

481/33
7-876

Н. М. Тугов
Б. А. Глебов
Н. А. Чарыков

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ

Под редакцией проф. В. А. Лабунцова

Допущено Государственным комитетом СССР по народному образованию в качестве учебника для студентов, обучающихся по специальности «Промышленная электроника»



МОСКВА
ЭНЕРГОАТОМИЗДАТ
1990



ББК 32.852

Т81

УДК 621.382.2/.3 (075.8)

Рецензенты: кафедра полупроводниковых приборов ЛЭТИ
(зав. кафедрой проф. Ю. М. Таиров), зав. кафедрой электроники ЛИТМО проф. В. В. Тогатов

Тугов Н. М. и др.

Т81 Полупроводниковые приборы: Учебник для вузов/Н. М. Тугов, Б. А. Глебов, Н. А. Чарыков; Под ред. В. А. Лабунцова. — М.: Энергоатомиздат, 1990. — 576 с.: ил.

ISBN 5-283-00554-2

Приведены физические свойства, характеристики и режимы работы полупроводниковых приборов, применяемых в устройствах промышленной электроники. Рассмотрены особенности применения приборов. Изложены физические основы построения моделей предельных режимов эксплуатации, методов управления и защиты полупроводниковых приборов.

Для студентов вузов, обучающихся по специальности «Промышленная электроника».

Г 2302030300-029
051(01)-90 225-89

ББК 32.852

Учебник

ТУГОВ НИКОЛАЙ МИХАЙЛОВИЧ
ГЛЕБОВ БОРИС АЛЕКСАНДРОВИЧ
ЧАРЫКОВ НИКОЛАЙ АНДРЕЕВИЧ

Полупроводниковые приборы

Зав. редакцией М. П. Соколова
Редактор В. А. Макаров
Редактор издательства Н. Б. Фоличева
Кругоместные редакторы В. А. Гозак, Г. И. Панфилова
Технический редактор О. Д. Кузнецова
Корректор Э. Б. Драиновская
ИБ № 650

Сдано в набор 12.04.89. Подписано в печать 19.10.89. Т-16648. Формат 84×108¹/₃₂. Бумага типографская № 2. Гарнитура литературная. Печать высокая. Усл. печ. л. 30,24. Усл. кр.-опт. 30,24. Уч.-изд. л. 32,45. Тираж 80 000 экз. Заказ 309. Цена 1 р. 40 к.

Энергоатомиздат, ПЗ114, Москва, М-114, Шлюзовая наб., 10

Владимирская типография Госкомитета СССР по печати
400000, г. Владимир, Октябрьский проспект, д. 7.

ISBN 5-283-00554-2

© Энергоатомиздат, 1990

ПРЕДИСЛОВИЕ

Учебник написан в соответствии с программой курса «Полупроводниковые приборы» для вузов по специальности «Промышленная электроника» и может быть использован в учебном процессе для других специальностей.

Курс «Полупроводниковые приборы» в учебном плане названной специальности является одним из базовых. Его основная задача — дать будущим инженерам в области разработки и применения современных электронных устройств преобразования информации (информационная электроника) и преобразования электроэнергии (энергетическая электроника или преобразовательная техника) необходимый объем сведений, обеспечивающий грамотное использование всей современной гаммы полупроводниковых приборов в схемах самого различного назначения. С этой точки зрения наиболее важным представляется развитие умения и навыков грамотного выбора полупроводникового прибора, исходя из требований к конкретному устройству, и правильного использования моделей приборов при проектировании этих устройств. Такой подход основывается на хорошо известном для инженеров-практиков положении, что надежность, быстродействие, энергетические показатели, габариты, масса и другие технико-экономические характеристики конкретных электронных устройств во многом зависят от выбора типа прибора и режима его эксплуатации.

Во введении определены функциональные возможности трех основных классов полупроводниковых приборов — диодов, транзисторов и тиристоров — и рассмотрены общие вопросы применения полупроводниковых приборов: особенности ключевого и усилительного режимов эксплуатации, влияние обратной связи между выходной и управляющей цепями управляемых приборов на параметры режима, об-

ласть безопасных режимов работы полупроводниковых приборов, особенности работы полупроводниковых приборов на комплексную нагрузку.

В гл. 1 введен параграф с кратким изложением основных положений физики полупроводников и полупроводниковых приборов, в частности изложены современные представления о физических эффектах в полупроводниках при больших концентрациях носителей заряда, сильных полях и высоких уровнях легирования, проведен анализ модели мощного диода с p^+-n-n^+ структурой при высоком уровне инжекции, рассмотрен режим двойной инжекции, влияние электроно-дырочного рассеяния и других нелинейных эффектов, рассмотрены вопросы физики переходов с барьером Шоттки и омических переходов, рассмотрены быстродействующие диоды Шоттки, перспективные в силовой электронике.

В гл. 2 традиционный материал по физике работы бездрейфовых транзисторов изложен в сокращенном виде, подробно рассмотрены вопросы физики современных $n^+-p-n-n^+$ транзисторов, включая эффект Кирка, квазинасыщения и переходные процессы в режиме больших токов, а также фундаментальные ограничения частотного диапазона мощных и высоковольтных транзисторов, рассмотрена область безопасной работы и физические явления, определяющие ее границы.

В гл. 3 проводится анализ статических состояний тиристора, методом заряда проводится расчет переходных процессов в режиме больших токов, определяются особенности построения цепи управления, приводятся схемотехнические методы защиты от эффектов du/dt , di/dt и повышения помехоустойчивости тиристоры. Рассмотрены разновидности тиристоры: симисторы, запираемые тиристоры, тиристоры-диоды.

По гл. 4 особый интерес представляет материал по мощным полевым транзисторам, имеющим значительные перспективы: дан анализ физических основ работы, основные характеристики и параметры, основы эксплуатации в ключевом и усилительном режимах, построение цепей управления ДМДП, ВМДП и СИТ. Следует подчеркнуть, что полевые приборы заметно потеснили биполярные на мировом рынке полупроводниковых приборов.

В гл. 5 рассмотрены основные свойства и параметры оптического излучения, параметры и характеристики полупроводниковых излучателей (излучающих диодов и лазеров),

параметры и характеристики фотоприемников и оптопар. Изложение материала в этих главах построено по единой методической схеме: прежде всего рассматриваются физические процессы в структуре прибора, являющиеся общими как для интегральных, так и для дискретных приборов. Далее проводится параметризация приборов в основных режимах эксплуатации. В заключение приводятся особенности мощных (силовых) полупроводниковых приборов. При изложении используется метод сравнения разных классов полупроводниковых приборов между собой с точки зрения их эксплуатации в том или ином режиме.

Большое место в учебнике занимают вопросы моделирования приборов с точки зрения их эксплуатации в электронном устройстве, классификация и области применения моделей, библиотека моделей приборов, критерии выбора моделей, методы идентификации и определения параметров моделей приборов (гл. 6).

Значительное внимание уделено таким общим для применения полупроводниковых приборов вопросам, как методы расчета тепловых режимов и работа прибора совместно с охладителем и системой охлаждения (гл. 7), а также особенностям работы при последовательно-параллельном соединении приборов (гл. 8).

Учебник в значительной части ориентирован на рассмотрение силовых полупроводниковых приборов, являющихся элементарной базой бурно развивающейся в последние годы области техники — энергетической (силовой) электроники. Физические основы работы слаботочных приборов (диодов, биполярных и полевых транзисторов и т. д.) достаточно полно отражены в имеющейся литературе, в то время как особенности работы силовых полупроводниковых приборов в учебной литературе необходимого отражения не нашли.

Принятые в книге условные обозначения параметров полупроводниковых приборов соответствуют действующим стандартам [18—23]; все параметры обозначаются буквами латинского алфавита, а в качестве индексов используются буквы русского алфавита. Условные обозначения основных величин приводятся в приложении.

Данная книга представляет собой фактически одну из первых попыток создания учебника по физическим основам применения полупроводниковых приборов в схеме (устройстве), т. е. предназначена прежде всего для студентов специальностей, занимающихся проектированием устройств на основе полупроводниковых приборов, но может быть полез-

на и студентам других специальностей, занимающихся проектированием собственно полупроводниковых приборов.

Авторы и редактор сознают, что при написании учебника не все из задуманного удалось реализовать в полной мере, и примут с признательностью все критические замечания.

Глава I написана Н. А. Чарыковым, гл. 2 — Б. А. Глебовым и Н. А. Чарыковым, гл. 3—8 и введение написаны Н. М. Туговым.

Много полезных замечаний по содержанию книги сделали рецензенты доктор техн. наук В. В. Тогатов и коллектив кафедры электроники ЛИТМО, коллектив кафедры диэлектриков и полупроводников ЛЭТИ им. В. И. Ульянова (Ленина), редактор канд. техн. наук В. А. Макаров. Всем им авторы выражают свою благодарность.

Авторы признательны канд. техн. наук И. Г. Недолужко, который представил ряд материалов по § 6.2—6.4, канд. техн. наук С. Г. Бузыкину за материалы § 2.3.3 и канд. техн. наук В. В. Попову за материалы по гл. 1.

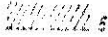
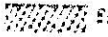







Доктор техн. наук *В. А. Лабунцов*

ОБОЗНАЧЕНИЯ ОСНОВНЫХ ВЕЛИЧИН И ПАРАМЕТРОВ

- D — коэффициент диффузии
- E — напряженность электрического поля, освещенность
- f — частота
- I — среднее и действующее значения тока
- U — среднее и действующее значения напряжения
- J — плотность тока
- L — индуктивность, диффузионная длина носителей заряда, длина канала полевых транзисторов
- M — коэффициент лавинного размножения, светимость
- n — концентрация электронов
- p — концентрация дырок
- N — концентрация примесей (N_A — акцепторная, N_D — донорная)
- n_0 — равновесная концентрация электронов
- p_0 — равновесная концентрация дырок
- p_i — собственная концентрация дырок
- n_i — собственная концентрация электронов
- Q — заряд
- q — элементарный заряд $\approx 1,6 \cdot 10^{-19}$ Кл
- Φ — поток излучения
- R — сопротивление (внешние цепи)
- r — сопротивление (параметры моделей приборов)
- S — площадь, крутизна, чувствительность фотоприемников
- T — температура, период колебаний
- TK — температурный коэффициент
- \mathcal{E} — энергия
- t — время

V — объем
 v — скорость
 α — коэффициент передачи тока эмиттера
 β — коэффициент передачи тока базы
 γ — коэффициент инжекции
 ϵ_a — абсолютная диэлектрическая проницаемость
 ϵ_r — относительная диэлектрическая проницаемость
 ϵ_0 — электрическая постоянная = $8,86 \times 10^{-14}$ Ф/см
 η — коэффициент полезного действия
 λ — длина волны, интенсивность отказов, коэффициент теплопроводности
 μ — подвижность
 ν — частота в оптическом диапазоне
 ρ — удельное сопротивление, плотность заряда
 σ — проводимость
 τ — время жизни носителей заряда, постоянная времени
 φ — потенциал
 ω — угловая частота
 A — работа
 E_K — ЭДС источника питания коллекторной цепи
 E_B — ЭДС источника питания базовой цепи
 E_C — ЭДС источника питания стоковой цепи
 E_Z — ЭДС источника питания затворной цепи
 E_A — ЭДС источника питания анодной цепи
 u_a, U_a — анодное напряжение
 i_a, I_a — анодный ток
 $U_{обр}$ — обратное напряжение
 $I_{обр}$ — обратный ток
 I_y — ток управления
 $U_{пер}$ — напряжение переключения
 $(du/dt)_{кр}$ — критическая скорость нарастания напряжения в закрытом состоянии
 $I_{пер}$ — ток переключения
 $I_{удерж}$ — ток удержания
 $I_{вкл}$ — ток включения
 $(di/dt)_{кр}$ — критическая скорость нарастания тока в открытом состоянии

$(du/dt)_{ном}$ — критическая скорость нарастания коммутационного напряжения
 I_K — коллекторный ток
 I_E — эмиттерный ток
 I_B — базовый ток
 $U_{проб}$ — напряжение пробоя
 $h_{11}, h_{12}, h_{22}, h_{21}$ — h -параметры биполярного транзистора
 C_K — емкость коллекторного перехода
 C_E — емкость эмиттерного перехода
 r_B — сопротивление базы
 $U_{си}$ — напряжение между стоком и истоком
 $U_{зи}$ — напряжение между затвором и истоком
 $U_{ип}$ — напряжение между истоком и подложкой
 i_c, I_c — ток стока
 i_z, I_z — ток затвора
 $U_{пор}$ — пороговое напряжение
 $U_{отс}$ — напряжение отсечки
 $I_{ф}$ — фототок
 I_T — темновой ток
 T_J — температура структуры
 T_{imax} — максимально допустимая температура структуры (пределная температура)
 T_K — температура корпуса
 T_0 — температура охладителя
 $T_{ср}$ — температура окружающей среды
 $R_{ст-к}$ — тепловое сопротивление структура—корпус
 $R_{ст-ср}$ — тепловое сопротивление структура—среда
 $R_{к-о}$ — тепловое сопротивление корпус—охладитель
 $R_{о-ср}$ — тепловое сопротивление охладитель—среда
 r_T — переходное тепловое сопротивление
 $P_{о,с}$ — мощность потерь в открытом состоянии прибора, среднее значение
 $P_{з,с}$ — среднее значение мощности потерь в закрытом состоянии прибора
 $P_{вкл}$ — среднее значение мощности потерь во время переходного процесса включения
 $P_{вкл}$ — среднее значение мощности потерь во

$t_{пер}$	время переходного процесса выключения
$U_{от}, r_{от}$	параметры аппроксимации ВАХ приборов в открытом состоянии; пороговое напряжение и динамическое сопротивление
	E_C свободная зона (зона проводимости)
	E_V валентная зона
	электроны
	дырки
	ионы (заряды ОПЗ)
	Перенос носителя заряда: диффузией
	дрейфом
	квант энергии
	рекомбинация

Введение

ОБЩИЕ ВОПРОСЫ ПРИМЕНЕНИЯ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Основные классы полупроводниковых приборов. Промышленные электронные устройства, как правило, содержат микроэлектронную систему управления, которая определяет логику работы устройства и строится на интегральных микросхемах, и мощную (исполнительную) схему, которая передает в нагрузку преобразованную электрическую мощность. В качестве нагрузки может выступать, например, электродвигатель, электронная вычислительная машина, громкоговоритель, лазер и т. п.

Электронное устройство (схема) состоит из электрически связанных между собой пассивных компонентов (резисторов, конденсаторов и индуктивностей) и активных компонентов — полупроводниковых приборов. В интегральной микросхеме активные и пассивные компоненты составляют единое целое, т. е. полупроводниковый прибор нельзя конструктивно выделить из ее корпуса как самостоятельный элемент; полупроводниковый прибор микросхемы будем в дальнейшем называть *интегральным прибором*. В мощной схеме полупроводниковый прибор — конструктивно самостоятельный элемент, и в этом применении его будем называть *дискретным прибором*.

Интегральный полупроводниковый прибор принципиально маломощный, так как основное требование к интегральной микросхеме — при минимальной массе и габаритах обеспечить максимальные функциональные возможности; одна микросхема содержит до 10^4 полупроводниковых приборов на 1 мм^2 , а потребляет мощность меньше 10^{-2} Вт . Современный дискретный полупроводниковый прибор мощной схемы преобразует мощность до 10 кВт и более.

Внутри электронного устройства полупроводниковый прибор выполняет две основные функции:

замыкает и размыкает цепь электрического тока, т. е. работает как *ключ*;

обеспечивает линейное усиление электрического сигнала, т. е. работает как *усилитель*.

Соответственно будем разделять *ключевой* и *усилительный режимы эксплуатации* полупроводникового прибора. В ключевом режиме прибор имеет два статических (длительно устойчивых) состояния: замкнутое (ключ открыт) — сопротивление прибора близко к нулю, и разомкнутое (ключ закрыт) — сопротивление прибора велико. Переход из одного статического состояния в другое обеспечивается управляющим сигналом, который должен быть больше некоторого граничного значения. Передаваемая через полупроводниковый ключ информация содержится в амплитуде выходного сигнала и может принимать только два значения, поэтому легко представляется в цифровой форме. Ключевой режим — рабочий режим полупроводниковых приборов в цифровых интегральных микросхемах и микроэлектронных устройствах, импульсных преобразователях и стабилизаторах, формирователях импульсов и других схемах.

