретает металлические свойства;
- 5) начиная с $E = E_{cr}$ электропроводность полупроводников возрастает за счет генерации дополнительных носителей заряда или изменяется за счет увеличения или уменьшения подвижности.

2. (А7ЭП). Почему концентрация носителей заряда в полупроводниках возрастает в сильных электрических полях?

Ответы:

1) в электрическом поле высокой напряженности повышается температура полупроводника за счет протекания электрического тока, что сопровождается тепловой ионизацией атомов кристаллической решетки;

2) во внешнем электрическом поле наблюдается термоэлектронная ионизация Френкеля, которая объясняется снижением потенциального барьера для электронов примесных атомов, вследствие чего тепловая ионизация происходит при пониженной температуре;

3) в электрическом поле электроны в зоне проводимости приобретают энергию, которая при соударении электрона с атомом кристаллической решетки достаточна для переброса нового электрона из валентной зоны в зону проводимости;

4) в сильном электрическом поле возможно туннелирование электронов из валентной зоны в зону проводимости без получения тепловой или световой энергии;

5) увеличение концентрации свободных носителей заряда в полупроводниках возможно только при поглощении квантов тепловой (фононов) или световой (фотонов) энергии. Поэтому если влияние электрического поля не сопровождается тепловыми или световыми эффектами, то увеличение концентрации носителей заряда в электрическом поле невозможно.

3. (Н7ПМ). Каким образом подвижность носителей заряда в полупроводниках связана с напряженностью внешнего электрического поля?

Ответы:

1) подвижность носителей заряда в полупроводниках не зависит от напряженности электрического поля, так как определяется только температурой;

2) зависимость подвижности от внешнего электрического поля объясняется соотношением средней тепловой v_T и дрейфовой v_d скоростей носителей заряда;

3) в слабых электрических полях ($E \ll E_{кр}$) в германии $v_d \ll v_T$, так как $v_d = \mu E \leq 10^4$ м/с, а $v_T \approx 10^5$ м/с. В этом случае подвижность носителей заряда будет определяться только тепловой скоростью, так как $\tau = l / (v_d + v_T)$. Таким образом, в слабых электрических полях $\mu(E) = \text{const}$;

4) в сильных электрических полях, когда $E > E_{кр}$, $v_d > v_T$, проявляется зависимость $\mu(E)$.

4. (ФЮМР). Какие из нижеприведенных положений необходимы для объяснения эффекта междолинного рассеяния?

Ответы:

1) в полупроводниках эффект междолинного рассеяния объясняется столкновениями электронов и дырок с фононами;

$E_{кр} q\tau/m^*$,

2) изменение кривизны зависимости $E(k)$ для зоны проводимости сложных полупроводников (GaAs, InP) рассматривается как существование долин, т.е. минимумов на кривой $E(k)$;

3) в долине 1 с большей кривизной зависимости $E(k)$ эффективная масса электронов меньше, чем в долине 2 с меньшей кривизной $E(k)$. Тогда электроны долины 1 называют «легкими», а долины 2 — «тяжелыми»;

4) в соответствии с формулой подвижность электронов в долинах 1 и 2 будет различной: $\mu_1 > \mu_2$;

5) подвижность электронов уменьшается при их переходе из долины 1 в долину 2.

5. (ММЭП). Выделите особенности электропроводности двухдолинных полупроводников:

1) в слабых электрических полях при $E < E_c$ электроны зоны проводимости находятся в долине 1. В этом случае выполняется закон Ома ($j \sim E$), так как подвижность всех «легких» электронов $\mu_1(E) = \text{const}$;

2) в электрическом поле при $E = E_c$ начинается переход электронов из долины 1 в долину 2. Эффективная масса электронов в долине 2 увеличивается, а подвижность уменьшается, вследствие чего замедляется скорость роста плотности тока при дальнейшем увеличении напряженности поля;

3) в электрических полях $E > E_c$ все электроны переходят в долину 2 одновременно, где их подвижность $\mu_2 = \text{const}$;

4) в электрических полях с $E > E_c$ количество электронов, переходящих из долины 1 в долину 2, возрастает в соответствии с ростом энергии, получаемой от электрического поля. Снижение скорости роста электрического тока сменяется его уменьшением при непрерывном увеличении напряженности электрического поля;

5) при $E = E_c$ все электроны оказываются «тяжелыми», т.е. перешедшими в долину 2, и в дальнейшем при увеличении E снова выполняется закон Ома ($j \sim E$). Однако наклон зависимости $j(E)$ будет значительно ниже, чем в области $E < E_c$.

6. (НОЭГ). Укажите пункты, которые объясняют физическую сущность эффекта Ганна:

1) эффект Ганна представляет собой изменение электропроводности полупроводника в высокочастотном магнитном поле;

2) эффект Ганна — это возникновение неоднородностей удельного сопротивления в полупроводнике при включении внешнего электрического поля;

3) эффект Ганна заключается в генерации СВЧ, т.е. сверхвысокочастотных колебаний электрического тока в полупроводнике, который находится в сильном электрическом поле;

4) в области неоднородностей удельного сопротивления на какой-либо микронеоднородности с максимальным сопротивлением, где обычно наибольшая напряженность электрического поля, возникают междолинные переходы электронов, что приводит к зарождению домена;

5) при наличии домена через полупроводник протекает минимальный электрический ток, так как в нем участвуют только «тяжелые» электроны;

6) домен разрушается при достижении им анода, и ток возрастает до мгновенного зарождения нового домена;

7) физические процессы, объясняющие генерацию СВЧ-колебаний электрического тока при увеличении напряженности электрического поля, можно разделить на следующие стадии: возникновение домена на соответствующей неоднородности — резкий спад тока — движение домена до анода — рост тока до зарождения нового домена.

7. (ЮВЭД). Укажите правильное определение электростатического домена в некоторых полупроводниках:

1) электростатический домен — это части объема полупроводника, самопроизвольно намагниченные до насыщения;

2) это двойной слой противоположных зарядов, возникающий в некоторых полупроводниках (GaAs, InP) в областях с микронеоднородностями удельного сопротивления вследствие междолинных переходов электронов при включении сильного электрического поля;

3) это микроскопическая область кристаллической решетки полупроводника с самопроизвольной ориентацией дипольных моментов молекул, входящих в его состав.

Глава 5. Генерация и рекомбинация носителей заряда

§ 5.1. Тепловая генерация

Как известно, при температуре абсолютного нуля в зоне проводимости полупроводника отсутствуют электроны, а в валентной зоне — свободные энергетические уровни — дырки (§ 1.6 и § 1.10). С повышением температуры (при $T > 0$) атомы кристаллической решетки совершают тепловые колебания. В процессе тепловых колебаний за счет взаимодействия атомы передают электронам часть тепловой энергии E . Если величина энергии, переданной донорному атому (в случае донорного полупроводника), соответствует соотношению

$$E \geq E_c - E_d, \quad (5.1)$$

то возникает так называемая генерация основных носителей заряда — электронов. Этот процесс представляет собой переход электронов с донорного уровня в зону проводимости (рис. 5.1,а) и называется монополярной генерацией электронов. В этом случае концентрация электронов описывается формулой (2.44).

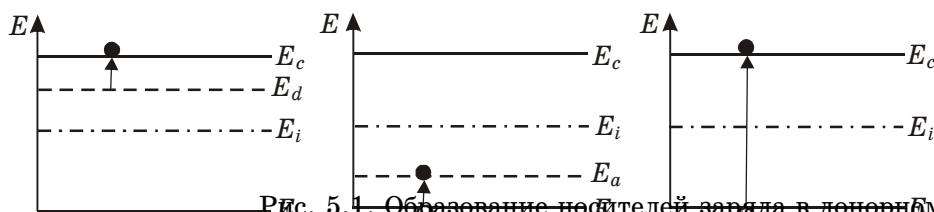


Рис. 5.1. Образование носителей заряда в донорном (а), акцепторном (б) и собственном (в) полупроводниках

В акцепторном полупроводнике электроны валентной зоны, получив тепловую энергию

$$E \geq E_a - E_v, \quad (5.2)$$

переходят на акцепторный уровень (рис. 5.1,б). Этот процесс называется монополярной тепловой генерацией дырок, в результате которой в полупроводнике p -типа генерируются основные свободные носители заряда — дырки, их концентрацию можно выразить формулой (2.48).

Средняя кинетическая энергия тепловых колебаний атомов в кристаллической решетке равна $(3/2)k_0T$. При комнатной температуре (300 К) эта величина составляет приблизительно 0,04 эВ, что соответствует условиям (5.1) и (5.2) и полной ионизации примесных атомов (§ 2.5), так как обычно энергия донорного (§ 1.11) и акцепторного (§ 1.12) уровней равна примерно 0,01 эВ.

Однако в общем случае средняя кинетическая энергия тепловых колебаний атомов при $T \approx 300$ К существенно меньше ширины запрещенной зоны ($\Delta E = E_c - E_v$). Но известно, что тепловая энергия неравномерно распределяется между частицами. В каждый момент времени имеется небольшое число атомов кристаллической решетки, у которых амплитуда и энергия тепловых колебаний значительно превышают средние значения. Именно за счет таких тепловых флуктуаций некоторые электроны могут перейти из валентной зоны в зону проводимости. Для такого электронного перехода необходимо, чтобы энергия тепловых колебаний (флуктуация энергии) удовлетворяла условию

$$E \geq E_c - E_v . \quad (5.3)$$

В этом случае генерируются одинаковые концентрации электронов и дырок, и на температурных зависимостях электропроводности наблюдаются области собственной проводимости, как, например, на рис. 3.9 и 3.10. Описанный выше процесс представляет тепловую генерацию электронно-дырочных пар, или так называемую биполярную генерацию (рис. 5.1,е). Концентрация электронно-дырочных пар выражается формулой (2.35).

Процесс тепловой генерации характеризуется скоростью генерации G_0 , т.е. количеством электронно-дырочных пар, создаваемых в единице объема в единицу времени. Соответственно единица измерения величины G_0 обозначается $\text{м}^{-3} \cdot \text{с}^{-1}$.

Распределение этих носителей по энергетическим состояниям в разрешенных зонах описывается равновесной функцией $f_{\text{М-В}}(E)$ или $f_{\text{Ф-Д}}(E)$, так как они находятся в термодинамическом равновесии с кристаллической решеткой. Соответственно такие носители называются равновесными.

§ 5.2. Неравновесные процессы в полупроводниках

Тепловая генерация не является единственным механизмом возникновения свободных носителей заряда. Они могут создаваться при воздействии на полупроводник света, корпускулярных потоков нейтронов, протонов, α -частиц, космических лучей и электрического поля, как показано в §§ 4.2–4.4. В перечисленных случаях состояние термодинамического равновесия в полупроводнике нарушается, и закон действующих масс, представленный выражениями (2.54) и (2.55), не будет выполняться. Подвижные носители заряда уже не будут находиться в термодинамическом равновесии, так как энергия электронов увеличивается за счет фотонов, а тепловая энергия кристаллической решетки остается практически неизменной. Такие носители называются неравновесными носителями заряда, а их концентрации n и p — неравновесными концентрациями.

На практике во многих полупроводниковых приборах (фоторезисторы, фотодиоды и др.) явление оптической генерации неравновесных носителей заряда играет главную роль. Механизм оптической генерации подобен рассмотренным в § 5.1 видам тепловой генерации. Главное отличие заключается в том, что переходы носителей, показанные на рис. 5.1, осуществляются за счет поглощения квантов света (фотонов) с энергией $E_{\text{ф}} \geq E_{\text{и}}$ равной или превышающей энергию ионизации примесного уровня.

При оптической генерации неравновесные носители заряда могут иметь сразу после генерации энергию выше средней тепловой энергии равновесных частиц. В результате рассеяния на дефектах избыточная энергия неравновесных носителей заряда передается кристаллической решетке. Уже через 10^{-10} с неравновесные носители заряда не будут отличаться от равновесных. Неравновесные электроны, рассеивая свою избыточную энергию, «опускаются» ко дну зоны проводимости, а неравновесные дырки «поднимаются» к потолку валентной зоны.

Таким образом, энергия неравновесных носителей в разрешенных зонах непрерывно уменьшается, что заканчивается ее выравниванием с энергией равновесных носителей заряда. В этом случае концентрация электронов

$$n = n_0 + \Delta n, \quad (5.4)$$

а дырок

$$p = p_0 + \Delta p, \quad (5.5)$$

где Δn и Δp — концентрации избыточных (неравновесных) электронов и дырок. Для неравновесных концентраций, как отмечалось выше, закон действующих масс (равенства (2.54) и (2.55)) не выполняется. В связи с этим выделяют следующие случаи:

- 1) если отклонение положительное

$$\Delta n > \Delta p, \quad (5.6)$$

то имеет место инжекция (ввод) неравновесных (избыточных) носителей заряда;

- 2) если отклонение отрицательное

$$np < n_i^2, \quad (5.7)$$

то наблюдается экстракция (вытяжка) носителей.

При наличии неравновесных носителей заряда формулы для определения концентрации электронов в зоне проводимости (2.29) и дырок в валентной зоне (2.30) не применимы. Так как в стационарном состоянии свойства избыточных носителей практически не отличаются от свойств равновесных носителей, то их концентрацию в неравновесных условиях удобно выражать формулами, подобными (2.29) и (2.30):

$$n = N_c \exp\left(-\frac{E_c - E_{F_n}}{k_0 T}\right), \quad (5.8)$$

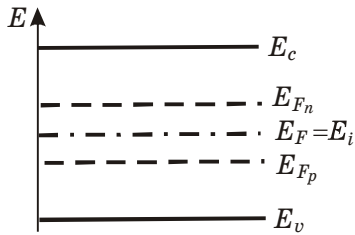


Рис. 5.2. Положение квазиуровней Ферми для неравновесного состояния собственного полупроводника

Величины E_{Fn} и E_{Fp} играют роль уровня Ферми для электронов и дырок и называются квазиуровнями Ферми. Они не совпадают друг с другом и с равновесным уровнем Ферми (рис. 5.2). К этому необходимо добавить, что в неравновесном состоянии равновесный уровень Ферми как бы расщепляется на два квазиуровня: для электронов E_{Fn} и для дырок E_{Fp} .

$$(5.9)$$

§ 5.3. Рекомбинация свободных носителей заряда

Носители заряда, созданные вследствие генерации, не могут бесконечно долго находиться в свободном состоянии в разрешенных зонах. Энергия электронов в зоне проводимости и дырок в валентной зоне уменьшается в результате процессов рассеяния. В итоге электроны оказываются на дне зоны проводимости, а дырки — у потолка валентной зоны. На последней стадии уменьшение энергии электрона может быть связано только с его переходом на свободный энергетический уровень в валентной зоне. При этом энергия электрона уменьшается скачком на величину ΔE (рис. 5.3). Таким образом, процесс хаотического уменьшения энергии электронов и дырок заканчивается исчезновением пары свободных носителей заряда, что называется рекомбинацией.

В рассмотренном случае, когда процесс представляет переход электрона из зоны проводимости в валентную зону, рекомбинация имеет название межзонной рекомбинации. Очевидно, что процесс рекомбинации является обратным процессу генерации.

Описанный выше процесс потери электронами избыточной энергии $\Delta E = E_c - E_v$ может сопровождаться:

- 1) излучением света, т.е. фотонами (излучательная рекомбинация);

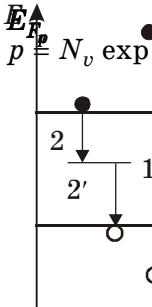


Рис. 5.3. Процесс перехода электронов из зоны проводимости в валентную зону: 1 — межзонная рекомбинация; 2, 2' — рекомбинация через промежуточный уровень

2) передачей энергии кристаллической решетке, что соответствует генерации фононов (безызлучательная рекомбинация);

3) передачей избыточной энергии другому электрону в зоне проводимости, вследствие чего он занимает более высокий энергетический уровень (ударная рекомбинация, или Оже-рекомбинация).

Кроме разделения рекомбинации по механизмам передачи избыточной энергии выделяют различные виды рекомбинации по способу перехода электронов из зоны проводимости в валентную зону. Здесь различают прямую межзонную рекомбинацию (переход 1 на рис. 5.3) и рекомбинацию через промежуточный уровень (переходы 2 и 2' на рис. 5.3). В последнем случае избыточная энергия будет выделяться двумя порциями, соответствующими переходам 2 и 2' (рис. 5.3). Промежуточные уровни обычно расположены вблизи середины запрещенной зоны и носят название глубоких или рекомбинационных уровней. По своей природе эти уровни могут быть обусловлены атомами примесей, например меди, золота, марганца, и другими точечными и линейными дефектами кристаллической решетки (вакансии, атомы в междоузлиях, дислокации и т.д.).

§ 5.4. Скорость рекомбинации и время жизни неравновесных носителей заряда

Процесс рекомбинации характеризуется скоростью рекомбинации R , которая определяет число электронно-дырочных пар, исчезающих (рекомбинирующих) в единичном объеме в единицу времени. Скорость рекомбинации электронов можно представить выражением

$$R_p = -\frac{d \Delta p}{dt} \quad (5.10)$$

Отрицательный знак в правой части указывает на уменьшение концентрации избыточных носителей заряда. Скорость рекомбинации дырок можно выразить формулой, подобной (5.10):

$$R_p = -\frac{d \Delta p}{dt} \quad (5.11)$$

За время $t = \tau$ после прекращения процесса генерации неравновесных носителей заряда рекомбинируют $\Delta n/\tau$ электронов и $\Delta p/\tau$ дырок. Тогда выражения (5.10) и (5.11) можно записать в виде

$$R_n = -\frac{d \Delta n}{dt} = \frac{\Delta n}{\tau}, \quad (5.12)$$

$$(5.13)$$

Решение уравнений (5.12) и (5.13) дает следующие формулы:

(5.14)

$$\Delta p(t) = \Delta p(0) e^{-\frac{t}{\tau}}. \quad (5.15)$$

При $t = \tau$ видно, что $\Delta n(\tau) = \Delta n(0)/e$ и $\Delta p(\tau) = \Delta p(0)/e$. Отсюда следует, что величина τ представляет собой время, за которое концентрация неравновесных носителей уменьшается за счет рекомбинации в $e \approx 2,71$ раза. Это время называется временем жизни неравновесных носителей заряда.

Физически время жизни представляет среднюю величину времени, в течение которого носитель (электрон) ведет себя как свободная частица до момента исчезновения в результате рекомбинации с носителем заряда противоположного знака (дыркой). Это время колеблется от нескольких миллисекунд до нескольких микросекунд и существенно превышает время между двумя столкновениями, которое при описании процессов рассеяния (см. § 3.3) называлось временем релаксации и обозначалось также буквой τ . Отметим, что до «опускания» неравновесного электрона на дно зоны проводимости и «подъема» дырки к потолку валентной зоны неравновесные носители испытывают 10^6 – 10^8 столкновений за время около 10^{-10} с.

§ 5.5. Рекомбинация через локальные уровни

Дефекты кристаллической решетки — примесные атомы, вакансии, дислокации и другие — создают в запрещенной зоне полупроводника локальные энергетические уровни.

Предположим, что в запрещенной зоне, вблизи ее середины, располагается свободный локальный уровень E_d (рис. 5.4,а). Такой уровень будет захватывать электрон из зоны проводимости (переход 1 на рис. 5.4,а) и превращаться в отрицательно заряженный ион. В дальнейшем захваченный электрон может быть переброшен назад, в зону проводимости (переход 1'), за счет тепловой энергии или перейти в валентную зону (переход 2). Первый процесс заканчивается перебросом электрона назад, в зону проводимости, второй — рекомбинацией электронно-дырочной пары. Переход 2 можно рассматривать как захват дырки из валентной зоны на локальный уровень E_d (переход 2').

Если локальный уровень располагается вблизи дна зоны проводимости (рис. 5.4,б), то вероятность обратного переброса электрона будет преобладать над вероятностью захвата дырки. Переход электрона с такого мелкого уровня (уровень расположен вблизи разрешенной зоны) в валентную зону практически столь же маловероятен, как и межзонный переход. Подобная ситуация наблюдается для уровня, расположенного вблизи потол-

$$\Delta n(t) = \Delta n(0) e^{-\frac{t}{\tau}}$$

ка валентной зоны (рис. 5.4,в). Такого рода мелкие локальные уровни называются уровнями прилипания, ловушками или уровнями захвата. Из вышеизложенного следует, что наличие в полупроводнике уровней прилипания может существенно увеличивать время жизни неравновесных носителей заряда.

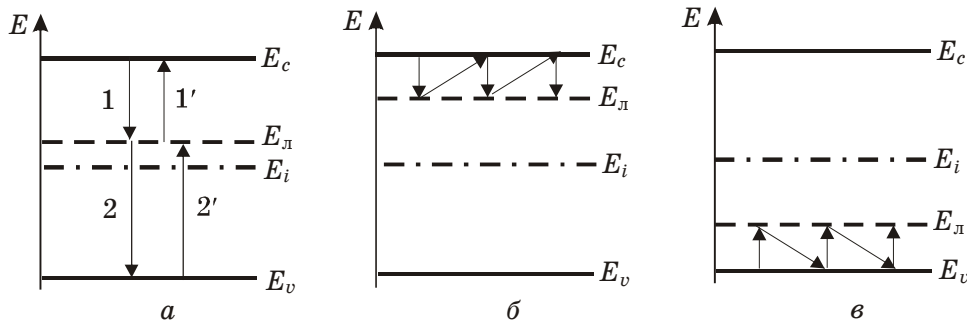


Рис. 5.4. Локальные уровни в полупроводнике:
 а — глубокий рекомбинационный уровень;
 б — уровень прилипания вблизи дна зоны проводимости;
 в — уровень прилипания у потолка валентной зоны

Для теплового переброса носителя с глубокого уровня назад, в зону проводимости, электрону требуется несколько фононов, и вероятность переброса в зону проводимости будет значительно меньше, чем переход в валентную зону, т.е. более вероятным будет захват дырки из валентной зоны на локальный уровень $E_{л}$. Поэтому глубокий локальный уровень $E_{л}$ (рис. 5.4,а) является эффективным центром (уровнем) рекомбинации и в полупроводниках с широкой запрещенной зоной рекомбинация через глубокие локальные уровни — основной вид рекомбинации. Увеличение концентрации глубоких рекомбинационных ловушек приводит к росту скорости рекомбинации и к соответствующему снижению времени жизни неравновесных носителей заряда, что широко используется в технологии изготовления полупроводниковых приборов.

§ 5.6. Зависимость времени жизни неравновесных носителей от концентрации основных носителей заряда при рекомбинации через локальные уровни

Теория рекомбинации через локальные уровни разработана американскими физиками Шокли, Ридом и Холлом, которые, рассматривая рекомбинацию через один локальный уровень, получили следующее выражение для времени жизни неравновесных носителей заряда:

$$(5.16)$$

где n_0, p_0 — равновесные концентрации электронов в зоне проводимости и дырок в валентной зоне соответственно; n_1, p_1 — равновесные концентрации электронов и дырок в соответствующих зонах, когда уровень Ферми совпадает с рекомбинационным уровнем; τ_{n0}, τ_{p0} — время жизни электронов и дырок соответственно.

Анализ зависимости времени жизни τ от равновесной концентрации электронов и дырок в соответствии с выражением (5.16) удобно проводить параллельно с анализом подобной зависимости положения уровня Ферми (рис. 5.5). За начало отсчета по оси абсцисс принята концентрация носителей в собственном полупроводнике (см. выражение (2.35)). Вправо от этой точки откладываются значения n_0/n_i , влево — значения p_0/n_i .

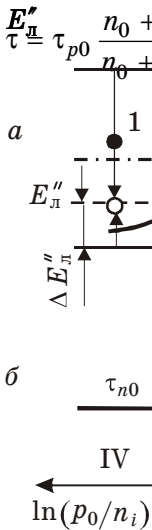
Как следует из рис. 5.5,а, в собственном полупроводнике уровень Ферми располагается приблизительно в середине запрещенной зоны (см. формулу (2.34)), в донорном полупроводнике — в верхней половине, в полупроводнике p -типа — в нижней половине запрещенной зоны (см. § 2.5). Можно выделить четыре характерные области на зависимости времени жизни от концентрации носителей заряда (рис. 5.5,б), которые соответствуют четырем областям изменения положения уровня Ферми (рис. 5.5,а).

Область I. Сильно легированный полупроводник n -типа. Уровень Ферми расположен ниже дна зоны проводимости, но выше энергетического уровня рекомбинационной ловушки, т.е. $E_c - \Delta E_r < E_F < E_c$. Здесь следует заметить, что на рис. 5.5,а и, следовательно, в дальнейшем будут рассматриваться только глубокие рекомбинационные уровни — в верхней и в нижней половине запрещенной зоны. Примесные мелкие уровни, соответствующие типу проводимости полупроводника, предполагаются, но на рис. 5.5,а не показаны, так как в рекомбинационных процессах прямого участия не принимают.

Рис. 5.5. Зависимости уровня Ферми (а) и времени жизни неосновных носителей (б) от равновесной концентрации основных носителей

В рассматриваемом случае справедливы следующие соотношения концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg n_1; \quad n_0 \gg p_1. \quad (5.17)$$



Учитывая эти неравенства, из выражения (5.16) получим

$$\tau = \tau_{p0}. \quad (5.18)$$

Этот вывод понятен из следующих рассуждений. В области I уровень Ферми лежит выше рекомбинационного уровня $E_{л}$, и, значит, уровень $E_{л}$ будет заполнен электронами, так как на уровне Ферми вероятность заполнения электронами равна $1/2$, а ниже уровня Ферми вероятность заполнения стремится к 1 и в пределах нескольких единиц k_0T ниже E_F вероятность заполнения практически равна 1 (см. рис. 2.1). Тогда каждый захват дырки на состоянии уровня $E_{л}$, целиком заполненные захваченными из зоны проводимости электронами, будет сопровождаться рекомбинацией, т.е. время жизни будет определяться только вероятностью захвата дырок и будет равно τ_{p0} , как следует из формулы (5.18).

Область II. Донорный полупроводник легирован слабо, так что уровень Ферми лежит ниже энергетического уровня рекомбинационных ловушек $E_{л}$, но выше середины запрещенной зоны, т.е. $E_i < E_F < E_{л}$. В этом случае имеют место следующие соотношения для концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg p_1; \quad n_0 < n_1. \quad (5.19)$$

Соответственно выражение для τ приобретет вид:

$$\tau = \tau_{p0} \frac{n_1}{n_0}. \quad (5.20)$$

Из соотношения (5.20) видно, что по мере понижения уровня Ферми время жизни τ растет по экспоненциальному закону. С физической точки зрения это объясняется следующим образом. При снижении уровня Ферми уменьшается степень заполнения ловушек электронами, вследствие чего вероятность захвата ловушкой дырок снижается, что ведет к увеличению времени жизни дырок.

Анализ зависимости времени жизни от положения равновесного уровня Ферми в полупроводнике p -типа удобно начать с области IV с последующим переходом к области III (рис. 5.5, а, б, слева от оси ординат).

Область IV. Сильно легированный полупроводник p -типа, для которого положение уровня Ферми определяется неравенством вида $E_v < E_F < E_v + k_0T$. В этом случае будут справедливы следующие соотношения концентраций:

$$p_0 \gg n_0; \quad p_0 \gg p_1; \quad p_0 \gg n_1. \quad (5.21)$$

Тогда из выражения (5.16) с учетом неравенств (5.21) следует, что

$$\tau = \tau_{p0}. \quad (5.22)$$

Соотношение (5.22) можно объяснить по аналогии с равенством (5.18). Когда уровень Ферми находится в области IV (рис. 5.5, а), то все рекомбинационные центры на уровне $E_{л}$ должны быть заполнены дырками, и вероятность рекомбинации определяется только захватом электронов

на уровень $E_{\text{л}}$. Каждый электрон, захваченный ловушкой, немедленно рекомбинирует с дыркой, количество которых в валентной зоне и, следовательно, на уровне $E_{\text{л}}$ велико.

Область III. Она начинается при концентрации дырок p_0 , когда уровень Ферми располагается выше уровня ловушек $E_{\text{л}}$. Дальнейший подъем уровня Ферми связан с уменьшением величины p_0 и вызывает падение степени заполнения ловушек дырками и уменьшение скорости их восстановления, что приводит к уменьшению скорости рекомбинации избыточных носителей n , следовательно, к увеличению времени их жизни.

Таким образом, из приведенного анализа изменения времени жизни с изменением положения уровня Ферми следует важный вывод: при ярко выраженном типе проводимости (области I и IV) время жизни носителей заряда в полупроводнике определяется неосновными носителями заряда. Поэтому для характеристики рекомбинационных процессов в полупроводнике пользуются понятием времени жизни неосновных носителей заряда.

§ 5.7. Температурная зависимость времени жизни носителей заряда

Рассмотрим процесс изменения времени жизни носителей заряда в донорном полупроводнике в зависимости от температуры. Температурная зависимость времени жизни носителей обусловлена изменением положения уровня Ферми. Поэтому по аналогии с § 5.6 будем анализировать параллельно две зависимости: $E_F(1/T)$ и $\tau(1/T)$, представленные на рис. 5.6. В донорном полупроводнике мелкий донорный уровень должен располагаться вблизи дна зоны проводимости в области между E_c и рекомбинационным уровнем $E_{\text{л}}$. На рис. 5.6,а донорный уровень не показан по причине, изложенной в § 5.6.

Анализ температурной зависимости положения уровня Ферми (рис. 5.6,а) не отличается от приведенного в § 2.5 (см. рис. 2.7). Предположим, что при низкой температуре все донорные атомы ионизированы, т.е. уровень Ферми находится ниже донорного уровня. Тогда в области I (рис. 5.6,б) при $T < T_{\text{л}}$ все ловушки рекомбинационного уровня $E_{\text{л}}$ будут заполнены электронами (уровень E_F находится выше $E_{\text{л}}$), и скорость рекомбинации определяется числом актов захвата дырок. Как показано в § 5.6, время жизни в этом случае определяется выражением (5.18), т.е. равно $\tau_{\text{л}}$ (рис. 5.6,б).

Область II начинается с температуры $T_{\text{л}}$, при которой уровень Ферми пересекает рекомбинационный уровень $E_{\text{л}}$, и продолжается до температуры T_i , выше которой свойства донорного полупроводника не отличаются от свойств собственного полупроводника. В этой области время жизни

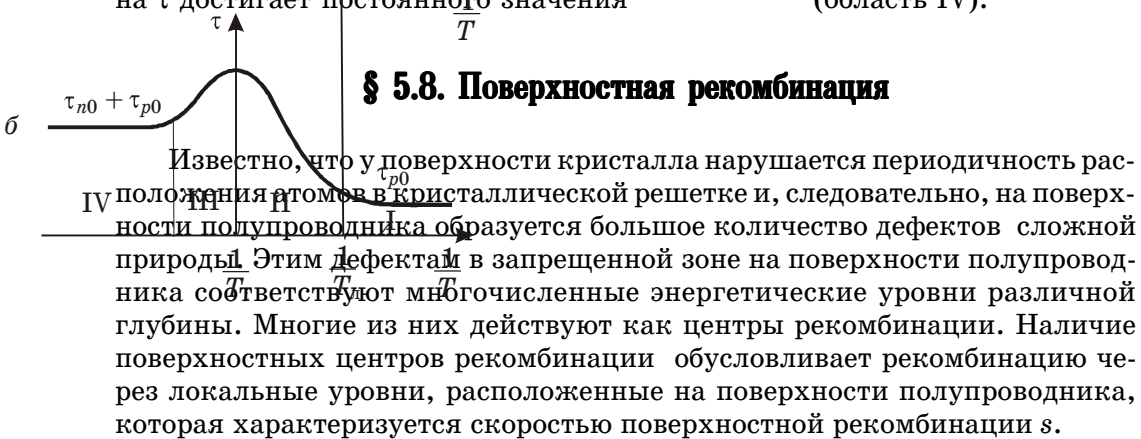
$E_{\text{л}}$

носителей увеличивается с ростом температуры, потому что в данной ситуации наблюдается интенсивный выброс электронов с уровня E_d в зону проводимости. Вследствие этого скорость рекомбинации дырок, забрасываемых на уровень E_d из валентной зоны, снижается, что обуславливает увеличение их времени жизни. Максимального значения величина τ достигает при температуре перехода полупроводника к собственной проводимости T_i . Таким образом, рассмотренные процессы в областях I и II относятся к области истощения донорной примеси (см. рис. 2.7, область 2), когда уровень Ферми продолжает с повышением температуры опускаться вниз к середине запрещенной зоны.

Дальнейшее повышение температуры вызывает интенсивное нарастание концентрации свободных носителей заряда (собственных электронов и дырок), что обсуждалось в § 2.5 (см. рис. 2.7,б), и увеличение скорости рекомбинации. Это обуславливает уменьшение времени жизни носителей заряда в области III (рис. 5.6,б).

Скорость рекомбинации в областях III и IV увеличивается с ростом температуры, а время жизни уменьшается пока не установится равновесие между числом носителей, захваченных рекомбинационными ловушками, и числом их тепловых выбросов из ловушек. В этом случае величина τ достигает постоянного значения (область IV).

Рис. 5.6. Зависимости уровня Ферми (а) и времени жизни носителей (б) от температуры полупроводника



Присутствие на поверхности огромного числа чужеродных атомов, нарушений, созданных при механической обработке кристалла, приводит к усилению роли поверхностной рекомбинации на фоне объемной.

Тогда при одновременном протекании в полупроводниковом кристалле объемной и поверхностной рекомбинаций время жизни неравновесных носителей заряда определяется временем жизни носителей в объеме τ_0 и скоростью поверхностной рекомбинации:

$$\frac{1}{\tau_{\text{эфф}}} = \frac{1}{\tau_0} + \frac{2s}{d}, \quad (5.23)$$

где d — толщина кристалла. Величина $\tau_{\text{эфф}}$ называется эффективным временем жизни носителей заряда.

Из выражения (5.23) видно, что уменьшение толщины кристалла соответствует возрастанию роли поверхностной рекомбинации и уменьшению времени жизни неравновесных носителей заряда.

Контрольные вопросы и задачи

1. (АУМГ). Укажите правильные определения механизма тепловой генерации носителей заряда в полупроводниках:

1) монополярная тепловая генерация электронов наблюдается в донорных полупроводниках, если атомы кристалла передают электронам донорного атома тепловую энергию, величина которой отвечает условию

$$E > E_c - E_d,$$

где E_c — энергия дна зоны проводимости; E_d — энергия донорного уровня;

2) монополярная тепловая генерация дырок наблюдается в акцепторных полупроводниках, когда электроны с акцепторного уровня переходят в валентную зону;

3) монополярная тепловая генерация дырок наблюдается при переходе электронов из валентной зоны на акцепторный уровень после передачи им тепловой энергии, величина которой соответствует условию

$$E > E_a - E_v,$$

где E_a — энергия акцепторного уровня; E_v — энергия потолка валентной зоны;

4) биполярная генерация пар электрон-дырка наблюдается при одновременных переходах электронов с донорного и акцепторного уровней при получении ими тепловой энергии, достаточной для ионизации указанных уровней;

5) биполярная генерация наблюдается при переходе электронов из валентной зоны в зону проводимости при получении ими энергии за счет тепловых флуктуаций, если выполняется условие

$$E \geq E_c - E_v, \quad \text{т.е. } E \geq \Delta E.$$

2. (ИЧИЗ). Какие носители заряда называются неравновесными? Укажите правильные ответы.

Ответы:

1) носители заряда, которые сразу после генерации имеют энергию выше средней тепловой энергии частиц, называются неравновесными;

2) дополнительные носители заряда, созданные при температурах выше комнатной (300 К), являются неравновесными;

3) носители заряда, у которых отсутствует термодинамическое равновесие с кристаллической решеткой, называются неравновесными;

4) свободные носители заряда, возникающие в полупроводнике при воздействии на него светом, корпускулярными потоками нейтронов, протонов, α -частиц, космических лучей и электрическим полем, называются избыточными;

5) носители заряда, для которых не выполняется закон действующих масс, т.е. $np \neq n_i^2$, называются неравновесными.

3. (ВНРЦ). Какой процесс в полупроводниках называется рекомбинацией носителей заряда? Укажите правильные ответы:

1) переход электронов из зоны проводимости в валентную зону, в результате чего исчезают свободные пары электрон-дырка;

2) уменьшение энергии электронов в зоне проводимости и дырок в валентной зоне в результате процесса хаотического рассеяния, вследствие чего электроны опускаются на дно зоны E_c , а дырки поднимаются к потолку нижней зоны E_v ;

3) исчезновение пары электрон-дырка, в результате чего происходит потеря носителями заряда энергии, величина которой равна ширине запрещенной зоны;

4) переход свободного электрона из зоны проводимости в валентную зону через промежуточный энергетический уровень вблизи середины запрещенной зоны с выделением энергии в виде двух частей.

4. (ЭДРЛ). Какие параметры определяют процесс рекомбинации неравновесных носителей заряда в полупроводниках?

Ответы:

1) время жизни неравновесных носителей заряда τ , за которое концентрация избыточных носителей уменьшается в $e \approx 2,71$ раза вследствие рекомбинации;

2) время релаксации τ , за которое свободный носитель заряда проходит путь, равный длине свободного пробега;

3) скорость рекомбинации, представляющая число электронно-дырочных пар, рекомбинирующих в единичном объеме за единицу времени.

5. (РАЛУ). Какие энергетические уровни в полупроводниках являются рекомбинационными и уровнями прилипания?

Ответы:

1) энергетическое положение локального уровня E_d вблизи середины запрещенной зоны E_i определяет приблизительно равные вероятности захвата электрона из зоны проводимости и дырки из валентной зоны, что приводит к исчезновению пар электрон-дырка. Следовательно, глубокие уровни, расположенные вблизи E_i , являются рекомбинационными;

2) энергетическое положение локального уровня вблизи дна зоны проводимости говорит о большей вероятности захвата электрона из зоны проводимости, чем электрона из валентной зоны. Захваченный электрон за счет тепловой энергии может перебрасываться назад, в зону проводимости. Такой уровень называется рекомбинационным;

3) энергетическое положение локального уровня вблизи потолка валентной зоны способствует наибольшей вероятности захвата электрона из валентной зоны, чем из зоны проводимости. Захваченный электрон за счет тепловой энергии затем может быть переброшен назад, в валентную зону. Такой уровень называется рекомбинационным, так как исчезают электрон с уровня E_d и дырка валентной зоны;

4) уровни, упомянутые в пунктах 2, 3, не могут быть рекомбинационными, так как в анализируемых процессах участвуют носители заряда только одного типа. Эти уровни называются уровнями прилипания.

6. (ИМВЖ). Как время жизни носителей заряда зависит от концентрации электронов в донорном полупроводнике?

Ответы:

1) в сильно легированном невырожденном полупроводнике n -типа, когда уровень Ферми занимает положение выше рекомбинационного уровня E'_d , последний полностью занят электронами, и вероятность рекомбинации определяется временем жизни свободной дырки в валентной зоне перед захватом ее на уровень E'_d ;

2) в слабо легированном полупроводнике n -типа, когда уровень Ферми находится ниже E'_d , время жизни электронно-дырочных пар определяется временем жизни неосновных носителей заряда, т.е. дырок τ_{p0} ,

и экспоненциально растет по закону $\tau = \tau_{p0} e^{\frac{E'_d - E_F}{k_0 T}}$ когда снижается уровень Ферми, т.е. уменьшается концентрация равновесных электронов. Это связано с уменьшением степени заполнения электронами рекомбинационного уровня;

3) время жизни носителей заряда при рекомбинации через локальный уровень определяется только его энергетическим положением в запрещенной зоне полупроводника и не зависит от равновесной концентрации основных носителей заряда.

