

Учреждение образования  
«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ  
ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

**Н. В. Богомазова**

# **ФИЗИЧЕСКАЯ ЭЛЕКТРОНИКА И ЭЛЕКТРОННЫЕ ПРИБОРЫ**

**Электронный конспект лекций  
по одноименной дисциплине для студентов  
специальности 1-48 01 01 «Химическая технология  
неорганических веществ, материалов и изделий»  
специализации 1-48 01 01 13 «Химическая технология  
квантовой и твердотельной электроники»**

Минск 2021

УДК 621.382:537.311.322(0.034.2 PDF)

ББК 32.852я73

Б74

Рассмотрен и рекомендован к изданию редакционно-издательским советом Белорусского государственного технологического университета.

**Р е ц е н з е н т ы :**

доктор физико-математических наук, профессор кафедры физики  
твёрдого тела Белорусского государственного университета

*В. Г. Шепелевич;*

кандидат технических наук, доцент, доцент кафедры  
микро- и наноэлектроники УО «Белорусский государственный  
университет информатики и радиоэлектроники»

*Д. А. Котов;*

кандидат технических наук, доцент, доцент кафедры автоматизации  
производственных процессов и электротехники

УО «Белорусский государственный технологический университет»

*И. О. Оробей*

**Богомазова, Н. В.**

Б74 Физическая электроника и электронные приборы : электронный конспект лекций по одноименной дисциплине для студентов специальности 1-48 01 01 «Химическая технология неорганических веществ, материалов и изделий» специализации 1-48 01 01 13 «Химическая технология квантовой и твердотельной электроники» / Н. В. Богомазова. – Минск : БГТУ, 2021. – 119 с.

ISBN 978-985-530-952-0.

В электронном конспекте представлены краткие сведения об основных положениях современной физической электроники, а также приведены данные о принципах работы, функциональных структурах и выходных параметрах важнейших полупроводниковых приборов, включая различные типы микро- и оптоэлектронных диодов, а также транзисторов.

Издание предназначено для студентов технологического профиля, обучающихся по направлению «Электроника», и других заинтересованных лиц.

УДК 621.382:537.311.322(0.034.2 PDF)

ББК 32.852я73

ISBN 978-985-530-952-0

© УО «Белорусский государственный  
технологический университет», 2021

© Богомазова Н. В., 2021

## ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ .....	5
1. ОБЩИЕ ПОЛОЖЕНИЯ СОВРЕМЕННОЙ ЭЛЕКТРОНИКИ .....	6
1.1. Классическая электронная теория Друде. Эффект Холла (лекция 1) .....	9
1.2. Квантовомеханическое описание электрона в твердом теле (лекции 2–3) .....	14
1.2.1. Теоретические основы квантовой механики .....	14
1.2.2. Энергетическое состояние электрона в твердом теле .....	17
1.2.3. Взаимодействие электрона с потенциальными барьерами. Туннельный эффект .....	20
1.2.4. Динамика и статистика носителей заряда .....	22
1.3. Электроника реальных полупроводников (лекция 4) .....	29
1.3.1. Зонная структура реальных полупроводников .....	29
1.3.2. Примесные и неупорядоченные полупроводники .....	30
1.3.3. Неравновесные носители заряда .....	34
1.3.4. Дрейфовый и диффузионный ток .....	36
2. ПОЛУПРОВОДНИКОВАЯ МИКРО- И НАНОЭЛЕКТРОНИКА ...	38
2.1. Физика полупроводниковых структур (лекция 5) .....	38
2.1.1. Общая характеристика функциональных контактов .....	38
2.1.2. Электроника <i>p-n</i> -перехода .....	40
2.2. Дискретные полупроводниковые диоды (лекции 6–9) .....	45
2.2.1. Выпрямительный диод .....	46
2.2.2. Стабилитрон .....	47
2.2.3. Тиристоры .....	50
2.2.4. Туннельный диод .....	51
2.2.5. Диод Ганна .....	56
2.2.6. Диод Шоттки .....	58
2.3. Дискретные транзисторы (лекции 10–11) .....	61
2.3.1. Биполярный транзистор .....	61
2.3.2. Полевые транзисторы .....	68
2.3.2.1. Полевые транзисторы с управляющим контактом ( <i>p-n</i> -переходом) .....	68
2.3.2.2. Полевые МОП-транзисторы .....	70
2.4. Полупроводниковые и гибридные интегральные схемы (лекция 12) .....	74
2.5. Полупроводниковая наноэлектроника (лекция 13) .....	77

3. ОПТИЧЕСКАЯ И КВАНТОВАЯ ЭЛЕКТРОНИКА.....	81
3.1. Физические основы оптической и квантовой электроники (лекция 14) .....	82
3.1.1. Основные физические процессы при взаимодействии света с кристаллом .....	83
3.1.2. Виды поглощения света в полупроводниках .....	84
3.1.3. Фотопроводимость и фото-ЭДС .....	87
3.1.4. Излучение и усиление света в твердом теле.....	92
3.2. Оптоэлектронные полупроводниковые приборы (лекции 15–16) .....	94
3.2.1. Светодиоды .....	94
3.2.2. Лазеры.....	99
3.2.3. Фоторезисторы .....	104
3.2.4. Фотодиоды .....	107
3.2.5. Фотоэлементы.....	109
3.2.6. Фототранзисторы.....	112
3.3. Оптроны (лекция 17).....	113
Рекомендуемая литература .....	118

## ПРЕДИСЛОВИЕ

В соответствии с действующим учебным планом 4-летней подготовки специалистов квалификации «инженер-химик-технолог» по специализации «Химическая технология квантовой и твердотельной электроники» изучение дисциплины «Физическая электроника и электронные приборы», которая долгое время преподавалась доцентом кафедры физики Почтенным А. Е., предусмотрено на 3-м курсе параллельно со специальными дисциплинами материаловедческой и технологической направленности (в частности, с дисциплинами «Основы материаловедения и защита материалов» и «Технология полупроводниковых материалов и приборов»).

Целью изучения данной дисциплины является систематизация и расширение знаний будущих инженеров-технологов об устройстве и функционировании наиболее важных современных изделий электронной техники на примере устройств полупроводниковой электроники, включая дискретные приборы и интегральные схемы. В соответствии с физико-технической направленностью, а также с учетом основных актуальных направлений развития электронной техники содержание дисциплины, предполагающей проведение 34 часов лекционных занятий, поделено на три раздела:

- 1) общие положения современной электроники (8 ч);
- 2) полупроводниковая микро- и наноэлектроника (18 ч);
- 3) оптическая и квантовая электроника (8 ч).

В связи с большим объемом учебного материала, предусмотренного рабочей программой, в данном конспекте реализован структурированный подход при изложении базового материала по приборам полупроводниковой микроэлектроники. Для рассматриваемых приборов учебный материал структурирован единообразно и включает следующие номинации:

- а) определение, отражающее тип функционального контакта, физический эффект функционирования и назначение функционирования;
- б) физические основы функционирования с учетом энергетической диаграммы функционального контакта или материала и вида вольт-амперной характеристики (ВАХ);
- в) принципиальная функциональная структура и материалы;
- г) основные выходные параметры.

При самостоятельной работе рекомендуется обращаться к первоисточникам, приведенным в списке рекомендуемой литературы, а в конспекте уделять особое внимание терминам, выделенным *курсивом*.

Сегодня под термином «электроника» понимают научно-техническую область, связанную с разработкой и производством устройств, функционирующих на основе физических эффектов взаимодействия электронов с электромагнитными полями в различных средах, включая газ (в том числе разреженный), жидкость, твердое тело или плазму. Иными словами, главным действующим объектом в электронных устройствах является электрон (от греческого слова «ηλεκτρον» – «янтарь»).

В несколько упрощенных представлениях можно считать, что электрон является частицей, которая обладает массой покоя ( $m_0 = 9,1 \cdot 10^{-31}$  кг), элементарным отрицательным зарядом ( $q = e = 1,6 \cdot 10^{-19}$  Кл), элементарным магнитным моментом, равным магнетону Бора ( $\mu = \mu_B = 9,3 \cdot 10^{-24}$  Дж/Тл =  $5,8 \cdot 10^{-5}$  эВ/Тл) и полуцелым спином (внутренне присущим моментом импульса, равным  $1/2$ ).

Первым на существование элементарного электрического заряда указал еще в 1874 г. ирландский физик Джордж Стони (или Стоуни) (1826–1911), который сформулировал теорию дискретности электричества. Он же в 1891 г. вычислил величину такого заряда и предложил назвать его «электроном». Но статус открывателя электрона получил современник Джорджа Стони английский физик Джозеф Томпсон (1863–1906), который в 1897 г. экспериментально подтвердил существование стоуновских «электронов» при изучении катодных лучей, образованных отрицательно заряженными частицами.

В череде «электроноткрывателей» редко упоминают уникального английского физика и химика Майкла Фарадея (1791–1867), который широко известен как основатель теории электромагнетизма, блестяще завершенной другим английским физиком Джеймсом Максвеллом (1831–1879), а также как открыватель магнитооптического эффекта. Между тем опубликование М. Фарадеем еще в 1836 г. законов электролиза можно считать ключевым шагом к открытию электрона и к последовавшему за этим созданию теории строения атома. Дело в том, что М. Фарадей открыл закон электролиза (уравнение (1.1)), т. е. процесса разложения вещества под действием электричества. При этом ученый определил количественное значение физико-химической константы  $F$ , которая наряду с молярной массой  $M_i$  и числом электронов  $z$  входит в выражение для электрохимического эквивалента  $i$ -того

вещества  $q_i$ . Константа  $F$  является коэффициентом пропорциональности между массой вещества  $m$ , прореагировавшей на электроде, и количеством электричества  $Q$ , участвовавшим в электролизе:

$$m_i = q_i Q = \frac{M_i}{zF} Q. \quad (1.1)$$

Сегодня значение этой константы Фарадея  $F = 96\,485$  Кл/моль может посчитать, хочется верить, любой ученик средней школы, умножив постоянную итальянского химика Амедео Авогадро (1776–1856)  $N_A = 6,022 \cdot 10^{23}$  моль<sup>-1</sup> на заряд электрона  $e = 1,602 \cdot 10^{-19}$  Кл, поскольку постоянная Фарадея представляет собой заряд одного моля электронов. Уникальность открытия Фарадея состоит в экспериментальном количественном определении этой константы за десятки лет до открытия электрона и величины его заряда.

Исторически первой группой электронных приборов, т. е. сложных устройств, выполняющих заданную функцию преобразования или обработки электромагнитного сигнала, еще в конце XIX в. стала вакуумная электроника (табл. 1.1, рис. 1.2). Твердотельная электроника проделала длительный путь от первых выпрямителей до миниатюрного в сравнении с вакуумной лампой германиевого транзистора, который открыл эру современной микро- и нанoeлектроники.

Нынешний этап развития общества характеризуется повсеместным использованием электронных устройств и информационных технологий. Американские эксперты оценивают долю электроники в валовом промышленном продукте на уровне 80%. Производство микро- и нанoeлектроники обеспечивает авангардное развитие инновационных технологий соответствующих отраслей промышленности. Так, в 2021 г. по заказу мировых лидеров производства электронных изделий корпораций Intel и Apple еще один лидер производства функциональных кристаллов компания TSMC (Taiwan Semiconductor Manufacturing Company) анонсировала освоение процесса с топологическими нормами 3 нм и разработку процесса 2 нм, в то время как дорожные карты развития собственного кристалльного производства Intel (рис. 1.1) предполагают переход от 7 нм к 3 нм только к 2025 г.



Рис. 1.1. Дорожная карта развития производства компании Intel

Таблица 1.1

### Характеристика устройств различных областей электроники

Область электроники	Основные функциональные объекты	Дата основания	Характеристики приборов		
			Примеры	Достоинства	Недостатки
Вакуумная и плазменная электроника	Электрон + вакуум (+ плазма) + электрическое поле	1873 (лампа накаливания)	Усилительные лампы, клистроны, электронно-лучевые приборы	Низкий уровень шумов	Повышенные размеры
Твердотельная микро- и наноэлектроника	Электрон + полупроводник + электрическое поле	1883 (выпрямитель) 1947 (биполярный германиевый транзистор)	Диоды, транзисторы, стабилитроны, тиристоры, интегральные схемы	Миниатюрность, широкий диапазон преобразований, интегральная технология	Повышенный уровень шумов
Оптическая и квантовая электроника	Электрон + активная среда + электрическое поле + электромагнитное излучение (свет)	1876 (селеновый фотоэлемент) 1955 (мазер)	Светодиоды, фотоэлементы, фоторезисторы, лазеры, оптоэлектронные интегральные схемы	Бесконтактное управление цепями, широкий диапазон параметров элементов и сигналов	Повышенная потребляемая мощность, невысокий КПД преобразования



а



б



в

Рис. 1.2. Внешний вид изделий вакуумной (а), полупроводниковой (б) и квантовой (в) электроники

## 1.1. Классическая электронная теория Друде. Эффект Холла (лекция 1)

После завершения работ по открытию электрона ученые-физики предложили ряд электронных теорий, в рамках которых предпринимались попытки математически описать процессы электропереноса в твердом кристаллическом теле, т. е. в среде, которая состоит из близко и упорядоченно расположенных и при этом химически связанных атомов, включающих те самые окончательно открытые отрицательно заряженные электроны.

Напомним, что первым количественным шагом на этом пути можно считать закон немецкого физика Георга Ома (1787–1854). Г. Ом в 1826 г. установил наличие возрастающей линейной зависимости между напряжением  $U$  (В) и силой тока  $I$  (А) в электрической цепи. В своих опытах в качестве электродвижущей силы Ом использовал термо-ЭДС, открытую чуть раньше в 1822 г. немецким физиком Томасом Зеебеком (1770–1831). Кстати, значимость открытия Ома, кажущегося на первый взгляд примитивным, поначалу оценили только лишь в Королевском научном обществе Великобритании.

Одной из классических электронных теорий, т. е. теорий, базирующихся на представлениях классической механики, заложенных английским физиком Исааком Ньютоном (1643–1727), была теория немецкого физика Пауля Друде (1863–1906), предложенная в 1900 г., через три года после открытия электрона.

Теория электропереноса в твердом теле Друде построена на ряде допущений (рис. 1.3).

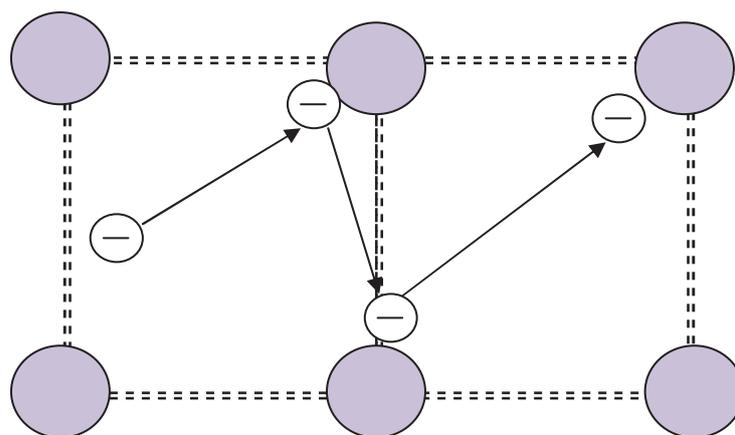


Рис. 1.3. Схема рассеяния электронов в кристаллической решетке в процессе электропереноса

1. Электрон проводимости является свободной (не связанной с атомом) частицей и соответственно обладает только кинетической энергией.

2. Электросопротивление материала (на примере металла) обусловлено столкновениями электронов проводимости с ионами решетки, при которых скорость электронов обнуляется.

3. Энергетическое распределение электронов подчиняется классической статистике Максвелла – Больцмана (британского физика Джеймса Максвелла (1831–1879) и австрийского физика Людвига Больцмана (1844–1906)) в виде экспоненциального множителя

$\exp(-\frac{\Delta E_{\text{акт}}}{kT})$ , описывающего, например, тепловое хаотическое распределение энергии частиц, способных преодолеть потенциальный барьер величиной  $\Delta E_{\text{акт}}$  при температуре  $T$  с учетом тепловой константы Больцмана  $k = 1,38 \cdot 10^{-23}$  Дж/К.

В рамках модели Друде введены важные физические понятия и математические выражения. Среди них:

1) *дрейфовая скорость* носителя заряда  $v_D$  (м/с), которая с учетом законов классической механики, в том числе 2-го закона Ньютона, при равноускоренном движении описывается выражением (1.2):

$$v_D = at = \frac{F_{\text{эл}}}{m_e} t = \frac{eE_U}{m_e} t, \quad (1.2)$$

где  $a$  – ускорение, м<sup>2</sup>/с, приобретаемое электроном под действием электрической силы  $F_{\text{эл}}$ , Н;  $t$  – продолжительность равноускоренного движения, с;  $m_e$  – масса электрона, кг;  $E_U$  – напряженность электрического ускоряющего поля, В/м;

2) *время релаксации*  $\tau$  (с) и *длина свободного пробега*  $L_E$  (м), отражающие соответственно время равноускоренного движения между двумя моментами рассеяния носителей заряда на ионах решетки, которое можно определить по формуле (1.3), и расстояние, пройденное носителем заряда за время  $\tau$  с учетом его дрейфовой  $v_D$  и тепловой  $v_T$  скоростей:

$$\tau = \frac{L_E}{v_D + v_T}; \quad (1.3)$$

3) *дрейфовая подвижность* носителя заряда  $\mu$  (м<sup>2</sup>/В·с или см<sup>2</sup>/В·с), которую обычно называют просто подвижностью носителя; она показывает, какую дрейфовую скорость приобретает носитель в поле напряженностью 1 В/м:

$$\mu = \frac{v_D}{E_U}. \quad (1.4)$$

Следует учитывать, что в металлах дрейфовая скорость электронов  $v_D \approx 5 \cdot 10^{-3}$  м/с значительно меньше, чем их тепловая скорость  $v_T$ , которую можно оценить, приравняв кинетическую и тепловую энергии электрона в виде уравнения (1.5), на уровне  $\approx 10^5$  м/с:

$$\frac{m_e v_T^2}{2} = \frac{3}{2} kT. \quad (1.5)$$

Такое соотношение дрейфовой и тепловой скорости электрона в металлах позволяет использовать равновесные статистические функции распределения при описании электропроводности.

Именно в модели Друде было получено наиболее употребительное выражение (1.6) для удельной электропроводности  $\sigma$  ( $\text{Ом}^{-1} \cdot \text{м}^{-1}$ ) или удельного электросопротивления  $\rho$  ( $\text{Ом} \cdot \text{м}$  или чаще  $\text{Ом} \cdot \text{см}$ ) материалов, которое может использоваться для описания и металлов, и полупроводников с учетом заряда носителей  $q$ , их концентрации  $n$  и подвижности  $\mu$ :

$$\sigma = qn\mu = en\mu = \frac{1}{\rho}. \quad (1.6)$$

Это несложное выражение мы часто используем в различных материаловедческих и технологических расчетах для полупроводников и поэтому называем «любимой» формулой.

Электронная теория Друде позволила определить ряд вышеуказанных понятий и неплохо описать процесс электропереноса в металлах, где наиболее корректно приближение «свободных» электронов. Однако для полупроводников эта теория не смогла обосновать существование положительно заряженных носителей заряда. В этой связи важно рассмотреть эффект, открытый американским физиком Эдвином Холлом (1855–1938) в 1879 г. и позволяющий определить заряд и концентрацию носителей в материале (рис. 1.4).

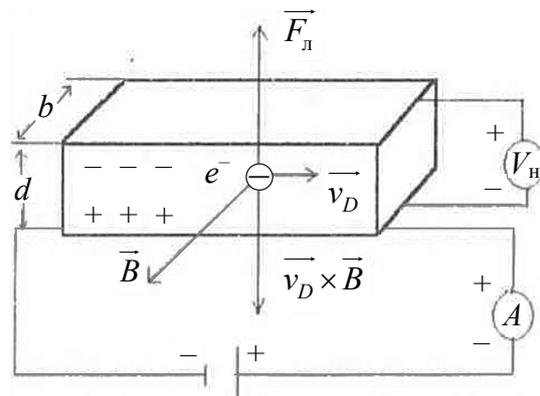


Рис. 1.4. Схема возникновения эффекта Холла

Эффект Холла заключается в возникновении постоянного электрического поля, или холловской разности потенциалов  $U_H$ , между гранями образца, к которому вдоль других граней приложено поперечное электромагнитное поле, т. е. поле, в котором вектор напряженности электрического поля  $\vec{E}$  и вектор магнитной индукции магнитного поля  $\vec{B}$  перпендикулярны друг другу.

Для примера на рис. 1.4 в приложенном к образцу электрическом поле возникает дрейфовая скорость для электрона  $v_D$ , направленная от отрицательного к положительному полюсу источника электрического поля. Со стороны приложенного магнитного поля с индукцией  $B$  на движущийся электрон действует сила Лоренца  $F_L$ , открытая нидерландским физиком Хендриком Лоренцем (1881–1928) и описываемая выражением

$$F_L = q[\vec{v}_D \times \vec{B}]. \quad (1.7)$$

Направление векторного произведения  $[\vec{v}_D \times \vec{B}]$  можно определить по правилу буравчика. На рис. 1.4 при вращении от вектора  $\vec{v}_D$  к вектору  $\vec{B}$  по кратчайшему расстоянию (по часовой стрелке) буравчик закручивается, т. е. векторное произведение направлено вниз, а сила Лоренца с учетом отрицательного заряда электрона направлена вверх. Важно заметить, что сила Лоренца отклоняет носители *обоих* знаков к одной и той же грани! В результате этого отклонения на противоположной грани (на рис. 1.4 – на нижней грани) скапливается заряд *нескомпенсированных неподвижных ионизованных примесных атомов* и возникает холловская разность потенциалов между двумя заряженными гранями (на рис. 1.4 – между нижней (ионизованные положительно заряженные доноры) и верхней (отрицательно заряженные электроны) гранями). Максимальное значение холловской разности потенциалов можно найти, приравняв электрическую силу холловского поля  $eE_H = eU_H/l$  и силу Лоренца  $e[v_D \times B]$ . Выражая из такого равенства холловскую разность потенциалов  $U_H$ , получим расчетную формулу (1.8):

$$U_H = \frac{IB}{enb}, \quad (1.8)$$

где  $I$  – сила тока, протекающего через образец, А;  $B$  – магнитная индукция приложенного магнитного поля, Тл;  $n$  – концентрация носителей заряда,  $\text{м}^{-3}$ ;  $b$  – размер образца в направлении индукции магнитного поля, м.

Измерения холловской разности потенциалов широко используются для определения заряда и концентрации носителей в полупроводниках. Алгоритм определения знака носителей включает 3 этапа:

1) определение грани, к которой сила Лоренца отклоняет **все** носители заряда (на рис. 1.4 – это верхняя грань);

2) определение полярности заряжения граней на основании знака измеренной холловской разности потенциалов и полярности подключения вольтметра (например, для рис. 1.4 отрицательное значение  $U_H$  при подключения положительного полюса вольтметра к верхней грани указывает, что верхняя грань образца заряжена отрицательно, а нижняя – положительно);

3) определение знака носителей заряда с учетом пунктов 1 и 2 (на рис. 1.4 по пункту 1 все носители отклоняются к верхней грани, по пункту 2 эта грань заряжена отрицательно, значит, основными носителями заряда в образце являются отрицательно заряженные электроны).

Для собственного полупроводника  $i$ -типа, в котором концентрации электронов проводимости и дырок близки,  $U_H \approx 0$ .

Таким образом, теории описания электрических явлений в проводниках и полупроводниках на основе представлений классической физики позволили охарактеризовать лишь некоторые грани процессов электропереноса. Несовершенство этих теорий обусловлено тем, что они не принимали во внимание фундаментальные особенности электрона, которые были учтены в квантовой механике. Среди этих отличительных особенностей электрона в сравнении с объектами классической механики:

– *корпускулярно-волновая двойственность* электрона, состоящая в способности проявлять как свойства частицы, например иметь массу покоя, так и свойства волны, например огибать препятствия при распространении, т. е. проявлять дифракцию;

– *квантование энергии* электрона, что требует при описании энергии электронов использовать набор не непрерывных, а дискретных значений, соответствующих разрешенным электронным состояниям;

– *вероятностный характер* описания электрона, позволяющий учитывать погрешности определения заданных параметров, например энергии, импульса, координат.

Кроме того, для реальных твердотельных материалов, в особенности полупроводников, корректное описание электропереноса предполагает количественный учет не только кинетической энергии ускоряемых электрическим полем носителей заряда, но и потенциальной энергии их взаимодействия как с собственными атомами, так и с ближайшими соседними атомами, которые в твердом теле находятся на расстоянии порядка 0,5 нм.

## 1.2. Квантовомеханическое описание электрона в твердом теле (лекции 2–3)

### 1.2.1. Теоретические основы квантовой механики

Создание квантовой (или волновой) механики называют революцией в физике, поскольку она соединила два подхода описания объектов: корпускулярный и волновой. Это позволило качественно и количественно объяснить многообразные электронные процессы в твердом теле и создать электронные устройства, без которых трудно представить современный мир. Теоретические основы квантовой механики заложены рядом нижеприведенных открытий, за которые практически все их авторы удостоились Нобелевских премий по физике. Авторство теорий отнесено к конкретным ученым, однако зачастую открытие является результатом множества находок ряда исследователей.

1. Квантовая теория немецкого физика, начинавшего жизненный путь как искусный пианист, Макса Планка (1858–1947), выдвигание которой в 1900 г. считают моментом основания квантовой механики, поскольку Планк выдвинул идею квантования энергии в процессах излучения и поглощения света твердым телом. Энергию кванта, т. е. некоторой порции света, М. Планк описал гениально простым выражением (1.9), полагая, что придуманные им кванты являются удачной теоретической моделью, но не энергетической реальностью:

$$E = h\nu \text{ или } E = \frac{hc}{\lambda}, \quad (1.9)$$

где  $h$  – постоянная Планка, или элементарный квант действия, который определяет границу между макромиром ньютоновской механики и микромиром квантовой механики,  $6,63 \cdot 10^{-34}$  Дж · с;  $\nu$  – частота излучения, Гц или  $\text{с}^{-1}$ ;  $c$  – скорость света в вакууме,  $3 \cdot 10^8$  м/с;  $\lambda$  – длина волны излучения, м.

2. Ряд физико-математических теорий ученого-скитальца Альберта Эйнштейна (1879, Германия – 1955, США), включая квантовую теорию света, а также специальную и общую теории относительности (1905), которые показали неразрывную связь между пространством и временем, а также между массой объекта  $m$  и энергией  $E$ , заключенной в этой массе, в рамках знаменитой формулы Эйнштейна (1.10):

$$E = mc^2. \quad (1.10)$$

3. Атомная теория датского ученого Нильса Бора (1885–1962), предложенная им в 1913 г. на базе планетарной, или ядерной, модели, в рамках которой учитывалась дискретность энергетических состояний отрицательно заряженных электронов, вращающихся вокруг положительно заряженного ядра. Н. Бор рассчитал радиус первой электронной орбитали в атоме водорода, который составил  $r_B = 0,5 \text{ \AA}$ .

4. Волновая теория французского физика Луи де Бройля (1892–1987), которая наделила волновыми свойствами все корпускулярные объекты, т. е. показала волновую природу материального мира. Волны де Бройля называют волнами вероятности, поскольку их квадрат отражает вероятность нахождения частицы в данной точке пространства. Длина волны  $\lambda_B$  была описана де Бройлем в 1924 г. также простым выражением (1.11):

$$\lambda_B = \frac{h}{p}, \quad (1.11)$$

где  $p$  – квазиимпульс частицы, кг/м · с, который в свою очередь может быть описан выражением (1.12):

$$p = \frac{h}{2\pi} k = \hbar k, \quad (1.12)$$

где  $\hbar = \frac{h}{2\pi}$  – приведенная константа Планка, Дж · с;  $k$  – волновое число, м<sup>-1</sup>, которое исходя из уравнений (1.11), (1.12) можно представить выражением (1.13):

$$k = \frac{2\pi}{\lambda_B}. \quad (1.13)$$

Выражение (1.13) определяет физический смысл волнового числа как величины, показывающей, сколько раз длина волны де Бройля укладывается в число  $2\pi$ .

5. Соотношение неопределенностей (или погрешностей) немецкого физика-теоретика Вернера Гейзенберга (1901–1976), которое было предложено 24-летним ученым в 1925 г. и создавало теоретическую основу для реализации вероятностного подхода описания корпускулярных объектов в рамках новой «матричной» механики. Это соотношение указывает на количественные ограничения при одновременном определении двух сопряженных физических величин, например координаты и импульса частицы, в виде неравенства (1.14):

$$\Delta p \Delta x \geq h. \quad (1.14)$$

Согласно соотношению (1.14), невозможно максимально точно определить и координату, и импульс, например электрона, поскольку произведение погрешностей определения этих величин не может быть меньше волновой константы Планка.

6. Волновая механика австрийского физика Эрвина Шрёдингера (1887–1961), который в 1926 г. предложил использовать для описания движения в силовом поле дискретных микрочастиц, например электрона, волновое уравнение (1.15), аналогичное уравнению электромагнитной волны, которое было получено гораздо раньше в 1865 г. английским физиком Джеймсом Максвеллом (1831–1879):

$$-\frac{\hbar^2}{2m}\Delta\Psi + U\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \quad (1.15)$$

где  $m$ ,  $\Psi$  и  $U$  – соответственно масса, волновая функция и потенциальная энергия частицы;  $\Delta$  – оператор Лапласа, равный в декартовых координатах сумме вторых производных данной функции, например  $\Psi$ , по трем пространственным координатам  $x$ ,  $y$ ,  $z$ :

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2};$$

$i$  – мнимая единица ( $i^2 = -1$ );  $t$  – время.

Характеристической величиной в уравнении Шрёдингера является волновая функция частицы  $\Psi$ , которая в общем случае зависит от четырех переменных  $\Psi = f(x, y, z, t)$  и должна быть однозначной, непрерывной и конечной. Физический смысл имеет произведение волновой функции на комплексно сопряженную ей величину  $\Psi\Psi^*$  или просто квадрат волновой функции  $\Psi^2$  в случае действительных значений  $\Psi$ . Такое произведение представляет собой плотность вероятности обнаружения частицы в данной точке в данный момент времени и соответственно подчиняется условию нормировки:

$$\int_V \Psi\Psi^* dV = 1.$$

В общем виде для многочастичной системы, например твердого тела, включающего множество электронов, постулированное уравнение Шрёдингера (1.15) не решается. На практике используют различные приближения для получения частных решений этого уравнения. Например, для стационарной системы в случае одномерного приближения уравнение Шрёдингера имеет частный вид (1.16):

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi}{dt^2} + U\Psi = E\Psi, \quad (1.16)$$

где  $E$  – полная энергия частицы, Дж.

### 1.2.2. Энергетическое состояние электрона в твердом теле

При описании движущегося электрона, например при воздействии на него электрического поля, можно выделить три варианта состояний:

- 1) «свободный» электрон, не обладающий потенциальной энергией;
- 2) электрон в изолированном атоме, который взаимодействует по большей части со своим ядром;
- 3) электрон в твердом теле, взаимодействующий не только с собственным атомом, но и с ближайшими соседними атомами (ионами).

В нашем рассмотрении мы будем под характеристикой состояния электрона понимать не привычный для электронов в атоме набор квантовых чисел (главного, орбитального, магнитного и спинового), а энергию электрона в данном состоянии.

«Свободный» электрон можно считать классическим объектом, его энергия может быть описана ньютоновским параболическим дисперсионным законом в виде следующей зависимости энергии  $E$  от импульса  $p$ :

$$E = E_{\text{кин}} = \frac{mv^2}{2} = \frac{p^2}{2m}, \quad (1.17)$$

где  $v$  – скорость движения электрона, м/с.

В изолированном атоме электроны в соответствии с атомной теорией Н. Бора и принципом швейцарского физика Вольфганга Паули (1900–1958) находятся в дискретных разрешенных энергетических состояниях, которые условно называют орбиталями. В таком случае движущиеся электроны могут быть описаны дискретным параболическим дисперсионным законом в виде набора точек, подчиняющихся уравнению (1.17) и соответствующих отдельным энергетическим уровням электронов в атоме, например  $1s$ ,  $2p$ ,  $3d$  и т. д.

Наиболее сложную задачу представляет описание движения электрона, в том числе процесса электропереноса, в многочастичном твердом теле в рамках современных представлений квантовой механики. Осознавая всю сложность математического аппарата квантовой механики, рассмотрим замечательную качественную модель, которая позволяет феноменологически и вместе с тем весьма наглядно описать

электронно-транспортные свойства различных материалов. Речь идет о *зонной модели*, на которой базируется современная электронная теория материалов, в частности полупроводников.

Зонная теория твердого тела была предложена Ф. Вильсоном в 1931 г. на базе метода молекулярных орбиталей (МО), описывающего образование молекулярных орбиталей из взаимодействующих атомных электронных уровней. Согласно зонной теории, при сближении атомов на достаточно малые расстояния (порядка  $5 \text{ \AA}$ ), характерные для твердого тела, в результате взаимодействия электронов с соседними атомами происходит расщепление *дискретных* атомных энергетических уровней в *квазинепрерывные* энергетические зоны, образованные большим количеством молекулярных орбиталей. В пределах таких разрешенных зон уровни разделены незначительными энергетическими зазорами, и электрон может перемещаться между ними при минимальной активации, например обладая некоторой тепловой энергией, среднее значение которой при температуре  $T$  равно  $E_T^{\text{cp}} = \frac{3}{2}kT$ .

Иными словами, электрон, находящийся в разрешенной зоне, при наличии в ней свободных уровней превращается в носитель заряда, обеспечивающий электроперенос в твердом теле. Графически эта модель иллюстрируется кривыми Вильсона (рис. 1.5).

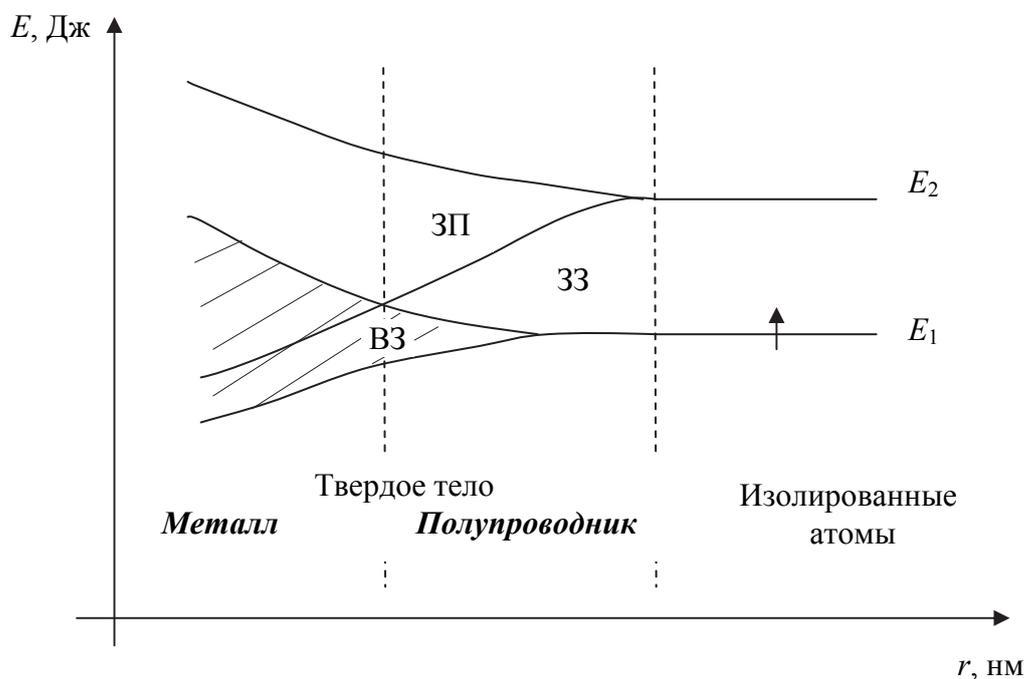


Рис. 1.5. Схема образования энергетических зон в твердом теле (кривые Вильсона)

Согласно кривым Вильсона, при сближении атомов сначала происходит расщепление внешних уровней с энергией  $E_2$  и при дальнейшем сближении расщепляются более глубокие уровни с энергией  $E_1$  (рис. 1.5). Между заполненной валентными электронами *валентной зоной* (ВЗ) и свободной от электронов *зоной проводимости* (ЗП) существует энергетический зазор запрещенных электронных уровней, которые образуют *запрещенную зону* (ЗЗ). Наличие запрещенной зоны, характерное для полупроводников и диэлектриков, обуславливает активационный характер проводимости. При достаточно сильном сближении атомов и расщеплении уровней происходит перекрывание валентной зоны и зоны проводимости, что характерно для металлов с безактивационной проводимостью.

Количественное описание зонной структуры твердого тела предложено, например, в рамках квантово-механической модели американского физика Ричарда Фейнмана (1918–1988). В этой модели использовалось уравнение Шрёдингера в операторной форме (1.18):

$$H\Psi = i\hbar \frac{\partial \Psi}{\partial t}, \quad (1.18)$$

где  $H$  – оператор полной энергии частицы в виде:

$$H = -\frac{\hbar^2}{2m} \Delta + U.$$

Опуская математические подробности решения уравнения Шрёдингера в модели Р. Фейнмана, отметим, что это решение было получено с использованием, во-первых, принципа суперпозиции для волновой функции электронов в многоэлектронной системе; во-вторых, разделения пространственной и временной частей волновой функции в приближении их независимости; в-третьих, исключения пространственной части волновой функции. Наиболее простой случай решения в виде гармонической функции при учете взаимодействия электрона с собственным атомом и двумя ближайшими атомами, расположенными на расстоянии  $b$ , соответствует уравнению (1.19):

$$E = E_0 - 2A \cos(kb), \quad (1.19)$$

где  $E$ ,  $E_0$  – энергия электрона с учетом его взаимодействия с соседними и с собственным атомом соответственно, Дж;  $A$  – константа, отражающая особенности потенциального поля данного твердого тела и зависящая, например, от степени ионности связи.