В усилительном режиме полупроводниковый прибор обеспечивает линейную передачу сигнала: выходной сигнал полностью повторяет форму управляющего (входного) сигнала, но имеет большую (усиленную) мощность. Передаваемая информация при этом заключена не только в амплитуде, но и в форме выходного сигнала. Усилительный режим — рабочий режим полупроводниковых приборов в аналоговых устройствах: в аналоговых интегральных микросхемах (прежде всего в операционных усилителях), в многокаскадных усилителях мощности звуковой частоты, непрерывных стабилизаторах напряжения и тока и др. Здесь сигнал на выходе — аналог входного сигнала, и амплитуды этих сигналов связаны прямо пропорциональной зависимостью.

По функциональным возможностям можно выделить три основных класса полупроводниковых приборов: *диоды*, *транзисторы* и *тиристоры*.

Диод — это электрический «вентиль», т. е. прибор, обеспечивающий однонаправленную передачу электрического сигнала: он открыт и проводит ток при прямом напряжении между его двумя электродами, и закрыт, когда напряжение между электродами имеет обратную полярность. Диод можно считать неуправляемым ключом, который не усиливает мощность передаваемого сигнала.

Транзистор — управляемый полупроводниковый прибор, который может работать в электронной схеме как в ключевом, так и в усилительном режимах. Это универсальный полупроводниковый прибор интегральных и мощных схем.

Тиристор — управляемый полупроводниковый прибор, который используется только в ключевом режиме прежде всего в мощных исполнительных устройствах.

Основные особенности режимов работы полупроводниковых приборов. С точки зрения применения полупроводниковых приборов, таким образом, прежде всего необходимо рассмотреть основные особенности усилительного и ключевого режимов. Это позволит в первом приближении определить требования к приборам для каждого из этих режимов. Такое рассмотрение логично провести на примере универсального, т. е. работающего в обоих режимах, а также наиболее распространенного полупроводникового прибора — транзистора. Пусть транзистор V_T и резистор R_n — нагрузка — соединены последовательно с источником постоянного напряжения E . Тогда через выходную цепь транзистора V_T и нагрузочный резистор R_n будет протекать ток I , который определяется из уравнения (рис. В.1, а)

$$I = (E - U)/R_n. \quad (B.1)$$

График функции (B.1) представляет собой прямую линию (рис. В.1, б) — это ВАХ резистора R_n . Типичные выходные ВАХ транзистора $I=f(U)$ (рис. В.1, б) имеют нелинейный характер; при заданном сигнале управления¹ выходной ток I определяется ординатой точки пересечения ВАХ резистора с выходной ВАХ транзистора (точки A , C_1 , C_2 , C_3 , B и т. д. на рис. В.1, б). Точки пересечения, таким образом, определяют рабочий электрический режим транзистора и нагрузки (ток и напряжение) — это *рабочие точки ВАХ*. Из рис. В.1, б легко получить *передаточную характеристику*, т. е. зависимость выходного напряжения $U_{\text{вых}}$ от управляющего (входного) сигнала (рис. В.1, в).

Рассмотрим *статические* и *динамические* (во время переходного процесса переключения транзистора) *положения рабочей точки* в ключевом режиме. Разомкнутому состоянию ключа соответствует точка B на характеристиках: че-

¹ В качестве управляющего сигнала используется ток в биполярных, напряжение в полевых и оптическое излучение в оптоэлектронных полупроводниковых приборах.

рез транзистор протекает минимальный ток, и все напряжение источника питания E прикладывается к транзистору — это точка «отсечки» выходного тока. Замкнутому (открытому) состоянию транзистора соответствует точка A : через транзистор протекает максимальный при данном значении сопротивления R_H ток, и на транзисторе минимальное па-

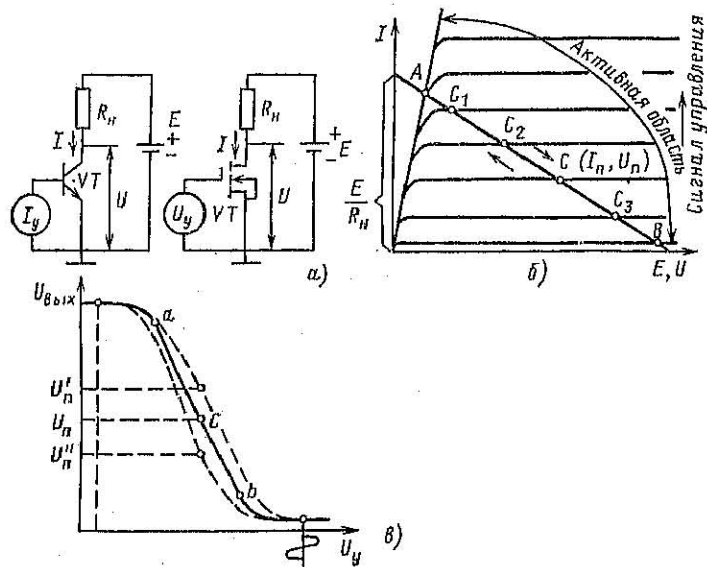


Рис. В.1. Схема включения (а), выходные ВАХ (б) и передаточная характеристика (в) в ключевом и усилительном режимах

дение напряжения — это точка «насыщения» выходного тока. Изменениям управляющего сигнала от значения, открывающего транзистор, до значения, закрывающего транзистор (или наоборот), соответствуют динамические (кратковременные) положения рабочей точки между статическими точками A и B . Направления перемещения рабочей точки при отпирании и запираии транзистора показаны стрелками (рис. В.1, б). Динамические рабочие точки проходят через активную область ВАХ.

В усилительном режиме эксплуатации транзистора рабочие точки в статическом и динамическом состояниях не выходят за пределы активной области, что обеспечивает

линейную передачу управляющего (входного) сигнала. Соответствующие рабочие точки на передаточной ВАХ находятся на участке с наибольшей крутизной характеристики, что обеспечивает наибольшее усиление сигнала. Статическое положение рабочей точки (точки покоя C) выбирается из условия наибольшей крутизны в этой точке.

Проведем краткую сравнительную оценку экономичности, быстродействия и стабильности выходных параметров в усилительном и ключевом режимах.

Статические состояния ключа характеризуются малыми потерями мощности, так как в открытом состоянии мало падение напряжения на приборе, а в закрытом ничтожно значение тока. Точка покоя в усилительном режиме характеризуется непрерывным потреблением мощности покоя $I_n U_n$ (рис. В.1, б).

Быстродействие в ключевом режиме определяется длительностью перехода из одного статического состояния в другое, при этом возможно ускорение этого перехода с помощью цепи управления без искажения передаваемой информации, которая содержится только в амплитуде выходного сигнала. В усилительном режиме передаваемая информация заключена в форме сигнала, и в понятие быстродействия здесь входит сохранение частотно-фазовых параметров сигнала, обеспечение минимальных частотных искажений и т. п.; в итоге инерционность транзистора в усилительном режиме как бы увеличивается, а ее оценка требует более точного учета паразитных реактивных элементов.

Передаточная характеристика (рис. В.1, в) хорошо иллюстрирует зависимость выходных параметров от режима. В ключевом режиме форма передаточной характеристики между точками A и B не влияет на амплитуду выходного сигнала, и транзисторный ключ малочувствителен к разбросу параметров транзистора, их температурному и временному дрейфу. Небольшие колебания входного напряжения U_y (около точки B , например) в результате внешних электромагнитных наводок или шумов практически не влияют на амплитуду выходного сигнала. В усилительном режиме используется линейный участок передаточной характеристики между точками a и b ; входные и выходные сигналы в пределах этого участка связаны друг с другом линейной зависимостью, и любая нестабильность характеристики на этом участке будет непосредственно сказывать-

ся на выходном сигнале (на рис. В.1, в показаны изменения напряжения покоя $U_{п}$).

Рассмотренный усилительный режим работы транзистора относится к так называемому *классу усиления А — однотактному усилительному режиму*. Такой режим неэкономичен, так как мощность теряется и в режиме покоя, КПД в классе А не более 50%. Поэтому в мощных усилителях в целях повышения КПД используют *комплементарную (дополняющую) пару транзисторов*. Каждый транзистор в такой паре работает в экономичном классе усиления (В или АВ — рис. В.2) и усиливает только одну половину

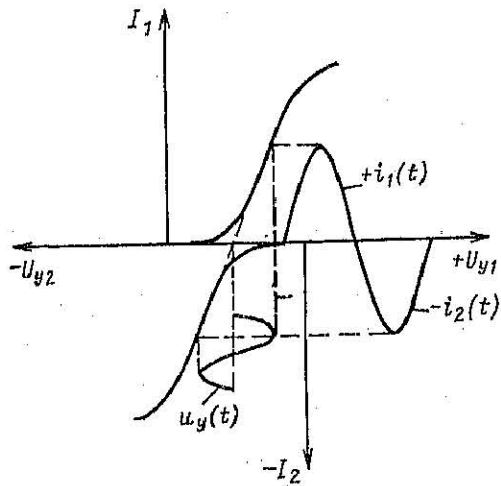


Рис. В.2. Передаточная характеристика комплементарной пары транзисторов в двухтактном усилительном режиме

периода передачи информационного сигнала (токи i_1 или i_2) — это *двухтактный усилительный режим*.

Из ключевого режима эксплуатации следует выделить так называемый *аналого-ключевой режим*, в котором открытый прибор служит для передачи аналогового сигнала в нагрузку (без усиления), а закрытый прибор такую передачу запрещает. Длительность открытого и закрытого состояний полупроводникового прибора в режиме такого «аналогового ключа» определяется сигналом управления (рис. В.3). Из схемы имеем при открытом состоянии ключа

$$u_{п} = u_{вх} / [1 + (r_{отк} / R_{н})], \quad (В.2)$$

т. е. погрешность передачи аналогового сигнала через полупроводниковый ключ тем меньше, чем меньше сопротивление прибора при открытом состоянии $r_{отк}$ и чем слабее зависимость $r_{отк}$ от передаваемого аналогового сигнала.

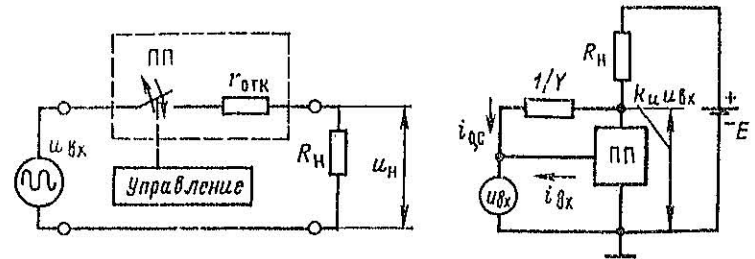


Рис. В.3. Схема включения полупроводникового прибора (ПП) в аналого-ключевом режиме

Рис. В.4. Обратная связь между выходом и входом полупроводникового прибора

Обратная связь выход—вход в полупроводниковых приборах. При работе управляемого полупроводникового прибора (транзистора, тиристора) в активной области (эта область является переходной для ключевого режима и основной для усилительного, см. рис. В.1, б) значительное влияние на энергетические параметры, быстродействие и надежность работы прибора может оказать *обратная связь между выходной и управляющей (входной) цепями прибора*. Пусть между управляющим и выходным электродами полупроводникового прибора ПП имеется комплексная проводимость Y (рис. В.4); при подаче на вход напряжения $u_{вх}$ на выходе в активной рабочей области имеем в $k_{н}$ раз усиленное напряжение. Тогда ток $i_{о,с}$, протекающий через Y , равен $(u_{вх} - k_{н} u_{вх}) Y$, а входная проводимость с учетом обратной связи $Y_{вх,с}$ (рис. В.4) равна:

$$Y_{вх,с} = \frac{i_{о,с} + i_{вх}}{u_{вх}} = (1 - k_{н}) Y + Y_{вх} = Y_{о,с} + Y_{вх}, \quad (В.3)$$

где $Y_{о,с}$ — эквивалентная проводимость, вносимая во входную цепь эффектом обратной связи; $Y_{вх}$ — входная проводимость без учета обратной связи. При работе транзистора в режиме усиления напряжения (с общим эмиттером для

↑ Это явление часто называют эффектом Миллера.

биполярного транзистора и с общим истоком для полевого) коэффициент усиления по напряжению отрицательный, причем $|k_u| \gg 1$ и эквивалентная проводимость $Y_{o,c}$ значительно превышает входную проводимость $Y_{вх}$. Если транзистор включен в схему повторителя (с общим коллектором для биполярного транзистора и с общим стоком для полевого), то $k_u > 0$ и меньше единицы, т. е. действие обратной связи выход—вход снижается.

Область безопасных режимов полупроводниковых приборов. Надежность работы полупроводниковых приборов

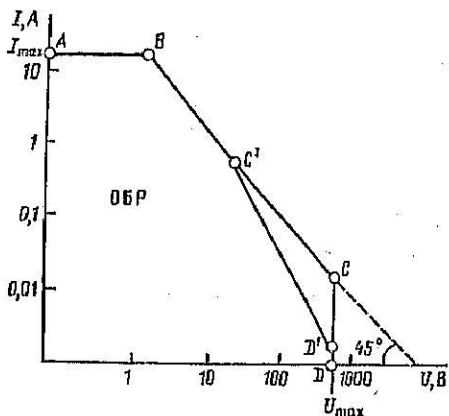


Рис. В.5. Диаграмма области безопасных режимов полупроводниковых приборов

характеризуют *областью безопасных режимов (ОБР)*. Эта область ограничена осями первого квадранта ВАХ полупроводниковых приборов; по оси ординат откладываются значения выходного тока прибора, по оси абсцисс — выходного напряжения в логарифмическом масштабе. Типичная ОБР полупроводникового прибора для статического режима эксплуатации (для медленно изменяющихся токов и напряжений) изображена на рис. В.5. Она ограничена прямолинейными отрезками АВ, ВС, CD. Горизонтальный отрезок АВ определяет значение *максимально допустимого выходного тока прибора I_{max}* . Вертикальный отрезок CD, ограничивающий ОБР, определяет значение *максимально допустимого выходного напряжения прибора U_{max}* .

Отрезок ВС характеризует ограничение электрического режима прибора по мощности: значения мощности, выделяющейся в приборе, не должны выходить за границу ВС ОБР. Это ограничение связано с тем, что температура

структуры прибора T_j не должна превышать *максимально допустимого значения T_{jmax}* (подробнее см. в гл. 7), т. е. выделяющаяся в приборе электрическая мощность P не должна превышать возможностей теплоотвода через тепловое сопротивление структура—окружающая среда $R_{ст-ср}$:

$$P = UI \leq (T_{jmax} - T_{ср}) / R_{ст-ср}, \quad (B.4)$$

где $T_{ср}$ — температура окружающей среды.

Переходя в выражении (B.4) к равенству и логарифмируя обе его части, получаем

$$\lg I = \lg \left(\frac{T_{jmax} - T_{ср}}{R_{ст-ср}} \right) - \lg U. \quad (B.5)$$

Таким образом, в логарифмическом масштабе граница ОБР по максимальной мощности (или по максимальной температуре T_{jmax}) — отрезок ВС — представляет собой прямую линию с наклоном под углом 45° к осям ординат.

Для большинства полупроводниковых приборов ОБР ограничивается дополнительно отрезком $C'D'$ и уменьшается. Это ограничение связано с воздействием электротепловой связи (в литературе используется также термин «тепловая неустойчивость»): мощность, выделяющаяся в приборе, и тепловое сопротивление увеличиваются с ростом температуры, т. е. нагрев прибора растет, а теплоотвод ухудшается; в результате температура структуры T_j либо ограничивается на определенном уровне (большем, чем в отсутствие обратной связи) или растет неограниченно. Выход рабочей точки за ОБР приводит, как правило, к необратимому ухудшению параметров прибора.