7. (ГЛТЗ). Укажите правильные ответы, объясняющие температурную зависимость времени жизни носителей заряда в донорных полупроводниках:

$$\tau = \tau_{p0} e^{\frac{E'_d - E_F}{k_0 T}}$$

1) температурная зависимость времени жизни $\tau(T)$ обусловлена изменением положения уровня Ферми и связанным с ним изменением плотности основных носителей заряда (электронов) на рекомбинационном уровне E_L ;

2) зависимость $\tau(T)$ обусловлена рассеянием носителей заряда на тепловых колебаниях атомов кристаллической решетки;

3) если в начальном состоянии все донорные атомы ионизированы, а уровень Ферми находится выше рекомбинационного уровня, то все рекомбинационные ловушки заполнены электронами и время жизни определяется временем жизни дырок (неосновных носителей заряда), т.е.

$$\tau = \tau_{p0};$$

4) при повышении температуры от $T = T_L$, при которой уровень Ферми пересекает рекомбинационный уровень, электроны за счет тепловой энергии начинают выбрасываться из рекомбинационных ловушек, вследствие чего скорость рекомбинации дырок, поступающих на уровень E_L из валентной зоны, уменьшается, а время жизни носителей заряда возрастает;

5) рост времени жизни наблюдается в области температур от T_L до T_i , при которой полупроводник приобретает свойства собственного полупроводника, и время жизни достигает максимального значения;

6) в области $T > T_i$ время жизни уменьшается за счет нарастания концентрации собственных носителей заряда до тех пор, пока не установится равновесие между числом носителей, захваченных рекомбинационными ловушками, и числом тепловых выбросов из них. Начиная с этой температуры величина

$$\tau = \tau_{n0} + \tau_{p0} = \text{const.}$$

8. (ОХПР). Какова природа поверхностной рекомбинации в полупроводниках и как такая рекомбинация влияет на время жизни носителей заряда?

Ответы:

1) поверхностная рекомбинация обусловлена наличием на поверхности полупроводников центров рекомбинации, связанных с дефектами кристаллической решетки, которые неизбежно образуются при технологической обработке образцов. Дефекты вызывают возникновение различных локальных уровней в запрещенной зоне на поверхности полупроводника. Глубокие энергетические уровни действуют как центры поверхностной рекомбинации;

2) носители заряда рекомбинируют только в объеме полупроводника, где они возникают при воздействии света, тепла и электрического поля;

3) объемная рекомбинация носителей заряда осуществляется через локальные рекомбинационные уровни, вызванные структурными дефектами в объеме полупроводника, параллельно с поверхностной рекомбинацией через поверхностные центры (структурные дефекты кристаллической решетки, чужеродные атомы, адсорбированные на поверхности);

4) время жизни носителей заряда определяется временем жизни при рекомбинации в объеме полупроводника τ_0 и скоростью поверхностной рекомбинации s . В этом случае эффективное время жизни $\tau_{\text{эфф}}$ можно найти из выражения

$$\frac{1}{\tau_{\text{эфф}}} = \frac{1}{\tau_0} + \frac{2s}{d},$$

где d — толщина полупроводникового образца.

Глава 6. Диффузия и дрейф носителей заряда в полупроводниках

§ 6.1. Диффузионные потоки носителей заряда

Пусть имеется одномерный кристалл донорного полупроводника, в котором концентрация основных носителей заряда уменьшается в направлении x , т.е. $n = f(x)$ (рис. 6.1). Такая ситуация может наблюдаться, например, при неоднородном распределении примесей в объеме полупроводника или при освещении его поверхности (плоскость $x = 0$) светом с энергией кванта $\hbar\omega \geq E_c - E_d$.

Рассмотрим движение носителей в бесконечно тонких слоях 1 и 2 (с толщиной dx). Вследствие теплового движения слои 1 и 2 будут обмениваться электронами. Так как тепловое движение является хаотическим, т.е. равновероятным во всех направлениях, то половина всех носителей из слоя 1 переместится в слой 2 и, наоборот, половина носителей из слоя 2 переместится в слой 1.

$I_{n \text{ диф}} =$

Но число носителей в слое 1 больше, чем в слое 2, и поток электронов из слоя 1 в слой 2 будет преобладать.

Если $n(x - dx/2)$ — средняя концентрация электронов в слое 1, а $n(x + dx/2)$ — в слое 2, то можно записать следующее выражение для разности концентраций электронов в этих слоях:

$$n\left(x - \frac{dx}{2}\right) - n\left(x + \frac{dx}{2}\right) = -\frac{dn}{dx} dx. \quad (6.1)$$

Из выражения (6.1) следует, что разность концентраций электронов в бесконечно тонких слоях 1 и 2 пропорциональна их градиенту — dn/dx , и результирующий поток электронов направлен в сторону уменьшения концентрации, т.е. из области 1 в область 2. Этот процесс и, следовательно, поток носителей называются диффузионными.

Для диффузионного потока электронов $I_{n \text{ диф}}$ можно записать выражение

$$(6.2)$$

где D_n — коэффициент диффузии электронов, $\text{м}^2/\text{с}$.

Рассматривая аналогично диффузию дырок, можно получить выражение для диффузионного потока дырок

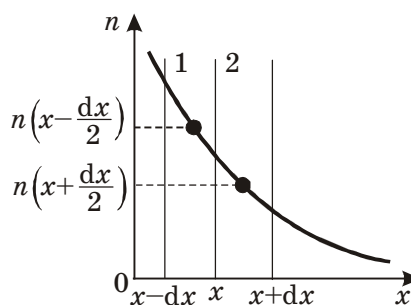


Рис. 6.1. Неравномерное распределение концентрации носителей заряда в направлении оси x

$$I_{p \text{ диф}} = -D_p \frac{d p}{d x}, \quad (6.3)$$

где D_p — коэффициент диффузии дырок, $\text{м}^2/\text{с}$.

Знак минус в выражениях (6.2), (6.3) указывает на то, что диффузионный поток носителей заряда направлен в сторону уменьшения концентрации.

§ 6.2. Диффузионный и дрейфовый токи носителей заряда

Диффузионным потокам заряженных частиц (6.2) и (6.3) соответствуют диффузионные токи электронов $j_{n \text{ диф}}$ и дырок $j_{p \text{ диф}}$, выражения для которых с учетом знака заряда носителей имеют вид:

$$j_{n \text{ диф}} = q D_n \frac{d n}{d x}; \quad (6.4)$$

$$j_{p \text{ диф}} = -q D_p \frac{d p}{d x}. \quad (6.5)$$

В электрическом поле \mathbf{E} электроны и дырки приобретают направленное движение, в результате чего возникают электронный и дырочный дрейфовые токи. Согласно формулам (3.8), (3.16) и (3.17) выражения для плотностей дрейфовых электронного $j_{n \text{ др}}$ и дырочного $j_{p \text{ др}}$ токов будут иметь вид:

$$(6.6)$$

$$(6.7)$$

Формулы (6.6) и (6.7) записаны без учета разделения носителей на основные и неосновные.

Полные электронный и дырочный токи в полупроводниках представляют сумму дрейфовых и диффузионных токов:

$$(6.8)$$

$$j_p = j_{p \text{ др}} + j_{p \text{ диф}} = q \mu_p p \mathbf{E} - q D_p \frac{d p}{d x}. \quad (6.9)$$

Плотность общего тока в неоднородном полупроводнике будет

$$j = j_n + j_p = q (\mu_n n + \mu_p p) \mathbf{E} + q \left(D_n \frac{d n}{d x} - D_p \frac{d p}{d x} \right). \quad (6.10)$$

Следует обратить внимание на то, что диффузионный ток существует только в полупроводниках, где возможно создание неоднородного распределения концентрации носителей заряда в различных точках кристалла,

$$j_{n \text{ др}} = \sigma_n \mathbf{E}$$

Интегрируя обе части уравнения (6.12) по dn в интервале от n_{01} до n_{02} и по dV от V_1 до V_2 , полагая $V_1 - V_2 = V$, получим:

$$n_{02} = n_{01} \exp\left(-\frac{\mu_n}{D_n} V\right). \quad (6.13)$$

Электроны в зоне проводимости полупроводника представляют собой невырожденный газ при определенной температуре, т.е. подчиняются статистике Максвелла — Больцмана. Из распределения Максвелла — Больцмана следует, что число электронов в левой области образца, способных преодолеть потенциальный барьер qV , будет равно

$$(6.14)$$

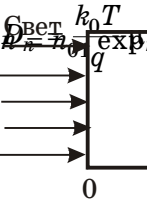
С другой стороны, очевидно, что в стационарном состоянии это число электронов должно быть равно числу электронов, которые могут под действием внутреннего поля перейти из правой области в левую, т.е. $n = n_{02}$.

Сравнивая выражения (6.14) и (6.13), выводим связь между коэффициентом диффузии и подвижностью носителей заряда, получившую название соотношения Эйнштейна:

$$(6.15)$$

Так же, как и подвижности (см. § 3.3), коэффициенты диффузии электронов и дырок имеют разные значения. Например, в германии при 20°C $D_n = 98 \text{ см}^2/\text{с}$, $D_p = 47 \text{ см}^2/\text{с}$.

Уравнение Эйнштейна имеет универсальный характер, так как применимо к носителям заряда любого типа: электронам, дыркам, ионам, ионным вакансиям. Оно справедливо как для равновесных, так и для неравновесных носителей заряда в полупроводниках и диэлектриках.



§ 6.4. Диффузионная длина носителей заряда

Рассмотрим полупроводник, поперечные размеры которого значительно меньше его длины (рис. 6.3,а).

Рис. 6.3. Одномерная модель полупроводникового образца при освещении (а) и распределение в нем избыточной концентрации носителей заряда (б)

Пусть левый торец образца (плоскость $x = 0$) освещается сильно поглощаемым светом так, что на поверхности происходит генерация избыточных электронов Δn и дырок Δp . Вследствие генерации концентрация электронов и дырок вблизи поверхности оказывается больше, чем в объеме, т.е. возникает градиент концентрации носителей заряда, обуславливающий диффузию носителей из области с повышенной концентрацией в область с меньшей концентрацией носителей заряда. По мере удаления от поверхности (плоскость $x = 0$) концентрация неравновесных носителей заряда будет уменьшаться из-за рекомбинации электронов и дырок по экспоненциальному закону согласно выражениям (5.14) и (5.15):

$$(6.16)$$

$$\Delta p(x) = \Delta p(0) \exp(-x/L_p), \quad (6.17)$$

где $\Delta n(0)$, $\Delta p(0)$ — концентрации избыточных электронов и дырок в области генерации (плоскость $x = 0$); L_n , L_p — диффузионная длина электронов и дырок соответственно.

За время жизни избыточные носители заряда диффундируют на расстояние L , которое называется диффузионной длиной. Из выражений (6.16) и (6.17) следует, что диффузионная длина — это расстояние, на котором концентрация неравновесных носителей заряда уменьшается в $e = 2,71$ раза.

Распределение концентрации неравновесных носителей заряда для одномерного случая представлено на рис. 6.3, б.

$\Delta n(x) = \sqrt{D_n \tau_n} \Delta n(0) \exp(-x/L_n)$,
 $\Delta p(x) = \sqrt{D_p \tau_p} \Delta p(0) \exp(-x/L_p)$.
 Диффузионная длина, время жизни и коэффициент диффузии связаны между собой следующими соотношениями:

$$L = \sqrt{D\tau}; \quad (6.18)$$

$$(6.19)$$

$$(6.20)$$

Диффузионная длина является важной характеристикой полупроводникового материала и определяет его пригодность для использования в производстве. Например, для изготовления биполярных транзисторов, фотоэлектрических приборов нужны материалы с большой диффузионной длиной; для изготовления импульсных быстродействующих диодов — с малой диффузионной длиной. У материалов, используемых в полупроводниковой электронике, диффузионная длина изменяется от десятых до тысячных долей сантиметра.

§ 6.5. Уравнение непрерывности

Рассмотрим поведение носителей заряда в полупроводнике при одновременном протекании процессов генерации, рекомбинации, диффузии и дрейфа. Очевидно, что в результате этих процессов концентрации носителей заряда в образце могут зависеть от координат и времени, т.е. для одномерного случая будут функциями типа $n(x, t)$ и $p(x, t)$. Для решения этой задачи выделим в длинном образце полупроводника (рис. 6.4) слой толщиной dx с сечением, равным единице площади ($S = 1$).

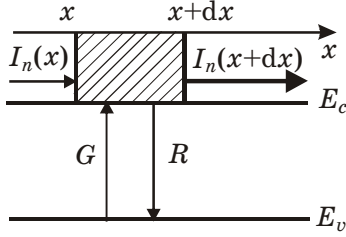


Рис. 6.4. К выводу уравнения непрерывности

Левая граница этого слоя имеет координату x , правая — $(x + dx)$, объем слоя равен dx (1 единица объема). В момент времени t количество электронов в выделенном слое равно $n(x, t) dx$, а в момент времени $t + dt$ эта величина составит $n(x, t + dt)dx$. Изменение числа электронов в слое dx за время dt будет равно

$$(6.21)$$

В общем случае это изменение может произойти за счет генерации, рекомбинации, диффузии и дрейфа носителей заряда. Рассмотрим возможный вклад этих процессов подробнее.

1. Механизмы генерации рассмотрены в § 5.1, § 5.2. Предположим, что в выделенном объеме протекает оптическая межзонная генерация. Пусть скорость генерации составляет G пар электрон-дырка в 1 см^3 за 1 с . Тогда световая генерация в объеме dx за время dt приведет к образованию свободных пар электрон-дырка в количестве

$$G dx dt. \quad (6.22)$$

2. Одновременно с процессом генерации в выделенном объеме (см. рис. 6.4) протекает рекомбинация, которая ограничивает рост числа свободных носителей заряда. Если R_n и R_p — скорости рекомбинации электронов и дырок, то количество электронов и дырок, рекомбинирующих в объеме dx за время dt , с учетом формул (5.4), (5.5), (5.12), (5.13) будет соответственно определяться выражениями:

$$-R_n dx dt = -\frac{n - n_0}{\tau_n} dx dt, \quad (6.23)$$

$$-R_p dx dt = -\frac{p - p_0}{\tau_p} dx dt. \quad (6.24)$$

3. Количество носителей заряда в объеме dx за время dt может изменяться за счет диффузии и дрейфа. Если $I_n(x, t)$ — втекающий поток элек-

$n(x, t + dt)$

тронов, а $I_n(x + dx, t)$ — вытекающий поток, то изменение количества электронов в выделенном объеме за время dt будет равно

$$I_n(x, t) dt - I_n(x + dx, t) dt = -\frac{\partial I_n}{\partial x} dx dt. \quad (6.25)$$

Суммируя выражения (6.21)–(6.25), получим полное изменение числа электронов в объеме dx за время dt :

$$\frac{\partial n}{\partial t} = G - \frac{\partial I_n}{\partial x} - \frac{n - n_0}{\tau_n}, \quad (6.26)$$

и дырок:

$$\frac{\partial p}{\partial t} = G - \frac{\partial I_p}{\partial x} - \frac{p - p_0}{\tau_p}. \quad (6.27)$$

Потоки электронов и дырок можно выразить через плотности тока:

$$-I_n = \frac{j_n}{q}, \quad (6.28)$$

$$I_p = \frac{j_p}{q}. \quad (6.29)$$

Тогда для стационарных условий, когда концентрации электронов и дырок не изменяются во времени, т.е. $dn/dt = dp/dt = 0$, полученные уравнения непрерывности (6.26) и (6.27) можно записать в виде

$$\frac{\partial j_n}{\partial x} = G - \frac{p - p_0}{\tau_p} - \frac{n_0}{\tau_n} - \frac{1}{q} \operatorname{div} j_p. \quad (6.30)$$

$$\frac{1}{q} \frac{\partial j_p}{\partial x} = G - \frac{p - p_0}{\tau_p}. \quad (6.31)$$

Уравнения непрерывности (6.30) и (6.31) выражают условия сохранения числа носителей заряда, т.е. в стационарном состоянии поток носителей заряда, вытекающих из объема, равен количеству носителей, созданных внешним возбуждением, за вычетом прорекомбинировавших носителей.

В общем случае, если изменение концентрации носителей заряда происходит по всем направлениям трехмерного пространства, уравнения непрерывности выглядят следующим образом:

$$\frac{\partial n}{\partial t} = G - \frac{n - n_0}{\tau_n} + \frac{1}{q} \operatorname{div} j_n, \quad (6.32)$$

$$(6.33)$$

Уравнение непрерывности является одним из фундаментальных в физике полупроводников. Оно позволяет найти распределение концентрации носителей заряда в образце при всех способах генерации, рекомбинации, диффузии и дрейфа носителей заряда.

Контрольные вопросы и задачи

1. (К8ДР). Какое выражение верно для определения дрейфового электронного тока?

Ответы:

- | | |
|----|--|
| 1) | 2) $j_{n\text{др}} = q\mu_n n \mathbf{E};$ |
| 3) | 4) $j_{n\text{др}} = qn \text{grad } V.$ |

2. (ОХДФ). С помощью какого выражения можно рассчитать диффузионный ток электронов?

Ответы:

- | | |
|---|--|
| 1) $j_{n\text{диф}} = qD_n \text{grad } n;$ | 2) $j_{n\text{диф}} = -qD_n \text{grad } n;$ |
| 3) $j_{n\text{диф}} = qn \text{grad } V;$ | 4) $j_{n\text{диф}} = n \exp(\pm qU/(kT)).$ |

3. (Л2ДТ). Какое выражение верно для определения диффузионного тока дырок?

Ответы:

- | | |
|--|---|
| 1) $j_{p\text{диф}} = -qD_p \text{grad } p;$ | 3) $j_{p\text{диф}} = qD_p \mu_p \mathbf{E};$ |
| 2) | 4) $j_{p\text{диф}} = q\mu_p \text{grad } p.$ |

4. (Ф2СЗ). Какова связь между диффузионной длиной и временем жизни носителей заряда?

Ответы:

- | | |
|------------------------|----|
| 1) $L = \sqrt{D\tau};$ | 2) |
| 3) | 4) |

5. (КГДИ). Каков физический смысл понятия «диффузионная длина»?

Ответы:

- 1) расстояние, проходимое носителем заряда за время жизни;
- 2) расстояние, на котором контактное электрическое поле уменьшается в 2,71 раза;
- 3) расстояние, на котором объемный заряд неравновесных носителей компенсируется объемным зарядом нескомпенсированных примесных атомов;

$$j_{n\text{др}} = q\mu_n n \mathbf{E};$$

4) длина области объемного заряда.

6. (МБКД). В германии при 20 °С подвижность электронов $\mu_n = 3700 \text{ см}^2/(\text{В} \cdot \text{с})$. Определите коэффициент диффузии электронов для этого образца.

Ответ, например 49,4 $\text{см}^2/\text{с}$, наберите после пароля в виде ряда символов 49,4 $\text{см}^2/\text{с}$.

7. (ГОУН). Укажите, какое выражение называется уравнением непрерывности:

$$1) \frac{\partial n}{\partial t} = G - \frac{n - n_0}{\tau_n} + \frac{1}{q} \operatorname{div} \mathbf{j}_n; \quad 2)$$

$$3) -R_n \, dx \, dt = -\frac{n - n_0}{\tau_n} \, dx \, dt; \quad 4) \frac{\partial p}{\partial t} = G - \frac{p - p_0}{\tau_p} - \frac{\partial I_p}{\partial x};$$

$$5) n(x, t + dt) \, dx - n(x, t) \, dx = \frac{\partial n}{\partial x} \, dx \, dt;$$

$$6) I_n(x, t) \, dt - I_n(x + dx, t) \, dt = -\frac{\partial I_n}{\partial x} \, dx \, dt.$$

8. (У4ИС). Для каких целей в физике твердого тела используется уравнение непрерывности?

Ответы:

1) с помощью этого уравнения можно определить время жизни нерав-

$\frac{\partial n}{\partial t} = G - \frac{n - n_0}{\tau_n}$ новесных носителей заряда;

2) это уравнение позволяет найти распределение концентрации избыточных носителей заряда в образце при всех способах генерации, рекомбинации, диффузии и дрейфа носителей заряда;

3) используя это уравнение, можно определить объемные заряды в образце.

9. (ИРДИ). С использованием необходимых данных из задачи 6 вычислите диффузионную длину электронов в германии при 20 °С, если их время жизни равно 5 нс.

Ответ с точностью до сотых долей множителя с указанием единицы физической величины, например $1,78 \cdot 10^{-3} \text{ см}$, наберите в виде ряда символов 1,78 – 3 см.

10. (МСПД). Определите диффузионную длину дырок, если концентрация неравновесных носителей заряда на расстоянии $x_1 = 2 \text{ мм}$ от места генерации $\Delta p_1 = 10^{14} \text{ см}^{-3}$, а на расстоянии $x_2 = 4,3 \text{ мм}$ $\Delta p_2 = 10^{13} \text{ см}^{-3}$.

11. (Г5СД). Вычислите скорости дрейфа электронов и дырок в собственном полупроводнике германии при 20 °С, если градиент потенциала в образце достигает 500 В/м, а подвижности электронов и дырок равны соответственно 3800 и 1800 $\text{см}^2/(\text{В} \cdot \text{с})$.

Ответ наберите в виде двух целых чисел последовательно для электронов и дырок с указанием в конце единицы физической величины.

12. (К8ДР). Выберите правильное определение понятия «дрейфовый ток»:

- 1) ток, возникающий при наличии градиента концентрации носителей заряда;
- 2) ток, возникающий под действием электрического поля;
- 3) ток, обусловленный поперечным магнитным полем;
- 4) движение свободных носителей заряда, возникающее при облучении полупроводника светом.

13. (Ю6МД). Какой эффект из нижеприведенных является следствием монополярной диффузии?

Ответы:

- 1) возникновение в образце магнитного поля;
- 2) возникновение в образце электрического поля;
- 3) усиление градиента концентрации носителей заряда в образце;
- 4) нарушение электронейтральности образца.

14. (ГБСД). Выберите из нижеприведенных соотношений выражения, правильно определяющие скорость диффузионного движения носителей заряда:

1) $v_{\text{диф}} = \frac{L}{\tau}$, где L — диффузионная длина носителя, τ — время жизни носителя в зоне;

2) $v_{\text{диф}} = \sqrt{\frac{2D}{\tau}}$, где D — коэффициент диффузии носителей заряда;

3) $v_{\text{диф}} = \sqrt{\frac{2k_0 T}{m^*}}$, где k_0 — постоянная Больцмана, T — температура, m^* — эффективная масса носителя;

4) $v_{\text{диф}} = \sqrt{\frac{k_0 T \mu}{q \tau}}$, где μ — подвижность носителей заряда;

5) $v_{\text{диф}} = \frac{q D}{k_0 T}$

$$v_{\text{диф}} = \frac{2k_0 T}{\sqrt{m^*}}$$

Глава 7. Контактные явления

§ 7.1. Работа выхода электрона

Контакты металл-полупроводник могут служить основой выпрямляющих диодов или осуществлять контакт с внешней цепью. Термоэлектрические генераторы, холодильные устройства тоже основаны на использовании свойств контактов металл-металл, металл-полупроводник или полупроводник-полупроводник. Контакты двух полупроводников широко используются в полупроводниковых приборах и интегральных схемах.

При сближении атомов и образовании из них твердого тела потенциальный барьер у поверхности кристалла остается практически таким же высоким, как и у изолированных атомов (рис. 7.1).

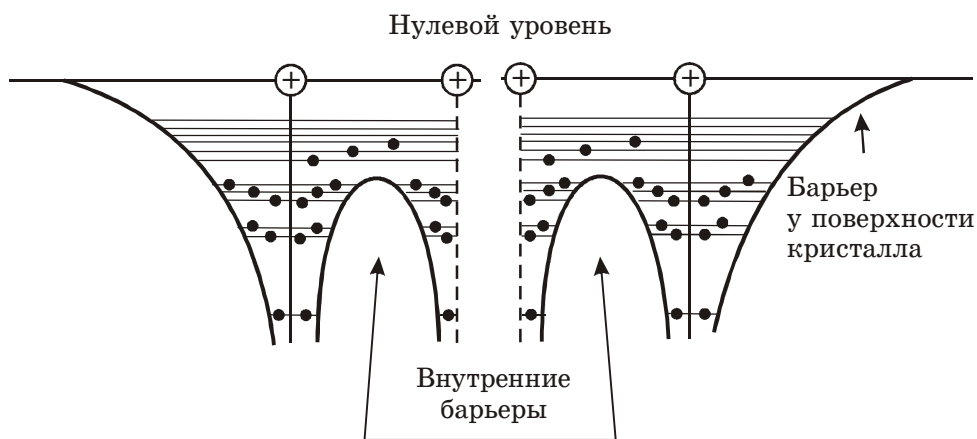


Рис. 7.1. Потенциальный барьер для электронов у поверхности кристалла

Чтобы удалить электроны из кристалла, требуется совершить определенную работу, называемую работой выхода электронов. В металлах электроны, участвующие в электропроводности, находятся на уровнях, расположенных вблизи уровня Ферми. Поэтому работа выхода электронов из металла χ , определяемая как работа, которую необходимо совершить для перевода электрона с уровня Ферми до нулевого уровня в вакууме (рис. 7.2), получила название термодинамической работы выхода.

За нулевой уровень вакуума принимается минимальная энергия, которую может иметь электрон в вакууме при отсутствии внешних воздействий.

В полупроводниках и диэлектриках на уровне Ферми нет электронов. В этих материалах электроны находятся либо в зоне проводимости, либо в валентной зоне. Для них вводятся понятия внешней и внутренней работы выхода (рис. 7.3).

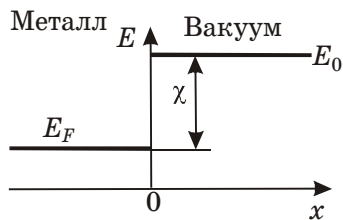


Рис. 7.2. Работа выхода электронов из металла

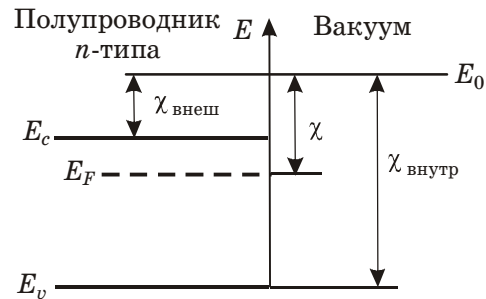


Рис. 7.3. Изображение термодинамической работы выхода на энергетической диаграмме полупроводников и диэлектриков

Внешняя работа выхода $\chi_{\text{внеш}}$ определяется как работа, которую нужно совершить, чтобы перевести электрон со дна зоны проводимости (с уровня E_c) на нулевой уровень вакуума E_0 . Эта величина $\chi_{\text{внеш}} = E_0 - E_c$ называется также электронным средством.

Внутренняя работа выхода $\chi_{\text{внутр}}$ — это работа, требуемая для перевода электрона с потолка валентной зоны E_v на нулевой уровень вакуума E_0 .

В среднем для полупроводника затраты энергии на преодоление потенциального поверхностного барьера при переводе электрона из твердого тела в вакуум оказываются равными термодинамической работе выхода χ , определяемой разностью энергий нулевого уровня вакуума E_0 и уровня Ферми E_F . Как у металлов, так и у полупроводников термодинамическая работа выхода составляет обычно несколько электронвольт. В полупроводниках положение уровня Ферми зависит от температуры, природы и концентрации введенных примесей. Следовательно, и значение термодинамической работы выхода для них будет зависеть от этих факторов. В литературе часто используется термин «работа выхода». Следует иметь в виду, что это относится к термодинамической работе выхода. Поскольку уровень Ферми в дырочном полупроводнике лежит ниже, чем в электронном (см. § 2.5), работа выхода электронов из полупроводника p -типа всегда больше, чем из того же полупроводника n -типа.

Большое влияние на термодинамическую работу выхода оказывают мономолекулярные слои, адсорбированные на поверхности металла. При адсорбции одновалентных электроположительных металлов на поверхности вольфрама образуется двойной заряженный электрический слой, поле которого облегчает выход электронов из вольфрама, т.е. работа выхода при этом уменьшается. Например, если поверхность вольфрама покрыть тонким слоем цезия, то работа выхода электронов уменьшится с 4,52 до 1,36 эВ.

При адсорбции кислорода наблюдается увеличение работы выхода электронов с поверхности вольфрама. Дело в том, что связь валентных

электронов в атоме кислорода значительно сильнее, чем в атомах электроположительных металлов. Поэтому при осаждении на поверхность вольфрама кислород не отдает, а наоборот, отбирает электрон у металла, превращаясь в отрицательно заряженный ион. Вследствие этого внешняя сторона двойного электрического слоя оказывается заряженной отрицательно, создавшееся электрическое поле тормозит выход электронов из металла и работа выхода увеличивается.

Уменьшение работы выхода под влиянием электроположительных элементов используется при изготовлении катодов электронных ламп, фотокатодов и т.п.

§ 7.2. Контактная разность потенциалов

Поскольку на границе твердого тела существует энергетический барьер, препятствующий выходу электронов из кристалла, то покинуть его могут лишь электроны, имеющие энергию, достаточную для преодоления этого барьера. Чем выше будет температура, тем большее количество электронов сможет преодолеть потенциальный барьер. Явление выхода электронов из вещества за счет тепловой энергии называется термоэлектронной эмиссией. Плотность тока термоэлектронной эмиссии описывается формулой Ричардсона — Дешмена

$$j = AT^2 e^{-\frac{\chi}{k_0 T}}, \quad (7.1)$$

где $A = \frac{4\pi m k_0^2 q}{h^3} = 12 \cdot 10^5 \frac{\text{А}}{\text{м}^2 \cdot \text{К}}$ — постоянная Ричардсона; χ — работа выхода электронов.

Из формулы (7.1) следует, что термоэлектронный ток быстро увеличивается с ростом температуры и с уменьшением работы выхода. Например, если температуру вольфрамового катода повысить от 1000 до 2500 К, то плотность тока термоэлектронной эмиссии увеличится примерно на 16 порядков; покрытие вольфрамового катода моноатомным слоем цезия, уменьшающим работу выхода с 4,52 до 1,36 эВ, вызывает увеличение плотности термоэлектронного тока примерно на 14 порядков. Рассмотрим процессы, происходящие при сближении двух металлов на расстояние d_0 , при котором вследствие термоэлектронной эмиссии возможен эффективный обмен электронами. Предположим, что работа выхода электронов χ_1 из металла 1 больше, чем χ_2 из металла 2. Энергетическая диаграмма для этого случая при расстоянии между металлами $d \gg d_0$ показана на рис. 7.4,а; при расстоянии d_0 (рис. 7.4,б) металлы могут обмениваться электронами.

Плотность тока термоэлектронной эмиссии из металла 1 в металл 2 в первый момент будет равна

(7.2)

а встречного тока из металла 2 в металл 1 при той же температуре —

$$j_{21} = AT^2 e^{-\frac{\chi_2}{k_0 T}}. \quad (7.3)$$

Так как $\chi_2 < \chi_1$ (см. рис. 7.4,а), то $j_{21} > j_{12}$, т.е. более интенсивный перенос электронов будет происходить из металла 2 в металл 1. Тогда поверхность первого металла будет заряжаться отрицательно, а второго — положительно. В зазоре возникает электрическое поле \mathbf{E}_k и соответствующая разность потенциалов

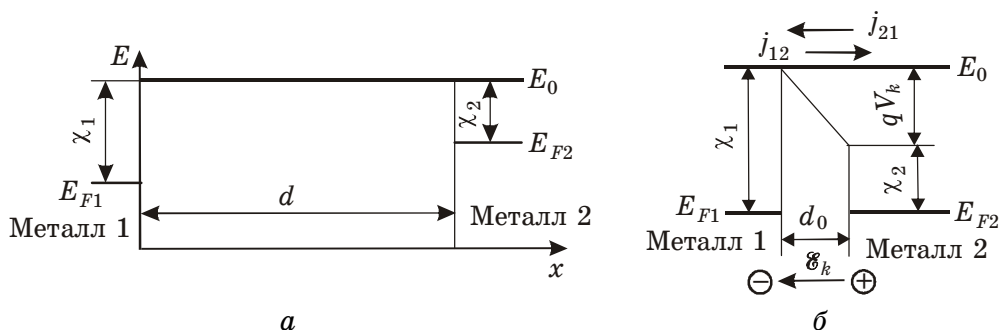


Рис. 7.4. К возникновению контактной разности потенциалов между различными металлами: а — $d \gg d_0$; б — $d = d_0$

Возникшее электрическое контактное поле является тормозящим для электронов, эмитируемых из второго металла, и ускоряющим для электронов из металла 1, вследствие чего токи выравниваются. Уровни Ферми в металлах будут иметь одинаковое положение, когда установится контактная разность потенциалов V_k , выравнивающая высоту потенциального барьера для электронов обоих металлов (рис. 7.4,б), т.е.

$$\chi_1 = \chi_2 + qV_k. \quad (7.4)$$

Таким образом, после установления равновесия контактная разность потенциалов определяется формулой

$$V_k = (\chi_1 - \chi_2)/q, \quad (7.5)$$

где V_k называется внешней контактной разностью потенциалов.

Оценим число электронов, переход которых из одного металла в другой создает обычно наблюдаемую контактную разность потенциалов. Пусть $\chi_1 - \chi_2 = 1$ эВ, а $d_0 = 10^{-7}$ см. Тогда $\mathbf{E}_k = V_k/d_0 = 10^7$ В/см. Если рассматриваемую систему представить в виде плоского конденсатора, то металл 1 и металл 2 будут обкладками воображаемого конденсатора, плотность заряда на которых определяется по формуле $Q_S = \epsilon_0 \mathbf{E}_k = 8,86 \cdot 10^{-14} \cdot 10^7 \approx 9 \times 10^{-7}$ Кл/см². Число электронов n , необходимое для создания Q_S :

$$V_k = \mathbf{E}_k d_0, \\ j_{12} = AT^2 e^{-\frac{\chi_1}{k_0 T}}$$

$$n = \frac{Q_S}{q} \approx \frac{9 \cdot 10^{-7}}{1,6 \cdot 10^{-19}} \approx 5,6 \cdot 10^{12} \text{ см}^{-2}.$$

Число свободных электронов в металле n_S , приходящихся на 1 см^2 поверхности, можно определить из объемной концентрации электронов зоны проводимости n_v :

$$n_S = (n_v)^{2/3} = (10^{22} \text{ см}^{-3})^{2/3} \approx 10^{15} \text{ см}^{-2}.$$

Если сравним n и n_S , то увидим, что $n/n_S = 5 \cdot 10^{-3}$. Это значит, что для создания контактной разности потенциалов в 1 В требуется лишь незначительная часть электронов ($\approx 10^{-3}$), приходящихся на единицу поверхности металла.

§ 7.3. Контакт металл-полупроводник

Физическая картина возникновения разности потенциалов в контакте металла с полупроводником принципиально не отличается от контакта двух металлов (§7.2).

Рассмотрим контакт металла с донорным полупроводником. Их энергетические диаграммы для случая, когда обмен электронами отсутствует, показаны на рис. 7.5,а. Предполагается, что термодинамическая работа выхода электронов из металла выше, чем из полупроводника, т.е.

$$j_{\text{ПМ}} = AT^2 e^{-\frac{\chi_{\text{П}}}{k_0 T}}, \quad \chi_{\text{М}} > \chi_{\text{П}}. \quad (7.6)$$

Тогда при контакте металла с полупроводником ток электронов, эмитированных из полупроводника в металл,

$$(7.7)$$

в первый момент превышает ток электронов $j_{\text{МП}}$, эмитированных из металла,

$$j_{\text{МП}} = AT^2 e^{-\frac{\chi_{\text{М}}}{k_0 T}}. \quad (7.8)$$

В результате металл заряжается отрицательно, а полупроводник — положительно. Возникшее электрическое поле препятствует переходу электронов из полупроводника в металл. Направленный поток электронов будет иметь место до тех пор, пока уровни Ферми в системе не сравняются (рис. 7.5,б), после чего установится динамическое равновесие, характеризующееся равенством токов

$$j_{\text{ПМ}} = j_{\text{МП}}. \quad (7.9)$$

Приравнявая выражения для токов из металла и полупроводника, определим значение контактной разности потенциалов:

$$j_{\text{ПМ}} = AT^2 e^{-\frac{\chi_{\text{П}} + \varphi_0}{k_0 T}} = j_{\text{МП}} = AT^2 e^{-\frac{\chi_{\text{М}}}{k_0 T}},$$

$$\varphi_0 = qV_k = \chi_{\text{М}} - \chi_{\text{П}}. \quad (7.10)$$

Так как напряженность электрического поля в слое объемного заряда $\mathbf{E}_k = Q_S / \epsilon \epsilon_0$, то при равенстве объемных зарядов (в металле и полупроводнике) разность потенциалов в контактирующих телах пропорциональна толщине слоя объемного заряда l , т.е.

$$V = \mathbf{E}_k l = \frac{Q_S}{\epsilon \epsilon_0} l. \quad (7.11)$$

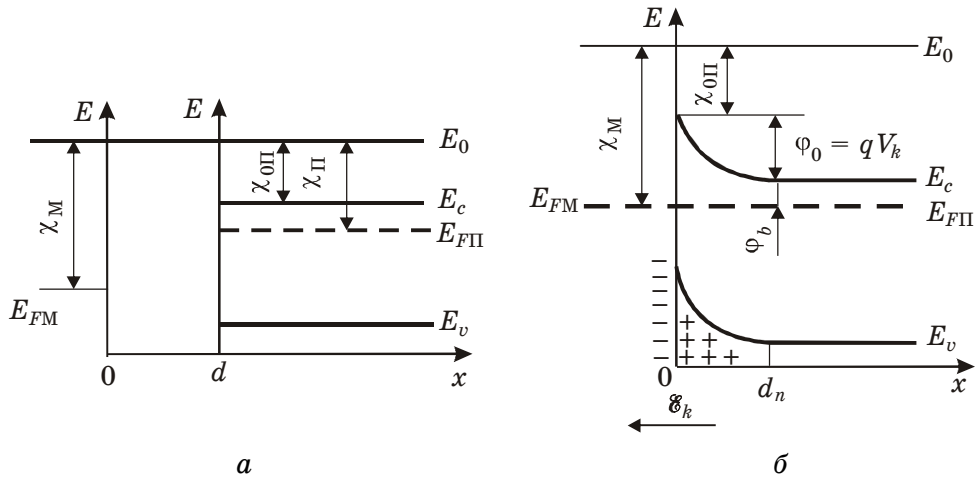


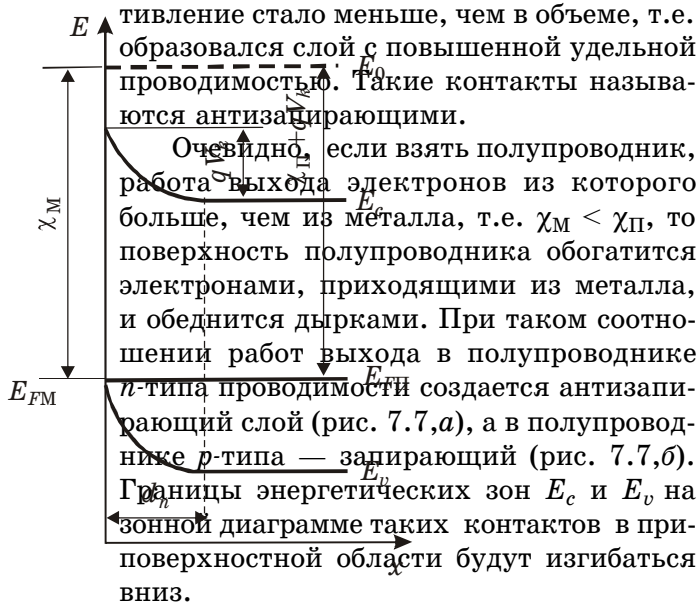
Рис. 7.5. Энергетические диаграммы металла и полупроводника до контактирования (а) и при тесном контакте (б)

Как было показано в § 7.2, объемный заряд, создающий контактную разность потенциалов, в металле сосредоточен в приповерхностном слое толщиной $\sim 10^{-7}$ см. В полупроводниках, в зависимости от концентрации примеси, плотность электронов в моноатомном слое может быть порядка $10^6 - 10^8 \text{ см}^{-2}$, что значительно меньше числа электронов, необходимых для создания контактной разности потенциалов ($n \approx 5 \cdot 10^{12} \text{ см}^{-2}$ при $V_k \approx 1 \text{ В}$). Поэтому для создания в металле необходимой поверхностной концентрации избыточных электронов требуется, чтобы в металл перешли электроны из глубины полупроводника. Это приведет к обеднению приконтактной области полупроводника подвижными носителями заряда.

Протяженность такого обедненного слоя составит от 10^2 до 10^6 атомных слоев при условии, что все электроны в этом объеме покинут атомы доноров и перейдут в металл. В свете сказанного можно сделать вывод, что протяженность области объемного заряда и разность потенциалов на

ней в металле ничтожно малы по сравнению с аналогичными величинами в полупроводнике. Следовательно, можно считать, что контактная разность потенциалов V_k практически полностью приложена к приконтактной области полупроводника.