Фактически выражение (1.19) представляет собой периодический дисперсионный закон тригонометрического вида для электрона в твердом теле с учетом его взаимодействия с двумя ближайшими соседними атомами. Из этого выражения вытекает, что при образовании твердого тела электронный уровень с энергией  $E_0$  расщепляется в разрешенную энергетическую зону шириной  $4A$ , в которой разрешенная энергия электрона составляет  $E_0 \pm 2A$  (рис. 1.6).

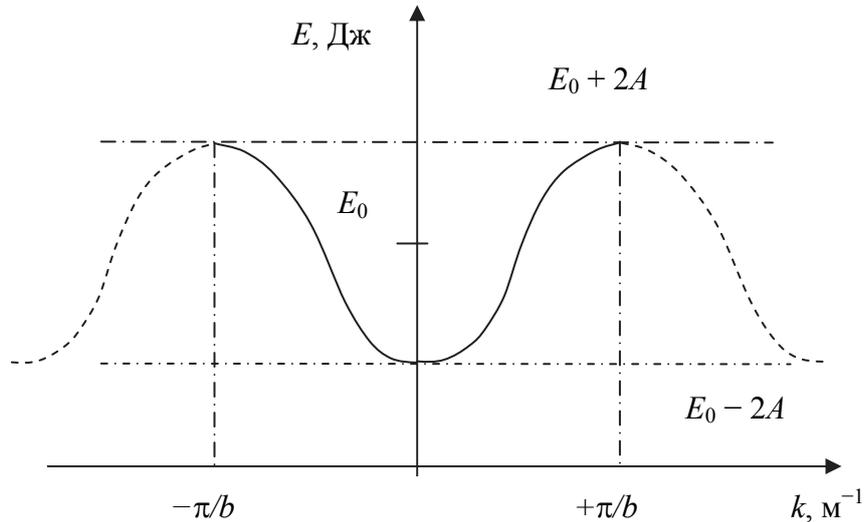


Рис. 1.6. Периодическая дисперсионная кривая для электрона в твердом теле

Периодический тригонометрический дисперсионный закон вида (1.19) может быть использован для описания энергетического состояния электронов в периодическом потенциальном поле кристаллической решетки твердого тела в пределах энергетических зон, задействованных в процессах электропереноса, а именно валентной зоны и зоны проводимости, разделенных запрещенной зоной.

### 1.2.3. Взаимодействие электрона с потенциальными барьерами. Туннельный эффект

С учетом периодического характера потенциального поля кристаллических твердотельных материалов и соответствующего периодического дисперсионного закона (рис. 1.6) энергию электрона в таких материалах можно описать набором потенциальных ям (энергетические минимумы), разделенных потенциальными барьерами (энергетические максимумы). В связи с этим для описания и последующего управления электронными процессами в твердом теле, например при создании полупроводниковых приборов, важное значение имеет опи-

сание особенностей взаимодействия электрона с потенциальными барьерами как в пределах данного тела, так и на границе контактов разных твердых тел или тела и окружающей среды.

Вероятностный характер описания объектов микромира, в частности электронов, в представлениях квантовой механики позволяет обосновать возможность специфических эффектов взаимодействия электрона с потенциальными барьерами, которые не имеют аналогов в классической механике. Примерами таких эффектов являются явления подбарьерного прохождения (или *туннелирования*) и надбарьерного отражения.

Эти квантовомеханические явления (рис. 1.7) состоят в преодолении частицей потенциального барьера шириной  $l$ , высота которого  $U_0$  больше энергии частицы  $E_1$ , т. е.  $E_1 < U_0$  (туннелирование), или в отражении частицы от барьера, высота которого  $U_0$  меньше энергии частицы  $E_2$ , т. е.  $E_2 > U_0$  (надбарьерное отражение).

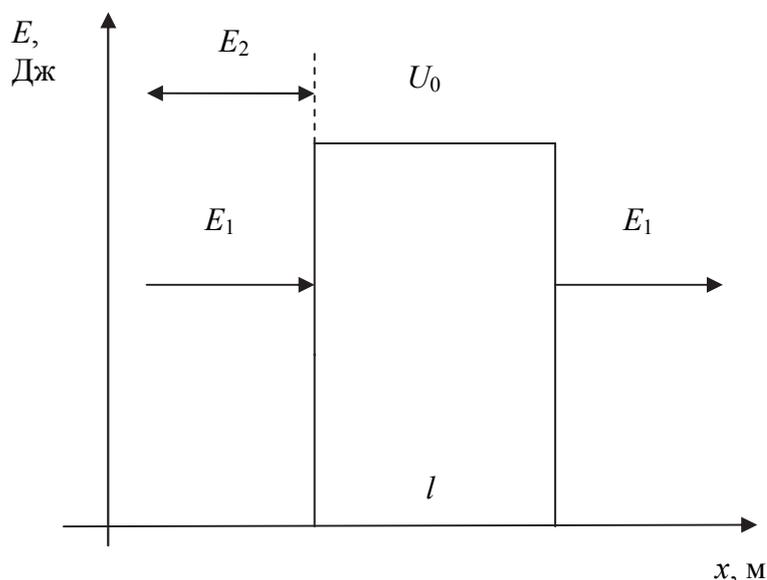


Рис. 1.7. Схема квантовомеханических явлений туннелирования (подбарьерного прохождения) и надбарьерного отражения

Обоснование этих эффектов возможно на основании соотношения неопределенностей Гейзенберга (1.13), которое в случае описания энергии частицы  $E$  в момент времени  $t$  может быть представлено неравенством (1.20):

$$\Delta E \Delta t \geq h. \quad (1.20)$$

Если максимально точно зафиксировать момент времени  $t$ , то погрешность определения момента времени  $\Delta t$  будет иметь минимальное

значение, а погрешность определения энергии  $\Delta E$  будет максимальна. Это означает, что в данный момент времени энергия частицы может существенно отличаться от начального значения и составлять

$$E_{\text{макс}} = E_1 + \Delta E_{\text{макс}} > U_0, \quad E_{\text{мин}} = E_2 - \Delta E_{\text{макс}} < U_0,$$

т. е. мгновенное значение энергии частицы может оказаться больше (туннелирование) или меньше (отражение) высоты потенциального барьера.

Количественное описание практически более важного явления туннелирования электронов с использованием их волновых функций для случая, например, прямоугольного одномерного потенциального барьера шириной  $l$  (рис. 1.7) приводит к выражению для коэффициента прохождения (или прозрачности) барьера  $D_T$  в виде (1.21):

$$D_T = D_0 \exp\left[-\frac{2l}{\hbar} \sqrt{2m(U_0 - E)}\right], \quad (1.21)$$

где  $D_0$  – коэффициент, который зависит от количественного соотношения высоты потенциального барьера  $U_0$  и энергии налетающей частицы  $E$  (в частности, для случая  $E = U_0/2$  можно принять  $D_0 = 1/4$ ).

Коэффициент прозрачности барьера  $D_T$ , имеющий физический смысл вероятности туннелирования частицы через барьер, принимает значения от 0 до 1. Описание туннелирования электронов через систему периодически расположенных барьеров, что соответствует случаю кристаллического твердого тела, приводит к решению, указывающему, что существует ряд квазистационарных энергетических состояний электрона, для которых вероятность туннелирования равна 1. Иными словами, для электрона в некоторых энергетических состояниях барьер оказывается абсолютно прозрачным, что означает возможность отсутствия рассеяния носителей заряда при движении по кристаллу.

#### 1.2.4. Динамика и статистика носителей заряда

Напомним, что в рамках рассматриваемого теоретического вступления перед изучением электронных приборов основной задачей для нас является описание процессов электропереноса или электронно-транспортных свойств твердотельных материалов с проводимостью I рода, т. е. связанной с движением электронов проводимости. Основной количественной характеристикой движущихся объектов в классической физике является их скорость  $v$ . При описании движения кван-

товомеханического объекта необходимо учесть его волновые свойства, что приводит к выражению (1.22):

$$v = \frac{1}{\hbar} \frac{\partial E}{\partial k}, \quad (1.22)$$

для использования которого следует задать в явном виде зависимость энергии  $E$  от волнового числа  $k$ , например в виде тригонометрического дисперсионного закона (1.19).

Однако выражением для скорости электрона с учетом его дисперсионного закона в твердом теле динамика электрона не исчерпывается. В квантовой механике специальной динамической характеристикой электрона является *эффективная масса*  $m^*$ , которая в отличие от массы покоя  $m_0$  не отражает ни инертных, ни гравитационных свойств электрона. Выражение для эффективной массы электрона может быть получено с использованием классической ньютоновской механики, где для расчета ускорения дифференцируется по времени скорость электрона (1.22) и в качестве ускоряющей силы учитывается внешняя электрическая сила. В такой модели эффективная масса электрона может быть описана выражением (1.23):

$$m^* = \frac{\hbar^2}{\partial^2 E / \partial k^2}. \quad (1.23)$$

При использовании фейнмановского дисперсионного закона выражение общего вида (1.23) принимает частный вид (1.24):

$$m^* = \frac{\hbar^2}{2Ab^2 \cos(kb)}. \quad (1.24)$$

Эффективная масса электрона является тензорной величиной и позволяет учесть влияние потенциального периодического поля кристаллического тела на движение носителя заряда в электрическом поле. В частности, анализ области значений  $m^*$  показывает, что в диапазоне значений  $kb$  от  $\pi/2$  до  $\pi$   $m^* < 0$ . Это означает, что движение носителей происходит в противоположном направлении, что возможно при изменении его заряда с отрицательного на положительный, т. е. при смене, например, электронной проводимости на дырочную. Кроме того, величина  $m^*$  обратно пропорциональна подвижности носителей в твердом теле.

Определившись с динамическими характеристиками электронов, обратимся к статистической характеристике, в качестве которой обычно рассматривают концентрации носителей заряда в твердом

теле. В общем случае эта величина зависит от множества факторов и изменяется во времени. В реальных полупроводниках носители заряда возникают благодаря наличию электроактивных собственных или примесных дефектов, способных при данных условиях ионизоваться с образованием подвижных зарядов, т. е. электронов в свободной зоне проводимости или дырок в заполненной валентной зоне. Более простой случай статистического описания носителей в твердом теле соответствует состоянию установившегося теплового равновесия между процессами генерации и рекомбинации носителей заряда, когда равновесная концентрация носителей заряда при постоянной температуре во времени не изменяется. Для начала остановимся на собственном полупроводнике, т. е. полупроводнике *i*-типа (от англ. *intrinsic* – собственный), в котором реализуется биполярная проводимость при одинаковой концентрации носителей разного знака  $n_e = n_h = n_i$ .

Для расчета концентрации носителей  $dn$  выделяют в разрешенной зоне твердого тела некоторый интервал энергий  $dE$  и с учетом величины плотности электронных состояний  $g(E)$ , отражающей количество электронных уровней в единичном энергетическом интервале, а также с учетом максимального числа электронов на одном уровне оценивают значение  $dn$  в виде (1.25):

$$dn = 2g(E)f(E,T)dE, \quad (1.25)$$

где  $f(E,T)$  – функция распределения электронов по данным энергетическим уровням в пределах  $dE$  при данной температуре  $T$ .

Выражение для описания плотности электронных состояний было получено немецким физиком и математиком Арнольдом Зоммерфельдом (1868–1951) в рамках модели «свободных» электронов, находящихся в бесконечно глубокой прямоугольной потенциальной яме. Напомним, что это приближение, которое использовалось и в классической модели Друде, более корректно описывает носители в металлических материалах.

В своих вычислениях Зоммерфельд использовал соотношение неопределенностей Гейзенберга (1.14) для случая трехмерного пространства, а точнее, для фазового пространства, т. е. абстрактного объединенного пространства координат и импульсов. В качестве дисперсионного закона использовался классический параболический непрерывный закон (1.17). Результатом вычислений явилось выражение (1.26), отражающее зависимость плотности энергетических состояний электрона  $g$  от энергии этих состояний  $E$ :

$$g(E) = 2\pi \left( \frac{2m}{h^2} \right)^{\frac{3}{2}} \sqrt{E}. \quad (1.26)$$

Из уравнения (1.26) следует, что плотность разрешенных энергетических уровней прямо нелинейно зависит от энергии этих уровней.

Для решения уравнения (1.25) остается оценить заселенность, или вероятность заполнения, электронных уровней, т. е. нужно определиться с видом функции распределения электронов по заданным энергетическим уровням  $f(E, T)$  при некоторой температуре  $T$ .

Классическая функция распределения Максвелла – Больцмана не может быть использована, поскольку она не учитывает квантование энергии электрона, а также квантовомеханический принцип неразличимости тождественных частиц, описываемых одинаковой волновой функцией  $\Psi$ . Для описания заполнения дискретных энергетических уровней используются два вида статистических функций: Ферми – Дирака и Бозе – Эйнштейна, авторы которых итальянский физик, создатель ядерного реактора Энрико Ферми (1901–1954), английский физик Поль Дирак (1902–1984), индийский физик Шатъендранат Бозе (1894–1974) и уже упоминавшийся Альберт Эйнштейн (1879–1955), кроме Ш. Бозе, получили Нобелевские премии по физике и окончили жизнь и труды в США. В общем виде функция распределения для квантующихся энергетических состояний имеет вид (1.27):

$$f(E, T) = \frac{1}{e^{\frac{E-E_0}{kT}} \pm 1}, \quad (1.27)$$

где  $E_0$  – характерная энергия уровня, который с одинаковой вероятностью заполнен или свободен, т. е. для этого уровня  $f = 1/2$ .

В выражении (1.27) знак «+» соответствует функции распределения Ферми – Дирака, а знак «–» – функции Бозе – Эйнштейна (табл. 1.2). Выбор частного вида функции распределения зависит от спинового состояния данной частицы. Так, частицы с полуцелым спином описываются антисимметричными или несимметричными волновыми функциями, для которых характерно изменения знака волновой функции на противоположный при перестановке частиц. Эти частицы подчиняются принципу Паули и находятся в энергетических состояниях с отличающимися наборами квантовых чисел, т. е. каждая частица в своем состоянии. Эти частицы называют

фермионами, к ним относится и ключевой объект наших рассуждений – электрон. Другой объект, открывший миру корпускулярно-волновую двойственность и являющийся квазичастицей, поскольку не имеет массы покоя, а представляет собой энергетический квант электромагнитного излучения оптического диапазона, т. е. фотон, относится к бозонам. Бозоны, в отличие от фермионов, наоборот, с наибольшей вероятностью заполняют уже заполненные энергетические состояния.

Таблица 1.2

**Сравнительная характеристика функций распределения Ферми – Дирака и Бозе – Эйнштейна**

Наименование функции распределения	Формула	Вид волновой функции	Выполнение принципа Паули, спин	Объекты описания	
				название	примеры
Ферми – Дирака	$f(E, T) = \frac{1}{e^{\frac{E-E_0}{kT}} + 1}$	Антисимметричная	Да, полуцелый	Фермион	Электрон, протон, нейтрон
Бозе – Эйнштейна	$f(E, T) = \frac{1}{e^{\frac{E-E_0}{kT}} - 1}$	Симметричная	Нет, 0 или целый	Бозон	Фотон, фонон, мезон

Но вернемся к задаче статистического описания носителей заряда в твердом теле. В нашем случае для электрона следует использовать функцию распределения Ферми – Дирака, в которой характерная энергия  $E_0$  получила название энергии Ферми или уровня Ферми  $E_F$ . Этот уровень, соответствующий  $f = 1/2$ , разделяет свободные и заполненные энергетические состояния электрона в твердом теле. Таким образом, для функции распределения мы можем использовать уравнение (1.28):

$$f(E, T) = \frac{1}{e^{\frac{E-E_F}{kT}} + 1}. \quad (1.28)$$

Теперь можно обсудить решение уравнения (1.25) с учетом уравнений (1.26) и (1.28) для теоретического случая *собственного* полупроводника, в котором отсутствуют примесные атомы. При этом можно использовать для плотности энергетических состояний выражение (1.26), полученное в модели Зоммерфельда для «свободных» электронов, но используя вместо массы покоя эффек-

тивную массу электрона  $m^*$ , отражающую влияние потенциального поля решетки на движение электрона, т. е. позволяющую учесть «несвободность» электрона в твердом теле.

Для получения решения уравнения (1.26) необходимо задать пределы интегрирования, которые соответствуют для электронов в зоне проводимости значениям от  $E_g$  до  $+\infty$  и для квазичастиц «дырок», фактически представляющих незаполненный электронный уровень в

валентной зоне, от 0 до  $-\infty$ . Особенностью описания дырок является то, что для них вероятность их наличия в валентной зоне равна вероятности отсутствия электрона в этой изначально заполненной зоне, т. е. равна не  $f$ , а  $1 - f$ . В интегральном виде уравнение (1.26) для концентрации электронов в зоне проводимости  $n_e$  и концентрации дырок в валентной зоне  $n_h$  имеет соответственно вид (1.29) и (1.30):

$$n_e = 4\pi \left( \frac{2m_e^*}{h^2} \right)^{\frac{3}{2}} \int_{E_F}^{\infty} \sqrt{E - E_g} f(E, T) dE; \quad (1.29)$$

$$n_h = 4\pi \left( \frac{2m_h^*}{h^2} \right)^{\frac{3}{2}} \int_{-\infty}^0 \sqrt{-E} (1 - f(E, T)) dE. \quad (1.30)$$

Поскольку уравнения (1.29) и (1.30) при подстановке функции распределения Ферми – Дирака в квадратурах не вычисляются, то с учетом того, что в собственном полупроводнике уровень Ферми расположен посередине запрещенной зоны (рис. 1.8) и энергетический зазор  $E - E_F \gg kT$ , используется математически более простая функция распределения Максвелла – Больцмана. В таких приближениях получены выражения для расчета концентрации носителей заряда в собственном полупроводнике в виде (1.31) и (1.32). Отметим, что при практическом использовании этих уравнений для полупроводников со сложной структурой разрешенных энергетических зон эффективная масса носителей заряда определяется как среднее геометрическое

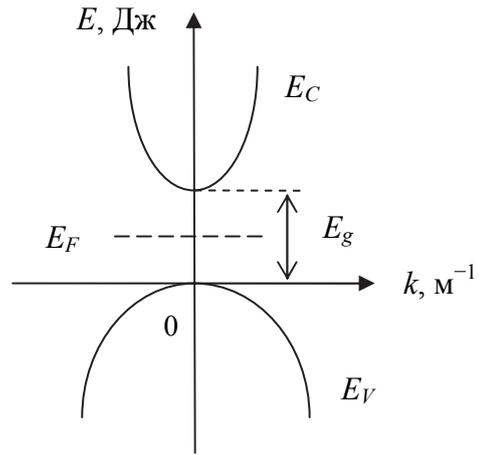


Рис. 1.8. Энергетическая структура собственного полупроводника

между различными значениями эффективной массы вдоль трех координат для данной зоны, которые определяются экспериментально методом циклотронного резонанса:

$$n_e = 2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{\frac{3}{2}} \exp \left( -\frac{E_g - E_F}{kT} \right); \quad (1.31)$$

$$n_h = 2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{\frac{3}{2}} \exp \left( -\frac{E_F}{kT} \right), \quad (1.32)$$

где предэкспоненциальные множители представляют собой эффективные плотности электронных состояний в зоне проводимости  $N_c$  и валентной зоне  $N_v$ .

Количественное значение энергии уровня Ферми для собственного полупроводника можно получить в виде выражения (1.33), приравняв друг к другу значения концентрации носителей заряда  $n_e = n_h$ :

$$E_F = \frac{E_g}{2} + \frac{3}{4} kT \ln \left( \frac{m_h^*}{m_e^*} \right). \quad (1.33)$$

Из уравнения (1.33) следует, что при увеличении температуры в полупроводнике уровень Ферми смещается от исходного при  $T = 0$  значения  $E_g/2$  (рис. 1.9, б) вверх и сила этого смещения зависит от соотношения эффективных масс носителей заряда.

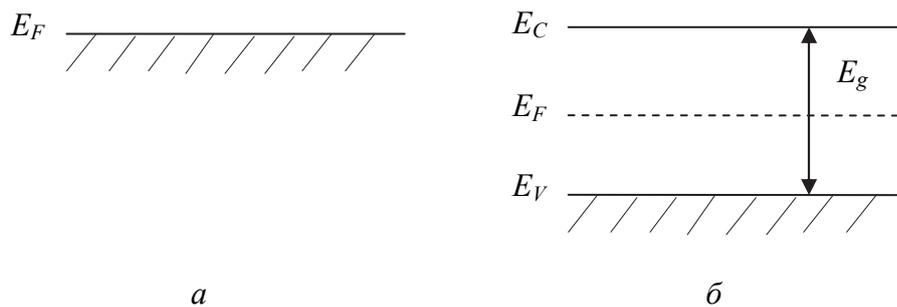


Рис. 1.9. Энергетическая диаграмма металла (а) и собственного полупроводника  $i$ -типа (б)

В металлах, где валентная зона и зона проводимости перекрываются (рис. 1.9, а), уровень Ферми совпадает с последним заполненным электронами энергетическим уровнем.

## 1.3. Электроника реальных полупроводников (лекция 4)

### 1.3.1. Зонная структура реальных полупроводников

Особенности зонной структуры полупроводников существенно влияют на процессы генерации носителей заряда, а также обуславливают наличие специфических электронных эффектов в конкретных материалах. Варианты зонной структуры различают в зависимости от количества и положения экстремумов дисперсионных кривых, описывающих границы ЗП и ВЗ.

В частности, вариант зонной структуры, приведенный на рис. 1.8 (с. 27), соответствует случаю *прямозонного однодолинного* полупроводника, в котором положение максимума дисперсионной кривой потолка валентной зоны  $E_V$  совпадает по волновому числу с положением минимума дисперсионной кривой дна зоны проводимости  $E_C$ . Такая зонная структура характерна, например, для GaN, InAs, GaSb, а также большей части халькогенидных полупроводников классов  $A^{II}B^{VI}$  и  $A^{IV}B^{VI}$ .

Более сложной зонной структурой обладают *прямозонные двудолинные* полупроводники (рис. 1.10), например GaAs, InP. В таких материалах экстремумы дисперсионных кривых  $E_C$  и  $E_V$  совпадают по  $k$ , но в зоне проводимости присутствуют два минимума, т. е. две потенциальные ямы, в которых стабильно могут находиться электроны проводимости. При этом эффективная масса в долинах 1 и 2 не совпадает, чаще  $m_{e1}^* < m_{e2}^*$ , следовательно,  $\mu_{e1} > \mu_{e2}$ .

В *прямозонных* полупроводниках при активации, например при поглощении фотонов, осуществляются *прямые* электронные переходы (рис. 1.10), при которых в результате поглощения фотонов с энергией  $h\nu \geq E_g$  электроны переходят из ВЗ в ЗП, повышая свою энергию на  $h\nu$  и не изменяя свой квазиимпульс  $p = \hbar k$  и соответственно волновое число.

Значительная часть полупроводников, в частности простые

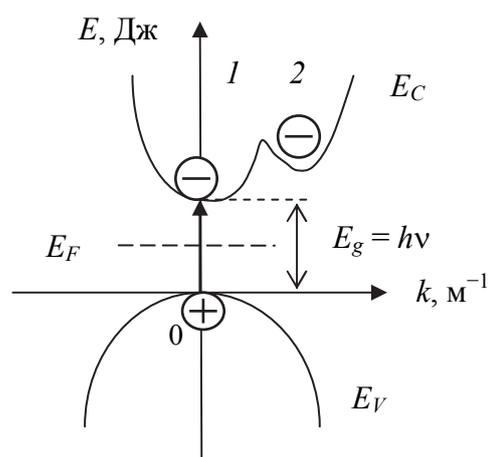


Рис. 1.10. Энергетическая структура прямозонного двудолинного полупроводника

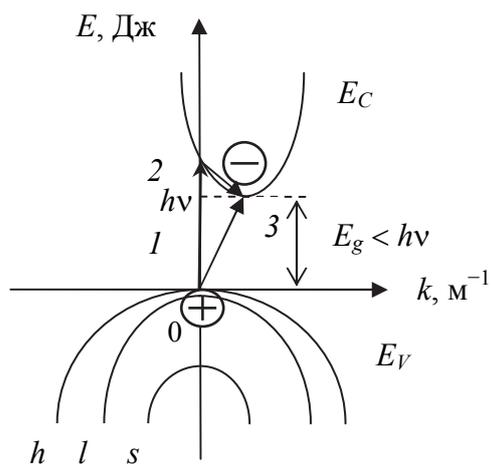


Рис. 1.11. Энергетическая структура непрямозонного однодолинного полупроводника

полупроводники Si и Ge, некоторые сложные полупроводники SiC, GaP, AlSb, относятся в *непрямозонным* материалам, в которых положение максимума дисперсионной кривой потолка валентной зоны  $E_V$  не совпадает по волновому числу с положением минимума дисперсионной кривой дна зоны проводимости  $E_C$ . В этом случае при поглощении полупроводником фотона реализуется *непрямой* электронный переход с изменением и энергии, и импульса электрона (рис. 1.11). Непрямой переход электрона (переход  $1 \rightarrow 3$ , рис. 1.11) осуществляется в два этапа:

1) поглощение электроном фотона с энергией  $h\nu > E_g$ , приводящее к переходу электрона из ВЗ в ЗП с сохранением величины импульса (рис. 1.11, переход  $1 \rightarrow 2$ );

2) взаимодействие электрона с фононом, сопровождающееся понижением энергии электрона и изменением его импульса при повышении энергии фонона (рис. 1.11, переход  $2 \rightarrow 3$ ).

Непрямые электронные переходы являются результатом трехчастичных взаимодействий электрон – фотон – фонон, вероятность которых существенно меньше двухчастичных взаимодействий электрон – фотон в случае прямых переходов. Вместе с тем участие фонона, обладающего большим импульсом, чем электрон, обеспечивает возможность изменения импульса электрона. В приборах, основанных на поглощении света полупроводником, предпочтительно использовать прямозонные материалы.

### 1.3.2. Примесные и неупорядоченные полупроводники

Собственный полупроводник, описанный в предыдущем пункте, является модельным объектом, под которым понимают идеальный монокристалл, не содержащий собственных и примесных дефектов. Реальные материалы характеризуются наличием дефектов кристаллической решетки различной мерности, а также наличием большого количества примесей разной химической природы.

Степень влияния примеси на процессы электропереноса в полупроводнике зависит от типа примеси, которые с этой точки зрения

разделяются на *электроактивные*, сильно увеличивающие проводимость, и *электронейтральные*, несильно уменьшающие проводимость (рис. 1.12).

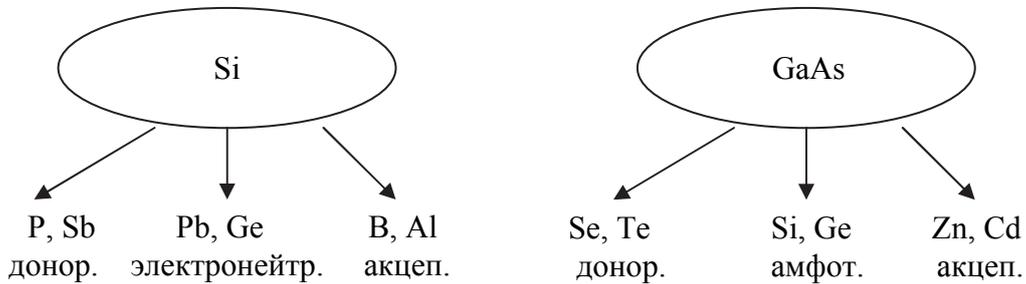


Рис. 1.12. Примеры примесей в простых и сложных полупроводниках

В свою очередь *электроактивные* примеси в зависимости от электронной функции, которую они могут выполнить в материале-матрице, (и следовательно, от типа возникающего носителя заряда) делятся на *донорные*, *акцепторные* и *амфотерные* (характерны только для сложных полупроводников).

*Донорные* примеси (рис. 1.13, а) – это примеси, за счет которых в зоне проводимости полупроводника могут появляться электроны проводимости. Эти примеси обуславливают в полупроводнике проводимость *n*-типа (от англ. *negative* – отрицательный (основной носитель)). *Донорными* являются примеси, валентность которых *больше*, чем валентность атома полупроводника, замещаемого этой примесью. Примесный электронный уровень *неионизованного донора* *заполнен*. Уровень Ферми  $E_F$  в полупроводнике *n*-типа при температуре 0 К находится в запрещенной зоне между последним заполненным разрешенным уровнем (донорным  $E_D$ ) и первым свободным разрешенным уровнем (дном зоны проводимости  $E_C$ ), т. е. смещен вверх.

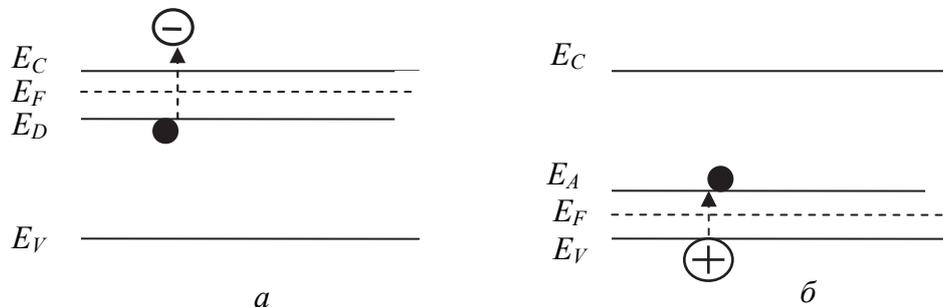


Рис. 1.13. Зонная диаграмма полупроводника *n*-типа (а) и *p*-типа (б)

Акцепторные примеси (рис. 1.13, б) – это примеси, за счет которых в валентной зоне полупроводника появляются квазичастицы в виде положительно заряженных дырок. Эти примеси обуславливают в полупроводнике проводимость *p*-типа (от англ. *positive* – положительный (основной носитель)). Акцепторными являются примеси, валентность которых меньше, чем валентность атома полупроводника, замещаемого этой примесью. Примесный электронный уровень неионизованного акцептора свободен. Уровень Ферми  $E_F$  в полупроводнике *p*-типа при температуре 0 К находится в запрещенной зоне между последним заполненным разрешенным уровнем (потолок валентной зоны  $E_V$ ) и первым свободным разрешенным уровнем (акцепторный  $E_A$ ), т. е. смещен вниз.

Амфотерные примеси характеризуются промежуточными значениями электроотрицательности (валентности) и геометрических размеров между атомами, образующими полупроводниковое соединение. Они могут встраиваться в разные положения кристаллической решетки и, соответственно, выполнять разную электронную функцию в полупроводнике (донорную или акцепторную). На практике амфотерные примеси предпочтительно занимают определенное положение и приводят к возникновению проводимости соответствующего типа (электронной или дырочной). Однако даже в таких случаях концентрация носителей заряда оказывается меньшей, чем концентрация амфотерной примеси.

Электронейтральными в простых и сложных полупроводниках являются примеси, замещающие атомы совпадающей валентности. Эти примеси уменьшают подвижность носителей заряда аналогично тому, как это происходит в легированных металлах. В результате электросопротивление полупроводника незначительно увеличивается.

В реальных материалах, даже таких высокочистых, как полупроводники для микроэлектронных устройств, присутствуют значительные количества (на уровне  $10^{-5}$  ат. %) различных примесей. Одновременное

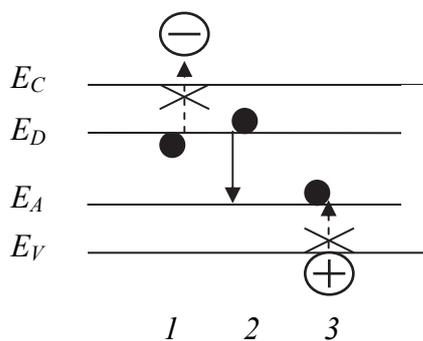


Рис. 1.14. Схема компенсации примесей в полупроводнике

присутствие донорных и акцепторных примесей приводит к процессу компенсации этих примесей в материале. В таком случае вместо генерации носителей (процессы 1 и 3 на рис. 1.14), которая возможна при повышении энергии электрона, происходит электронный переход 2 с понижением энергии, в результате которого электрон из одного локализованного состояния ( $E_D$ ) переходит в другое

локализованное состояние ( $E_A$ ), т. е. носитель заряда (делокализованная заряженная частица) не образуется. Реальные полупроводники в основном являются компенсированными материалами, которые количественно характеризуются степенью компенсации ( $K_{\text{ком}} < 1$ ), равной отношению концентрации примесей различного типа и соответственно отражающей долю атомов некомпенсированных примесей, поставляющих носители заряда и определяющих  $n$ - или  $p$ -тип полупроводника.

При определении типа проводимости полупроводника, например при комнатной температуре, кроме типа и концентрации примесей следует учитывать глубину залегания примесных уровней (или энергию ионизации этих примесей). По глубине залегания создаваемых в ЗЗ уровней примеси делятся на *мелкие* и *глубокие* (рис. 1.15). Глубина залегания донорной примеси  $\Delta E_D$  отсчитывается от дна зоны проводимости  $E_C$ , акцепторной примеси  $\Delta E_A$  – от потолка валентной зоны  $E_V$ .

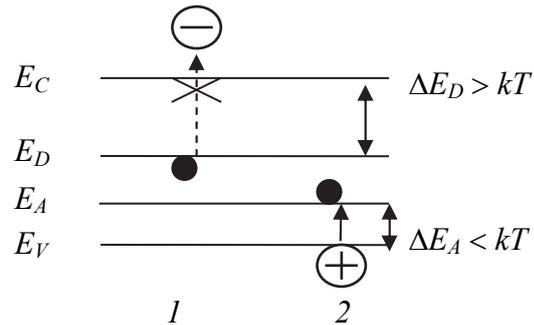


Рис. 1.15. Зонная диаграмма полупроводника, содержащего глубокую и мелкую примеси

Мелкими являются примеси, энергия ионизации которых меньше либо равна тепловой энергии частицы (например, при данной температуре  $E_T = 3/2 kT = 0,025$  эВ). При *комнатной температуре мелкие примеси ионизованы* в результате перехода 2 (рис. 1.15), а *глубокие – не ионизованы*, т. е. переход 1 (рис. 1.15) не происходит без дополнительного повышения температуры.

Для расчета концентрации носителей заряда в полупроводниках  $n$ - и  $p$ -типа могут быть использованы уравнения (1.31) и (1.32), полученные для собственного полупроводника, с учетом положения уровня Ферми, соответствующего данному типу материалов (рис. 1.13). В реальном полупроводнике обычно присутствуют носители обоих знаков, т. е. реализуется биполярная проводимость с преобладанием носителей определенного типа: электронного ( $n$ -тип с преобладанием донорной примеси) или дырочного ( $p$ -тип с преобладанием акцепторной примеси). В любом полупроводнике выполняется закон сохранения заряда в виде (1.34):

$$n_e n_h = n_i^2. \quad (1.34)$$

Это означает, что при введении в собственный полупроводник доноров или акцепторов увеличивается концентрация соответственно электронов проводимости или дырок, а концентрация носителей противоположного знака пропорционально уменьшается.

Обсуждая особенности реального полупроводника, который обычно является примесным компенсированным материалом, следует остановиться и на неупорядоченных системах, к которым относят аморфные или сильно разупорядоченные, например сильно легированные, кристаллические тела. Описание процессов электропереноса в этих материалах, обладающих упорядочением только в пределах ближнего порядка, базируется на модели локализации электронных состояний американского физика Филипа Андерсона (1923–2020).

Идея состоит в том, что в неупорядоченных системах, в отличие от упорядоченных кристаллов, происходит локализация волновой функции электрона в окрестности конкретного узла системы. В этом случае вместо *зонного* механизма электропереноса по делокализованным в пределах тела состояниям реализуется *прыжковый* механизм движения носителей по локализованным состояниям, который характеризуется пониженной подвижностью носителей. Для реализации в материале прыжковой проводимости расстояние между центрами локализации, например между примесными атомами, не должно превышать величину перколяционного радиуса  $r_{п}$ , связанного с концентрацией центров локализации  $N_{ц}$  выражением (1.35):

$$r_{п} = \frac{0,865}{N_{ц}^{\frac{1}{3}}}. \quad (1.35)$$

В представлениях перколяционной теории о протекании тока по телу при условии образования бесконечного перколяционного кластера с учетом уравнения (1.35) получена зависимость прыжковой электропроводности (или электросопротивления  $\rho$ ) от концентрации центров локализации, которая имеет экспоненциальный вид (1.36):

$$\rho = \rho_0 \exp \left( \frac{1,73}{a_{л} N_{ц}^{\frac{1}{3}}} \right), \quad (1.36)$$

где  $a_{л}$  – радиус локализации носителя заряда, например электрона, м.

### 1.3.3. Неравновесные носители заряда

В полупроводнике, находящемся при некоторой постоянной температуре, устанавливается равновесие между процессами генерации и рекомбинации носителей заряда, что соответствует некоторой посто-

янной во времени концентрации этих носителей, называемых *равновесными* (или тепловыми). Если в материале при активации проводимости отсутствует равновесие, т. е. скорость генерации и рекомбинации носителей не равны между собой, то образующиеся при этом носители называют *неравновесными*.