Особенности работы полупроводниковых приборов на комплексную нагрузку. Для простоты анализ режимов эксплуатации полупроводниковых приборов обычно проводится при резистивной нагрузке. В реальных промышленных электронных устройствах характер нагрузки чаще всего комплексный, что сказывается на работе полупроводниковых приборов. Рассмотрим качественно особенности эксплуатации полупроводниковых приборов на нагрузке активно-индуктивную (RL-нагрузку) и на активно-емкостную (RC-нагрузку) в ключевом режиме по сравнению с R-нагрузкой (рис. В.6). При R-нагрузке изменения тока $i(t)$ и напряжения $u(t)$ во время переходных процессов определяются быстроедействием прибора ПП и имеют одинаковую форму: $u(t) = E - i(t)R$ (рис. В.6, б). Переходный процесс

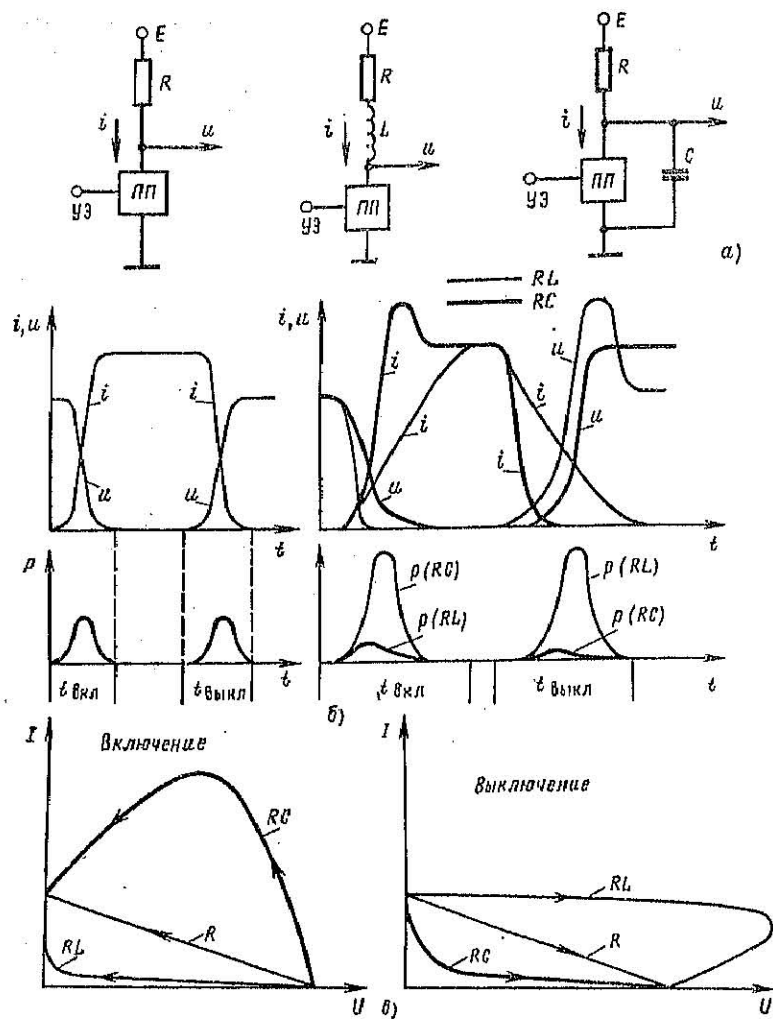


Рис. В.6. Схемы включения (а), переходные характеристики тока i , напряжения u и мощности потерь P (б) и траектории динамической рабочей точки (в) при работе полупроводникового прибора на комплексную нагрузку

включения на RL -нагрузку характеризуется снижением скорости нарастания тока, которая теперь задается постоянной времени индуктивности $\tau_L = L/R$; скорость спада напряжения при этом изменяется мало и определяется быстродействием прибора (рис. В.6, б). При включении на RC -нагрузку нарастания на приборе замедляется, а скорость нарастания тока задается прибором, причем амплитуда тока возрастает за счет энергии, запасенной в конденсаторе (рис. В.6, б). Во время переходного процесса выключения полупроводникового прибора RL - и RC -нагрузки в своем воздействии на переходные характеристики $i(t)$, $u(t)$ меняются местами: при RC -нагрузке прибор быстро выключается «по току» (длительность этого процесса определяется собственным быстродействием прибора), а нарастание напряжения затягивается во времени; при RL -нагрузке спад тока замедляется (определяется постоянной времени τ_L), а скорость изменения напряжения задается прибором, причем амплитуда напряжения возрастает за счет энергии, запасенной в индуктивности нагрузки (рис. В.6, б). Траектории динамической рабочей точки во время переходных процессов включения и выключения полупроводникового прибора показаны на рис. В.6, в. Таким образом, переходный процесс включения полупроводникового прибора на RC -нагрузку характеризуется повышенными потерями мощности (рис. В.6, б) и перегрузкой по току, включение на RL -нагрузку энергетически наиболее благоприятно для прибора; наихудший с этих позиций режим выключения полупроводникового прибора имеет место, наоборот, при RL -нагрузке; RC -нагрузка обеспечивает при выключении наименьшую мощность потерь в приборе. Соответственно должна меняться оценка надежности работы полупроводникового прибора с учетом влияния комплексной нагрузки. Если допустимые сочетания тока и напряжения полупроводникового прибора ограничить ОБР, то при работе на комплексную нагрузку возможны кратковременные «выходы из ОБР» (рис. В.7), что может привести к нарушению нормальной работы прибора.

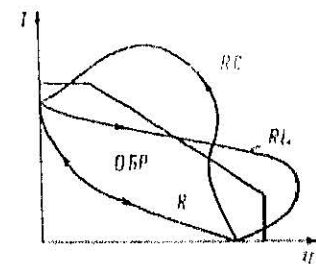


Рис. В.7. Траектории рабочей точки и область безопасной работы для разных видов нагрузки полупроводникового прибора

1. Определите основные функции, выполняемые полупроводниковыми приборами — диодом, транзистором, тиристором — в электронной схеме.
2. Проведите сравнительную оценку экономичности, быстродействия и стабильности выходных параметров полупроводникового прибора (транзистора) в усилительном и ключевом режимах эксплуатации.
3. Приведите диаграммы изменения тока и напряжения полупроводникового прибора во время переходного процесса включения на активно-индуктивную и активно-емкостную нагрузки; оцените энергетические потери и надежность работы полупроводникового прибора при этом.
4. Выполните то же самое (см. п. 3) для переходного процесса выключения.
5. Определите воздействие эффекта Миллера на быстродействие управляемого полупроводникового прибора.
6. Определите коротко основные физические явления, ограничивающие область безопасных режимов (ОБР) полупроводниковых приборов.

Глава первая

ОСНОВЫ ФИЗИКИ ПОЛУПРОВОДНИКОВ И ЭЛЕКТРОННО-ДЫРОЧНЫХ ПЕРЕХОДОВ. ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

1.1. ОСНОВЫ ФИЗИКИ ПОЛУПРОВОДНИКОВ

Основным элементом большинства полупроводниковых приборов является электрический переход — структура, содержащая переходный слой в полупроводниковом материале между двумя областями с различными типами электропроводности или разными значениями удельной электрической проводимости, причем одна из областей может быть металлом.

Электрический переход между двумя областями полупроводника, одна из которых имеет электропроводность n -типа (электронную), а другая — p -типа (дырочную), называют *электронно-дырочным переходом*, или, кратко, *$p-n$ переходом*. Электрические переходы между двумя областями одного и того же типа электропроводности (n - или p -типа), с различающимися значениями удельной электрической проводимости, называют *изотипными* переходами. Изотипные переходы могут быть электронно-электронными ($n-n^+$) или дырочно-дырочными ($p-p^+$). Знаком $+$ условно отмечается область с более высокой удельной электрической проводимостью.

В зависимости от используемых для образования электрического перехода полупроводниковых материалов различают *гомогенный переход* (*гомопереход*), образованный в одном полупроводниковом материале — германии (Ge), кремнии (Si), арсениде галлия (GaAs) и др., и *гетерогенный переход* (*гетеропереход*), образованный смежными областями полупроводниковых материалов с различной шириной запрещенной зоны: германий—кремний, германий—

арсенид галлия и многие другие. В настоящее время доминирующее, с практической точки зрения, положение занимают гомопереходы на кремнии.

Электрические переходы между металлом и полупроводником являются неотъемлемым элементом любого, без исключения, полупроводникового прибора. Прежде всего переходы металл—полупроводник (М—П) используют для создания токоподводящих электродов к прибору—*омических переходов* М—П, электрическое сопротивление которых мало и в заданном рабочем диапазоне токов практически не зависит от направления тока (дополнительные требования к свойствам омических переходов будут рассмотрены далее подробнее). Создание таких омических переходов представляет сложную научно-техническую задачу. Переходы М—П, обладающие выпрямляющими свойствами, т. е. имеющие резко несимметричную ВАХ, называют *переходами Шоттки*, а диоды на их основе—*диодами Шоттки*.

Полупроводниковые диоды на основе *p-n* перехода должны содержать как минимум три электрических перехода: собственно *p-n* переход и два омических перехода М—П, используемых в качестве электрических выводов (электродов). Диоды Шоттки содержат как минимум два электрических перехода: собственно переход Шоттки и омический переход М—П. Перечисленные электрические переходы образуют так называемую *структуру* полупроводникового прибора—совокупность граничащих друг с другом областей или слоев полупроводника и омических переходов (контактов) М—П. Современные диоды могут иметь достаточно сложную структуру, например М— $p^+—n—n^+—M$. Обычно наличие крайних, выводных (электродных) слоев металла в условном обозначении структуры не указывается, т. е. обычно записывают так: диод $p^+—n—n^+$ типа.

При создании *p-n* переходов используют обширный арсенал современных технологических методов, связанных с введением различных атомов легирующей примеси в исходный полупроводниковый материал путем диффузии, имплантации, сплавления, эпитаксии для получения заданного профиля распределения легирующей примеси в полупроводнике, обеспечивающего требуемые свойства *p-n* перехода.

Свойства электрических переходов и диодов на их основе описываются совокупностью различных характеристик и параметров. *Вольт-амперные характеристики* (ВАХ) описывают функциональную зависимость между током дио-

да и приложенным к диоду напряжением. *Переходные характеристики* описывают зависимость тока (или напряжения) диода от времени в нестационарном режиме. *Температурные характеристики* определяют зависимость тока (при заданном напряжении), напряжения (при заданном токе) либо какого-нибудь параметра диода от температуры. Имеется еще ряд специфических характеристик, связанных с изменением параметров приборов во времени (старением), воздействием разнообразных внешних факторов и т. д.

1.1.1. ОСНОВНЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

По значению проводимости σ полупроводниковые материалы занимают промежуточное положение между металлами ($\sigma = 10^6 \div 10^9 \text{ Ом}^{-1} \cdot \text{см}^{-1}$) и диэлектриками ($\sigma < 10^{-10} \text{ Ом}^{-1} \cdot \text{см}^{-1}$). Проводимость полупроводников сильно зависит от присутствия чужеродных атомов, неидеальности кристаллической структуры и подвержена сильному влиянию различных факторов, таких, как температура, электрическое поле, освещение, магнитное поле, давление и т. п. Эти свойства широко используются для создания ПП с различными функциональными свойствами.

Описание природы носителей заряда в полупроводниках и законов их движения производят с помощью зонной теории твердого тела. Электроны в полупроводнике могут иметь некоторые фиксированные значения энергии или, как говорят, занимают определенные энергетические уровни, образующие области близкорасположенных дискретных значений полной энергии электронов—*разрешенные энергетические зоны*. Разрешенные энергетические зоны разделены интервалами энергий, которыми электроны не могут обладать и которые называются *запрещенными энергетическими зонами*. При абсолютном нуле температуры электроны заполняют самые нижние разрешенные уровни, где их энергия минимальна. Согласно принципу Паули в каждом энергетическом состоянии может находиться только один электрон. Поэтому электроны заполняют несколько нижних зон, а вышележащие зоны остаются пустыми. Верхняя из полностью заполненных электронами разрешенных зон полупроводника называется *валентной*, а следующая за ней незаполненная—*зоной проводимости*. Структура энергетических зон дана на рис. 1.1, где обозначено \mathcal{E}_c —энергия дна зоны проводимости, \mathcal{E}_v —энергия потолка валентной зоны, $\mathcal{E}_g = \mathcal{E}_c - \mathcal{E}_v$ —ширина запрещенной зоны.

При $T > 0$ К тепловое движение приводит к разрыву части валентных связей и электроны из валентной зоны переходят в зону проводимости, где они становятся свободными носителями заряда. В валентной зоне при этом образуются дырки проводимости, которые могут также свободно перемещаться по кристаллу. Этот процесс называют

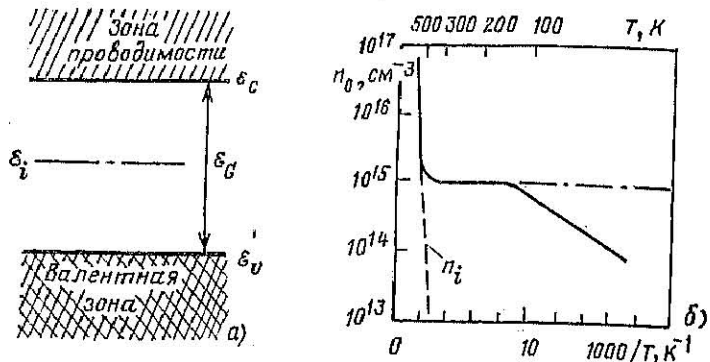


Рис. 1.1. Схематическое изображение энергетических зон в полупроводнике (а) и зависимость концентрации электронов от температуры для кремния с $N_D = 10^{15} \text{ см}^{-3}$ (б)

тепловой генерацией носителей заряда. Существует и противоположный конкурирующий процесс, когда свободный электрон проводимости возвращается в незаполненную валентную связь. Этот процесс называется рекомбинацией электрона с дыркой. При заданной температуре T осуществляется термодинамическое равновесие между тепловой генерацией и рекомбинацией носителей заряда, в результате чего в зоне проводимости устанавливается некоторая, вполне определенная концентрация свободных электронов n_0 , а в валентной зоне — дырок проводимости p_0 . Свободные носители заряда, возникающие в результате теплового возбуждения и находящиеся с решеткой полупроводника в термодинамическом равновесии, называются равновесными или тепловыми. Их концентрация отмечается индексом 0. Теория дает следующие выражения для равновесной концентрации электронов и дырок в невырожденных полупроводниках:

$$n_0 = N_c \exp \left[-(\mathcal{E}_c - \mathcal{E}_F)/(kT) \right]; \quad (1.1)$$

$$p_0 = N_v \exp (\mathcal{E}_v - \mathcal{E}_F)/(kT) \quad (1.2)$$

где N_c — эффективная плотность состояний в зоне проводимости; N_v — эффективная плотность состояний в валентной зоне; $k = 1,38 \cdot 10^{-23} \text{ Дж/К} = 8,62 \cdot 10^{-5} \text{ эВ/К}$ — постоянная Больцмана.

Значения N_c и N_v при 300 К в кремнии примерно равны $2 \cdot 10^{19} \text{ см}^{-3}$.

Входящая в (1.1) и (1.2) величина \mathcal{E}_F есть энергия Ферми или электрохимический потенциал, т. е. работа, которую необходимо затратить для изменения числа частиц в системе на единицу при условии постоянства объема и температуры. Для определения значения \mathcal{E}_F необходима дополнительная информация о свойствах полупроводникового материала.