Напряженность контактного электрического поля E_k обычно не превышает 10^6 В/см, а напряженность периодического поля решетки составляет 10^8 В/см. Поэтому небольшая добавка за счет контактного поля не может изменить ширину запрещенной зоны полупроводника, но приведет к искривлению энергетических зон относительно уровня Ферми, а это вызовет изменение концентрации электронов и дырок в его приповерхностном слое. В рассматриваемом случае контакта металла и полупроводника n -типа проводимости при $\chi_M > \chi_{II}$ полупроводник заряжается положительно, и границы зон искривляются кверху. При этом концентрация электронов у поверхности понижается, а дырок — увеличивается по сравнению с объемом (см. рис. 7.5,б). Электроны, являясь основными носителями заряда, определяют сопротивление полупроводника. Так как их концентрация в поверхностном слое уменьшилась, сопротивление слоя возросло. Такие слои с пониженной удельной проводимостью называются запирающими или блокирующими. Если в рассматриваемом контакте металла с полупроводником возьмем проводник p -типа (при условии, что $\chi_M > \chi_{II}$ (рис. 7.6)), то окажется, что поверхностный слой будет обогащен основными носителями заряда дырками. Его сопротивление стало меньше, чем в объеме, т.е.



Итак, можно сделать вывод, что свойства контакта определяются соотношением работ выхода электронов из металла

Рис. 7. 6. Образование антизапирающего контакта металл-полупроводник p -типа при $\chi_M > \chi_{II}$

и полупроводника и характером электропроводности полупроводника (*n*-тип, *p*-тип). Мы рассмотрели качественную картину процессов в контакте металл-полупроводник. Для количественных расчетов необходимо решить уравнение Пуассона

$$(7.12)$$

где $\varphi(x)$ — изменение потенциальной энергии электрона при его перемещении в области, занятой объемными зарядами; $\rho(x)$ — объемная плотность заряда; ϵ — относительная диэлектрическая проницаемость полупроводника; ϵ_0 — диэлектрическая постоянная вакуума.

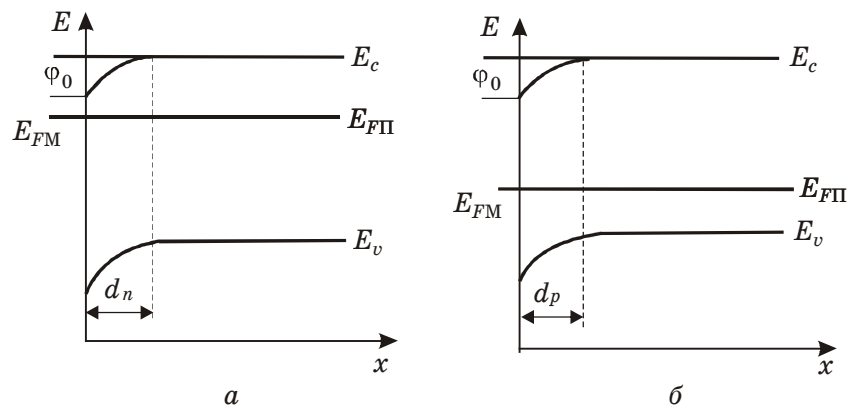


Рис. 7.7. Зонные диаграммы контакта металл-полупроводник при $\chi_M < \chi_{II}$ в равновесном состоянии:
a — полупроводник *n*-типа; *б* — полупроводник *p*-типа

Для вычисления объемной плотности заряда примем, что при данной температуре все атомы доноров ионизированы, т.е. все электроны из слоя объемного заряда перешли в металл. Тогда $\rho = qN_d$; подставив это значение в (7.12), получим

$$\frac{d^2 \varphi(x)}{dx^2} = \frac{q^2}{\epsilon \epsilon_0} N_d. \quad (7.13)$$

При решении этого уравнения следует иметь в виду, что на расстоянии $x \geq d_n$ контактное поле в полупроводнике отсутствует, поэтому граничные условия можно записать следующим образом:

$$\varphi(d) = 0; \quad \left(\frac{d\varphi}{dx} \right)_{x=d_n} = 0, \quad (7.14)$$

где d_n — толщина слоя объемного заряда в полупроводнике *n*-типа.

Интегрирование (7.13) с граничными условиями (7.14) дает следующий результат:

$$\varphi(x) = \frac{q^2 N_d}{2\varepsilon\varepsilon_0} (d_n - x)^2. \quad (7.15)$$

Из выражения (7.15) следует, что при увеличении x (при удалении от поверхности вглубь полупроводника) потенциал контактного поля убывает по параболическому закону. Для $x = 0$ будем иметь $\varphi_0 = \chi_M - \chi_{\text{П}}$. Подставив φ_0 в выражение (7.15), найдем толщину обедненного слоя:

$$d_n = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{q^2 N_d} \right)^{1/2} = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{q^2 n_n} \right)^{1/2}, \quad (7.16)$$

где $n_n = N_d$ — концентрация электронов в полупроводнике n -типа проводимости.

Из выражения (7.16) видно, что толщина обедненного слоя возрастает при увеличении разности работ выхода электронов из металла и полупроводника и уменьшается при увеличении концентрации примеси в полупроводнике.

§ 7.4. Влияние внешнего электрического поля на высоту потенциального барьера и толщину запирающего слоя

$\Phi_k(V) = \varphi_0 - qV$ Энергетическая зонная диаграмма для равновесного состояния контакта металл-полупроводник показана на рис. 7.5,б. Потенциальный барьер для электронов, переходящих из металла в полупроводник, равен разности $\chi_M - \xi_{0\text{П}}$; для электронов, переходящих из полупроводника в металл, он составляет $\varphi_0 = qV_k$. В состоянии равновесия поток электронов из полупроводника равен потоку электронов из металла, поэтому результирующий ток через контакт равен нулю. После приложения к контакту внешней разности потенциалов в прямом направлении (на металле — положительный потенциал, на полупроводнике — отрицательный) металл заряжается положительно, а полупроводник — отрицательно относительно друг друга. В этом случае направления векторов напряженности внешнего \mathbf{E} и внутреннего контактного полей не совпадают (рис. 7.8,а), вследствие чего контактная разность потенциалов, т.е. высота потенциального барьера φ_0 , уменьшается на величину внешнего напряжения

$$(7.17)$$

На такую же величину поднимаются все энергетические уровни, включая уровень Ферми, в полупроводнике относительно уровней в металле.

Уменьшение высоты потенциального барьера для электронов, переходящих из полупроводника в металл, вызовет увеличение потока электронов из полупроводника. При этом часть ионов доноров, создающих

объемный заряд в приконтактной области полупроводника, оказывается скомпенсированной подошедшими к поверхности электронами, и протяженность области пространственного заряда уменьшается. Математически этот факт отражается заменой в формуле (7.16) φ_0 на $\varphi = \varphi_0 - qV$:

$$d_n = \left(\frac{2\epsilon\epsilon_0 (\varphi_0 - qV)}{q^2 n_n} \right)^{1/2}. \quad (7.18)$$

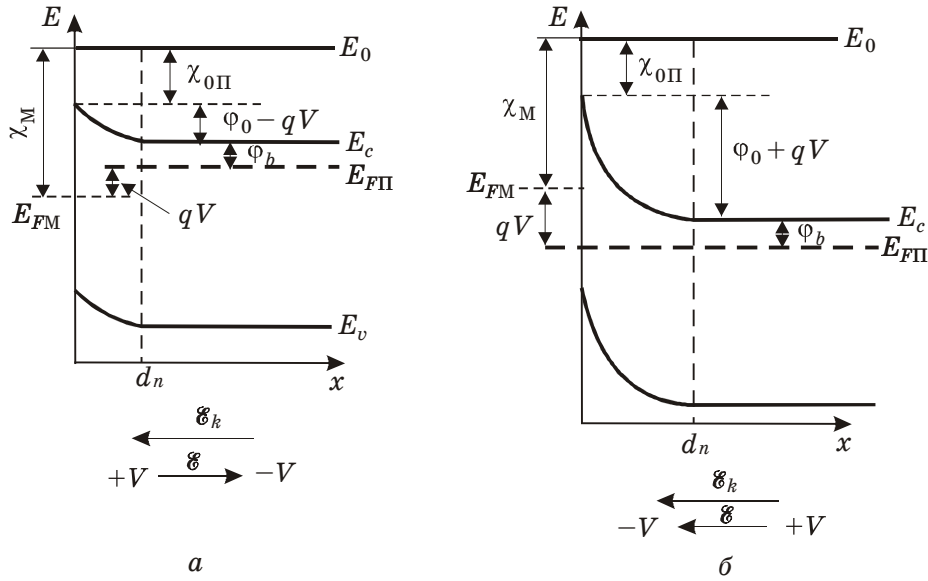


Рис. 7.8. Зонные диаграммы контакта металл-полупроводник n -типа ($\chi_M > \chi_{П}$) при наличии внешнего электрического поля:
 $a - V > 0$; $b - V < 0$

В случае приложения к контакту внешней разности потенциалов в обратном (запирающем) направлении, т.е. отрицательный потенциал — на металле, положительный — на полупроводнике (рис. 7.8, б), внешнее электрическое поле совпадает с внутренним, и все энергетические уровни в полупроводнике опускаются вниз. Потенциальный барьер возрастает на величину внешней разности потенциалов

$$(7.19)$$

В отличие от прямого смещения ширина области объемного заряда при обратном напряжении увеличивается за счет выталкивания электронов, не имеющих достаточной энергии для преодоления возросшего потенциального барьера, в глубь полупроводника:

$$d_n = \left[\frac{2\epsilon\epsilon_0 (V_k + V)}{qn_{n0}} \right]^{1/2} = \left[\frac{2\epsilon\epsilon_0 (\varphi_0 + qV)}{q^2 n_{n0}} \right]^{1/2}. \quad (7.20)$$

Из выражений (7.18) и (7.20) видно, что внешняя разность потенциалов, приложенная к контакту металл-полупроводник в прямом направлении, уменьшает толщину запирающего слоя — область пространственного заряда (ОПЗ), а в обратном направлении — увеличивает (см. рис. 7.8).

§ 7.5. Вольт-амперная характеристика контакта металл-полупроводник

В § 7.3 показано, что в равновесных условиях, т.е. при $V = 0$, равенство токов термоэлектронной эмиссии из металла в полупроводник $j_{МП}$ и из полупроводника в металл $j_{ПМ}$ обусловлено возникновением потенциального барьера высотой ϕ_0 .

Используя выражения (7.7) и (7.8), вышеизложенное можно записать в математической форме следующим образом:

$$j_{МП} = j_{ПМ}, \quad (7.21)$$

где

$$j_{МП} = \frac{4\pi m k_0^2 q T^2}{h^3} e^{-\frac{\chi_M}{k_0 T}}, \quad (7.22)$$

$$j_{ПМ} = \frac{4\pi m k_0^2 q T^2}{h^3} e^{-\frac{\chi_{0П} + \phi_0 + \phi_b}{k_0 T}}. \quad (7.23)$$

Пусть на контакт металл-полупроводник в прямом направлении подаю внешнее смещение V , которое приводит к снижению высоты потенциального барьера на величину, равную qV . Тогда ток термоэлектронной эмиссии из полупроводника в металл в соответствии с (7.7) будет зависеть от напряжения смещения:

$$j_{ПМ}(V) = \frac{4\pi m k_0^2 q T^2}{h^3} e^{-\frac{\chi_{0П} + \phi_0 - qV + \phi_b}{k_0 T}}. \quad (7.24)$$

Из (7.24) следует, что ток из полупроводника в металл увеличился в $e^{\frac{qV}{k_0 T}}$ раз, а ток из металла в полупроводник не изменился и по-прежнему описывается формулой (7.8), так как для электронов, переходящих из металла в полупроводник, высота потенциального барьера не зависит от приложенного напряжения и условия перехода электронов при изменении напряжения не изменяются.

Результирующий ток, который называется прямым током $j_{пр}$, равен разности токов $j_{ПМ}$ и $j_{МП}$ (выражения (7.23) и (7.22)):

$$j_{пр} = j_{ПМ} - j_{МП} = \frac{4\pi m k_0^2 q T^2}{h^3} e^{-\frac{\chi_{0П} + \phi_0 + \phi_b}{k_0 T}} \left(e^{\frac{qV}{k_0 T}} - 1 \right). \quad (7.25)$$

Если выражение перед скобками обозначить через j_s ,

(7.26)

то формула для прямого тока через контакт металл-полупроводник будет иметь вид

$$j_{\text{пр}} = j_{\text{ПМ}} - j_{\text{МП}} = j_s \left(e^{\frac{qV}{k_0T}} - 1 \right). \quad (7.27)$$

В случае приложения к контакту внешней разности потенциалов в обратном направлении высота потенциального барьера увеличивается, и термоэлектронный ток из полупроводника в металл уменьшается, так как экспоненциально уменьшается число электронов полупроводника, способных преодолеть возросший потенциальный барьер. Термоэлектронный ток из металла не изменяется. Следовательно, результирующий ток при такой полярности напряжения изменит направление. Проведя расчет, аналогичный предыдущему, получим выражение для обратного тока $j_{\text{обр}}$, который соответствует включению внешней разности потенциалов в обратном направлении:

(7.28)

Объединив выражения (7.27) и (7.28), получим уравнение статической вольт-амперной характеристики контакта металл-полупроводник:

(7.29)

Вольт-амперная характеристика (ВАХ) контакта металл-полупроводник показана на рис. 7.9. Для прямого смещения $V > 0$, при обратном

смещении $V < 0$. При обратном смещении величина $j_{\text{обр}}$ быстро уменьшается до нуля, и обратный ток при изменении смещения остается равным току насыщения j_s (плотность тока насыщения). Учет диффузионных процессов при больших обратных смещениях показывает слабую (линейную) зависимость плотности тока насыщения j_s от приложенного напряжения.

При прямом смещении плотность тока увеличивается в $e^{\frac{qV}{k_0T}}$ раз по сравнению с j_s , т.е. прямой ток увеличивается экспоненциально при $V > 0$. Из ВАХ (см. рис. 7.9) видно, что контакт металл-полупроводник обладает

$$j_{\text{обр}} = j_s \left(e^{-\frac{qV}{k_0T}} - 1 \right)$$

выпрямляющим действием: он пропускает ток в прямом направлении и значительно хуже — в обратном. Отношение $j_{пр} / j_{обр}$ при одной и той же величине V называется коэффициентом выпрямления, который для качественных контактов может быть выше 10^5 .

Потенциальный барьер в выпрямляющем контакте металла с полупроводником называется барьером Шоттки, а приборы на основе такого контакта, обладающие ВАХ типа приведенной на рис. 7.9, — диодами Шоттки, по имени немецкого физика В. Шоттки, который первым построил теорию полупроводниковых диодов с таким барьером.

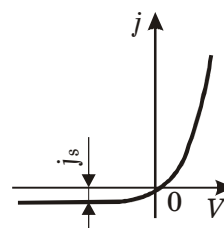


Рис. 7.9.
Вольт-амперная характеристика диода Шоттки

Контрольные вопросы и задачи

1. (Ш4РВ). Выберите правильное определение понятия «термодинамическая работа выхода»:

- 1) работа, которую необходимо совершить, чтобы перевести электрон из твердого тела в вакуум;
- 2) работа, которую необходимо совершить, чтобы перевести электрон из зоны проводимости в вакуум;
- 3) работа, которую необходимо совершить, чтобы перевести электрон из валентной зоны в вакуум;
- 4) работа, которую необходимо совершить, чтобы перевести электрон с уровня Ферми на нулевой уровень энергии в вакууме.

2. (ФЯКП). Введение акцепторной примеси в донорный полупроводник увеличивает или уменьшает работу выхода?

Ответы:

- 1) уменьшает, поскольку введение акцепторной примеси в донорный полупроводник снижает концентрацию основных носителей заряда;
- 2) увеличивает, поскольку введение акцепторной примеси в донорный полупроводник понижает уровень Ферми;
- 3) увеличивает, поскольку введение акцепторной примеси в донорный полупроводник изменяет энергию электронного сродства;
- 4) уменьшает, поскольку введение акцепторной примеси в донорный полупроводник увеличивает концентрацию дефектов в полупроводнике.

3. (ВССТ). Установление стационарного состояния при контакте двух различных материалов происходит в результате:

- 1) диффузионного перехода части носителей из материала с меньшей работой выхода в контактирующий материал, что ведет к созданию контактной разности потенциалов и появлению встречного диффузионного дрейфового потока носителей;

2) нарушения электронейтральности каждого из контактирующих материалов, что сопровождается возникновением контактного электрического поля, препятствующего дальнейшему нарушению электронейтральности;

3) выравнивания диффузионного и дрейфового потоков носителей заряда через контакт;

4) изменения удельной проводимости прилегающей к контакту области полупроводника.

4. (ФЮПЗ). Из перечисленных ниже параметров укажите по порядку те, которые определяют толщину области пространственного заряда в диоде Шоттки:

1) внешнее напряжение на диоде;

2) время жизни неравновесных носителей заряда;

3) эффективные массы носителей заряда;

4) концентрация основных носителей заряда;

5) механическая твердость полупроводника;

6) абсолютная диэлектрическая проницаемость полупроводника;

7) контактная разность потенциалов;

8) полный заряд носителей;

9) подвижность носителей заряда.

5. (ЮБПЗ). Оцените ширину области объемного заряда контакта металл-полупроводник, находящегося в равновесном состоянии, если $\phi_0 = 1$ эВ. Относительная диэлектрическая проницаемость полупроводника равна 11,8, концентрация легирующей примеси составляет 10^{22} м⁻³. Контакт находится при температуре выше температуры истощения примеси, но меньше температуры наступления собственной проводимости.

Ответ с точностью до десятых долей сомножителя с указанием единицы физической величины, например $5,93 \cdot 10^{-8}$ м, наберите после кода задания в виде ряда символов: 5 , 9 –8 м.

6. (ШРЗК). Какие контакты металл-полупроводник называются запирающими?

Ответы:

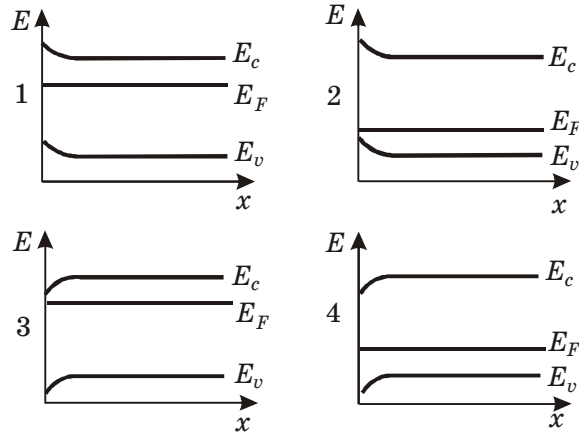
1) контакты, в которых отсутствует заметный ток при подаче внешнего напряжения любой полярности;

2) контакты, в которых сопротивление слоя объемного заряда при $V < 0$ оказывается значительно больше сопротивления остальной части полупроводника;

3) контакты, у которых сопротивление области объемного заряда при $V > 0$ значительно меньше сопротивления остальной части полупроводника;

4) контакт, обогащенный основными носителями заряда, перешедшими из металла в приповерхностный слой полупроводника.

7. (МКЗД). Укажите верное изображение энергетической зонной диаграммы запирающего контакта металл-полупроводник n -типа проводимости.



8. (Б5ВК). Для каких целей используются в радиоэлектронике запирающие контакты металл-полупроводник?

Ответы:

- 1) для стабилизации напряжения на нагрузочном резисторе;
- 2) для выпрямления тока, в качестве высокоскоростных логарифмирующих элементов;
- 3) в качестве тонкопленочных конденсаторов;
- 4) в качестве управляемых резисторов.

9. (Г5АК). Для чего используются антизапирающие контакты металл-полупроводник?

Ответы:

- 1) эти контакты имеют нелинейную несимметричную вольт-амперную характеристику, их используют для выпрямления тока;
- 2) вследствие того что такие контакты обладают малым сопротивлением и имеют почти линейную вольт-амперную характеристику, их используют как омические контакты;
- 3) в качестве управляемой напряжением емкости (варикапы);
- 4) поскольку сопротивление контакта зависит от величины и полярности приложенного напряжения, такие контакты можно использовать в качестве переменных резисторов.

10. (НДОЗ). Как влияет концентрация легирующей примеси на ширину области объемного заряда контакта металл-полупроводник?

Ответы:

- 1) увеличение концентрации примеси приводит к уменьшению ширины области объемного заряда;

2) при увеличении концентрации примеси возрастает концентрация основных носителей заряда, поэтому ширина области объемного заряда увеличивается;

3) концентрация примеси не оказывает влияния на ширину области объемного заряда;

4) увеличение концентрации примеси вызывает уменьшение времени жизни носителей заряда, поэтому ширина области объемного заряда увеличивается.

11. (Л2КП). Определите контактную разность потенциалов, которая возникнет между металлами с работами выхода $\chi_1 = 4,5$ эВ и $\chi_2 = 2,5$ эВ. Ответ наберите с указанием единицы измерения в СИ.

Глава 8. Физические явления в электронно-дырочном переходе

§ 8.1. Электронно-дырочный переход в равновесном состоянии

Электронно-дырочный переход, или p - n -переход, образуется на границе раздела двух полупроводников с разным типом проводимости (электронным и дырочным). Для образования p - n -перехода необходим очень плотный контакт между такими полупроводниками. Зазор между ними должен быть сравним с межатомным расстоянием. Очевидно, что никакая механическая обработка поверхности контактирующих материалов не позволит достичь такого зазора. Поэтому p - n -переходы получают с помощью специальных технологических приемов в одном монокристалле полупроводника. Наиболее распространенными методами получения p - n -переходов являются диффузия примесей, эпитаксиальное наращивание, ионное легирование. В зависимости от технологии можно получить различный характер распределения примесей вблизи границы раздела электронного и дырочного полупроводников (рис. 8.1). Если концентрация примесей изменяется резко, скачкообразно, то имеем дело с резким, или ступенчатым, p - n -переходом (рис. 8.1,а). Такое распределение примесей возможно получить при эпитаксиальном наращивании.

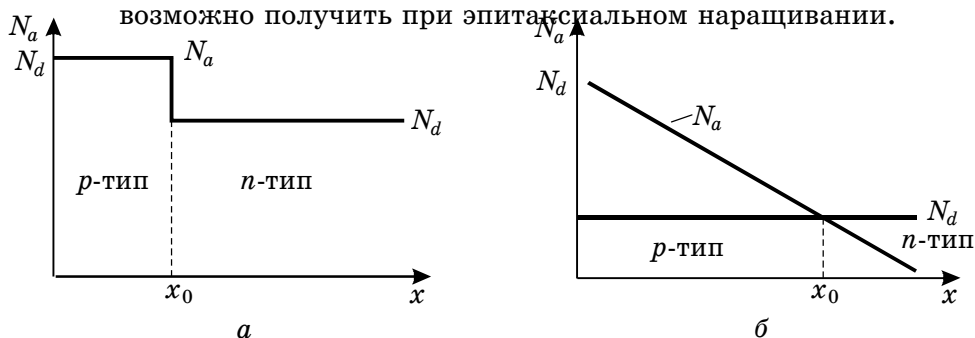


Рис. 8.1. Распределение примесей в резком (а)
и плавном (б) p - n -переходах

При диффузии примесей получают плавные p - n -переходы (рис. 8.1,б). В технике находят применение и те и другие p - n -переходы. Можно классифицировать p - n -переходы по соотношению концентраций примесей в n - и p -областях. Если концентрации примесей сравнимы, т.е. $N_d \approx N_a$, то такой p - n -переход называется симметричным, если $N_d > N_a$ или $N_a > N_d$, то имеем дело с несимметричными p - n -переходами. В электронике несимметричные переходы являются основой полупроводниковых

диодов, эмиттерными переходами в биполярных транзисторах, затворами в полевых транзисторах. Симметричными являются коллекторные p - n -переходы дрейфовых биполярных транзисторов и т.п.

Рассмотрим образование ступенчатого симметричного p - n -перехода. Для простоты рассуждений будем считать, что изменение концентрации носителей заряда происходит вдоль одной оси трехмерного пространства (вдоль оси x). Это возможно, если области p - и n -типов проводимости полубесконечны. Такое предположение позволяет не учитывать процессы генерации и рекомбинации носителей заряда на поверхности кристаллов и явления, происходящие в месте выхода p - n -перехода на поверхность.

Будем считать, что рассматриваемые полупроводники разделены плоскостью, расположенной перпендикулярно рисунку при $x = 0$ (рис. 8.2,а).

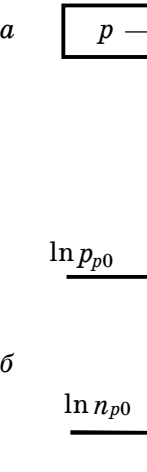


Рис. 8.2. Характеристики p - n -перехода в равновесном состоянии:

- a — внутренняя граница в полупроводнике между областями с проводимостью p - и n -типа;
- $б$ — распределение основных и неосновных носителей заряда;
- $в$ — распределение плотности объемных зарядов;
- $г$ — распределение напряженности электрического поля

Слева от границы раздела находится полупроводник p -типа проводимости (например, p -германий) с концентрацией акцепторов N_a , справа — полупроводник n -типа проводимости (n -германий) с концентрацией доноров N_d . Изменение концентраций примесей и равновесных носителей заряда вдоль оси x показано на рис. 8.2,б. Концентрация акцепторов N_a постоянна в полупроводнике p -типа и скачком уменьшается до нуля в точке $x = 0$, а концентрация доноров в этой же точке увеличивается

от нуля до N_d и остается постоянной в полупроводнике n -типа. Пусть температура кристалла такова, что примеси полностью ионизированы, тогда $p_{p0} \approx N_a$, а $n_{n0} \approx N_d$. Концентрация неосновных носителей заряда определится в соответствии с законом действующих масс:

$$n_{n0} = n_i^2 / p_{p0}.$$

Например, при $N_d = N_a = 10^{23} \text{ м}^{-3}$, $p_{p0} = n_{n0} = 10^{23} \text{ м}^{-3}$; для германия $n_i^2 = 10^{38} \text{ м}^{-6}$, тогда $p_{n0} = 10^{38} / 10^{23} = 10^{15} \text{ м}^{-3}$, $n_{p0} = 10^{15} \text{ м}^{-3}$.

Сравнивая концентрации одноименных носителей по обе стороны от границы раздела ($x = 0$), видим, что $p_{p0} \gg p_{n0}$, $n_{n0} \gg n_{p0}$, т.е. в плоскости $x = 0$ возникает градиент концентраций носителей заряда одного знака, который обуславливает диффузию носителей из области с большей концентрацией в область с меньшей концентрацией носителей заряда. Дырки начинают диффундировать из p -области в n -область, электроны — из n -области в p -область. Так как электроны и дырки являются подвижными носителями заряда, то их концентрации не могут изменяться скачком и будут изменяться плавно от p_{p0} до p_{n0} и от n_{n0} до n_{p0} (см. рис. 8.2, б). По мере ухода дырок из приграничной области p -полупроводника в этой области остаются нескомпенсированные ионы акцепторов, создающие отрицательный объемный заряд. В приграничном слое n -полупроводника после ухода электронов остаются нескомпенсированные ионы доноров, создающие положительный объемный заряд. Распределение неподвижных объемных зарядов показано на рис. 8.2, в (ρ — объемная плотность зарядов). Через d обозначена протяженность области объемного заряда: d_p — в p -области, d_n — в n -области. Очевидно, что $d = d_n + d_p$.

В приграничных областях d_p и d_n концентрация основных носителей заряда оказывается значительно меньше, чем в объеме полупроводника. С достаточной степенью точности можно считать, что объемные заряды создаются только нескомпенсированными ионами примесей, зарядом подвижных носителей (дырок в n -области и электронов в p -области) пренебрегают.

Пространственное разделение зарядов приводит к возникновению контактного электрического поля с напряженностью и разностью потенциалов V_k . Это поле максимально на границе раздела ($x = 0$) и убывает к границам областей объемных зарядов (см. рис. 8.2, г). На рисунке стрелкой показано направление напряженности электрического поля. Возникшее электрическое поле по-разному действует на электроны и дырки, перемещающиеся через границу раздела: для основных носителей заряда это поле является тормозящим. Перейти через область объемного заряда могут только те основные носители, которые имеют достаточную энергию для преодоления противодействия контактного поля, т.е. энергия таких носителей должна превышать qV_k . Поэтому по мере увеличения объемных зарядов контактного электрического поля диффузия основных носителей заряда через границу раздела будет уменьшаться. Для неосновных

носителей, попавших за счет теплового хаотического движения в область объемных зарядов, контактное поле является ускоряющим, оно не препятствует их движению в соседнюю область. Так как поток основных носителей через переход уменьшается, а неосновных носителей заряда остается постоянным, то с течением времени оба эти потока уравновесят друг друга: диффузионный поток дырок из p -области в n -область станет равным встречному дрейфовому потоку дырок из n -области в p -область; диффузионный поток электронов из n -области в p -область уравновесится дрейфовым потоком электронов из p -области в n -область. Математически это можно записать следующим образом:

$$(8.1)$$

Таким образом, в контакте двух полупроводников устанавливается равновесное состояние.

Представим энергетические диаграммы полупроводников p - и n -типа проводимости до соприкосновения (рис. 8.3).

$$\begin{aligned} \frac{j_{p \text{ диф}}}{E_c} &= \frac{j_{n \text{ др}}}{E_v} \\ j_{p \text{ диф}} &= j_{n \text{ др}} \end{aligned}$$

$$\begin{aligned} E_{Fp} & \text{ —————} \\ E_a & \text{ - - - - -} \\ E_v & \text{ —————} \end{aligned}$$

Рис. 8.3. Энергетические диаграммы p - n -перехода:
a — неравновесная диаграмма (до начала обмена носителями заряда между p - и n -областью); *б* — равновесная диаграмма (после образования стационарных объемных зарядов в переходе)

Границы энергетических зон E_c и E_v изображены прямыми линиями. Это означает, что энергия электронов в зоне проводимости или дырок в валентной зоне во всех частях полупроводника одинакова. После установления равновесного состояния p - n -перехода вид диаграммы существенно изменяется (рис. 8.3, *б*). За пределами p - n -перехода электрическое поле отсутствует, поэтому вид энергетических диаграмм в глубине p - и n -областей остается без изменений. Вследствие обмена носителями заря-

да и выравнивания потоков носителей уровень Ферми устанавливается единым для всей системы, поэтому границы энергетических зон в n -области расположены на φ_0 ниже, чем в p -области. Это означает, что энергия электронов у дна зоны проводимости полупроводника n -типа оказывается на φ_0 меньше, чем на аналогичных уровнях в полупроводнике p -типа. Следовательно, из n -области в p -область смогут перебраться только такие электроны, которые имеют энергию, достаточную для преодоления потенциального барьера φ_0 , т.е. с энергией, превышающей E_c на величину φ_0 . Аналогичное рассуждение можно провести и для дырок, перемещающихся из p -области в n -область. Из p -области в n -область смогут продифундировать только дырки, обладающие избыточной энергией ($\geq \varphi_0$) относительно потолка валентной зоны в p -области. Действительно, потенциальный барьер, контактное поле p - n -перехода препятствуют перемещению основных носителей через область объемного заряда в соседний слой. Если энергия электронов n -области будет меньше $E_c + \varphi_0$, а энергия дырок p -области меньше $E_v - \varphi_0$, то такие основные носители заряда не смогут преодолеть потенциальный барьер, отразятся от него и возвратятся назад в ту же область.

Иначе выглядит движение неосновных носителей. Если электроны p -области n_{p0} подходят к границе области объемного заряда, то они переходят в n -область без затраты энергии. Аналогично дырки n -области p_{n0} переходят в p -область, не встречая противодействия. Движение носителей заряда легко представить себе, если электроны уподобить металлическим шарикам, а изменение положения дна зоны проводимости в p - n -переходе — наклонной плоскости. Шарик скатиться вниз по наклонной плоскости легко (перейти из p -области в n -область); «подняться» сможет далеко не каждый (из n -области в p -область). Дырки можно представить воздушными пузырьками, перемещающимися в сосуде, наполненном жидкостью. «Всплыть» пузырек может легко (дырка переходит из n -области в p -область); «утопить» пузырек (перевести дырку из p -области в n -область) значительно труднее.

Из рассмотрения зонной диаграммы можно прийти к выводу, что барьер в p - n -переходе равен

$$\varphi_0 = E_{F_n} - E_{F_p}, \quad (8.2)$$

т.е. определяется разностью положений уровней Ферми в исходных полупроводниках.

Воспользуемся известными соотношениями

$$\left. \begin{aligned} n_n &= n_i \exp\left(-\frac{E_{F_i} - E_{F_n}}{k_0 T}\right), \\ p_p &= p_i \exp\left(+\frac{E_{F_i} - E_{F_p}}{k_0 T}\right). \end{aligned} \right\} \quad (8.3)$$

Выразим отсюда

$$E_{F_n} = E_{F_i} + k_0 T \ln \frac{n_n}{n_i},$$

$$E_{F_p} = E_{F_i} + k_0 T \ln \frac{p_p}{p_i}.$$

Подставим эти значения в (8.2):

$$\varphi_0 = k_0 T \ln \frac{p_p n_n}{n_i^2}. \quad (8.4)$$

Учтем, что $n_i^2 = p_p n_p = n_n p_n$, тогда

$$\varphi_0 = q V_k = k_0 T \ln \frac{p_p}{p_n} = k_0 T \ln \frac{n_n}{n_p}. \quad (8.5)$$

Следовательно, контактная разность потенциалов увеличивается при увеличении концентрации примеси. Максимальное значение φ_0 определяется шириной запрещенной зоны полупроводника. При увеличении температуры $n_i^2 \propto e^{-\frac{\Delta E}{k_0 T}}$, а $p_p n_n \approx \text{const}$, значит, φ_0 будет уменьшаться и обратиться в нуль при $p_p n_n = n_i^2$.

Если в выражение (8.4) подставить и учесть, что концентрации основных носителей заряда значительно меньше эффективной плотности состояний в соответствующих зонах, то получим

$$\varphi_0 = \Delta E - k_0 T \ln \frac{N_c N_v}{n_n p_p}, \quad (8.6)$$

т.е., действительно, φ_0 уменьшается при увеличении температуры.

Определим φ_0 для контакта двух германиевых полупроводников различного типа проводимости при $N_a = N_d = 10^{23} \text{ м}^{-3}$ при 20°C :

$$\varphi_0 = 0,025 \ln \frac{10^{23}}{10^{15}} = 0,46 \text{ эВ}.$$

Положение уровня Ферми в запрещенной зоне определяется температурой и концентрацией примеси. Приведенная на рис. 8.3,б энергетическая диаграмма соответствует симметричному p - n -переходу, так как энергетические зазоры $E_F - E_v$ в p -области и $E_c - E_F$ в n -области равны.

В несимметричном p - n -переходе соответствующие зазоры различны, причем, чем больше концентрация примеси, тем меньше расстояние между уровнем Ферми и соответствующей границей зоны.

Для поиска вида функции $\varphi(x)$, которая определяет изменение потенциальной энергии электрона при переходе его из n -области в p -область

$$\frac{E_{F_n}, E_{F_p}}{n_i} = \left(\frac{N_c N_v}{N_c N_v} \right)$$

(или дырки при переходе ее из p -области в n -область), необходимо решить уравнение Пуассона

$$\frac{d^2 \varphi}{d x^2} = \frac{q}{\varepsilon \varepsilon_0} \rho(x), \quad (8.7)$$

где q — заряд электрона; ε — относительная диэлектрическая проницаемость полупроводника; ε_0 — абсолютная диэлектрическая проницаемость вакуума; $\rho(x)$ — объемная плотность зарядов, создающих поле.

Считаем, что объемные заряды создаются ионизированными донорами и акцепторами, слой d_n покинули все электроны, d_p — все дырки (см. рис. 8.2, б). Тогда в n -области при $0 \leq x \leq d_n$ $\rho(x) \approx qN_d = +qn_n$, а в p -области при $d_p \leq x \leq 0$ $\rho(x) = -qN_a \approx -qp_p$. Подставляя эти значения объемной плотности зарядов в уравнение (8.7), получим

$$\frac{d^2 \varphi_n}{d x^2} = \frac{q^2}{\varepsilon \varepsilon_0} N_d \quad (0 < x < d_n), \quad (8.8)$$

$$\frac{d^2 \varphi_p}{d x^2} = \frac{q^2}{\varepsilon \varepsilon_0} N_a \quad (-d_p < x < 0). \quad (8.9)$$

За пределами области объемного заряда ($x > d_n$ и $x < -d_p$) контактное поле в полупроводнике отсутствует, тогда решение уравнений (8.8) и (8.9) проведем при следующих граничных условиях:

$$\left. \frac{d \varphi_n}{d x} \right|_{x=d_n} = \frac{q^2}{2\varepsilon \varepsilon_0} N_d \left(\frac{d \varphi_p}{d x} \right)_{x=-d_p}^2 \quad (0 \leq x < d_n); \quad \varphi_n(x)|_{x=d_n} = 0, \quad \varphi_p(x)|_{x=-d_p} = \varphi_0; \quad (8.10)$$

$$\left. \frac{d \varphi_p}{d x} \right|_{x=-d_p} = \frac{q^2}{2\varepsilon \varepsilon_0} N_a \left(\frac{d \varphi_n}{d x} \right)_{x=0}^2 \quad (-d_p < x < 0); \quad (8.11)$$

Решение уравнений (8.8) и (8.9) с граничными условиями (8.10) и (8.11) дает следующие результаты:

$$(8.12)$$

$$\varphi_p = \varphi_0 = \frac{q^2}{2\varepsilon \varepsilon_0} N_a (d_p + x)^2 \quad (-d_p < x < 0); \quad (8.13)$$

$$d = d_n + d_p = \left\{ \begin{aligned} & \left(\frac{2\varepsilon \varepsilon_0 \varphi_0}{q^2} \frac{N_d + N_a}{N_d N_a} \right)^{1/2} = \left(\frac{2\varepsilon \varepsilon_0 \varphi_0}{q^2} \frac{n_n + p_p}{n_n p_p} \right)^{1/2} \\ & d_n / d_p = N_a / N_d. \end{aligned} \right\}; \quad (8.14)$$

Из выражений (8.12) и (8.13) видно, что $\varphi(x)$ является квадратичной функцией координаты x . Толщина слоя объемного заряда d (см. выражение (8.14)) обратно пропорциональна концентрации основных носителей заряда, равной концентрации примеси. Глубина проникновения

контактного поля больше в той области полупроводника, которая легирована меньше. Для несимметричного $p-n$ -перехода при $N_a \ll N_d$ соотношение (8.14) примет вид

$$(8.15)$$

Например, для германия при $N_a = 10^{21} \text{ м}^{-3}$, $N_d = 10^{23} \text{ м}^{-3}$ имеем $d_n = 8 \cdot 10^{-9} \text{ м}$, $d_p = 8 \cdot 10^{-7} \text{ м}$, т.е. практически весь $p-n$ -переход сосредоточен в более высокоомной области (той, где меньше концентрация примеси).

Приведенные формулы и расчеты относятся к резкому $p-n$ -переходу, в котором концентрация примеси на границе раздела ($x = 0$) меняется скачкообразно (см. рис. 8.1,а). Для плавных переходов (см. рис. 8.1,б), в которых концентрация примеси изменяется по линейному закону, получаем из решения уравнения Пуассона следующее выражение для толщины слоя объемного заряда:

$$d = \left[12 \epsilon \epsilon_0 \phi_0 / (q^2 a) \right]^{1/3}, \quad (8.16)$$

где a — коэффициент пропорциональности в выражении линейного закона изменения концентрации некомпенсированной примеси $N_a - N_d$ в переходе: $N_a - N_d = a(x_0 - x)$ (см. рис. 8.1,б).

§ 8.2. Вольт-амперная характеристика $p-n$ -перехода

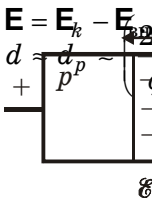
Посмотрим, как ведет себя электронно-дырочный переход при подаче внешнего напряжения. Подсоединим к p -области плюс, а к n -области минус источника напряжения. Напряжение такой полярности считается положительным. Поскольку $p-n$ -переход обеднен подвижными носителями заряда, его сопротивление значительно превышает сопротивление объемов p - и n - типов проводимости. Поэтому практически все приложенное напряжение V падает на $p-n$ -переходе. При указанной полярности напряжения направления напряженности контактного и внешнего электрического полей противоположны (рис. 8.4). Результирующая напряженность поля в переходе \mathbf{E} оказывается меньше, чем в равновесном состоянии:

$$(8.17)$$

Уменьшение напряженности электрического поля \mathbf{E} вызывает уменьшение высоты потенциального барьера $p-n$ -перехода ϕ :

$$\phi = \phi_0 - qV. \quad (8.18)$$

Рис. 8.4. $P-n$ -переход при положительном смещении



Так как основные носители заряда, диффундируя через p - n -переход, должны преодолевать потенциальный барьер, то уменьшение высоты барьера обусловит увеличение потока основных носителей заряда, т.е.