На практике генерацию неравновесных носителей заряда осуществляют, например, при облучении материала светом. Это приводит к генерации неравновесных фотоносителей и возникновению фотопроводимости  $\Delta\sigma$ , т. е. избыточной проводимости по сравнению с тепловой  $\sigma_T$ . В таком облучаемом материале суммарный ток обусловлен участием равновесных и неравновесных носителей:  $\sigma_\Sigma = \sigma_T + \Delta\sigma$ .

Количественное описание процессов нарастания  $\Delta\sigma_\uparrow$  и спада  $\Delta\sigma_\downarrow$  фотопроводимости во времени  $t$  соответственно при наличии и отсутствии освещения подчиняется уравнениям (1.37) и (1.38):

$$\Delta\sigma_\uparrow = \Delta\sigma_{\text{стац}} \left( 1 - \exp\left(-\frac{t}{\tau}\right) \right); \quad (1.37)$$

$$\Delta\sigma_\downarrow = \Delta\sigma_{\text{стац}} \exp\left(-\frac{t}{\tau}\right), \quad (1.38)$$

где  $\Delta\sigma_{\text{стац}}$  – стационарное (постоянное во времени) значение фотопроводимости, См/м;  $\tau$  – время жизни неравновесных носителей заряда от момента генерации до момента рекомбинации, с.

Величина  $\tau$  может быть определена по кривым релаксации фотопроводимости при импульсном облучении полупроводника (рис. 1.16).

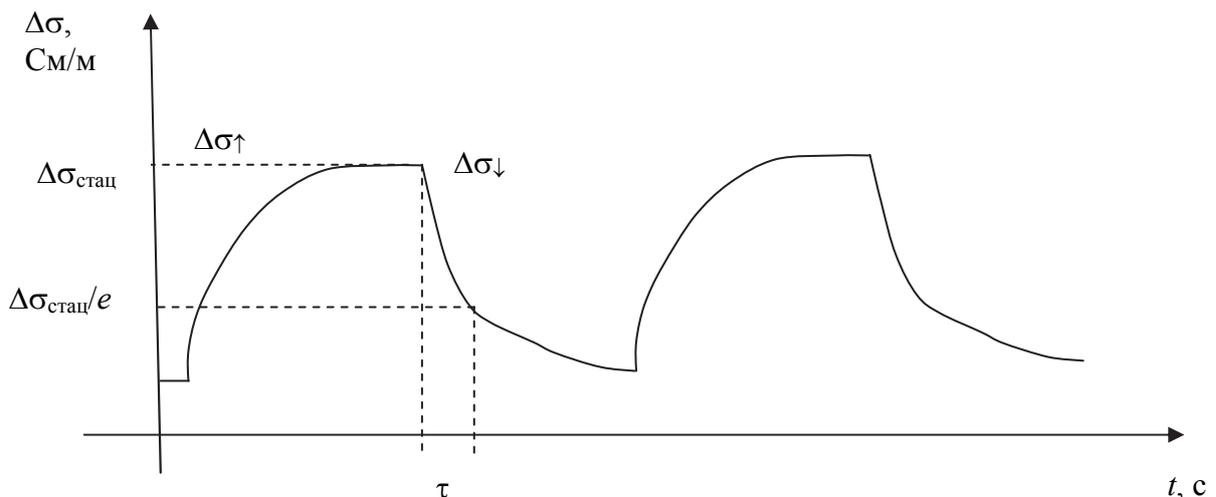


Рис. 1.16. Кривые релаксации фотопроводимости

Условием достоверного определения  $\tau$  является достаточная продолжительность облучения, обеспечивающая достижение неизменяющегося значения  $\Delta\sigma_{\text{стац}}$ . В таком случае для момента времени  $t = \tau$  в соответствии с уравнением спада фотопроводимости (1.38) текущее значение фотопроводимости будет равно  $\Delta\sigma = \Delta\sigma_{\text{стац}}/e$ . Следовательно, определив текущий момент времени, для которого  $\Delta\sigma_{\text{стац}}$  уменьшилось в  $e$  раз, можно графически на кривой спада фотопроводимости определить соответствующий временной отрезок  $\tau$ , начиная от момента прекращения облучения (рис. 1.16).

Величина  $\tau$  решающим образом влияет на быстроедействие детекторов оптического излучения.

### 1.3.4. Дрейфовый и диффузионный ток

Для последующего рассмотрения электронных процессов в функциональных структурах полупроводниковых приборов необходимо обсудить еще понятия «дрейфового» и «диффузионного» тока.

*Дрейфовый* ток возникает в материале с однородным распределением носителей заряда, например в металле, при приложении внешней электродвижущей силы в виде электрического поля напряженностью  $E$  (В/м), т. е. напряжения. Плотность дрейфового тока  $J_E$  (А/м<sup>2</sup>) была описана еще в модели Друде в виде выражения (1.39):

$$J_E = \sigma E. \quad (1.39)$$

В полупроводнике, где концентрация носителей заряда существенно зависит от ряда факторов, например температуры, облучения, концентрации примесей, на практике наблюдается или специально создается неравномерное распределение носителей заряда, приводящее к возникновению диффузионного тока, протекающего в направлении уменьшения концентрации носителей. Плотность диффузионного тока можно описать на основании 1-го закона Фика с учетом коэффициента диффузии соответствующих носителей заряда  $D_{e,h}$  (м<sup>2</sup>/с) выражением (1.40):

$$J_D = -eD_{e,h}\text{grad } n. \quad (1.40)$$

При этом коэффициент диффузии соответствующих носителей заряда  $D_{e,h}$  может быть определен по формуле Эйнштейна (1.41), которая отражает линейную зависимость между диффузионной характеристикой носителей  $D_{e,h}$  и их дрейфовой характеристикой в виде подвижности  $\mu_{e,h}$ :

$$D_{e,h} = \frac{kT}{e} \mu_{e,h}. \quad (1.41)$$

В условиях теплового равновесия в неоднородном полупроводнике дрейфовый  $I_E$  и диффузионный  $I_D$  токи компенсируют друг друга и результирующий ток через материал не протекает (рис. 1.17).

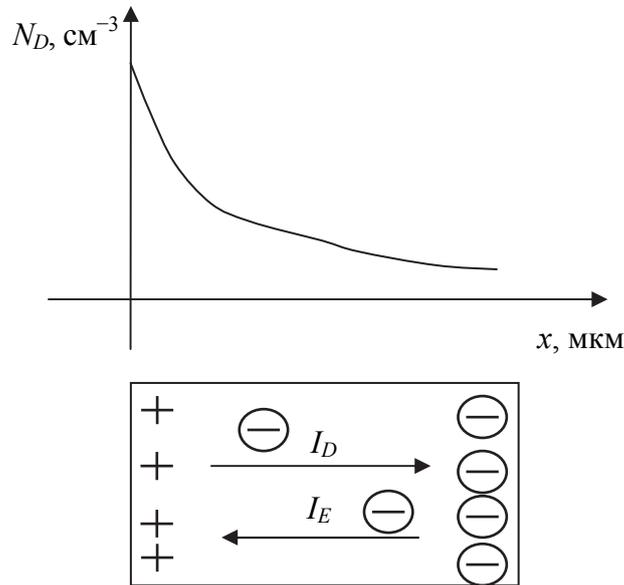


Рис. 1.17. Схема возникновения диффузионного и дрейфового токов в неоднородном полупроводнике

При описании диффузионного электропереноса в полупроводнике важное значение имеет величина диффузионной длины носителей  $L_{e,h}$ , которая равна расстоянию, преодолеваемому носителем посредством диффузии до момента рекомбинации.

Эта величина в соответствии с уравнением (1.42) прямо пропорциональна времени жизни носителя  $\tau_{e,h}$  и его коэффициенту диффузии:

$$L_{e,h} = \sqrt{D_{e,h} \tau_{e,h}}. \quad (1.42)$$

Подводя итог обсуждению особенностей реальных полупроводников, отметим, что на практике функциональными материалами приборов являются полупроводники, во-первых, с различной структурой энергетических зон, включая двудольные и непрямозонные полупроводники, во-вторых, с наличием большого числа зачастую неравномерно распределенных примесей и структурных дефектов, в-третьих, с носителями разных знаков (*электроны* и *дырки*), разной концентрации (*основные* и *неосновные*), в разном состоянии (*равновесные* и *неравновесные*).

## 2.1. Физика полупроводниковых структур (лекция 5)

### 2.1.1. Общая характеристика функциональных контактов

Приступая непосредственно к знакомству с важнейшими полупроводниковыми приборами, прежде всего мы обсудим в самых общих чертах функциональные структуры, которые обеспечивают реализацию функциональных электронных процессов в устройствах электроники. Дело в том, что большая часть полупроводниковых приборов основана на электронных процессах не в однородном полупроводнике, а в контактных структурах, возникающих на границе контакта различных материалов. Эти явления связаны с возникновением и изменением *потенциальных барьеров* для носителей заряда в функциональных контактах. В данном случае процесс электропереноса через контакт не описывается законом Ома в интегральном виде, а регулируется величиной потенциального барьера, который в свою очередь зависит от приложенного напряжения, освещения и других факторов.

Среди актуальных функциональных структур можно выделить 4 вида контактов:

1) ***p-n-переход***, представляющий собой контакт областей одного полупроводника с разным типом проводимости (дырочной или электронной), например контакт  $\text{Si}^{\text{IV}} + \text{B}^{\text{III}}$  (*p*-тип) /  $\text{Si}^{\text{IV}} + \text{P}^{\text{V}}$  (*n*-тип) или  $\text{Ga}^{\text{III}}\text{As}^{\text{V}} + \text{Zn}^{\text{II}}$  (*p*-тип) /  $\text{Ga}^{\text{III}}\text{As}^{\text{V}} + \text{Te}^{\text{VI}}$  (*n*-тип);

2) контакт **металл – полупроводник**, или контакт Шоттки, который может создаваться с участием полупроводника как *p*-, так и *n*-типа, например контакт  $\text{Mo} / \text{Si}^{\text{IV}} + \text{B}^{\text{III}}$  (*p*-тип) или  $\text{W} / \text{Si}^{\text{IV}} + \text{P}^{\text{V}}$  (*n*-тип);

3) **гетеропереход**, представляющий собой контакт полупроводников с различной шириной запрещенной зоны  $E_{g1} \neq E_{g2}$  одного типа проводимости, например *изотипный P-p*-гетеропереход  $\text{Si}^{\text{IV}} + \text{B}^{\text{III}}$  (*P*-тип) /  $\text{Ge}^{\text{IV}} + \text{Ga}^{\text{III}}$  (*p*-тип) ( $E_{g\text{Si}} > E_{g\text{Ge}}$ ) или *анизотипный p-N*-гетеропереход  $\text{Ga}^{\text{III}}\text{As}^{\text{V}} + \text{Zn}^{\text{II}}$  (*p*-тип) /  $\text{Ga}^{\text{III}}_{1-x}\text{Al}^{\text{III}}_x\text{As}^{\text{V}} + \text{Te}^{\text{VI}}$  (*N*-тип) ( $E_{g\text{GaAs}} \ll E_{g\text{Ga}_{1-x}\text{Al}_x\text{As}}$ );

4) **МДП-структура**, представляющая собой контакт трех материалов *металл – диэлектрик – полупроводник*, в котором роль диэлектрика чаще выполняет оксид, например МОП-структура состава  $\text{Al} / \text{SiO}_2 / \text{Si}^{\text{IV}} + \text{P}^{\text{V}}$  (*n*-тип).

Изначально полупроводниковые приборы развивались на основе функциональных  $p$ - $n$ -переходов, на которых основаны выпрямительные (рис. 2.1, *а*) и туннельные диоды, стабилитроны, тиристоры, биполярные транзисторы, фотодиоды, светодиоды, фотоэлементы. Контакт металл/полупроводник послужил для создания диода Шоттки (рис. 2.1, *б*), а в настоящее время находит применение в рамках совершенствования ряда диодных и триодных приборов. Гетеропереходы являются авангардным направлением развития полупроводниковой микро- и нанoeлектроники, поскольку позволяют улучшить показатели традиционных микро- и оптоэлектронных приборов (рис. 2.1, *в*), а также решить виртуозные задачи создания нанoeлектронных устройств нового поколения. За разработку тематики гетеропереходов в 2000 г. Нобелевской премии был удостоен российский физик, родившийся в БССР, Жорес Иванович Алферов (1930–2019). Разработка МОП-структуры позволила создать бюджетные субмикронные полевые транзисторы (рис. 2.1, *г*).

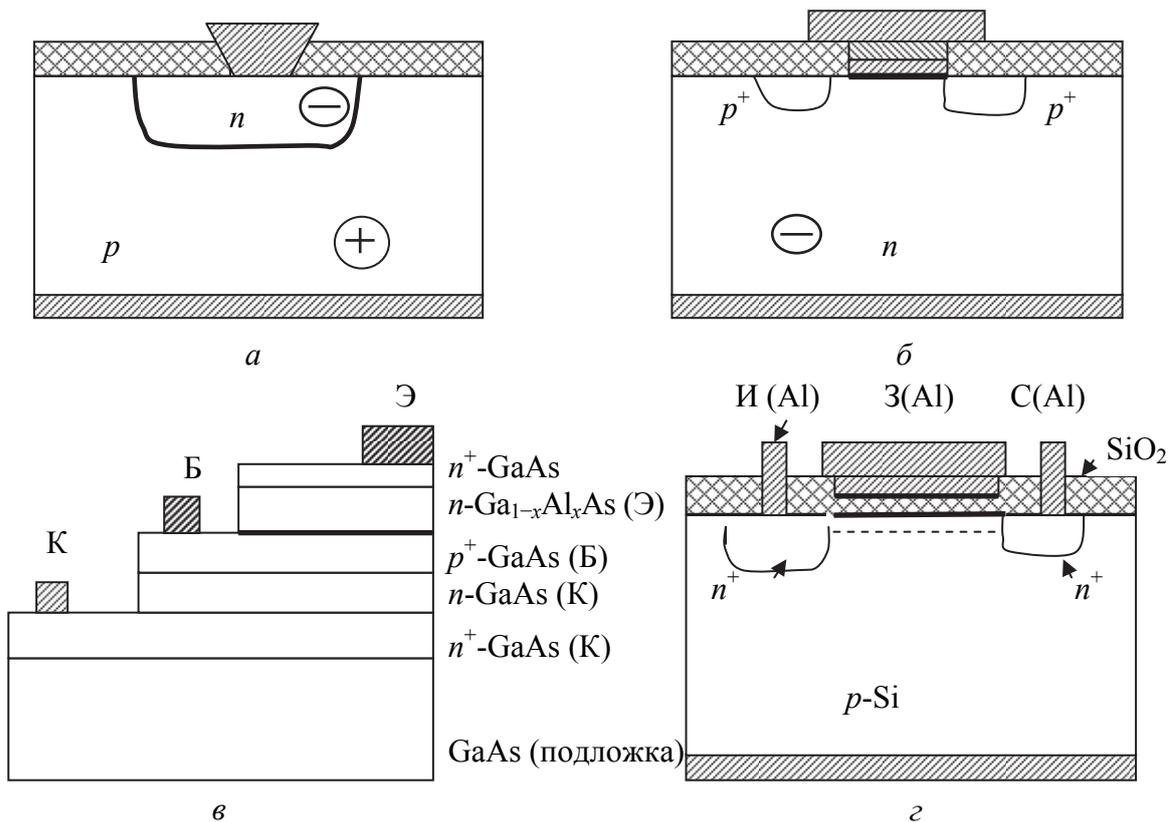


Рис. 2.1. Принципиальные схемы функциональных структур выпрямительного диода (*а*), диода Шоттки (*б*), гетеробиполярного транзистора (*в*), МОП-транзистора (*г*)

## 2.1.2. Электроника *p-n*-перехода

При всей альтернативности функциональных контактов *p-n*-переход все еще сохраняет свое доминирующее положение среди контактных структур полупроводниковых приборов, поэтому электронные процессы в *p-n*-переходе мы рассмотрим отдельно от конкретных приборов.

При рассмотрении *p-n*-перехода важно выделять три отдельных состояния: *равновесие*, *прямое смещение*, *обратное смещение*.

Сначала разберемся, что происходит на границе контакта полупроводников *p*- и *n*-типа без приложения внешнего электрического поля в *равновесии*. В области контакта полупроводников с электронным и дырочным типом проводимости возникает *градиент концентрации* носителей разного знака и соответственно происходит процесс *диффузии* основных носителей через границу контакта (рис. 2.2).

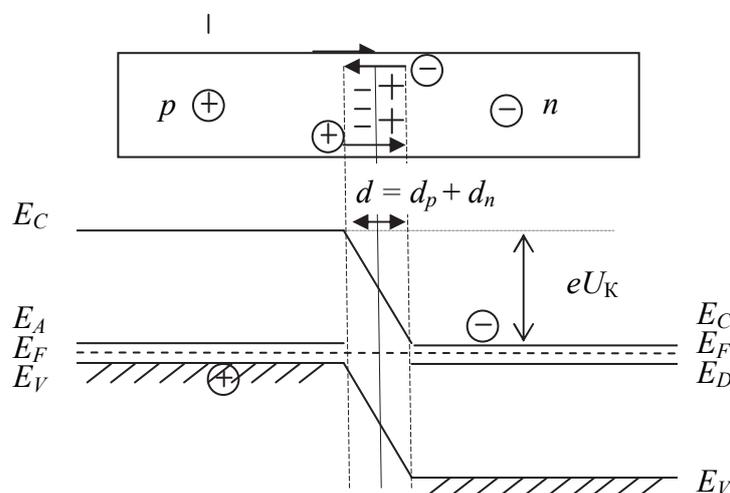


Рис. 2.2. Энергетическая диаграмма равновесного *p-n*-перехода

В результате удаления подвижных носителей из соответствующей области полупроводника вблизи границы контакта остаются электрически *нескомпенсированные неподвижные ионизованные примесные атомы*, которые формируют области пространственного заряда (ОПЗ), прилегающие к границе контакта. В пределах ОПЗ энергия электронов не соответствует энергии в объеме материала. На энергетических диаграммах это изменение энергии отражается *искривлением зон*, которое приводит к возникновению *потенциального барьера* для движения определенных носителей. Контактное поле, возникающее между противоположно заряженными ОПЗ, препятствует дальнейшей диффузии носителей через *p-n*-переход. В равновесном состоянии на границе *p-n*-перехода устанавливается постоянное значение *контактной разности потенциалов*  $U_K$ . Таким

образом, мы замкнули логическую цепочку, описывающую процессы в равновесном  $p$ - $n$ -переходе: *контакт областей*  $\rightarrow$  *градиент концентрации носителей*  $\rightarrow$  *диффузия носителей через контакт*  $\rightarrow$  *образование ОПЗ* *вблизи контакта*  $\rightarrow$  *искривление зон*  $\rightarrow$  *потенциальные барьеры для носителей*  $\rightarrow$  *контактная разность потенциалов*.

При построении энергетической диаграммы  $p$ - $n$ -перехода удобным ориентиром является уровень Ферми  $E_F$ . В случае равновесия  $E_F$  имеет одинаковое значение для всех точек диаграммы (рис. 2.2). Относительно  $E_F$  изображают уровни разрешенных зон и примесей за пределами ОПЗ. Внутри ОПЗ характер искривления зон определяется типом ОПЗ. Различают три типа ОПЗ: *обедненный* ( $n_{\text{ОПЗ}} < n_V$ ), *обогащенный* ( $n_{\text{ОПЗ}} > n_V$ ) и *инверсный*, в котором происходит инверсия знака основных носителей заряда по сравнению с объемом полупроводника. В пределах **обедненного** ОПЗ, образующегося в  $p$ - $n$ -переходах, край зоны, в которой находятся основные носители заряда, искривляется **от** уровня Ферми (на рис. 2.2  $E_C$  в  $n$ -области и  $E_V$  в  $p$ -области искривляются от  $E_F$ ). Для **обогащенного** ОПЗ искривление происходит **к** уровню Ферми. Инверсия знака основных носителей сопровождается смещением  $E_F$  в пределах запрещенной зоны полупроводника.

При прямом смещении  $p$ - $n$ -перехода (рис. 2.3) полярность приложенного напряжения совпадает с полярностью основных носителей, но не совпадает с полярностью образовавшихся ОПЗ. Это означает, что внешнее электрическое поле компенсирует контактное поле.

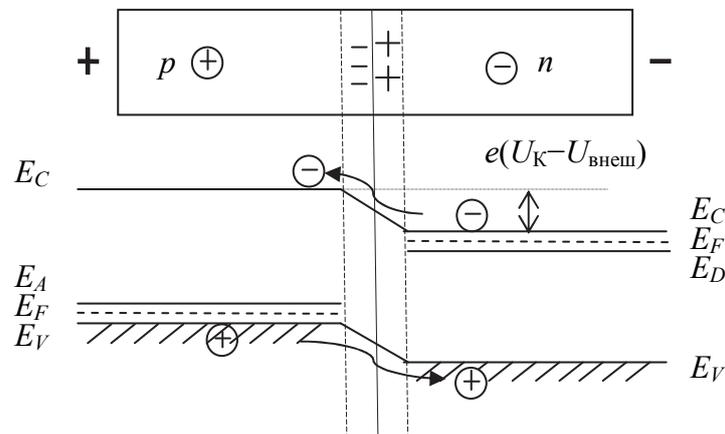


Рис. 2.3. Энергетическая диаграмма  $p$ - $n$ -перехода при прямом смещении

В результате величина искривления зон, а следовательно, и потенциальный барьер в  $p$ - $n$ -переходе уменьшаются. Другими словами,

прямосмещенный  $p$ - $n$ -переход находится в низкоомном состоянии, через него протекают большие токи основных носителей заряда.

При построении энергетической диаграммы смещенного  $p$ - $n$ -перехода, т. е. в случае приложения внешнего электрического поля, удобно сначала изобразить диаграмму одной из областей, например левой, за пределами ОПЗ. Далее при изображении другой области, например правой, необходимо сместить уровень  $E_F$  в соответствии с направлением искривления зон. В частности, при прямом смещении величина искривления зон, как указано выше, уменьшается, потенциальный барьер по сравнению с состоянием равновесия (рис. 2.2) уменьшается от значения  $eU_K$  до  $e(U_K - U_{\text{внеш}})$ , следовательно, для правой области на рис. 2.3 уровень  $E_F$  смещается вверх относительно левой области.

При обратном смещении  $p$ - $n$ -перехода (рис. 2.4) полярность приложенного напряжения не совпадает с полярностью основных носителей в соответствующих областях, но совпадает с полярностью ОПЗ. Это означает, что внешнее электрическое поле «усиливает» действие контактного поля. В результате величина искривления зон, а следовательно, и потенциальный барьер увеличиваются, но только для основных носителей.

Таким образом, обратносмещенный  $p$ - $n$ -переход находится в высокоомном состоянии, через него протекают малые токи неосновных носителей заряда (рис. 2.4).

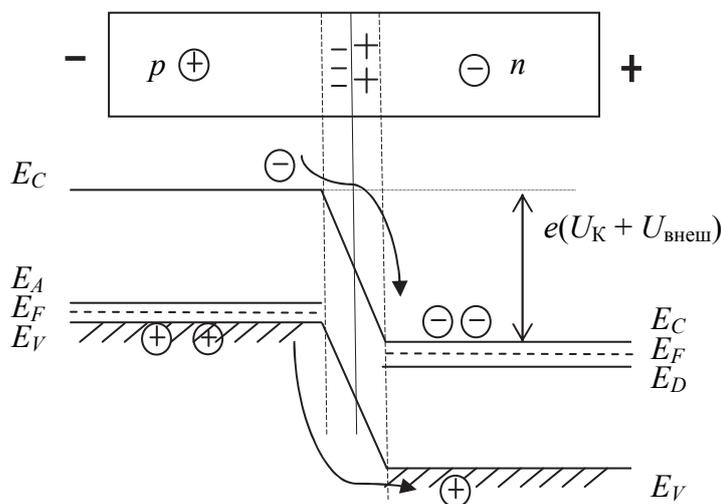


Рис. 2.4. Энергетическая диаграмма  $p$ - $n$ -перехода при обратном смещении

Обсудив качественное состояние  $p$ - $n$ -перехода для случаев равновесия, прямого и обратного смещения, приступим к знакомству с количественными выражениями. В первую очередь для описания  $p$ - $n$ -перехода необходимо выражение для расчета величины контактной

разности потенциалов, которая существенно влияет на высоту потенциального барьера, управляющего электропереносом через  $p$ - $n$ -переход. Значение  $U_K$  можно получить на основании закона сохранения заряда, в соответствии с которым заряды противоположных ОПЗ с толщиной  $d_p$  и  $d_n$  компенсируют друг друга. Отсюда вытекает соотношение (2.1), по которому толщины слоев ОПЗ  $d_p$ ,  $d_n$  обратно пропорциональны концентрациям электроактивных примесей  $N_A$ ,  $N_D$ , т. е. в сильнолегированных полупроводниках толщины ОПЗ невелики:

$$N_A d_p = N_D d_n. \quad (2.1)$$

Вид количественного выражения для  $U_K$  зависит от типа  $p$ - $n$ -перехода. Рассмотрим две классификации  $p$ - $n$ -переходов:

1) по соотношению величины концентраций основных носителей заряда (*симметричный*  $n_e^n = n_h^p$  и *асимметричный*  $n_e^n > n_h^p$ ,  $n_e^n < n_h^p$ );

2) по соотношению толщины  $p$ - $n$ -перехода  $d_{p-n}$  и градиентной или переходной области на границе контакта материалов  $d_{\Delta N}$  (*резкий* ( $d_{\Delta N} < d_{p-n}$ ) и *плавный* ( $d_{\Delta N} > d_{p-n}$ )).

В реальных структурах наблюдаются обычно асимметричные предпочтительно резкие  $p$ - $n$ -переходы. Для резкого  $p$ - $n$ -перехода выражение для  $U_K$  имеет вид (2.2):

$$U_K = \frac{e}{2\varepsilon\varepsilon_0} \frac{N_A N_D}{N_A + N_D} d_{p-n}^2, \quad (2.2)$$

где  $\varepsilon$  – относительная диэлектрическая проницаемость материала. Величина  $U_K$  зависит от  $E_g$  и на практике составляет около 1 В.

При смещении  $p$ - $n$ -перехода через него протекает ток, который описывается уравнением Шокли, или уравнением выпрямителя, при прямом ( $U > 0$ ) или обратном ( $U < 0$ ) смещении в виде (2.3):

$$I = I_0 \left[ \exp \frac{eU}{kT} - 1 \right], \quad (2.3)$$

где  $I_0$  – обратный ток насыщения  $p$ - $n$ -перехода, т. е. максимальный ток, протекающий при обратном смещении, А.

Еще одной важной особенностью  $p$ - $n$ -перехода является возникновение *диффузионной емкости* при прямом смещении и *барьерной емкости*  $C_B$  (Ф) при обратном смещении, которая для случая резкого  $p$ - $n$ -перехода имеет вид (2.4):

$$C_B = \sqrt{\frac{e\varepsilon\varepsilon_0 S^2}{2(U_K - U)} \frac{N_A N_D}{N_A + N_D}}, \quad (2.4)$$

где  $S$  – площадь контакта областей в  $p$ - $n$ -переходе,  $\text{м}^2$ ;  $U$  – напряжение смещения  $p$ - $n$ -перехода (в данном случае отрицательное), В.

Измерение барьерной емкости  $p$ - $n$ -перехода в виде вольт-фарадных характеристик  $C_B = f(U)$  (ВФХ) позволяет определить величину  $U_K$ , а также широко используется при производстве в координатах  $1/C_B^2 = f(U_{\text{обр}})$  для проверки качества  $p$ - $n$ -перехода, который в предпочтении должен быть резким (рис. 2.5).

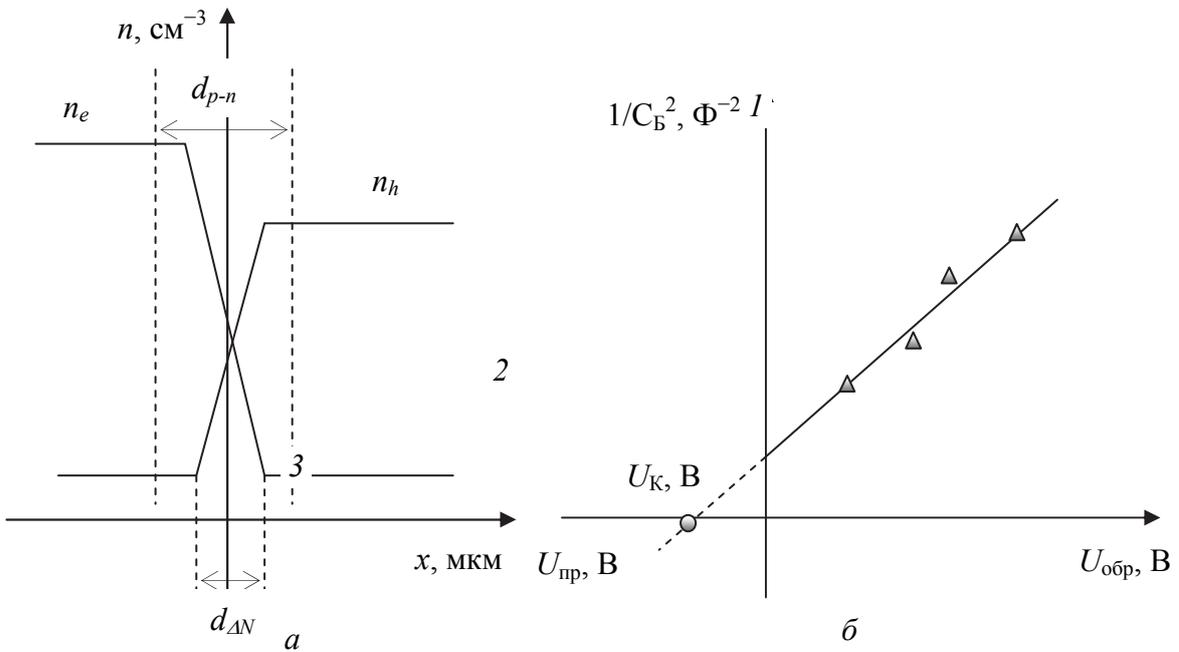


Рис. 2.5. Концентрационный профиль (а) и ВФХ (б) асимметричного ( $n_e^n > n_h^p$ ) резкого ( $d_{\Delta N} < d_{p-n}$ )  $p$ - $n$ -перехода

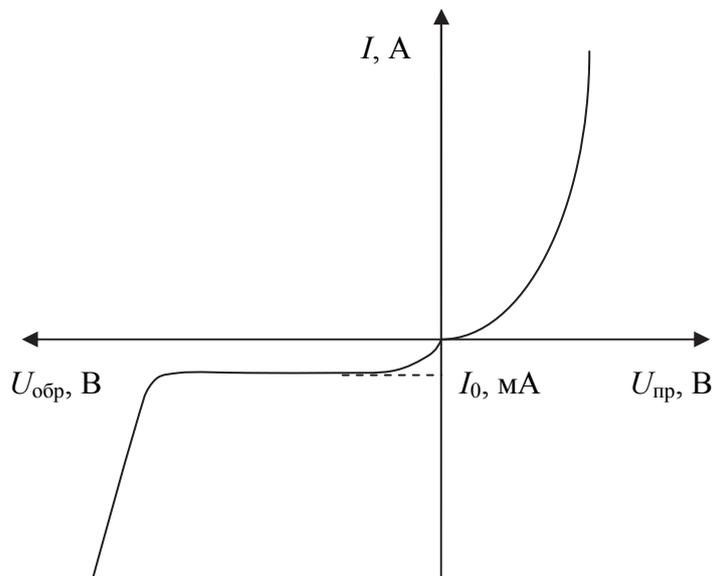


Рис. 2.6. Полная ВАХ  $p$ - $n$ -перехода

В завершение заметим, что индивидуальный  $p$ - $n$ -переход в отличие от безконтактного полупроводника обладает нелинейной (или немонотонной) *выпрямляющей* вольт-амперной характеристикой (ВАХ), которая в широком диапазоне прямых и обратных смещений включает участки прохождения тока (рис. 2.6, 1), насыщения тока (рис. 2.6, 2) и пробоя (рис. 2.6, 3).

## 2.2. Дискретные полупроводниковые диоды (лекции 6–9)

В соответствии с учебной программой дисциплины «Физическая электроника и электронные приборы» в группе диодов рассматриваются 6 микроэлектронных приборов, которые различаются типом функциональных контактов и особенностями электрических эффектов взаимодействия этих контактов с управляющими электрическими сигналами (табл. 2.1).

Таблица 2.1

Примеры функциональных структур полупроводниковых диодов

Название и обозначение прибора	Тип и состояние структуры	Эффект	Назначение
1. Выпрямительный диод	$p$ - $n$ -переход, прямое и обратное смещение	Вентильный	Выпрямление переменного сигнала
2. Стабилитрон	$p$ - $n$ -переход, обратное смещение	Лавинный или туннельный электрический пробой	Стабилизация монополярного сигнала
3. Динистор	$p_1$ - $n_1$ - $p_2$ - $n_2$ -структура		Переключение электрических цепей
4. Туннельный диод	$p^+$ - $n^+$ -переход, прямое смещение	Туннелирование электронов	Генерация СВЧ-сигналов; переключение электрических цепей
5. Диод Ганна	Двудолинный $n/n^+$ -полупроводник	Междолинный переход электронов	Генерация СВЧ-сигналов
6. Диод Шоттки	Металл/ $n$ -полупроводник, прямое и обратное смещение	Вентильный	Выпрямление переменного сигнала

Основным видом функциональных контактов в диодах является  $p$ - $n$ -переход, который в различных приборах может отличаться концентрацией и характером распределения электроактивной примеси в контактирующих областях полупроводника (например, обычный  $p$ - $n$ , вырожденный  $p^+$ - $n^+$ , аномальный  $p^+$ - $n$ -переход и др.).

Как уже было указано в предисловии, информация о каждом конкретном приборе структурирована и включает определение (1), пояснения принципа работы с учетом энергетического состояния и ВАХ функционального контакта (2), принципиальную схему функциональной структуры (3) и примеры значений выходных параметров (4).

### 2.2.1. Выпрямительный диод

1. Выпрямительный диод – это полупроводниковый прибор, основанный на *p-n-переходе*, работающий на *вентильном* эффекте, предназначенный для *выпрямления* переменного электрического сигнала.

2. Принцип работы выпрямительного диода фактически рассмотрен нами в пункте 2.1.2. Выпрямление переменного сигнала при пропускании его через *p-n-переход* вытекает из ВАХ *p-n-перехода* (рис. 2.7) и происходит благодаря изменению потенциального барьера при прямых и обратных смещениях. Эффект выпрямления называют *вентильным*, поскольку *p-n-переход* пропускает токи при прямом смещении (рис. 2.8), находясь в низкоомном состоянии ( $R_{p-n} \downarrow$ ), и практически не пропускает токи при обратном смещении, находясь в высокоомном состоянии ( $R_{p-n} \uparrow$ ).

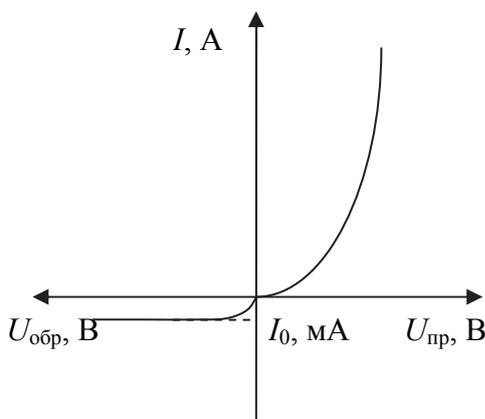


Рис. 2.7. ВАХ выпрямительного диода

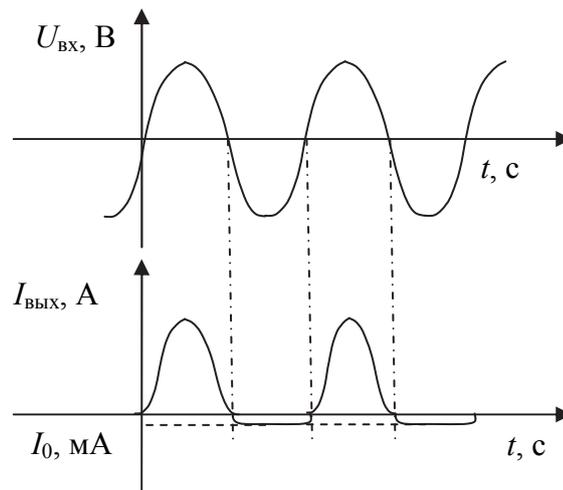


Рис. 2.8. Входной и выходной сигналы выпрямительного диода

3. Принципиальная функциональная структура выпрямительного диода представлена на рис. 2.1, а. Основным материалом для этих приборов является кремний, например в виде подложки КДБ-10 (111), поскольку необходимо обеспечить стойкость к пробоям, которой обладают более широкозонные материалы. Продолжается также выпуск германиевых диодов. При необходимости повышения быстродействия прибора возможно использование GaAs, который приблизительно в 5 раз дороже кремния.

4. Выходные параметры выпрямительного диода зависят от полупроводникового материала (табл. 2.2).