В чистом (бездефектном и без присутствия посторонних примесей) полупроводнике — его называют собственным — при возбуждении свободных электронов с концентрацией n_i в валентной зоне возникает такое же количество дырок проводимости p_i , т. е. полупроводник остается в целом электронейтральным. Привлекая условие электронейтральности, т. е. приравнявая (1.1) и (1.2), определяем

$$\mathcal{E}_F = \mathcal{E}_i = \frac{\mathcal{E}_c + \mathcal{E}_v}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c}. \quad (1.3)$$

Учитывая, что отношение N_v/N_c (пропорциональное отношению эффективных масс дырки и электрона) не сильно отличается от единицы, делаем вывод, что энергетический уровень Ферми в собственном полупроводнике расположен практически в середине запрещенной зоны. Это позволяет упрощенно трактовать энергию Ферми как среднюю энергию, приходящуюся на один носитель заряда. Теперь, подставив (1.3) в (1.1), получим

$$n_i = p_i = \sqrt{N_c N_v} \exp (-\mathcal{E}_g/(2kT)). \quad (1.4)$$

Это важнейшее выражение физики полупроводников показывает, что собственная концентрация электронов резко (экспоненциально) возрастает с увеличением температуры, причем n_i при заданной абсолютной температуре T тем меньше, чем больше ширина запрещенной зоны. При расчетах n_i следует иметь в виду, что $N_c N_v \sim T^3$, а \mathcal{E}_g зависит от температуры: $\mathcal{E}_g = \mathcal{E}_{g0} - k_1 T$, где \mathcal{E}_{g0} — значение, экстраполированное к $T = 0$ К.

Практические расчеты n_i удобно проводить по формуле

$$n_i(T) = VT^{3/2} \exp(-\mathcal{E}_{\epsilon 0}/(2kT)). \quad (1.5)$$

Для кремния $V=3,873 \cdot 10^{16} \text{ К}^{-3/2} \cdot \text{см}^{-3}$, $\mathcal{E}_{\epsilon 0}=1,21 \text{ эВ}$; для германия $V=1,760 \cdot 10^{16} \text{ К}^{-3/2} \cdot \text{см}^{-3}$, $\mathcal{E}_{\epsilon 0}=0,785 \text{ эВ}$. При «комнатной» температуре (условно $T=300 \text{ К}$) для кремния $n_i=1,4 \cdot 10^{10} \text{ см}^{-3}$, а для германия $n_i=2,3 \cdot 10^{13} \text{ см}^{-3}$.

Для целенаправленного изменения свойств полупроводника легируют, т. е. вводят в небольшом количестве атомы примесей, позволяющие управлять типом электропроводности (электронный, дырочный) и удельной электрической проводимостью. Элементарные полупроводники — кремний, германий, а также алмаз принадлежат IV группе Периодической системы элементов Д. И. Менделеева. Каждый атом такого элементарного полупроводника имеет на внешней орбите по четыре валентных электрона, которые в тетраэдрической кристаллической решетке (четыре ближайших соседа) образуют заполненные ковалентные связи. При введении в такой полупроводник примесных атомов V группы — фосфора (P), мышьяка (As) или сурьмы (Sb) — атом примеси замещает атом полупроводника. Например, для фосфора в кремнии четыре из пяти валентных электронов фосфора образуют с соседними четырьмя атомами кремния ковалентные связи, а энергия связи пятого электрона, равная энергии ионизации, оказывается малой и составляет соотнеси доли электрон-вольта. Поэтому уже при температурах более 100 К такие электроны покидают примесные атомы (говорят, что примесь полностью ионизована) и становятся электронами проводимости, полупроводник имеет преобладающую электронную проводимость n -типа. Легирующую примесь такого сорта называют *донорной*, а атомы — *донорами*. В диапазоне комнатных температур каждый донор дает свободный электрон в зону проводимости, а сам становится положительно заряженным ионом. Это сдвигает термодинамическое равновесие: дырок становится меньше, чем в собственном полупроводнике, а электронов больше. В полупроводнике n -типа электропроводности электроны называют *основными носителями*, а дырки — *неосновными носителями заряда*.

Атомы III группы периодической системы элементов — бор, алюминий, галлий, индий — в кремнии и германии являются *акцепторами*. Захватывая один из валентных электронов кремния в дополнение к своим трем валентным электронам, они образуют четыре ковалентные связи с ближай-

шими соседями — атомами кремния — и становятся отрицательно заряженными ионами. В покинутой связи остается положительно заряженная дырка проводимости, свободно перемещающаяся по связям. Это также меняет соотношение между концентрациями дырок и электронов по сравнению с собственным полупроводником. Дырок становится больше, они чаще рекомбинируют с электронами, уменьшая концентрацию последних. Однако и в акцепторном (p -типа), и в донорном (n -типа) полупроводнике, как следует из рассмотрения (1.1) и (1.2),

$$n_0 p_0 = N_0 N_i \exp(-\mathcal{E}_G/(kT)) = n_i^2. \quad (1.6)$$

В дырочном полупроводнике дырки будут основными, а электроны — неосновными носителями.

Для нахождения энергии Ферми в примесном полупроводнике, содержащем или доноры, или акцепторы, или те и другие одновременно, необходимо использовать условие электронейтральности, записанное с учетом того, что при $T > 100 \text{ К}$ все доноры с концентрацией N_D и акцепторы с концентрацией N_A полностью ионизованы:

$$p_0 - n_0 + N_D - N_A = 0. \quad (1.7)$$

Это условие устанавливает, что сумма всех зарядов отрицательных (электронов, ионов акцепторов) и положительных (дырок, ионов доноров) равна нулю. Обычно вместо $N_D - N_A$ вводится результирующая (разностная) концентрация $N = N_D - N_A$, знак которой определяется (условно) знаком заряда ионов, преобладающих в полупроводнике. Уравнения (1.6) и (1.7) дают систему уравнений для определения концентрации дырок и электронов. Ее решение:

$$n_0 = \frac{N}{2} + \sqrt{\left(\frac{N}{2}\right)^2 + n_i^2}. \quad (1.8)$$

Из этого выражения следует, что если $N > 0$, то $n_0 > n_i$, а в соответствии с (1.6) $p_0 < n_i$, т. е. имеем примесный полупроводник n -типа. В противоположном случае $n_0 < n_i$, а $p_0 > n_i$. Если $N = 0$, то $p_0 = n_0 = n_i$, и в этом случае говорят о беспримесном или полностью компенсированном полупроводнике.

Комбинируя (1.1), (1.2) и (1.3), получаем

$$n_0 = n_i \exp\left(\frac{\mathcal{E}_F - \mathcal{E}_i}{kT}\right); \quad p_0 = n_i \exp\left(\frac{\mathcal{E}_i - \mathcal{E}_F}{kT}\right). \quad (1.9)$$

Теперь можно определить энергию Ферми в примесном

полупроводнике для состояния термодинамического равновесия:

$$\mathcal{E}_F - \mathcal{E}_i = kT \ln(n_0/n_i). \quad (1.10)$$

Для нахождения \mathcal{E}_F в это выражение нужно подставить значение n_0 , рассчитанное по (1.8), откуда следует, что для донорного полупроводника, когда $(N_D - N_A) > n_i$, уровень Ферми расположен в верхней половине запрещенной зоны и тем выше, чем больше концентрация доноров. С ростом температуры возрастает n_i и уровень Ферми стремится к уровню Ферми в собственном полупроводнике E_i , т.е. к центру запрещенной зоны. Температурная зависимость концентрации электронов в кремнии n -типа с $N_D = 10^{15} \text{ см}^{-3}$, рассчитанная по (1.8) с учетом (1.6), изображена на рис. 1.1, б. Эти формулы не описывают зависимость $n_0(T)$ только в области криотемператур ($T < 120 \text{ К}$), где сказывается неполная ионизация примеси.

В полупроводниковых приборах используются полупроводники с ярко выраженной примесной проводимостью, когда в рабочем диапазоне температур ($200 \text{ К} < T < 450 \text{ К}$) $n_i(T_i)$ несколько меньше $|N_D - N_A|$, где T_i — температура наступления собственной проводимости (верхняя граница диапазона); при больших температурах $n_0 \approx n_i$ и примесный полупроводник теряет индивидуальные свойства, обусловленные легированием. Полагая $n_i(T_{max}) = 0,1 |N_D - N_A|$, с помощью (1.6) получаем

$$T_{max} = \frac{\mathcal{E}_G}{k \ln(100 N_c N_v / (N_D - N_A)^2)}. \quad (1.11)$$

Например, для n -кремния ($N_c \approx N_v \approx 2 \cdot 10^{19} \text{ см}^{-3}$) только при $N \geq 10^{14} \text{ см}^{-3}$ значения $T_{max} \geq 450 \text{ К}$. Таким образом, эта формула приближенно определяет нижний уровень концентрации легирующей примеси в приборе.

1.1.2. ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКОВ

Движение электронов в кристалле складывается из беспорядочного теплового движения и упорядоченного движения под действием внешнего электрического поля. В результате происходит медленное перемещение всей совокупности свободных электронов с некоторой средней скоростью. Направленное движение совокупности свободных электронов во внешнем электрическом поле носит название *дрейфа*, а скорость их направленного движения называется дрейфовой

скоростью. Плотность дрейфового тока (заряд электронов с концентрацией n , проходящий за 1 с через поперечное сечение площадью 1 см^2) определяется выражением

$$J = q n v. \quad (1.12)$$

Экспериментально установлено, что дрейфовая скорость зависит от напряженности электрического поля E . Например, на рис. 1.2 изображена зависимость $v(E)$ для электронов и дырок в кремнии при 300 К. В области слабых полей ($E < 3 \cdot 10^3 \text{ В/см}$) $v \sim E$, а далее скорость стремится к насыщению. Дрейфовая скорость насыщения $v_{ns} \approx v_{ps} \approx 10^7 \text{ см/с}$. Подвижность электронов μ_n — это коэффициент пропорциональности в зависимости $v_n = \mu_n E$, измеряется в $\text{см}^2/(\text{В} \cdot \text{с})$. В области слабых полей подвижность от поля не зависит. Ее значение определяется совместным влиянием механизмов рассеяния носителей заряда на колебаниях решетки и на ионах примеси. В практике расчетов полупроводниковых приборов обычно используются многочисленные эмпирические формулы для подвижности. Наиболее простая из них имеет вид

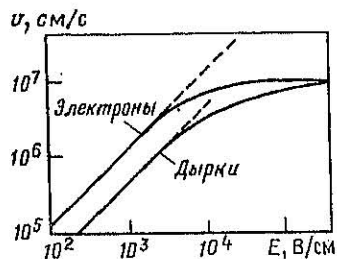


Рис. 1.2. Зависимость дрейфовой скорости носителей заряда в кремнии от напряженности электрического поля при 300 К

$$\mu(N_\Sigma, T) = \left[\mu_1 + \frac{\mu_2}{1 + (N_\Sigma/N_{10})^a} \right] \left(\frac{T}{T_0} \right)^b, \quad (1.13)$$

где $N_\Sigma = N_D + N_A$ — суммарная концентрация ионов доноров и акцепторов; T — абсолютная температура; $T_0 = 300 \text{ К}$; остальные величины являются эмпирическими постоянными.

Таблица 1.1. Значения эмпирических постоянных в зависимости $\mu(N_\Sigma)$ для кремния при 300 К

Носители заряда	μ_1	μ_2	$N_{10} \cdot 10^{14} \text{ см}^{-3}$	a
	$\text{см}^2/(\text{В} \cdot \text{с})$			
Электроны	65	1265	$8,5 \cdot 10^{14}$	0,72
Дырки	47,7	447	$6,3 \cdot 10^{16}$	0,76

ми, соответствующими $T=300$ К, и приведены в табл. 1.1 и 1.2.

Таблица 1.2. Значения эмпирических постоянных в зависимости $\mu(T)$ для кремния

Носители заряда	$N_{\Sigma}, \text{см}^{-3}$	b
Электроны	$N \leq 1 \cdot 10^{15}$ $10^{15} < N < 1 \cdot 10^{20}$ $N \geq 1 \cdot 10^{20}$	-2,6 $0,356 \ln N - 14,9$ +1,5
Дырки	$N \leq 1 \cdot 10^{16}$ $1 \cdot 10^{16} < N < 7 \cdot 10^{21}$ $N \geq 7 \cdot 10^{21}$	-2,3 $0,282 \ln N - 12,7$ +1,5

Зависимости дрейфовой скорости и подвижности носителей заряда от поля удовлетворительно описываются (с погрешностью менее 25 %) следующими выражениями:

$$v = \mu_0 E / (1 + \mu_0 E / v_s); \quad (1.14)$$

$$\mu(E) = v/E = \mu_0 / (1 + \mu_0 E / v_s), \quad (1.15)$$

где для кремния при 300 К дрейфовая скорость насыщения $v_s \approx 10^7$ см/с, а значение μ_0 может быть рассчитано по (1.13).

В области слабых полей (выполняется закон Ома) выражение для плотности дрейфового тока, например, электронов можно представить в виде

$$J_n = q n \mu_n E = \sigma_n E = E / \rho_n. \quad (1.16)$$

Величина $\sigma_n = q \mu_n n$ называется удельной проводимостью и измеряется в $(\text{Ом} \cdot \text{см})^{-1}$, а ρ_n — удельное сопротивление, Ом·см. Удельное сопротивление является основным электрофизическим параметром полупроводника. На рис. 1.3 изображена зависимость удельного сопротивления от концентрации легирующей примеси в кремнии и германии.

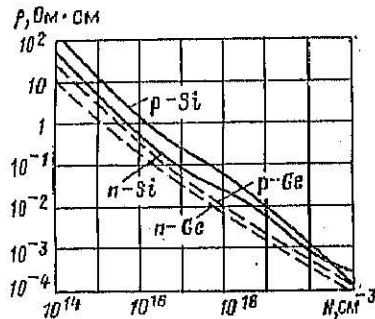


Рис. 1.3. Зависимость удельного сопротивления от концентрации легирующей примеси в кремнии и германии при 300 К

В области очень сильных электрических полей (в полях более 10^5 В/см) проводимость начинает возрастать из-за увеличения концентрации носителей заряда, что связано с ударной ионизацией полупроводника. Это происходит, если носитель заряда набрал энергию, достаточную для разрыва валентной связи — составляющую примерно $1,5 \mathcal{E}_g$. В процессе ударной ионизации электрон из валентной зоны переводится в зону проводимости, а в валентной зоне остается дырка. После акта ионизации ионизирующий носитель также должен остаться в свободном состоянии. Количественно интенсивность ударной ионизации характеризуется коэффициентом ударной ионизации носителей заряда. Он определяется числом электронно-дырочных пар, образуемых носителями заряда на единице пути их движения, в направлении электрического поля в пересчете на один носитель. В ограниченном диапазоне полей для коэффициента ударной ионизации α используются следующие простые аппроксимации:

$$\alpha(E) = A \exp(-B/E); \quad (1.17)$$

$$\alpha(E) \approx A_1 E^n, \quad (1.18)$$

где E — напряженность электрического поля, а величины A, B, A_1 являются параметрами аппроксимации.

В принципе коэффициенты ударной ионизации для электронов α_n и дырок α_p различны (в кремнии $\alpha_n > \alpha_p$), однако для упрощения расчетов считают, что $\alpha_n = \alpha_p$ и равны некоторому эффективному коэффициенту ионизации $\alpha_{эф}$, параметры аппроксимации которого для кремния $A = 7,94 \cdot 10^5 \text{ см}^{-1}$, $B = 1,49 \cdot 10^6 \text{ В/см}$.

1.1.3. ОСНОВНАЯ СИСТЕМА УРАВНЕНИЙ

Физические процессы в любом полупроводниковом приборе могут быть описаны системой нелинейных дифференциальных уравнений в частных производных. Эта система в общем виде содержит два уравнения для плотностей токов электронов и дырок, два уравнения непрерывности электронов и дырок и четыре уравнения Максвелла и описывает поведение носителей заряда в пространстве и во времени.