диффузионный ток через p - n -переход возрастет в $e^{\frac{qV}{k_0T}}$ раз. Условия для перемещения неосновных носителей заряда практически не изменяются, поэтому дрейфовый ток остается неизменным. Если равновесие токов нарушится, через переход потечет ток, обусловленный превышением диффузионной составляющей над дрейфовой:

$$j = j_{\text{диф}} - j_{\text{др}} \quad (8.19)$$

Изменим полярность приложенного напряжения. Подадим на p -область минус, а на n -область плюс от источника напряжения. Смещение такой полярности считается отрицательным. При этом направления напряженности контактного и внешнего электрического полей совпадают, напряженность поля в переходе \mathbf{E} и высота потенциального барьера ϕ увеличиваются:

$$(8.20)$$

$$\phi = \phi_0 + qV \quad (8.21)$$

Вследствие увеличения высоты потенциального барьера уменьшится

$\mathbf{E} = \mathbf{E}_k + \mathbf{E}_{\text{внеш}}$ в $e^{\frac{qV}{k_0T}}$ раз число основных носителей заряда, способных преодолеть его, уменьшится диффузионный ток через p - n -переход. Условия для перемещения неосновных носителей заряда остаются практически неизменными, дрейфовая составляющая тока не изменяется. Так как диффузионный ток уменьшается и обращается в нуль при увеличении напряжения такой полярности, ток через переход изменит направление (по сравнению с направлением тока при положительном напряжении) и будет обусловлен дрейфовой составляющей, т.е. неосновными носителями заряда:

$$j = j_{\text{диф}} - j_{\text{др}} \approx -j_{\text{др}} \quad (8.22)$$

Выведем соотношение, определяющее зависимость тока от приложенного напряжения, т.е. вольт-амперную характеристику.

В равновесном состоянии $j_{\text{диф}} = j_{\text{др}}$ и

$$j = j_{\text{диф}} - j_{\text{др}} = 0.$$

При подаче напряжения изменяется диффузионная составляющая тока. Дрейфовая составляющая остается неизменной. Обозначим неизменяющуюся с напряжением составляющую тока через j_s и назовем ее током насыщения.

Тогда в равновесном состоянии

$$j_{\text{диф}} = j_{\text{др}} = j_s \quad (8.23)$$

Диффузионная составляющая тока зависит от высоты потенциального барьера по закону

$$j_{\text{диф}} = j_s e^{\pm \frac{qV}{k_0T}}$$

Дрейфовый ток, т.е. ток неосновных носителей, от приложенного напряжения не зависит и равен $j_{\text{др}} = j_s$.

Результирующий ток равен

$$j = j_{\text{диф}} - j_{\text{др}} = j_s \left(e^{\pm \frac{qV}{k_0T}} - 1 \right), \quad (8.24)$$

или можем записать выражение ВАХ в виде

$$(8.25)$$

имея в виду, что в (8.25) нужно подставлять напряжение со своим знаком. Как следует из (8.25), при $V > 0$ ток через переход экспоненциально

возрастает, при $V > (3-4)k_0T/q$. При 20°C $k_0T/q = 0,025$ В, сле-

довательно, при $V > 0,1$ В.

При отрицательном напряжении такой же величины ($V = -0,1$ В) экспоненциальным членом в (8.25) можно пренебречь по сравнению с единицей, и $j = -j_s$.

Изобразим графически вольт-амперную характеристику (рис. 8.5). Из рисунка видно, что зависимость тока от напряжения нелинейна и несимметрична. Следовательно, на основе p - n -перехода можно создать прибор, пригодный для выпрямления тока. Таким прибором является полупроводниковый диод. Он содержит p - n -переход, снабженный омическими (антизапирающими) контактами металл-полупроводник с линейной вольт-амперной характеристикой для подачи внешнего напряжения (рис. 8.6). Положительное напряжение, при возрастании которого ток резко (экспоненциально) увеличивается, называется прямым, или пропускным, ток при этом — прямой, или пропускной. Отрицательное напряжение, при котором ток меняет направление и остается практически постоянным, называется обратным, или запирающим; ток при этом напряжении — обратный. При построении ВАХ масштабы прямых и обратных токов и напряжений различны. Прямой ток может превышать обратный в тысячу и более раз. Отношение этих токов при одном и том же напряжении разной полярности называется коэффициентом выпрямления. Прямое

$$j = j_s \left(e^{\frac{qV}{k_0T}} - 1 \right)$$

напряжение, подаваемое на диод, не превышает 1,5–2 В; обратное напряжение может составлять сотни вольт.

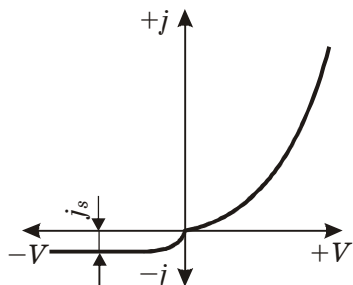


Рис. 8.5. Вольт-амперная характеристика p - n -перехода

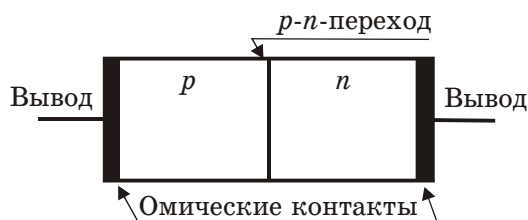


Рис. 8.6. Полупроводниковый диод

Следует иметь в виду, что при выводе выражения для вольт-амперной характеристики не учитывались генерация и рекомбинация носителей в p - n -переходе, явления на поверхности кристалла.

§ 8.3. Определение тока насыщения

Определим ток насыщения p - n -перехода, который создается неосновными носителями заряда. Для вывода формулы обратимся к плоскостной модели p - n -перехода (рис. 8.7).

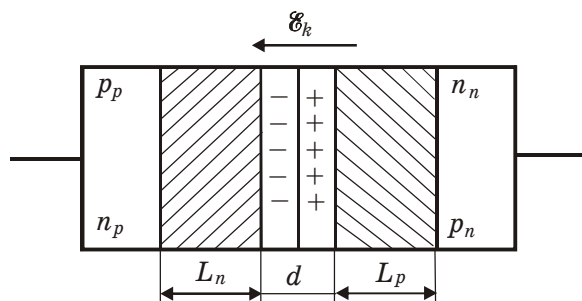


Рис. 8.7. К определению тока насыщения

Как было показано ранее (см. § 8.2), обратный ток создается неосновными носителями заряда — электронами p -области и дырками n -области. Очевидно, что в образовании тока принимают участие только те неосновные носители заряда, которые за время жизни успевают достичь границы действия электрического поля, т.е. дойти до p - n -перехода. Таким образом, величина обратного тока p - n -перехода будет определяться диффузионными потоками неосновных носителей из объема полупроводника на границу слоя объемного заряда: $j_s = j_{sn_p} + j_{sp_n} = qn_p v_{n \text{ диф}} + qp_n v_{p \text{ диф}}$.

Поскольку $v_{n \text{ диф}} = L_n / \tau_n$, $v_{p \text{ диф}} = L_p / \tau_p$, то

$$j_s = qn_p \frac{L_n}{\tau_n} + qp_n \frac{L_p}{\tau_p}. \quad (8.26)$$

Выражение (8.26) можно трактовать следующим образом. Ток насыщения создается неосновными носителями заряда, возникающими в объемах, прилегающих к p - n -переходу, со скоростью $1/\tau$. Полный ток насыщения равен

$$I_s = j_s S, \quad (8.27)$$

где S — площадь p - n -перехода.

Формулу (8.26) можно преобразовать. Известно, что $p_n = n_i^2 / n_n$, $L = \sqrt{D\tau}$. Подставив эти значения в (8.26), получим

$$(8.28)$$

Так как $n_i^2 \propto e^{-\frac{\Delta E}{k_0 T}}$, то с ростом температуры ток насыщения будет увеличиваться экспоненциально. Очевидно, что при увеличении ширины запрещенной зоны ток насыщения уменьшается экспоненциально. Именно поэтому германиевые диоды имеют значительно большие токи насыщения, чем кремниевые. Величина этого тока зависит от удельного сопротивления материала:

$$\rho_n = \frac{1}{qn_n \mu_n}; \quad \rho_p = \frac{1}{qp_p \mu_p}.$$

С учетом этих выражений формулу (8.28) можем записать следующим образом:

$$(8.29)$$

С ростом удельного сопротивления полупроводника ток насыщения увеличивается. Физический смысл этой зависимости заключается в том, что при увеличении удельного сопротивления уменьшается концентрация основных носителей заряда и увеличивается концентрация неосновных носителей, создающих обратный ток. Чем меньше ток насыщения, тем лучшими выпрямляющими свойствами обладает p - n -переход.

§ 8.4. Пробой p - n -перехода

Зависимость тока от напряжения (8.25) оказывается справедливой при незначительных обратных напряжениях. При увеличении запирающего напряжения на p - n -переходе выше некоторого критического значения

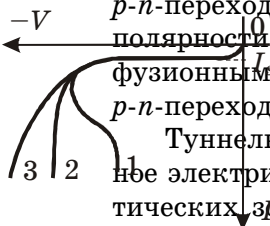
$$\frac{j_{sp}}{j_s} = \frac{n_i^2}{n_n^2} \left(\frac{p_1}{p} \right)^2$$

происходит резкое увеличение обратного тока — пробой p - n -перехода (рис. 8.8). Различают три основных вида пробоя: 1) тепловой; 2) лавинный; 3) туннельный. Первый вид пробоя, как правило, необратим, тогда как второй и третий обратимы и используются в полупроводниковых стабилизаторах напряжения — стабилитронах.

Тепловой пробой получил такое название из-за того, что рост обратного тока связан с неограниченным ростом температуры. Последнее обусловлено тем, что выделяющаяся мощность в p - n -переходе больше, чем отводимая от него. Выделяющаяся мощность равна сумме мощностей, передаваемых от электрического поля кристаллической решетке при протекании дрейфового тока, т.е. тока неосновных носителей заряда, при прямом и обратном включении. За счет этого повышается температура p - n -перехода. При тепловом пробое мощность, выделяющаяся при обратном включении, как правило, больше мощности, выделяющейся при прямом включении, поскольку обратное смещение, подаваемое на p - n -переход, может быть много больше, чем прямое. Прямое напряжение всегда меньше контактной разности потенциалов p - n -перехода V_k , а величина дрейфового тока не зависит от изменения полярности приложенного к p - n -переходу напряжения и определяется диффузионным потоком неосновных носителей в область объемного заряда p - n -перехода.

Туннельный и лавинный механизмы обусловлены следующим. Сильное электрическое поле в p - n -переходе вызывает большой наклон энергетических зон. В результате начинается туннельный переход электронов из валентной зоны в зону проводимости, концентрация свободных носителей возрастает, что приводит к резкому увеличению обратного тока. Это так называемый Зенеровский пробой. Другой механизм пробоя связан с тем, что свободные носители заряда могут создаваться в результате разрыва валентных связей за счет кинетической энергии носителя заряда, полученной им от электрического поля при прохождении запирающего слоя (p - n -перехода). Возникающие свободные носители заряда в свою очередь тоже ускоряются полем и могут принять участие в генерации электронно-дырочных пар. Происходит лавинообразное увеличение концентрации свободных носителей, обратный ток увеличивается. Такой вид пробоя называется лавинным.

Рис. 8.8. Обратная ветвь ВАХ p - n -перехода при различных механизмах пробоя: 1 — тепловом; 2 — лавинном; 3 — туннельном



§ 8.5. Энергетические диаграммы p - n -перехода при наличии внешнего электрического поля

Физические явления, описанные в § 8.2, можно представить на энергетических диаграммах. Построение зонной диаграммы p - n -перехода в равновесном состоянии основывалось на единстве уровня Ферми в n - и p -областях, так как устанавливалось равенство потоков носителей заряда, перемещающихся из одной области в другую. Наличие внешнего напряжения нарушает равновесие токов. Высота потенциального барьера изменяется, уровни Ферми в n - и p -областях устанавливаются на разной высоте. Разрыв в положении этих уровней происходит в p - n -переходе, величина его равна qV . Пусть на p - n -переход подано прямое напряжение. При этом высота потенциального барьера понижается (см. уравнение (8.18)). Так как внешнее приложенное напряжение падает на p - n -переходе, то границы энергетических зон в n - и p -областях не будут претерпевать изменений. Если мы «закрепим» зонную диаграмму полупроводника p -типа (см. рис. 8.3), то, «разорвав» уровень Ферми в области объемного заряда и сместив вверх зонную диаграмму полупроводника n -типа на qV , получим энергетическую диаграмму p - n -перехода при прямом напряжении. Следует учесть при построении диаграммы, что ширина области объемного заряда d будет изменяться по сравнению с равновесным состоянием, так как в выражении (8.14) вместо ϕ_0 следует подставлять $\phi = \phi_0 - qV$. В результате получим диаграмму, изображенную на рис. 8.9,а.

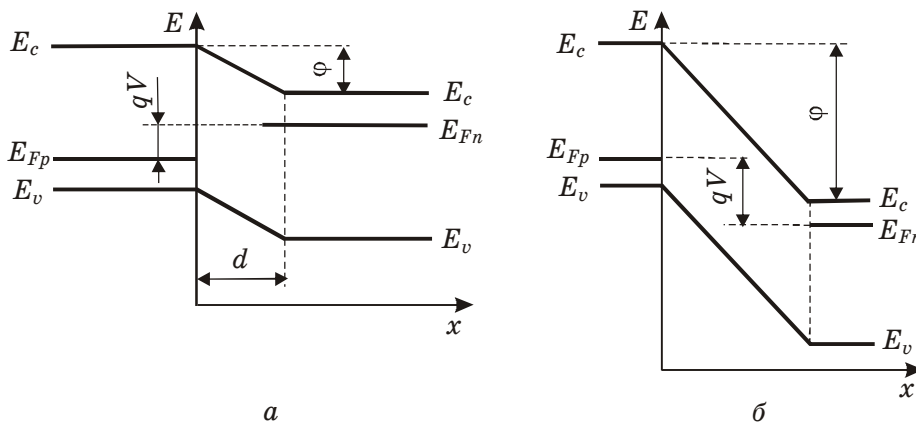


Рис. 8.9. Энергетические диаграммы p - n -перехода
при прямом (а) и обратном (б) напряжениях

При обратном напряжении энергетическая диаграмма n -области понижается на величину qV относительно энергетической диаграммы p -области (рис. 8.9,б). По диаграммам можно объяснить изменение потоков основных носителей при подаче напряжения разной полярности.

§ 8.6. Емкость $p-n$ -перехода

При изменении напряжения, приложенного к $p-n$ -переходу, изменяются ширина области объемного заряда, величина объемных зарядов, следовательно, $p-n$ -переход ведет себя подобно емкости, которую называют зарядной или барьерной емкостью C_n $p-n$ -перехода. Кроме того, при протекании прямого тока наблюдается введение неосновных неравновесных носителей заряда — электронов в p -область и дырок в n -область, называемое инжекцией. При инжекции заряд «впрыскиваемых» носителей изменяется при изменении напряжения. Значит, в диоде возникает еще одна составляющая емкости — диффузионная $C_{\text{диф}}$, обусловленная изменением объемного заряда инжектированных носителей при изменении напряжения. Полная емкость C определяется суммой

$$C = C_n + C_{\text{диф}} . \quad (8.30)$$

Определим каждую составляющую емкости. При прямом смещении высота потенциального барьера понижается (см. § 8.2), увеличивается поток основных носителей заряда через $p-n$ -переход, часть ионов доноров и ионов акцепторов оказывается скомпенсированной, что приводит к уменьшению и ширины, и объемного заряда $p-n$ -перехода (рис. 8.10). При подаче обратного смещения увеличивается потенциальный барьер, вследствие чего возрастают ширина и объемный заряд в $p-n$ -переходе (см. выражения (8.14), (8.21)). На рис. 8.10 сплошной линией показаны границы $p-n$ -перехода в равновесии, штриховыми — при подаче напряжения.

Так как область $p-n$ -перехода обеднена подвижными носителями, ее можно представить как диэлектрик, разделяющий обкладки, на которых сосредоточен отрицательный заряд в p -области и положительный — в n -области. Расстояние между обкладками d зависит от величины и полярности напряжения, приложенного к $p-n$ -переходу. Тогда емкость перехода C_n можно определить как $C_n = dQ / dV$ или по формуле плоского конденсатора

$$(8.31)$$

$$Q_n = qN_d S_{p-n} d_n \quad \text{или} \quad Q_p = qN_a S_{p-n} d_p . \quad (8.32)$$

Продифференцировав эти выражения по напряжению, получим значение емкости перехода. Или в (8.31) подставим значение d (выражение (8.14)), заменив ϕ_0 на ϕ :

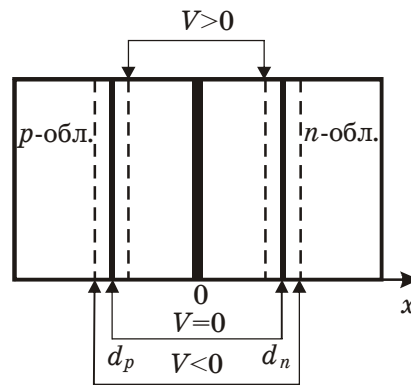


Рис. 8.10. Изменение ширины области объемного заряда $p-n$ -перехода при прямом и обратном напряжении

(8.33)

Для несимметричного p - n -перехода $N_a \ll N_d$, тогда

$$C_n = S_{p-n} \left(\frac{q^2 \varepsilon \varepsilon_0 p_p}{2\varphi} \right)^{1/2}. \quad (8.34)$$

Для плавного p - n -перехода

$$C_n = S_{p-n} \left[\frac{q^2 a (\varepsilon \varepsilon_0)^2 p_p}{12\varphi} \right]^{1/3}. \quad (8.35)$$

Из приведенных выражений следует, что емкость резкого перехода возрастает при увеличении концентрации примеси и пропорциональна корню квадратному из напряжения. Для плавного перехода емкость зависит от градиента концентрации примеси a и пропорциональна $V^{1/3}$. Следовательно, по зависимости емкости C_n от приложенного напряжения можно определить характер распределения примеси.

Диффузионная емкость может быть определена по изменению заряда инжектированных носителей, т.е.

$$C_{\text{диф}} = dQ / dV.$$

Рассмотрим несимметричный p - n -переход. Пусть $N_a \ll N_d$. В этом случае более высокоомная (имеющая большее сопротивление) p -область будет называться базой диода. При подаче положительного напряжения прямой ток переносится в основном электронами из n -области в p -область (так как $n_n \gg p_p$). Встречной инжекцией дырок из p -области в n -область в силу принятого условия можно пренебречь. Пришедшие в p -область электроны, являясь неосновными неравновесными носителями, нарушают электрическую нейтральность объема базы. Для восстановления электрической нейтральности из внешней цепи вводится такое же количество основных неравновесных носителей. При изменении приложенного напряжения изменяется заряд инжектированных и введенных из внешней цепи носителей, т.е., действительно, база диода ведет себя подобно емкости. Так как в основе накопления заряда лежит диффузия носителей в базу, то эту емкость и называют диффузионной. Для определения $C_{\text{диф}}$ найдем заряд носителей, инжектированных в базу (рис. 8.11):

$$Q_n = q S_{p-n} \int_0^{\infty} n'(x) dx, \quad (8.36)$$

$$n'(x) = n'(0) \exp\left(-\frac{x}{L_n}\right), \quad (8.37)$$

$$C_n = \frac{\varepsilon \varepsilon_0 S_p}{(2\varepsilon \varepsilon_0 \varphi)}$$

где $n'(0)$ — концентрация неравновесных электронов, «впрыснутых» в p -область при $x = 0$,

Таким образом,

$$Q_n = qn'(0)L_nS_{p-n} = qn_p e^{\frac{qV}{k_0T}} L_n S_{p-n}. \quad (8.38)$$

Продифференцировав это выражение по напряжению и преобразовав, получим

$$C_{\text{диф}} = \frac{q}{k_0T} (I + I_s) \tau_n. \quad (8.39)$$

Как следует из (8.39), при прямом напряжении диффузионная емкость возрастает пропорционально протекающему току, при обратном — уменьшается, обращаясь в нуль при $I = -I_s$.

Анализ выражений для определения емкости диода показывает, что изменяя напряжение, подаваемое на p - n -переход, можно управлять его емкостью. Приборы, основанные на использовании этого свойства p - n -перехода, называются варика-

Рис. 8.11. К определению диффузионной емкости



§ 8.7. Структура биполярного транзистора

Транзистором называется полупроводниковый прибор с электронно-дырочными переходами, предназначенный для усиления или генерирования электрических сигналов и имеющий три или более выводов. Цепь одного из этих выводов является входной, или управляющей, цепь другого — выходной, или управляемой. Кроме усиления и генерирования, транзистор может выполнять и другие функции, которые не являются для него основными и характерными.

В зависимости от способа управления выходным сигналом транзисторы делятся на транзисторы с инжекцией, у которых управление выходным током осуществляется за счет инжекции неосновных неравновесных носителей, и транзисторы без инжекции, у которых управление выходным сигналом происходит под действием электрического поля, поэтому их называют полевыми. Принцип действия полевых транзисторов будет рассмотрен в следующей главе.

Транзисторы с инжекцией называются биполярными, так как их нормальная работа обеспечивается носителями заряда двух знаков — электронами и дырками, основными и неосновными носителями.

Биполярные транзисторы — это приборы с двумя электронно-дырочными переходами и тремя выводами. С этой точки зрения их часто называют плоскостными триодами. Транзистор образован тремя областями полупроводника. Крайние области имеют один тип проводимости (n или p), средняя область — другой. На рис. 8.12 изображены плоскостные модели транзисторов $p-n-p$ - и $n-p-n$ -типов. Левая область называется эмиттером, средняя — базой, правая — коллектором. Электронно-дырочный переход между базой и эмиттером — эмиттерный, между базой и коллектором — коллекторный. Назначение эмиттера — инжектировать носители в базу, поэтому на эмиттерный $p-n$ -переход подается положительное напряжение. База должна с наименьшими потерями передавать инжектированные эмиттером носители коллекторному $p-n$ -переходу. Для этого ширину базы W делают значительно меньше диффузионной длины неосновных неравновесных носителей. Коллектор собирает дошедшие до него носители, в связи с этим на коллекторный переход подается отрицательное напряжение и площадь его превышает площадь эмиттерного перехода.

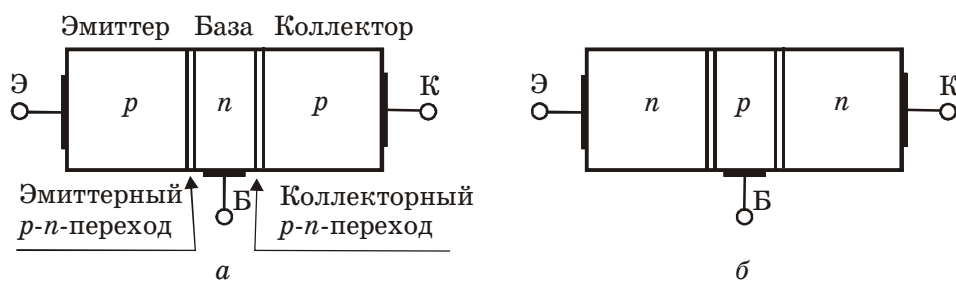


Рис. 8.12. Структура биполярных транзисторов типа $p-n-p$ (а) и $n-p-n$ (б)

В зависимости от механизма переноса носителей заряда через область базы различают бездрейфовые транзисторы, у которых инжектированные эмиттером носители перемещаются к коллектору под действием градиента концентрации, и дрейфовые — у них неравновесные носители двигаются от эмиттера к коллектору не только за счет диффузии, но и под действием внутреннего электрического поля, возникающего в базе вследствие неравномерного распределения примеси.

§ 8.8. Принцип работы биполярного транзистора

Рассмотрим транзистор $p-n-p$ -типа. Так как транзистор представляет собой трехполюсник, то при включении его в цепь один из выводов оказывается общим для входной и выходной цепей. В связи с этим различают три схемы включения транзистора: с общей базой, общим эмиттером и общим коллектором. С принципом действия транзистора познакомимся

на примере схемы включения с общей базой (рис. 8.13). При подаче положительного напряжения на эмиттерный p - n -переход его потенциальный барьер понижается, и дырки начинают инжектироваться из эмиттера в базу. Концентрация неравновесных дырок в базе около эмиттерного перехода $p'(0)$ определяется выражением

$$(8.40)$$

где p_n — равновесная концентрация неосновных носителей в базе.

При $V_{эб} > k_0T/q$ будет $p'(0) \gg p_n$. Как и для одиночного p - n -перехода, приложенное внешнее напряжение $V_{эб}$ падает на эмиттерном переходе, а $V_{кб}$ — на коллекторном. Поэтому электрическое поле в базе отсутствует. «Впрыснутые» эмиттером в базу дырки диффундируют через базу от эмиттера к коллектору. Как уже отмечалось ранее (см. § 8.7), ширина базы W значительно меньше диффузионной длины неосновных неравновесных носителей заряда L_p . Вследствие этого при диффузии дырок от эмиттера к коллектору через базу почти не происходит потери дырок на рекомбинацию с электронами.

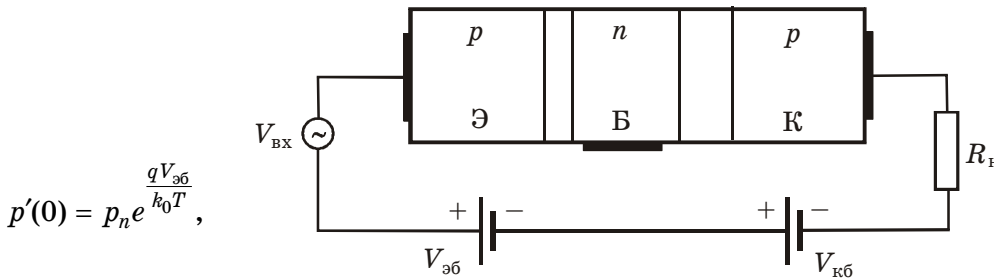


Рис. 8.13. Включение транзистора по схеме с общей базой

Коллекторный p - n -переход смещен в обратном направлении. Ток через него создается неосновными носителями заряда. Продиффундировавшие через базу дырки, подходя к области объемного заряда коллектора, будут втягиваться его полем и переноситься в коллектор. Ток в коллекторе по величине будет незначительно отличаться от тока в эмиттере.

Необходимо отметить, что концентрация неравновесных неосновных носителей в базе около коллектора определяется выражением, аналогичным (8.40):

$$p'(W) = p_n e^{\frac{qV_{кб}}{k_0T}}. \quad (8.41)$$

Так как $V_{кб} \gg k_0T/q$ и коллектор смещен в обратном направлении, т.е. $V_{кб} < 0$, то $p'(W) = 0$.

Физически это означает, что как только неравновесные неосновные носители оказываются на границе области объемного заряда коллекторного p - n -перехода, они подхватываются полем этого перехода и уносятся

в коллектор. Это положение можно пояснить с помощью следующей аналогии (рис. 8.14). Через щель в стенке ящика песок высыпается на площадку небольшой ширины и, достигая края этой площадки, падает вниз (рис. 8.14,а). Ширина щели в этом случае тождественна высоте потенциального барьера эмиттерного p - n -перехода, определяющей начальную концентрацию, ширина площадки — ширине базовой области W , а движение песка в поле тяготения за пределами площадки — движению зарядов в электрическом поле коллекторного p - n -перехода. Очевидно, что закон распределения песка на площадке шириной W будет зависеть и от ширины щели и от ширины площадки. При одинаковой ширине щели распределение песка на площадках разной ширины W существенно различается (рис. 8.14,а,б).

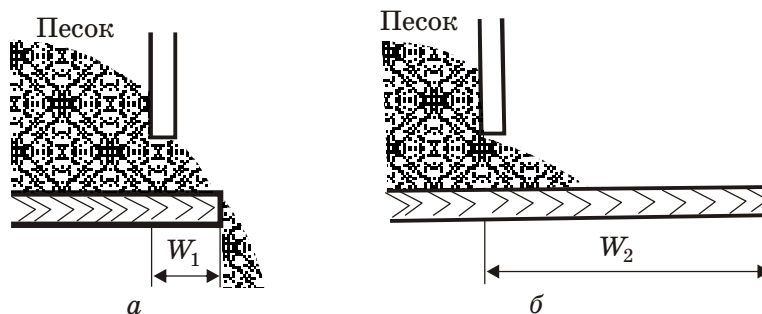


Рис. 8.14. Схематическое изображение распределения песка, проходящего через щель в ящике, аналогичное распределению неравновесных носителей в базе транзистора:

а — площадка малой ширины W_1 ; б — площадка большой ширины W_2

Для оценки активных свойств транзистора необходимо оценить коэффициент усиления по мощности K_p , который определяется отношением мощности, выделяемой на нагрузочном резисторе, $P_{\text{ВЫХ}}$, к мощности, затрачиваемой на управление входной цепью, $P_{\text{ВХ}}$.

Во входной цепи протекает эмиттерный ток $I_э$, а сопротивление определяется сопротивлением прямосмещенного перехода $R_э$. Поэтому $P_{\text{ВХ}} = R_э I_э^2$. В выходной цепи протекает ток $I_к$, и $P_{\text{ВЫХ}} = R_н I_к^2$. Следовательно,

$$K_p = \frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}}} = \frac{R_н I_к^2}{R_э I_э^2}. \quad (8.42)$$

Поскольку $I_к \approx I_э$, то $K_p \approx R_н / R_э$. Нагрузочный резистор включается последовательно обратносмещенному коллекторному переходу и выбирается достаточно большим: $R_э < R_н$, следовательно, $K_p > 1$, т.е., действительно, транзистор является активным элементом. Схема с общей базой обеспечивает усиление по мощности и по напряжению. По току усиление не происходит.

§ 8.9. Энергетические зонные диаграммы биполярного транзистора

Рассмотрим энергетическую диаграмму бездрейфового транзистора типа $p-n-p$ в равновесном состоянии. Эта диаграмма может быть составлена из аналогичных диаграмм эмиттерного и коллекторного $p-n$ -переходов. При построении нужно учесть, что в равновесном состоянии токи в транзисторе равны нулю, поэтому уровень Ферми является единым для всех трех областей — эмиттера, базы и коллектора (рис. 8.15). Выделим области объемных зарядов эмиттерного и коллекторного $p-n$ -переходов (на рисунке это сделано вертикальными штриховыми линиями). За пределами этих областей электрическое поле отсутствует. Так как полупроводники эмиттера, базы и коллектора однородны, то энергетические зазоры между уровнем Ферми и границами зон E_c , E_v не изменяются в этих областях.

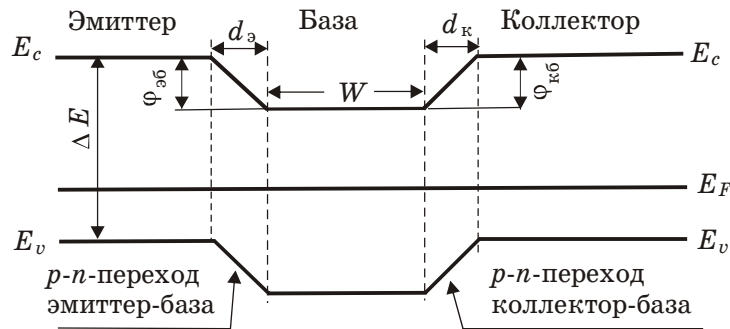


Рис. 8.15. Энергетическая диаграмма биполярного транзистора
при отсутствии внешнего смещения
на эмиттерном и коллекторном переходах

При построении учитываем, что в эмиттере и коллекторе уровень Ферми расположен ближе к потолку валентной зоны, а в базе — ближе к дну зоны проводимости. Так как концентрации акцепторов в эмиттере и коллекторе $N_{aэ}$, $N_{ак}$ обычно значительно превышают концентрацию доноров в базе $N_{dб}$, энергетические зазоры $E_F - E_v$ в эмиттере и коллекторе должны быть меньше, чем $E_c - E_F$ в базе. Уровни, соответствующие дну зоны проводимости в эмиттере и коллекторе, проводим на расстоянии ΔE от потолка валентной зоны и аналогично располагаем уровень E_v в области базы. Соединим границы соответствующих зон в областях объемных зарядов эмиттерного и коллекторного $p-n$ -переходов. Высота изгиба зон определяет равновесное значение потенциальных барьеров $\phi_{эб}$ и $\phi_{кб}$. При подаче положительного напряжения на эмиттерный переход уменьшается высота потенциального барьера $\phi_{эб}$ и ширина области объемного заряда эмиттерного перехода $d_э$.

На коллекторный переход подается отрицательное напряжение, поэтому $\phi_{кб}$ и ширина области объемного заряда d_k возрастают (рис. 8.16). При рассмотрении рис. 8.16 можно сделать вывод, что уменьшившийся потенциальный барьер $\phi_{эб}$ преодолевает большее число основных носителей заряда, поэтому ток эмиттера возрастает. Поскольку эмиттер легирован больше, чем база, этот ток будет в основном представлен дырочной составляющей $I_{pэ}$ (т.е. дырки будут диффундировать из эмиттера в базу). В базе электрическое поле отсутствует, так как внешние напряжения падают на $p-n$ -переходах. Поэтому инжектированные эмиттером в базу дырки перемещаются к коллектору под действием градиента концентрации.

Рис. 8.16. Энергетическая диаграмма биполярного транзистора при подаче положительного напряжения на эмиттерный $p-n$ -переход и отрицательного — на коллекторный

Потенциальный барьер в коллекторе $\phi_{кб}$, возросший за счет внешнего напряжения, не могут преодолеть основные носители. Ток переносится неосновными носителями заряда — дырками базы (равновесными и неравновесными) и электронами коллектора.

§ 8.10. Коэффициент усиления биполярного транзистора по току

Транзистор — это активный элемент электрической цепи, поэтому одним из основных его параметров является коэффициент усиления по току. Для схемы с общей базой он обозначается через α и определяется отношением изменения тока коллектора I_k к вызвавшему его изменению тока эмиттера при постоянном напряжении на коллекторном $p-n$ -переходе:

$$(8.43)$$

$$\alpha = \frac{\partial I_k}{\partial I_{э}} \bigg|_{V_{кб} = \text{const}}$$

E_c —
 E_F —
 E_v —

Анализируя физические явления в транзисторе, можно установить, что коэффициент α определяется результирующим действием нескольких факторов: способностью эмиттера инжектировать неосновные неравновесные носители в базу, способностью базы донести (или передать) эти носители до коллекторного перехода, способностью коллектора увеличить ток через переход. В свете сказанного можно записать

$$\alpha = \gamma\beta\alpha_i, \quad (8.44)$$

где γ — эффективность эмиттера; β — коэффициент переноса; α_i — эффективность коллектора.

Эффективность эмиттера (или коэффициент инжекции) γ характеризует способность эмиттера инжектировать неосновные неравновесные носители в базу. Для $p-n-p$ -транзистора γ определяется отношением дырочной составляющей тока эмиттера $I_{pэ}$ к полному току через эмиттерный $p-n$ -переход. Для p^+-n -перехода (p^+ — сильнолегированная p -область)

$$(8.45)$$

Подставляя в (8.45) выражения для $I_{pэ}$, $I_{nэ}$ и учитывая, что $p_n = \frac{n_i^2}{n_n}$,

$$\gamma = \frac{I_{pэ}}{I_{pэ} + I_{nэ}} = \frac{\mu_{pэ} \frac{n_i^2}{p_p} \frac{1}{L_p}}{\mu_{pэ} \frac{n_i^2}{p_p} \frac{1}{L_p} + \mu_{nэ} \mu_{pэ} \sigma_6 \frac{W}{L_3}} = \frac{1}{1 + \frac{I_{nэ}}{I_{pэ}}}. \quad I_{nэ} \ll I_{pэ}, \text{ получим} \quad \gamma = 1 - \frac{\mu_{nэ} \mu_{pэ} \sigma_6 W}{\mu_{nэ} \mu_{pэ} \sigma_3 L_3}. \quad (8.46)$$

Как следует из (8.46), для того чтобы приблизить эффективность эмиттера к единице, необходимо эмиттер легировать больше, чем базу ($\sigma_3 \gg \sigma_6$) и делать базу тонкой ($W \ll L_p$). Например, для германиевых сплавных $p-n-p$ -транзисторов типичными значениями являются $\sigma_3 = \sigma_p \approx 10^3 \text{ См} \cdot \text{см}^{-1}$; $\sigma_6 = \sigma_n \approx 1 \text{ См} \cdot \text{см}^{-1}$. При этом $\gamma = 0,9996$, и при инженерных расчетах часто принимают $\gamma = 1$. Необходимо учитывать, что эффективность эмиттера ухудшается с ростом эмиттерного тока. Коэффициент переноса β (его иногда называют коэффициентом передачи неосновных носителей заряда через базу транзистора) характеризует потери неосновных неравновесных носителей (дырок для $p-n-p$ -транзистора) на рекомбинацию при их движении от эмиттера к коллектору и является отношением дырочной составляющей коллекторного тока $I_{рк}$ к дырочной составляющей эмиттерного тока при постоянном напряжении на коллекторном $p-n$ -переходе, т.е.

$$\beta = \left. \frac{\partial I_{рк}}{\partial I_{пэ}} \right|_{V_{кб} = \text{const}}. \quad (8.47)$$

Вследствие того что часть дырок при движении через базу рекомбинирует с электронами, $\beta < 1$. При $W \ll L_p$

$$\beta \approx 1 - \frac{1}{2} \left(\frac{W}{L_p} \right)^2. \quad (8.48)$$

Выражение (8.48) позволяет определить потери неосновных носителей в результате их рекомбинации в объеме базовой области. Однако рекомбинация протекает и на поверхности базы, преимущественно вблизи эмиттера.

Общее выражение для коэффициента переноса β с учетом рекомбинации носителей в объеме и на поверхности можно записать в виде

$$(8.49)$$

где s — скорость поверхностной рекомбинации; A_s — поверхность рекомбинации, равная площади кольца шириной W , окружающего эмиттерный p - n -переход; A_p — площадь эмиттерного p - n -перехода; D_p — коэффициент диффузии дырок в базе.

Следует иметь в виду, что выражение (8.49) справедливо для низких частот и малых уровней сигналов. Из (8.49) видно, что для увеличения коэффициента переноса необходимо уменьшать ширину базы W и скорость поверхностной рекомбинации s . Типичные значения $\beta \approx 0,94-0,98$.

Эффективность коллектора

$$(8.50)$$

$$\beta \approx 1 - \frac{\partial I_{sA_s}}{\partial I_{nA_p}} \left(\frac{V_k}{V_{к проб}} \right)$$

Эффективность α_i определяется умножением носителей заряда вследствие ударной ионизации в коллекторном p - n -переходе и увеличением коллекторного тока за счет неосновных носителей, дрейфующих из коллектора в базу. С учетом этих физических явлений

$$\alpha_i = M \alpha^*, \quad (8.51)$$

где M — коэффициент лавинного умножения; α^* — собственный коэффициент усиления коллекторного p - n -перехода, $\alpha^* \approx 1$.

Коэффициент лавинного умножения равен:

$$(8.52)$$

где V_k — напряжение, прикладываемое к коллекторному переходу; $V_{к проб}$ — напряжение лавинного пробоя коллекторного p - n -перехода; n — показатель степени, зависящий от материала и типа перехода. Для

германиевых переходов n^+ - p -типа $n \approx 6$, а для переходов p^+ - n -типа $n \approx 3$. Для кремниевых p^+ - n -переходов $n \approx 3,5$, для n^+ - p -переходов $n \approx 2$. Анализируя выражение (8.52), можно сделать вывод, что при $V_k < V_{k \text{ проб}}$ коэффициент $M = 1$.

Собственный коэффициент усиления коллекторного перехода

$$(8.53)$$

Если дырочная составляющая коллекторного тока dI_{pk} вызывает изменение электронной составляющей dI_{nk} , то величина α^* может превышать единицу. Но обычно указанное влияние мало и значение $\alpha^* \approx 1$.