Таблица 2.2

Выходные параметры выпрямительного диода

Название параметра	Значения параметра		
	Ge	Si	GaAs
Характеристики полупроводника: $E_g$ , эВ $\mu_e$ , см <sup>2</sup> /В·с	0,65 3800	1,1 1500	1,4 8500
1. Максимальный прямой ток $I_{\text{макс}}$ , А	До 100	До 1500	До 1000
2. Напряжение $U_{\text{макс}}$ при $I_{\text{макс}}$ , В	0,5	1	1
3. Ток насыщения $I_0$ , мкА	10–100	5–20	10–100
4. Максимальное обратное напряжение $U_{\text{обр}}^{\text{макс}} = U_{\text{пр}}/2$ , В	100	1000	400
5. Максимальная рабочая температура $T_{\text{раб}}^{\text{макс}}$ , °С	80	250	400
6. Максимальная частота сигнала $f_{\text{макс}}$ , Гц	$10^4$	$10^3$	$10^6$

К достоинствам выпрямительных диодов на основе кремния можно отнести повышенную максимальную токовую нагрузку и рабочую температуру, а также пониженный ток насыщения, что обеспечивает повышение мощности прибора при улучшении качества выпрямления.

### 2.2.2. Стабилитрон

1. Стабилитрон – это полупроводниковый диод, основанный на обратносмещенном *p-n-переходе*, работающий на эффекте лавинного и/или туннельного пробоя, предназначенный для стабилизации монополярного электрического сигнала.

2. Работа стабилитрона основана на эффектах туннельного (горизонтального) или лавинного (вертикального) электрических пробоев *p-n-перехода*, заключающихся в резком увеличении тока при приложении достаточно большого обратного смещения (рис. 2.9).

Реализация конкретного вида пробоя зависит от уровня легирования функциональных областей полупроводника. В частности, при повышенной концентрации электроактивных примесей, согласно выражению (2.1), толщина *p-n-перехода* уменьшается и соответственно увеличивается вероятность туннелирования носителей (рис. 2.9, а), т. е. горизонтального перехода носителей заряда через такой *p-n-переход* без возбуждения.

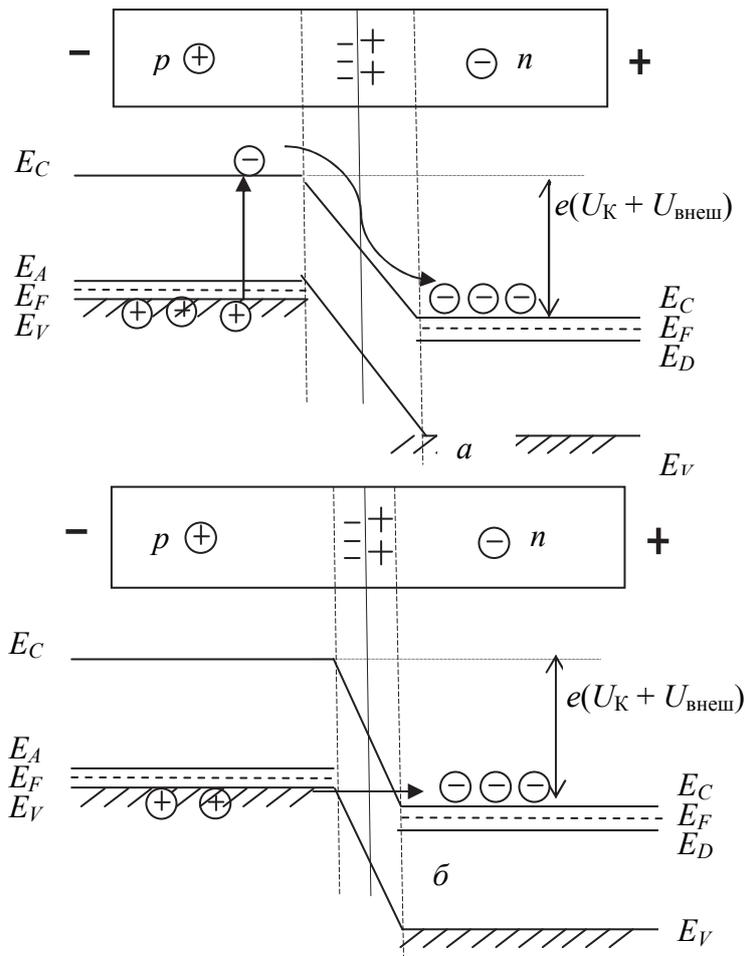


Рис. 2.9. Энергетическая диаграмма  $p$ - $n$ -перехода в состоянии туннельного (а) и лавинного (б) пробоя

В этом случае происходит туннельный пробой  $p$ - $n$ -перехода, который наблюдается при пониженных обратных смещениях  $p$ - $n$ -перехода  $U_{\text{пр1}}$  (рис. 2.10) и характеризуется пониженной крутизной нарастания тока пробоя.

Реализация лавинного пробоя достигается в полупроводниках при пониженной концентрации примесей, т. е. при повышенной толщине  $p$ - $n$ -перехода, в результате приложения достаточного большого обратного смещения, обеспечивающего лавинную генерацию электронно-дырочных пар (вертикальный переход) за счет ионизации собственных атомов полупроводника (рис. 2.9, б).

ВАХ стабилитрона фактически представляет собой часть обратной ветви полной ВАХ  $p$ - $n$ -перехода (рис. 2.6) и отличается для случаев туннельного и лавинного пробоя крутизной нарастания тока при достижении напряжения пробоя (рис. 2.10).

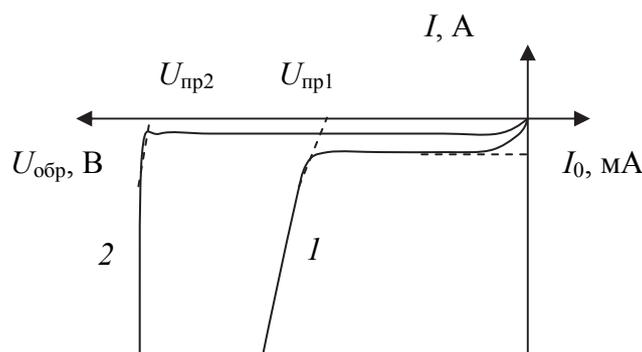


Рис. 2.10. ВАХ стабилитрона, основанного на туннельном (1) или лавинном (2) пробое

3. Принципиальная функциональная структура стабилитрона аналогична структуре выпрямительного диода, представленной на рис. 2.1, а. Доминирующим материалом для стабилитронов является структурно наиболее совершенный кремний, например в виде подложки КДБ-10 (111), поскольку необходимо обеспечить воспроизводимость по параметрам пробоя, которая существенно зависит от дефектности полупроводника. Уровень легирования электроактивными примесями для приборов с различным видом пробоя может составлять от  $10^{15}$  до  $10^{18}$  см<sup>-3</sup>.

4. Выходные параметры стабилитрона зависят от вида пробоя, реализующегося в функциональной структуре (табл. 2.3).

Таблица 2.3

**Выходные параметры стабилитронов различной мощности**

Название параметра	Значения параметра		
	$P_{\text{ном}} \downarrow$ , туннельный пробой	$P_{\text{ном}} \downarrow \uparrow$ , смешанный пробой	$P_{\text{ном}} \uparrow$ , лавинный пробой
1. Ток стабилизации $I_{\text{стаб}}$ , мА	5–20		
2. Напряжение стабилизации, В	<6	6–12	>12
3. Концентрация примесей, см <sup>-3</sup>	$>10^{17}$	$10^{15}–10^{17}$	$<10^{15}$
4. Входное сопротивление $R_{\text{вх}}$ , В	$10^3$	$10^3–10^6$	$10^6$
5. Температурный коэффициент напряжения стабилизации $\alpha_U$ , К <sup>-1</sup>	-0,01–0,00001	–	+0,01–0,00001

Ток в цепи стабилизации для всех стабилитронов составляет около 10 мА. Напряжение стабилизации, практически равное напряжению пробоя, имеет повышенные значения в случае лавинного пробоя, что обеспечивает повышенную номинальную мощность прибора  $P_{\text{ном}} = I_{\text{стаб}} U_{\text{стаб}}$ . Обратите внимание, что температурная зависимость напряжения пробоя

имеет противоположный характер для различных механизмов пробоя, что используют для идентификации вида пробоя в структуре.

### 2.2.3. Тиристоры

1. Тиристоры – группа полупроводниковых приборов, включающая *динисторы* (диоды) и *тринисторы* (триоды), основанных на трех и более последовательно включенных *p-n*-переходах, работающих на эффекте лавинного пробоя рабочего перехода, предназначенных для переключения постоянных электрических сигналов.

2. Работа тиристоров, как и стабилитронов, основана на эффекте электрического пробоя, в частности вертикального лавинного пробоя (рис. 2.9, б), который в тиристорной структуре реализуется для центрального обратносмещенного *p-n*-перехода П2 (рис. 2.11, а). Два периферийных прямосмещенных *p-n*-перехода П1 и П3 находятся в низкоомном состоянии. Таким образом, при недостаточно большом смещении тиристорной структуры в ней протекают малые токи, не превышающие ток насыщения центрального обратносмещенного *p-n*-перехода П2. При пробое перехода П2 ток достаточно резко увеличивается. Такая ситуация описывается S-образной ВАХ, ключевой точкой которой является напряжение переключения структуры из исходного высокоомного состояния в низкоомное (рис. 2.11, б). Для снижения этого напряжения вместо диодного подключения динистора используют триодное подключение тринистора, что позволяет существенно снизить энергозатраты на переключение цепей, так как  $U_T < U_D$ .

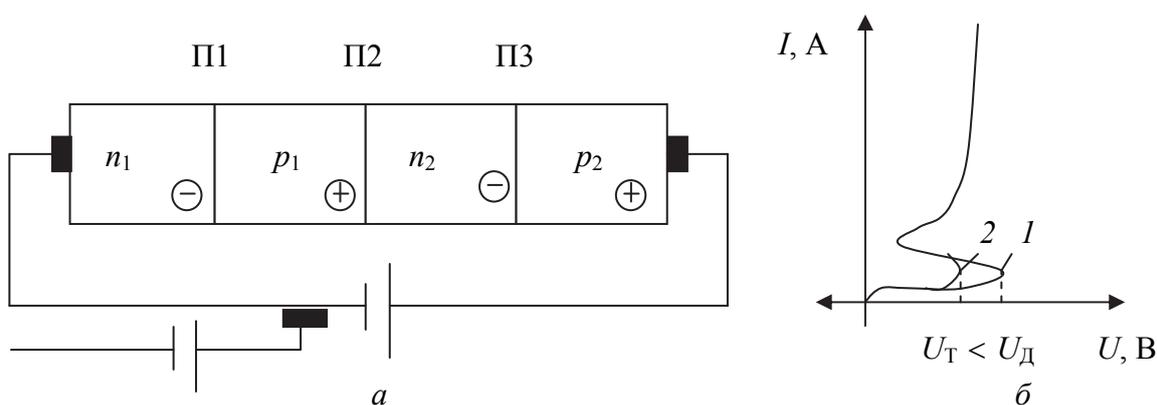


Рис. 2.11. Принципиальная схема тиристорной структуры на примере тринистора (а) и ВАХ динистора (б, 1) и тринистора (б, 2)

3. Принципиальная функциональная структура тиристоров, которая представлена на примере тринистора, включает ряд эпитаксиальных слоев

и легированные области полупроводника (рис. 2.12). Наиболее высокий уровень легирования создается в локально легированных областях. Два из трех  $p$ - $n$ -переходов тиристорной структуры формируются в рамках эпитаксиальных пленочных структур. В реальном приборе реализуется многовыводной вариант параллельного подключения на планарной стороне кристалла. Основным материалом для динисторов и тринисторов является кремний, который обеспечивает надежное переключение сигнала, в том числе достаточно мощного, но проявляет повышенную инерционность.

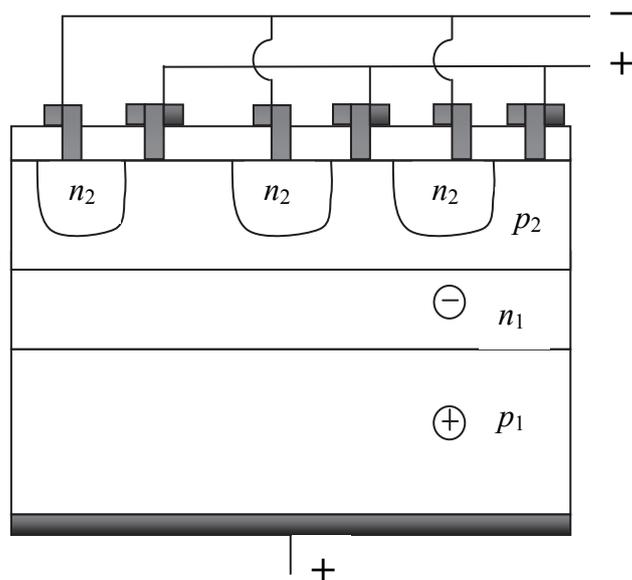


Рис. 2.12. Функциональная структура тринистора

4. К основным выходным параметрам тиристоров относят напряжение переключения  $U_T$  или  $U_D$ , составляющее величину от 1 до нескольких сотен вольт, максимальный прямой ток и соответствующее максимальное прямое напряжение (до сотен ампер и до 10 В). Время переключения цепи из закрытого в открытое состояние при подаче управляющего импульса изменяется от 1 мкс до 1 мс. Максимальная мощность тиристора может составлять величину от нескольких ватт до 100 кВт.

#### 2.2.4. Туннельный диод

1. Туннельный диод – многофункциональный полупроводниковый прибор, основанный на прямосмещенном специфическом  $p^+ - n^+$ -переходе, работающий на эффекте туннелирования электронов, предназначенный для генерации СВЧ-колебаний, переключения и усиления электрического сигнала.

2. Работа туннельного диода основана на эффекте туннелирования электронов через тонкий потенциальный барьер  $p^+-n^+$ -перехода, который создается с использованием высоколегированных (или вырожденных) полупроводников. Энергетическая диаграмма таких материалов имеет две особенности:

1) образование примесных зон в результате перекрывания примесных уровней ввиду высокой концентрации примеси (более  $10^{18} \text{ см}^{-3}$ ), перекрывание примесных и собственных зон полупроводника, в результате чего уровень Ферми смещается из запрещенной зоны в соответствующие разрешенные зоны и часть зоны проводимости оказывается заполненной, а часть валентной зоны – пустой;

2) большая высота потенциального барьера, но малая его толщина, равная толщине  $p^+-n^+$ -перехода, составляющей величину  $<1 \text{ мкм}$ , что увеличивает вероятность туннелирования электронов через такой переход.

Теперь можно обсудить энергетическое состояние и электронные процессы в  $p^+-n^+$ -переходе. В равновесном  $p^+-n^+$ -переходе (рис. 2.13) заполненные электронные состояния зоны проводимости  $n^+$ -области соответствуют по энергии заполненной части валентной зоны  $p^+$ -области, т. е. туннелирование электронов невозможно.

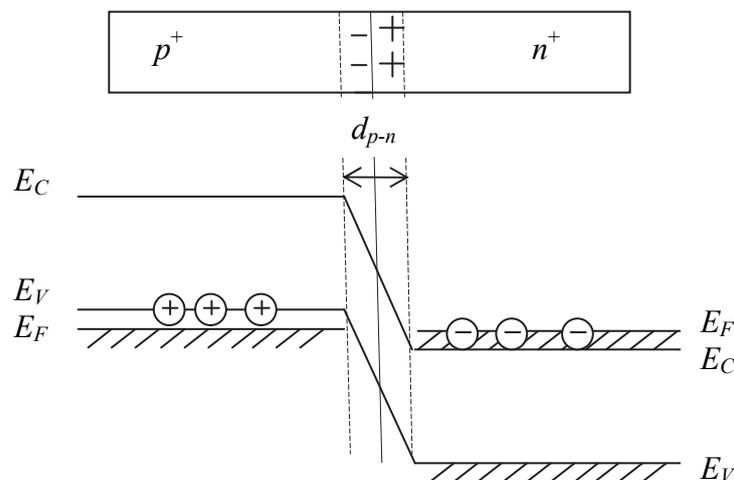


Рис. 2.13. Энергетическая диаграмма равновесного  $p^+-n^+$ -перехода

При прямом смещении  $p^+-n^+$ -перехода, сопровождающегося понижением потенциального барьера (правая часть диаграммы смещается вверх), возникает возможность туннелирования электронов, поскольку при смещении заполненные электронные состояния зоны проводимости  $n^+$ -области соответствуют по энергии незаполненной

части валентной зоны  $p^+$ -области. Нарастание потока туннелирующих электронов при увеличении напряжения прямого смещения будет продолжаться до тех пор, пока все электроны проводимости  $n^+$ -области не окажутся изоэнергетичными всем свободным уровням валентной зоны  $p^+$ -области (рис. 2.14).

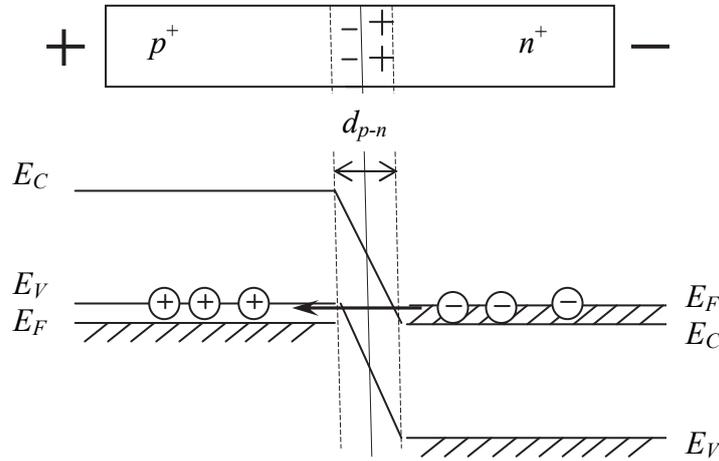


Рис. 2.14. Энергетическая диаграмма  $p^+ - n^+$ -перехода при пониженном прямом смещении

При дальнейшем увеличении напряжения прямого смещения поток туннелирующих электронов начнет уменьшаться, поскольку энергия некоторой части электронов зоны проводимости  $n^+$ -области будет соответствовать энергии запрещенных уровней, образующих запрещенную зону  $p^+$ -области (рис. 2.15).

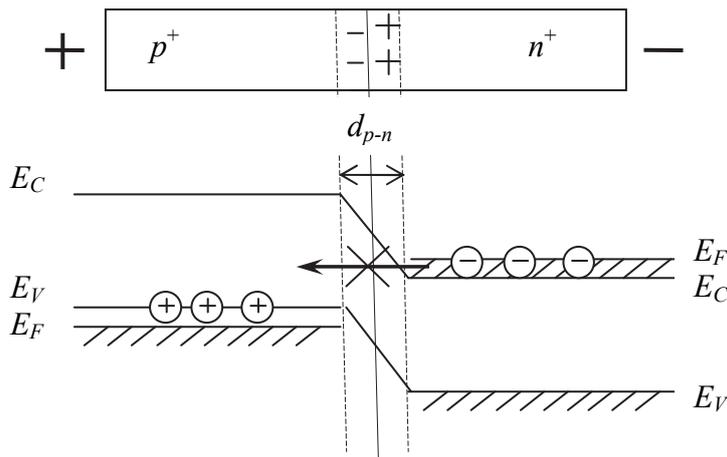


Рис. 2.15. Энергетическая диаграмма  $p^+ - n^+$ -перехода при повышенном прямом смещении

Дальнейшее увеличение прямого смещения приведет к снижению потенциального барьера на границе  $p^+ - n^+$ -перехода до значений, сопоставимых со средней тепловой энергией частиц  $3/2kT$ , что означает возможность обычного инжекционного переноса электронов из  $n^+$ -области в  $p^+$ -область (рис. 2.16).

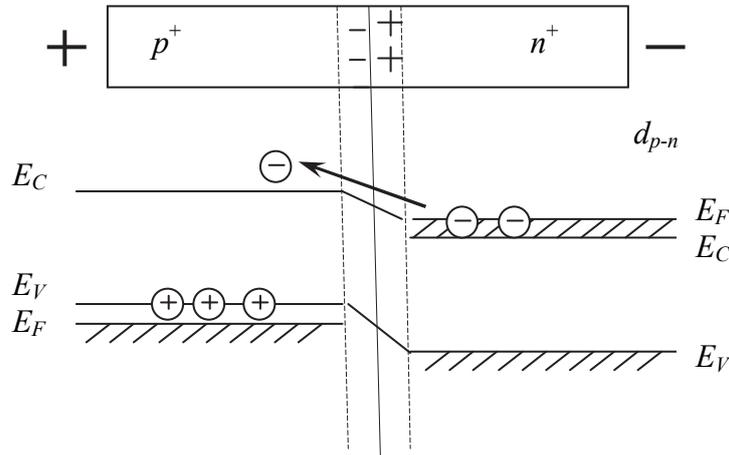


Рис. 2.16. Энергетическая диаграмма  $p^+ - n^+$ -перехода при высоком прямом смещении

Результатом рассмотренных электронных процессов в  $p^+ - n^+$ -переходе является  $N$ -образная ВАХ туннельного диода (рис. 2.17), на которой выделяют 4 участка: нарастание тока, спад тока, насыщение по току, повторное нарастание тока. Характерной особенностью всех  $N$ -образных ВАХ является наличие участка с отрицательным дифференциальным (изменяющимся) сопротивлением (рис. 2.17, участок 2), т. е. участка, на котором при увеличении напряжения ток падает. Наличие такого участка на ВАХ позволяет использовать прибор как генератор колебаний или усилитель в электрической цепи. Кроме того, возможность реализации низкоомного (максимум участка 1, рис. 2.17) и высокоомного (участок 3, рис. 2.17) состояния  $p^+ - n^+$ -перехода позволяет с помощью туннельного диода выполнять функцию переключения электрических сигналов.

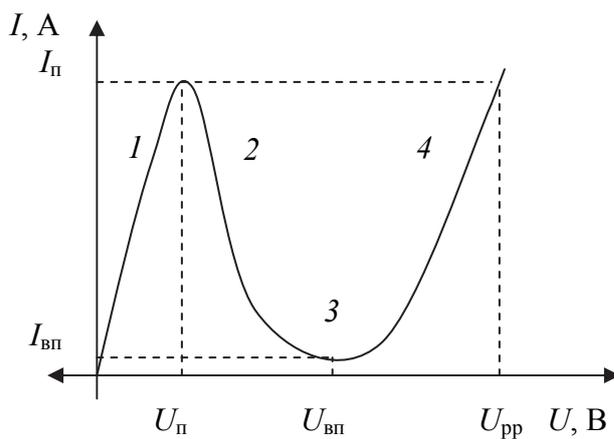


Рис. 2.17. ВАХ туннельного диода

Наличие такого участка на ВАХ позволяет использовать прибор как генератор колебаний или усилитель в электрической цепи. Кроме того, возможность реализации низкоомного (максимум участка 1, рис. 2.17) и высокоомного (участок 3, рис. 2.17) состояния  $p^+ - n^+$ -перехода позволяет с помощью туннельного диода выполнять функцию переключения электрических сигналов.

3. Поскольку прибор используется как генератор СВЧ-колебаний в электрических цепях и быстродействующий переключатель, то для его производства применяют полупроводники с повышенной подвижностью носителей заряда, например Ge, GaAs, GaSb.

В реальных приборах создается высоколегированный точечный  $p^+ - n^+$ -переход с малой площадью контакта для обеспечения высокой рабочей плотности тока (рис. 2.18). Для формирования таких контактов используется технология легирования при сплавлении, которая плохо интегрируется с современными процессами интегральной планарной технологии производства функциональных полупроводниковых кристаллов.

4. Основные выходные параметры туннельного диода приведены в табл. 2.4. Рабочим участком ВАХ является участок с отрицательным дифференциальным сопротивлением (рис. 2.17, 2). К достоинствам этого прибора, кроме многофункциональности, относят высокое быстродействие и высокую плотность рабочего тока, а также термическую стабильность (ввиду вырожденного состояния полупроводников) и пониженный уровень шумов. Среди недостатков, кроме сплавной технологии, выделяют невысокую мощность прибора, обусловленную небольшими значениями рабочих токов и напряжений.

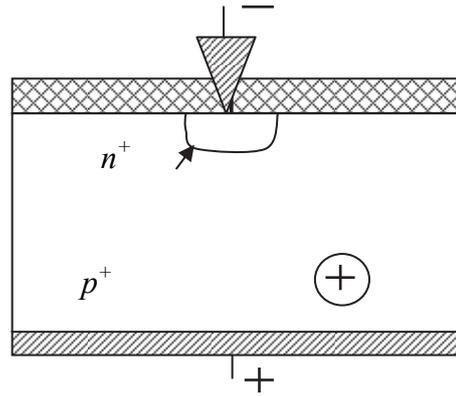


Рис. 2.18. Функциональная структура туннельного диода

Таблица 2.4

**Выходные параметры туннельного диода**

Название параметра	Значения параметра	
	Ge	GaAs
Характеристики полупроводника:		
$E_g$ , эВ	0,65	1,4
$\mu_e$ , см <sup>2</sup> /В·с	3800	8500
1. Ток пика $I_p$ , А	До 100	До 1
2. Напряжение пика $U_p$ , В	0,05	0,15
3. Ток впадины $I_{вп}$ , А	10–100	10–100
4. Напряжение впадины $U_{вп}$ , В	0,2	0,5
5. Отношение токов $I_p/I_{вп}$	<5	>10
6. Напряжение раствора $U_{рр}$ , В	<1	>1
7. Максимальная частота рабочего сигнала $f_{\max}$ , Гц	$10^9$	$10^{11}$

## 2.2.5. Диод Ганна

1. Диод Ганна – полупроводниковый прибор, основанный на использовании двудолинного полупроводника, работающий на эффекте междолинного перехода электрона (эффекте Ганна), предназначенный для генерации СВЧ-колебаний в электрической цепи.

2. В диоде Ганна не используются функциональные контакты. Его работа основана на особенностях энергетического состояния электронов в зоне проводимости двудолинного полупроводника. Напомним, что в этих долинах отличается эффективная масса, а значит, и подвижность электрона, а именно  $\mu_1 > \mu_2$ .

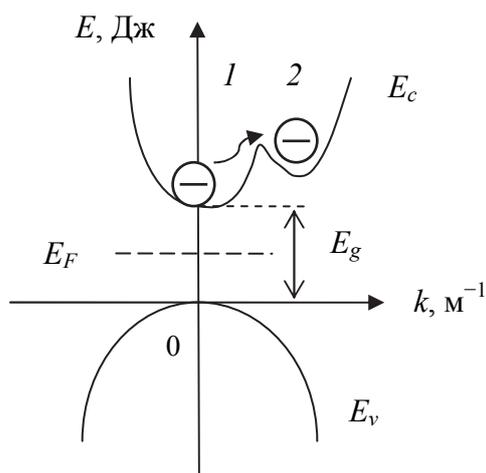


Рис. 2.19. Схема междолинного перехода электрона в двудолинном полупроводнике

В таком случае электрон, находясь в двудолинной зоне проводимости, при приложении электрического поля может участвовать в трех процессах (рис. 2.19): дрейф в пределах долины 1, междолинный переход  $1 \rightarrow 2$ , дрейф в пределах долины 2. В результате двудолинный полупроводник без наличия функциональных контактов с управляемыми барьерами обладает N-образной ВАХ, аналогичной туннельному диоду (рис. 2.20), позволяющей использовать прибор на основе такого материала в качестве генератора колебаний в электрических цепях.

ВАХ диода Ганна практически воспроизводит экстремальную зависимость дрейфовой скорости электронов в двудолинном материале от напряженности приложенного электрического поля  $v_D = f(E)$ . Наличие на этой зависимости убывающего участка, который соответствует диапазону напряжений от  $U_1$  до  $U_2$  (рис. 2.20), приводит при указанных напряжениях к формированию в материале особенной области, которая называется «электрический домен» и представляет собой область скопления замедляемых (а не ускоряемых) полем электронов. Зарождение этой

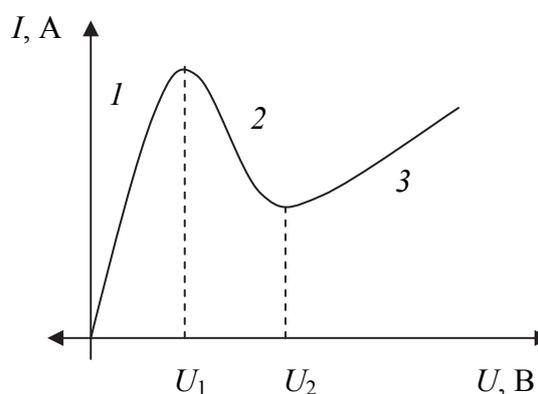


Рис. 2.20. ВАХ диода Ганна

области начинается с обычной флуктуации распределения электронов, которая разрастается благодаря тому, что электроны попадают в такой электрический домен и не могут его покинуть, замедляясь все больше за счет увеличения напряженности поля в растущем домене и пропорционального уменьшения дрейфовой скорости электронов. Интересно, что второй домен в материале не образуется, поскольку напряженность поля за пределами домена снижается по мере роста первого домена. В результате электроперенос в функциональном слое полупроводника представляет собой процесс движения разрастающегося электрического домена от отрицательного к положительному полюсу прибора. При выходе домена во внешнюю цепь возникает импульс тока, иными словами, при приложении к диоду Ганна постоянного напряжения (в диапазоне от  $U_1$  до  $U_2$ ) на выходе получаем импульсный ток, частота генерации которого определяется временем  $\tau_\phi$  прохождения домена с некоторой дрейфовой скоростью через функциональный слой толщиной  $d_\phi$ . Учитывая количественную связь между дрейфовой скоростью  $v_D$  и подвижностью  $\mu$  носителей заряда (уравнение (1.4)), можно получить выражение для расчета частоты выходного сигнала  $f_\Gamma$  в диоде Ганна (2.5):

$$f_\Gamma = \frac{1}{\tau_\phi} = \frac{\mu_e U_{\text{вх}}}{d_\phi d_{\text{подл}}}, \quad (2.5)$$

где  $d_{\text{подл}}$  – общая толщина полупроводниковой подложки, м.

3. Структура диода Ганна (рис. 2.21) включает монокристаллическую низкоомную подложку на основе двудолинного полупроводника, например GaAs или InP, с функциональным эпитаксиальным слоем, толщина которого  $d_\phi$  определяет частоту выходного сигнала прибора.

4. Основными выходными параметрами диода Ганна являются рабочий диапазон управляющих напряжений от  $U_1$  до  $U_2$ , составляющий от 10 до 100 В, амплитуда выходного тока, достигающая 100 А, а также рабочая частота генерируемого импульсного сигнала, которая может достигать десятков гигагерц, т. е. величины порядка  $10^{10}$  Гц, что соответствует, например, частоте работы радара.

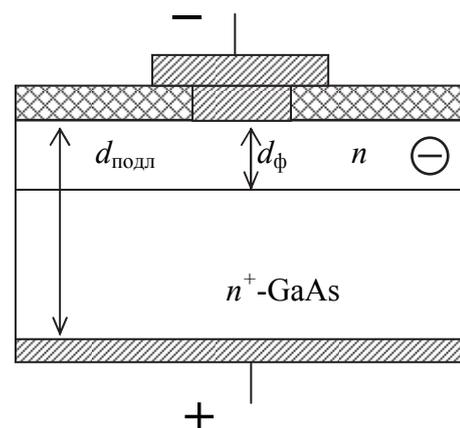


Рис. 2.21. Функциональная структура диода Ганна

## 2.2.6. Диод Шоттки

1. Диод Шоттки – полупроводниковый прибор, основанный на контакте металл – полупроводник, работающий на вентильном эффекте, предназначенный для выпрямления переменного сигнала повышенной частоты и мощности.

2. Функциональный контакт диода Шоттки, выполняющего функцию, аналогичную выпрямительному диоду на  $p-n$ -переходе, имеет две основные особенности в сравнении с последним:

1) ОПЗ образуется только с одной стороны от контакта, а именно в полупроводнике;

2) потенциальные барьеры для электронов и дырок имеют различную величину.

Рассмотрим для начала равновесный контакт металл – полупроводник. Для определения вида ОПЗ, образующегося в полупроводнике, необходимо учесть, во-первых, соотношение работы выхода электрона из контактирующих материалов, поскольку электроны переходят из материала с пониженной работой выхода в материал с повышенной работой выхода. На практике в качестве контактного металла на кремниевых пластинах чаще используют молибден и вольфрам. Кремний обладает пониженной работой выхода по сравнению с этими металлами, поэтому на контакте будет происходить переход электронов проводимости из кремния в молибден или вольфрам. Во-вторых, следует учесть тип проводимости полупроводника. В реальных структурах используют  $n$ -Si, что позволяет повышать частоту рабочего сигнала с учетом повышенной подвижности электронов проводимости по сравнению с дырками. В результате на контакте Шоттки, например в варианте Mo/ $n$ -Si, со стороны полупроводника образуется ОПЗ обедненного типа. В области  $n$ -Si потенциальный барьер для электронов проводимости будет равен величине искривления зон, а для дырок, неосновных носителей в  $n$ -Si, будет значительно больше.

При прямом смещении контакта металл – полупроводник (рис. 2.22), как и в  $p-n$ -переходе, искривление зон уменьшается, что приводит для основных носителей заряда к понижению потенциального барьера до значений, близких к тепловой энергии. Для неосновных носителей барьер тоже понижается, но остается непреодолимым для большей части этих носителей. В результате через контакт Шоттки протекают большие токи основных носителей заряда за счет их *односторонней* инжекции, а инжекции неосновных носителей не происходит.

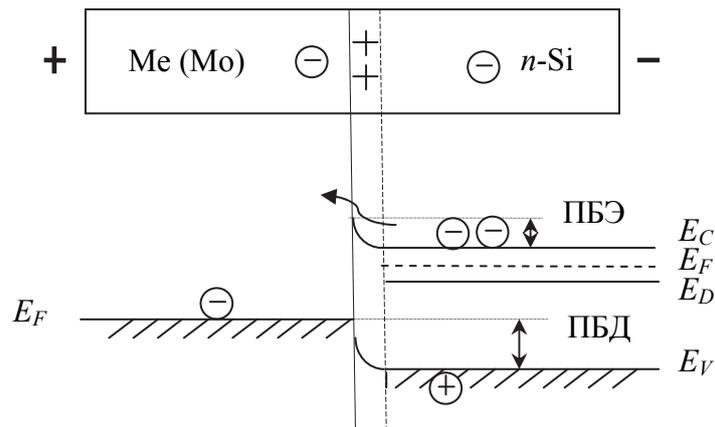


Рис. 2.22. Энергетическая диаграмма контакта Шоттки при прямом смещении

ВАХ диода Шоттки имеет выпрямляющий вид (рис. 2.23) и характеризуется в сравнении с ВАХ выпрямительного диода повышенными значениями как прямых (достоинство), так и обратных (недостаток) токов.

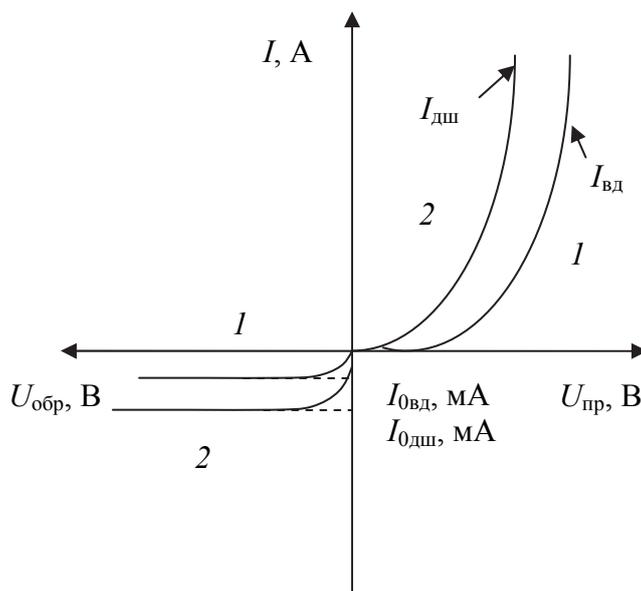


Рис. 2.23. ВАХ выпрямительного диода (1) и диода Шоттки (2)

Важной положительной особенностью ВАХ диода Шоттки является то, что эта зависимость подчиняется уравнению Шокли (2.3) практически без дополнительных эмпирических коэффициентов, что

позволяет различить выпрямительные приборы на основе  $p$ - $n$ -перехода и контакта Шоттки.

3. Структура диода Шоттки (рис. 2.24) включает функциональный контакт металл – полупроводник, который формируется обычно в высокоомном эпитаксиальном слое, выращенном на низкоомной пластине того же типа проводимости. На сегодняшний день диоды Шоттки производят не только на основе кремния, но и на основе SiC, который обладает в сравнении с кремнием более широкой запрещенной зоной, более высокой теплопроводностью, термо- и радиостойкостью. Технология производства диодов Шоттки на SiC гораздо сложнее, чем на традиционном кремнии в связи с высокой тугоплавкостью, твердостью и химической стойкостью SiC.