Ток носителей заряда определяется их диффузией и дрейфом в электрическом поле. Уравнение для плотности токов электронов \vec{J}_n и дырок \vec{J}_p содержит две составляющие — дрейфовую (первый член) и диффузионную (второй

член):

$$\vec{J}_n = q\mu_n n \vec{E} + qD_n \text{grad } n; \quad (1.19)$$

$$\vec{J}_p = q\mu_p p \vec{E} - qD_p \text{grad } p, \quad (1.20)$$

где $q=1,602 \cdot 10^{-19}$ Кл — элементарный заряд; n, p — концентрации электронов и дырок соответственно; μ_n, μ_p — подвижность электронов и дырок; D_n, D_p — коэффициенты диффузии электронов и дырок; \vec{E} — вектор напряженности электрического поля в декартовой системе координат:

$$\vec{E} = \vec{e}_1 E_x + \vec{e}_2 E_y + \vec{e}_3 E_z. \quad (1.21)$$

Величина \vec{E} связана с потенциалом ϕ соотношением

$$\vec{E} = - \text{grad } \phi. \quad (1.22)$$

В (1.21) E_x, E_y, E_z — составляющие вектора напряженности электрического поля по осям x, y, z , а $\vec{e}_1, \vec{e}_2, \vec{e}_3$ — единичные орты, совпадающие по направлению с осями x, y, z . Градиент концентрации электронов

$$\text{grad } n = \vec{e}_1 \frac{\partial n}{\partial x} + \vec{e}_2 \frac{\partial n}{\partial y} + \vec{e}_3 \frac{\partial n}{\partial z}. \quad (1.23)$$

Аналогичным образом выражается градиент концентрации дырок.

Для невырожденного полупроводника (в кремнии при концентрации носителей заряда меньше концентрации вырождения $\sim 2 \cdot 10^{19}$ см $^{-3}$ для $T=300$ К) коэффициент диффузии связан с подвижностью носителей заряда соотношением Эйнштейна

$$D_n = \mu_n kT/q = \mu_n \varphi_T; \quad (1.24)$$

$$D_p = \mu_p kT/q = \mu_p \varphi_T, \quad (1.25)$$

где $k=1,38 \cdot 10^{-23}$ Дж/К $= 8,62 \cdot 10^{-5}$ эВ/К — постоянная Больцмана; T — абсолютная температура, К; $\varphi_T = kT/q$ — температурный потенциал, составляет 0,0258 В при 300 К.

Уравнение непрерывности для электронов имеет вид

$$\frac{1}{q} \text{div } \vec{J}_n - R_n + G_n = \frac{\partial n}{\partial t}. \quad (1.26)$$

Это уравнение описывает сохранение общего количества электронов и связывает изменение концентрации электронов n в некотором заданном элементарном объеме во време-

мени (правая часть уравнения) с изменением n в этом объеме за счет протекания тока электронов (первый член левой части), а также за счет генерации G_n или рекомбинации R_n электронов. Величина R_n называется темпом (скоростью) рекомбинации электронов и определяется уменьшением концентрации электронов в элементарном объеме в единицу времени вследствие рекомбинации. Темп генерации определяется увеличением концентрации электронов за счет теплового, ударного, оптического и других механизмов генерации. В условиях термодинамического равновесия рекомбинация электронов полностью уравновешивает их тепловую генерацию, поэтому $R_n = G_n$. Если нет ударной и оптической генерации, то генерация электронов возможна только за счет тепловой энергии. В этом случае можно говорить о результирующем эффекте генерации — рекомбинации, введя обозначение $RG = R_n - G_n$.

Аналогично (1.26) записывается уравнение непрерывности для дырок:

$$-\frac{1}{q} \text{div } \vec{J}_p - R_p + G_p = \frac{\partial p}{\partial t}. \quad (1.27)$$

Напомним, что дивергенция векторной величины определяется выражением

$$\text{div } \vec{J} = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z}. \quad (1.28)$$

Из многих известных механизмов рекомбинации в кремнии и германии доминирующим является механизм рекомбинации носителей заряда через ловушки. Ловушками (центрами рекомбинации) могут служить атомы ряда элементов таблицы Д. И. Менделеева (золото, платина, медь, серебро и др.), дефекты кристаллической решетки. Ловушки, как правило, создают ряд уровней в запрещенной зоне полупроводника, однако наиболее эффективен с точки зрения рекомбинации один из уровней, который наиболее близок к середине запрещенной зоны. В модели рекомбинации, учитывающей процессы рекомбинации носителей заряда через ловушки, имеющие один-единственный энергетический уровень ξ в запрещенной зоне (модель Шокли-Рида-Холла), темп генерации — рекомбинации носителей заряда определяется выражением:

$$RG = \frac{p_0 n_1^2}{\tau_{n0}(p_0 + p_1) + \tau_{p0}(n_0 + n_1)}, \quad (1.29)$$

где $\tau_{n0} = (N_l \sigma_n v_l)^{-1}$; $\tau_{p0} = (N_l \sigma_p v_l)^{-1}$; N_l — концентрация ловушек; σ_n, σ_p — сечения захвата электрона или дырки на ловушку соответственно; v_l — тепловая скорость носителя заряда, приблизительно равная 10^7 см/с при $T = 300$ К; концентрации n_l и p_l определяются выражениями (1.1) и (1.2), если в них вместо \mathcal{E}_F подставить \mathcal{E}_l .

Температурная зависимость v_l описывается выражением $v_l(T) = v_l(T_0) (T/T_0)^{1/2}$, а сечение захвата

$$\sigma(T) = \sigma(T_0)(T/T_0)^a \sigma. \quad (1.30)$$

Показатель степени a_n для ловушек, находящихся в нейтральном состоянии, близок к нулю. При захвате на притягивающие центры $a_n = -(2,5 \div 4,0)$ и сечение захвата уменьшается с ростом температуры. Притягивающие центры имеют большее сечение захвата, поэтому влияние и температурная зависимость этих центров являются определяющими. Для нейтральных центров сечение захвата в первом приближении определяется размером атома и по порядку величины составляет $(10^{-8} \text{ см})^2 = 10^{-16} \text{ см}^2$. Сфера действия притягивающих центров за счет действия кулоновской силы значительно больше, а сечение захвата для них в 10—100 раз больше. Например, для акцепторного уровня золота в кремнии (расположенного примерно в середине запрещенной зоны) сечение захвата дырки составляет $\sigma_{p2} = 1,15 \cdot 10^{-14} \text{ см}^2$, причем $a_{\sigma p2} = -4$.

Пусть в полупроводник n -типа с равновесными концентрациями n_0 и p_0 введены избыточная концентрация неравновесных дырок Δp и электронов Δn , тогда полная концентрация дырок и электронов $p = p_0 + \Delta p$ и $n = n_0 + \Delta n$. Допустим, что избыточные дырки и электроны введены в равном количестве: $\Delta p \approx \Delta n$. Для определенности рассмотрим примесный полупроводник n -типа, причем $n_0 \gg n_i \gg p_0$. При таких условиях (1.29) можно упростить. Допустим, что $\Delta p \ll n_0$, т. е. избыточная концентрация неосновных носителей — дырок — много меньше равновесной концентрации основных носителей — электронов. В этом случае говорят о *низком уровне инжекции* (НУИ) неосновных носителей заряда, при котором выражение (1.29) можно привести к виду

$$RG = (p - p_0)/\tau_p = \Delta p/\tau_p, \quad (1.31)$$

где τ_p — время жизни дырок при НУИ.

При *высоком уровне инжекции* (ВУИ) дырок, когда

$\Delta p \gg n_0$, преобразование (1.29) дает

$$RG = \Delta p/\tau_{pB}, \quad (1.32)$$

где $\tau_{pB} = \tau_{p0} + \tau_{n0}$ — время жизни дырок при ВУИ.

Физический смысл времени жизни дырок следующий. Если в полупроводнике создать избыточную концентрацию дырок Δp_1 , а затем возбуждение прекратить, то избыточная концентрация дырок за счет их рекомбинации уменьшится во времени по экспоненциальному закону

$$\Delta p(t) = \Delta p_1 \exp(-t/\tau_p).$$

Постоянная τ_p есть постоянная времени спада, или, упрощенно, среднее время существования избыточных дырок в валентной зоне. Аналогичные соображения полностью применимы к электронам.

Время жизни неосновных носителей заряда является важным электрофизическим параметром конкретного полупроводникового материала и указывается в технических условиях. К сожалению, значения τ в процессе изготовления полупроводникового прибора сильно меняются и, как правило, уменьшаются. Диапазон возможных значений τ в кремнии велик и простирается от единиц наносекунд до сотен микросекунд. Экспериментально установлено, что значения τ коррелируют с суммарной концентрацией легирующих примесей $N_\Sigma = N_D + N_A$, и они тем меньше, чем больше N_Σ . Имеется много эмпирических формул, аппроксимирующих эту зависимость. Приведем одну из них:

$$\tau = \tau_0 / (1 + N_\Sigma / N_{\tau 0}). \quad (1.33)$$

Диапазон возможных значений $N_{\tau 0}$ велик (10^{15} — 10^{17} см^{-3}) и определяется как свойствами исходного кремния, так и уровнем технологии. Часто указывается значение $\sim 10^{16} \text{ см}^{-3}$.

В общем случае систему уравнений (1.19), (1.20), (1.26) и (1.27) необходимо дополнить четырьмя уравнениями Максвелла. Если не рассматривать влияние внешних магнитных полей и предположить, что собственное магнитное поле, обусловленное протеканием тока через прибор, мало (для кремния и германия в рабочем диапазоне температур это допущение справедливо), то достаточно рассматривать только уравнение Пуассона $\text{div } \vec{D} = \rho$, которое в случае изотропного полупроводника имеет вид

$$\epsilon_r \epsilon_0 \text{div } \vec{E} = q(p - n + N_D - N_A). \quad (1.34)$$

Это уравнение справедливо при условии, что доноры и акцепторы полностью ионизованы, а их концентрация много больше концентрации примеси с глубокими уровнями, определяющей рекомбинационные процессы; это условие, как правило, хорошо выполняется на практике.

Таким образом, основная система уравнений состоит из пяти перечисленных уравнений, которые являются основой анализа процессов в полупроводниковых приборах. Эту систему уравнений часто называют *фундаментальной системой уравнений* (ФСУ) физики полупроводниковых приборов. Для выяснения сущности физических процессов и анализа большого числа полупроводниковых приборов достаточно ограничиться одномерным приближением, для которого неизвестные пока величины n , p , J_n , J_p , E (или φ) меняются в одном направлении (вдоль оси x). В одномерном приближении ФСУ принимает вид

$$J_n = q\mu_n nE + qD_n \frac{\partial n}{\partial x}; \quad (1.35)$$

$$J_p = q\mu_p pE - qD_p \frac{\partial p}{\partial x}; \quad (1.36)$$

$$\frac{1}{q} \frac{\partial J_n}{\partial x} - \frac{n - n_0}{\tau_n} = \frac{\partial n}{\partial t}; \quad (1.37)$$

$$-\frac{1}{q} \frac{\partial J_p}{\partial x} = \frac{p - p_0}{\tau_p} = \frac{\partial p}{\partial t}; \quad (1.38)$$

$$\epsilon_r \epsilon_0 \frac{\partial E}{\partial x} = q(p - n + N_D - N_A). \quad (1.39)$$

При записи этих уравнений мы опустили индексы x у переменных n , кроме того, считали, что τ_n или τ_p постоянны, что справедливо либо при низком, либо при высоком уровне инжекции.

Решение системы указанных нелинейных дифференциальных уравнений в аналитическом виде невозможно. Тем не менее, принимая некоторые физически обоснованные и подкрепленные экспериментом допущения и упрощения, получаем аналитические решения для большого числа частных случаев. Это оказывается возможным на основе так называемых *региональных методов*, суть которых состоит в том, что структуру полупроводникового прибора разбивают на ряд областей, для которых уравнения системы удается свести к виду, допускающему аналитические решения (как правило, линейному). Другой подход основан на численном решении ФСУ с применением ЭВМ.

1.1.4. ЭФФЕКТЫ В ПОЛУПРОВОДНИКАХ ПРИ ВЫСОКИХ УРОВНЯХ ЛЕГИРОВАНИЯ И БОЛЬШИХ КОНЦЕНТРАЦИЯХ НОСИТЕЛЕЙ ЗАРЯДА

Экспериментально установлено, что при высоком уровне легирования, когда концентрация примеси достаточно большая, ширина запрещенной зоны полупроводника начинает уменьшаться с ростом концентрации примеси. Это связано прежде всего с тем, что при концентрациях примеси более 10^{18} см^{-3} среднее расстояние между атомами примеси становится меньше $10^{-6} \text{ см} = 10 \text{ нм}$, что сравнимо с длиной волны электрона. При таких концентрациях имеет место перекрытие волновых функций электронов, что приводит к расщеплению примесных уровней в примесные подзоны и образованию «хвостов» плотности состояний в запрещенной зоне, как показано на рис. 1.4, б. Единый примесный уровень доноров с энергией \mathcal{E}_D (рис. 1.4, а) при высоком

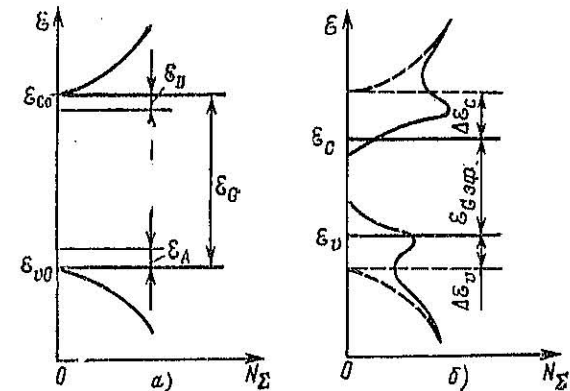


Рис. 1.4. Структура энергетических зон в слаболегированном (а) и сильнолегированном (б) полупроводниках

легированию расщепляется в примесную подзону, которая перекрывается с зоной проводимости. Ширина запрещенной зоны при этом уменьшается на некоторую величину $\Delta \mathcal{E}_G$ и принимает значение $\mathcal{E}_{G,эф}$, которое уменьшается с ростом уровня легирования. В практических расчетах используют эмпирические аппроксимации зависимости *сужения запрещенной зоны* $\Delta \mathcal{E}_G$ от суммарной концентрации легирующей примеси $N_\Sigma = N_D + N_A$. Одна из таких аппроксимаций для

кремния при 300 К имеет вид

$$\Delta \mathcal{E}_a = \mathcal{E}_a - \mathcal{E}_{a-1} = V_1 \ln(N_a/N_a), \quad (1.40)$$

где сужение ширины запрещенной зоны $\Delta \mathcal{E}_a$ измеряется в эВ; $V_1 = 0,018$ эВ, а N_a — в см^{-3} ; эмпирическая постоянная $N_a \approx 10^{17} \text{ см}^{-3}$. Начиная со значения N_a , запрещенная зона существенно сужается, что приводит к ряду новых физических эффектов.

Как следует из рассмотрения рис. 1.4, б, эффективное значение энергии дна зоны проводимости уменьшается на величину $\Delta \mathcal{E}_c$, а эффективное значение энергии потолка валентной зоны повышается на величину $\Delta \mathcal{E}_v$.

Если электроны и дырки подчиняются статистике Максвелла—Больцмана, то выражения для их равновесных концентраций можно записать в форме, аналогичной (1.1) и (1.2):

$$n_i = N_i \exp\left(-\frac{\mathcal{E}_{i0} - \Delta \mathcal{E}_c - \mathcal{E}_F}{kT}\right); \quad (1.41a)$$

$$p_i = N_i \exp\left(\frac{\mathcal{E}_{i0} + \Delta \mathcal{E}_v - \mathcal{E}_F}{kT}\right). \quad (1.41б)$$

С учетом очевидного соотношения $\Delta \mathcal{E}_a = \Delta \mathcal{E}_c + \Delta \mathcal{E}_v$

$$p_0 n_0 = n_{i\Phi}^2 \exp(\Delta \mathcal{E}_a / (kT)). \quad (1.42)$$

Таким образом, с ростом уровня легирования эффективная собственная концентрация $n_{i\Phi}$ сильно возрастает; резко увеличивается и концентрация неосновных носителей заряда. Например, для *n*-кремния с $N_D = 10^{19} \text{ см}^{-3}$ при $T = 292$ К и $N_a = 10^{17} \text{ см}^{-3}$ получим $\Delta \mathcal{E}_a = 0,083$ эВ, $n_{i\Phi}/n_i = 5,24$, а равновесная концентрация дырок p_0 оказывается в 27,5 раза выше, чем без учета эффектов высокого уровня легирования (ВУЛ).