В свете сказанного можно сделать вывод, что транзистор работает при напряжениях на коллекторном переходе значительно ниже пробивного, то есть

$$\alpha_i \leq 1. \quad (8.54)$$

Следовательно, в схеме с общей базой усиление по току не происходит. Необходимо отметить, что коэффициент усиления по току зависит от режима работы транзистора. При изменении тока эмиттера изменяется градиент концентрации инжектированных в базу носителей. Напряжение на коллекторном p - n -переходе влияет на ширину области объемного заряда этого перехода, изменяя ширину базы W .

$$\alpha^* = 1 + \frac{d I_{nk}}{d I_{pk}}.$$

§ 8.11. Выходные вольт-амперные характеристики биполярного транзистора

Исходя из схемы включения транзистора и выбора зависимых и независимых переменных (токов и напряжений) на входе и выходе, можно получить для одного транзистора большое количество различных вольт-амперных характеристик. Рассмотрим только выходные характеристики транзистора, включенного по схеме с общей базой. Эти характеристики представляют зависимость выходного тока (в цепи коллектор-база) от выходного напряжения (на зажимах коллектор-база) при различных значениях входного тока (в цепи эмиттер-база). Они наиболее интересны, так как отражают основные процессы управления коллекторным током с помощью тока эмиттера.

Для одиночного коллекторного p - n -перехода зависимость тока от напряжения имеет вид

$$I_k = I_{k0} \left(e^{\frac{qV_{кб}}{k_0T}} - 1 \right), \quad (8.55)$$

где I_{k0} — обратный ток перехода.

Влияние эмиттерного перехода на ток коллектора отразится введением в формулу (8.55) слагаемого αI_3 . Тогда полный ток

$$(8.56)$$

Общий характер зависимости $I_K(V_{кб})$ аналогичен обратной ветви вольт-амперной характеристики диода, так как коллекторный переход включен в обратном направлении.

При построении графика для удобства расположим характеристики в первом и втором квадранте (рис. 8.17).

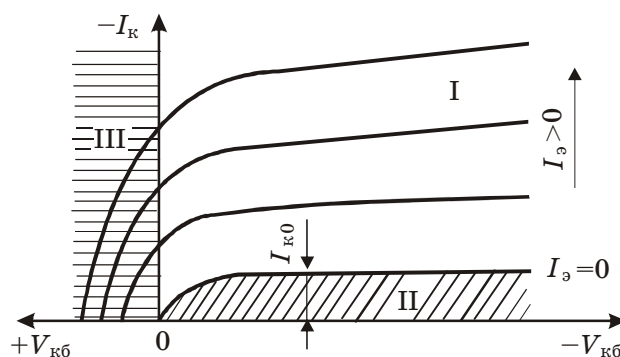


Рис. 8.17. Выходные характеристики транзистора, включенного по схеме с общей базой

Первую кривую получим, полагая $I_3 = 0$. При этом если $V_{кб} < 0$, то $I_K = I_{к0}$. При увеличении тока эмиттера выходные статические характеристики смещаются вверх в выбранной системе координат, что соответствует принципу действия транзистора. Особенностью рассматриваемых зависимостей является то, что через коллекторный $p-n$ -переход протекает ток $I = \alpha I_3$ при напряжении на коллекторе, равном нулю, т.е. при коротком замыкании коллектора с базой. Это объясняется наличием градиента концентрации неосновных носителей заряда в базе транзистора при инжекции из эмиттера. Чтобы ток коллектора стал равен нулю, на коллектор необходимо подать положительное напряжение. При этом $\phi_{кб}$ понизится, увеличится диффузионная составляющая тока, которая и уравновесит возросшую дрейфовую составляющую:

$$I_K = \alpha I_3 - I_{к0} e^{\frac{qV_{кб}}{k_0T}} = 0.$$

Если на коллектор подать достаточно большое запирающее напряжение, наклон характеристик будет быстро увеличиваться вследствие лавинообразного нарастания концентрации неосновных носителей заряда, т.е. произойдет пробой $p-n$ -перехода коллектора.

$$\begin{aligned} I_K(V_{кб})|_{I_3=0} \\ I_K = \alpha I_3 - I_{к0} e^{\frac{qV_{кб}}{k_0T}} \end{aligned}$$

Анализируя вид выходных вольт-амперных характеристик, отметим, что транзистор может работать при различных сочетаниях напряжений на эмиттерном и коллекторном переходах. В области I (см. рис. 8.17) коллекторный переход смещен в обратном направлении ($V_{кб} < 0$), эмиттерный — в прямом ($I_э > 0$). Именно этот режим работы, называемый активным, анализировался выше. При работе в активном режиме током коллектора можно управлять с помощью тока эмиттера, он обеспечивает усилительные свойства транзистора.

Во области II (заштрихована под углом на рис. 8.17) оба перехода смещены в обратном направлении, через них протекают обратные токи $I_{э0}$, $I_{к0}$. Следовательно, при работе в этом режиме, называемом режимом отсечки, отсутствует инжекция неравновесных носителей, ток эмиттера не управляет током коллектора.

В области III (заштрихована горизонтально на рис. 8.17) оба перехода смещены в прямом направлении и в рассматриваемом случае инжектируют неосновные носители в базу. Этот режим называется режимом насыщения. Ток коллектора определяется выражением (8.56).

Увеличение тока эмиттера не приводит к увеличению коллекторного тока, возрастает концентрация неравновесных носителей, накапливаемых в базе транзистора, увеличиваются его инерционные свойства. Следовательно, при работе в усилительных схемах нужно, чтобы транзистор работал в активном режиме. Режимы отсечки и насыщения реализуются при работе транзистора в импульсных схемах.

Контрольные вопросы и задачи

1. (ЮСЭД). Что такое электронно-дырочный переход?

Ответы:

- 1) переход электронов валентной зоны в зону проводимости;
- 2) переходный слой, разделяющий две области полупроводника с различными типами проводимости;
- 3) энергетический зазор, разделяющий две области разрешенных состояний энергий;
- 4) переход электронов или дырок с примесных уровней в соответствующую разрешенную зону.

2. (ЦЯРП). Что такое резкий p - n -переход?

Ответы:

- 1) переход, в котором концентрации дырок и электронов изменяются резко (ступенчато);
- 2) переход, в котором градиент концентрации примесей стремится к бесконечности;
- 3) переход, в котором высота потенциального барьера много больше k_0T ;

4) переход, в котором концентрация основных подвижных носителей заряда уменьшается в _____ раз;

5) переход, у которого обратный ток при пробивном напряжении резко увеличивается.

3. (НЭПЛ). Что такое плавный p - n -переход?

Ответы:

1) переход, у которого обратный ток при пробивном напряжении меняется плавно;

2) переход, у которого ширина области объемного заряда зависит от напряжения пропорционально выражению $(\varphi_k - qV)^{1/3}$;

3) переход, у которого прямой ток плавно зависит от напряжения при $qV > \varphi_k$;

4) переход, в котором концентрации подвижных носителей заряда изменяются плавно;

5) переход, который изготавливается методом диффузии.

4. (ФЭСП). Что такое сплавной p - n -переход?

Ответы:

1) переход, получаемый сплавлением донорного и акцепторного полупроводников;

2) переход, получаемый вплавлением в полупроводник сплава, содержащего примесь противоположного типа, чем в полупроводнике;

3) переход, получаемый эпитаксией из расплава;

4) переход, к которому припаян омический контакт;

5) переход, в котором концентрация доноров в полупроводнике n -типа много больше концентрации акцепторов в полупроводнике p -типа.

5. (ФЕДП). Что такое диффузионный p - n -переход?

Ответы:

1) переход, в котором при прямом включении ток определяется диффузионной составляющей;

2) переход, в котором обратный ток определяется диффузионным потоком неосновных носителей к p - n -переходу;

3) переход, в котором распределение примесей плавное;

4) переход, изготавливаемый диффузией примесей;

5) переход, для которого справедлива диффузионная теория выпрямления.

6. (ОЦДК). Что такое диффузионная составляющая тока в p - n -переходе?

Ответы:

1) составляющая, обусловленная переходом основных носителей через потенциальный барьер;

2) составляющая, обусловленная диффузией неосновных носителей в области объемного заряда p - n -перехода;

3) составляющая, обусловленная диффузией неосновных носителей в базе диода при прямом включении;

$$\exp[\varphi_k / (k_0 T)]$$

4) составляющая, обусловленная градиентом концентрации подвижных носителей заряда;

5) составляющая, определяемая выражением

7. (АУДР). Что такое дрейфовая составляющая тока в p - n -переходе?

Ответы:

1) составляющая, определяемая выражением

2) составляющая тока, определяемая градиентом концентрации неосновных носителей заряда;

3) составляющая тока, определяемая дрейфом основных носителей заряда при обратном включении p - n -перехода;

4) ток, протекающий при обратном напряжении более 0,1 В;

5) составляющая тока, уменьшающаяся до нуля при прямом включении.

8. (ТЮПТ). Как зависит диффузионная составляющая тока p - n -перехода от напряжения?

Ответы:

1) уменьшается;

2) увеличивается;

3) растет по экспоненте;

$$j_{\text{диф}} = \pm q \left(\frac{D_p n_p}{L_p} + \frac{D_n p_p}{L_n} \right) e^{\pm \frac{qV}{k_0 T}};$$

5) прямо пропорциональна высоте потенциального барьера.

9. (ЛШДК). Как зависит дрейфовая составляющая тока p - n -перехода от напряжения?

Ответы:

1) пропорциональна напряженности электрического поля в p - n -переходе;

2) не зависит;

3) пропорциональна $\exp[qV/(k_0 T)]$;

4) пропорциональна напряжению;

5) пропорциональна высоте потенциального барьера.

10. (ОМКР). Чем определяется контактная разность потенциалов в резком p - n -переходе?

Ответы:

1) разностью внешних работ выхода электронов полупроводников n - и p -типа;

2) концентрацией примесей;

3) выражением

4) выражением $\varphi_k = \frac{1}{q} \left(\Delta E - k_0 T \ln \frac{N_c N_v}{p_p n_n} \right);$

5) шириной запрещенной зоны, деленной на заряд электрона.

11. (Ш2ПТ). Как зависит прямой ток *p-n*-перехода от концентраций основных носителей заряда в *p*- и *n*-областях?

Ответы:

- 1) пропорционально;
- 2) обратно пропорционально;
- 3) не зависит, а определяется приложенным напряжением;
- 4) зависит не от концентрации основных носителей заряда, а от их градиента;
- 5) зависит только от температуры *p-n*-перехода.

12. (ЯЦЭП). Какое направление у контактного электрического поля в *p-n*-переходе?

Ответы:

- 1) от *p*-типа к *n*-типу;
 - 2) от *n*-типа к *p*-типу;
 - 3) определяется приложенным напряжением;
 - 4) от *p*-типа к *n*-типу только при положительном напряжении;
 - 5) от *n*-типа к *p*-типу только при отрицательном напряжении.
- 13. (НМОЗ).** Каково соотношение между зарядами в области объемного заряда *n*- и *p*-типа в *p-n*-переходе?

Ответы:

- 1) в высоколегированной *p*-области зарядов больше, чем в *n*-области;
- 2) в высоколегированной *p*-области зарядов меньше, чем в *n*-области;
- 3) заряды равны;
- 4) при положительном напряжении почти весь заряд накапливается в слаболегированной области;
- 5) при отрицательном напряжении почти весь заряд накапливается в области с меньшей концентрацией примеси.

14. (ЗЯРП). Почему отличаются количественные соотношения, описывающие распределение потенциала в резком и плавном *p-n*-переходах?

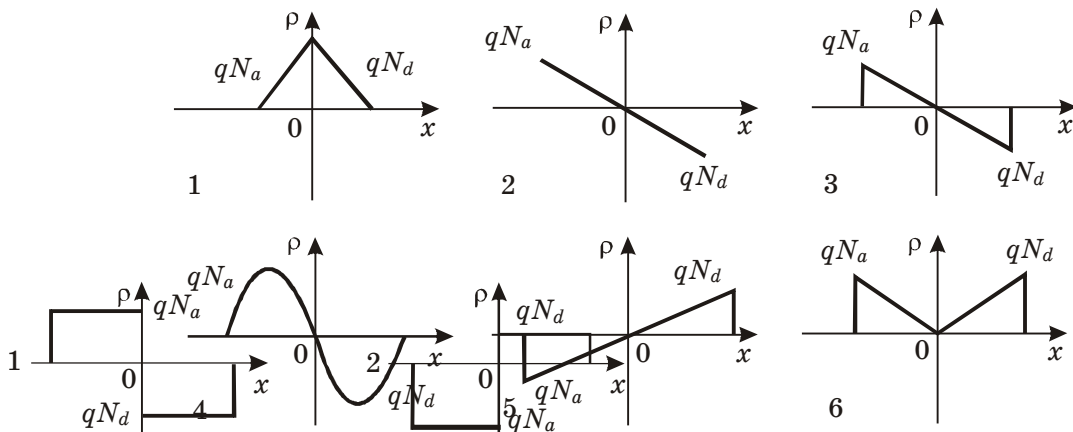
Ответы:

- 1) так как отличается распределение плотности объемного заряда;
- 2) так как напряженность электрического поля в резком *p-n*-переходе выше;
- 3) так как выше контактная разность потенциалов в резком *p-n*-переходе;
- 4) так как зарядная емкость плавного *p-n*-перехода меньше;
- 5) так как пробивное напряжение плавного *p-n*-перехода выше.

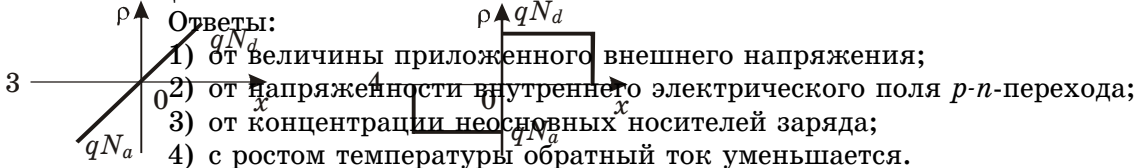
15. (ЦОПЗ). Каково распределение плотности объемного заряда в резком *p-n*-переходе?

Укажите номер правильного рисунка.

16. (ТЭПЛ). Укажите рисунок, соответствующий распределению плотности объемного заряда в плавном $p-n$ -переходе?



17. (УЦТН). От чего зависит обратный ток $p-n$ -перехода в области насыщения?



Ответы:

- 1) от величины приложенного внешнего напряжения;
- 2) от напряженности внутреннего электрического поля $p-n$ -перехода;
- 3) от концентрации неосновных носителей заряда;
- 4) с ростом температуры обратный ток уменьшается.

18. (ДЦТН). Как ток насыщения $p-n$ -перехода зависит от температуры?

Ответы:

- 1) линейно увеличивается с ростом температуры;
- 2) уменьшается с ростом температуры;
- 3) экспоненциально возрастает при увеличении температуры;
- 4) не меняется при изменении температуры.

19. (АЮКР). Как изменяется контактная разность потенциалов $p-n$ -перехода при увеличении температуры (в интервале рабочих температур $p-n$ -перехода)?

Ответы:

1) увеличивается в соответствии с выражением

$$V_k = \frac{k_0 T}{q} \ln \frac{n_n}{n_p};$$

2) уменьшается, так как концентрация основных носителей заряда остается практически постоянной, а концентрация неосновных носителей увеличивается экспоненциально;

3) увеличивается экспоненциально из-за увеличения ширины области объемного заряда;

4) уменьшается пропорционально температуре.

20. (ЖХБЕ). Какова природа барьерной емкости p - n -перехода?

Ответы:

1) емкость обусловлена накоплением объемных зарядов в базе диода;

2) емкость обусловлена перераспределением неосновных носителей заряда в p - n -переходе под действием приложенного напряжения;

3) емкость обусловлена изменением толщины слоя объемных зарядов в p - n -переходе при изменении приложенного напряжения;

4) емкость связана с накоплением электронов и дырок у барьера.

21. (ЮТЭМ). Что называется эмиттером?

Ответы:

1) область транзистора с большой концентрацией примеси;

2) область транзистора, назначением которой является инжекция в базу неосновных носителей;

3) область транзистора со стороны открытого p - n -перехода;

4) область транзистора, назначением которой является экстракция из базы неосновных носителей.

22. (ХЯИЭ). Укажите тип носителей, которые инжектируются в базы эмиттерами n - p - n - и p - n - p -транзисторов:

1) дырки в n - p - n - и электроны в p - n - p -транзисторах;

2) эмиттеры в n - p - n -транзисторах инжектируют в базу электроны, а в p - n - p -транзисторах — дырки;

3) в любых биполярных транзисторах эмиттерные переходы инжектируют в базу только дырки;

4) в любых биполярных транзисторах эмиттеры инжектируют в базы только электроны.

23. (ЯТКО). Что называется коллектором?

Ответы:

1) область транзистора со средней концентрацией примеси;

2) область транзистора, назначением которой является инжекция неосновных носителей в базу;

3) область транзистора со стороны закрытого p - n -перехода;

4) область транзистора, назначением которой является экстракция из базы неосновных носителей.

24. (ШЮОТ). Какими носителями заряда, основными или неосновными, обусловлен в биполярных транзисторах обратный ток коллектора?

Ответы:

- 1) основными носителями заряда в области базы;
- 2) неосновными носителями заряда, имеющимися в области базы и в области коллектора;
- 3) основными носителями заряда в области коллектора;
- 4) правильного ответа нет.

25. (ХЯНП). Укажите правильные ответы, объясняющие природу коллекторного тока в биполярных диффузионных транзисторах:

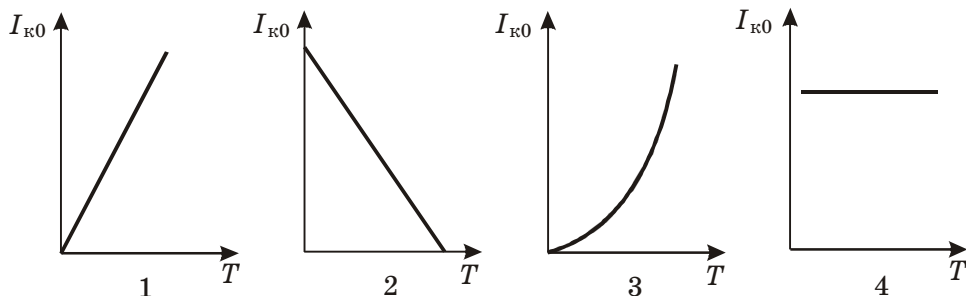
1) в диффузионном $n-p-n$ -транзисторе дырки, инжектированные из эмиттера в базу, увлекаются электрическим полем в базе и через коллекторный $p-n$ -переход уходят во внешнюю цепь;

2) в диффузионном $n-p-n$ -транзисторе электроны при прямом смещении эмиттерного $n-p$ -перехода инжектируются в базу, что приводит к увеличению в ней вблизи перехода концентрации неосновных носителей заряда. Вследствие возникшего градиента концентрации dn/dx электроны диффундируют к обратно смещенному коллекторному $p-n$ -переходу, захватываются им и перебрасываются во внешнюю цепь;

3) в диффузионном $p-n-p$ -транзисторе электроны, инжектированные из эмиттера в базу, дрейфуют в электрическом поле к коллектору, который находится под обратным напряжением, и свободно уходят во внешнюю цепь;

4) в диффузионном $p-n-p$ -транзисторе при включении $p-n$ -перехода в прямом направлении высота потенциального барьера для основных носителей в эмиттерной области (дырок) понижается на величину прямого напряжения qV , вследствие чего из p -области в базу (n -область) перетекает дополнительное количество дырок, которые там будут неосновными носителями. Инжектированные таким образом дырки за счет диффузии перемещаются по базе к обратно смещенному коллекторному $n-p$ -переходу, в котором отсутствует потенциальный барьер для неосновных носителей заряда. Дырки подхватываются электрическим полем коллекторного $n-p$ -перехода и через омический контакт выталкиваются во внешнюю цепь.

26. (РГТК). Как зависит обратный ток коллектора $I_{к0}$ от температуры T ? Укажите номер правильного рисунка.



27. (ОЮТБ). Укажите причины, объясняющие малую толщину базы и низкую концентрацию в ней основных носителей заряда для диффузионных транзисторов:

1) при уменьшении толщины базы снижается количество носителей заряда, рекомбинирующих в ней при диффузии к коллектору;

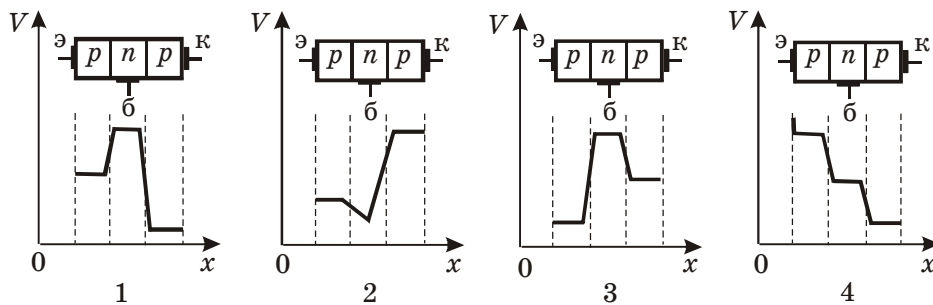
2) толщина базы определяется шириной областей эмиттерного и коллекторного p - n -переходов и поэтому зависит только от концентраций в них основных носителей заряда;

3) снижение в базе концентрации основных носителей заряда уменьшает рекомбинационные потери носителей, инжектированных эмиттерным p - n -переходом, что повышает коллекторный ток;

4) малая концентрация основных носителей заряда в базе соответствует ее большому электрическому сопротивлению, что обуславливает меньшее электрическое взаимодействие эмиттера и коллектора;

5) выбор материала базы с пониженной концентрацией основных носителей заряда позволяет повысить эффективность эмиттера.

28. (Н9РП). Укажите номер графика распределения потенциалов в бездрейфовом транзисторе типа p - n - p в активном режиме.



29. (ЛШЭК). Что характеризует эффективность коллектора в транзисторе типа p - n - p ?

Ответы:

1) увеличение дырочной составляющей коллекторного тока по отношению ко всему току;

2) пробивная прочность коллектора;

3) увеличение тока коллектора из-за электронной составляющей и лавинного умножения;

4) обратный ток коллектора.

30. (ОТЭМ). Что показывает эффективность эмиттера?

Ответы:

1) степень насыщенности эмиттера примесью;

2) долю тока, создаваемого вводимыми в базу неосновными носителями заряда, в общем токе эмиттера;

3) как велика доля вводимых в базу неосновных носителей в сравнении с неосновными носителями в коллекторе;

4) долю дырочной составляющей коллекторного тока, обусловленную током эмиттера.

31. (Ф1КП). Что показывает коэффициент переноса?

Ответы:

1) долю неосновных носителей, захваченных коллекторным переходом, в общем количестве неосновных носителей, инжектированных в базу эмиттером;

2) долю неосновных носителей, захваченных коллекторным переходом, в общем количестве носителей, движущихся через коллекторный переход;

3) долю основных носителей в общем количестве носителей, имеющих в базе;

4) долю основных носителей в общем количестве носителей, рекомбинирующих в базе.

32. (ДФТЭ). Что больше, полный ток эмиттера или ток инжекции неосновных носителей из эмиттера в базу?

Ответы:

1) ток инжекции больше;

2) ток эмиттера больше;

3) оба тока равны.

33. (ФЯКП). Как математически определяется коэффициент переноса? Укажите номер формулы:

$$\beta = \frac{\partial I_{\text{к}}}{\partial I_{\text{рк}}} \Big|_{V_{\text{эб}}=\text{const}} \quad 1) \quad ; \quad 2) \quad \beta = \frac{\partial I_{\text{рк}}}{\partial I_{\text{к}}} \Big|_{V_{\text{кб}}=\text{const}} \quad ;$$
$$3) \quad \beta = \frac{\partial I_{\text{пэ}}}{\partial I_{\text{рк}}} \Big|_{V_{\text{эб}}=\text{const}} \quad ; \quad 4) \quad \beta = \frac{\partial I_{\text{рк}}}{\partial I_{\text{пэ}}} \Big|_{V_{\text{кб}}=\text{const}} \quad .$$

34. (Ф9КП). Почему коэффициент передачи тока эмиттера меньше единицы?

Ответы:

1) вследствие инжекции неосновных носителей из коллектора в базу;

2) из-за большого сопротивления базы;

3) вследствие рекомбинации в базе части неосновных носителей, инжектированных эмиттером, и поскольку эффективность эмиттера меньше единицы;

4) вследствие расширения коллекторного *p-n*-перехода при подаче запирающего напряжения.

35. (Ё5КО). Почему коэффициент передачи тока базы больше единицы?

Ответы:

1) потому что приращение тока коллектора больше приращения тока эмиттера;

2) потому что коллекторный p - n -переход открыт положительным напряжением;

3) потому что изменение тока базы составляет малую часть изменения тока эмиттера, определяющего приращение коллекторного тока.

36. (ОХЭФ). Может ли быть эффективность коллектора больше единицы?

Ответы:

1) не может;

2) может за счет лавинного умножения в коллекторном переходе;

3) может за счет большой площади коллекторного перехода;

4) может при малых коллекторных напряжениях за счет малой высоты потенциального барьера коллекторного перехода.

37. (УЯЭЭ). Как добиться, чтобы эффективность эмиттера была возможно большей?

Ответы:

1) сделать удельное сопротивление эмиттера много меньшим удельного сопротивления базы;

2) увеличить площадь эмиттерного перехода;

3) уменьшить ширину базы;

4) увеличить удельное сопротивление базы в сравнении с удельным сопротивлением эмиттера.

38. (ЖНАЭ). В какой из областей транзистора больше атомов примеси, в области базы или в области эмиттера?

Ответы:

1) в области базы;

2) в области эмиттера;

3) равное количество в обеих частях.

39. (ГКБС). Почему концентрация примесей в базе намного меньше концентрации примеси в эмиттере?

Ответы:

1) чтобы увеличить эффективность эмиттера;

2) чтобы увеличить сопротивление базы;

3) чтобы увеличить постоянную времени эмиттерной и коллекторной емкостей;

4) чтобы уменьшить вероятность рекомбинации основных носителей в базе.

40. (ЕЯКП). Может ли быть коэффициент передачи тока эмиттера больше единицы и почему?

Ответы:

1) может из-за умножения тока неосновных носителей в эмиттерном переходе;

2) может из-за лавинного умножения в коллекторном переходе;

3) не может быть больше единицы, так как часть неосновных носителей рекомбинирует в базе.

41. (ЦМИЗ). Чем объяснить изменение коэффициента передачи тока эмиттера (или базы) при изменении коллекторного напряжения?

Ответы:

- 1) увеличением инжекции носителей в базу;
- 2) уменьшением противотока носителей из коллектора в базу;
- 3) изменением в основном ширины коллекторного перехода и соответственно изменением толщины базы.

42. (ДФКН). Чем ограничены значения коллекторного напряжения?

Ответы:

- 1) величина коллекторного напряжения ограничивается пробоем коллекторного перехода;
- 2) коэффициент переноса становится слишком большим;
- 3) ширина коллекторного перехода в области коллектора становится слишком большой;
- 4) коэффициент обратной связи достигает предельно допустимой величины.

43. (БРТК). Будет ли в цепи коллектора транзистора в схеме с общей базой протекать ток, если $I_3 = 0$, а $V_{кб} < 0$?

Ответы:

- 1) не будет из-за отсутствия инжекции дырок в базу;
- 2) будет из-за наличия тока неосновных носителей базы и коллектора через коллекторный переход;
- 3) ток коллектора будет определяться инжекцией носителей из коллектора в базу;
- 4) ток коллектора будет определяться инжекцией носителей из базы в коллектор.

44. (ФЮКЗ). Когда ток в цепи коллектора может быть равен нулю при включении транзистора типа $p-n-p$ по схеме с общей базой и с эмиттерным переходом, смещенным в прямом направлении?

Ответы:

- 1) при $V_{кб} = 0$ и отсутствии тока основных носителей через коллекторный переход;
- 2) при $V_{кб} > 0$ вследствие компенсации в коллекторном переходе токов основных и неосновных носителей заряда;
- 3) при $V_{кб} = 0$ и отсутствии тока неосновных носителей;
- 4) при $V_{кб} < 0$ вследствие увеличения высоты потенциального барьера перехода коллектор-база.

Глава 9. Поверхностные явления в полупроводниках

§ 9.1. Поверхностные состояния

Образование поверхности любого кристаллического вещества неизбежно сопровождается разрывом химических связей между атомами в кристаллической решетке и, следовательно, нарушением периодичности потенциального поля $U(x)$, представление о котором рассматривалось в главе 1.

Известно, что нарушения периодичности потенциала, вызванные дефектами в объеме кристаллического полупроводника (примесные атомы в узлах решетки, вакансии и др.), обуславливают создание разрешенных энергетических уровней в запрещенной зоне. Тогда поверхность в целом представляет собой дефект, который объясняет возникновение разрешенных энергетических уровней в запрещенной зоне поверхностного слоя полупроводника (плоскость $x = 0$). Эти уровни называются поверхностными уровнями, которым соответствуют поверхностные состояния с концентрацией, равной плотности поверхностных атомов ($\approx 10^{15} \text{ см}^{-2}$).

Существование электронных состояний, обусловленных нарушением периодичности потенциала на поверхности полупроводников, показал в 1932 г. советский физик-теоретик И.Е. Тамм.

У атомов, расположенных на поверхности четырехвалентного полупроводника (германий, кремний), одна валентная связь не укомплектована электроном. На рис. 9.1 изображена плоская модель решетки германия, где плоскость AA представляет поверхность полупроводника. Каждый атом в объеме связан с четырьмя ближайшими соседями ковалентными связями, в которых участвуют четыре пары обобществленных валентных электронов. У атомов на поверхности AA одна валентная связь оказывается разорванной, а электронная пара — неукомплектованной. Создание устойчивой восьмиэлектронной конфигурации может быть связано с переходом электрона от атомов из более глубоких слоев (см. рис. 9.1).

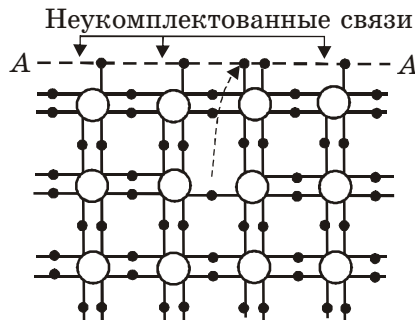
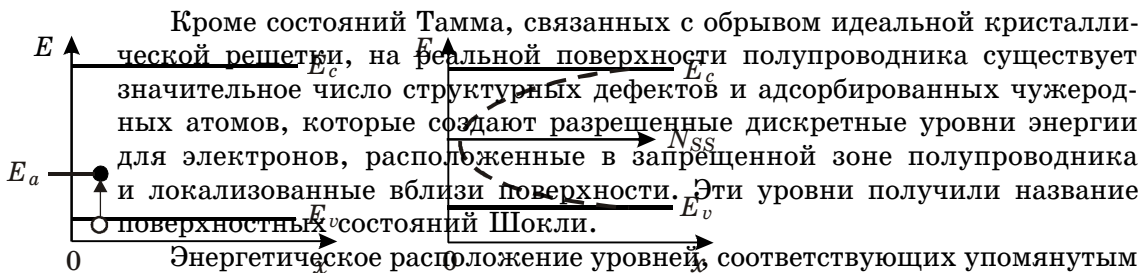


Рис. 9.1. Образование поверхностных состояний Тамма

Таким образом, поверхностные атомы ведут себя как акцепторы, которым в запрещенной зоне на поверхности соответствует акцепторный поверхностный уровень E_a . Действительно, в 1959 г. на германии n -типа был экспериментально обнаружен поверхностный слой ярко выраженного p -типа проводимости. Электроны, захваченные на эти уровни, закрепляются в первом от поверхности атомном слое (рис. 9.2,*a*). В валентной зоне вблизи плоскости $x = 0$ возникают дырки, что соответствует дырочной проводимости поверхностного слоя полупроводника.

Рис. 9.2. Поверхностные состояния в полупроводнике:
a — показан один тип поверхностного акцепторного уровня E_a
 с плотностью N_{SS} ; *b* — экспериментальная зависимость
 плотности поверхностных состояний N_{SS}
 от их положения в запрещенной зоне



Кроме состояний Тамма, связанных с обрывом идеальной кристаллической решетки, на реальной поверхности полупроводника существует значительное число структурных дефектов и адсорбированных чужеродных атомов, которые создают разрешенные дискретные уровни энергии для электронов, расположенные в запрещенной зоне полупроводника и локализованные вблизи поверхности. Эти уровни получили название поверхностных состояний Шокли.

Энергетическое расположение уровней, соответствующих упомянутым поверхностным состояниям, зависит от природы поверхности и частиц, адсорбированных на ней. Атомы кислорода, адсорбированные на поверхности германия, создают акцепторные уровни; молекулы воды — донорные. Учитывая многообразие типов химических элементов, которые могут быть адсорбированы на поверхности полупроводника в реальных условиях, распределение поверхностных состояний в запрещенной зоне может быть как дискретным типа E_a (рис. 9.2,*a*), E_a , так и непрерывным, т.е. монотонно изменяющимся по запрещенной зоне (рис. 9.2,*b*). На практике чаще всего наблюдается увеличение плотности поверхностных состояний N_{SS} в запрещенной зоне к краям разрешенных зон. Зависимость $N_{SS}(E)$ в запрещенной зоне показана на рис. 9.2,*b* штриховой линией.

Поверхностные состояния, природа которых связана со структурными дефектами поверхности, находятся в хорошем контакте с объемом полупроводника, вследствие чего время их обмена носителями заряда

с разрешенными зонами в объеме оказывается около 10^{-7} с. Такие состояния принято называть быстрыми поверхностными состояниями (БПС). Они имеют плотность порядка 10^{11} см⁻², зависящую от характера обработки поверхности, и обладают большими сечениями захвата как для электронов, так и для дырок, вследствие чего могут служить эффективными центрами рекомбинации.

В обычных условиях полупроводник находится в среде атмосферных газов, поэтому поверхность всегда покрыта слоем окисла толщиной в несколько нанометров и более. На внешней поверхности и в объеме окисла адсорбируются примесные атомы, создающие внешние, или медленные, поверхностные состояния (МПС) (рис. 9.3).

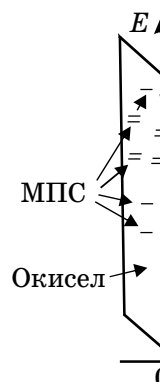
Плотность МПС изменяется в диапазоне $10^{13}-10^{14}$ см⁻² в зависимости от природы и состояния внешней среды. В связи с удаленностью время установления равновесия МПС с объемом полупроводника значительно больше, чем для БПС, и может изменяться от микросекунд до суток. Это обусловлено малой вероятностью прохождения носителей заряда

Рис. 9.3. Быстрые и медленные поверхностные состояния

сквозь окисный слой ($x < 0$), который является изолятором. Поэтому с увеличением толщины окисной пленки постоянная времени перезарядки МПС заметно возрастает.

§ 9.2. Приповерхностный слой объемного заряда

Накопление заряда в поверхностных состояниях сопровождается его нейтрализацией противоположным зарядом в приповерхностном слое полупроводника. Процесс нейтрализации заключается в подтягивании в приповерхностную область из объема или выталкивании из нее в объем свободных носителей заряда. Отрицательный заряд электронов, захваченных в акцепторные поверхностные состояния, нейтрализуется подтягиванием в приповерхностную область свободных дырок из объема полупроводника и отталкиванием из приповерхностной области в объем свободных электронов. В результате неподвижный отрицательный заряд в акцепторных поверхностных состояниях уравнивается, т.е. нейтрализуется со стороны полупроводника положительным зарядом в приповерхностной области дырок и примесных ионов, оставленных электронами, ушедшими в объем. Так называемые плоские энергетические диаграммы, характерные для идеального объема полупроводника (см. гл. 2), искривляются



у поверхности полупроводника под действием электрического поля, созданного поверхностными зарядами. Зонная энергетическая диаграмма для вышеописанного случая нейтрализации заряда в акцепторных поверхностных состояниях донорного полупроводника показана на рис. 9.4,а.

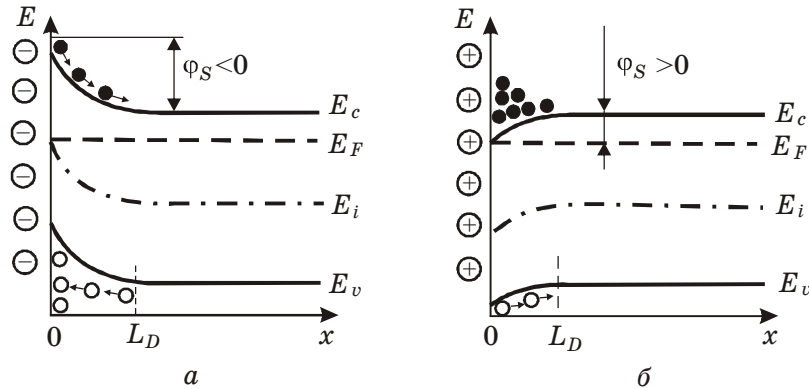


Рис. 9.4. Изгиб зон у поверхности полупроводника при наличии на поверхности заряженных состояний акцепторного типа (а), донорного типа (б)

Положительный заряд в донорных поверхностных состояниях уравновешивается, т.е. компенсируется отрицательным зарядом в приповерхностной области полупроводника n -типа (рис. 9.4,б) зарядом основных носителей (электронов), которые подтягиваются к поверхности из объема. Дырки (неосновные носители заряда) отталкиваются в объем.

Количество носителей заряда, необходимых для нейтрализации заряда в поверхностных состояниях, определяет ширину области приповерхностного заряда (ОПЗ), которая называется дебаевской длиной экранирования L_D . Дебаевская длина L_D — это расстояние, на которое электрическое поле, связанное с зарядом в поверхностных состояниях, проникает в полупроводник. На расстоянии L_D потенциал, соответствующий данному полю, уменьшается в e раз.

Таким образом, приповерхностный слой полупроводника является областью пространственного заряда, экранирующей объем полупроводника от поля поверхностного заряда, причем экранировка осуществляется за счет того, что равновесные концентрации электронов и дырок в ОПЗ отличаются от объемных. Это приводит к появлению потенциального барьера в приповерхностной области полупроводника. Его высота и знак характеризуются поверхностным потенциалом ϕ_S , который однозначно связан с величиной и знаком поверхностного заряда. Длину экранирования Дебая можно найти при интегрировании уравнения Пуассона (8.7) для соответствующих рассматриваемому случаю граничных условий:

— для полупроводника n -типа

$$L_D = \left[\frac{\varepsilon\varepsilon_0 k_0 T}{q^2 n_{n0}} \right]^{1/2}; \quad (9.1)$$

— для полупроводника p -типа

$$L_D = \left[\frac{\varepsilon\varepsilon_0 k_0 T}{q^2 p_{p0}} \right]^{1/2}. \quad (9.2)$$

Из выражений (9.1) и (9.2) видно, что длина экранирования увеличивается при уменьшении концентрации свободных носителей заряда. Это обусловлено тем, что для экранировки поверхностного заряда необходимое число свободных носителей заряда может быть собрано из слоя большей толщины. В полупроводниках величина L_D изменяется в интервале 10^{-6} – 10^{-4} см.

§ 9.3. Заряд в приповерхностной области полупроводника

Заряд в приповерхностной области определяется величиной и знаком поверхностного потенциала, за который можно принять величину φ_S , численно равную изгибу зон, т.е. $\varphi_S = -qV_S$ (см. рис. 9.4). Изгиб зон вверх соответствует отрицательным значениям φ_S , вниз — положительным.