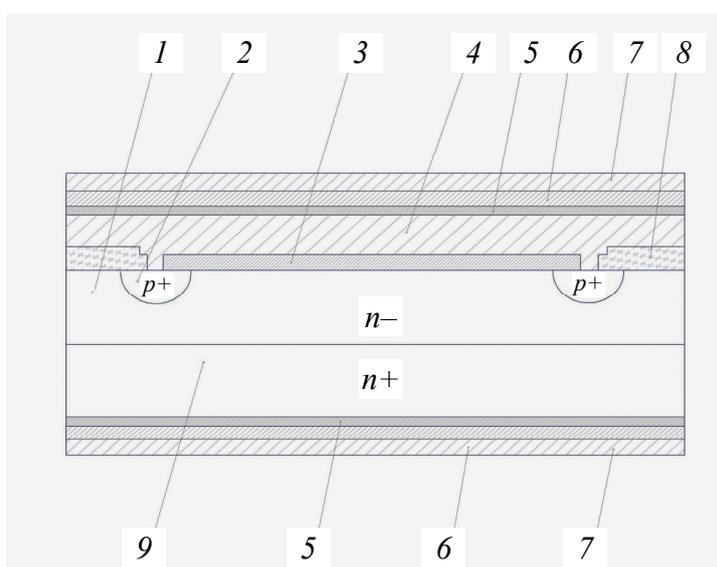


Рис. 2.24. Структура диода Шоттки:

- 1 – эпитаксиальный слой ( $n$ -Si);
- 2 – охранный слой; 3 – контактный слой (Mo, W);
- 4 – буферный слой (Al + Si);
- 5 – адгезионный слой (Ti);
- 6 – монтажный слой (Ni + V);
- 7 – барьерный слой (Ag);
- 8 – изолирующий слой ( $\text{SiO}_2$ );
- 9 – подложка ( $n^+$ -Si)

4. Основные выходные параметры диодов Шоттки зависят от химической природы функционального полупроводника. В целом, по сравнению с выпрямительными диодами на  $p$ - $n$ -переходе, они позволяют выпрямлять более мощный сигнал повышенной частоты и достигать рабочих температур порядка  $600^\circ\text{C}$  (в случае SiC).

## 2.3. Дискретные транзисторы (лекции 10–11)

*Транзисторы* – это полупроводниковые триоды, основанные на различных функциональных контактах, работающие на эффектах перераспределения энергии между управляющим постоянным и рабочим переменным электрическим полем, предназначенные для усиления переменного электрического сигнала или выполнения логических операций. Основной классификацией транзисторов является их разделение на *биполярные* (БТ) и *полевые* (ПТ) транзисторы (табл. 2.5).

Таблица 2.5

**Сравнительная характеристика биполярных и полевых транзисторов**

Биполярный транзистор	Полевой транзистор
<i>Управляющий сигнал</i>	
Ток базы $I_B$	Напряжение на затворе $U_3$
<i>Участие носителей в функциональном электропереносе</i>	
И электроны, и дырки (биполярная проводимость в транзисторной структуре), <i>n-p-n, p-n-p</i>	Или электроны, или дырки (моно- или униполярная проводимость в канале), <i>n</i> -канальный, <i>p</i> -канальный
<i>Название выводов</i>	
Эмиттер, база, коллектор	Исток, сток, затвор
<i>Функциональные контакты или структуры</i>	
Последовательно включенные <i>p-n</i> -переходы	Управляющий переход ( <i>p-n</i> или Шоттки, или гетеро-) или МДП (МОП)-структура
<i>Функции преобразования сигнала</i>	
Усиление Логические операции	
–	Стабилизация
<i>Достоинства</i>	
<ul style="list-style-type: none"> <li>↑ Быстродействие</li> <li>↑ Термостабильность ВАХ</li> <li>↑ Стабильность во времени</li> <li>↑ Надежность и воспроизводимость</li> </ul>	<ul style="list-style-type: none"> <li>↑ Входное сопротивление</li> <li>↓ Потребляемая мощность</li> <li>↓ Размеры</li> <li>↓ Число операций в маршруте</li> <li>↓ Стоимость</li> </ul>

### 2.3.1. Биполярный транзистор

1. Биполярный транзистор (БТ) – полупроводниковый триод, основанный чаще на двух последовательно включенных *p-n*-переходах, находящихся в различном омическом состоянии, работающий на эффекте управления перераспределением энергии между постоянным и пере-

менным электрическим полем с помощью тока базы, предназначенный для усиления переменного сигнала и выполнения логических операций.

2. Принципиальная структура БТ включает два *p-n*-перехода, которые образуются последовательно контактирующими областями эмиттера – базы – коллектора и называются соответственно эмиттерный и коллекторный переходы. Эмиттерный переход (ЭП) смещается прямо и находится в низкоомном состоянии, а коллекторный переход (КП) смещается обратно и находится в высокоомном состоянии. Перенос носителей заряда в пределах такой структуры включает три основных процесса:

1) эффективная *инжекция* основных носителей заряда через прямосмещенный ЭП из области эмиттера в область базы;

2) *диффузия* инжектированных (ставших неосновными) носителей через базу;

3) *экстракция* продифундировавших носителей через обратносмещенный КП в область коллектора.

Основным нежелательным процессом в этом электропереносе является процесс *рекомбинации* инжектированных носителей в области базы, который уменьшает выходной ток структуры – ток коллектора – по сравнению с входным током структуры – током эмиттера.

Соотношение токов в функциональной транзисторной структуре  $I_K < I_{\text{Э}} \gg I_B$ , независимо от схемы подключения этой структуры в заданном рабочем режиме при эксплуатации прибора, определяется двумя параметрами – коэффициентом передачи тока эмиттера  $\alpha_{\text{Э}}$  и коэффициентом передачи тока базы  $\beta_B$ :

$$\alpha_{\text{Э}} = \frac{I_K}{I_{\text{Э}}} < 1; \beta_B = \frac{I_K}{I_B} > 1. \quad (2.6)$$

При эксплуатации БТ подключается во входную и выходную цепи с использованием общего вывода. Наиболее распространенными являются схемы с общим эмиттером (ОЭ) и общей базой (ОБ). Работа транзистора в различных схемах описывается коэффициентами усиления (2.7) по току  $K_I$ , по напряжению  $K_U$  и по мощности  $K_P$  (2.7):

$$K_I = \frac{I_{\text{ВЫХ}}}{I_{\text{ВХ}}}; K_U = \frac{U_{\text{ВЫХ}}}{U_{\text{ВХ}}}; K_P = \frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}}} = \frac{I_{\text{ВЫХ}} U_{\text{ВЫХ}}}{I_{\text{ВХ}} U_{\text{ВХ}}}. \quad (2.7)$$

Рассмотрим работу БТ, включенного по *схеме ОБ* (рис. 2.25). В этом случае входной цепью является цепь эмиттер – база, а выходной – коллектор – база. Усиление по току в этой схеме не происходит, поскольку  $I_K(\text{ВЫХ}) < I_{\text{Э}}(\text{ВХ})$ .

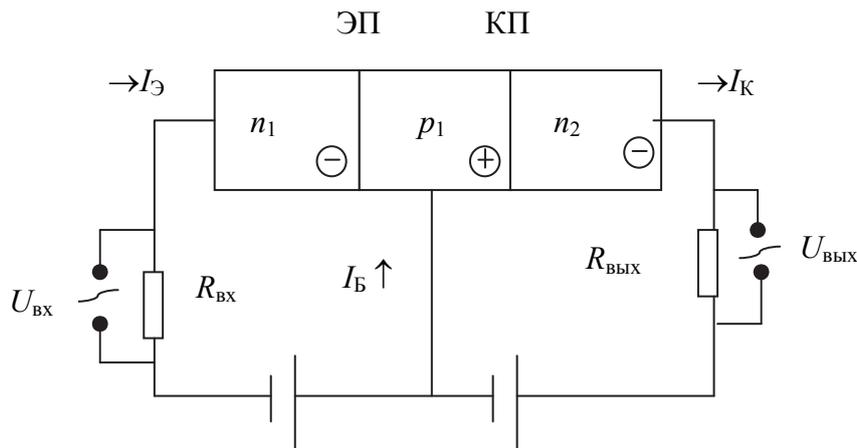


Рис. 2.25. Функциональная схема биполярного транзистора, включенного по схеме ОБ

Усиление по напряжению в схеме ОБ фактически определяется количественным соотношением нагрузочных сопротивлений во входной  $R_{\text{вх}}$  и в выходной  $R_{\text{вых}}$  цепях. При этом величина сопротивления  $R_{\text{вх}}$  должна быть не больше сопротивления низкоомного эмиттерного перехода, и величина сопротивления  $R_{\text{вых}}$  должна быть не больше сопротивления высокоомного коллекторного перехода, т. е. с учетом соотношения сопротивлений функциональных  $p$ - $n$ -переходов  $R_{\text{ЭП}} \ll R_{\text{КП}}$  получаем  $R_{\text{вх}} \ll R_{\text{вых}}$ . Это означает, что в схеме ОБ возможно многократное усиление по напряжению (и по мощности) в соответствии с выражением (2.8):

$$K_U = \frac{U_{\text{вых}}}{U_{\text{вх}}} = \frac{U_{\text{КБ}}}{U_{\text{ЭБ}}} = \frac{I_{\text{К}} R_{\text{вых}}}{I_{\text{Э}} R_{\text{вх}}} \approx \frac{R_{\text{вых}}}{R_{\text{вх}}} \gg 1. \quad (2.8)$$

Более универсальной является *схема ОЭ*, которая позволяет усиливать переменный электрический сигнал и по току, и по напряжению, и соответственно по мощности. В схеме ОЭ входной цепью является цепь база – эмиттер, а выходной – коллектор – эмиттер (рис. 2.26). Качественные представления об усилении тока в схеме ОЭ связаны с тем, что небольшой управляющий ток базы  $I_{\text{Б}}$  изменяет количество носителей в области базы и тем самым понижает потенциальный барьер на эмиттерном переходе, что приводит к увеличению потока инжектированных в базу и экстрагированных в коллектор носителей заряда. Количественно усиление по току описывается коэффициентом  $K_I$ :

$$K_I = \frac{I_{\text{вых}}}{I_{\text{вх}}} = \frac{I_{\text{КЭ}}}{I_{\text{БЭ}}} = \beta_{\text{Б}} \gg 1. \quad (2.9)$$

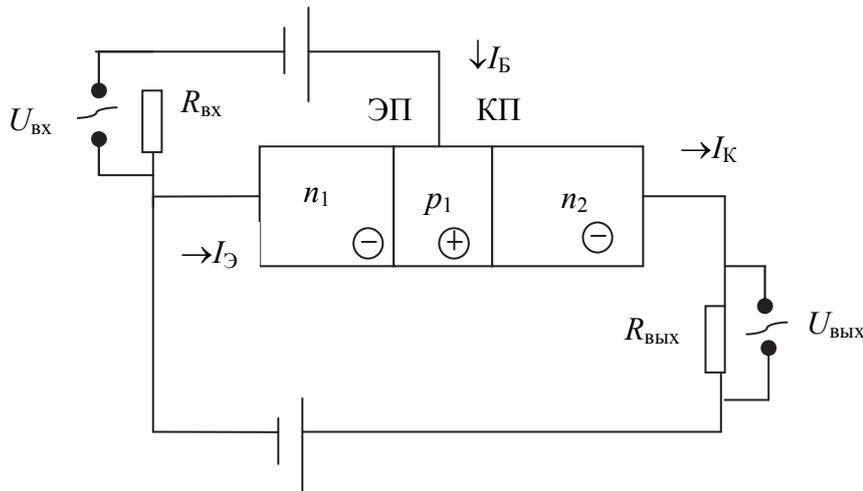


Рис. 2.26. Функциональная схема биполярного транзистора, включенного по схеме ОЭ

Биполярный транзистор описывается набором ВАХ, среди которых выделяют входные ВАХ  $I_{\text{ВХ}} = f(U_{\text{ВХ}})$  и выходные ВАХ  $I_{\text{ВЫХ}} = f(U_{\text{ВЫХ}})$ . Выходные ВАХ (рис. 2.27) в отсутствие сигнала во входной цепи представляют собой обратную ветвь выпрямляющей ВАХ, поскольку выходной ток  $I_{\text{К}}$  лимитируется высоким сопротивлением обратносмещенного коллекторного перехода. Причем величина выходного тока и наклон ВАХ БТ в схеме ОЭ больше (рис. 2.27, б), поскольку при подаче напряжения смещения на коллекторный переход эмиттерный переход оказывается смещенным прямо, что увеличивает поток носителей в транзисторной структуре. При подаче входного тока выходные ВАХ в обеих схемах смещаются вверх, поскольку выходной ток увеличивается. Отметим здесь, что при увеличении температуры выходные ВАХ также смещаются вверх, причем в схеме ОЭ более сильно, что указывает на пониженную термостабильность этой схемы.

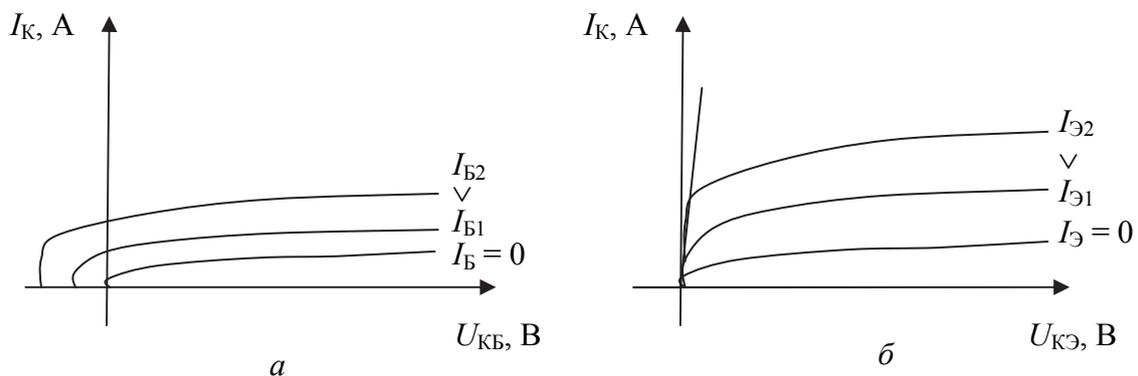


Рис. 2.27. Выходные ВАХ биполярного транзистора, включенного по схеме ОБ (а) и ОЭ (б)

3. Функциональная структура дискретного (неинтегрального) БТ включает два  $p-n$ -перехода и формируется на эпитаксиальной пластине, что позволяет, с одной стороны, обеспечить высокий потенциальный барьер на коллекторном переходе, представляющем контакт среднелегированных полупроводниковых областей, а с другой стороны, существенно снизить потребляемую мощность прибора за счет использования высоколегированной низкоомной подложки (рис. 2.28).

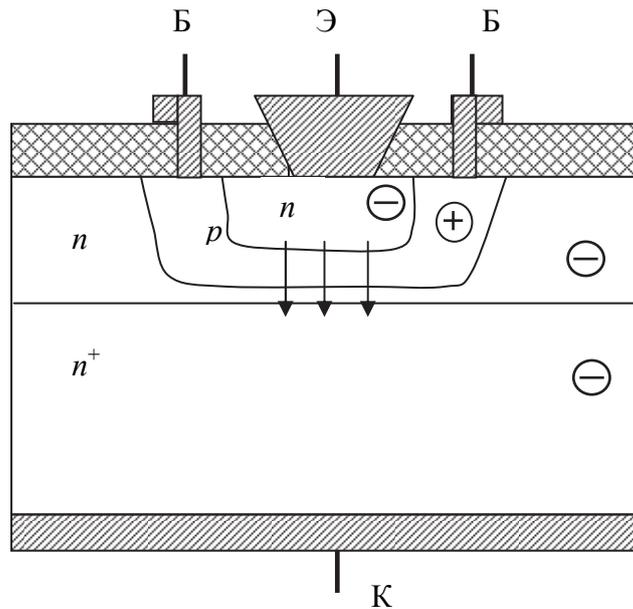


Рис. 2.28. Функциональная структура биполярного транзистора  $n-p-n$ -типа

Область эмиттера легируется высоко с целью обеспечения эффективной инжекции носителей заряда в область базы, которая легируется на среднем уровне, чтобы уменьшить долю рекомбинации инжектированных носителей в базе. При этом толщина базы в активной части, где локализуются функциональные электронные процессы транзисторной структуры, не превышает 1 мкм. Область базы снабжается двумя параллельными выводами для повышения эффективности управления электронными процессами в транзисторной структуре, где управляющим сигналом является микроамперный ток базы  $I_B$ . Основным материалом современных БТ, которые изначально формировались на германии, является кремний. Для СВЧ-транзисторов используется GaAs, обладающий повышенной подвижностью носителей заряда.

4. Среди ключевых параметров БТ можно выделить коэффициенты усиления, а также входные и выходные сопротивления. Данные табл. 2.6 показывают, что наиболее существенное усиление мощности входного сигнала достигается в БТ, включенном по схеме ОЭ.

Таблица 2.6

**Выходные параметры биполярного транзистора**

Название параметра	Значения параметра		
	ОБ	ОЭ	ОК
1. Коэффициент усиления по току $K_I$	–	10–100	10–100
2. Коэффициент усиления по напряжению $K_U$	10–1000	10–100	–
3. Коэффициент усиления по мощности $K_P$	10–1000	100–100 000	10–100
4. Сопротивление входное $R_{вх}$ , Ом	10–100	>100	10 000
5. Сопротивление выходное $R_{вых}$ , Ом	>100 000	>10 000	10–100

Отдельно рассмотрим вопрос частотной зависимости усилительных свойств БТ, поскольку частота рабочего переменного сигнала может существенно влиять на величину соответствующих коэффициентов усиления. Общей тенденцией является эффект уменьшения коэффициентов усиления при увеличении частоты рабочего переменного сигнала (рис. 2.29).

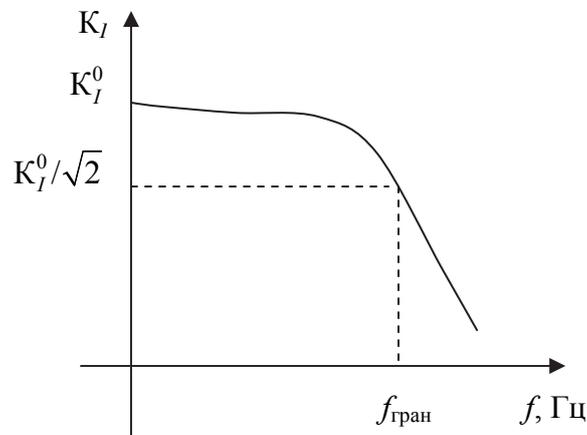


Рис. 2.29. Частотная зависимость коэффициента усиления по току БТ

При этом существенное ухудшение усилительных свойств начинает проявляться при достижении некоторого характерного значения частоты, которое называют граничной частотой работы транзистора  $f_{\text{гран}}$ .

Такое явление связано с тем, что величины коэффициентов передачи токов в биполярной структуре  $\alpha$  и  $\beta$  существенно уменьшаются при увеличении частоты входного переменного сигнала в соответствии с выражениями (2.10), что в свою очередь приводит к ухудшению усилительных свойств этих транзисторов:

$$\alpha = \frac{\alpha_0}{\sqrt{1 + \left(\frac{f}{f_\alpha}\right)^2}}; \beta = \frac{\beta_0}{\sqrt{1 + \left(\frac{f}{f_\beta}\right)^2}}; f_\alpha = (1 + \beta)f_\beta; f_\beta = (1 - \alpha)f_\alpha, \quad (2.10)$$

где  $\alpha_0$  и  $\beta_0$  – величины коэффициентов передачи токов эмиттера и базы при низких частотах;  $f$  – текущая частота входного сигнала, Гц;  $f_\alpha$  и  $f_\beta$  – *предельные*, или *граничные*, частоты усиления схемы ОБ и ОЭ соответственно, при которых коэффициенты  $\alpha_0$  и  $\beta_0$  уменьшаются в  $\sqrt{2}$  раз.

Можно выделить две основные причины наличия частотной зависимости усилительных свойств БТ:

1) фиксированное время пролета  $\tau_{пр}$  инжектированных носителей через базу, которое должно быть меньше полупериода переменного рабочего сигнала с тем, чтобы инжектированные носители успевали преодолеть базу за время монополярности рабочего сигнала, т. е. условием сохранения величины коэффициентов усиления является соотношение  $\tau_{пр} \leq T/2$ ;

2) наличие диффузионной емкости прямосмещенного эмиттерного перехода  $C_d$  и барьерной емкости обратосмещенного коллекторного перехода  $C_b$ , с которыми связано увеличение ответвляемого тока при увеличении частоты сигнала по причине уменьшения емкостного сопротивления.

Среди способов улучшения частотных свойств БТ можно назвать:

1) уменьшение времени пролета  $\tau_{пр}$  инжектированных носителей через базу за счет использования, например, *n-p-n*-транзисторов вместо *p-n-p*, полупроводников с повышенной подвижностью носителей, например GaAs вместо Si, а также уменьшения толщины базы;

2) уменьшение емкостей переходов в транзисторной структуре за счет уменьшения площади ЭП и КП или увеличения их толщины;

3) использование дрейфовых транзисторов с неоднородно легированной базой, где возникает диффузионное электрическое поле и перенос носителей осуществляется как диффузией, так и дрейфом по направлению от эмиттера к коллектору.

## 2.3.2. Полевые транзисторы

Полевые транзисторы (ПТ) могут создаваться на основе всех функциональных контактов и структур, которые мы обсуждали в пункте 2.1.1. Они делятся на следующие:

1) ПТ с *управляющим контактом* (*p-n*-переходом, гетеропереходом или контактом Шоттки);

2) ПТ с *изолированным затвором*, которые создаются на МДП- или МОП- (с оксидом в качестве диэлектрика) структуре и которые в свою очередь делятся на ПТ со *встроенным* и *индуцированным* каналами.

Выводы ПТ носят название *исток* (И), *сток* (С) и *затвор* (З). На затвор подается управляющий потенциал. Он регулирует поток носителей заряда в проводящем канале, находящемся между истоком и стоком, которые обеспечивают соответственно ввод и вывод рабочего сигнала.

### 2.3.2.1. Полевые транзисторы с управляющим контактом (*p-n*-переходом)

1. Полевой транзистор с управляющим *p-n*-переходом – это полевой транзистор, основанный на обратносмещенном *p-n*-переходе, работающий на управлении толщиной проводящего канала с помощью управляющего напряжения, предназначенный для усиления переменного сигнала и выполнения логических операций.

2. Работа этого вида транзисторов основана на регулировании с помощью потенциала на затворе толщины рабочей части проводящего канала за счет изменения толщины обедненных ОПЗ, прилегающих в функциональному обратносмещенному *p-n*-переходу. При увеличении управляющего поля толщина ОПЗ увеличивается, а толщина проводящей части канала уменьшается, что приводит к уменьшению тока в транзисторе. Иными словами, полевой транзистор с управляющим *p-n*-переходом работает в режиме *обеднения* и описывается убывающей ВАХ с характерным значением управляющего напряжения  $U_{отс}$ , при котором рабочая толщина канала уменьшается в два раза (рис. 2.30, а). Особенностью всех полевых транзисторов является важное значение как *переходной* (или *передаточной*) ВАХ, отражающей зависимость выходного тока от управляющего (входного) напряжения, так и выходной ВАХ (рис. 2.30). Отметим, что полевые транзисторы обычно используются в схеме с общим истоком (ОИ), которая позволяет усиливать напряжение и мощность благодаря тому, что управляющее входное напряжение существенно меньше напряжения в канале. Выполнение логических операций по двоичному коду  $0 (\uparrow R) \leftrightarrow 1 (\downarrow R)$  возможно благодаря наличию двух состояний транзисторной структуры: высокоомного (при  $U_{упр} > U_{отс}$ ) и низкоомного (при  $U_{упр} < U_{отс}$ ).

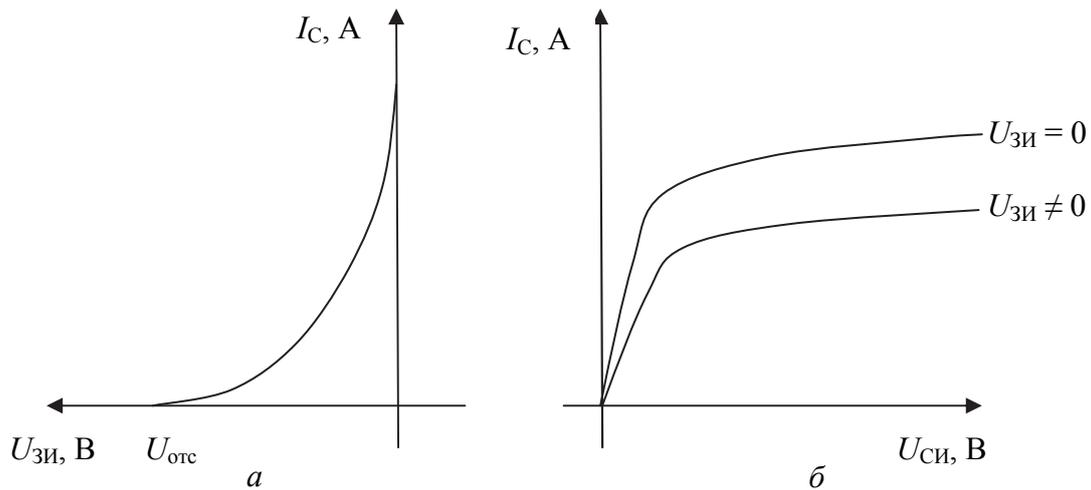


Рис. 2.30. Переходная (а) и выходные (б) ВАХ полевого транзистора с управляющим  $p-n$ -переходом

3. Функциональная структура ПТ с управляющим  $p-n$ -переходом может создаваться как в вертикальном, так и в горизонтальном направлении (рис. 2.31).

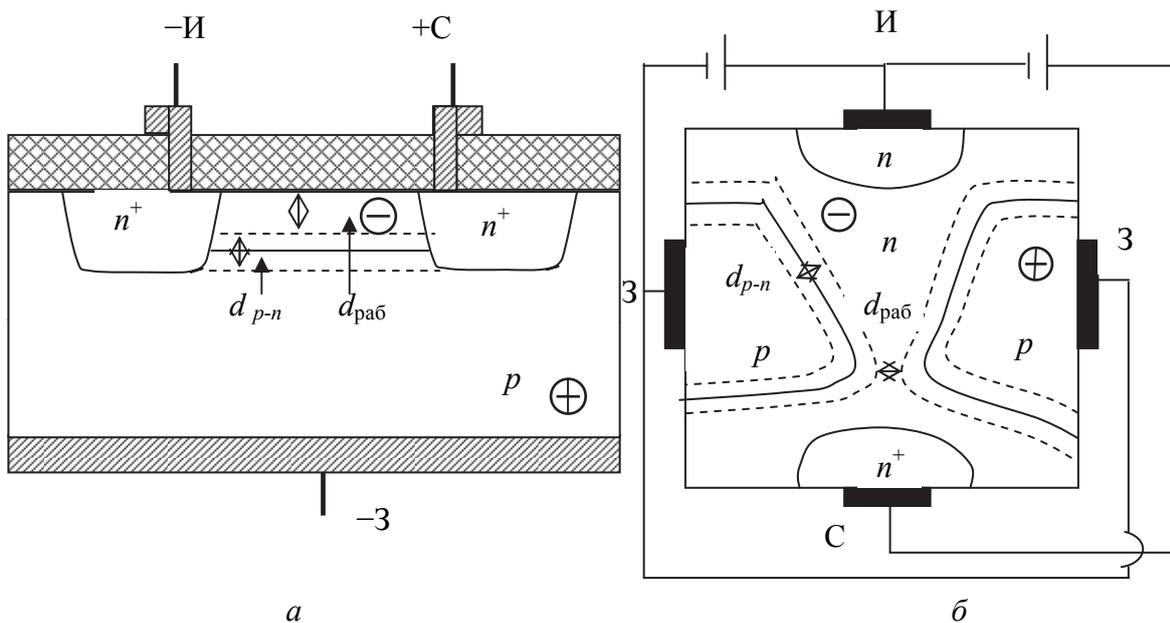


Рис. 2.31. Вертикальная (а) и горизонтальная (б) функциональные структуры  $n$ -канального полевого транзистора с управляющим  $p-n$ -переходом

В любом случае на затвор подается потенциал, обеспечивающий обратное смещение управляющего  $p-n$ -перехода и открытое (низкоомное) или закрытое (высокоомное) состояние транзистора,

а на исток – потенциал, обеспечивающий перенос по каналу основных носителей канала. Базовым материалом для таких транзисторов является кремний.

4. К выходным параметрам полевых транзисторов относят усилительные коэффициенты, сопротивления, а также характерное значение управляющего напряжения. Для транзисторов, работающих в режиме обеднения, таким напряжением является *напряжение отсечки*  $U_{отс}$  (рис. 2.30), которое соответствует двукратному уменьшению рабочей толщины проводящего канала и переходу транзистора в высокоомное состояние. Напряжение отсечки  $U_{отс}$  описывается выражением (2.11):

$$U_{отс} = \frac{eNd_p^2}{2\epsilon\epsilon_0}, \quad (2.11)$$

где  $N$  – концентрация ионизованных примесей в канале, т. е. концентрация мелких акцепторов в канале  $p$ -типа или концентрация мелких доноров в канале  $n$ -типа,  $m^{-3}$ ;  $d_p$  – рабочая толщина канала, которая при напряжении отсечки составляет половину от его исходной толщины, т. е.  $d_p = 0,5d_0$ .

Для ПТ  $U_{отс}$  может составлять величину от 1 до 100 В, а коэффициент усиления по напряжению  $K_U$  – от 10 до 1000.

### 2.3.2.2. Полевые МОП-транзисторы

1. МОП-транзистор – полупроводниковый триод, основанный на МДП-структуре (в варианте МОП-структуры), работающий на эффекте управления концентрацией основных носителей заряда в проводящем канале с помощью управляющего напряжения, предназначенный для усиления переменного сигнала и выполнения логических операций.

2–3. В случае МОП-транзистора прежде всего необходимо рассмотреть энергетическую диаграмму МОП-структуры, на которой ОПЗ присутствует только в области полупроводника.

Обсудим для начала диаграмму МОП-структуры  $Al/SiO_2/p-Si$  в равновесии (рис. 2.32). Направление переноса носителей заряда в реальной МОП-структуре без смещения определяется фактическими дефектами контактирующих слоев. Так, диэлектрические слои  $SiO_2$  обычно содержат донорные примеси и дефекты, которые обуславливают переход электронов из диэлектрика в полупроводник, что приводит к рекомбинации основных носителей (дырок на рис. 2.32) и образованию ОПЗ обедненного типа ( $E_V$  искривляется от  $E_F$ ).

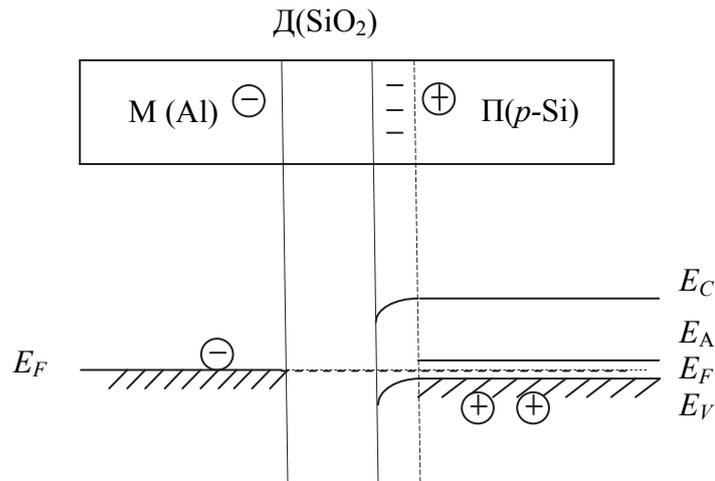


Рис. 2.32. Равновесная энергетическая диаграмма МОП-структуры Al/SiO<sub>2</sub>/p-Si

В МОП-структуре в качестве прямого смещения принят случай, когда полярность смещения совпадает с полярностью ОПЗ в полупроводнике и соответственно не совпадает с полярностью основных носителей в полупроводнике. При таком смещении внешнее и внутреннее поле усиливают друг друга и искривление зон увеличивается (рис. 2.33). Обеднение в ОПЗ углубляется, уровень Ферми на границе Д/П (SiO<sub>2</sub>/p-Si) смещается вверх и при некотором смещении происходит *инверсия* носителей заряда в ОПЗ полупроводника. Такая ситуация на диаграмме соответствует смещению уровня Ферми к дну зоны проводимости в нашем случае. При дальнейшем увеличении управляющего напряжения режим обеднения канала собственными носителями сменится режимом обогащения канала инверсными носителями.

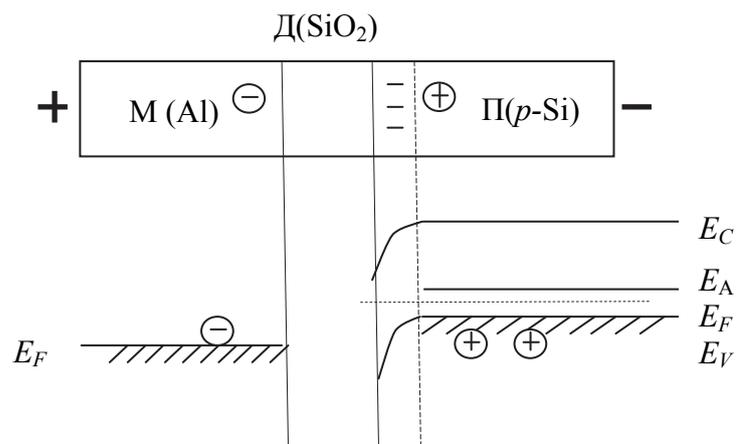
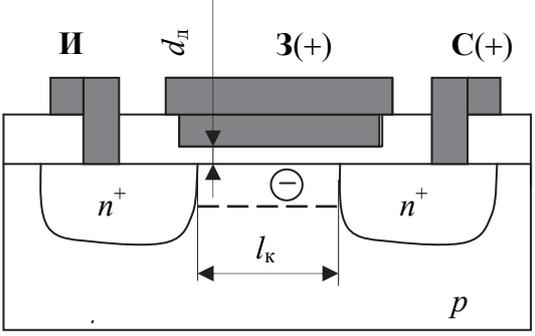
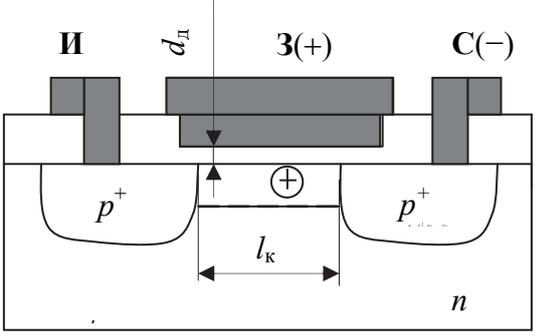
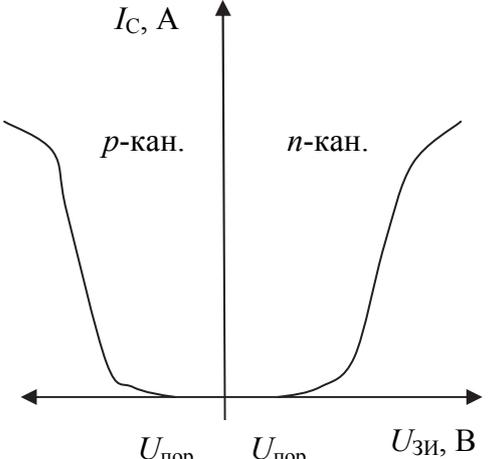
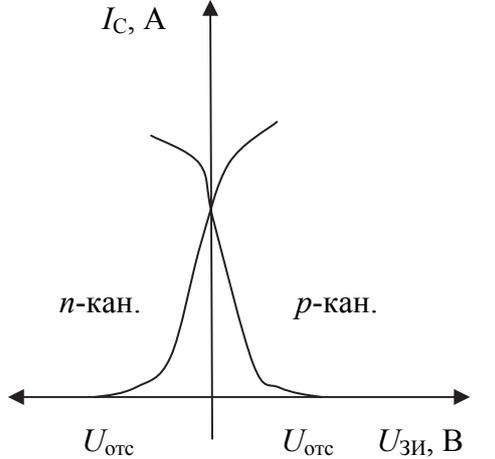


Рис. 2.33. Энергетическая диаграмма МОП-структуры Al/SiO<sub>2</sub>/p-Si при прямом смещении

На основе описанных процессов работает МОП-транзистор с индуцированным каналом. МОП-транзистор со встроенным каналом работает в режиме обеднения основными носителями канала. Удобно рассмотреть особенности структуры и работы различных МОП-транзисторов при сравнении (табл. 2.7).

Таблица 2.7

**Сравнительная характеристика МОП-транзисторов с индуцированным и встроенным каналом**

МОП-транзистор с индуцированным каналом	МОП-транзистор со встроенным каналом
<i>Функциональная структура</i>	
 <p style="text-align: center;"><i>n</i>-канальный</p>	 <p style="text-align: center;"><i>p</i>-канальный</p>
<i>Рабочий концентрационный режим канала</i>	
<b>Обогащение</b> основными носителями индуцированного канала (неосновными носителями подложки)	<b>Обеднение</b> основными носителями встроенного в подложку канала
<i>Полярность управляющего напряжения</i>	
Обеспечивает отталкивание основных носителей подложки и притягивание инверсных носителей канала	Обеспечивает отталкивание основных носителей встроенного канала
<i>Переходные ВАХ</i>	
	

Выходные ВАХ МОП-транзисторов аналогичны кривым для полевого транзистора с управляющим  $p$ - $n$ -переходом (рис. 2.30, б). При формировании функциональных структур МОП-транзисторов в случае *индуцированного* канала проводят подлегирование подложки примесью, соответствующей типу проводимости *подложки*, а в случае *встроенного* канала проводят компенсирующее легирование примесью, соответствующей типу проводимости *канала*.