Рассмотрим поведение неоднородного полупроводника с непрерывно меняющимся уровнем легирования и шириной запрещенной зоны. Вначале представим такой полупроводник в виде трех изолированных областей 1—3, не взаимодействующих друг с другом, как показано на энергетической диаграмме на рис. 1.5, а.

Уровни энергии \mathcal{E}_0 определяют энергию электрона, который вышел из полупроводника в вакуум и является свободным относительно некоторого отчетного уровня $\mathcal{E}_{отсч}$. Будем полагать, что изменение ширины запрещенной зоны обусловлено эффектами ВУЛ, а для третьей области уро-

вень легирования невысок и величина \mathcal{E}_a равна исходному значению в слаболегированном полупроводнике (например, для кремния при 300 К $\mathcal{E}_a = 1,11$ эВ). На рис. 1.5, а величина χ_n представляет собой энергию электронного сродства, являющуюся истинной работой выхода электрона из полупроводника. В слаболегированном кремнии $\chi_n = 4,05$ эВ.

Если все три области объединены в систему, находящуюся в термодинамическом равновесии, то области начина-

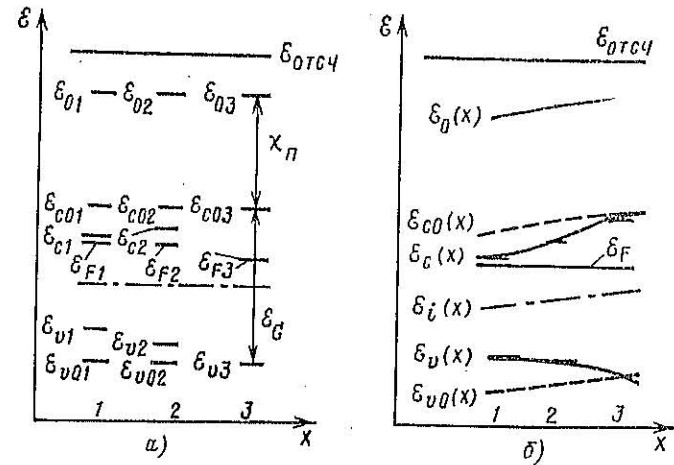


Рис. 1.5. Энергетическая зонная диаграмма трех областей полупроводника до (а) и после (б) их объединения

ют обмениваться электронами, в результате чего все уровни энергии Ферми выравниваются. На рис. 1.5, б показано положение уровней в результате установления равновесия, и толстыми линиями отмечен их ход в неоднородном полупроводнике с непрерывно меняющимися свойствами. Преимущественный переход электронов слева направо (в направлении *x* из области 1 в область 2 и далее из области 2 в область 3) приведет к тому, что правый край образца будет обогащен электронами, а левый обеднен. Это и определяет изменение энергетического уровня вакуума $\mathcal{E}_0(x)$. Разделение электрического заряда приводит к появлению электрического поля в вакууме с напряженностью *E*, совпадающего с направлением оси *x*, а значит, и к изменению

потенциала. На краю области β , где накапливается отрицательный заряд, потенциал понижается, а энергия электронов, обладающих отрицательным зарядом, увеличивается. Именно поэтому все энергетические уровни в области β оказываются выше относительно уровней в области I .

Разность потенциалов между областями β и I называется *контактной разностью потенциалов*. Учитывая, что разность потенциалов измеряется работой, которую надо совершить, чтобы перенести единичный положительный заряд из области I в область β , а заряд электрона составляет $-q$, можно записать $\varphi_\beta - \varphi_I = -(\mathcal{E}_{0\beta} - \mathcal{E}_{0I})/q$.

Напряженность электростатического поля, обусловленная градиентом потенциала,

$$E = -\frac{d\varphi}{dx} = \frac{1}{q} \frac{d\mathcal{E}_0(x)}{dx}. \quad (1.43a)$$

Значительно сложнее для понимания вопрос о перемещении электрона в зоне проводимости из области I в область β . Изменение ширины запрещенной зоны означает изменение силы связи электрона с решеткой кристалла, меняющейся от области к области. В полупроводнике все свободные электроны сосредоточены у дна зоны проводимости и имеют энергию, не сильно отличающуюся от величины $E_c(x)$ в каждой точке x . Поэтому сила, действующая на электроны, определяется градиентом энергии электронов и с учетом соотношения $\mathcal{E}_c = \mathcal{E}_{c0} - \Delta\mathcal{E}_c$

$$F_n = -\frac{d\mathcal{E}_c(x)}{dx} = -\frac{d\mathcal{E}_{c0}(x)}{dx} + \frac{d\Delta\mathcal{E}_c(x)}{dx}.$$

Этой силе можно сопоставить некоторую эффективную полную напряженность электрического поля с помощью соотношения

$$E_n = -\frac{F_n}{q} = \frac{1}{q} \frac{d\mathcal{E}_{c0}}{dx} - \frac{1}{q} \frac{d\Delta\mathcal{E}_c(x)}{dx}. \quad (1.43b)$$

Первый член этого выражения является истинной напряженностью электростатического поля и определяется формулой (1.43a), так как уровни энергии $\mathcal{E}_0(x)$, $\mathcal{E}_{c0}(x)$, $\mathcal{E}_{v0}(x)$ эквидистантны, т. е. расстояния между ними неизменны (рис. 1.5, б). Второй член отражает наличие добавочной силы, связанной с изменением валентных сил в кристалле, и в данном случае связана с изменением уровня легирования от области к области. Важно отметить, что поток электронов связан именно с силой F_n , а значит, с полной напряженностью поля E_n , действующей на электроны. Поэтому плотность дрейфового тока электронов с учетом сужения запрещенной зоны

$$J_n = q\mu_n n E_n = q\mu_n n \left(E - \frac{1}{q} \frac{d\Delta\mathcal{E}_c(x)}{dx} \right). \quad (1.44)$$

Аналогично можно получить выражение для плотности дрейфового тока дырок:

$$J_p = q\mu_p p E_p = q\mu_p p \left(E + \frac{1}{q} \frac{d\Delta\mathcal{E}_v(x)}{dx} \right). \quad (1.45)$$

Сдвиг края зоны проводимости можно связать с сужением запрещенной зоны соотношением $\Delta\mathcal{E}_c = A\Delta\mathcal{E}_G$, аналогично $\Delta\mathcal{E}_v = (1-A)\Delta\mathcal{E}_G$, где A называют коэффициентом асимметрии в сужении. В настоящее время окончательно вопрос о величине A не решен, но на основании сопоставления теории с экспериментом многие исследователи полагают $A=0.5$. Если $A=0.5$, то выражения (1.41) можно представить в следующем виде:

$$n = n_{i0p} \exp((\mathcal{E}_F - \mathcal{E}_i)/(kT)); \quad (1.46a)$$

$$p = n_{i0p} \exp((\mathcal{E}_i - \mathcal{E}_F)/(kT)). \quad (1.46b)$$

Таким образом, в варьированном полупроводнике (с меняющейся величиной \mathcal{E}_G) силовое воздействие поля на дырки и электроны различается: в данном случае (рис. 1.5, б) для электронов роль дрейфовой составляющей тока увеличивается, а для дырок — ослабляется. Наличие эффектов ВУЛ может качественно изменить характер движения дырок. Например, если при отсутствии эффекта сужения запрещенной зоны дырка, вводимая в образец со стороны области β (рис. 1.5, б), попадет в тормозящее поле, то с учетом эффекта ВУЛ наклон зоны \mathcal{E}_v таков, что дырка начинает не тормозиться, а ускоряться, что оказывает существенное влияние на токоперенос.

Выразив сужение запрещенной зоны через потенциал $\Delta\mathcal{E}_G = A\Delta\mathcal{E}_G/q$, полные плотности токов электронов и дырок с учетом дрейфовых и диффузионных составляющих представим в виде

$$J_n = q\mu_n n \left(E - \frac{d(-A\Delta\mathcal{E}_G)}{dx} \right) + qD_n \frac{dn}{dx}; \quad (1.47a)$$

$$J_p = q\mu_p p \left(E + \frac{d((1-A)\Delta\mathcal{E}_G)}{dx} \right) - qD_p \frac{dp}{dx}. \quad (1.47b)$$

При большой концентрации введенных в полупроводник носителей заряда возникает ряд новых эффектов, влияние которых несущественно при низкой плотности носителей. Одним из таких эффектов является *электронно-дырочное рассеяние*. Конечная подвижность носителей заряда в полупроводниках связана с их рассеянием на различного вида неоднородностях, нарушениях и дефектах кристаллической решетки. При малых концентрациях новых легирующей примеси и носителей заряда основную роль играет рассеяние носителей на тепловых колебаниях решетки (акустиче-

ских фононах). Это рассеяние определяет значения подвижности электронов и дырок. Например, в кремнии при комнатной температуре подвижность электронов μ составляет примерно $1000 \text{ см}^2/(\text{В}\cdot\text{с})$. Электроны, двигаясь с тепловой скоростью $\approx 10^7 \text{ см/с}$, за время свободного пробега $\approx 10^{-12} \text{ с}$ проходят среднее расстояние $l \approx 10^{-5} \text{ см}$, называемое *длиной свободного пробега*, испытывают столкновения с фононами и, меняя направление и значение скорости, продолжают двигаться далее. Время существования электрона в зоне проводимости (время жизни) $\tau_n \approx 10^{-6} \text{ с}$, и электрон за свое время жизни испытывает до 10^6 актов рассеяния.

Положение меняется, если концентрация примеси такая, что расстояние между ионами примеси $a \sim N_D^{1/3}$ сравнимо или меньше l . Электрон подвергается воздействию кулоновского поля иона, изменяет направление своего движения и проходит в направлении силовой линии поля путь, меньший, чем ранее, а поэтому значение μ уменьшается. Эти эффекты начинают играть роль при концентрации ионов $N > 10^{15} \text{ см}^{-3}$ (при комнатных температурах). Характерной особенностью как решеточного, так и ионного рассеяния является то, что эти виды рассеяния одинаково влияют на подвижность μ и коэффициент диффузии D носителей заряда, что количественно выражается в соблюдении соотношения Эйнштейна $D = \mu kT/q$.

Электронно-дырочное рассеяние, т. е. рассеяние движущихся электронов на дырках, и наоборот, рассеяние дырок на электронах, имеет много общего с рассеянием носителей на ионах примеси. Рассеяние происходит под действием кулоновских сил. Пороговая концентрация начала электронно-дырочного рассеяния также порядка 10^{15} см^{-3} . Однако имеются и отличия. В заданном электрическом поле электроны и дырки движутся навстречу друг другу. Их эффективные массы соизмеримы между собой. Поэтому дрейфовый поток электронов сильно «тормозится» дрейфовым потоком дырок и дрейфовая подвижность сильно уменьшается с ростом концентрации носителей заряда. В то же время влияние электронно-дырочного рассеяния на коэффициенты диффузии электронов и дырок слабое. В большинстве полупроводниковых приборов возникают условия, при которых в большей части прибора избыточные концентрации $\Delta n \approx \Delta p$. Диффузионные потоки дырок и электронов направлены в одну и ту же сторону, что приводит к перераспределению направленного импульса между этими носителями, но не к его затуханию. Перераспределение им-

пульса выражается в выравнивании диффузионных скоростей носителей заряда и приводит к равенству $D_n = D_p = D$. На рис. 1.6 показаны зависимости подвижности и коэффициентов диффузии электронов и дырок в кремнии в функции их концентрации при условии высокого уровня инжекции $n \approx p$.

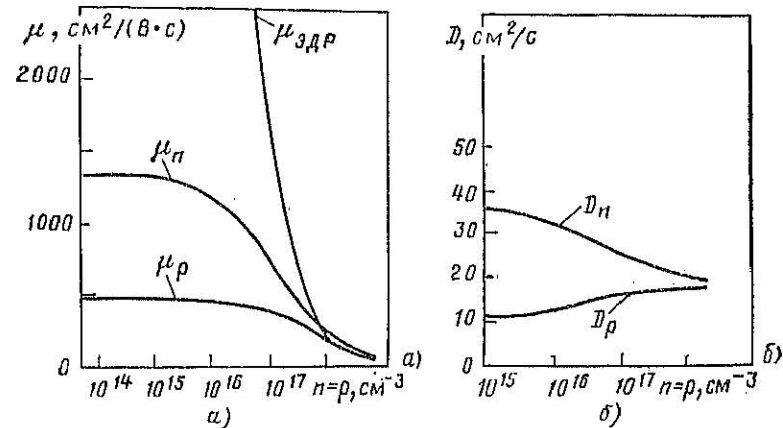


Рис. 1.6. Зависимость подвижности (а) и коэффициентов диффузии электронов и дырок (б) от концентрации носителей заряда в кремнии при 300 К

При больших концентрациях электронов и (или) дырок в полупроводниках вступает в действие еще один механизм рекомбинации — *межзонная ударная рекомбинация (Оже-рекомбинация)*, при этом происходит столкновение одновременно двух свободных электронов и одной дырки или двух дырок и одного свободного электрона, в результате чего имеет место рекомбинация электрона и дырки и переход третьего носителя на более высокий уровень энергии в соответствующей зоне. Этот третий носитель в результате столкновений с решеткой передает ей избыточную энергию и приходит в равновесное состояние с решеткой.

Вероятность столкновения пары электрон—дырка со свободным электроном пропорциональна $n^2 p$, а с дыркой $p^2 n$. Поэтому темп Оже-рекомбинации определяется выражением

$$RG = (pn - n_{\text{эф}}^2)(C_n n + C_p p). \quad (1.48a)$$

Значения коэффициентов ударной рекомбинации C_n и C_p , приведенные различными исследователями, сильно различаются. Например, для кремния $C_n = 1 \cdot 10^{-31} + 5 \cdot 10^{-30} \text{ см}^6/\text{с}$. В последнее время на основании современных исследований силовых ПП можно считать, что $C_n \approx 6 \cdot 10^{-31} \text{ см}^6/\text{с}$, $C_p \approx 3 \cdot 10^{-31} \text{ см}^6/\text{с}$ в кремнии при $T = 300 \text{ К}$.

В сильнолегированном n^+ -слое при низком уровне инжекции ($\Delta p \ll n_0 \approx N_D$) выражение для темпа генерации-рекомбинации можно упростить:

$$RG \approx C_n N_D^2 \Delta p = \Delta p / \tau_{pA}. \quad (1.48б)$$

В этом выражении $\tau_{pA} = (C_n N_D^2)^{-1}$ — время жизни неравновесных дырок при Оже-рекомбинации, которое резко уменьшается при увеличении степени легирования полупроводника.

В сильнолегированном p^+ -полупроводнике

$$RG \approx C_p N_A^2 \Delta n = \Delta n / \tau_{nA}, \quad (1.48в)$$

где $\tau_{nA} = (C_p N_A^2)^{-1}$ — время жизни неравновесных электронов при низком уровне инжекции.

При высоком уровне инжекции носителей заряда $n \approx p \gg N_A, N_D$ и время жизни неравновесных носителей заряда при Оже-рекомбинации

$$\tau_A \approx ((C_p + C_n) n^2)^{-1} = ((C_p + C_n) p^2)^{-1}. \quad (1.48г)$$

1.2. ПЕРЕХОДЫ МЕТАЛЛ — ПОЛУПРОВОДНИК

1.2.1. ВЫСОТА ПОТЕНЦИАЛЬНОГО БАРЬЕРА

Переходы металл—полупроводник используются для создания полупроводниковых диодов (диодов Шоттки), а также для создания омических переходов (контактов), являющихся неотъемлемой частью любого полупроводникового прибора.

Рассмотрим систему, состоящую из металла и полупроводника n -типа электропроводности, для случая, когда работа выхода электронов из металла A_M больше работы выхода электрона из полупроводника $A_{П}$. Для металла работа выхода электрона измеряется минимально необходимой энергией для удаления электрона из металла в вакуум. Как принято в физике твердого тела, будем измерять эту энергию в электрон-вольтах (эВ).