В полупроводнике n -типа изгиб зон вниз ($\varphi_S > 0$) является следствием подтягивания основных носителей заряда в ОПЗ (см. рис. 9.4,б), при этом увеличение φ_S сопровождается ростом концентрации электронов n_S в приповерхностной области. Отрицательный заряд в ОПЗ (Q_S), равный в этом

случае qn_S , возрастает с ростом поверхностного потенциала (рис. 9.5). Приповерхностный слой, в котором концентрация основных носителей выше, чем в объеме, называется обогащенным. При этом в электронном полупроводнике

$$n_S > n_{n0}. \quad (9.3)$$

Изгиб зон вверх, т.е. при $\varphi_S < 0$ (см. рис. 9.4,а), свидетельствует об отталкивании основных носителей заряда (в данном случае электронов) в объем полупроводника. В ОПЗ остаются положительные донорные ионы, которые

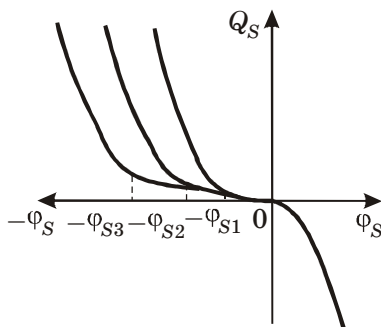


Рис. 9.5. Зависимость заряда Q_S в приповерхностной области от величины изгиба зон φ_S для донорного полупроводника

определяют заряд приповерхностной области в интервале поверхностного потенциала от 0 до $+Q_{S1}$. В этом интервале приповерхностный слой характеризуется концентрацией основных носителей, величина которой меньше, чем в объеме,

$$n_S < n_0, \quad (9.4)$$

но тип электропроводности не изменяется, т.е. $n_S > p_S$. Такая ОПЗ называется обедненной областью. В ней концентрация неосновных носителей заряда (дырок) с приближением к поверхности возрастает, но уровень Ферми в слое толщиной L_D находится в верхней половине запрещенной зоны (см. рис. 9.4,а), что характерно для полупроводников n -типа. При $\varphi_S = \varphi_{S1}$ уровень E_F на поверхности полупроводника совпадает с серединой запрещенной зоны, и при дальнейшем увеличении изгиба вверх уровень Ферми в приповерхностном слое толщиной L_{inv} (рис. 9.6) перемещается в нижнюю половину запрещенной зоны. В этом слое, который называется инверсионным, выполняется следующее соотношение между концентрациями свободных носителей заряда:

$$p_S > n_S. \quad (9.5)$$

Рост изгиба зон вверх соответствует быстрому увеличению концентрации неосновных носителей p_S и положительного заряда дырок в ОПЗ (см. рис. 9.5).

Повышение концентрации основных носителей заряда в объеме приводит к смещению уровня Ферми к соответствующей разрешенной зоне (§ 2.12). В данном случае рост концентрации электронов в объеме n_{n0} равнозначен смещению уровня Ферми ко дну зоны проводимости и соответствующему увеличению поверхностного потенциала, при котором возникает инверсионный слой (инверсия). На рис. 9.5 величины $\varphi_{S1} < \varphi_{S2} < \varphi_{S3}$ соответствуют разным поверхностным потенциалам, при которых достигается инверсия в полупроводниках с равновесными концентрациями электронов $n_{n1} < n_{n2} < n_{n3}$.

В полупроводнике p -типа слой, обогащенный основными носителями, возникает при отрицательных значениях поверхностного потенциала, а обедненный и инверсионный слои формируются при изгибах зон вниз. Заряд в указанных слоях при этом будет отрицательным и состоящим из зарядов акцепторных ионов и электронов. С указанными уточнениями объяснение природы заряда в ОПЗ в полупроводниках p -типа будет аналогичным изложенному выше для донорных полупроводников.

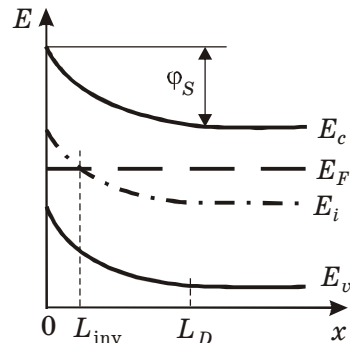


Рис. 9.6. Возникновение инверсионного слоя в донорном полупроводнике

§ 9.4. Емкость структуры металл-диэлектрик-полупроводник

Рассмотрим явления в приповерхностном слое полупроводника на примере структуры металл-диэлектрик-полупроводник (МДП), состоящей из металлического и диэлектрического слоев, нанесенных на полупроводник с омическим контактом, в котором отсутствует потенциальный барьер (рис. 9.7,а).

Рис. 9.7. МДП-структура (а) и ее эквивалентная схема (б)

МДП-структура, таким образом, представляет собой плоский конденсатор, одной обкладкой которого является полевой металлический электрод, другой — полупроводник. При внешней разности потенциалов V_g на обкладках индуцируется заряд Q , величина которого пропорциональна емкости конденсатора C и разности потенциалов:

$$Q = C V_g . \quad (9.6)$$

Заряд, индуцированный на полевом электроде, как отмечалось в § 8.3, сосредоточен непосредственно на поверхности металла, а в полупроводнике — в ОПЗ шириной L_D (см. рис. 9.4 и 9.6). Тогда емкость МДП-структуры можно представить в виде последовательного соединения емкости диэлектрика C_d и емкости ОПЗ C_S , как показано на рис. 9.7,б.

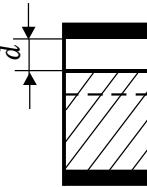
Емкость диэлектрика определяется его толщиной d и диэлектрической проницаемостью ϵ :

$$C_d = \frac{\epsilon \epsilon_0 S}{d} . \quad (9.7)$$

Емкость ОПЗ можно представить, исходя из общего определения емкости, формулой:

$$C_S = \frac{d Q_S}{d \varphi_S} . \quad (9.8)$$

Тогда зависимость $C_S(\varphi_S)$ для полупроводника n -типа, полученная дифференцированием кривой $Q_S(\varphi_S)$, показанной на рис. 9.5, будет иметь вид, представленный на рис. 9.8. Емкость МДП-структуры согласно эквивалентной схеме (рис. 9.7,б), равна



(9.9)

Емкость диэлектрика не зависит от внешней разности потенциалов на полевом электроде. В то же время зависимость $C_S(V_g)$ будет иметь вид, близкий к представленному на рис. 9.8.

На рис. 9.9,а показаны зависимости $C_d(V_g)$ и $C_S(V_g)$. Зависимость емкости идеальной МДП-структуры (не имеющей поверхностного заряда) от напряжения при положительных значениях последнего отвечает накоплению электронов в ОПЗ у поверхности полупроводника. В этом режиме дифференциальная емкость полупроводника C_S значительно больше емкости диэлектрика, поэтому полная емкость структуры близка к C_d (см. выражение (9.9) и рис. 9.9,а). При подаче отрицательного напряжения на металлический электрод в приповерхностном слое полупроводника образуется обедненная область, которую можно трактовать как добавочный слой диэлектрика. Это приводит к уменьшению полной емкости МДП-структуры. При дальнейшем увеличении отрицательного смещения у границы раздела полупроводника с диэлектриком образуется инверсионный слой, т.е. слой, обогащенный дырками, дифференциальная емкость которого также значительно превышает емкость диэлектрика. Поэтому в режиме инверсии емкость МДП-структуры будет определяться емкостью диэлектрика.

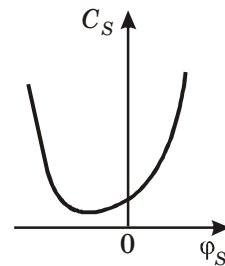


Рис. 9.8. Зависимость емкости ОПЗ от ϕ_S

$$C = \frac{C_d C_S}{C_d + C_S}$$

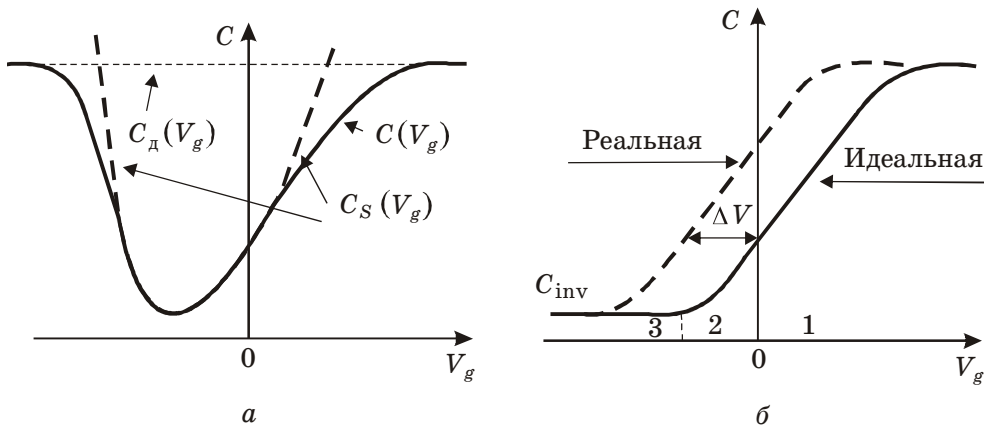


Рис. 9.9. Вольт-фарадная характеристика МДП-структуры с донорным полупроводником при низкочастотном (а) и высокочастотном (б) напряжении

Необходимо отметить, что нарастание емкости при отрицательных напряжениях (режим инверсии) зависит от того, успевает ли концентрация

инверсионных дырок следовать за изменениями приложенного к структуре переменного напряжения, с помощью которого измеряется емкость. Описанный режим осуществляется лишь при сравнительно малых частотах, когда скорость генерационно-рекомбинационных процессов, ответственных за изменение концентрации неосновных носителей заряда (в рассматриваемом случае — дырок), достаточна, чтобы дырочная плотность изменялась в фазе с напряжением измерительного сигнала. Экспериментально установлено, что для структуры металл-оксид кремния-кремний соответствующая частотная граница лежит в диапазоне 5–100 Гц. При более высоких частотах увеличение дифференциальной емкости МДП-структуры при отрицательных напряжениях не наблюдается, и вольт-фарадная характеристика имеет вид, представленный на рис. 9.9,б (сплошная кривая).

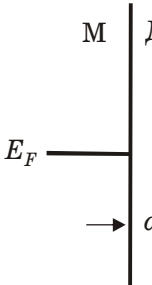
Как уже отмечалось, в идеальной МДП-структуре отсутствует поверхностный заряд, поэтому при нулевом напряжении на затворе поверхностный потенциал равен нулю и не наблюдается искривление энергетических зон в приповерхностной области полупроводника (так называемое состояние «плоских зон», рис. 9.10).

Существование начального заряда в поверхностных состояниях полупроводника и в диэлектрике приводит к изгибу энергетических зон, обусловленному наличием начального поверхностного потенциала. Его компенсация (нейтрализация) внешним напряжением V_g выражается в смещении реальных кривых $C(V_g)$ по оси напряжений. В зависимости от знака

исходного заряда в МДП-структуре наблюдается сдвиг кривых $C(V_g)$ в сторону отрицательных или положительных значений V_g . Первый случай соответствует существованию положительного начального заряда МДП-структуры, второй — отрицательного. По величине смещения ΔV на уровне емкости плоских зон при $V_g = 0$ на теоретической зависимости $C(V_g)$, как показано на рис. 9.9,б, можно определить величину начального заряда в МДП-структуре. В соответствии с формулой (9.6)

$$Q = C_d \Delta V. \quad (9.10)$$

Рис. 9.10. Энергетическая диаграмма идеальной МДП-структуры при нулевом напряжении на затворе



На рис. 9.9,б область 1 соответствует обогащению ОПЗ, а максимальная величина — емкости C_d . Области 2 и 3 связаны соответственно с формированием обедненного и инверсионного слоев.

Вид кривой $C(V_g)$ зависит от типа проводимости полупроводника. В акцепторном полупроводнике инверсионный слой формируется в области положительных, обогащенный — в области отрицательных потенциалов на полевом электроде.

§ 9.5. Эффект поля

Внешнее электрическое поле, направление которого совпадает с нормалью к поверхности полупроводника, изменяет изгиб энергетических зон φ_S и, следовательно, концентрацию свободных носителей заряда в приповерхностном слое, что аналогично процессам, описанным в предыдущих параграфах.

Рассмотрим полупроводниковый образец n -типа, по которому в продольном направлении пропускается постоянный электрический ток (рис. 9.11). Такой образец, фактически, подобен МДП-структуре (см. рис. 9.7,а). Поперечное электрическое поле создается подачей на полевой электрод относительно полупроводника внешней разности потенциалов V_g .

Рис. 9.11. Схема для наблюдения поверхностной проводимости полупроводника

В режиме обогащения концентрация основных носителей заряда (электронов) в ОПЗ будет превышать объемную на величину Δn_S . Полная концентрация электронов в ОПЗ будет равна

$$n_S = n_{n0} + \Delta n_S. \quad (9.11)$$

Изгиб энергетических зон φ_S и изменение концентрации носителей заряда в приповерхностном слое полупроводника приведет к изменению его электропроводности, которое выражается формулой

$$\Delta \sigma_S = q \mu_{nS} \Delta n_S, \quad (9.12)$$

где μ_{nS} — подвижность электронов в слое ОПЗ.

Обычно поверхностная подвижность электронов и дырок меньше, чем объемная, вследствие дополнительного рассеяния носителей на поверхностных дефектах в тонком слое ОПЗ.

Величина $\Delta \sigma_S$ представляет собой поверхностную проводимость, которая возрастает при увеличении обогащающего напряжения на полевом электроде (рис. 9.12).

Обеднение приповерхностного слоя полупроводника n -типа под воздействием отрицательного напряжения на полевом электроде объясняется уменьшением концентрации основных носителей заряда за счет их

Рис. 9.12. Зависимость поверхностной проводимости от смещающего напряжения для донорного полупроводника

выталкивания из приповерхностного слоя в объем внешним электрическим полем. При увеличении отрицательного потенциала $-V_g$ поверхностная проводимость уменьшается, что подчеркивается отрицательным знаком $\Delta\sigma_S$ на рис. 9.12 в области обеднения. Повышение отрицательного потенциала приводит к состоянию ОПЗ, когда концентрация неосновных носителей заряда (дырок) сначала уравнивается с концентрацией электронов, а при дальнейшем увеличении отрицательного напряжения превышает ее. Поверхностная проводимость становится дырочной, т.е. происходит инверсия поверхностной проводимости.

Поверхностная проводимость, в которой участвуют основные и неосновные носители заряда, описывается формулой

$$\Delta\sigma_S = q\mu_{nS}\Delta n_S + q\mu_{pS}\Delta p_S. \quad (9.13)$$

Так как обычно $\mu_{pS} < \mu_{nS}$, то минимум на так называемой кривой эффекта поля $\Delta\sigma_S = f(V_g)$ наблюдается уже в состоянии инверсии, т.е. при $p_S > n_S$. Изменение поверхностной проводимости полупроводника под действием поперечного электрического поля называется эффектом поля, хотя и изменение емкости МДП-структуры в аналогичных обстоятельствах (§ 9.4) также можно представить эффектом поля.

Зависимость $\Delta\sigma_S(V_g)$ для дырочного полупроводника показана на рис. 9.13. Режим обогащения дырками приповерхностного слоя полупроводника наблюдается при отрицательных потенциалах на полевом электроде, режим обеднения и инверсии — при положительных.

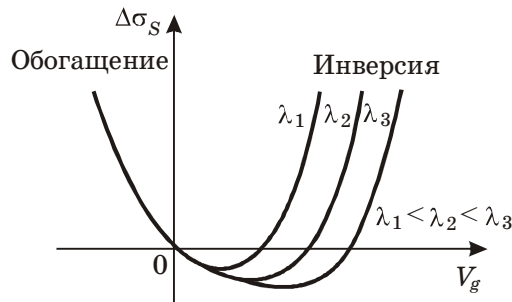


Рис. 9.13. Зависимость поверхностной проводимости от смещающего напряжения для акцепторного полупроводника

На практике кривые эффекта поля можно наблюдать на установке, упрощенная принципиальная схема которой показана на рис. 9.11.

Изменение тока через образец, равное

$$\Delta j = \Delta\sigma_S \mathbf{E}, \quad (9.14)$$

можно регистрировать по изменению падения напряжения ΔV_H на сопротивлении нагрузки R_H .

Положение минимума $\Delta\sigma_S$ по оси V_g (см. рис. 9.12 и 9.13) легко определить из условия $\frac{d\Delta\sigma_S}{dV_g} = 0$, полагая, что в соответствии с распределением Максвелла — Больцмана Δn_S и Δp_S равны:

При этом покажем, что величина $V_g < 0$ при изгибе зонной диаграммы в приповерхностной области вверх и $V_g > 0$ при изгибе зонной диаграммы вниз.

Расчет показывает, что положение минимума на кривой $\Delta\sigma_S(V_g)$ зависит от концентрации примеси в полупроводнике:

$$V_g^{\min} = \frac{k_0 T}{2g} \ln \left[\frac{\lambda^2}{b} \right], \quad (9.15)$$

где λ — показатель уровня легирования полупроводника; n_i — собственная концентрация носителей заряда в полупроводнике; n_0 — равновесная концентрация электронов в объеме полупроводника; $b = \frac{\mu_n}{\mu_p}$ — отношение подвижностей электронов и дырок.

Из выражения (9.15) следует, что для донорного полупроводника, когда $\lambda < 1$, а $b \approx 1$, величина V_g^{\min} отрицательна и возрастает по абсолютному значению при увеличении степени легирования, т.е. при уменьшении λ (см. рис. 9.12). Для акцепторного полупроводника $\lambda > 1$, $b \approx 1$. Следовательно, V_g^{\min} будет положительной величиной, возрастающей при увеличении концентрации акцепторной примеси в полупроводнике (см. рис. 9.13).

Смещение минимума поверхностной проводимости с увеличением степени легирования объясняется тем, что при возрастании концентрации примеси увеличивается энергетическое расстояние между уровнем Ферми и серединой запрещенной зоны и поэтому увеличивается потенциал, который необходимо приложить к поверхности полупроводника относительно объема для создания в приповерхностной области инверсного состояния.

Следует отметить, что положение минимума кривой $\Delta\sigma_S(V_g)$ зависит от параметров полупроводника в объеме (n_i/n_0 , μ_n/μ_p) и от свойств поверхности. Наличие поверхностных заряженных состояний приводит к возникновению на поверхности потенциала ϕ_S и созданию внутреннего электрического поля в приповерхностной области полупроводника. В результате при измерении $\Delta\sigma_S(V_g)$ действующее значение потенциала на

поверхности полупроводника равно $V_S + V_g$, т.е. положение минимума зависимости $\Delta\sigma_S(V_g)$ оказывается смещенным по оси V_g на величину V_S . Таким образом, сравнение положений минимума теоретической зависимости $\Delta\sigma_S(V_g)$ и экспериментально измеренной $\Delta\sigma_S(V_g + V_S)$ позволяет определить величину и знак поверхностного потенциала V_S .

В целом из измерений эффекта поля можно определить тип поверхностной проводимости, изгиб зон на поверхности полупроводника в зависимости от приложенного напряжения, величину захваченного поверхностными состояниями заряда, оценить энергетическое положение этих состояний и их концентрацию.

§ 9.6. Полевые транзисторы с изолированным затвором

Структура металл-диэлектрик-полупроводник является основой для изготовления приборов с переносом заряда, полевых транзисторов с изолированным затвором, элементов памяти, твердотельных матричных видеоконвертеров и др. Во всех этих приборах используется изменение концентрации носителей заряда в приповерхностной области полупроводника под действием внешнего поля или других факторов. Рассмотрим один из типов названных выше приборов — полевой транзистор с изолированным затвором, который является униполярным транзистором, так как ток в нем создается носителями заряда одного знака. Структуры двух разновидностей таких транзисторов и их условное обозначение на схемах показаны на рис. 9.14.

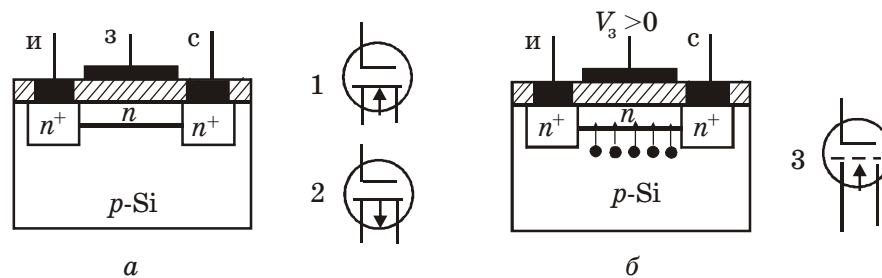


Рис. 9.14. Структура полевых МДП-транзисторов со встроенным (а) и индуцированным (б) каналами и их условные обозначения:

- 1 — МДП-транзистор со встроенным каналом n -типа;
- 2 — со встроенным каналом p -типа;
- 3 — с индуцированным каналом n -типа

В кристалле полупроводника с относительно высоким удельным сопротивлением, называемым подложкой (на рис. 9.14 подложка p -типа), созданы две сильнолегированные области противоположного типа про-

димости (n^+ -типа). Эти области с нанесенными металлическими электродами получили названия: одна — исток (и), другая — сток (с). Расстояние между истоком и стоком составляет несколько микрометров. Поверхность кристалла полупроводника между истоком и стоком покрыта тонким (порядка 0,1 мкм) слоем диэлектрика. На слой диэлектрика нанесен металлический электрод, называемый затвором (з). Получается структура, состоящая из металла, диэлектрика и полупроводника, т.е. МДП-структура. В последнее время в качестве затвора используется высоколегированный поликристаллический кремний (поликремний), а также некоторые комбинации поликремния и силицидов.

Полевой транзистор с изолированным затвором, в котором в качестве изолирующего слоя между затвором и подложкой использован диэлектрик, называют полевым транзистором типа металл-диэлектрик-полупроводник или МДП-транзистором.

Выпрямляющие электрические переходы под истоком и стоком могут быть выполнены не только в виде p - n -переходов, но и в виде выпрямляющих переходов Шоттки, т.е. путем нанесения металлических электродов истока и стока непосредственно на подложку (создаются запирающие контакты металл-полупроводник). Использование выпрямляющих переходов Шоттки под истоком и стоком может обеспечить ряд преимуществ в технологии изготовления таких транзисторов, а также улучшить их характеристики. К настоящему времени известны МДП-транзисторы, изготовленные на различных полупроводниковых материалах, таких как германий, кремний, арсенид галлия, с использованием различных диэлектрических слоев (диоксид кремния, нитрид кремния, диоксид алюминия); наибольшее распространение получили МДП-транзисторы на кремнии-диоксиде кремния. Иногда такие транзисторы называют полевыми транзисторами типа металл-оксид-полупроводник или МОП-транзисторами. Однако чаще используют более общий термин — МДП-транзисторы.

Существует две разновидности таких транзисторов — с индуцированным и со встроенным каналом (рис. 9.14, *а, б*). В МДП-транзисторах с индуцированным каналом (рис. 9.14, *б*) проводящий канал между сильнолегированными областями стока и истока и, следовательно, заметный ток стока появляются при напряжении на затворе определенной величины и полярности относительно истока, называемом пороговым напряжением $U_{зи пор}$.

В МДП-транзисторах со встроенным каналом (рис. 9.14, *а*) у поверхности полупроводника под затвором при нулевом напряжении на затворе относительно истока существует инверсионный слой — канал, который соединяет исток со стоком.

Изображенные на рис. 9.14 структуры МДП-транзисторов имеют подложку p -типа, а встроенный или индуцированный канал — n -типа. Если же аналогичные транзисторы созданы на подложке n -типа, то канал у них будет проводимости p -типа.

§ 9.7. Принцип работы и характеристики МДП-транзистора со встроенным каналом

При включении МДП-транзистора в цепь к стоку присоединяют плюс источника напряжения V_c относительно подложки, которая вместе с истоком находится под нулевым потенциалом (рис. 9.15).

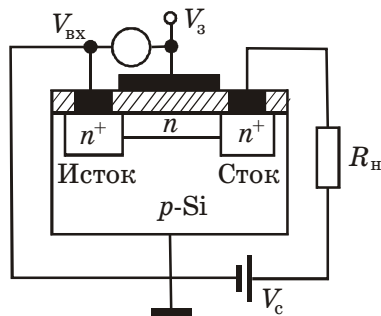


Рис. 9.15. Схема включения МДП-транзистора со встроенным каналом

Если при напряжении затвора $V_3 = 0$ между истоком и стоком приложить напряжение V_c , то по каналу потечет ток электронов. Через кристалл (подложку) ток не пойдет, так как $p-n$ -переход между ним и каналом, а также $p-n$ -переход стока находятся под обратным напряжением. При возрастании напряжения V_c от нуля ток растет пропорционально напряжению, а затем скорость роста уменьшается, так как канал начинает сужаться в первую очередь около стока. На рис. 9.16,а стоковая характеристика, т.е. зависимость $I_c(V_c)$, выходит на насыщение при $V_{c \text{ нас}}$.

Природа насыщения связана с расширением ОПЗ в обратносмещенном $p-n$ -переходе, отделяющей канал от подложки вблизи стока за счет дополнительного обратного смещения, создаваемого падением напряжения на сопротивлении канала при протекании тока стока. Ток стока изменяется слабо при $V_{c \text{ нас}} < V_c < V_{c \text{ проб}}$, где $V_{c \text{ проб}}$ — пробивное напряжение.



Рис. 9.16. Стоковые характеристики МДП-транзистора со встроенным каналом (а) и характеристика передачи (б)

При подаче на затвор отрицательного напряжения (относительно истока) в канале создается поперечное электрическое поле, под влиянием которого носители заряда (электроны) выталкиваются из канала. Канал

обедняется электронами, вследствие чего его сопротивление увеличивается и ток стока уменьшается. Чем больше отрицательное смещение затвора, тем меньше ток I_c . Такой режим работы МДП-транзистора называется режимом обеднения.

Положительное напряжение на затворе ($V_3 > 0$) создает электрическое поле, под влиянием которого электроны подтягиваются в канал, что приводит к росту его проводимости и увеличению тока стока. Этот режим называется режимом обогащения.

Рассмотренный транзистор может, таким образом, работать как в режиме обеднения, так и в режиме обогащения. Это справедливо только для транзистора со встроенным каналом. Зависимость $I_c(V_3)/V_c = \text{const}$, называемая характеристикой передачи, представлена на рис. 9.16,б. Статические характеристики передачи выходят из точки на оси абсцисс, соответствующей напряжению отсечки $V_{зи\text{ отс}}$, т.е. напряжению между затвором и истоком МДП-транзистора со встроенным каналом, работающего в режиме обеднения, при котором ток стока достигает заданного низкого значения.

§ 9.8. Принцип работы и характеристики МДП-транзистора с индуцированным каналом

Рассмотрим полевой транзистор, который отличается от описанного в § 9.7 тем, что имеет проводящий канал индуцированного типа. Структура и условное графическое изображение такого транзистора показаны ранее (см. условное обозначение на рис. 9.14,б). Природа возникновения инверсии поверхностной проводимости подробно рассмотрена в § 9.5, откуда следует, что инверсионный слой на подложке p -типа может возникнуть только при положительном смещении на полевом электроде (затворе) или при наличии в диэлектрике встроенного положительного заряда.

В транзисторе с индуцированным каналом при $V_3 = 0$ инверсионный слой на подложке p -типа, а значит, и канал n -типа отсутствуют. В этом состоянии при включении транзистора в цепь по схеме рис. 9.15 сопротивление между истоком и стоком очень большое, так как p - n -переход стока находится под обратным смещением V_c . При положительном напряжении на затворе электроны из подложки будут перемещаться в ОПЗ на границе полупроводника p -типа с диэлектриком, как показано на рис. 9.14,б. Когда напряжение на затворе превысит так называемое пороговое напряжение $V_{\text{пор}}$, при котором концентрации электронов и дырок в поверхностном слое становятся одинаковыми, т.е. $n_s = p_s$, и обычно составляющее единицы вольт, то в слое ОПЗ концентрация электронов станет выше концентрации дырок, т.е. произойдет инверсия проводимости и образуется инверсионный (индуцированный) канал n -типа. Ток

стока, начиная с порогового напряжения на затворе, будет быстро увеличиваться с ростом величины V_3 (рис. 9.17,а) при неизменном напряжении на стоке. При увеличении напряжения стока ток стока увеличивается при неизменном напряжении на затворе и характеристики $I_c(V_3)$ смещаются вверх по оси токов.

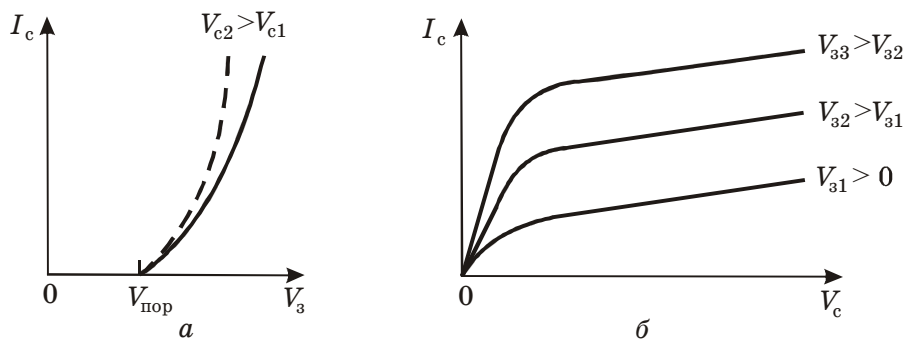


Рис. 9.17. Характеристики МДП-транзистора с индуцированным каналом: а — передачи, б — стоковая

Таким образом, МДП-транзистор с индуцированным каналом может работать только в режиме обогащения, что следует из его выходных характеристик (рис. 9.17,б). Обогащением называется такой режим, когда увеличение напряжения на затворе вызывает увеличение тока стока. Как следует из принципа работы транзистора, управление током стока, т.е. током в цепи нагрузки, от внешнего относительно мощного источника питания происходит при изменении напряжения на затворе. В связи с большим сопротивлением слоя диэлектрика, отделяющего затвор от канала, и малостью токов смещения через диэлектрик, мощность, необходимая для управления током стока и потребляемая от источника сигнала в цепи затвора, оказывается ничтожно малой. Поэтому полевой транзистор обеспечивает усиление как по мощности, так и по току и напряжению.

Усилительные свойства МДП-транзистора оцениваются так называемой крутизной характеристики S , которая представляет собой отношение изменения тока стока к изменению напряжения на затворе при коротком замыкании по переменному току на выходе транзистора с общим истоком (см. рис. 9.15):

$$S = \left. \frac{d I_c}{d V_{зи}} \right|_{V_c}$$

Проводящий канал может иметь электропроводность как n -, так и p -типа. Все полярности напряжений смещения, подаваемых на электроды полевых транзисторов с n - и p -каналом, противоположны. Полевые транзисторы, имеющие проводящие каналы с разными типами электропро-

водности, называют комплементарными, т.е. дополняющими друг друга по типу электропроводности проводящих каналов. Такие приборы оказываются очень удобными для построения ряда схем.

От биполярных полевые транзисторы отличаются принципом действия: в биполярном транзисторе управление выходным сигналом осуществляется входным током, а в полевом транзисторе — входным напряжением или электрическим полем (напряжением на затворе). Полевые МДП-транзисторы имеют значительно большие входные сопротивления, определяемые сопротивлением подзатворного диэлектрика, малый уровень шумов, высокую противорадиационную устойчивость, обладают слабой зависимостью параметров от температуры, высокой надежностью и удобством применения в интегральных схемах.

Недостатком МДП-транзисторов является возможность электрического пробоя тонкого слоя подзатворного диэлектрика под действием статических электрических зарядов, которые могут возникнуть на изолированном затворе.

Контрольные вопросы и задачи

1. (ШВПС). Что такое поверхностные состояния?

Ответы:

- 1) загрязнения на поверхности полупроводника;
- 2) дополнительные разрешенные энергетические уровни во всех зонах, расположенные на поверхности полупроводника;
- 3) энергетические уровни в объеме полупроводника, вносимые донорными или акцепторными примесными атомами в зоне проводимости или валентной зоне соответственно;
- 4) энергетические уровни, возникающие при ионизации атомов основного вещества.

2. (ГКБС). Какие поверхностные состояния называются быстрыми?

Ответы:

- 1) состояния, обусловленные структурными дефектами поверхности и находящиеся в хорошем контакте с объемом полупроводника, время их обмена носителями заряда с разрешенными зонами в объеме составляет около 10^{-7} с;
- 2) состояния, создаваемые на внешней поверхности окисла, покрывающего полупроводник, вследствие чего время обмена электронами не превышает 10^{-3} с;
- 3) состояния, возникающие при подаче измерительного сигнала высокой частоты на полупроводник;
- 4) состояния, возникающие при подаче положительного напряжения на затворе.

3. (ЛКМС). Какие поверхностные состояния называются медленными?

Ответы:

- 1) состояния, обусловленные структурными дефектами поверхности и имеющие малое (около 10^{-7} с) время обмена носителями заряда с разрешенными зонами;
- 2) состояния, возникающие на внешней поверхности окисла, покрывающего полупроводник вследствие адсорбции примесных атомов;
- 3) состояния, имеющие малую скорость рекомбинации;
- 4) состояния, увеличивающие скорость перехода носителей из валентной зоны в зону проводимости на поверхности.

4. (Т7АК). Как возникает обогащение приповерхностного слоя полупроводника?

Ответы:

- 1) при подаче напряжения электроны и дырки из объема притягиваются в поверхностный слой. Концентрация носителей заряда у поверхности возрастает;
- 2) обогащение возникает при небольших отрицательных смещениях на затворе $-V_{см}$ относительно полупроводника n -типа;
- 3) обогащение возникает при положительном напряжении на затворе $+V_{см}$ относительно полупроводника n -типа. При этом положительный заряд на затворе компенсируется подвижными электронами, которые подтягиваются из объема в приповерхностный слой полупроводника;
- 4) обогащение возникает при подаче переменного напряжения от генератора за счет лавинообразного увеличения концентрации подвижных носителей заряда в сильном электрическом поле.

5. (Я9ДП). Что такое обеднение приповерхностного слоя полупроводника?

Ответы:

- 1) уменьшение концентрации подвижных носителей заряда, возникающее в объеме полупроводника при подключении напряжения к структуре металл-диэлектрик-полупроводник;
- 2) уменьшение концентрации электронов в приповерхностном слое полупроводника при небольших отрицательных смещениях $-V_{см}$ на затворе относительно полупроводника n -типа;
- 3) уменьшение сопротивления поверхностного слоя полупроводника при подаче внешнего напряжения;
- 4) уменьшение концентрации доноров в полупроводнике n -типа.

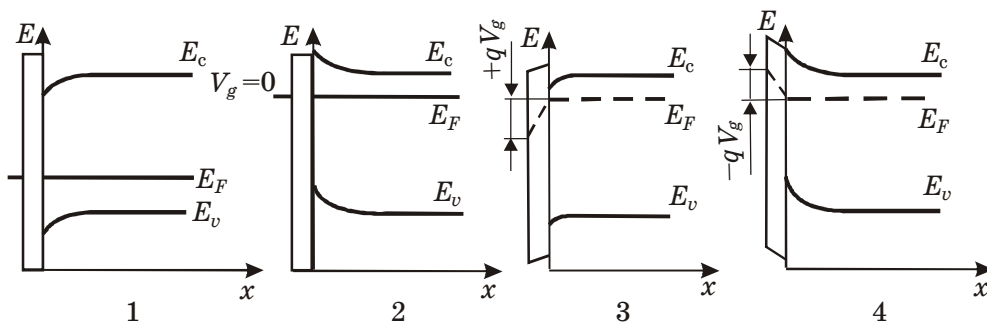
6. (Х9ИН). Что такое инверсия проводимости?

Ответы:

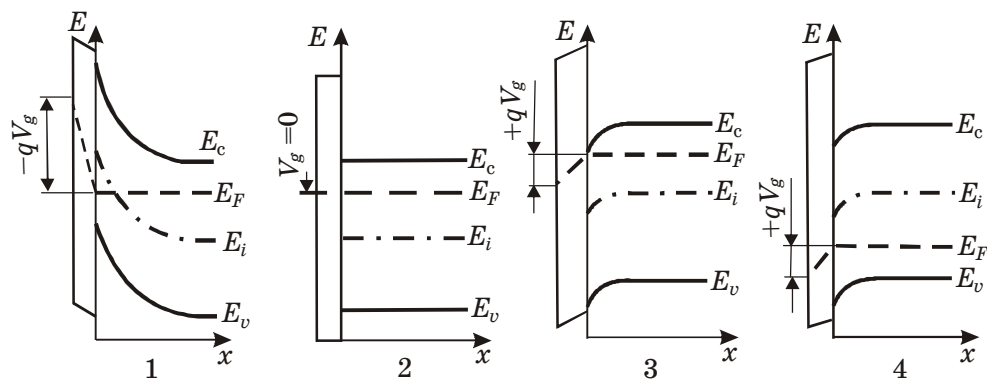
- 1) увеличение ионной проводимости полупроводника под действием внешнего напряжения;
- 2) уменьшение проводимости полупроводника под действием внешнего напряжения;
- 3) изменение типа проводимости;

4) изменение емкости МДП-структуры при изменении частоты внешнего напряжения.

7. (У6СД). На нижеприведенных рисунках укажите верное изображение энергетической зонной диаграммы структуры металл-диэлектрик-полупроводник n -типа проводимости, соответствующее режиму обогащения.



8. (Ю4ЭД). На рисунках, приведенных ниже, укажите верное изображение энергетической зонной диаграммы МДП-структуры, соответствующее режиму инверсии.



9. (Я9ЭП). Что называется эффектом поля?

Ответы:

- 1) изменение поверхностной проводимости полупроводника под действием поперечного электрического поля;
- 2) возникновение электрического поля при освещении полупроводника;
- 3) возникновение поперечной разности потенциалов при пропускании тока через полупроводниковый образец в продольном направлении;
- 4) возникновение разности потенциалов вдоль полупроводникового образца при наличии градиента температуры.

10. (И2ПБ). В чем отличие полевых и биполярных транзисторов?

Ответы:

1) в биполярных транзисторах имеются два $p-n$ -перехода, а в полевых $p-n$ -переходы отсутствуют;

2) ток в полевых транзисторах переносится носителями заряда одного знака, а в биполярных — носителями заряда разных знаков;

3) у биполярных транзисторов меньше изменяются параметры при изменении температуры окружающей среды, уровня радиации, ионизирующего излучения, чем у полевых транзисторов;

4) входное сопротивление большинства биполярных транзисторов превышает входное сопротивление полевых транзисторов.

11. (Г7ИК). Какие полевые транзисторы называются транзисторами с индуцированным каналом?

Ответы:

1) транзисторы, в которых сечение проводящего канала изменяется при изменении приложенного напряжения;

2) транзисторы, в которых проводящий канал возникает только при подаче на затвор напряжения определенной величины и полярности;

3) транзисторы, в которых канал индуцируется за счет изменения температуры полупроводника;

4) транзисторы, длина проводящего канала у которых изменяется при подаче напряжения.

12. (РТВК). Что такое полевой транзистор со встроенным каналом?

Ответы:

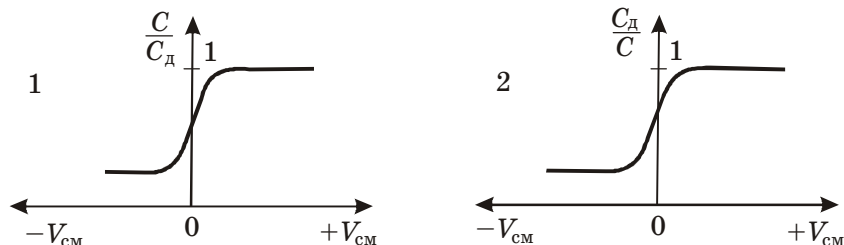
1) полевой транзистор, в котором у поверхности полупроводника под затвором существует инверсионный слой, соединяющий области стока и истока при нулевом напряжении на затворе;

2) полевой транзистор, у которого затвор от канала отделен запирающим слоем $p-n$ -перехода;

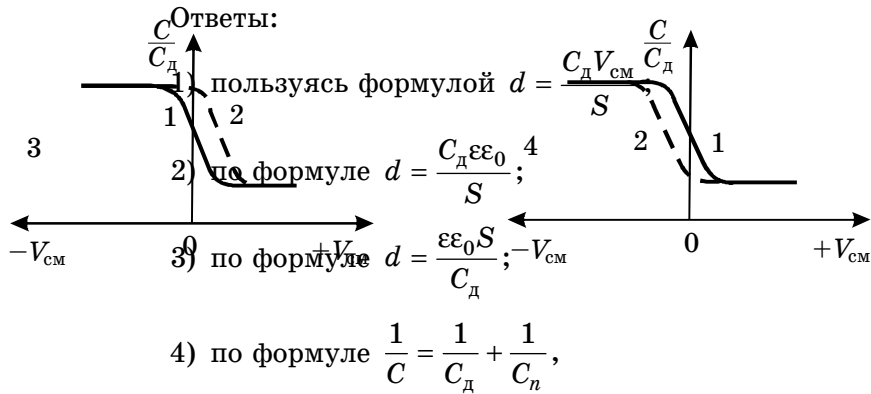
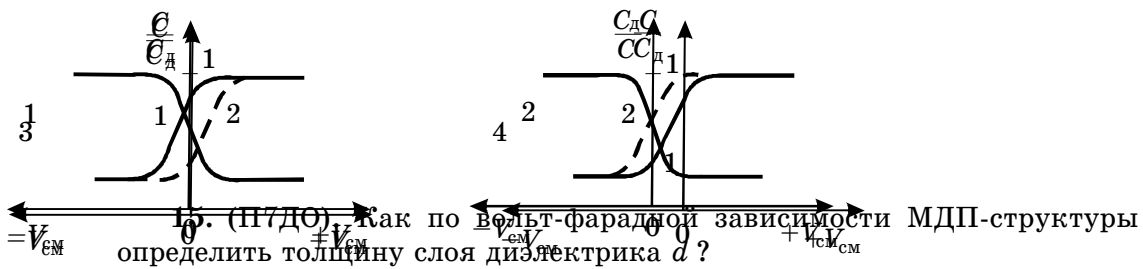
3) полевой транзистор, управление током стока которого осуществляется за счет инжекции носителей из истока;

4) полевой транзистор, у которого проводящий канал встраивается при подаче на затвор напряжения выше порогового значения.