4. Среди выходных параметров МОП-транзисторов можно выделить следующие:

1) характерное значение управляющего напряжения, при котором происходит переход между открытым и закрытым состоянием  $U_{отс}$  (для работы в режиме обеднения) или  $U_{пор}$  (для работы в режиме обогащения), составляющее от 1 до 100 В;

2) крутизна переходной ВАХ:

$$S_p = \frac{dI_{\text{ВЫХ}}}{dU_{\text{ВХ}}} \approx \frac{\Delta I_{\text{ВЫХ}}}{\Delta U_{\text{ВХ}}}. \quad (2.12)$$

Для схемы ОИ:

$$S_p \approx \frac{\Delta I_c}{\Delta U_{\text{зи}}}, \quad (2.13)$$

где  $\Delta I_c$  – изменение выходного тока стока, А;  $\Delta U_{\text{зи}}$  – изменение управляющего напряжения между затвором и истоком, В.

Величина  $S_p$  составляет величину порядка 1 А/В и для конкретного прибора может быть рассчитана по формуле

$$S_p = \frac{C_{\text{МОП}} U_{\text{ВЫХ}} \mu}{l_k^2}, \quad (2.14)$$

где  $C_{\text{МОП}}$  – электрическая емкость МОП-структуры, Ф;  $U_{\text{ВЫХ}}$  – выходное напряжение между стоком и затвором, В;  $\mu$  – подвижность основных носителей канала,  $\text{м}^2/\text{В} \cdot \text{с}$ ;  $l_k$  – длина проводящего канала, м;

3) коэффициент усиления по напряжению, достигающий 1000:

$$K_U = S_p R_{\text{ВЫХ}}; \quad (2.15)$$

4) напряжение пробоя подзатворного диэлектрика должно существенно превышать величину  $U_{отс}$  или  $U_{пор}$ ;

5) максимальная рабочая частота сигнала не превышает 10 кГц в приборах на основе традиционной структуры  $\text{Al}(\text{ПКК})/\text{SiO}_2/p\text{-Si}$ , но существенно выше в инновационных структурах  $\text{Cu}/\text{HfO}_2/\text{Si}$ .

## 2.4. Полупроводниковые и гибридные интегральные схемы (лекция 12)

*Интегральные схемы (ИС)* – это сложные электронные устройства, включающие множество коммутированных элементов и/или компонентов, которые выполняют заданную более или менее сложную функцию преобразования электрического сигнала и рассматриваются как единое целое.

ИС включают неотделимые от подложки *элементы* и *навесные компоненты*, которые обратимо монтируются на подложку с использованием монтажных технологий, например пайки или сварки. При этом элементы и компоненты разделяют на *активные*, или нелинейные (диоды, транзисторы, ИС), и *пассивные*, или линейные (резисторы, конденсаторы).

ИС классифицируются по различным критериям:

1) по конструкторско-технологическим особенностям:

– *полупроводниковые* (содержат активные и пассивные элементы, встроенные в полупроводниковую подложку);

– *гибридные* (содержат пассивные пленочные элементы, сформированные на диэлектрической подложке, и навесные активные и пассивные компоненты));

– *вакуумные* (содержат в качестве активных компонентов вакуумные лампы и другие приборы);

2) по виду рабочего сигнала:

– *аналоговые* (преобразуют непрерывный электрический сигнал, выполняют функцию усиления);

– *цифровые* (преобразуют дискретный электрический сигнал, выполняют функцию логических операций).

Основной идеей создания ИС являлось стремление интегрировать в компактном устройстве выполнение все более сложных функций преобразования сигнала. Количественно эта интегрированность отражается двумя основными показателями ИС.

1. Степень интеграции  $N_{\text{инт}} = \lg N_{\text{эл/ком}}$ , которая учитывает количество активных элементов или компонентов  $N_{\text{эл/ком}}$  в составе ИС и варьируется от 2 (малые интегральные схемы МИС) до более 9 (например, в новейших процессорных схемах содержится более 2 миллиардов транзисторов). Все ИС с  $N_{\text{инт}} > 4$  относят в настоящее время к большой группе сверхбольших ИС (СБИС).

2. Плотность упаковки элементов, которая отражает количество активных элементов на единице площади подложки. Анонсируемый

мировыми лидерами к настоящему времени уровень плотности упаковки составляет, например, от компании TSMC 170 миллионов транзисторов на квадратный миллиметр, т. е.  $170 \text{ МТр/мм}^2$  в рамках технологического процесса с топологическими нормами 5 нм, от компании Samsung в рамках аналогичного процесса –  $125\text{--}130 \text{ МТр/мм}^2$ , а от компании Intel – около  $100 \text{ МТр/мм}^2$  в рамках процесса с нормами 10 нм.

Рассмотрим более подробно полупроводниковые интегральные схемы (ПИС), которые называют микросхемами, поскольку они характеризуются непревзойденной компактностью и сегодня уже смело могут называться наносхемами. К их достоинствам, кроме миниатюрности, можно отнести высокую надежность, пониженное энергопотребление и материалоемкость, универсальность в решении поставленных технических задач. По виду основного активного элемента микросхемы разделяют на *биполярные* (на основе биполярных транзисторов) и *КМОП* (на основе комплиментарной пары *n*- и *p*-канальных МОП-транзисторов).

Биполярные микросхемы в качестве основного активного элемента включают интегральный биполярный транзистор, который отличается от дискретного транзистора тем, что содержит верхний вывод коллектора на планарной стороне пластины. Это ухудшает параметры работы транзистора и усложняет технологию его формирования. Кроме того, требуется надежная изоляция транзисторов и других элементов, которые формируются в отдельных карманах как друг от друга, так и от подложки. На рис. 2.34 представлен вариант кармана, в котором сформирована структура интегрального БТ, совмещенного с диодом Шоттки в области базы и изолированного от соседних карманов локальным диэлектриком  $\text{SiO}_2$ , а от подложки – *p-n*-переходом.

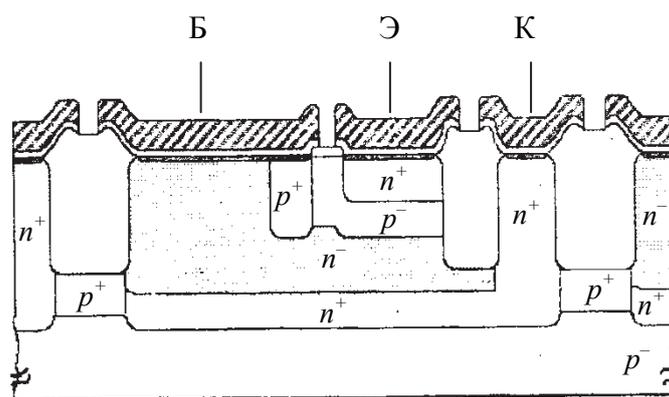


Рис. 2.34. Функциональная структура интегрального биполярного транзистора в варианте комбинированной изоляции



## 2.5. Полупроводниковая наноэлектроника (лекция 13)

Современный этап развития электроники, особенно полупроводниковой электроники, можно охарактеризовать как переход от микроэлектроники к наноэлектронике, в рамках которой создаются устройства с характерными размерами менее 100 нм. Переход к наноэлектронным приборам стал возможен в результате решения триединой задачи: во-первых, изучения наноматериалов, во-вторых, описания электронных процессов в функциональных наноструктурах, и в-третьих, что самое главное для реализации идей, создания нанотехнологий, способных воспроизводимо, масштабно и рентабельно изготавливать наноустройства.

Наноэлектроника основана на наноматериалах, под которыми понимают материалы и структуры с размером элемента неоднородности (например, размером кристаллита или толщиной пленки) такого же порядка, как характерный размерный параметр того или иного физического явления (например, длина волны излучения или длина свободного пробега носителей заряда). Основой существенного изменения свойств в наноструктурированных твердых телах является представление о решающей роли поверхностных атомов в наноматериалах. Размер элемента структуры в них составляет 1–100 нм. Верхний предел этого диапазона принят достаточно условно и вытекает из технически интересных возможностей изменения физико-механических свойств наноматериалов (например, прочности, твердости, коэрцитивной силы и др.) при уменьшении размера зерна ниже 100 нм. Нижний предел указанного диапазона ограничен критическим размером элемента кристаллической решетки, в пределах которого сохраняются все присущие данному типу кристаллов элементы симметрии.

В наноразмерных материалах доминирующими являются *квантово-механические* (т. е. волновые) явления и закономерности, что определяет их электронные, оптические и другие свойства. В таких квантоворазмерных структурах (или частицах) в роли характерной физической длины, т. е. длины, с которой сравнивают размеры образца, выступают квантовые, а не классические величины. Для электронов в твердом теле такой длиной является длина волны де Бройля. Если характерный геометрический размер (например, толщина активного слоя  $d$ ) соизмерим или меньше длины волны де Бройля для электрона, то проявляются эффекты размерного квантования, т. е. характер кван-

тования энергии электрона, ее величина зависят не только от природы объекта, но и от его размеров. Это значит, что энергией уровней и волновыми функциями электронов можно управлять, изменяя размеры объекта (например, толщину активного слоя или размер зерна).

Нанoeлектронный функциональный элемент в современной полупроводниковой электронике обычно состоит из набора квантовых ям и потенциальных барьеров (рис. 2.36). В такой системе добавление лишь одного электрона существенно меняет энергетическое состояние (диаграмму) системы в целом. Каждый наноструктурированный элемент системы (например, отдельный слой гетероструктуры) является многофункциональным и способен выполнять сложные динамические функции.

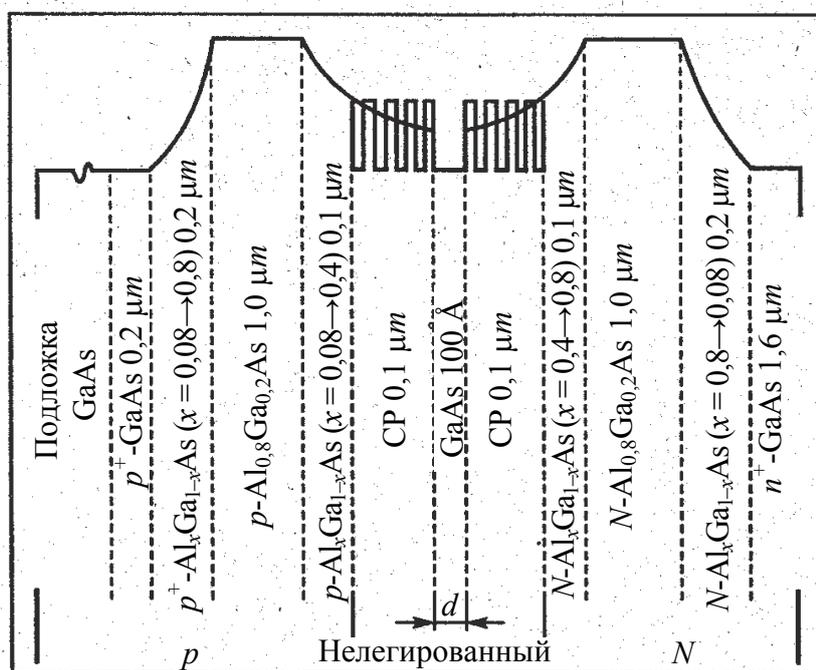


Рис. 2.36. Структура РОДГС-лазера с квантовой ямой и параболическим волноводом

Фактически объекты нанoeлектроники представляют собой функциональные структуры, которые изготавливаются с топологическими нормами менее 100 нм, что уменьшает размеры и энергопотребление электронных приборов. Рекордные нанопараметры миниатюризации достигнуты на монополярных структурах, в которых основным активным элементом является полевой транзистор. Изначально уменьшение его размеров решалось путем уменьшения длины проводящего канала, что приводило к увеличению токов утечки в подложку. В качестве

решения этой проблемы предложено использование КНИ-структур (кремний на изоляторе), что позволило в 2007–2009 гг. достигнуть уровня топологических норм 45–32 нм.

Следующий этап развития микроэлектроники в нанoeлектронику – переход от двумерных планарных 2D-структур к трехмерным объемным 3D-структурам на основе данных математического моделирования электромагнитных процессов в объемных функциональных структурах, что дало возможность уменьшить топологические нормы до 14–22 нм начиная с 2013 г. Использование объемных структур позволило снизить напряжение питания до 2,5–3,0 В. Примером объемных транзисторов являются различные варианты FinFET (field effect transistor).

Функциональная 3D-структура такого транзистора включает вертикальный канал, который возвышается над поверхностью подложки, как плавник акулы над гладью океана. Затвор выполняется в виде «подковы», которая охватывает канал с трех сторон (рис. 2.37). Таким образом, по мере снижения топологических норм производства затворы прошли путь от планарных структур, которые могли контролировать одну единственную область под затвором, до вертикальных каналов, окруженных затвором с трех сторон, и, наконец, приблизились к переходу на каналы, окруженные затворами со всех четырех сторон. Увеличение площади затвора вокруг управляемого канала позволяло снижать питание транзисторов без ущерба для их токовых характеристик, что вело к увеличению производительности транзисторов и к снижению токов утечек. Окруженные затвором каналы могут выпускаться как в виде тонких нанопроводов (вариант GAAFET (gate-all-around FET), так и в виде широких мостов или наностраниц (MBCFET (Multi Bridge Channel FET)) (рис. 2.37). Широкие страницы каналов, по мнению компании-разработчика Samsung, обеспечат значительные токи, а также, по сравнению с 7-нм FinFET техпроцессом, переход на нормы 3 нм, снижение потребления на 50%, увеличение производительности на 30% и уменьшение площади чипов на 45%, возможно, в 2022 г.

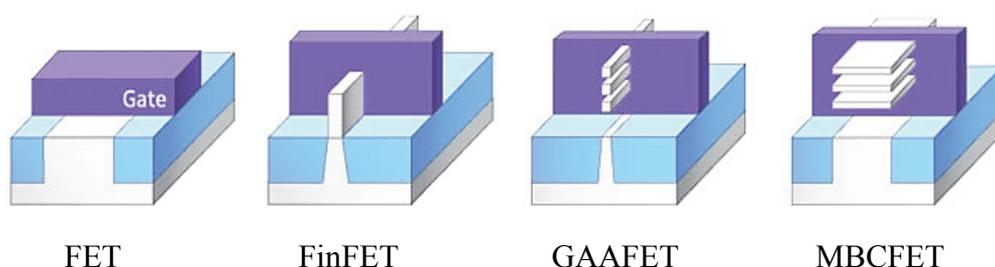


Рис. 2.37. Этапы развития функциональных структур полевого транзистора

Есть и другие предложения для техпроцесса 3 нм. Так, группа китайских исследователей из Института микроэлектроники китайской академии наук разработала транзистор, который в отличие от 3-нанометровой структуры транзистора компании Samsung, предполагающей переход на полностью окруженные затворами каналы в виде наностраниц, выполнен в виде каналов из вертикальных FinFET-ребер, окруженных затворами только с трех сторон. Другое отличие китайской разработки заключается в материале, из которого изготавливается транзистор. Это ферроэлектрик, и в этом суть изобретения.

В целом проблема при изготовлении 3-нанометрового транзистора даже не в том, что его размеры становятся слишком маленькими (сравнимыми, например, с нитью ДНК). Препятствием для уменьшения размера транзистора является так называемая больцмановская тирания (Boltzmann Tyranny). Это фундаментальное ограничение, которое сопровождается снижением рассеивания мощности в процессе работы электронного прибора. Попросту говоря, после определенного уменьшения размера транзистора он перестает рассеивать рабочее тепло и, следовательно, сгорает. Чтобы этого не произошло, необходимо снижать питание, но ниже порогового значения опуститься нельзя. Это противоречит физике процессов в полупроводниках. И тогда на помощь приходят ферроэлектрики. Точнее, такое теоретически известное и парадоксальное явление в ферроэлектриках, как отрицательная емкость. Устойчивый отрицательный конденсатор, в котором по мере роста напряжения емкость не увеличивается, а уменьшается, впервые представлен физически лишь в 2021 г. Это позволит снизить напряжение питания ниже порогового значения. Китайские разработчики сумели воплотить эффект отрицательной емкости в конструкции 3-нанометрового транзистора и снизить в два раза по сравнению с теоретическим минимумом напряжение питания транзисторов.

Мировым лидером производства полупроводниковых структур по уровню технологий и по объему выпуска является тайваньская компания TSMC (Taiwan Semiconductor Manufacturing Company), которая отвечает за выпуск 92% продукции с использованием передовых литографических технологий. Техпроцесс 5 нм этой компании может обеспечить около 170 МТр/мм<sup>2</sup>. Такой же процесс 5 нм компании Samsung обеспечивает плотность только 125–130 МТр/мм<sup>2</sup>, а 10-нанометровые нормы компании Intel – только 100 МТр/мм<sup>2</sup>. Ожидается, что компания TSMC будет производить транзисторы GAAFET в техпроцессе 2 нм.

Под устройствами оптоэлектроники понимают электронные приборы, работа которых основана на эффектах взаимодействия электрона с оптическим излучением и электрическим или электромагнитным полем. В этих приборах реализуются различные виды оптического поглощения, эффекты электролюминисценции, фотопроводимости, фото-ЭДС и другие явления, связанные с определенной областью оптического диапазона спектра электромагнитного излучения (рис. 3.1).



Рис. 3.1. Схема поддиапазонов оптического диапазона спектра электромагнитных излучений

Среди достоинств оптоэлектроники можно отметить следующие.

1. *Высокая информационная емкость.* Высокие частоты электромагнитного излучения оптического диапазона ( $10^{13}$ – $10^{15}$  Гц, т. е. в 1000–100 000 раз выше, чем в радиодиапазоне) позволяют резко увеличить количество информационных каналов передающего устройства или передающей линии. Соответственно, малые длины волн используемого излучения обеспечивают высокую плотность записи информации в запоминающих устройствах (порядка  $10^8$  бит/см<sup>2</sup>).

2. *Острая направленность светового излучения.* При передаче сигнала световым пучком, угловая расходимость которого может быть менее 1', энергия передается концентрированно и с малыми потерями в заданную область пространства, например на фоточувствительные площадки микронных размеров.

3. *Возможность двойной (временной и пространственной) модуляции светового луча.* Это позволяет проводить не только последова-

тельную, но и параллельную обработку информации, что важно, к примеру, для увеличения быстродействия компьютеров.

4. *Использование электрически нейтральных фотонов.* Электро-нейтральность фотонов обеспечивает механическую и электрическую бесконтактность связи, идеальную гальваническую развязку входа и выхода электронной схемы, однонаправленность потока информации и отсутствие обратной реакции приемника сигнала на его источник, помехозащищенность линий связи.

5. *Возможность непосредственного оперирования со зрительно воспринимаемыми образами.* Поскольку зрение является нашим основным органом чувств, зрительное представление количественной информации (например, цифры на табло вольтметра вместо отклонения стрелки), да и любой другой, воспринимается наиболее естественно.

6. *Возможность создания устройств подлинно функциональной микроэлектроники.* Оптоэлектроника с самого начала ставила своей целью избавиться от таких громоздких устройств, как трансформаторы, реле, герконы, кабели, переменные резисторы, разъемы. В настоящее время решению этой проблемы посвящены активные разработки в области так называемых *микроэлектромеханических систем* (МЭМС).

## 3.1. Физические основы оптической и квантовой электроники (лекция 14)

Физическую основу оптоэлектроники составляют процессы преобразования электрических сигналов в оптические или, наоборот, процессы распространения излучения в различных средах, а также эффекты взаимодействия электромагнитных излучений оптического диапазона с веществом, включая физиологические. Квантовая электроника является разделом оптоэлектроники, в который выделены процессы излучения электромагнитного поля (в виде квантов энергии) средой, например твердым телом.

Среди оптоэлектронных приборов выделяют *источники* излучения (светодиоды, лазеры) и *приемники* излучения (фоторезисторы, фотодиоды, фотоэлементы (солнечные элементы), фототранзисторы и др.). Источники излучения преобразуют электрическую энергию управляющих сигналов в когерентное (лазеры) или некогерентное (светодиоды) оптическое излучение. Эти приборы ранее изготавливались на *p-n*-переходах, а теперь чаще основаны на гетеропереходах, включающих прямозонные (GaAs, GaN) и непрямозонные (GaP, SiC) полу-

проводники. Приемники излучения преобразуют световую энергию в электрическую и используются как *фотодетекторы* (например, фоторезисторы, фотодиоды) или как фотопреобразователи (фотоэлементы). Основным материалом для этих приборов пока является кремний.

### 3.1.1. Основные физические процессы при взаимодействии света с кристаллом

Физические процессы, протекающие в кристаллическом теле при его взаимодействии с электромагнитным излучением в оптическом диапазоне длин волн, могут сопровождаться сохранением энергетического кванта излучения (фотона) или передачей энергии фотона твердому телу. К первым относятся процессы пропускания, отражения, рассеяния (преломления) света, вращения плоскости поляризации и другие процессы, которые изучаются классической оптикой. В результате вторых процессов за счет поглощения фотона в твердом теле происходит генерация заряженных частиц (электрон проводимости, дырка) или незаряженных квазичастиц (фотон, фонон, экситон) (рис. 3.2).

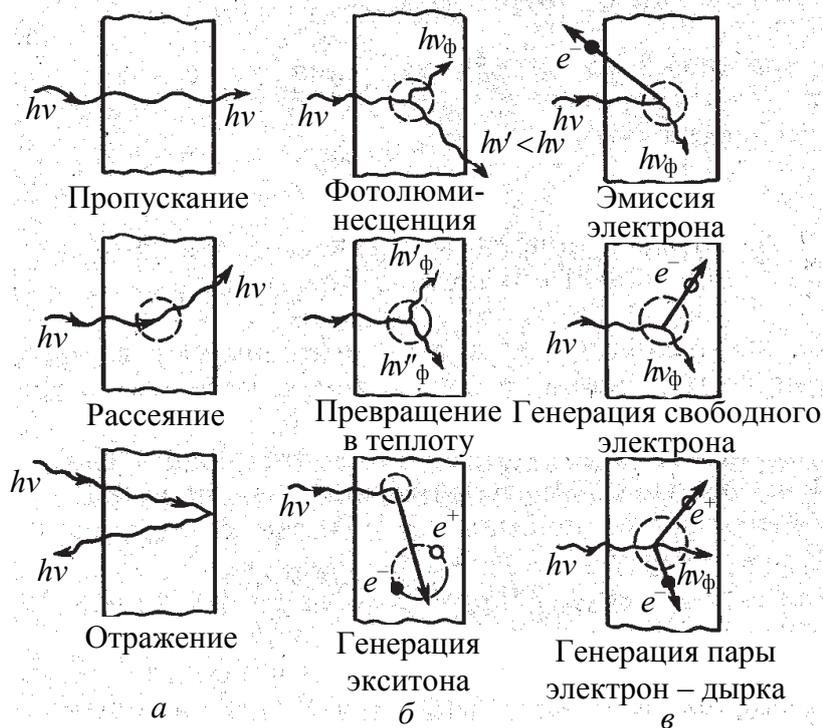


Рис. 3.2. Взаимодействия света с твердым телом (неэлектрические и электрические соответственно):  
 а – процессы с сохранением энергии кванта;  
 б и в – процессы с поглощением кванта

В результате процессов второго типа, сопровождающихся поглощением фотонов в твердом теле, изменяется энергия свободных или связанных электронов либо колебательная энергия атомов. Взаимосвязь электронно-атомных возбуждений, происходящих при поглощении кванта энергии, отражает *принцип Франка – Кондона*, по которому вследствие большой разницы в массах между ядерным остовом и возбуждаемым электроном электронный переход происходит так быстро, что ядра при этом можно считать неподвижными. Это значит, что электронный переход является вертикальным, поскольку ядра сохраняют свои положения. Однако если энергетические кривые (например, дисперсионные кривые, отражающие зависимость энергии от волнового числа) смещены друг относительно друга, электронный переход приводит к сжатому или растянутому состоянию молекулы, которая начинает колебаться, т. е. электронный переход сопровождается колебательным возбуждением молекулы.

### 3.1.2. Виды поглощение света в полупроводниках

Поглощение фотонов твердым телом сопровождается передачей энергии фотона электрону или атомам. Интенсивность светового потока  $J$  при прохождении в теле расстояния  $x$  уменьшается от исходного значения  $J_0$  экспоненциально в соответствии с законом Бугера – Ламберта:

$$J = J_0 \exp(-\alpha x), \quad (3.1)$$

где  $\alpha$  – коэффициент поглощения, пропорциональный вероятности поглощения фотона на единичной длине,  $\text{см}^{-1}$  (если толщина образца измеряется в сантиметрах).

Вид спектра оптического поглощения полупроводника, т. е. зависимость  $\alpha$  от частоты  $\nu$  или длины волны падающего света  $\lambda$ , определяется характером и концентрацией поглощающих центров,

которыми могут быть свободные и связанные электроны, а также частицы кристаллической решетки (рис. 3.3).

*Собственное* поглощение в полупроводнике сопровождается образованием пар неравновесных носителей заряда электрон – дырка и возникновением фото-

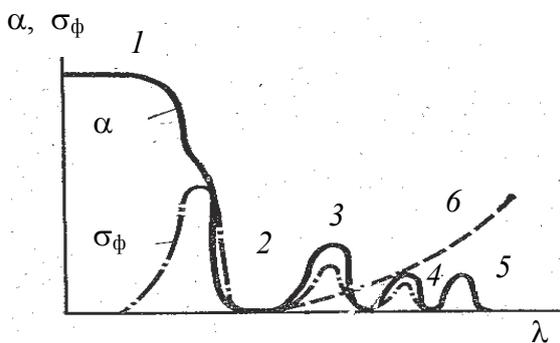


Рис. 3.3 Спектры поглощения (1, 2) и фотопроводимости (3–6)

проводимости  $\sigma_{\phi}$ . На спектре оптического поглощения полупроводника такое явление обуславливает проявление ступеньки в области граничной частоты, которую называют краем собственного поглощения (рис. 3.3, участок 1–2). Это позволяет определять ширину запрещенной зоны материала  $E_g$  по граничной частоте  $E_g/h$ . Смещение края собственного поглощения от длины волны, соответствующей  $h/E_g$ , происходит в соответствии с особенностями зонной структуры полупроводника. Так, в спектре поглощения двудолинных полупроводников (GaAs, InP) присутствуют два края поглощения, обусловленные переходами электрона в разные долины зоны проводимости. Для непрямозонных материалов (Si, Ge, GaP и др.) край поглощения смещен в коротковолновую область спектра, поскольку более вероятным прямым (вертикальным) переходам соответствует энергия бóльшая, чем зазор между потолком валентной зоны и дном зоны проводимости (рис. 3.4).

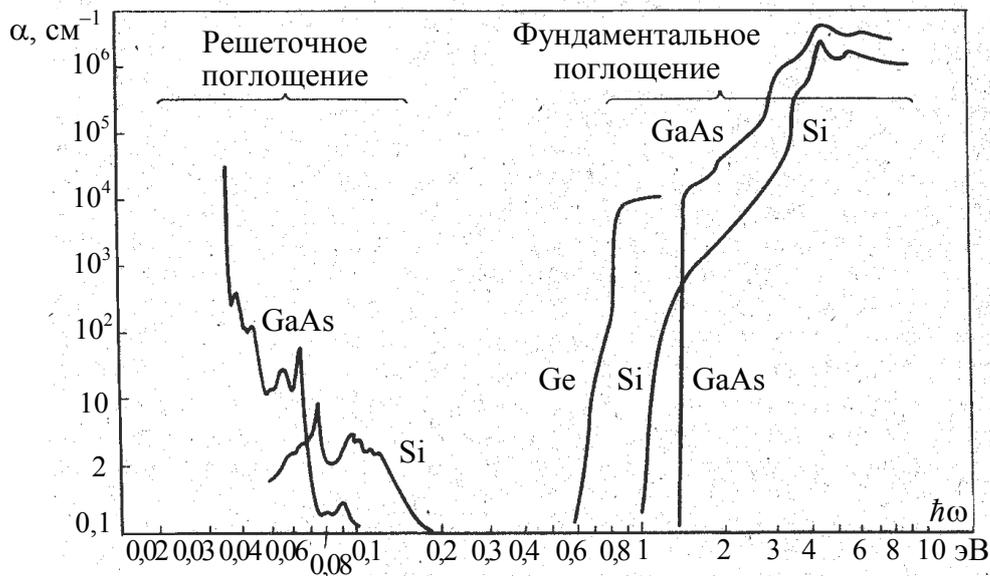


Рис. 3.4. Спектры собственного поглощения полупроводников

*Примесное* поглощение приводит к возбуждению или ионизации примеси (рис. 3.3, области 4 и 3 соответственно). Если электрон в таком случае переходит в зону (проводимости) или из зоны (валентной), то происходит генерация неравновесных носителей заряда и увеличение проводимости материала (рис. 3.3, область 3). Однако при изучении спектров поглощения примесных полупроводников следует помнить, что полосы примесного поглощения проявляются только для

глубоких примесей, не ионизованных в условиях эксперимента за счет тепловой энергии.

*Экситонное* поглощение сопровождается возникновением возбужденного электрона, связанного с образовавшейся дыркой. Такое возбуждение, которое может распространяться по кристаллу, называется *экситоном*. Его энергия немного меньше ширины запрещенной зоны, которая определяет минимальную энергию образования пары носителей заряда (несвязанные электрон и дырка). Образование экситонов характерно для полупроводников и диэлектриков. Однако в последних притяжение между электроном и дыркой велико, что обуславливает повышенную устойчивость экситона. В полупроводниках же экситонные орбиты охватывают несколько элементарных ячеек кристалла (десятки нанометров) и легко распадаются на дефектах с образованием двух носителей, что приводит к изменению проводимости и появлению узких пиков у края собственного поглощения (в области 2 на рис. 3.3).

Поглощение *свободными носителями* заряда в полупроводниках проявляется на спектрах оптического поглощения при достаточно высокой концентрации носителей в виде участков монотонного возрастания коэффициента поглощения вплоть до длин волн порядка 100 мкм (неселективное поглощение, рис. 3.3, область б) или в виде узких полос (селективное поглощение). В первом случае не соблюдаются правила отбора, согласно которым реализуются электронные переходы между разрешенными энергетическими уровнями с сохранением квазиимпульса электрона. Такое неселективное поглощение возможно, если кроме процесса поглощения фотона в полупроводнике происходит поглощение (или испускание) фонона или рассеяние носителей на ионизованных примесях, поскольку именно электрон-фононное взаимодействие обеспечивает соблюдение закона сохранения импульса при переходах с изменением квазиимпульса электрона. В этом случае коэффициент поглощения  $\alpha$  связан с проводимостью  $\sigma$  соотношением

$$\alpha = \sigma / c \epsilon_0 n, \quad (3.2)$$

где  $c$  – скорость света в вакууме ( $3,00 \cdot 10^8$  м/с);  $\epsilon_0$  – электрическая постоянная ( $8,85 \cdot 10^{-12}$  Кл<sup>2</sup>/Н · м<sup>2</sup>);  $n$  – показатель преломления.

Селективное поглощение происходит с выполнением правил отбора и может быть связано, например, с переходами электрона между подзонами валентной зоны, образующимися за счет спин-орбитального взаимодействия электронов. Поглощение носителями заряда при-

водит к возникновению тока высокой (оптической) частоты и в конечном итоге выделяется в виде джоулева тепла.

*Междолинное* поглощение характерно для многодолинных полупроводников (рис. 1.10). Его можно рассматривать как разновидность неселективного поглощения носителями заряда. Особенностью этого поглощения является переход электрона из одного устойчивого (с минимальной потенциальной энергией) состояния в другое, которые отличаются эффективной массой и, следовательно, подвижностью носителей, что отражается на проводимости полупроводника.

*Решеточное* поглощение особенно важно для веществ с достаточной долей ионности связи, поскольку генерация оптических фононов (изменение энергии колебания атомов или ионов) приводит к заметному изменению вектора поляризации диполей в материале. Наиболее сильно поглощаются фотоны, энергия которых близка к энергии собственных колебаний диполей и составляет десятки миллиэлектронвольт, что соответствует далекой инфракрасной области электромагнитного спектра (рис. 3.3, область 5). Импульс фотона  $h/\lambda$  (где  $\lambda$  в оптическом диапазоне спектра электромагнитного излучения составляет величину от 1 мм до сотен нанометров) существенно меньше квазиимпульса фонона  $h/a$  (где  $a$  – постоянная решетки, в большинстве случаев имеющая значение менее 1 нм). Поэтому закон сохранения импульса соблюдается при выделении нескольких фононов на 1 поглощенный фотон, что усложняет структуру спектра решеточного поглощения. Данный вид поглощения определяет окраску материалов.

*Плазменное* поглощение обусловлено взаимодействием фотонов с плазменными волнами. Последние представляют собой согласованные колебания напряженности электрического поля и объемной плотности заряда, в режиме когда эти колебания взаимно поддерживают друг друга и затухают со временем за счет омических потерь. По сути, плазменные волны представляют собой один из типов нормальных колебаний поля в среде со свободными зарядами. Данный тип поглощения проявляется на так называемой плазменной частоте, которая для полупроводников составляет величину около  $10^{-12} \text{ с}^{-1}$ .

### 3.1.3. Фотопроводимость и фото-ЭДС

Одним из многочисленных фотоэлектрических явлений в полупроводниках является *фоторезистивный* эффект, который заключается в изменении электрического сопротивления при воздействии

электромагнитного излучения. Добавочную или избыточную проводимость  $\Delta\sigma$ , возникающую в результате поглощения фотонов, называют *фотопроводимостью*. Фотопроводимость обусловлена в первую очередь внутренним фотоэффектом, который сопровождается возбуждением электронов, но не приводит к выходу электронов из твердого тела, как это происходит при внешнем фотоэффекте. В результате внутреннего фотоэффекта увеличивается концентрация носителей заряда по механизму собственного или примесного поглощения. Кроме того, возможно изменение подвижности (эффективной массы) носителей при возбуждении в пределах разрешенной зоны. Такой вклад в фотопроводимость реален, если «горячие» электроны не успевают термализоваться за время пребывания (время жизни) в разрешенной зоне, т. е. их энергия остается избыточной и не соответствует распределению по энергиям и импульсам равновесных носителей в данной зоне. Это имеет значение для невырожденных слаболегированных полупроводников, которые характеризуются малой эффективной массой носителей заряда при невысоких температурах и энергиях фотонов.

На практике при облучении тела фотопроводимость со временем увеличивается до некоторого стационарного значения, соответствующего равенству между скоростью генерации и рекомбинации неравновесных носителей заряда. Стационарная фотопроводимость зависит от времени жизни  $\tau$  и подвижности  $\mu$  соответствующих неравновесных носителей заряда, от коэффициента поглощения  $\alpha$  и квантового выхода  $\beta$ , равного числу пар или единиц носителей, образовавшихся при поглощении одного кванта излучения.  $\beta$  приближается к единице, когда энергия фотона составляет величину  $1-2 E_g$ . При больших значениях энергии фотонов  $h\nu$  квантовый выход может превышать единицу за счет вторичного явления ударной ионизации атомов «горячими» электронами с образованием дополнительных носителей.

Рост и спад фотопроводимости  $\Delta\sigma$  во времени происходят по экспоненциальному закону с константой  $\tau$ , равной времени жизни неравновесных носителей заряда (рис. 1.14 на с. 32), и описываются кривыми релаксации фотопроводимости (рис. 1.16 на с. 35) в соответствии с уравнениями (1.37) и (1.38), которые обсуждались в пункте 1.3.3.

Время жизни неравновесных носителей  $\tau$  определяется особенностями рекомбинационных процессов, а также участием ловушек

(центров временной локализации носителей), которые увеличивают эффективное время жизни неравновесных носителей и фотопроводимость.

Изменение концентрации носителей при освещении полупроводника приводит не только к изменению электропроводности, но и к возникновению электродвижущих сил (фото-ЭДС) в случае пространственного разделения фотогенерированных носителей заряда. Общей *причиной* возникновения фото-ЭДС в наиболее важных известных случаях является *диффузия* фотоэлектронов и фотодырок от места генерации (поверхности) в объем материала. Эта диффузия биполярна и характеризуется биполярным коэффициентом диффузии  $D_{\pm}$  неравновесных носителей, который связан с коэффициентами диффузии электронов  $D_e$  и дырок  $D_h$  соотношением (3.3):

$$D_{\pm} = (D_e\sigma_h + D_h\sigma_e)/(\sigma_h + \sigma_e), \quad (3.3)$$

где  $\sigma_h$  и  $\sigma_e$  – равновесные удельные электропроводности соответствующего типа.

Можно отметить следующие основные *условия* возникновения фото-ЭДС достаточно большой величины:

- 1) интенсивное поглощение фотонов, наблюдающееся в материалах в определенном диапазоне спектра электромагнитного излучения;
- 2) фотогенерация носителей обоих знаков или носителей знака, противоположного знаку основных равновесных носителей;
- 3) неоднородность полупроводника, что часто проявляется отличием концентрации носителей в поверхностном освещаемом слое от концентрации таковых в объеме материала.

В зависимости от структуры полупроводника и условий возникновения различают несколько типов фото-ЭДС. В однородном полупроводнике при освещении может возникать *ЭДС Дембера* в результате диффузии с разной скоростью фотоэлектронов (чаще более подвижных частиц) и фотодырок (рис. 3.5). Величина ЭДС Дембера  $E_d$  пропорциональна разности коэффициентов диффузии носителей, а следовательно, разности их подвижностей, и компенсирует разность диффузионных токов дырок и электронов. При этом возникающее электрическое поле ускоряет менее подвижные частицы и замедляет быстрые.