На рис. 1.7, а энергетические зонные диаграммы металла и полупроводника рассмотрены изолированно, металл

и полупроводник не приведены в соприкосновение друг с другом и их система не находится в термодинамическом равновесии. Для полупроводника минимальная энергия для удаления электрона со дна зоны проводимости в вакуум называется энергией электронного средства $\chi_{П}$, или *истинной работой выхода*.

Величина $A_{П}$, равная разности энергий, соответствующих уровню Ферми полупроводника \mathcal{E}_F и уровню энергии

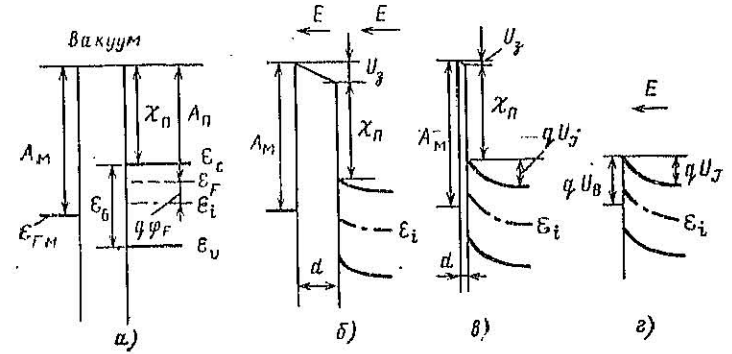


Рис. 1.7. Энергетические зонные диаграммы перехода металл — полупроводник с различной шириной зазора

вакуума, называется *термоэлектронной работой выхода*. Если обозначить разницу между уровнями Ферми \mathcal{E}_F и уровнем Ферми в собственном полупроводнике \mathcal{E}_i через $q\varphi_F = \mathcal{E}_F - \mathcal{E}_i$, то, как следует из рис. 1.7, а, величины $\chi_{П}$ и $A_{П}$ связаны между собой выражением (для n -полупроводника)

$$A_{П} = \chi_{П} + \mathcal{E}_G/2 - q\varphi_F. \quad (1.49)$$

Таким образом, термоэлектронная работа выхода $A_{П}$ в отличие от $\chi_{П}$ зависит от степени легирования полупроводника через зависимость

$$\varphi_F = \varphi_T \ln(N_D/n_i). \quad (1.50)$$

Если металл и полупроводник соединить внешней цепью, то по ней из полупроводника в металл перетечет некоторое количество электронов, установится термодинамическое равновесие, энергия уровня Ферми станет постоянной во всей системе (рис. 1.7, б). В зазоре между металлом и полупроводником и в приповерхностном слое полупроводника

возникнет электрическое поле, обусловленное тем, что на внутренней поверхности металла накопятся электроны, а на поверхности полупроводника обнажится заряд нескомпенсированных положительных ионов доноров. Разность потенциалов в зазоре будет определяться соотношением $U_3 = dE_3$, где d — толщина зазора, E_3 — напряженность электрического поля в зазоре. Обеднение поверхности полупроводника электронами в приповерхностном слое приведет к изгибу зон вверх.

При сближении металла и полупроводника $U_3 \rightarrow 0$, так как E_3 — конечная величина (рис. 1.7, в). Когда металл и полупроводник в конце концов придут в контакт, то барьер, обусловленный зазором, исчезнет ($U_3 = 0$), и получится идеальный переход металл—полупроводник (рис. 1.7, г). Высота потенциального барьера для электронов, переходящих из металла в полупроводник, будет в этом случае определяться соотношением

$$U_B = (A_M - \chi_{П})/q. \quad (1.51)$$

Для большинства реальных контактов идеальная ситуация, изображенная на рис. 1.7, г, никогда не выполняется, так как на поверхности полупроводника обычно существует тонкий диэлектрический слой толщиной 0,5—2 нм, который называют *промежуточным*. Для кремния промежуточный слой является слоем диоксида кремния (оксида) и обусловлен технологическими причинами. Таким образом, реальные контакты больше соответствуют случаю, приведенному на рис. 1.7, в. Однако толщина барьера для электронов, обусловленного слоем оксида, настолько мала, что электроны могут легко туннелировать сквозь него. Кроме того, падение напряжения U_3 на пленке оксида настолько мало, что уравнение (1.51) является хорошей аппроксимацией для высоты барьера.

Высота потенциального барьера для электронов, переходящих из полупроводника в металл, qU_j , определяется изгибом зон в полупроводнике (рис. 1.7, г), где U_j носит название *контактной разности потенциалов*. Ее величина, как следует из рассмотрения рис. 1.7,

$$U_j = (A_M - A_{П})/q. \quad (1.52)$$

При принятом соотношении $A_M > A_{П}$ приповерхностная область полупроводника обедняется электронами. Под действием электрического поля электроны уходят из этой области, оставляя нескомпенсированный заряд положитель-

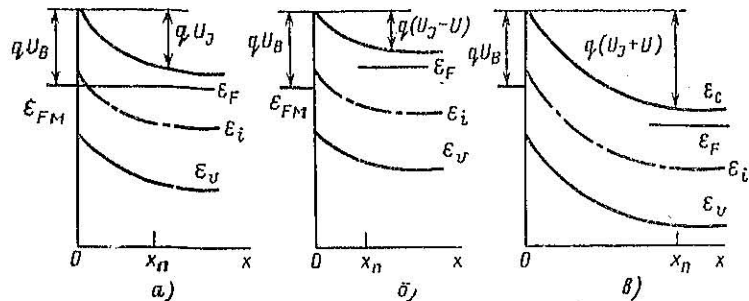


Рис. 1.8. Энергетические зонные диаграммы перехода металл—полупроводник:

а — $U=0$; б — $U>0$; в — $U<0$

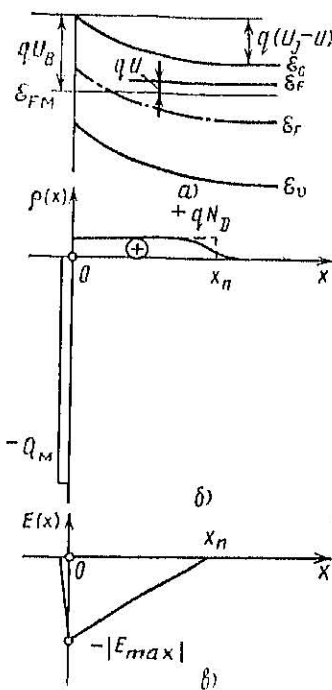


Рис. 1.9. Энергетическая зонная диаграмма перехода М—П (а), распределение объемного заряда (б) и электрического поля (в)

ных ионов доноров. При достаточно большой высоте барьера концентрация электронов, а также дырок в некоторой приповерхностной области полупроводника (при $0 < x < x_n$) оказывается много меньше концентрации ионов доноров N_D . Поэтому такую область называют *областью пространственного заряда* (ОПЗ) или *обедненной областью*.

Рассмотрим, что произойдет, если к переходу металл—полупроводник приложить внешнее напряжение U . Так как

ОПЗ обеднена носителями заряда, ее сопротивление велико и практически все напряжение U прикладывается к этой области. Если к металлу подключить отрицательный полюс источника напряжения U , то напряженность электрического поля в ОПЗ увеличится, разность потенциалов на ОПЗ возрастет до величины $(U_j + U)$. Возросшее электрическое поле будет оттягивать вблизи границы x_n электроны из

квазинейтральной области полупроводника и переносить их через внешнюю цепь в металл, вследствие чего толщина ОПЗ увеличится, как показано на рис. 1.8, в. Такое включение перехода М—П называют *обратным*.

Если к металлу подключен положительный полюс источника напряжения U , то поле в ОПЗ перехода уменьшится, высота потенциального барьера станет равной $(U_J - U)$ и толщина ОПЗ сократится. Это случай *прямого смещения* перехода.

Толщину области пространственного заряда x_n , распределение электрического поля и потенциала в ОПЗ можно определить с помощью решения уравнения Пуассона, предположив, что в ОПЗ концентрация дырок и электронов много меньше концентрации доноров N_D , граница между ОПЗ и нейтральной областью резкая и напряженность электрического поля $E(x_n) = 0$, как показано на рис. 1.9. Обычно в рабочем диапазоне температур ($-60 \div +125^\circ\text{C}$) атомы донорной примеси полностью ионизованы, поэтому объемный заряд в ОПЗ $\rho = +qN_D$ и уравнение Пуассона (1.34) принимает вид

$$\epsilon_r \epsilon_0 \frac{dE}{dx} = qN_D. \quad (1.53)$$

Взяв интеграл с помощью разделения переменных, можно получить

$$E(x) = qN_D (x - x_n) / (\epsilon_r \epsilon_0). \quad (1.54)$$

Распределение $E(x)$ изображено на рис. 1.9, в. Разность потенциалов на ОПЗ составляет $(U_J - U)$, поэтому

$$U_J - U = - \int_0^{x_n} E(x) dx = \frac{|E_{\max}| x_n}{2} = \frac{qN_D x_n^2}{2\epsilon_r \epsilon_0}. \quad (1.55)$$

Из последнего соотношения можно определить

$$x_n = \sqrt{2\epsilon_r \epsilon_0 (U_J - U) / (qN_D)}; \quad (1.56)$$

$$E_{\max} = 2(U_J - U) / x_n = \sqrt{2qN_D (U_J - U) / (\epsilon_r \epsilon_0)}. \quad (1.57)$$

1.2.2. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ПЕРЕХОДА МЕТАЛЛ—ПОЛУПРОВОДНИК (ДИОДА ШОТТКИ)

Перейдем к рассмотрению вольт-амперной характеристики (ВАХ) диода Шоттки, т. е. функциональной зависимости тока диода от приложенного напряжения. Теория

ВАХ диода Шоттки базируется на теории термоэлектронной эмиссии электронов.

Обозначим плотность тока термоэлектронной эмиссии электронов из металла в полупроводник J_{M-P} и, наоборот, из полупроводника в металл J_{P-M} . Как известно, плотность тока термоэлектронной эмиссии из металла в вакуум определяется формулой Ричардсона — Дэшмана

$$J = A_R T^2 \exp(-A_M / (kT)), \quad (1.58)$$

где $A_R = 4\pi q m^* k^2 / h^3$ — постоянная Ричардсона; m^* — эффективная масса электрона; $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана; $h = 6,62 \cdot 10^{-34}$ Дж·с — постоянная Планка; T — абсолютная температура, К.

В приближении свободных электронов $m^* = m_0 = 9,1 \cdot 10^{-31}$ кг значение $A_R = 120$ А/(см²·К²).

Если промежуточный диэлектрический слой между металлом и полупроводником туннельно тонок, то для перехода из металла в полупроводник электронам нужно преодолеть потенциальный барьер высотой qU_B , как показано на рис. 1.7, в. Воспользовавшись формулой (1.58) с учетом того, что в данном случае высота потенциального барьера равна не A_M , а qU_B , запишем

$$J_{M-P} = A_R T^2 \exp(-qU_B / (kT)). \quad (1.59)$$

Поскольку все внешнее напряжение U падает на запирающем слое (ОПЗ), то величина qU_B со стороны металла, а значит, и J_{M-P} не зависят от U . В состоянии термодинамического равновесия $J_{M-P} = J_{P-M}$, и поэтому $J_{P-M} = A_R T^2 \exp(-qU_B / (kT))$. При наличии внешнего прямого смещения высота потенциального барьера для электронов, текущих из полупроводника в металл, уменьшится на величину U , поэтому

$$J_{P-M} = A_R T^2 \exp(-q(U_B - U) / (kT)). \quad (1.60)$$

Результирующая плотность тока через диод $J = J_{P-M} - J_{M-P}$. Учитывая, что площадь перехода М—П равна S , окончательно записываем выражение для ВАХ диода Шоттки:

$$I = SJ_s (\exp(qU / (kT)) - 1) = I_s (\exp(U / \Phi_T) - 1). \quad (1.61)$$

В последнем выражении предэкспоненциальный множитель определяется формулой

$$J_s = A_R T^2 \exp(-qU_B / (kT)). \quad (1.62)$$

В соответствии с (1.61) при прямом смещении ($U > 0$) прямой ток экспоненциально увеличивается с ростом U . При обратном включении $U < 0$ и ток диода стремится к постоянному значению $-I_s = -SI_s$.

В реальных диодах Шоттки зависимость высоты барьера qU_B значительно слабее зависит от работы выхода A_M , чем это следует из (1.51). Это обусловлено тем, что в нижней трети запрещенной зоны полупроводников имеется большое количество поверхностных состояний, которые могут принять на «себя» — захватить — большое количество электронов, переходящих из металла. В этом случае поверхностные состояния, захватывая электроны из объема полупроводника, создают обеднение на поверхности и достаточно сильный изгиб зон, фиксируя уровень Ферми полупроводника в нижней трети запрещенной зоны. В системе с металлом эти поверхностные состояния, захватывая электроны металла, как бы экранируют объем полупроводника, в результате чего высота барьера qU_B слабо зависит от значения A_M металла и не обнаруживает четких изменений в функции концентрации доноров. Другое явление заключается во влиянии сил «изображения» при переходе электронов из полупроводников в металл. Это эквивалентно понижению высоты потенциального барьера, которое наиболее сильно сказывается при обратных смещениях перехода М — П, приводя к нарушению постоянства I_s .

1.2.3. ОМИЧЕСКИЕ ПЕРЕХОДЫ

На основе переходов металл — полупроводник изготавливают омические контакты к любым полупроводниковым приборам. Сопротивление такого перехода должно быть мало и практически не должно зависеть от направления и значения тока в рабочем диапазоне токов. Омический переход не должен инжектировать неосновные носители заряда, иметь стабильные электрические и механические свойства. Металл перехода должен обладать высокой электро- и теплопроводностью и температурным коэффициентом расширения (ТКР), близким к ТКР полупроводника. Наиболее сложно создать хорошие омические контакты к относительно слаболегированным полупроводникам ($N < 10^{17} \text{ см}^{-3}$). Например, с кремнием n -типа электропроводности практически все металлы дают барьер Шоттки с относительно большой высотой барьера (более 0,5 эВ) и контакт получается выпрямляющим. Выходом из положения

является создание на поверхности n -полупроводника n^+ -слоя с концентрацией доноров выше 10^{19} см^{-3} , на который наносится металл. Энергетическая диаграмма такого контакта изображена на рис. 1.10.

Поскольку степень легирования n^+ -слоя высокая (более 10^{19} см^{-3}) и полупроводник вырожден, то толщина ОПЗ чрезвычайно мала (несколько нанометров) и носители заряда (электроны) беспрепятственно преодолевают барьер за счет туннельного эффекта.

Наиболее часто в качестве металла для омических контактов на кремнии используют алюминий, сплавы металлов (платины, никеля), эвтектику золото—кремний. В частности, алюминий, являясь акцепторной примесью, дает хорошие омические контакты на p -кремнии. Подавлению инжекции неосновных носителей заряда способствует нарушение поверхности полупроводника (например, шлифовкой), за счет чего образуются дефекты кристаллической решетки, являющиеся эффективными центрами рекомбинации. Подобную роль в кремнии выполняют золото и платина.

Важной характеристикой омического перехода является удельное сопротивление контакта ρ_c (сопротивление, умноженное на площадь). Хороший омический контакт должен иметь $\rho_c < 10^{-3} \text{ Ом} \cdot \text{см}^2$.

1.3. ДИОДЫ НА ОСНОВЕ ЭЛЕКТРОННО-ДЫРОЧНЫХ ПЕРЕХОДОВ

Полупроводниковым диодом называют полупроводниковый прибор с одним электрическим переходом и двумя выводами.