13. (У2ВФ). На рисунке укажите верное изображение высокочастотной вольт-фарадной зависимости МДП-структуры, изготовленной на основе полупроводника проводимости n -типа. C_d — емкость диэлектрика. Зависимость измерена на высокой частоте.



14. (Н2П3). Какая из нижеприведенных зависимостей $C(V)$ верно отражает накопление положительного заряда в окисле и на поверхностных состояниях полупроводника n -типа проводимости?



где C_d — емкость диэлектрика; C_n — емкость полупроводника; ε — относительная диэлектрическая проницаемость диэлектрика; ε_0 — абсолютная диэлектрическая проницаемость вакуума.

16. (ФТСО). Как по вольт-фарадной зависимости МДП-структуры определить емкость диэлектрика?

Ответы:

- 1) по величине емкости в режиме обогащения;
- 2) по величине емкости в режиме обеднения;
- 3) по величине емкости в режиме инверсии;
- 4) по величине емкости при нулевом напряжении.

17. (ЖУТД). Определите толщину диэлектрического слоя структуры металл-диэлектрик-полупроводник, в которой используется диэлектрик с относительной диэлектрической проницаемостью, равной 9,8.

Электрическая постоянная составляет величину $8,85 \cdot 10^{-12}$ Ф/м, а диаметр полевого электрода равен 1 мм. На экспериментальной вольт-фарадной кривой для данной МДП-структуры высокочастотные емкости в областях инверсии, обогащения и при нулевом внешнем смещении ($V_g = 0$) равны соответственно 97, 520 и 480 пФ.

Ответ в микрометрах с точностью до двух знаков после запятой с указанием единицы длины, например 0,15 мкм, наберите после кода задания в виде ряда символов.

18. (ПМЦО). По высокочастотной вольт-фарадной зависимости МДП-структуры определите емкость слоя диэлектрика (минимальная емкость равна 150 пФ, максимальная емкость — 350 пФ, емкость при напряжении, равном нулю, — 180 пФ, площадь затвора составляет $7,85 \cdot 10^{-7}$ м², относительная диэлектрическая проницаемость — 3,5).

Ответ наберите в пикофарадах после кода задания.

19. (СОЕМ). Чему будет равна максимальная емкость МДП-структуры (см. задачу 18), если диэлектрик заменить вакуумом? Ответ дайте в пикофарадах.

20. (Г5ДВ). Рассчитайте коэффициент перекрытия емкости $k = C_{\max} / C_{\min}$, если в МДП-структуре (см. задачу 18) диэлектрик заменить вакуумом. Ответ укажите с точностью до двух значащих цифр после запятой и наберите его за кодом данной задачи.

Глава 10. Гальваномагнитные, термомагнитные и термоэлектрические явления в полупроводниках

§ 10.1. Эффект Холла

Гальваномагнитными и термоэлектрическими называют явления, связанные с влиянием магнитного поля или тепловых процессов на электрические свойства твердых тел, в том числе полупроводников. Основная причина гальваномагнитных явлений — искривления траекторий носителей заряда (электронов и дырок) под действием силы Лоренца (эффект Холла, магниторезистивный эффект).

Термоэлектрические явления обусловлены нарушением теплового равновесия в потоке носителей заряда (эффекты Зеебека, Пельтье, Томсона).

Американский физик Э. Холл в 1879 г. открыл один из важнейших гальваномагнитных эффектов — возникновение в проводнике с током, помещенном в магнитное поле, электрического поля, направленного перпендикулярно к направлениям тока и магнитного поля (эффект Холла). Рассмотрим прямоугольный образец полупроводника n -типа, по которому в направлении оси x пропускается электрический ток j_x . Магнитное поле действует в направлении оси z , т.е. вектор магнитной индукции B_z направлен перпендикулярно вектору скорости дрейфа v_x электронов (рис. 10.1). На дрейфующие электроны действует сила Лоренца

$$F_{\text{Л}} = -q[v_x B_z]. \quad (10.1)$$

Под действием силы Лоренца электроны, дрейфующие вдоль оси x , отклоняются на боковую грань 2 образца в соответствии с направлением вектора $F_{\text{Л}}$. Это направление определяется согласно правилу векторного произведения с учетом отрицательного знака заряда электрона в выражении (10.1). На рис. 10.1 направление вектора j_x соответствует техническому направлению тока (противоположно вектору скорости дрейфа электронов).

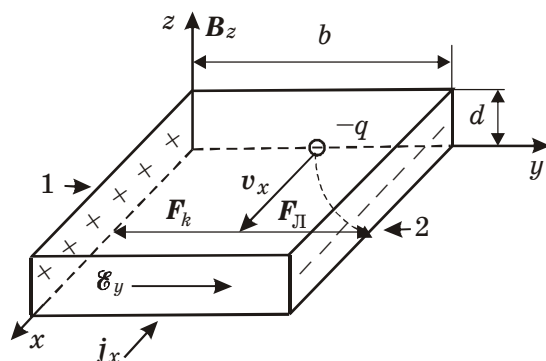


Рис. 10.1. К объяснению эффекта Холла
в донорном полупроводнике

Таким образом, на грани 2 образца будет создаваться избыточный отрицательный заряд отклоняемых магнитным полем электронов, на грани 1 будет накапливаться положительный заряд ионизированных донорных атомов из-за недостатка вблизи этой грани электронов. Вследствие такого разделения зарядов в образце возникает электрическое поле E_y , направленное вдоль оси y и называемое полем Холла. Поле Холла действует на электроны с силой eE_y направленной против силы Лоренца.

Процесс разделения зарядов вдоль оси y будет продолжаться до уравновешивания силы Лоренца и кулоновской силы со стороны поля Холла:

$$eE_y = ev_x B_z \quad (10.2)$$

Из равенства (10.2) следует, что в образце прекратится накопление зарядов на боковых гранях, когда напряженность поля Холла будет равна

$$E_y = [v_x B_z], \quad (10.3)$$

или с учетом направления векторов v_x и B_z равенство (10.3) в скалярной форме будет иметь вид

$$E_y = -v_x B_z. \quad (10.4)$$

Напряженности электрического поля E_y соответствует разность потенциалов (ЭДС Холла)

$$V_y = E_y b \quad (10.5)$$

где b — ширина образца (см. рис. 10.1).

Вычислим ЭДС Холла для полупроводника проводимости n -типа. Плотность электрического тока через образец в соответствии с выражением (3.9) позволяет определить дрейфовую скорость электронов v_x по экспериментально измеренной величине полного тока I :

$$I = qn_0 b d v_x \quad (10.6)$$

где d — толщина образца.

Тогда из выражений (10.4)–(10.6) можно получить выражение для ЭДС Холла:

$$V_y = \frac{1}{qn_0} \frac{IB_z}{d}. \quad (10.7)$$

Выражение (10.7) обычно записывается в виде

$$V_y = +R_X \frac{IB_z}{d}, \quad (10.8)$$

где

$$R_X = -\frac{1}{qn_0} \quad (10.9)$$

есть постоянная величина для данного полупроводника при неизменных внешних условиях, которая называется постоянной Холла.

При выводе выражения (10.9) мы не учитывали распределение носителей по скоростям. Поэтому выражение (10.9) справедливо для металлов и вырожденных полупроводников n -типа, электропроводность которых обусловлена носителями, обладающими практически одной и той же энергией.

Если учесть распределение носителей заряда по скоростям в невырожденных полупроводниках, то вместо (10.9) получим

$$R_X = -\frac{A}{qn_0}, \quad (10.10)$$

где A — коэффициент (Холл-фактор), учитывающий действующий в полупроводнике механизм рассеяния носителей заряда. При этом для атомных полупроводников:

$A = 1,17$ — при рассеянии на фононах,

$A = 1,93$ — при рассеянии на ионах примесей,

$A = 1$ — при рассеянии на нейтральных атомах, а также при всех видах рассеяния в сильных магнитных полях.

§ 10.2. Использование эффекта Холла для определения основных параметров полупроводников

Нетрудно видеть практическую пользу измерений ЭДС Холла. Как следует из (10.10), концентрация основных носителей в полупроводнике n -типа определяется по формуле

$$n_0 = -\frac{A}{qR_X}. \quad (10.11)$$

Измерив ЭДС Холла, из выражения (10.8) можно найти постоянную Холла

$$R_X = \frac{V_y d}{IB_z} \quad (10.12)$$

и далее концентрацию электронов в полупроводнике n -типа в соответствии с (10.11).

При одновременном измерении постоянной Холла и проводимости полупроводника можно, если известен Холл-фактор, определить подвижность носителей

$$\mu_n = \frac{\sigma_0}{qn_0} = -\frac{\sigma_0 R_X}{A}. \quad (10.13)$$

Проведя такие измерения при различных температурах, можно рассчитать температурные зависимости концентрации и подвижности носителей, определить энергию ионизации примеси и ширину запрещенной зоны полупроводника.

И наконец, по полярности ЭДС Холла можно определить тип проводимости полупроводника. Физическую природу эффекта Холла в полупроводнике *p*-типа объясним с помощью рис. 10.2. Вектор дрейфовой скорости дырок совпадает с техническим направлением тока. В результате сила Лоренца отклоняет дырки, движущиеся вдоль оси *x*, на грань 2. Таким образом, как следует из рис. 10.1 и 10.2, при одинаковой ориентации и направлении тока в образцах разного типа проводимости полярность ЭДС Холла различается, поскольку электроны и дырки при этом отклоняются в одну сторону. Поэтому в собственном полупроводнике величина ЭДС Холла значительно меньше, чем в примесных образцах, так как накопление противоположных зарядов у одной грани образца в одинаковых количествах приводит к их взаимной компенсации. Но вследствие более высокой подвижности электронов по сравнению с подвижностью дырок (см. § 3.3) грань 2 образца (рис. 10.2) будет иметь небольшой отрицательный потенциал. Величина постоянной Холла для полупроводников, содержащих носители заряда обоих знаков (электроны и дырки), определяется выражением

Рис. 10.2. К объяснению эффекта Холла в дырочном полупроводнике

димости полярность ЭДС Холла различается, поскольку электроны и дырки при этом отклоняются в одну сторону. Поэтому в собственном полупроводнике величина ЭДС Холла значительно меньше, чем в примесных образцах, так как накопление противоположных зарядов у одной грани образца в одинаковых количествах приводит к их взаимной компенсации. Но вследствие более высокой подвижности электронов по сравнению с подвижностью дырок (см. § 3.3) грань 2 образца (рис. 10.2) будет иметь небольшой отрицательный потенциал. Величина постоянной Холла для полупроводников, содержащих носители заряда обоих знаков (электроны и дырки), определяется выражением

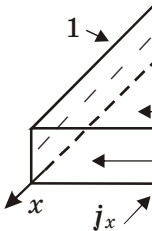
$$R_x = \frac{A}{q} \frac{\mu_p^2 p_0 - \mu_n^2 n_0}{(\mu_p p_0 - \mu_n n_0)^2}. \quad (10.14)$$

Очевидно, что в случае собственного полупроводника $n_0 = p_0 = n_i$ и выражение (10.14) примет вид

$$R_x = \frac{A}{q n_i} \frac{\mu_p - \mu_n}{\mu_p + \mu_n}. \quad (10.15)$$

Таким образом, эффект Холла может использоваться в полупроводниковом материаловедении как для исследования новых материалов, так и для контроля параметров полупроводников, выпускаемых промышленностью.

Эффект Холла широко применяется и в других областях. Например, на его основе создаются полупроводниковые датчики магнитного поля, устройства для бесконтактного измерения силы тока по магнитным полям рассеяния, устройства для перемножения двух сигналов и т.д. Для изготовления высокочувствительных датчиков на эффекте Холла обычно используются полупроводники с высокой подвижностью носителей заряда: германий, кремний, антимонид индия, арсенид индия и др.



§ 10.3. Магниторезистивный эффект. Термомагнитные явления

В полупроводниковом образце при протекании тока в направлении, перпендикулярном магнитному полю, возникает ЭДС Холла. Сила, действующая на носители заряда со стороны поля Холла, и сила Лоренца уравновешивают друг друга только для носителей, скорость направленного перемещения которых равна скорости дрейфа.

Более быстрые носители заряда ($v > v_{др}$) будут отклоняться к одной боковой грани образца, а более медленные ($v < v_{др}$) — к другой (рис. 10.3). Искривление траекторий этих носителей должно, очевидно, сопровождаться уменьшением длины свободного пробега в направлении тока. В итоге наблюдается уменьшение дрейфовой подвижности носителей заряда и соответственно проводимости в полупроводнике, зависящее от напряженности магнитного поля (магниторезистивный эффект). Величина, на которую в поперечном магнитном поле увеличивается сопротивление полупроводника, называется магнитосопротивлением.

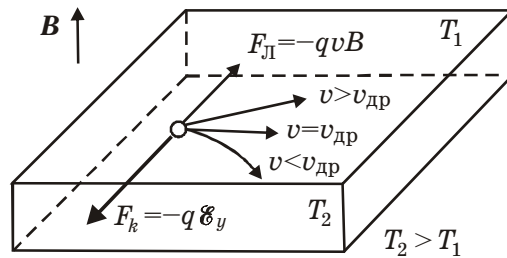


Рис. 10.3. Возникновение магнитосопротивления и эффекта Эттингсгаузена в полупроводнике

Поскольку в полупроводниках при протекании тока в поперечном магнитном поле на одну из боковых граней образца относительно направления тока (см. рис. 10.3) попадают более горячие носители, а на другую — более холодные, то между этими гранями должна возникать разность температур (эффект Эттингсгаузена).

Направленное перемещение носителей заряда в полупроводнике может быть вызвано не электрическим полем, а градиентом температуры (диффузионный ток). При этом в поперечном к направлению диффузионного тока магнитном поле будут наблюдаться те же явления, что рассматривались выше (возникновение поперечной разности потенциалов, поперечного и продольного градиентов температур и пр.). Эта группа эффектов имеет общее название: термомагнитные явления.

§ 10.4. Эффект Зеебека

Немецкий физик Т. Зеебек в 1821 г. открыл явление термоэлектричества, когда обнаружил, что в замкнутой цепи, состоящей из двух различных проводников 1 и 2 (рис. 10.4), один контакт которых находится при повышенной температуре T_{Γ} , другой — при пониженной T_{\times} , возникает термоэлектрический ток. Если цепь разомкнуть, то на ее концах наблюдается термо-ЭДС V_{T} :

$$V_{\text{T}} = \alpha(T_{\Gamma} - T_{\times}). \quad (10.16)$$

В относительно узком диапазоне температур величина термо-ЭДС пропорциональна разности температур с коэффициентом пропорциональности

$$\alpha = \frac{dV_{\text{T}}}{dT}, \quad (10.17)$$

который называется дифференциальной или удельной термо-ЭДС. Он зависит от природы контактирующих проводников и от температуры.

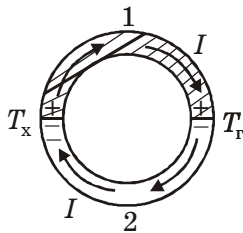


Рис. 10.4.
К эффекту Зеебека

Термо-ЭДС в цепи из различных материалов возникает в результате трех процессов и состоит соответственно из трех слагаемых:

1) объемной термо-ЭДС, которая является результатом диффузионного потока носителей заряда, порожденного градиентом температуры;

2) контактной термо-ЭДС, обусловленной различным положением уровня Ферми в материалах при различных температурах;

3) термо-ЭДС, связанной с увеличением носителей заряда фононами, которые движутся направленным потоком под влиянием градиента температур.

Рассмотрим эти процессы подробнее.

1. Объемная термо-ЭДС. На рис. 10.5 показана контактная цепь из невырожденного полупроводника n -типа и металла (T_{\times} , T_{Γ} — температуры холодного и горячего спаев соответственно). Концентрация и средняя скорость свободных носителей заряда на горячем конце образца $n_{0\Gamma}$ будут больше, чем на холодном $n_{0\times}$. Градиент концентрации и скорость электронов вызывает их диффузионный поток от горячего спаива к холодному, что приводит к увеличению концентрации электронов на холодном конце полупроводника, т.е. на холодном конце возникает отрицательный электрический заряд. На горячем конце концен-

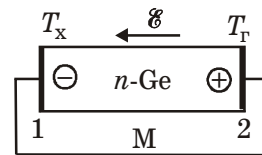


Рис. 10.5. Возникновение объемной термо-ЭДС в полупроводнике n -типа при создании в нем градиента температуры

трация электронов уменьшается на такую же величину, т.е. появляется положительный заряд нескомпенсированных донорных ионов. Накопление отрицательного и положительного зарядов на концах полупроводника продолжается до тех пор, пока возникшее электрическое поле напряженностью не вызовет равный диффузионному, но встречно направленный дрейфовый поток электронов. Таким образом, стационарное состояние наступает при выполнении условия

$$j_{\text{диф}} + j_{\text{др}} = 0. \quad (10.18)$$

Возникшая между концами полупроводника разность потенциалов носит название объемной термо-ЭДС.

Аналогичные процессы протекают в дырочном полупроводнике при наличии в нем градиента температуры. Но в дырочном полупроводнике холодный конец будет заряжаться положительно за счет диффузии положительно заряженных дырок. Таким образом, величина и знак объемной термо-ЭДС в полупроводниках определяются зависимостью концентрации носителей от температуры и знаком заряда основных носителей. В общем случае в полупроводниках с двумя типами носителей дифференциальная термо-ЭДС будет равна

$$(10.19)$$

где α_p и α_n — значения дифференциальной термо-ЭДС для дырочного и электронного полупроводников соответственно.
 $\alpha = \frac{\alpha_p p + \alpha_n n}{p + n}$ — контактная термо-ЭДС. На контакте двух материалов, имеющих разное положение уровня Ферми E_{F1} и E_{F2} , всегда возникает контактная разность потенциалов

$$V_k = \frac{E_{F1} - E_{F2}}{q}. \quad (10.20)$$

Положение уровня Ферми в разных материалах зависит от температуры различным образом. При отсутствии градиента температуры между контактами (см. рис. 10.4) векторы контактного электрического поля в обоих контактах равны по величине и направлены в сторону материала с большей работой выхода, т.е. навстречу друг другу. При этом если цепь металла (см. рис. 10.5) разомкнуть, то на концах разомкнутой цепи разность потенциалов не возникает.

Ситуация меняется, если температуру одного из контактов изменить. На этом контакте изменится положение уровней Ферми E_{F1} и E_{F2} и, следовательно, изменится значение контактной разности потенциалов. Очевидно, что на концах разомкнутой цепи в этом случае возникнет контактная термо-ЭДС, равная

$$V_k = V_{k1} - V_{k2}, \quad (10.21)$$

где V_{k1} , V_{k2} — контактная разность потенциалов на первом и втором контактах цепи (см. рис. 10.4).

3. Термо-ЭДС, обусловленная увлечением носителей заряда фононами. Концентрация фононов в твердых телах при низких температурах возрастает с температурой пропорционально T^3 , а при высоких температурах — пропорционально T . Следовательно, в образце с градиентом температуры будет наблюдаться градиент концентрации фононов и поэтому возникнет направленный от горячего конца к холодному диффузионный поток фононов. В результате рассеяния фононов на электронах и передачи электронам импульса от фононов появляется направленное движение электронов в сторону более холодной части образца, где в итоге скапливается отрицательный объемный заряд. Разделение зарядов создает на концах образца термо-ЭДС, препятствующую дальнейшему переносу электронов вследствие их увлечения фононами.

Этот механизм создания термо-ЭДС становится заметным по сравнению с другими только при низких температурах, когда процесс рассеяния фононов на электронах доминирует над процессом фонон-фононного рассеяния.

В полупроводниках концентрация и средняя энергия носителей заряда сильно зависят от температуры. У металлов концентрация электронов и их средняя энергия от температуры практически не зависят. Поэтому в металлических парах значения дифференциальной термо-ЭДС на несколько порядков меньше, чем у полупроводниковых пар или пар полупроводник-металл.

Эффект возникновения термо-ЭДС (эффект Зеебека) широко используется в технике, особенно в случае новых полупроводниковых материалов, обладающих высоким коэффициентом термо-ЭДС — порядка сотен и тысяч микровольт на градус (например, полупроводниковое тройное соединение ZnSnAs_2). Такие материалы применяются в измерительных термопарах и в термогенераторах для непосредственного преобразования тепловой энергии в электрическую.

§ 10.5. Эффект Пельтье

В 1834 г. французский физик Ж. Пельтье обнаружил, что в местах спаев двух различных металлов в зависимости от направления электрического тока происходит выделение или поглощение тепла. Теплота Пельтье $Q_{\text{П}}$ пропорциональна величине постоянного тока I , протекающего через спай в течение времени t :

$$Q_{\text{П}} = \pm \Pi I t, \quad (10.22)$$

где коэффициент пропорциональности Π называется коэффициентом Пельтье.

Теплота Пельтье выделяется или поглощается дополнительно к джоулевой теплоте, которая пропорциональна квадрату тока I^2 и только выде-

ляется. Коэффициент Пельтье связан с дифференциальной термо-ЭДС соотношением

$$\Pi = \alpha T . \quad (10.23)$$

Теплота Пельтье выделяется на спае вследствие передачи части энергии носителей заряда кристаллической решетке при переходе их из одного материала в другой. При обратном направлении тока решетка передает часть своей энергии носителям заряда и охлаждается. Физический смысл этого процесса можно понять из энергетической диаграммы контакта металл-полупроводник, которая в простейшем виде с учетом уже известных диаграмм (см. главы 2 и 4) представлена на рис. 10.6. Более полная теория контакта металл-полупроводник рассмотрена в главе 7.

В условиях равновесия уровни Ферми в металле и в полупроводнике занимают одно положение E_F . Во внешнем слабом электрическом поле заданного направления (рис. 10.6), когда сдвигом уровней Ферми и деформацией краев энергетических зон можно пренебречь, переход электрона из полупроводника в металл сопровождается уменьшением его энергии на величину ΔE_k . Эта энергия равна теплоте Пельтье:

$$\Delta E_k = Q_{\Pi} = (E_c - E_F) + E_T, \quad (10.24)$$

где E_T — средняя тепловая энергия свободных электронов в зоне проводимости или, другими словами, кинетическая энергия электронов проводимости.

Уменьшение энергии электрона при его переходе из полупроводника в металл связано с ее поглощением атомами решетки металла в результате процесса рассеяния. Так как электроны приходят в тепловое равновесие в результате небольшого числа столкновений в непосредственной близости от контакта, то практически вся теплота Пельтье выделяется на самом контакте, который в этом случае нагревается.

При противоположном направлении внешнего электрического поля (положительный потенциал на полупроводнике) электроны металла переходят в полупроводник, только преодолев энергетический барьер $E_c - E_F$. Энергию для таких переходов электроны могут получить от кристаллической решетки, вследствие чего металл в области контакта охлаждается.

Явление охлаждения контакта при прохождении электрического тока из металла в полупроводник позволяет создавать термоэлектрические холодильные устройства, например устройства для охлаждения бортовой радиоэлектронной аппаратуры, различные электрические термостаты

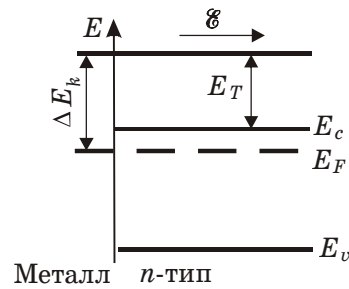


Рис. 10.6. Зонная диаграмма контакта металла и полупроводника с равными работами выхода (случай плоских зон)

и т.д. Изменяя направление тока, холодильник можно превращать в нагреватель и наоборот.

Наиболее эффективными материалами для создания термоэлектрических холодильников и нагревателей являются полупроводники с высокими значениями коэффициента Пельтье (удельной термо-ЭДС) и удельной электропроводности. Эти требования обусловлены стремлением сделать теплоту Пельтье преобладающей над джоулевым нагревом.

§ 10.6. Эффект Томсона

В 1856 г. английский физик У. Томсон (Кельвин) открыл термоэлектрический эффект, заключающийся в следующем: если вдоль проводника, по которому течет ток, существует перепад температур, то кроме джоулевой теплоты в объеме проводника в зависимости от направления тока выделяется или поглощается еще некоторое количество тепла Q_T (теплота Томсона). Величина Q_T пропорциональна току I , времени t , перепаду температур $T - T_0$ и коэффициенту τ_T (коэффициент Томсона), зависящему от природы материала:

$$Q_T = \tau_T (T - T_0) It. \quad (10.25)$$

Физическую природу эффекта Томсона можно понять, рассматривая перенос тепла свободными носителями во внешнем электрическом поле. Если вдоль проводника имеется градиент температур, то электроны, переходя под действием внешнего электрического поля от горячего конца к холодному, отдают избыточную энергию окружающим атомам (выделяется теплота). При обратном направлении тока электроны, переходя из более холодного участка в горячий, ускоряются полем термо-ЭДС и пополняют свою энергию за счет энергии окружающих атомов (поглощение тепла).

Поскольку все термоэлектрические явления связаны между собой и обусловлены зависимостью функции распределения носителей заряда от температуры и вида материала, то все они проявляются одновременно. При этом между коэффициентами термо-ЭДС, Пельтье и Томсона существует связь, устанавливаемая на основе законов термодинамики:

$$\frac{d\Pi}{dT} = \alpha + \tau_{T1} - \tau_{T2}, \quad (10.26)$$

где τ_{T1} и τ_{T2} — коэффициенты Томсона для двух контактирующих проводников.

Контрольные вопросы и задачи

1. (ЖРЭХ). Почему ЭДС Холла обычно мала в собственных полупроводниках?

Ответы:

1) потому что в них концентрация носителей заряда меньше, чем в примесных полупроводниках;

2) в собственных полупроводниках концентрация носителей заряда выше, чем в примесных, и это приводит к малой величине постоянной Холла;

3) в собственных полупроводниках имеются в равных количествах носители заряда противоположных знаков, тогда как их подвижности обычно отличаются. Поэтому постоянная Холла оказывается малой величиной;

4) при протекании тока в собственном полупроводнике электроны и дырки отклоняются внешним магнитным полем в одну и ту же сторону, что приводит к взаимной компенсации их зарядов, вследствие чего ЭДС Холла оказывается малой.

2. (ЛТПХ). У двух различных собственных полупроводников постоянные Холла на два порядка отличаются по величине. Укажите возможные причины этого.

Ответы:

1) это связано с различием Холл-факторов для выбранных полупроводников;

2) это обусловлено тем, что в одном полупроводнике подвижности электронов и дырок близки, а в другом — сильно отличаются;

3) это возможно при одинаковых значениях подвижностей электронов и дырок в полупроводниках, если при этом концентрации собственных носителей заряда в них отличаются на два порядка;

4) это может произойти, если выбранные полупроводники имеют запрещенные зоны различной ширины, а подвижности электронов и дырок в них не равны.

3. (У1ДХ). Укажите, что нужно сделать для повышения чувствительности холловского датчика магнитного поля?

Ответы:

1) выбрать слаболегированный донорный полупроводник с большой шириной запрещенной зоны, что дает низкую концентрацию носителей заряда;

2) увеличить созданную в образце полупроводника в направлении оси x (см. рис. 10.1) напряженность электрического поля, определяющую величину тока, протекающего через полупроводник;

3) уменьшить толщину образца (размер d на рис. 10.1);

4) выбрать примесный полупроводник с высоким значением подвижности носителей заряда;

5) увеличить ширину образца (размер b на рис. 10.1).

4. (ДЗТЭ). Укажите причину малого значения дифференциальной термо-ЭДС, наблюдаемой при эффекте Зеебека в замкнутой цепи из двух металлов.

Ответы:

1) положение уровня Ферми в металлах слабо зависит от температуры;

2) среднее значение энергии электронов в зонах проводимости металлов слабо зависит от температуры;

3) концентрация электронов в зоне проводимости металла практически не зависит от температуры;

4) эффект фононного увлечения электронов не проявляется в металлах.

5. (РГТК). Определите единицы измерения в СИ термоэлектрических коэффициентов Зеебека (α), Пельтье (Π) и Томсона (τ_T) и укажите правильный ответ из нижеприведенных:

1)

2) α [В], Π [В], τ_T $\left[\frac{\text{В}}{\text{К}} \right]$;

3)

4)

5) α [В], Π $\left[\frac{\text{В}}{\text{К}} \right]$, τ_T $\left[\frac{\text{В}}{\text{К}} \right]$.

6. (Я1КП). Укажите пары материалов, для которых следует ожидать более высокие коэффициенты Пельтье, учитывая, что между коэффициентами Пельтье и Зеебека существует связь, определяемая формулой

где T — средняя температура пары материалов.

Ответы:

1) металл-металл;

2) металл-диэлектрик;

3) металл-собственный полупроводник;

4) диэлектрик-диэлектрик;

5) дырочный полупроводник-электронный полупроводник.

$$\alpha \left[\frac{\text{В}}{\text{К}} \right], \frac{\Pi}{T} \left[\frac{\text{В}}{\text{К}} \right]$$

7. (ЛДТЭ). В каком случае электрический ток может выделять количество тепла, превосходящее выделяемое по закону Джоуля — Ленца? Выберите правильные ответы:

1) если при пропускании тока через контакт двух проводников электроны будут переходить с более высоких энергетических уровней на более низкие;

2) если при пропускании электрического тока через контакт двух проводников электроны с низких энергетических уровней переходят на более высокие;

3) если при пропускании тока через полупроводник с градиентом температуры средняя энергия носителей заряда по направлению тока будет понижаться;

4) если при пропускании электрического тока в направлении начального градиента температуры образца средняя энергия носителей заряда повышается.

8. Образец германия с длиной $l = 2,5$ мм и с поперечными размерами $b = 0,25$ мм, $d = 50$ мкм (см. рис. 10.1), с продольным током $I = 2$ мА под действием внешнего напряжения $V = 85$ мВ помещается в магнитное поле с индукцией $B_z = 0,5$ Тл перпендикулярно его силовым линиям, вследствие чего возникает холловская разность потенциалов $V_x = -1,25$ мВ.

Определите:

1) (Б4ТП) тип проводимости полупроводника (ответ наберите после кода вопроса в виде одного слова из указанных ниже: собственный, электронный, дырочный);

2) (АУКН) концентрацию основных носителей заряда (ответ, например $6 \cdot 10^{21} \text{ м}^{-3}$, наберите в виде последовательности символов: 61021м-3);

3) (Ю5ПХ) коэффициент Холла (ответ, например $3,11 \cdot 10^{-8} \text{ м}^3/\text{Кл}$, наберите в виде последовательности символов: 3,1110-8м3/Кл);

4) (К2МЮ) подвижность основных носителей заряда (ответ, например $3,05 \text{ м}^2/(\text{В} \cdot \text{с})$, наберите в виде последовательности символов: 3,05м2/В·с).

9. (ЦТЭХ). Имеется образец германия с толщиной $0,1$ см и концентрацией акцепторной примеси 10^{17} см^{-3} . Вдоль образца протекает ток $0,6$ А. Перпендикулярно направлению тока действует магнитное поле с индукцией $0,5$ Тл. Вычислите ЭДС Холла между контактами на узких боковых гранях образца. Ответ наберите в милливольтгах (мВ) с точностью до десятых долей.

Глава 11. Фотоэлектрические явления в полупроводниках

§ 11.1. Спектры отражения и поглощения

Фотоэлектрическими явлениями называются изменения электрических свойств твердого тела (электропроводности, диэлектрической проницаемости, появление фото-ЭДС и пр.) под действием электромагнитного излучения. Эти явления обычно обусловлены внутренним фотоэффектом, т.е. перераспределением электронов по энергетическим состояниям в полупроводнике или диэлектрике, происходящем при поглощении света. Свет, или оптическое излучение, охватывает очень широкую область длин волн и включает в себя ультрафиолетовый ($\lambda \approx 0,01-0,38$ мкм), видимый ($\lambda \approx 0,38-0,77$ мкм) и инфракрасный ($\lambda \approx 0,77-1000$ мкм) диапазоны частот.

Рассмотрим основные процессы, происходящие в полупроводниках при воздействии света, и принципы работы некоторых фотоэлектрических приборов.

Оптическое излучение характеризуется интенсивностью J_0 . Интенсивностью излучения называется поток энергии излучения, проходящий за единицу времени через единичную площадку в направлении нормали к ней.

При облучении твердых тел световым потоком происходит как отражение световой энергии от поверхности, так и поглощение ее некоторой части телом.

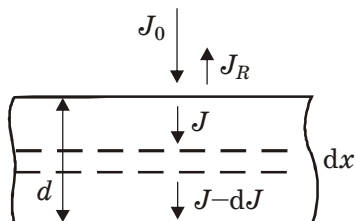


Рис. 11.1. Прохождение оптического излучения через пластину полупроводника

Пусть пучок монохроматического света интенсивностью J_0 падает на поверхность полупроводника (рис. 11.1). Тогда доля световой энергии, отраженной поверхностью, будет равна величине R , которая называется коэффициентом отражения:

$$R = \frac{J_R}{J_0}. \quad (11.1)$$

Зависимость коэффициента отражения от частоты $R(\omega)$ или от длины волны $R(\lambda)$ называется спектром отражения. Вид спектра отражения зависит от материала и состояния отражающей поверхности.

Таким образом, через поверхность внутрь твердого тела проникает доля излучения

$$J = (1 - R)J_0. \quad (11.2)$$

При прохождении через полупроводник интенсивность излучения уменьшается вследствие поглощения. Пусть на слой dx падает излучение интенсивностью J , а из слоя выходит излучение $J - dJ$. Количество поглощенной в слое световой энергии пропорционально количеству падающей энергии и толщине слоя dx :

$$dJ = -\alpha J dx. \quad (11.3)$$

Коэффициент пропорциональности α , численно равный количеству поглощенной энергии из пучка единичной интенсивности в слое единичной толщины, носит название коэффициента поглощения, единица измерения которого в СИ равна обратной длине (m^{-1}).

Знак минус в выражении (11.3) показывает, что эта часть энергии поглощается. Проинтегрировав уравнение (11.3)

$$(11.4)$$

получаем закон Бугера — Ламберта

$$J = J_0(1 - R)e^{-\alpha d}. \quad (11.5)$$

Величина α является характеристикой поглощающей среды и зависит от длины волны излучения. Зависимость коэффициента поглощения от частоты $\alpha(\omega)$ или длины волны $\alpha(\lambda)$ называется спектром поглощения.

Поскольку в процессах поглощения, отражения и излучения световая энергия может изменяться лишь на величину, кратную энергии фотона $h\nu$, то поток фотонов $J_0 - RJ_0$ в таких взаимодействиях удобно рассматривать как поток фотонов. Тогда коэффициент поглощения есть вероятность поглощения фотона на единице длины:

$$(11.6)$$

где $\sigma(\lambda)$ — эффективное сечение поглощения фотона отдельным центром; N — концентрация поглощающих центров; $l_\phi(\lambda)$ — средняя длина пробега фотона в среде до его поглощения.

Если в полупроводнике имеется несколько типов поглощающих центров, то полный спектр поглощения складывается из спектров поглощения отдельных центров. В качестве центров поглощения в полупроводниках обычно выступают свободные и связанные электроны, а также атомы решетки.

Тогда

$$(11.7)$$

где i — порядковый номер типа поглощающих центров.

В процессе поглощения фотон исчезает, его энергия и импульс переходят к поглощающему центру. Очевидно, что при этом должны выполняться фундаментальные законы сохранения энергии и импульса, которые для случая взаимодействия двух частиц (квазичастиц) имеют вид:

$$(11.8)$$

$$(11.9)$$

где E' , p' — энергия и квазиимпульс центра после взаимодействия с фотоном, имеющим энергию E и квазиимпульс p . E и p — энергия и квазиимпульс центра до взаимодействия с фотоном.

Исходя из равенства (11.8), электрон полупроводника может поглотить фотон с энергией E только при условии, что имеется свободное разрешенное состояние в зонной энергетической структуре с энергией E' , где E — энергия электрона до взаимодействия с фотоном.

Поскольку квазиимпульс фотона p на несколько порядков меньше квазиимпульса электрона, то условие (11.9) переходит в равенство

$$p' \approx p, \quad (11.10)$$

или с учетом известного соотношения (1.2) получаем

$$(11.11)$$

где k' — волновой вектор электрона после поглощения фотона; k — волновой вектор электрона до взаимодействия с фотоном.

То есть при поглощении фотона электрон должен осуществлять переходы без изменения волнового числа (вертикальные переходы на зонной диаграмме $E(k)$). Такие переходы называются прямыми.

Однако при взаимодействии излучения с полупроводником возможны процессы, в которых участвуют больше двух частиц, например фотон, электрон и фонон (или несколько фононов). При этом оказываются возможными не прямые переходы электрона при поглощении фотона, т.е. переходы с изменением волнового числа, поскольку волновой вектор фонона по порядку величины сравним с волновым вектором электрона.

В полупроводниках в зависимости от природы поглощающих центров различают пять основных механизмов оптического поглощения: собственное, экситонное, примесное, поглощение свободными носителями заряда и решеточное.

На рис. 11.2 представлены типичный спектр поглощения полупроводника и схема энергетических переходов электронов, обуславливающих этот спектр поглощения.

Рассмотрим основные механизмы поглощения оптического излучения в полупроводниках подробнее.

1. Собственное поглощение наблюдается в области коротких длин волн, когда энергия фотона больше или равна ширине запрещенной зоны полупроводника. При этом фотоны взаимодействуют с электронами валентной зоны, переводя их в зону проводимости (переход 1 на рис. 11.2,б). Минимальная энергия, которая требуется для такого перехода, равна ширине запрещенной зоны. Поэтому собственное поглощение начинается при

$$\frac{E_p}{\hbar k} = \frac{E_g}{\hbar k} + \frac{2\pi\hbar c}{\lambda}$$

(11.12)

и с увеличением энергии фотона коэффициент поглощения резко увеличивается и достигает $10^4-10^6 \text{ см}^{-1}$. Столь большое значение коэффициента поглощения обусловлено высокой концентрацией поглощающих центров — валентных электронов (см. выражение (11.6)).