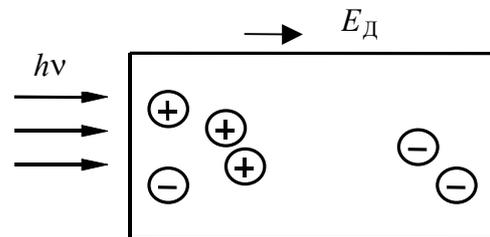


Рис. 3.5. Схема возникновения фото-ЭДС Дембера в однородном полупроводнике

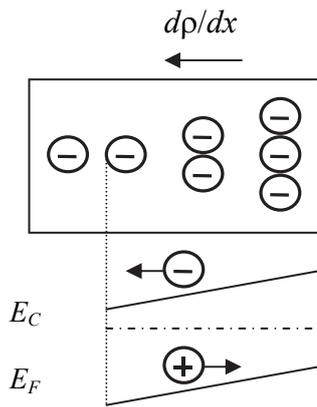


Рис. 3.6. Схема возникновения объемной фото-ЭДС в неоднородном полупроводнике

центрации активной примеси, а также квадрату диффузионной длины неосновных носителей. Для наиболее распространенных полупроводников значения вышеуказанных фото-ЭДС составляют единицы милливольт.

Наиболее важным с практической точки зрения типом фото-ЭДС является *вентильная* фото-ЭДС (фотогальванический эффект) (рис. 3.7). Она проявляется при освещении полупроводника и генерации неравновесных носителей заряда вблизи потенциального барьера, в качестве которого могут выступать функциональные контакты и просто поверхность полупроводника, где наблюдается приповерхностное искривление энергетических зон. Отличие этого типа фото-ЭДС состоит в том, что потоки диффундирующих неравновесных носителей заряда разделяются барьером, а не в результате различия свойств самих носителей (фото-ЭДС Дембера) или характеристик материала (объемная фото-ЭДС). Дополнительным условием возникновения вентильной фото-ЭДС является то, что расстояние от области генерации неравновесных носителей до барьера, например *p-n*-перехода, не должно превышать диффузионной длины этих носителей. В этом случае контактное поле разделяет диффундирующие фотоносители (пропускает неосновные и задерживает основные носители), что приводит к возникновению вентильной фото-ЭДС  $V_0$ ,

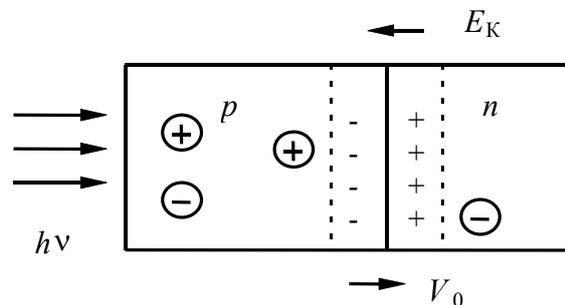


Рис. 3.7. Схема возникновения вентильной фото-ЭДС

полярность которой противоположна полярности контактной разности потенциалов  $U_K$  в  $p$ - $n$ -переходе. Максимальная величина  $V_0$  не превышает  $U_K$ , поскольку вентильная фото-ЭДС снижает потенциальный барьер на границе  $p$ - $n$ -перехода. Величина  $V_0$  описывается основным уравнением теории вентильных фотоэлементов (3.4) и может быть увеличена при обратном смещении  $p$ - $n$ -перехода, что используется в фотодиодах:

$$V_0 = \frac{kT}{e} \ln \left( 1 + \frac{I_c}{I_0} \right), \quad (3.4)$$

где  $I_c$  – световой ток неравновесных носителей заряда (фотоносителей) через  $p$ - $n$ -переход, А;  $I_0$  – ток насыщения равновесных носителей.

Спектр вентильной фото-ЭДС содержит максимум, что можно пояснить на основании уравнения (3.4). На возрастающем участке спектра при увеличении энергии падающих фотонов (и соответственно уменьшении длины волны падающего света) световой ток  $I_c$  в соответствии с уравнением (3.5) значительно увеличивается, поскольку растет темп генерации неравновесных носителей заряда  $g$ :

$$I_c = egS_{hv}\beta = ev_r\beta, \quad (3.5)$$

где  $g$  – темп генерации фотоносителей при освещении полупроводника, шт./м<sup>2</sup> · с;  $S_{hv}$  – площадь освещаемой поверхности полупроводника, м<sup>2</sup>;  $v_r$  – скорость генерации фотоносителей, шт./с;  $\beta$  – доля неосновных фотоносителей, дошедших до  $p$ - $n$ -перехода без рекомбинации.

На этом же участке ток насыщения  $I_0$  также увеличивается, однако существенно меньше, чем  $I_c$ , поскольку подавляющая часть равновесных носителей не могут преодолеть потенциальный барьер на границе  $p$ - $n$ -перехода. В результате на этом участке спектра  $I_c > I_0$  и с ростом энергии фотонов  $V_0$  увеличивается до некоторого максимального значения. При этом значении  $V_0$  потенциальный барьер на границе  $p$ - $n$ -перехода понизится до величины, сравнимой с тепловой энергией носителей заряда, и через  $p$ - $n$ -переход будет протекать большой ток равновесных носителей, концентрация которых значительно выше числа неравновесных носителей, и следовательно  $I_c < I_0$ . Дальнейшее увеличение энергии фотонов приведет к снижению  $V_0$ . Абсолютные значения максимальной вентильной фото-ЭДС наблюдаются в легированных полупроводниках, когда уровень Ферми близок к краю разрешенной зоны, и составляют величину порядка  $E_g/2$ , т. е. 0,3 В для германия и 0,5 В для кремния. Это на несколько порядков больше, чем величины других типов фото-ЭДС.

### 3.1.4. Излучение и усиление света в твердом теле

Излучение фотонов полупроводником может наблюдаться в результате процессов рекомбинации носителей заряда, непременно сопровождающих генерацию (и равновесную, и неравновесную), поскольку процесс возбуждения электронов является обратимым. Процессы излучения могут быть тепловыми (безызлучательная рекомбинация) и нетепловыми (излучательная рекомбинация или люминесценция). Под люминесценцией понимают излучение, избыточное над тепловым при данной температуре и обладающее длительностью, значительно превышающей период световых колебаний. Эта длительность называется временем послесвечения. Она составляет не более  $10^{-8}$  с для флуоресценции и значительно больше (до минут и часов) для фосфоресценции. В зависимости от способа возбуждения материала различают фото-, электро- и другие виды люминесценции. Поскольку в чистых материалах люминесцентные свойства проявляются обычно слабо, для интенсификации процессов излучения в возбужденном твердом теле используются активаторы в виде примесей, создающих разрешенные состояния в запрещенной зоне основного материала. Энергия излучения определяется природой и состоянием активатора. Такие легированные источники люминесцентного свечения называют кристаллофосфорами, которые включают широкозонный материал, активатор и легкоплавкие добавки.

При рассмотрении процесса усиления света в материале следует различать явления спонтанного и вынужденного излучения. Спонтанное излучение наблюдается в результате самопроизвольного перехода возбужденных, например светом, электронов в валентную зону, что сопровождается излучением фотонов с энергией  $h\nu = E_g$ . Таким образом, тепловое равновесие в теле поддерживается «порхающими» в нем фотонами, которые непрерывно излучаются и поглощаются при электронных переходах. Это пронизывающее равновесную систему излучение называется тепловым, и его спектральная плотность  $\rho(\nu)$ , т. е. объемная плотность энергии фотонов, приходящаяся на единичный интервал частот вблизи заданной частоты  $\nu$ , описывается знаменитой формулой Планка (3.6):

$$\rho(\nu) = \frac{8\pi^3 h \nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1}. \quad (3.6)$$

Вероятность теплового излучения описывается коэффициентом спонтанного излучения  $\alpha_{\text{сп}}$ , который является коэффициентом пропорциональности между скоростью излучения фотонов (шт./с) и концентрацией возбужденных электронов, находящихся в зоне проводимости полупроводника. Поскольку спонтанные переходы не связаны, то спонтанное излучение не когерентно.

*Вынужденное* излучение происходит в неравновесном материале с *инверсной заселенностью* разрешенных энергетических уровней, при которой часть зоны проводимости заполнена, а часть валентной зоны не заполнена (рис. 3.8). Такое состояние создают с помощью накачки материала светом, полем и другими воздействиями. При облучении такого инверсно заселенного материала фотонами с энергией  $h\nu = E_g$  поглощение этих фотонов невозможно, поскольку энергия разрешенного перехода больше, чем  $E_g$ , а испускание фотонов с энергией  $h\nu \geq E_g$  весьма вероятно. Иными словами, на выходе из материала поток фотонов будет больше, чем на входе. При таком вынужденном излучении выполняется соотношение Эйнштейна (3.7) между коэффициентом спонтанного излучения  $\alpha_{\text{сп}}$  и коэффициентом вынужденного излучения  $\beta_{\text{вын}}$ :

$$\alpha_{\text{сп}} = \beta_{\text{вын}} \frac{8\pi n^3 h\nu^3}{c^3}. \quad (3.7)$$

В результате наблюдается явление, которое по-английски называется *Light Amplification by Stimulated Emission of Radiation* (усиление света за счет испускания вынужденного излучения). Значит, мы получим усилитель света, который по первым буквам в английском наименовании явления называется *лазер* (laser). Лазер работает не только как усилитель, но и как генератор оптического излучения. Энергия излучения совпадает с энергией падающего света, т. е. излучение будет монохроматическим. Кроме того, если учесть, что фотоны являются бозонами (а не фермионами, как электроны), то при появлении фотона (бозона) в определенном квантовом состоянии сразу повышается вероятность

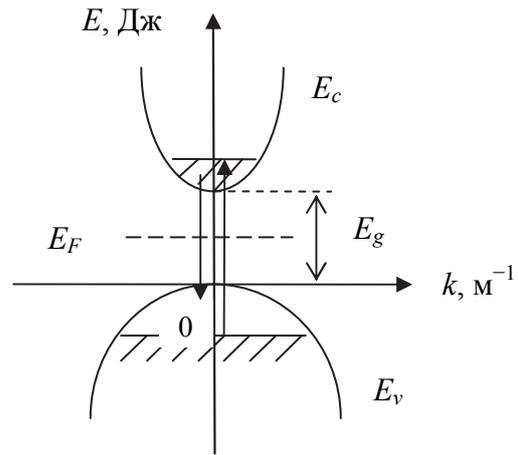


Рис. 3.8. Схема накачки и вынужденного излучения в материале с инверсной заселенностью уровней

появления в системе других бозонов в таком же квантовом состоянии. Процесс развивается очень быстро, поэтому все возникшие фотоны (бозоны) излучаются практически одновременно, т. е. лазерное излучение является упорядоченным, или когерентным, а также поляризованным, поскольку поляризация является одним из квантовых состояний фотона.

## 3.2. Оптоэлектронные полупроводниковые приборы (лекции 15–16)

Оптоэлектронные полупроводниковые приборы (ОПП) излучают или преобразуют электромагнитное излучение оптического диапазона (с длиной волны от 1 мм до 1 нм). Эти приборы можно разделить на следующие:

- *полупроводниковые излучатели* (когерентные (лазеры) и некогерентные (светодиоды));
- *приемники излучения* (фотодиоды, фототранзисторы, фоторезисторы, фототиристоры, солнечные элементы и др.);
- *оптроны* (оптопары и оптоэлектронные интегральные схемы).

Заметим, что среди указанных оптоэлектронных устройств целый ряд приборов на практике изготавливается на основе полупроводниковых материалов. Это светодиоды, фотодиоды, фототранзисторы, фоторезисторы, множество устройств интегральной оптики. Вместе с тем при создании различных типов лазеров (например, так называемых твердотельных) широко используются легированные кристаллические и аморфные диэлектрики. В данном пункте мы рассмотрим полупроводниковые электронные приборы, которые характеризуются технологической преемственностью с микроэлектронными приборами.

### 3.2.1. Светодиоды

Полупроводниковые излучатели используются для преобразования электрической энергии (сигнала) в световую. Обычно такое преобразование в приборе осуществляется посредством межзонной излучательной рекомбинации в полупроводнике неосновных носителей заряда, инжектированных через *p-n*-переход. Некогерентное излучение позволяют получать *светодиоды*, а когерентное – *полупроводниковые (и другие) лазеры*.

*Светодиоды* – это полупроводниковые приборы, основанные на *прямосмещенном контакте* (например, *p-n*-переходе), работающие на

эффекте спонтанной излучательной рекомбинации инжектированных носителей заряда (электролюминесценции) и предназначенные для генерации некогерентного оптического излучения. Светодиоды, излучающие в видимом диапазоне, называют *светоизлучающими диодами (СИД)*. Они используются большей частью как устройства отображения информации (для визуализации сигналов). Эти приборы изготавливаются на основе полупроводников, ширина запрещенной зоны которых больше 1,7 эВ (согласно соотношению  $E_g = h\nu = hc/\lambda$ , где  $\lambda = 0,38-0,78$  мкм соответствует видимому диапазону оптического излучения). Приборы, работающие в инфракрасной области, называют *ИК-диодами*. В них используются более узкозонные полупроводники ( $E_g < 1,7$  эВ).

ВАХ светодиода фактически представляет собой прямую ветвь *p-n*-перехода (рис. 2.7), а спектральные характеристики являются экстремальными.

Основными характеристиками светодиодов являются: длина волны генерируемого излучения  $\lambda$  (в зависимости от цвета свечения от 400 до 1000 нм), максимальное прямое напряжение  $U_{пр}$  (от 1,5 до 4,5 В) и максимальное обратное напряжение  $U_{обр}$  (около 5 В), прямой ток  $I$  (от 10 мкА до 10 мА) или соответствующая плотность тока  $j$  (А/м<sup>2</sup>) и яркость свечения  $B$  (кд/м<sup>2</sup>). Для современных светодиодов чаще используют более универсальную величину – световой поток  $F_{hv}$  (лм (люмен)):

$$B = \frac{I_{hv}}{S_{hv}}; F_{hv} = I_{hv} \Omega, \quad (3.8)$$

где  $I_{hv}$  – сила света, кд (кандела);  $S_{hv}$  – площадь излучающей поверхности, м<sup>2</sup>;  $\Omega$  – телесный угол, определяемый углом свечения  $\theta$ , ср (стерадиан):

$$\Omega = 2\pi(1 - \cos \frac{\theta}{2}). \quad (3.9)$$

Длина волны генерируемого излучения и соответственно цвет свечения определяется особенностями зонной структуры функционального полупроводника, а именно шириной запрещенной зоны или глубиной залегания рекомбинационных ловушек.

С развитием светодиодных приборов для оценки энерго-излучательной эффективности светодиодов получила распространение величина светоотдачи  $\eta_E$  (лм/Вт):

$$\eta_E = \frac{F_{hv}}{P}, \quad (3.10)$$

где  $F_{h\nu}$  – световой поток генерируемого светодионом излучения, лм;  $P$  – мощность, потребляемая светодиодом обычно в виде электрической энергии, Вт.

Светоотдача современных светодиодов превышает 100 лм/Вт, что приблизительно в 5 раз больше, чем у ламп накаливания.

Характеристикой эффективности полупроводниковых излучателей является также коэффициент полезного действия  $\eta$  (уравнение (3.11)), который равен отношению энергии генерированного излучения к энергии управления прибором:

$$\eta = \eta_{\text{вн}} \frac{h\nu}{eU_p} = \eta_{\text{вн}} \frac{E_g}{eU_p}, \quad (3.11)$$

где  $\eta_{\text{вн}}$  – внешний квантовый выход излучения:

$$\eta_{\text{вн}} = \eta_{h\nu} \eta_I \eta_0, \quad (3.12)$$

где  $\eta_{h\nu}$  – внутренний квантовый выход (отношение числа генерированных квантов излучения к числу носителей заряда, инжектированных через функциональный контакт);  $\eta_I$  – коэффициент инжекции носителей (отношение тока инжектированных носителей заряда к полному току через функциональный контакт);  $\eta_0$  – коэффициент вывода излучения (отношение числа фотонов, вышедших из прибора, к общему числу генерированных квантов излучения);  $E_g$  – ширина запрещенной зоны рабочего полупроводника, Дж;  $U_p$  – рабочее напряжение (прямое смещение функционального контакта), В.

Для повышения  $\eta_{h\nu}$  целесообразно использовать материалы с минимальной структурной и примесной дефектностью, которая способствуют реализации безызлучательной рекомбинации. Также предпочтительно применение прямозонных полупроводников, в которых скорость излучательной рекомбинации выше скорости безызлучательной. Теоретически перспективными являются прямозонные цинк- и кадмийсодержащие материалы класса  $A^{II}B^{VI}$ , которые находятся в стадии технологического освоения. Основными материалами для светодиодов являются полупроводники класса  $A^{III}B^V$ . ИК-диоды изготавливаются на основе эпитаксиальных пленок арсенида галлия ( $E_g = 1,4$  эВ), имеющих низкую дефектность. Эффективные СИД получают при использовании фосфида галлия ( $E_g = 2,3$  эВ) и твердых растворов на его основе ( $GaAs_{1-x}P_x$ ,  $Ga_xIn_{1-x}P$ ,  $Ga_xIn_{1-x}As_{1-y}P_y$ ). Действие таких приборов основано на излучательной рекомбинации экситонов, связанных на изоэлектронных ловушках. Значимый прогресс светодиодов

достигнут с использованием гетероструктур на базе нитрида галлия ( $E_g = 3,4$  эВ), который является прямозонным широкозонным полупроводником (табл. 3.1).

Таблица 3.1

Параметры светодиодов

Материал полупроводника	Примесь (или значения $x$ )	Цвет свечения	Спектральный максимум, нм	$\eta_e$ , %
GaAs	Si	ИК	950	12–50
GaAs	Zn	ИК	900	12–50
GaP	ZnO	Красный	690	7
$\text{GaAl}_x\text{As}_{1-x}$	0,3	Красный	675	1,3
$\text{GaAs}_x\text{P}_{1-x}$	0,39	Красный	660	0,5
$\text{Ga}_{1-x}\text{In}_x\text{P}$	0,6	Янтарный	617	0,1
GaP	N	Зеленый	550	0,7
GaN	–	Синий	466	$\approx 1$

Наиболее эффективное увеличение коэффициента инжекции  $\eta_i$  обеспечивается в гетеропереходах (из широкозонного в узкозонный полупроводник). В гомо- $p-n$ -переходах следует добиваться наибольших значений диффузионной длины неосновных носителей заряда и обеспечивать более низкий уровень легирования в активной области светодиода. Повышение  $\eta_0$  реализуется оптимизацией конструкции светодиодов при переходе от плоской к сферической и эллиптической форме (рис. 3.9), что уменьшает внутреннее отражение на границе полупроводник – окружающая среда. Кроме того, активная область прибора должна иметь толщину порядка нескольких микрон для предотвращения активного поглощения генерированных фотонов в объеме полупроводника. Однако даже при всех приведенных усовершенствованиях  $\eta_0$  не превышает 30%.

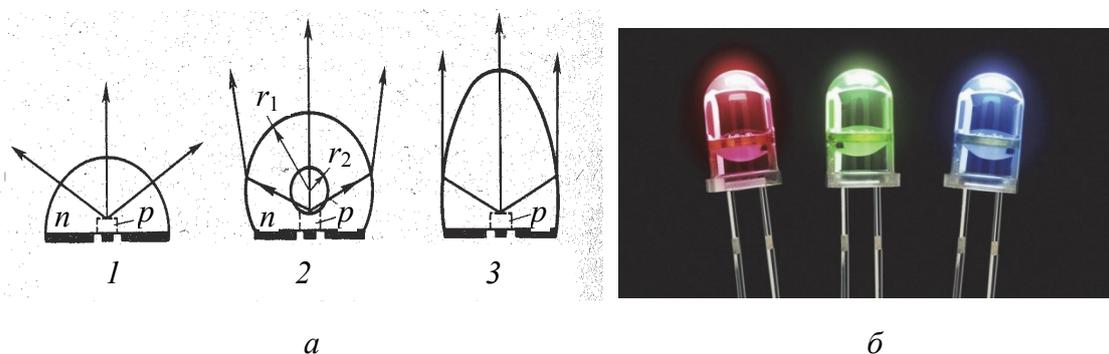


Рис. 3.9. Конфигурации (а) и внешний вид (б) светодиодов:  
1 – полусфера; 2 – усеченная сфера; 3 – эллипсоид

Спектр излучения светодиода определяется шириной запрещенной зоны используемого полупроводника, природой и количеством примесей, механизмом излучательной рекомбинации. Использование твердых растворов и индивидуальных соединений позволяет создавать светодиоды для различных диапазонов спектра электромагнитного излучения (рис. 3.10).

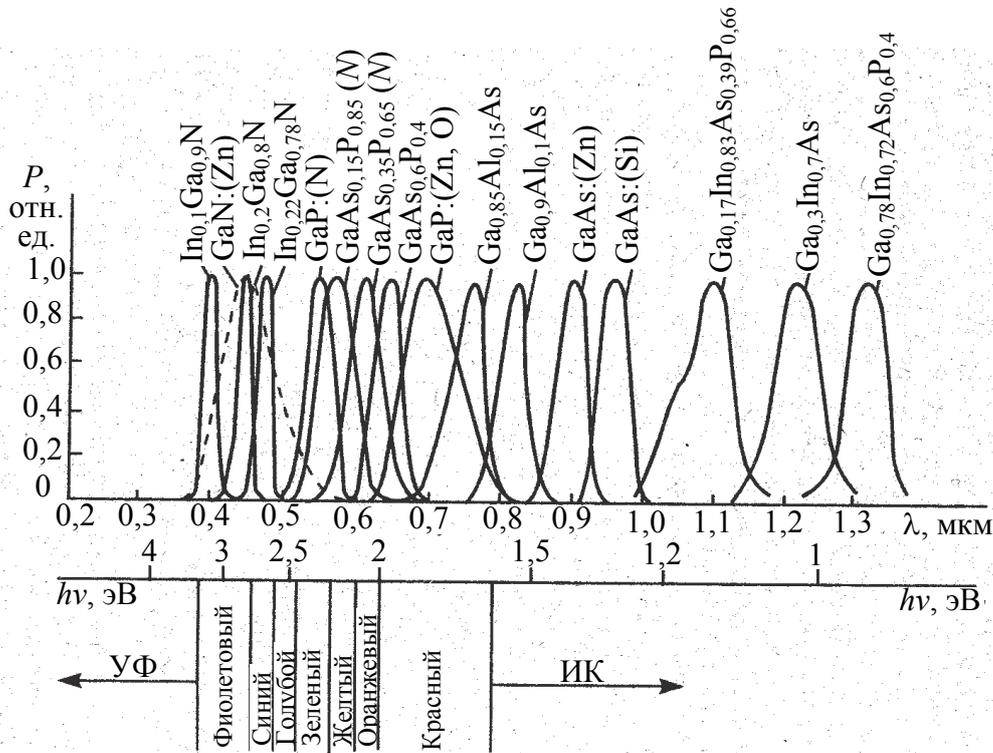


Рис. 3.10. Спектры излучения светодиодов на основе различных полупроводниковых материалов

В наиболее эффективных арсенидгаллиевых светодиодах активную  $p$ -область создают введением цинка или кремния (табл. 3.1). В первом случае достигается наибольшее быстродействие прибора, а во втором – эффективность. Очень важно, что спектр излучения GaAs-светодиодов хорошо соответствует (имеет максимум при той же длине волны) спектру фоточувствительности наиболее распространенных Si-фотодиодов.

Некогерентное оптическое излучение может быть получено также в помощью порошковых и пленочных *электролюминесцентных излучателей*. Такие приборы (рис. 3.11) представляют собой многослойную структуру, включающую два электрода (обычно оксидных), слой люминофора (чаще сульфид цинка, легированный марганцем) и

защитный диэлектрический слой (для порошковых приборов). Возбуждение электронов в материале происходит при наложении электрического поля и сопровождается последующей рекомбинационной люминесценцией.

Длительность послесвечения определяется временем жизни неосновных носителей заряда и наличием ловушек захвата. Порошковые приборы отличаются от пленочных тем, что в первом случае частички электролюминофора распределены в диэлектрической связке, что обуславливает необходимость использования в качестве возбуждающего сигнала переменного напряжения до 200 В с частотой порядка 1 кГц.

Пленочные излучатели на основе однородной вакуумной пленки являются менее экономичными, однако имеют высокую разрешающую способность, длительность послесвечения и стабильность. Среди общих достоинств электролюминесцентных излучателей можно отметить возможность высвечивания больших площадей (десятки квадратных метров), многоцветность, компактность, яркость, простоту изготовления.

### 3.2.2. Лазеры

*Лазеры*, в том числе и *полупроводниковые*, представляют собой квантовые генераторы или усилители оптического когерентного излучения, основанные на использовании вынужденной излучательной рекомбинации. Они преобразуют электрическую энергию или энергию некогерентного излучения в энергию когерентного излучения.

Как уже отмечалось, условием усиления света за счет вынужденного излучения является наличие материала с инверсной заселенностью электронных уровней (рис. 3.8). Такой материал называют лазерной средой, которая может быть и твердой, и жидкой, и газовой.

Первая удачная попытка создания когерентного излучателя была осуществлена в 1954 г. одновременно в Физическом институте Академии наук (СССР) советскими физиками Николаем Басовым

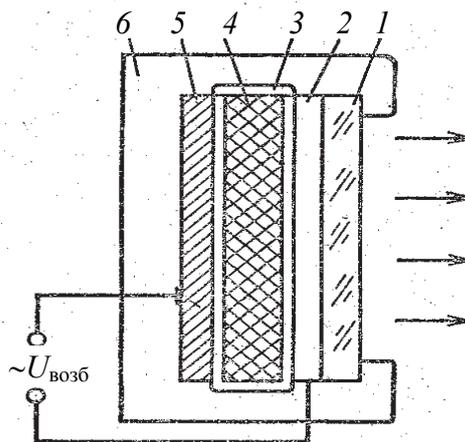


Рис. 3.11. Структура электролюминесцентного индикатора:

- 1 – стекло; 2 – прозрачный электрод;
- 3 – изолирующая пленка; 4 – слой люминофора; 5 – металлический электрод; 6 – корпус

(1922–2001), Александром Прохоровым (1916–2002) и в Колумбийском университете американским физиком Чарлзом Таунсом (1915–2015), за что трое исследователей были удостоены в 1964 г. Нобелевской премии по физике. В качестве рабочего вещества в этом первом устройстве использовался аммиак. Правда, излучал он не свет, а радиочастотные волны (microwave), и был назван Таунсом мазером. Благодаря способности усиливать радиоволны мазеры сразу начали применяться в радиотелескопах.

Первый собственно лазер был создан в 1960 г. американским физиком Теодором Мейманом (1927–2007). Рабочим телом в этом лазере был стержень из диэлектрического рубина, активированного ионами хрома ( $\text{Al}_2\text{O}_3:\text{Cr}^{3+}$ ). Широкое распространение имеет твердотельный лазер на иттрий-алюминиевом гранате  $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Nd}^{3+}$  благодаря низким пороговым энергиям накачки (200–400 Вт), высокой теплопроводности и твердости матрицы, малым оптическим потерям, простоте конструкции, высокой мощности и частоте следования импульсов излучения (до 10 кГц). Общим недостатком диэлектрических лазерных матриц, активированных редкоземельными металлами, является отсутствие широких полос поглощения. Поэтому в такие матрицы, кроме ионов-активаторов, вводят ионы-сенсбилизаторы, например ионы  $\text{Cr}^{3+}$ , которые поглощают энергию в широком спектральном диапазоне и отдают ее ионам-активаторам.

Особый класс твердотельных лазеров представляют полупроводниковые лазеры (рис. 3.12). *Полупроводниковые лазеры* – это квантово-электронные приборы, основанные на использовании полупроводника с *инверсной заселенностью электронных уровней*, работающие на эффекте *вынужденной излучательной рекомбинации* носителей заряда и предназначенные для получения *когерентного оптического излучения*.

Эти лазеры отличаются тем, что они миниатюрны, высокоэффективны, легко управляются электрическим сигналом, работают при низких напряжениях, устойчивы к механическим воздействиям, имеют высокое быстродействие, большой срок службы и, что очень важно, изготавливаются по стандартной технологии микроэлектроники и поэтому дешевы.

В полупроводниковых лазерах состояние инверсной заселенности электронных уровней можно создать за счет инжекции носителей через прямосмещенный *p-n*-переход (рис. 3.12). В таком случае около перехода появляется область перекрытия, где велики концентрации как электронов, так и дырок. Эта область называется активной (так как именно здесь происходит излучательная рекомбинация электронно-дырочных пар). Чтобы поддерживать процесс излучения, необхо-

димо компенсировать убыль рекомбинировавших электронов и дырок, пропуская через устройство электрический ток для инжекции в активную область новых носителей. Такой прибор называют «лазерным диодом» (laser diode)). Для обеспечения эффективного взаимодействия света с активной средой в нем необходимо совместить область инверсной заселенности с областью светового излучения, т. е. локализовать в одном слое неравновесные носители заряда и фотоны.

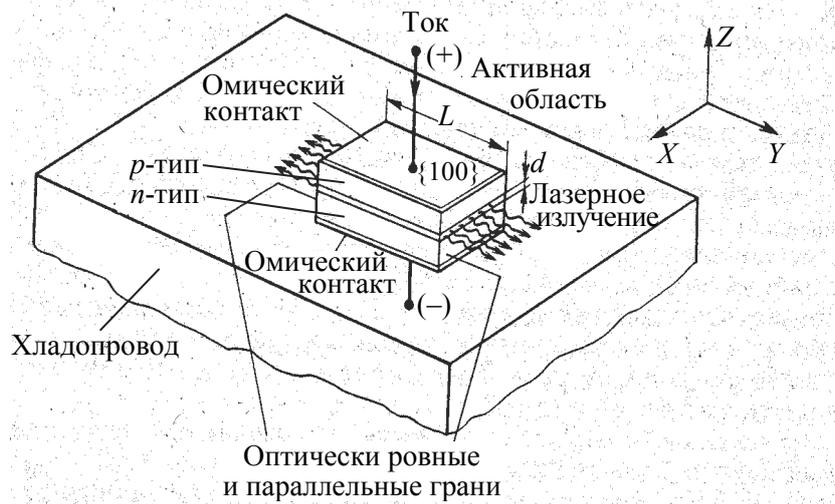


Рис. 3.12. Упрощенная схема инжекционного лазера

В структурах на основе гомопереходов эта задача реализуется неоптимально, поэтому в настоящее время наибольшая доля промышленного выпуска приходится на инжекционные лазеры на основе двойных гетероструктур (ДГС) (рис. 3.13) в сложных системах  $Al_xGa_{1-x}As/GaAs$  и  $Ga_xIn_{1-x}As_{1-y}P_y/InP$ . В таких структурах эффекты односторонней инжекции, сверхинжекции, волноводный эффект позволяют обеспечить не только электронное (носители), но и оптическое (фотоны) пространственное ограничение в пределах активной области. В частности, за счет эффекта сверхинжекции концентрация электронов в активной области выше, чем в  $n$ -эмитторе. Вместе с тем потенциальный барьер на границе  $p-p^+$  запирает электроны в пределах узкозонной активной области гетероструктуры. Аналогичный эффект наблюдается для дырок. Эффективное оптическое ограничение достигается, поскольку показатель преломления полупроводников резко меняется на границах гетеропереходов и активный слой ведет себя как волновод.

На современном этапе полупроводниковые лазеры создаются большей частью на основе полосковых или скрытых ДГС (рис. 3.13), а также ДГС с распределенной обратной связью (РОДГС).

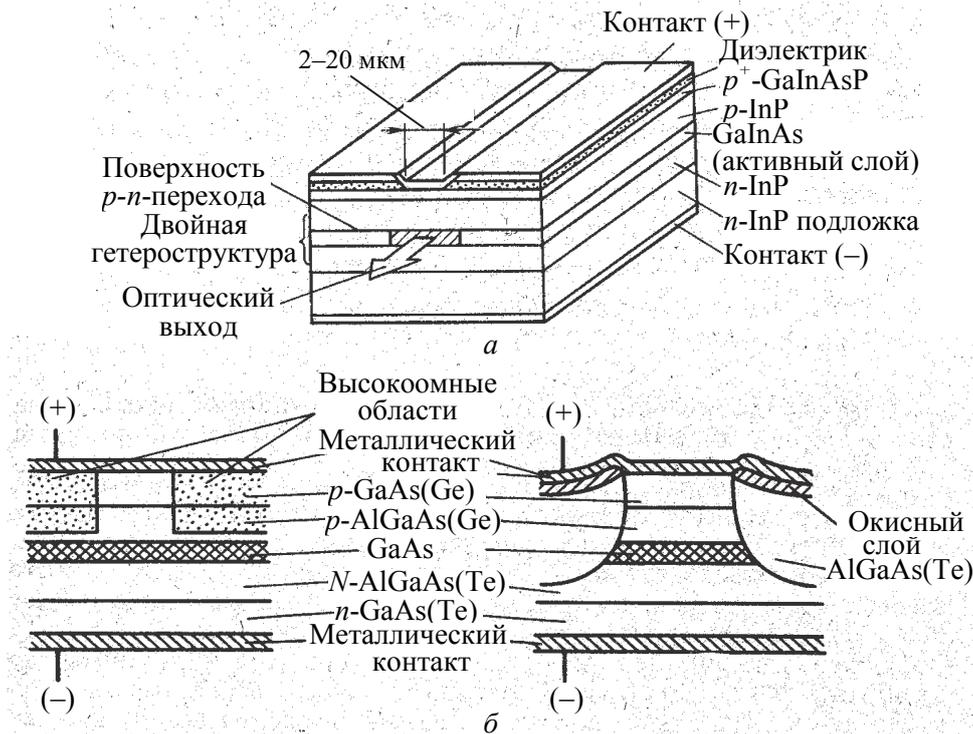


Рис. 3.13. Структуры контактных ДГС-полосковых лазеров: *a* – контактная структура с однокоординатным ограничением (вдоль оси *z*); *б* – нитевидная структура с двухкоординатным ограничением (вдоль осей *x* и *z*)

Рабочая длина волны излучения лазера определяется составом активного слоя полупроводника (табл. 3.2). Она зачастую выбирается в соответствии с тремя областями пропускания оптических волокон. Используя твердые растворы, можно гибко варьировать характеристики лазеров.

Таблица 3.2

Длина волны излучения полупроводниковых лазеров

Состав активной среды	Длина волны излучения, мкм
$Al_{0,05}Ga_{0,95}As$ или GaAs	0,84
$Ga_{0,28}In_{0,72}As_{0,6}P_{0,4}$	1,3
$Ga_{0,4}In_{0,6}As_{0,88}P_{0,12}$	1,55

В ДГС-полосковом лазере (рис. 3.13, *a*) ширина активной области ограничена полосковым омическим контактом, что обеспечивает понижение рабочего тока и повышает устойчивость работы лазера. В такой многослойной структуре три центральных слоя образуют рабочую ДГС-структуру, включающую широкозонный *N*-эмиттер, активный

излучающий нелегированный слой, широкозонный  $P$ -эмиттер. Данные структуры обеспечивают электронное и оптическое ограничение только вдоль оси  $z$ . Лучшими характеристиками обладает лазер с двухкоординатным ограничением и вдоль оси  $z$ , и в плоскости инжекционного  $p$ - $n$ -перехода, т. е. вдоль оси  $x$  (рис. 3.13, б).

Ключевой выходной характеристикой лазеров является ватт-амперная характеристика (рис. 3.14), отражающая зависимость выходной мощности  $P_{\text{вых}}$  (в виде потока излучения) от рабочего тока  $I_p$  в лазерной структуре. Эта зависимость резко увеличивает свою крутизну при достижении значений порогового тока накачки  $I_{\text{пор}}$ , при котором генерируемая энергия превышает энергетические затраты на активацию (накачку) лазерной среды.

Дальнейшее совершенствование характеристик инжекционного полупроводникового лазера было достигнуто в ДГС-лазерах с разделным электронным и оптическим ограничением (РОДГС-лазеры). В таких структурах, пример которых приведен в пункте 2.5 (рис. 2.36), внутри нелегированного активного слоя, обеспечивающего оптическое ограничение, формируют периодические структуры, например сверхрешетки, где обеспечивается электронное ограничение. При пропускании тока в прямом направлении электроны и дырки из широкозонных эмиттеров инжектируются в протяженный нелегированный слой и затем попадают в тонкий еще более узкозонный активный слой, в котором собственно и происходит излучательная рекомбинация. В современных РОДГС-лазерах в качестве тонких (5–20 нм) активных слоев, где происходит накопление неравновесных носителей и их излучательная рекомбинация, используются квантоворазмерные структуры (например, квантовые ямы). В квантоворазмерных структурах для электрона характерной физической длиной является длина волны де Бройля. Если характерный геометрический размер (например, толщина активного слоя  $d$  на рис. 2.36) соизмерим или меньше длины волны де Бройля для электрона, то проявляются эффекты размерного квантования, т. е. характер квантования энергии электрона, ее величина зависят не только от природы материала, но и от его размеров. Это значит, что энергией уровней и волновыми функциями элект-

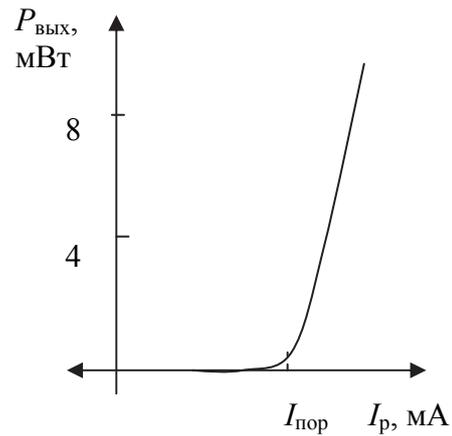


Рис. 3.14. Ватт-амперная характеристика лазера

тронов можно управлять, изменяя размеры объекта (например, активного слоя). Применение квантоворазмерных структур в активном слое РОДГС-лазера позволяет:

- уменьшить пороговую плотность тока накачки (около  $100 \text{ А/см}^2$ , что на порядок меньше, чем в ДГС-лазерах);
- ослабить температурную зависимость плотности тока накачки;
- увеличить коэффициент усиления излучения на единицу длины активной области;
- улучшить спектральные параметры генерируемого излучения.