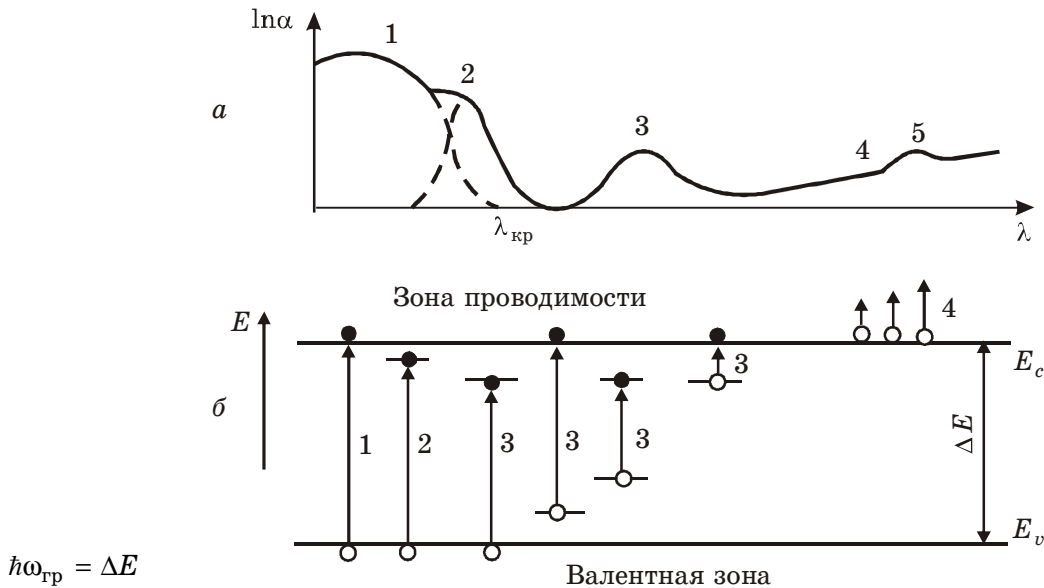


Рис. 11.2. Основные полосы поглощения в полупроводнике (а) и соответствующие им переходы на зонной диаграмме (б)

В области собственного поглощения глубина проникновения фотонов в полупроводник составляет всего $0,01-1,0 \text{ мкм}$, т.е. свет поглощается в тонком приповерхностном слое полупроводника.

Длина волны $\lambda_{кр}$, определяемая из условия (11.12), носит название «красной» границы собственного (фундаментального) поглощения (рис. 11.2,а).

Рост концентрации свободных электронов и дырок при собственном поглощении фотонов приводит к увеличению проводимости полупроводника.

2. Экситонное поглощение (переход 2 на рис. 11.2,б) наблюдается в области длин волн, чуть больших «красной» границы собственного поглощения. При поглощении фотонов с $h\nu \leq h\nu_{гр}$ в полупроводнике могут возникать возбужденные состояния атомов, при этом валентный электрон переходит не в зону проводимости, а на уровень возбуждения, лежащий в запрещенной зоне вблизи дна зоны проводимости. Электрон, перешедший на уровень возбуждения, не теряет связи со своим атомом, т.е. при

поглощении фотона образуется связанная электронейтральная пара электрон-дырка, получившая название экситона. Созданные экситоны могут перемещаться в кристалле, при этом связанная пара электрон-дырка переходит от одного атома к другому. Спустя время жизни, электронно-дырочная пара может аннигилировать, передав запасенную при создании экситона энергию дефектам, тепловым колебаниям кристаллической решетки или излучив фотон. Возможен и другой вариант событий. Электрон экситона, поглотив фотон или фонон, переходит в зону проводимости, т.е. происходит распад экситона на свободные электрон и дырку. Такой распад будет сопровождаться изменением проводимости полупроводника.

Экситонные полосы поглощения лежат вблизи края собственного поглощения, и при комнатных температурах выделить эти полосы в чистом виде не удастся из-за расширения экситонных полос, связанного с тепловыми колебаниями решетки, и их перекрытия с краем фундаментального поглощения. Величина коэффициента поглощения в экситонной полосе оказывается сравнима с величиной коэффициента собственного поглощения (рис. 11.2,а, полосы 2 и 1).

3. Примесное поглощение (полоса 3 на рис.11.2,а и переходы 3 на рис. 11.2,б) обусловлено переходами электронов из валентной зоны на примесные уровни, расположенные в нижней половине запрещенной зоны, или с примесных уровней, расположенных выше середины запрещенной зоны, в зону проводимости. Возможно также внутрицентровое поглощение, связанное с переходом электрона в примесном центре из основного в возбужденное состояние. Поскольку примесные уровни расположены в запрещенной зоне полупроводника, то примесные полосы поглощения лежат в длинноволновой области от края собственного поглощения, т.е.

$$\nu_{\text{прим}} < \nu_{\text{гр}}$$

(11.13)

Коэффициент примесного поглощения зависит от вида и концентрации примесных атомов и по своей величине на несколько порядков меньше коэффициента собственного поглощения (см. рис. 2.9).

Введение донорной и акцепторной примесей может менять положение края собственного поглощения из-за заполнения вакантных состояний вблизи дна зоны проводимости и потолка валентной зоны, а также из-за перекрытия при больших концентрациях примеси примесной подзоны с одной из разрешенных зон энергетического спектра (см. рис. 2.9,з). Если примесное поглощение сопровождается переходом электронов с примесного уровня в зону проводимости или из валентной зоны на примесный уровень, то наблюдается увеличение проводимости полупроводника.

4. Поглощение света свободными носителями заряда наблюдается в инфракрасной области спектра, когда энергия фотона $h\nu \ll \Delta E$. При поглощении фотонов малых энергий электроны переходят на более высокие энергетические уровни в пределах одной зоны (зоны проводимости

или валентной). Поскольку энергетическое расстояние между разрешенными уровнями в пределах одной зоны очень мало (квазинепрерывное распределение уровней в зоне), то спектр поглощения фотонов свободными носителями оказывается сплошным. Коэффициент поглощения пропорционален концентрации свободных носителей и, следовательно, сильно зависит от температуры полупроводника, которая определяет значение равновесной концентрации носителей заряда. С увеличением длины волны излучения коэффициент поглощения свободными носителями возрастает по закону $\alpha \approx \lambda^p$, где показатель степени p зависит от преобладающего механизма рассеяния носителей заряда.

5. Решеточное поглощение наблюдается в инфракрасной области спектра и сопровождается передачей энергии фотонов атомам кристаллической решетки. При этом происходит усиление колебаний атомов около положений равновесия, т.е. генерация фононов.

§ 11.2. Фотопроводимость полупроводников

Воздействие светом в области собственного и примесного поглощения может вызывать в полупроводнике генерацию неравновесных электронов Δn и дырок Δp , что сопровождается увеличением проводимости. При этом полная проводимость будет равна

$$\sigma(\lambda) = q\mu_n \Delta n + q\mu_p \Delta p + \sigma_0, \quad (11.14)$$

где $\sigma_0 = q\mu_n n_0 + q\mu_p p_0$ — проводимость полупроводника при отсутствии светового облучения; $\Delta\sigma$ — фотопроводимость,

$$\Delta\sigma = q\mu_n \Delta n + q\mu_p \Delta p. \quad (11.15)$$

Таким образом, фотопроводимость — это добавочная проводимость, обусловленная неравновесными носителями, созданными при оптической генерации. Основная проводимость σ_0 , вызванная равновесными носителями заряда, называется темновой проводимостью.

Поскольку фотопроводимость обычно обусловлена увеличением концентрации носителей в зонах под действием света, то ее часто называют концентрационной фотопроводимостью. Она возникает в результате нескольких процессов (рис. 11.3):

- 1) при поглощении фотонов электроны переходят из валентной зоны в зону проводимости, при этом наблюдается одновременное возрастание концентрации электронов и дырок (собственная фотопроводимость);
- 2) электроны под действием света переходят из валентной зоны на свободные акцепторные уровни, при этом увеличивается число дырок в валентной зоне (дырочная примесная фотопроводимость);
- 3) электроны под действием света переходят с донорных уровней в зону проводимости (электронная примесная фотопроводимость).

Очевидно, что концентрационная фотопроводимость может возникать только тогда, когда энергия фотона превышает либо ширину запрещенной зоны, либо энергетическое расстояние между одной из зон и примесным уровнем.

На рис. 11.4 приведен характерный вид спектра фотопроводимости полупроводника. В целом он достаточно хорошо коррелирует со спектром поглощения. Однако в области собственного поглощения наблюдается плавный спад фотопроводимости (ФП) с уменьшением длины волны при росте коэффициента поглощения. Этот спад обусловлен тем, что фотогенерация неравновесных носителей происходит с уменьшением длины волны все в более тонкой приповерхностной области полупроводника, где очень высока скорость рекомбинации носителей. Кроме того, спектры поглощения и фотопроводимости отличаются в случаях, если в процессе поглощения света не создаются неравновесные свободные носители заряда в разрешенных зонах. Например, в спектре фотопроводимости не проявляются полосы экситонного возбуждения и примесного поглощения, обусловленные внутрицентровыми процессами.

Изменение проводимости полупроводника при поглощении света может быть связано не с увеличением концентрации носителей в зонах, а с изменением подвижности носителей заряда.

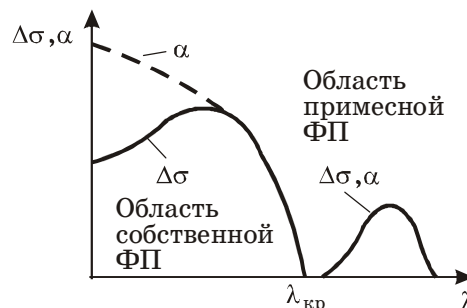
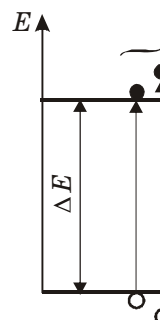


Рис. 11.3. Создание свободных носителей заряда при поглощении света в полупроводнике

Рис. 11.4. Спектр фотопроводимости полупроводника



При этом, поглощая фотоны, носители переходят внутри зоны на более высокие (для электронов) или более низкие (для дырок) энергетические уровни и, следовательно, изменяют свою эффективную массу и подвижность. Поскольку $\sigma = qn\mu$, то изменение подвижности ведет к изменению проводимости. Ранее (см. § 4.6) подобные процессы обсуждались при рассмотрении эффекта Ганна. Однако оптически возбужденные носители заряда за очень короткое время ($\sim 10^{-12}$ с) теряют свою избыточную энергию при столкновениях — термализуются. Поэтому вклад этого вида фотопроводимости в общую проводимость обычно мал.

§ 11.3. Фотопроводимость при импульсном освещении

Рассмотрим изменение фотопроводимости при освещении полупроводника n -типа импульсом света прямоугольной формы (рис. 11.5,а).

При освещении полупроводника генерация электронов с донорных уровней сопровождается рекомбинацией, и изменение концентрации избыточных электронов описывается уравнением непрерывности

$$\frac{d \Delta n}{d t} = G - \frac{\Delta n}{\tau_n}, \quad (11.16)$$

где Δn — концентрация неравновесных носителей; G — скорость фотогенерации носителей; τ_n — скорость рекомбинации электронов для случая низкого уровня возбуждения; τ_n — время жизни неравновесных носителей в зоне проводимости.

Решение уравнения (11.16) в предположении, что полупроводник начинает освещаться в момент времени $t = 0$ светом одинаковой интенсивности и, следовательно, при $t_1 = 0$ $\Delta n = 0$, имеет вид

$$\Delta n = G \tau_n (1 - e^{-t/\tau_n}). \quad (11.17)$$

Из выражения (11.17) следует, что начиная с момента включения импульса света ($t = 0$), концентрация неравновесных электронов увеличивается по экспоненциальному закону и при $t \rightarrow \infty$ достигает стационарной величины $\Delta n_{ст}$ (рис. 11.5,б):

$$\Delta n_{ст} = G \tau_n. \quad (11.18)$$

Скорость фотогенерации носителей будет равна

$$G = \beta \Phi_0. \quad (11.19)$$

где β — квантовый выход фотоэффекта, т.е. величина, равная отношению числа образующихся носителей заряда к общему числу поглощенных фотонов. Тогда

$$\Delta n_{ст} = \beta \Phi_0 \tau_n. \quad (11.20)$$

С момента прекращения освещения t_2 процесс генерации неравновесных носителей выключен, и в уравнении непрерывности остается только член, учитывающий рекомбинацию:

$$\frac{d \Delta n}{d t} = - \frac{\Delta n}{\tau_n}. \quad (11.21)$$

Решение простого уравнения (11.21) при условии $t_2 = 0$, $\Delta n = n_{ст}$ имеет вид

$$(11.22)$$

Из анализа выражения (11.22) следует, что, начиная с момента прекращения импульса света ($t = t_2 = 0$), концентрация избыточных электронов уменьшается по экспоненциальному закону до тех пор, пока все неравновесные носители не рекомбинируют. Фотопроводимость полупроводника n -типа пропорциональна концентрации неравновесных электронов, что следует из формулы (11.15). В связи с этим вид зависимости $\Delta\sigma(t)$ на рис. 11.5,в не отличается от кривой $\Delta n(t)$, показанной на рис. 11.5,б.

Рис. 11.5. Изменение во времени интенсивности света (а), концентрации носителей заряда (б) и фотопроводимости (в)

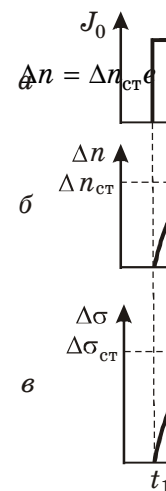
Кривые нарастания и спада являются кривыми релаксации фотопроводимости. Время релаксации на участках

подъема и спада фотопроводимости, как видно из уравнений (11.17) и (11.22), представляется одной величиной τ_n , т.е. является временем жизни неравновесных носителей заряда (в данном случае электронов). Тогда из выражений (11.17), (11.22) и (11.15) можно получить формулу для экспериментального определения времени жизни неравновесных носителей заряда по релаксации фотопроводимости:

$$\tau = \frac{t_2 - t_1}{\ln \Delta\sigma_1 - \ln \Delta\sigma_2}, \quad (11.23)$$

где $\Delta\sigma_1$ — фотопроводимость в момент времени t_1 ; $\Delta\sigma_2$ — фотопроводимость в момент t_2 .

В данном случае моменты времени t_1 и t_2 относятся к участку спада или подъема фотопроводимости. Таким образом, из рассмотренного в данном параграфе материала следует, что увеличение времени жизни избыточных носителей заряда замедляет нарастание и спад фотопроводимости, что прямо отражается на инерционности фотоприемников излучения.



§ 11.4. Фотоэлектродвижущая сила

Фотоэлектродвижущей силой называется разность потенциалов, возникающая между различными участками твердого тела в результате пространственного разделения созданных при фотогенерации неравновесных носителей заряда различного знака.

Такое разделение может произойти при равномерном освещении полупроводника с неоднородным распределением примесей, при неравномерном освещении полупроводника однородного состава, при фотогенерации носителей в областях полупроводника с внутренним электрическим полем (*p-n*-переход, область обеднения МДП-структуры и т.п.).

В результате освещения локального участка собственного полупроводника при условии, что $h\nu > \Delta E$, в освещенном месте возникает избыточная концентрация свободных носителей заряда (электронов и дырок), которые начинают диффундировать в неосвещенную область с более низкой концентрацией свободных носителей (рис. 11.6).

Вследствие различных скоростей диффузии электронов и дырок происходит их пространственное разделение и возникает фото-ЭДС. Поскольку обычно подвижность дырок меньше подвижности электронов, то освещенная поверхность полупроводника заряжается положительно, а неосвещенная — отрицательно. Такой механизм возникновения фото-ЭДС получил название эффекта Дембера.

При этом разность потенциалов освещенной и неосвещенной областей полупроводника может достигать единиц милливольт.

При локальной фотоионизации примесных состояний в полупроводнике возникает градиент неравновесных носителей одного знака между освещенным и неосвещенным участками, что ведет к появлению фото-ЭДС вследствие монополярной диффузии неравновесных носителей в неосвещенную область.

При равномерном освещении полупроводника с неоднородным распределением примеси светом из области примесного поглощения скорость фотоионизации примесных состояний будет выше в области с более высокой концентрацией примеси. Возникший градиент концентрации неравновесных носителей заряда обуславливает их диффузию и появление фото-ЭДС между участками образца с различным содержанием примеси.

При облучении *p-n*-перехода фотонами с энергией, большей ширины запрещенной зоны полупроводника, электроны и дырки, созданные светом, ускоряются полем контактной разности потенциалов и переводятся на разные стороны *p-n*-перехода (рис. 11.7), т.е. происходит пространственное разделение неравновесных электронов и дырок, что обуславливает возникновение фото-ЭДС, получившей название вентильной фото-ЭДС. Направление поля вентильной фото-ЭДС противоположно направлению поля контактной разности потенциалов *p-n*-перехода. Следовательно, высота потенциального барьера на *p-n*-переходе при появлении фото-ЭДС уменьшается на величину $qV_{\text{ф}}$, где $V_{\text{ф}}$ — значение вентильной фото-ЭДС. Зонная диаграмма *p-n*-перехода при появлении вентильной фото-ЭДС приобретает вид диаграммы *p-n*-перехода при прямом смещении (пунктир

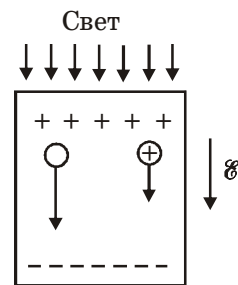


Рис. 11.6.
К эффекту Дембера

на рис. 11.7). При замыкании p - и n -областей внешней электрической цепью в ней начинает протекать электрический ток (фототок), зависящий от интенсивности освещения p - n -перехода, его температуры и свойств полупроводника.

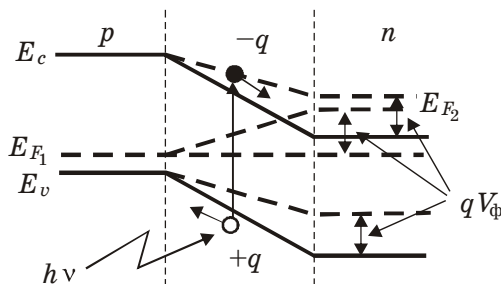


Рис. 11.7. Возникновение вентильной фото-ЭДС при облучении светом p - n -перехода

Эффект возникновения вентильной фото-ЭДС при освещении p - n -перехода получил название фотогальванического или вентильного фотоэффекта. На его основе созданы приборы для регистрации и измерения мощностей светового излучения, т.е. выполняющие роль фотоприемных устройств. Вентильная фото-ЭДС может создаваться не только в p - n -переходах, но и в других структурах, содержащих области с контактной разностью потенциалов. Например, для создания фотоприемников часто используются структуры металл-полупроводник (диоды Шоттки) и диоды на основе p - i - n -структур и гетеропереходов.

§ 11.5. Использование фотоэлектрических явлений в полупроводниковых приборах

На основе фотоэлектрических явлений созданы различные фоточувствительные и светоизлучающие полупроводниковые устройства (фоторезисторы, фото- и светодиоды, фототранзисторы, фотогальванические элементы и т.д.), предназначенные для регистрации и измерения световых потоков, управления характеристиками электрических цепей за счет световых воздействий или преобразования световой энергии в электрическую. Рассмотрим принципы работы некоторых фотоэлектрических устройств.

1. Фоторезисторы представляют собой пластину полупроводника (монокристаллическую или в виде тонкой пленки) с омическими контактами на противоположных концах. Схема включения фоторезистора показана на рис. 11.8. При включении источника питания через фоторезистор протекает темновой ток, равный согласно закону Ома

$$I_0 = \sigma_0 S \frac{V}{l}, \quad (11.24)$$

где l — длина полупроводника, заключенного между металлическими контактами; S — площадь поперечного сечения активной части фоторезистора.

При облучении фоторезистора квантами света с энергией $h\nu$ в цепи будет протекать ток

$$I = I_0 + \Delta I_\phi, \quad (11.25)$$

где $\Delta I_\phi = \Delta\sigma_\phi S \frac{V}{l}$ — величина приращения темнового тока, пропорциональная фотопроводимости.

На сопротивлении нагрузки при протекании тока в цепи с фоторезистором при освещении можно зарегистрировать падение напряжения

$$V_H = (\sigma_0 + \Delta\sigma_\phi) \frac{V}{l} R_H S. \quad (11.26)$$

Облучению светом переменной интенсивности на сопротивлении нагрузки будет соответствовать переменный сигнал с частотой световой модуляции. Амплитуда модулированного сигнала равна

$$\Delta V_\phi = \Delta\sigma_\phi \frac{V}{l} R_H S. \quad (11.27)$$

Параметры фоторезисторов (спектральная чувствительность, время фотоответа, обнаружительная способность и др.) зависят от свойств используемого полупроводникового материала и конструктивных особенностей прибора. Например, область спектральной чувствительности фоторезистора определяется спектром фотопроводимости полупроводника, а время фотоответа — дрейфовой скоростью носителей и расстоянием между омическими контактами фоторезистора.

Фоторезисторы используются в качестве датчиков световых потоков в различных устройствах.

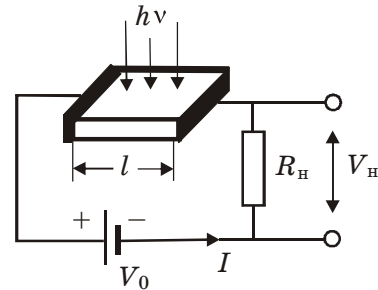


Рис. 11.8. Схема включения фоторезистора

2. Фотодиоды строятся на основе p - n -переходов, p - i - n -диодов, диодов со структурой металл-полупроводник (диоды Шоттки) и диодов на гетеропереходах.

При создании фотодиодов одной из основных задач, которую необходимо решить разработчику, является создание такой конструкции, чтобы поглощение фотонов происходило в основном в области контактного потенциального барьера. На рис. 11.9 приведены две возможные схемы размещения p - n -перехода фотодиода относительно светового потока.

Если p - n -переход освещается так, что лучи света направлены вдоль металлургической границы p - n -перехода (рис. 11.9,*а*), то свет с энергией кванта $h\nu > \Delta E$ полностью поглощается в очень тонком приповерхностном слое полупроводника, при этом область p - n -перехода в глубине кристалла не участвует в формировании фото-ЭДС. Кроме того, в создании фото-ЭДС могут принимать участие только неосновные носители заряда, возникшие при фотоионизации в слое толщиной не более диффузионной длины от области объемных зарядов p - n -перехода. Остальные носители, созданные светом, не успевают за время жизни дойти до p - n -перехода и исчезают в результате рекомбинации.

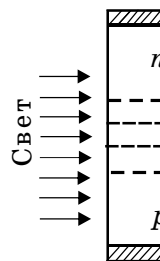
Рис. 11.9. Конструкции фотодиодов, в которых световой поток параллелен (*а*) и перпендикулярен (*б*) металлургической границе p - n -перехода

Эти причины приводят к низкой эффективности фотодиода, представленного на рис. 11.9,*а*. На рис. 11.9,*б* приведена более удачная конструкция фотодиода. Свет падает перпендикулярно плоскости p - n -перехода и поглощается в базе вблизи поверхности. Если толщина базы W меньше диффузионной длины неосновных носителей, то в создании вентильной фото-ЭДС могут принимать участие почти все дырки, генерируемые при фотопоглощении.

Таким образом, для повышения интегральной чувствительности фотодиодов необходимо, чтобы освещаемая база p - n -перехода была как можно тоньше, а полупроводниковый материал выбирают с возможно большей диффузионной длиной неосновных носителей, что обуславливает значительную инерционность фотодиодов.

Определим величину фото-ЭДС, создаваемую при освещении p - n -перехода. При отсутствии освещения и при разомкнутой внешней цепи p - n -переход должен находиться в равновесном состоянии, т.е. токи основных и неосновных носителей через переход будут равны:

$$I_{\text{осн}} = I_{\text{неосн}} = I_s . \quad (11.28)$$



При облучении p - n -перехода появляется фототок I_{ϕ} , обусловленный неравновесными неосновными носителями, который приводит к уменьшению высоты потенциального барьера на величину qV_{ϕ} , где V_{ϕ} — значение фото-ЭДС.

Очевидно, что в стационарном состоянии в освещенном p - n -переходе фототок и ток основных носителей, обусловленный прямым смещением p - n -перехода из-за возникающей фото-ЭДС, должны компенсировать друг друга. Тогда

$$I_{\phi} = I_s \left[\exp\left(\frac{qV_{\phi}}{k_0T}\right) - 1 \right]. \quad (11.29)$$

Откуда

$$(11.30)$$

Если замкнуть внешнюю цепь освещенного p - n -перехода на нагрузочное сопротивление $R_{\text{н}}$ (рис. 11.10), то в этой цепи возникнет ток $I_{\text{н}}$, обусловленный неравновесными неосновными носителями, созданными светом и переброшенными контактным полем через p - n -переход. При этом прямое смещение p - n -перехода за счет создаваемой фото-ЭДС уменьшится до величины $V_{\phi.\text{н}}$. Очевидно, что в этом случае должно выполняться следующее равенство для токов:



$$(11.31)$$

$$V_{\phi.\text{н}} = \frac{k_0T}{q} \ln\left(\frac{I_{\phi} - I_{\text{н}}}{I_s} + 1\right). \quad (11.32)$$

При коротком замыкании p - n -перехода все неравновесные неосновные носители, создающие фототок, будут уходить во внешнюю цепь, вследствие чего $V_{\phi.\text{н}} = 0$, а ток во внешней цепи будет равен фототоку.

Из выражений (11.30) и (11.32) следует, что при прочих равных условиях большая фото-ЭДС будет создаваться в фотодиодах с малым темновым током насыщения I_s , который определяется площадью p - n -перехода и равновесными концентрациями неосновных носителей в n - и p -областях диода.

Различают два режима работы фотодиодов:

Рис. 11.10. Схема включения фотодиода при отсутствии внешнего смещения

1) режим фотоэлемента, когда фотоприемник используется без внешнего смещения. Этот режим эффективен для прямого преобразования световой энергии в электрическую, но фотоэлементы отличаются высокой инерционностью и нелинейной световой характеристикой;

2) фотодиодный режим, когда на диод подается внешнее отрицательное смещение, превышающее значение фото-ЭДС. Подача внешнего обратного смещения на фотодиод позволяет снизить емкость p - n -перехода и уменьшить время пролета неосновных носителей через барьер. Поэтому фотодиодный режим обычно обеспечивает большее быстродействие, и, кроме того, в этом режиме фотодиоды имеют линейную связь между мощностью светового излучения и силой тока во внешней цепи фотоприемника.

Чтобы осуществить равномерное освещение p - n -перехода и не допустить произвольных потерь световой энергии при поглощении в базовых областях полупроводника, для создания вентиляльных фотоэлементов часто используют гетеропереходы.

Гетеропереходом называется p - n -переход между электронным и дырочным полупроводниками с различными значениями ширины запрещенной зоны (рис. 11.11).

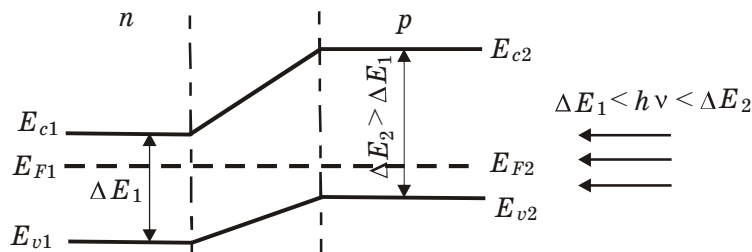


Рис. 11.11. Зонная диаграмма гетероперехода

Если освещать гетеропереход со стороны широкозонного полупроводника светом с энергией квантов $\Delta E_1 < h\nu < \Delta E_2$, то свет будет поглощаться и создавать электронно-дырочные пары только в узкозонном материале в непосредственной близости от p - n -перехода. Кроме того, фотоэлементы на основе гетеропереходов имеют более широкую область спектральной чувствительности (рис. 11.12), обычно охватывающую энергетический интервал от ΔE_1 до ΔE_2 (ΔE_1 и ΔE_2 — ширина запрещенной зоны полупроводников в n - и p -областях гетероперехода). Использование гетеропереходов позволяет существенно повысить КПД фотоэлемента. КПД определяется отношением максимальной мощности электрического тока, которую можно получить от фотоэлемента, к мощности излучения, падающего на фотоэлемент. В фотоэлементах на основе гетеропереходов КПД достигает 20 %, тогда как фотоэлементы на основе p - n -переходов имеют КПД порядка 12 %.

Для создания гетеропереходов используются обычно пары полупроводниковых материалов, таких как германий — арсенид галлия (Ge — GaAs), пары полупроводниковых бинарных соединений арсенид галлия — фосфид галлия (GaAs — GaP), и пары тройных соединений, например $\text{Al}_{x_1}\text{Ga}_{1-x_1}\text{As} - \text{Al}_{x_2}\text{Ga}_{1-x_2}\text{As}$. В последнем соединении ширина запрещенной зоны зависит от состава и изменяется от 1,43 эВ для $x = 0$ до 2,16 эВ, если $x = 1$.

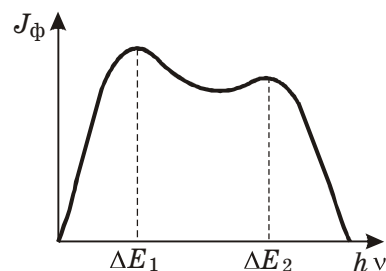


Рис. 11.12. Спектральная характеристика фотодиода на основе гетероперехода

Контрольные вопросы и задачи

1. (ШОСП). Разместите нижеперечисленные полосы поглощения, наблюдаемые в полупроводниках, в порядке убывания длины волны:

- 1) поглощение свободными носителями заряда;
- 2) собственное поглощение;
- 3) примесное поглощение;
- 4) экситонное поглощение.

2. (ЭЗФП). Определите длину волны излучения, при которой должна наблюдаться красная граница фундаментального поглощения в кремнии при $T = 300 \text{ К}$ ($\Delta E = 1,124 \text{ эВ}$). Ответ представьте в цифровом виде « x, y » в микрометрах.

3. (А7СЧ). Расставьте фоторезисторы, изготовленные из нелегированных кремния (Si, $\Delta E = 1,12 \text{ эВ}$), германия (Ge, $\Delta E = 0,7 \text{ эВ}$), арсенида галлия (GaAs, $\Delta E = 1,42 \text{ эВ}$) и фосфида галлия (GaP, $\Delta E = 2,24 \text{ эВ}$) в порядке убывания коротковолновой границы их спектральной чувствительности:

- 1) Si, Ge, GaP, GaAs;
- 2) GaP, GaAs, Si, Ge;
- 3) Ge, Si, GaAs, GaP;
- 4) GaAs, GaP, Ge, Si.

4. (Н9ФО). Укажите из нижеприведенных правильные определения фотопроводимости:

- 1) приращение проводимости, вызванное освещением полупроводника;
- 2) проводимость освещенного полупроводника;
- 3) проводимость, обусловленная неравновесными носителями заряда, созданными при освещении полупроводника;
- 4) проводимость, обусловленная равновесными и неравновесными носителями заряда освещенного полупроводника.

5. (ИТПЧ). Укажите причины возрастания пороговой чувствительности фотоприемных приборов с понижением температуры (под пороговой

чувствительностью понимают минимальный уровень световых сигналов, которые после преобразования в электрический сигнал могут быть уверенно выделены на фоне собственных электрических шумов фотоприемника).

Ответы:

- 1) увеличение ширины запрещенной зоны полупроводникового материала;
- 2) снижение величины темнового тока;
- 3) увеличение числа неравновесных носителей заряда, создаваемых освещением;
- 4) уменьшение концентрации равновесных носителей заряда пропорционально $\exp(-\Delta E/2k_0T)$.

6. (Д7ЭС). Когда в полупроводнике при освещении возникает внутреннее электрическое поле (фото-ЭДС)?

Ответы:

- 1) при равномерном освещении полупроводника с однородным распределением донорной или акцепторной примеси светом из области примесного поглощения;
- 2) при равномерном освещении полупроводника с неоднородным распределением донорной или акцепторной примеси светом из области примесного поглощения;
- 3) при освещении *p-n*-перехода светом из области фундаментального поглощения;
- 4) при освещении приповерхностной области обеднения полупроводника светом из области фундаментального поглощения;
- 5) при освещении приповерхностной области обогащения в полупроводнике светом из области собственного поглощения;
- 6) при неравномерном освещении полупроводника с однородным распределением примеси светом из области примесного поглощения.

7. (С1ФГ). Рассчитайте скорость фотогенерации электронов и дырок в собственном германии ($\Delta E = 0,7$ эВ) при воздействии на него инфракрасного излучения с длиной волны $\lambda = 1$ мкм и интенсивностью светового потока $J_0 = 2,21$ Дж/(м³·с). При этом считайте, что поглощение 10 фотонов сопровождается генерацией 9 пар носителей заряда, а коэффициент поглощения $\alpha = 10^6$ см⁻¹.

Ответ $x \cdot 10^y$, где x и y — целые числа, например $5 \cdot 10^{27}$ м⁻³·с⁻¹, наберите как последовательность символов 5 27 м -3 с -1.

8. (М9РФ). Световое возбуждение создает в образце кремния стационарную концентрацию избыточных носителей заряда 10^{17} см⁻³, при этом скорость рекомбинации равна 10^{16} см⁻³·с⁻¹.

Определите время, которое должно пройти с момента выключения освещения, чтобы концентрация избыточных носителей снизилась до 10^{16} см⁻³.

Ответ к задаче после вышеприведенного кода наберите в виде двузначного числа с единицей физической величины.

9. (Ю4РФ). Известно, что после включения освещения кремниевго образца n -типа светом из области примесного поглощения его проводимость за 2 мс изменилась от 20,2 до 21 Ом⁻¹·см⁻¹. Темновая проводимость образца составляла 20 Ом⁻¹·см⁻¹. Определите время жизни неравновесных носителей заряда.

Ответ наберите в виде десятичного четырехзначного числа с соответствующей единицей физической величины.

10. (Д7РФ). Для условий задачи 9 определите время между моментом включения света и моментом первого измерения проводимости образца, если известно, что скорость фотогенерации носителей равна $3 \cdot 10^{19}$ см⁻³·с⁻¹, а их подвижность составляет 500 см²/(В·с).

В ответе наберите цифры x , y с указанием степени z и единицы измерения, например $1,7 \cdot 10^{-6}$ с необходимо набрать как последовательность символов 1,7 -6 с.

Приложение

Некоторые сведения о строении твердых тел

Введение

Различают четыре агрегатных состояния вещества — газ, жидкость, твердое тело и плазма. В газах какой-либо порядок в расположении атомов отсутствует из-за очень большого расстояния между атомами, что приводит к отсутствию взаимодействия между ними. В жидкостях расстояние между атомами значительно меньше. Атомы начинают взаимодействовать между собой, и это ведет к возникновению ближнего порядка в их расположении: соблюдается определенный порядок в пространственном расположении ближайших к каждому атому атомов, т.е. атомов, относящихся к первой конфигурационной сфере. В твердых телах расстояние между атомами еще меньше. В результате атомы взаимодействуют не только с ближайшими соседями, но и с более далекими атомами, что может приводить к возникновению не только ближнего, но и дальнего порядка в пространственном расположении атомов. Плазма в общем случае представляет систему, состоящую из ионов, нейтральных атомов и электронов и обычно характеризуется значительным расстоянием между элементами системы и отсутствием какого-либо порядка в их пространственном расположении. Поскольку в микроэлектронике основную роль играют твердотельные компоненты, то в дальнейшем мы будем вести речь только о твердых телах.

Твердыми телами часто называют тела, сохраняющие свою форму и объем при отсутствии внешних воздействий. Различают три вида твердых тел: аморфные, поликристаллические и монокристаллические. Аморфные твердые тела (вар, воск, стекло и пр.) по своим свойствам близки к очень вязким жидкостям, и в них наблюдается только ближний порядок в расположении атомов. Поэтому свойства аморфных тел одинаковы во всех направлениях, т.е. для них характерна изотропия свойств. Поликристаллические твердые тела представляют собой объединение мелких кристаллов различной пространственной ориентации. Поэтому для них в пределах каждого кристалла соблюдается как ближний, так и дальний порядок в расположении атомов, но в целом из-за различной ориентации отдельных кристаллов свойства поликристаллического твердого тела не зависят от направления, т.е. они изотропны. Для монокристаллических твердых тел в каждом направлении соблюдается свой порядок в расположении атомов. Поэтому свойства монокристаллов зависят от направления, т.е. монокристаллы обладают анизотропией свойств.

Ближний и дальний порядок в пространственном расположении атомов в монокристаллах обусловлен тем, что атомы находятся в равновесном состоянии.

III. Кристаллическая структура твердых тел

III.1. Равновесное состояние атомов в монокристалле

Помимо сил химической связи, притягивающих атомы, между атомами действуют силы отталкивания, обусловленные кулоновским взаимодействием внутренних электронных оболочек или ядер. Обе эти силы зависят от расстояния. Однако силы притяжения являются дальнедействующими и возрастают сравнительно медленно с уменьшением расстояния между атомами, в то время как силы отталкивания короткодействующие и очень резко возрастают при уменьшении расстояния.

Современная теория дает приближенную формулу для силы притяжения:

$$F_{\text{пр}}(r) = -\frac{b}{r^{m+1}}, \quad (\text{П1})$$

где b и m — постоянные, величина которых зависит от вида химической связи; r — расстояние между атомами. Так, например, в ионных кристаллах и в металлах $m = 1$, а в кристаллах с молекулярной связью $m = 6$. Силы притяжения и соответствующая им потенциальная энергия обычно записываются с отрицательным знаком.

Силы отталкивания выражаются сложной математической зависимостью от расстояния. В грубом приближении они могут быть описаны фор-

$$U(r) = \frac{C_{\text{мул}}}{r^n} - \frac{B}{r^m}, \quad F_{\text{отт}}(r) = \frac{c}{r^{n+1}}, \quad (\text{П2})$$

где c и n — постоянные, величина которых зависит от природы взаимодействующих атомов. Для некоторых кристаллов полагают $n = 12$. При определенном расстоянии между атомами $r = r_0$ силы притяжения и отталкивания уравниваются друг друга, в результате чего атомы оказываются в равновесном состоянии.

Перейдем к рассмотрению энергии, обусловленной силами притяжения и отталкивания. Поскольку $F = -\frac{dU}{dr}$, составляющие потенциальной

энергии взаимодействия атомов будут иметь вид $U_{\text{пр}}(r) = -\frac{B}{r^m}$ и

$$U_{\text{отт}}(r) = \frac{C}{r^n}, \quad \text{где } B = \frac{b}{m}, \quad C = \frac{c}{n}.$$

Тогда результирующую потенциальную энергию запишем как

$$(\text{П3})$$

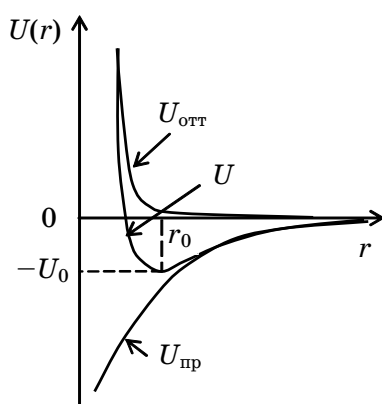


Рис. П1. Зависимость потенциальной энергии сил притяжения ($U_{\text{пр}}$), отталкивания ($U_{\text{отт}}$) и результирующей силы $U(r)$ от расстояния между атомами

На рис. П1 показаны составляющие потенциальной энергии, а также результирующая энергия U взаимодействия атомов. Здесь энергия свободных атомов принята за нуль.

Как видно из рисунка, равносному состоянию атомов, когда сила притяжения равна силе отталкивания, соответствует минимальное отрицательное значение потенциальной энергии U_0 . Следовательно, равновесное состояние является устойчивым. При всяком малом отклонении от равновесного состояния атомы возвращаются в исходное состояние.

Следует отметить, что мы рассмотрели равновесное состояние весьма упрощенно. Фактически атомы в кристаллической решетке испытывают значи-

тельно большее число взаимодействий. Наиболее важным из них является влияние атомов, окружающих данную пару атомов. Если даже рассматривать линейную цепочку атомов, то уже очевидно, что на данную пару действуют силы взаимодействия с соседними атомами, которые в свою очередь связаны с последующими атомами, и т.д. В трехмерных кристаллических решетках картина оказывается еще более сложной. Задача отыскания условия равновесного состояния атомов в кристаллической решетке существенно усложняется и является, по сути, задачей многих тел. Однако, несмотря на это, качественная картина равновесного устойчивого состояния атомов в монокристаллах остается той же.

В реальных кристаллах атомы практически никогда не находятся в положении устойчивого равновесия, так как они непрерывно совершают колебательные движения вокруг этого устойчивого состояния. Даже при абсолютном нуле температуры атомы совершают так называемые нулевые колебания и обладают некоторой конечной «нулевой» кинетической энергией. При повышении температуры интенсивность колебаний атомов вокруг устойчивого состояния увеличивается. При дальнейшем повышении температуры может наступить момент, когда энергия теплового колебательного движения превысит энергию связи частиц U_0 . В этом случае, очевидно, частица не возвратится в исходное состояние и химическая связь будет разорвана.

Многие элементы и химические соединения в силу сложной структуры электронных оболочек могут иметь не один, а два и более минимума энергии, причем последовательность этих минимумов может располагаться не только в радиальных направлениях, но и по угловым координатам.