За многие десятилетия с момента создания лазеров они нашли широкое применение в самых различных областях человеческой деятельности. Мощное излучение лазеров используется в промышленных и медицинских технологиях, в устройствах волоконно-оптической связи. Узкая направленность лазерного луча позволяет измерить расстояние до Луны с точностью до сантиметра. Голография начала по-настоящему развиваться только с применением монохроматического лазерного излучения.

### 3.2.3. Фоторезисторы

*Фоторезистор* – оптоэлектронный прибор, основанный на использовании *фоточувствительного материала*, работающего на *фоторезистивном эффекте*, предназначенный для *детектирования оптического излучения*. В зависимости от реализуемого в полупроводнике вида фотопоглощения выделяют собственные фоторезисторы, функционирующие в видимой и ближней ИК-области спектра, и примесные фоторезисторы, работающие в средней и дальней области при глубоком охлаждении. Эти приборы отличаются простой конструкцией, включающей активный слой полупроводника с высоким исходным сопротивлением  $R_{\text{темн}}$  (чаще всего соединение класса  $A^{II}B^{VI}$ ) и электроды (рис. 3.15).



Рис. 3.15. Схема фоторезистора

Фототок в фоторезисторе определяется тремя основными процессами:

- неравномерная по толщине чувствительного слоя генерация неравновесных носителей заряда под действием света;
- перенос носителей заряда посредством дрейфа и диффузии с возможным умножением по внутренним механизмам;
- взаимодействие тока с внешней цепью, которое обеспечивает выходной сигнал фоторезистора.

ВАХ фоторезистора в большей части линейна (рис. 3.16, *a*). Отклонение от линейности при малых напряжениях в случае использования поликристаллического полупроводника связано с контактными явлениями на межкристаллитной границе. Нелинейность в области больших напряжений обусловлена нагревом полупроводника. Спектральная характеристика (рис. 3.16, *б*) характеризуется наличием максимума, положение которого для собственных фоторезисторов определяется шириной запрещенной зоны полупроводника. Спад фототока при уменьшении длины волны ниже порогового значения связан с уменьшением глубины проникновения фотонов в полупроводник и увеличением поверхностной рекомбинации фотоносителей, что уменьшает время их жизни.

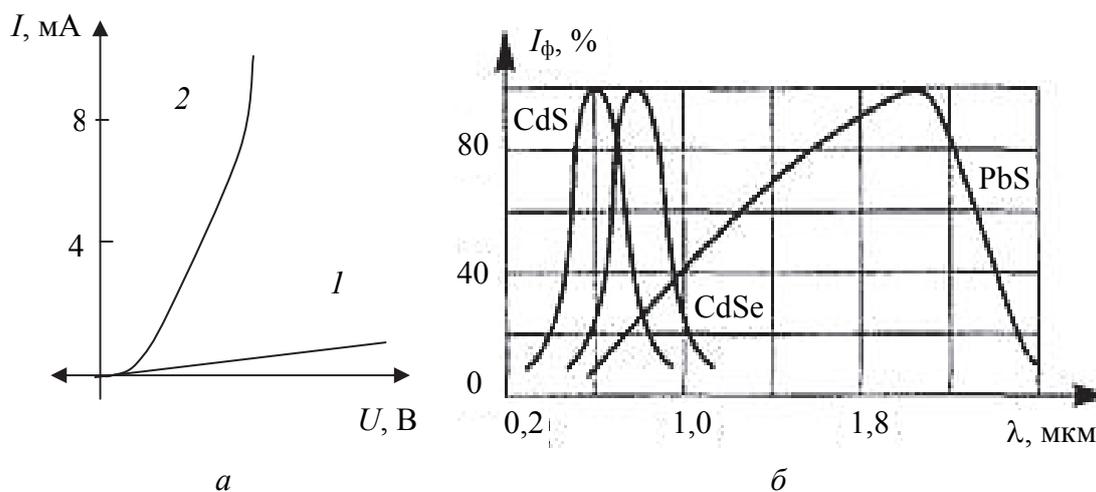


Рис. 3.16. ВАХ (*a*) и спектральные характеристики (*б*) фоторезисторов

Эффективность прибора определяется коэффициентом усиления фототока  $K_I$ , который пропорционален отношению числа носителей, образующих фототок, к числу поглощенных фотонов. С другой стороны,  $K_I$  прямо пропорционален времени жизни неосновных носителей и обратно пропорционален времени пролета носителей через активную область фоторезистора (между электродами). Таким образом,

увеличение  $K_I$  достигается ростом времени жизни носителей (например, за счет создания ловушек), однако при этом увеличивается постоянная времени  $\tau$  (время изменения тока фоторезистора в  $e$  раз), т. е. снижается быстродействие прибора.

При оценке эффективности фотоэлектрических детекторов (фоторезисторов и фотодиодов) наиболее информативной величиной является детектирующая способность прибора  $D_{hv}$  ( $\text{см} \cdot \sqrt{\text{Гц}}/\text{Вт}$ ), которая может быть определена по формуле

$$D_{hv} = \frac{\sqrt{S_{hv}} \sqrt{\Delta f}}{P_{hv} K_{ш}}, \quad (3.13)$$

где  $S_{hv}$  – площадь поглощающей поверхности полупроводника,  $\text{см}^2$ ;  $\Delta f$  – частотная полоса пропускания прибора, определяемая рабочим спектральным диапазоном, Гц;  $P_{hv}$  – мощность поглощаемого излучения, Вт;  $K_{ш}$  – коэффициент шума, определяемый отношением шумового сигнала к рабочему сигналу (току или напряжению).

В целом можно отметить, что достоинствами фоторезисторов как детектирующих приборов является низкая стоимость и достаточно высокий коэффициент усиления (порядка 1000–10 000). Однако большая инерционность прибора ограничивает его применение в оптоэлектронике. Тем не менее целый ряд фоторезисторов находит практическое применение для регистрации достаточно медленно протекающих процессов (частота модуляции светового потока до 10 кГц) в различных областях спектра (табл. 3.3).

Таблица 3.3

**Параметры фоторезисторов**

Материал	$T_{\text{раб}}$ , К	$\Delta\lambda_{\text{раб}}$ , мкм	$\lambda_{\text{макс}}$ , мкм	$D$ , $\text{см} \cdot \text{Гц}^{1/2}/\text{Вт}$	$\tau$ , мс	$R_{\text{темн}}$ , МОм	Примеры
CdS	295	0,4–0,8	0,6	$10^{14}$	10–200	0,1–100,0	ФСК-1
CdSe	295	0,4–1,0	0,7	$10^{13}$	10–100	2–1000	ФСД-1
PbS	295 80	1–3 1–4	2,2 3,2	$10^{11}$ $2 \cdot 10^{11}$	0,4 3,0	0,02–5,00	ФСА-1 ФСВ-18
PbSe	295 80	0,6–4,5 1,0–6,5	3,7 5	$10^{10}$ $2 \cdot 10^{10}$	0,01 0,04	0,2–1,0	СФ-1Д
InSb	80	2,0–5,6	5,2	$5 \cdot 10^{10}$	0,005	0,002–0,020	ФРО-146
Ge (Au)	80	2,4–7,6	5,9	$3 \cdot 10^9$	30 нс	0,03–1,00	ФСГ-22
Ge (Cu)	15	2–28	24	$2 \cdot 10^{10}$	10 нс		ФСГ-28
Hg <sub>0,8</sub> Cd <sub>0,2</sub> Te	80	8–13	12	$5 \cdot 10^{10}$	0,001	50–2000 Ом	ФРО-Х1

### 3.2.4. Фотодиоды

Фотодиод – полупроводниковый диод, основанный на *обратно-смещенном контакте* (*p-n-переходе*, барьере Шоттки, гетеропереходе), работающий на возникновении *вентильной фото-ЭДС* и предназначенный для *детектирования оптического излучения*. Фактически обратный ток фотодиода зависит от освещенности (рис. 3.17).

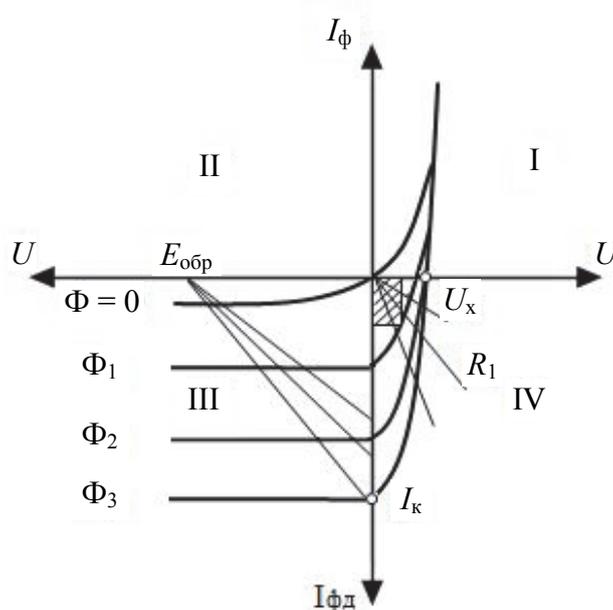


Рис. 3.17. ВАХ фотодиода (область III) и фотоэлемента (область IV)

Работа этого прибора связана с генерацией фотоносителей вблизи потенциального барьера на расстоянии меньшем, чем диффузионная длина, и с разделением потока этих носителей за счет возникновения вентильной фото-ЭДС. В результате обратный ток фотодиода при освещении увеличивается (рис. 3.17). Спектральные характеристики фотодиодов схожи с фоторезистивными (рис. 3.16). Для обеспечения высокой чувствительности и быстродействия фотодиода необходимо, чтобы генерация электронно-дырочных пар происходила в ОПЗ.

В качестве быстродействующих фотоприемников в диапазоне частот до нескольких гигагерц (табл. 3.4) наиболее эффективно использование *p-i-n-* и лавинных фотодиодов (рис. 3.18). При обратном смещении *p-i-n-*структуры обедненный слой захватывает всю *i-*область, что снижает емкость перехода, расширяет диапазон поглощения света и повышает чувствительность прибора. Фотоносители ускоряются в

обедненном слое электрическим полем (до скорости дрейфа  $\approx 10^7$  см/с) и тормозятся в  $p$ - и  $n$ -областях (до скорости диффузии  $\approx 10^4$  см/с), что уменьшает быстродействие и за счет диффузионной рекомбинации – квантовый выход. Таких нежелательных эффектов избегают, уменьшая толщину «входного» слоя до долей микрона и увеличивая толщину  $i$ -области свыше длины поглощения света ( $1/\alpha$ ), чтобы поглощение происходило в обедненном слое.

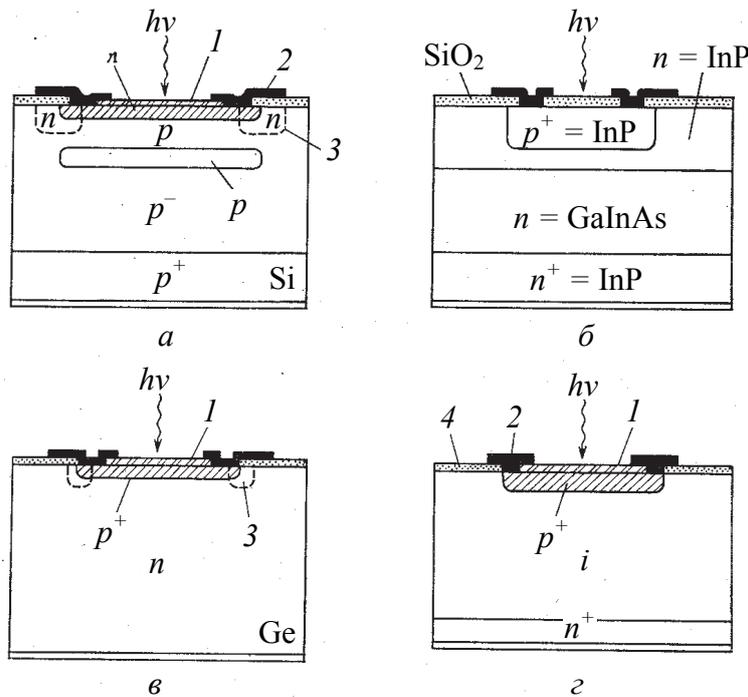


Рис. 3.18. Конструкции лавинных (а, б, в) и  $p$ - $i$ - $n$ - (г) фотодиодов:  
 1 – просветляющая пленка; 2 – электрод;  
 3 – защитное кольцо; 4 – изоляция

Лавинный фотодиод имеет более сложную структуру (рис. 3.18, б). Он содержит вблизи возникающего при обратном смещении прибора обедненного слоя (области поглощения) область с повышенной концентрацией носителей (высоким электрическим полем), в которой происходит лавинное умножение фотоносителей за счет ударной ионизации. Эффект лавинного умножения фотоносителей позволяет повысить быстродействие прибора. Реальные лавинные и  $p$ - $i$ - $n$ -фотодиоды создаются на базе различных структур (например,  $n^+p^-pp^-p^+$ ), которые включают активные области, электроды, просветляющую пленку (для предотвращения отражения), защитное кольцо (для увеличения напряжения пробоя и обеспечения однородности лавинного умножения).

Параметры фотодиодов

Материал, тип диода	$T_{\text{раб}}, \text{K}$	$\Delta\lambda_{\text{раб}},$ мкм	$\lambda_{\text{макс}},$ мкм	$D,$ $\text{см} \cdot \text{Гц}^{1/2}/\text{Вт}$	$\tau, \text{мс}$	Примеры
Si	295	0,4–1,0	–	–	–	ФДК,
<i>p-n</i>	–	–	0,85	$3 \cdot 10^{12}$	1–10	ФД-10КП,
<i>p-i-n</i>	–	–	0,9	$3 \cdot 10^{13}$	0,002–0,010	ФД-256
ЛФД	–	–	0,9	$10^{14}$	0,002–0,010	
Ge	295	0,5–1,8	1,55	–	–	ФД-1,
<i>p-n</i>	–	–	–	$5 \cdot 10^{10}$	1–10	ФД-4Г
ЛФД	–	–	–	$10^{12}$	<0,1	
GaAs	300	0,5–0,9	0,8	$10^{12}$	0,1	–
InSb	80	2,2–5,4	5,2	$5 \cdot 10^{10}$	5	ФДО-117
$\text{Ga}_x\text{In}_{1-x}\text{As}/\text{InP}$	300	1,3–1,6	1,5	$10^{11}$	0,005	–
$\text{Cd}_x\text{Hg}_{1-x}\text{Te}$	80	2–30	8–14	$10^{10}$	0,01	–

Высокое быстродействие (доли наносекунд), простота и широкая спектральная характеристика отличает фотодиоды на основе контакта металл – полупроводник. Широкий спектральный диапазон такого прибора обусловлен, с одной стороны, возможностью поглощения в верхнем металлическом слое (электроде) фотонов с энергией, меньшей ширины запрещенной зоны полупроводника. С другой стороны, с увеличением энергии квантов и коэффициента поглощения продолжается эффективное поглощение фотонов в области пространственного заряда полупроводника.

### 3.2.5. Фотоэлементы

*Фотоэлемент* (или *солнечный элемент*) – полупроводниковый диод, основанный на *функциональном контакте* (*p-n*-переходе, барьере Шоттки, гетеропереходе), работающий на возникновении *вентильной фото-ЭДС* и предназначенный для прямого преобразования световой энергии в электрическую.

Фотоэлемент подобно фотодиоду является примером фотоэлектронного приемника, работающего в режиме генерации вентильной фото-ЭДС (или в фотовольтаическом режиме). Возникновение в структуре прибора при освещении фото-ЭДС эквивалентно некоторому прямому смещению функционального контакта. ВАХ фотоэлемента (рис. 3.17, область IV) подчиняется уравнению (3.16), в основе которого лежит уравнение Шокли (2.3), описывающее ВАХ *p-n*-перехода:

$$I = I_{ph} - I_0 \left[ \exp \frac{e(U + IR_s)}{nkT} - 1 \right] - \frac{U + IR_s}{R_{sh}}, \quad (3.14)$$

где  $I$  – ток, протекающий в приборе, мА;  $I_{ph}$  – фототок, генерируемый светом, т. е. ток неравновесных носителей, мА;  $I_0$  – ток насыщения функционального контакта, т. е. ток равновесных носителей, мА;  $U$  – напряжение, генерируемое прибором, В;  $R_s$  – последовательное сопротивление, Ом;  $n$  – коэффициент идеальности диодной структуры;  $R_{sh}$  – сопротивление шунтирования, Ом.

Основными параметрами фотоэлементов (рис. 3.17) являются ток короткого замыкания  $I_k$  (или плотность тока  $i_k$ ), напряжение холостого хода  $U_x$ , коэффициент заполнения световой ВАХ КЗ

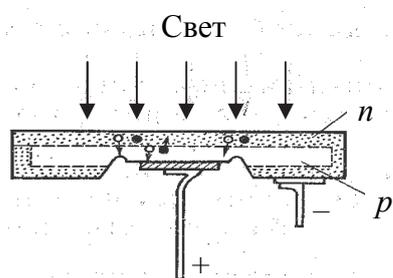


Рис. 3.19. Структура кремниевого фотоэлемента

(определяется по заштрихованной области рис. 3.17), максимальная мощность  $P_{max}$  и коэффициент полезного действия  $\eta$ , отражающий эффективность преобразования световой энергии. Максимальная величина возникающей при облучении в приборе фотоЭДС соответствует точке пересечения ВАХ с осью напряжения  $U_x$ . Она зависит от освещенности фотоэлемента, т. е. интенсивности облучающего светового потока, от коэффициента поглощения света материалом и может быть увеличена за счет повышения площади рабочей поглощающей поверхности (рис. 3.19).

Величины  $I_k$  и  $U_x$  определяют выходную мощность солнечного преобразователя, которая в наиболее эффективных приборах достигает нескольких киловатт. Важнейшей итоговой характеристикой фотоэлемента является коэффициент  $\eta$ , который определяется отношением выходной мощности прибора к входной мощности светового потока. Разработанные к настоящему времени фотоэлементы имеют невысокий КПД (от 0,1 до 20,0%) за счет значительного влияния процессов поверхностного отражения света, поверхностной и объемной рекомбинации неравновесных носителей, а также потерь мощности при прохождении тока в объеме полупроводника (уравнение (3.14)).

Наиболее распространенным материалом для солнечных преобразователей пока остается кремний (моно-, поликристаллический или даже аморфный). Это обусловлено тем, что максимум спектральной

характеристики кремниевых фотоэлементов соответствует максимуму спектрального распределения солнечной энергии. Перспективы полупроводниковой солнечной энергетики связывают с приборами на основе функциональных гетеропереходов. Создание фотоэлементов на базе гетеропереходов (например, GaAs/Ge, GaAs/GaP) позволяет увеличить КПД до 30–45%, а также расширить спектральную характеристику фотоэлемента. В таком случае максимум спектральной кривой преобразуется в плато, соответствующее диапазону энергий между величинами ширины запрещенной зоны используемых полупроводников. Наиболее высокий КПД (порядка 45%) к настоящему времени зафиксирован в фотовольтаических структурах на основе монокристаллического арсенида галлия. Однако такие материалы слишком дороги для массового применения.

Активно совершенствуются эффективные и экономичные тонкопленочные солнечные элементы на основе поликристаллических халькогенидных полупроводников, например CdTe,  $\text{CuIn}_{1-x}\text{Ga}_x\text{Se}_2$ . Одним из направлений такого развития является замена токсичных селена и кадмия, а также дорогих индия и теллура, например, на олово (II) и серу в составе поглощающих слоев SnS,  $\text{Cu}_2\text{ZnSnS}_4$  и других материалов, которые предлагается использовать в составе многослойных гетероструктур (рис. 3.20).

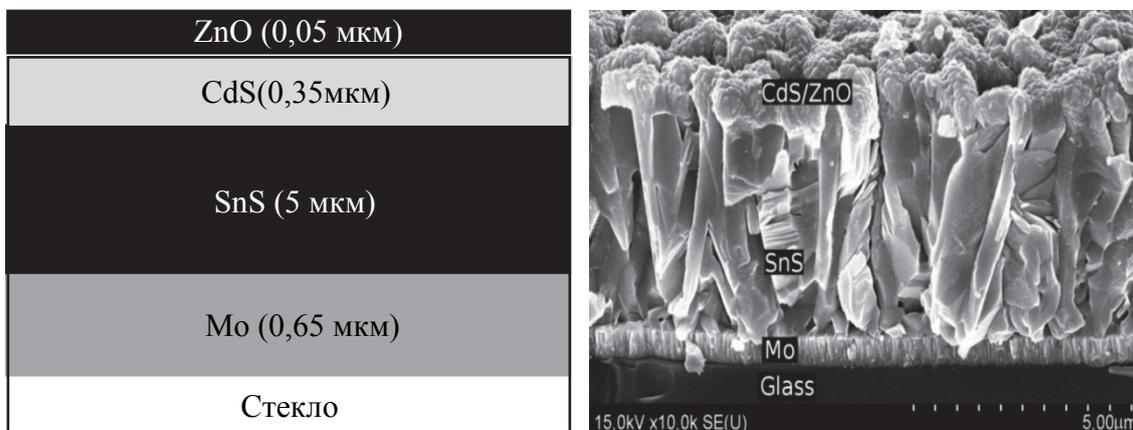


Рис. 3.20. Принципиальная схема и электронномикроскопическое изображение фотовольтаической гетероструктуры Mo/p-SnS/n-CdS/ZnO на основе гетероперехода SnS/ZnO

В таких гетероструктурах для солнечных элементов на основе поглощающего слоя SnS в качестве слоя, выполняющего функцию широкозонного окна, активно изучаются такие широкозонные полупроводники, как халькогениды CdS, ZnS,  $\text{Cd}_x\text{Zn}_{1-x}\text{S}$ , а также оксиды ZnO,  $\text{TiO}_2$ ,

SnO<sub>2</sub> и другие материалы. Совершенствование фотовольтаических гетероструктур продолжается как в направлении усложнения состава функциональных полупроводников, так и в направлении увеличения количества слоев в гетероструктуре. Показано, например, что существенного повышения квантовой эффективности SnS-элементов можно достичь при использовании широкозонного слоя состава ZnIn<sub>2</sub>Se<sub>4</sub> (табл. 3.5).

Таблица 3.5

**Фотовольтаические показатели гетероструктур на основе SnS**

Состав гетероструктуры	$U_k$ , мВ	$i_k$ , мА/см <sup>2</sup>	КЗ	$\eta$ , %
SnS/CdS	170	7,2	0,28	1
SnS/ZnO	83	2,7	1,4	<1
SnS/ZnIn <sub>2</sub> Se <sub>4</sub>	472	16,1	0,38	2,9

Максимальное значение КПД порядка 5% для фотоэлементов на основе SnS, которые называют устройствами SnS-энергетики, достигнуто к настоящему времени для гетероструктуры SnS/TiO<sub>2</sub>, полученной недорогим методом химического осаждения.

### 3.2.6. Фототранзисторы

*Фототранзистор* – оптоэлектронный триод, основанный на использовании биполярной или униполярной транзисторной структуры, работающий на комбинированном эффекте управления как омическим состоянием, так и потоком носителей с помощью света и предназначенный для детектирования оптического сигнала и усиления электрического сигнала.

В биполярном фототранзисторе (рис. 3.21), обычно включенном по схеме с общим эмиттером, освещается часть базы, где происходит генерация фотоносителей.

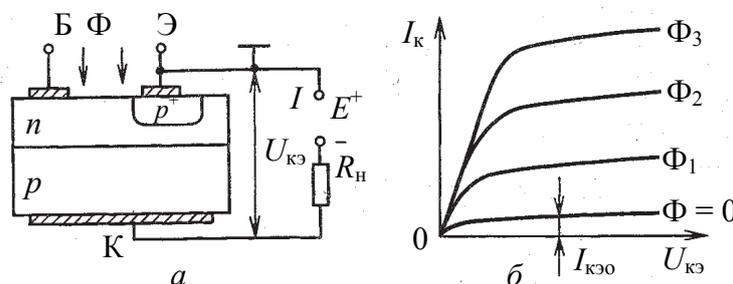


Рис. 3.21. Структура (а) и ВАХ (б) биполярного фототранзистора

Одни носители преодолевают обратносмещенный коллекторный  $p$ - $n$ -переход (на рис. 3.21 – дырки), а другие носители (на рис. 3.21 – электроны) скапливаются на границе прямосмещенного эмиттерного перехода, уменьшая контактное поле и увеличивая инжекцию носителей.

В результате входной сигнал (фототок базы) зависит от освещенности прибора и многократно усиливается, что обуславливает значительно большую чувствительность биполярного фототранзистора по сравнению с фотодиодами.

В качестве примера полевого фототранзистора на рис. 3.22 приведен транзистор с управляющим  $p$ - $n$ -переходом. В нем световой поток генерирует фотоносители в области затвора и  $p$ - $n$ -перехода. Поле последнего разделяет потоки фотоносителей, увеличивая концентрацию носителей и в канале, и в затворе. Фототранзисторы используются как фоточувствительные элементы повышенной чувствительности в оптопарах и других фотоприемных системах, в частности в волоконно-оптических линиях связи.

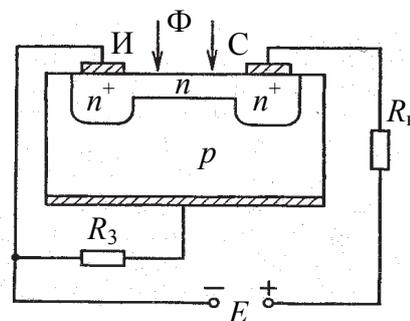


Рис. 3.22. Структура полевого фототранзистора

Выходные параметры фототранзисторов аналогичны, но превосходят параметры фотодиодов.

### 3.3. Оптроны (лекция 17)

*Оптронами* называют оптоэлектронные приборы, которые содержат и источник, и приемник излучения, связанные между собой электрооптически и конструктивно.

Элементарным оптроном является *оптопара*. Этот оптоэлектронный полупроводниковый прибор (рис. 3.23) включает один

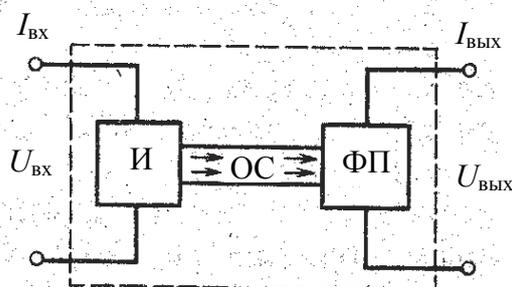


Рис. 3.23. Общая схема оптрона

светоизлучатель и один фотоприемник, оптически связанные и электрически изолированные друг от друга. В качестве излучателя в оптопарах чаще используются некогерентные полупроводниковые излучатели-светодиоды, которые имеют высокий КПД преобразования электрической энер-

гии в световую, длительно и надежно работают с малыми токами и напряжениями. В некоторых случаях применяют лампы накаливания, электролюминесцентные излучатели, а также полупроводниковые гетероструктурные лазеры (в быстродействующих системах). В качестве оптической среды можно использовать воздух, стекло, а также перспективные прозрачные компаунды (эпоксидные смолы, оптические клеи, вазелиноподобные полимеры), которые одновременно осуществляют механическое крепление элементов оптопары. В зависимости от прибора – приемника излучения выделяют резисторные, диодные, транзисторные и другие оптопары. *Оптоэлектронная интегральная схема* содержит ряд схем или компонентов, связанных между собой оптической связью с тем, чтобы избежать электрической развязки. Такая схема состоит из оптронов и усилительных или согласующих устройств, электрически связанных между собой.

Основными элементами оптрона (рис. 3.23), кроме излучателя (И) и фотоприемника (ФП), являются оптический канал (оптическая среда ОС) и устройство управления световым потоком (на рисунке не изображен). Электрический сигнал поступает на светоизлучатель, который преобразует его в световой сигнал. По оптическому каналу происходит передача сигнала на фотоприемник, которая при необходимости сопровождается изменением величины светового потока соответствующим устройством. Фотоприемник преобразует световой сигнал в соответствующий электрический. Таким образом, осуществляется преобразование входного электрического сигнала ( $I_{вх}, U_{вх}$ ) в выходной электрический ( $I_{вых}, U_{вых}$ ) через оптический сигнал.

При отсутствии электрической связи между излучателем и приемником оптрон исполняет в электрической цепи роль переключателя (элемента гальванической развязки). Наличие обратной электрической и оптической связи позволяет использовать оптроны для генерации и усиления электрических и оптических сигналов, для создания запоминающих устройств. Среди достоинств оптоэлектронных интегральных схем, обусловленных реализацией электронейтрального фотонного переноса информации, в сравнении с традиционными электронными схемами можно отметить:

- обеспечение качественной электрической (гальванической) развязки между входом и выходом (т. е. коммутации элементов с различными электрическими параметрами);
- высокую помехозащищенность (невосприимчивость к электромагнитным полям);

- реализацию бесконтактного оптического управления электронными элементами (разнообразие управляющих цепей);
- широкий диапазон рабочих частот (переменный и постоянный сигнал);
- физическую и конструктивно-технологическую совместимость с другими полупроводниковыми микроэлектронными устройствами;
- возможность управления выходным сигналом путем воздействия на оптический канал (оптические датчики).

Нерешенными проблемами оптоэлектронных интегральных схем остаются:

- значительная потребляемая мощность (из-за двойного преобразования сигнала);
- невысокий КПД (несколько процентов);
- низкая термостабильность;
- временная деградация параметров устройства;
- высокий уровень собственных шумов;
- необходимость использования гибридной непланарной технологии.

Общепринятой классификацией оптронов является их разделение в зависимости от типа фотоприемника на резисторные, диодные, транзисторные, тиристорные, диодно-транзисторные и т. д. (табл. 3.6).

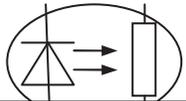
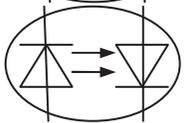
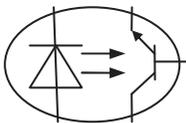
Наиболее важными параметрами оптрона являются:

- время переключения ( $\tau$ ), определяющее быстродействие прибора;
- коэффициент передачи по току ( $K_I = I_2/I_1$ );
- параметры изоляции  $R$  (темновое  $R_T$  и световое  $R_C$  сопротивление прибора, сопротивление изоляции  $R_{и}$ , проходная емкость между входом и выходом  $\approx 1$  пФ).

Среди различных типов оптопар широкое применение на практике в качестве элементов развязки высокоомных и низкоомных цепей находят диодные оптопары по причине их высокого быстродействия (до 1 нс), однако низкий коэффициент передачи тока в таких системах требует дополнительных усилительных устройств для выходного сигнала. Достоинствами резисторных оптопар является линейность и симметричность ВАХ, отсутствие внутренних фото-ЭДС и пониженный уровень собственных шумов. Низкое быстродействие этих устройств ограничено инерционностью фоторезисторов.

Таблица 3.6

## Характеристика некоторых типов оптронов

Название	Обозначение	Элементы		Параметры			Применение
		излучатель	приемник	$\tau$ , мс	$R$	$K_I$	
Резисторные		Светодиод (GaP, GaP <sub>1-x</sub> As <sub>x</sub> ) или лампа накаливания	Фоторезистор (CdS, CdSe)	200	$R_T/R_C = 10^4-10^7$	0,50-0,05	Бесконтактная коммутация и управление в цепях переменного тока
Диодные		Светодиод (GaAs, Ga <sub>x</sub> Al <sub>1-x</sub> As, GaP <sub>1-x</sub> As <sub>x</sub> , GaAs/Ga <sub>x</sub> Al <sub>1-x</sub> As)	<i>p-i-n</i> -фотодиод (Si)	$10^{-3}-10^{-4}$	$R_{и} = 10^{10}$ Ом	0,01	Оптоэлектронные импульсные трансформаторы, элементы согласования периферии с центром ЭВМ, схемы защиты от перегрузки
Транзисторные		GaAs, твердые растворы соединений A <sup>III</sup> B <sup>V</sup>	Фототранзистор <i>n-p-n</i> (Si)	$10^{-1}-10^{-3}$	$R_{и} = 10^9-10^{11}$ Ом	0,5-20,0	Сильноточные оптоэлектронные реле, аналоговые и ключевые коммутаторы сигналов, схемы согласования датчиков с измерителями, элементы гальванической развязки линий связи

Существует большое число других разновидностей оптронов:

- дифференциальные оптопары, в которых один излучатель воздействует на несколько приемников;
- волстоны, своеобразная длинная диодная оптопара, где излучатель и приемник связаны между собой волоконно-оптическим кабелем (волоконным световодом из тонкой светопроводящей нити, выполняющей функцию диэлектрического волновода);
- преобразователи изображений (оптрон с внутренней электрической и внешней оптической связью, включающий, например, многоэлементный приемник на основе PbS-фоторезистора и излучатель на базе GaP-светодиода и осуществляющий преобразование инфракрасного излучения в видимое);
- оптические датчики (например, оптопара с открытым оптическим каналом и оптопара с управляемым оптическим каналом, которые позволяют преобразовывать внешние воздействия в удобные для регистрации, а также обработки оптические и электрические сигналы).

## РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Щука, А. А. Электроника: учеб. пособие для вузов / А. А. Щука. – 2-е изд. – СПб.: БВХ-Петербург, 2008. – 739 с.
2. Борисенко, В. Е. Нанoeлектроника: учеб. пособие для вузов / В. Е. Борисенко, А. И. Воробьева, Е. А. Уткина. – Минск: БИНОМ. Лаборатория знаний, 2009. – 223 с.
3. Бонч-Бруевич, В.Л. Физика полупроводников: учебник для вузов / В. Л. Бонч-Бруевич, С. Г. Калашников. – М.: Наука, 1977. – 638 с.
4. Хандогин, М. С. Электронные приборы: учеб. пособие / М. С. Хандогин, В. Б. Рожанский, С. В. Дробот. – Минск: БГУИР, 2011. – 310 с.
5. Пасынков, В. В. Полупроводниковые приборы: учеб. пособие для вузов / В. В. Пасынков, Л. К. Чиркин. – 3-е изд. – СПб.: Лань, 2001. – 367 с.
6. Колосницын, Б. С. Полупроводниковые приборы и интегральные микросхемы: учеб. пособие / Б. С. Колосницын, П. П. Стешенко, В. В. Шульгов. – Минск: Амалфея, 2002. – 272 с.
7. Щука, А. А. Нанoeлектроника: учеб. пособие для вузов / А. А. Щука. – М.: Физматкнига, 2007. – 463 с.
8. Лачин, В. И. Электроника: учеб. пособие для вузов / В. И. Лачин, И. С. Савелов. – 7-е изд. – М.: Феникс, 2009. – 703 с.
9. Пихтин, А. Н. Оптическая и квантовая электроника: учеб. пособие для вузов / А. Н. Пихтин. – М.: Высшая школа, 2001. – 573 с.
10. Светцов, В. И. Оптическая и квантовая электроника: учеб. пособие для вузов / В. И. Светцов. – Минск: БГТУ, 2000. – 111 с.
11. Ткаченко, Ф. А. Электронные приборы и устройства: учебник / Ф. А. Ткаченко. – Минск: Новое знание; М.: ИНФРА-М, 2020. – 682 с.
12. Зи, С. Физика полупроводниковых приборов: пер. с англ.: в 2 кн. / С. Зи. – 2-е изд. – М.: Мир, 1984. – Кн. 1. – 457 с.; Кн. 2. – 456 с.
13. Розеншер, Э. Оптоэлектроника: пер. с англ. / Э. Розеншер, Б. Винтер. – 2-е изд. – М.: Техносфера, 2006. – 592 с.
14. Джонс, М. Х. Электроника – практический курс: учеб. пособие для вузов / М. Х. Джонс. – М.: Высш. шк., 2006. – 510 с.
15. Бузанева, Е. В. Микроструктуры интегральной техники / Е. В. Бузанева. – М.: Радио и связь, 1990. – 335 с.
16. Богомазова, Н. В. Полупроводниковые приборы: физико-химия и технология: сб. задач / Н. В. Богомазова, О. А. Алисиенок. – Минск: БГТУ, 2016. – 85 с.

Учебное издание

**Богомазова Наталья Валентиновна**

**ФИЗИЧЕСКАЯ ЭЛЕКТРОНИКА  
И ЭЛЕКТРОННЫЕ ПРИБОРЫ**

Электронный конспект лекций

Редактор *О. П. Приходько*  
Компьютерная верстка *О. П. Приходько*  
Корректор *О. П. Приходько*

Издатель:

УО «Белорусский государственный технологический университет».

Свидетельство о государственной регистрации издателя,  
изготовителя, распространителя печатных изданий

№ 1/227 от 20.03.2014.

Ул. Свердлова, 13а, 220006, г. Минск.