1.3.1. ЭНЕРГЕТИЧЕСКАЯ ДИАГРАММА p - n ПЕРЕХОДА. КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ

Электрический переход между двумя областями полупроводника, одна из которых имеет электропроводность n -типа, а другая p -типа, называется электронно-дырочным

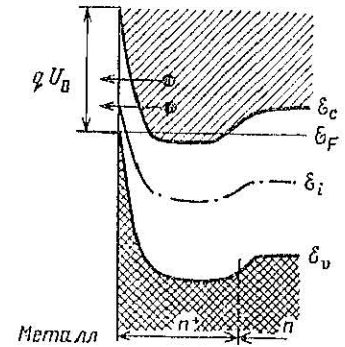


Рис. 1.10. Энергетическая зонная диаграмма омического перехода

переходом (p - n переходом). Диоды на основе электронно-дырочного перехода нашли наибольшее применение. Электронно-дырочный переход можно создать внутри полупроводника, если ввести в одну его область донорную примесь, а в другую акцепторную (рис. 1.11, а). В рабочем диапазоне температур атомы примесей полностью ионизованы, концентрации электронов в n -области, а дырок в p -области, т. е. концентрации основных носителей заряда вдали от границы раздела p - и n -области, можно считать равными концентрации соответствующей примеси: $n_{n0} \approx N_D$, $p_{p0} \approx N_A$. Равновесные концентрации неосновных носителей будут равны: $p_{p0} = n_i^2 / p_{p0}$, $n_{n0} = n_i^2 / n_{n0}$. Будем считать, что полупроводник достаточно сильно легирован, т. е. $n_{n0} \gg n_i$, $p_{p0} \gg n_i$, и поэтому $p_{n0} \ll n_i$, $n_{p0} \ll n_i$. Для определенности положим, что $N_A \gg N_D$, т. е. будем рассматривать резконе-

симметричный p^+-n переход.

Рассмотрим физические процессы, происходящие в областях p - n перехода в условиях термодинамического равновесия, т. е. при постоянной температуре и отсутствии внешнего напряжения смещения. Так как концентрация дырок в p -области много больше концентрации дырок в n -области, то дырки из p -области будут диффундировать в n -область, при этом в p -области у границы раздела останутся неподвижные отрицательные ионы акцепторов и возникнет отрицательный объемный заряд $-qN_A$. Дырки, переходя в n -область, рекомбинируют с электронами, в результате чего концентрация электронов справа от границы уменьшается (рис.

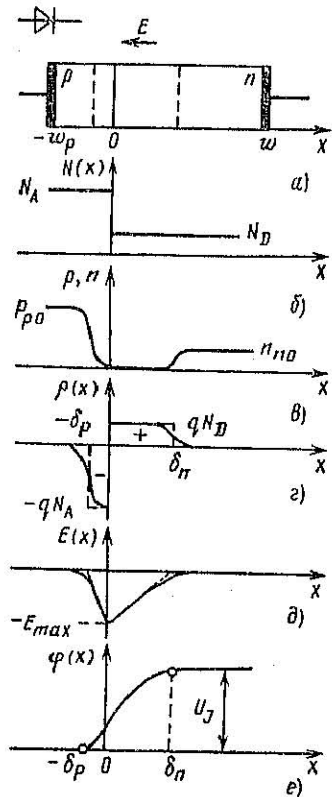


Рис. 1.11. Резкий p^+-n переход в состоянии термодинамического равновесия

1.11, а). Аналогично электроны из n -области (где их много) диффундируют в p -область (где их мало), при этом в n -области остаются нескомпенсированные положительные ионы доноров и возникает положительный объемный заряд $+qN_D$. Электроны, переходящие в p -область, рекомбинируют с дырками, что также приводит к образованию нескомпенсированного отрицательного заряда ионов акцепторов вблизи границы раздела. В результате описанного выше процесса вблизи границы раздела образуется ОПЗ, в которой концентрация электронов и дырок понижена. ОПЗ имеет высокое электрическое сопротивление, и ее называют также *запирающим слоем*. Распределение объемного заряда в ОПЗ $\rho(x)$ показано на рис. 1.11, в. Электронно-дырочные переходы с распределением легирующей примеси, изображенным на рис. 1.11, б, называют *резкими* переходами.

При разделении положительных и отрицательных зарядов в ОПЗ образуется электрическое поле, направленное против оси x , с увеличением которого увеличивается сила, препятствующая диффузионному перемещению дырок из p - в n -область и электронов из n - в p -область. В результате конкурирующих процессов диффузии и дрейфа носителей заряда между p - и n -областями перехода устанавливается некоторая равновесная разность потенциалов (рис. 1.11, е), которая носит название *контактной разности потенциалов* и связана с распределением напряженности электрического поля в ОПЗ перехода соотношением

$$U_j = - \int E(x) dx. \quad (1.63)$$

Рассмотрим зонную энергетическую диаграмму p - n перехода, представленную на рис. 1.12. В состоянии термодинамического равновесия энергия уровня Ферми постоянна в любой точке кристалла. В приконтактной области концентрации дырок и электронов меняются, меняется потенциал, что обуславливает изгиб энергетических зон. В глубине p - и n -областей, т. е. на большом удалении от границы их раздела, концентрации электронов и дырок соответствуют своим равновесным значениям. С помощью (1.46 б) можно найти связь величины $(\mathcal{E}_{ip} - \mathcal{E}_F)$ с равновесной концентрацией дырок p_{p0} в глубине p -области

$$\mathcal{E}_{ip} - \mathcal{E}_F = kT \ln(p_{p0} / n_{ip} p_p). \quad (1.64)$$

Аналогично разницу энергий $(\mathcal{E}_F - \mathcal{E}_{in})$ в глубине n -области можно выразить формулой

$$\mathcal{E}_F - \mathcal{E}_{in} = kT \ln(n_{n0} / n_{in} p_n). \quad (1.65)$$

Эффективная собственная концентрация в p -области $n_{эфp}$ может сильно отличаться от величины $n_{эфn}$ в n -области ввиду более сильного легирования p^+ -области. Более того, если сужение запрещенной зоны в p^+ -области существует, то на энергетической диаграмме уровни энергии \mathcal{E}_c и \mathcal{E}_v должны иметь скачки при $x=0$, равные $\Delta\mathcal{E}_c$ и $\Delta\mathcal{E}_v$ соответственно (рис. 1.12, б). Независимо от этого обстоятельства контактную разность потенциалов p - n перехода определим через изменение

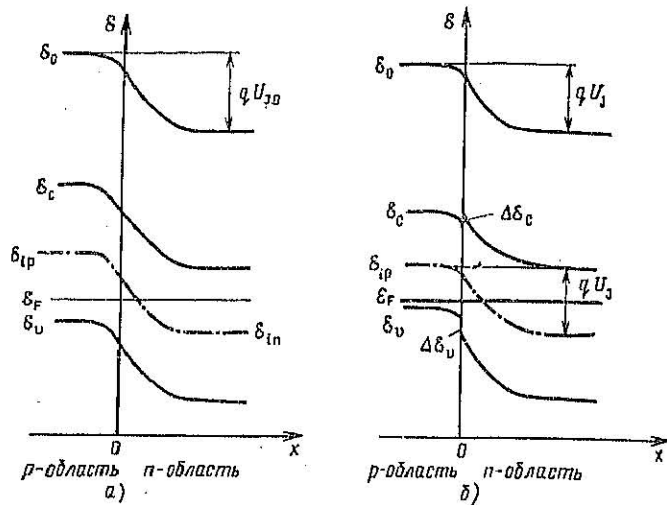


Рис. 1.12. Энергетическая зонная диаграмма p - n перехода без учета (а) и с учетом (б) сужения запрещенной зоны

уровня вакуума, или, что то же самое, через изменение уровня энергии \mathcal{E}_i следующим образом: $qU_J = \mathcal{E}_{ip} - \mathcal{E}_{in}$. Используя формулы (1.64) и (1.65), получим

$$U_J = \frac{kT}{q} \ln \frac{p_{p0} n_{n0}}{n_{эфp} n_{эфn}}. \quad (1.66)$$

В отсутствие сужения запрещенной зоны контактная разность потенциалов определяется выражением

$$U_{J0} = \frac{kT}{q} \ln (p_{p0} n_{n0} / n_i^2). \quad (1.67)$$

Следует особо подчеркнуть, что наличие эффекта СЗЗ приводит к различным значениям высоты потенциального барьера для электронов и дырок. В зоне проводимости появляется «пикок» высотой $\Delta\mathcal{E}_c$ (рис. 1.12, б). Если ширина пика в направлении оси x невелика, то электро-

ны могут свободно туннелировать сквозь него. Тогда высота потенциального барьера для электронов в зоне проводимости составит $q\phi_{ПБn} = qU_J - \Delta\mathcal{E}_c$. В валентной зоне за счет разрыва в энергии, равного $\Delta\mathcal{E}_v$, высота потенциального барьера для дырок будет больше, чем для электронов, и составит $q\phi_{ПБp} = qU_J + \Delta\mathcal{E}_v$.

Из (1.67) следует, что контактная разность потенциалов тем больше, чем сильнее легированы области перехода. При одинаковых степенях легирования U_J больше в полупроводниках с большей шириной запрещенной зоны. Для большинства германиевых p - n переходов $U_J = 0,25 \div 0,45$ В ($\mathcal{E}_G = 0,66$ эВ), для кремниевых $U_J = 0,5 \div 0,9$ В ($\mathcal{E}_G = 1,11$ эВ).

При приложении к p - n переходу внешнего напряжения U практически все оно падает на ОПЗ, так как ОПЗ обеднена носителями заряда и имеет высокое по сравнению с другими областями электрическое сопротивление. Если полярность напряжения U такова, что создаваемое им электрическое поле совпадает по направлению с внутренним электрическим полем контактной разности потенциалов, то высота потенциального барьера повысится и разность потенциалов будет равняться сумме U_J и обратного напряжения. Этому случаю соответствует такая полярность внешнего напряжения, при которой положительный полюс источника напряжения подключен к n -области, а минус — к p -области. Такое включение p - n перехода в электрическую цепь называют *обратным*.

При *прямом* направлении к p -области прикладывается положительный потенциал, а к n -области — отрицательный. Поэтому высота потенциального барьера понижается и становится примерно равной $(U_J - U)$. В дальнейшем для общности будем считать, что при любом направлении смещения высота потенциального барьера равна $(U_J - U)$, но для этого будем полагать, что обратное напряжение $U < 0$.

1.3.2. ТОЛЩИНА ОБЛАСТИ ПРОСТРАНСТВЕННОГО ЗАРЯДА И БАРЬЕРНАЯ ЕМКОСТЬ РЕЗКОГО p - n ПЕРЕХОДА

Диоды с p - n переходами обладают емкостными свойствами, что связано с пространственным перераспределением заряда электронов и дырок при приложении внешнего напряжения к p - n переходу. Емкость p - n перехода в сильной степени определяет быстродействие переходов и приборов на их основе.

В полном объеме задача о нахождении перераспреде-

ления дырок и электронов при приложении к p - n переходу внешнего напряжения сводится к решению фундаментальной системы нелинейных дифференциальных уравнений (1.35)–(1.39) и может быть решена только численно, методами математического моделирования, с помощью ЭВМ. Настоятельная потребность получения аналитических решений, пусть даже приближенных, привела к созданию физической теории, которая построена на ряде упрощающих допущений. Сущность этой теории заключается в том, что p - n переход разбивается на ряд областей, внутри каждой из которых фундаментальную систему уравнений можно заменить некоторой более простой системой, — как правило, системой линейных уравнений, допускающей аналитические решения. Такой подход называют методом региональных аппроксимаций (приближений). Непременным условием физической теории является проверка ее результатов путем сопоставления их с результатами эксперимента.

Первый шаг к региональному приближению был сделан при рассмотрении энергетической диаграммы p - n перехода (см. рис. 1.11, z), когда предполагалось, что плотность объемного заряда в p -области постоянна и равна $\rho_p = -qN_A$ при $-\delta_p < x < 0$, а в n -области $\rho_n = +qN_D$ при $0 < x < \delta_n$, где δ_p и δ_n — некоторые искусственно введенные границы ОПЗ, для которой предполагается, что она полностью обеднена электронами и дырками.

Указанные допущения позволяют для отыскания зависимостей $E(x)$ и $\varphi(x)$ внутри ОПЗ использовать только одно уравнение ФСУ — уравнение Пуассона:

$$\epsilon_r \epsilon_0 \frac{dE}{dx} = q(p - n + N(x)), \quad (1.68)$$

где $N(x) = -N_A$ в p -области и $N(x) = +N_D$ в n -области перехода, а $p \cong n \cong 0$ (точнее, $p, n \ll N_A$ в p -области и $p, n \ll N_D$ в n -области перехода).

Таким образом, надо решить, по сути дела, два уравнения Пуассона: для левой ($-\delta_p < x < 0$) и правой ($0 < x < \delta_n$) частей ОПЗ, а затем «сшить» решения, учтя непрерывность электрического поля E и потенциала φ в точке $x=0$. При этом предполагается, что электрическое поле вне ОПЗ и на ее границах отсутствует: $E(-\delta_p) = 0$, $E(\delta_n) = 0$. Первое интегрирование уравнения (1.68) с учетом принятых допущений определяет линейную зависимость $E(x)$, как показано на рис. 1.11, d . Зависимость $E(x)$ определяется выражениями

$$E(x) = -\frac{d\varphi}{dx} = -\frac{qN_A}{\epsilon_r \epsilon_0} (x + \delta_p) \quad \text{при } -\delta_p < x < 0; \quad (1.69)$$

$$E(x) = -\frac{d\varphi}{dx} = \frac{qN_D}{\epsilon_r \epsilon_0} (x - \delta_n) \quad \text{при } 0 < x < \delta_n. \quad (1.70)$$

Из условия непрерывности поля E при $x=0$ вытекает, что объемные заряды ионов акцепторов и доноров слева и справа от границ раздела p - и n -областей равны между собой:

$$Q = SqN_A \delta_p = SqN_D \delta_n, \quad (1.71)$$

где S — площадь p - n перехода.

Уровень отсчета потенциала можно выбрать произвольно. Положим $\varphi(-\delta_p) = 0$. Так как разность потенциалов на ОПЗ полагается равной $(U_J - U)$, то второе граничное условие для потенциала имеет вид $\varphi(\delta_n) = U_J - U$, где для прямого смещения перехода $U > 0$.

Интегрирование (1.69) и (1.70) дает квадратичную зависимость $\varphi(x)$ и позволяет с учетом условия непрерывности потенциала при $x=0$ найти

$$\frac{qN_A}{2\epsilon_r \epsilon_0} \delta_p^2 + \frac{qN_D}{2\epsilon_r \epsilon_0} \delta_n^2 = U_J - U. \quad (1.72)$$

Решая систему алгебраических уравнений (1.71) и (1.72), получаем

$$\delta_p = \sqrt{\frac{2\epsilon_r \epsilon_0 N_D (U_J - U)}{qN_A (N_A + N_D)}}; \quad \delta_n = \sqrt{\frac{2\epsilon_r \epsilon_0 N_A (U_J - U)}{qN_D (N_A + N_D)}}. \quad (1.73)$$

Полная толщина ОПЗ

$$\delta = \delta_p + \delta_n = \sqrt{2\epsilon_r \epsilon_0 (U_J - U) / qN^*}, \quad (1.74)$$

где для сокращения записи введено обозначение $N^* = N_A N_D / (N_A + N_D)$ — приведенная концентрация примеси.

Как следует из (1.71), ОПЗ расширяется в основном в область с меньшей концентрацией легирующей примеси. В резконецентричном p^+-n переходе при $N_A \gg N_D$ $N^* \approx N_D$ и практически вся ОПЗ располагается в слаболегируемой n -области.

Максимальная напряженность поля E_{max} (см. рис. 1.11, d) связана с напряжением U соотношением

$$E_{max} = \frac{2(U_J - U)}{\delta} = \sqrt{\frac{2qN^* (U_J - U)}{\epsilon_r \epsilon_0}}. \quad (1.75)$$

Приведенные выражения показывают, что при обратном направлении смещения p - n перехода ($U < 0$) высота потенциального барьера повышается, толщина ОПЗ и значение E_{max} увеличиваются, а при прямом ($U > 0$) уменьшаются.

Изменение толщины ОПЗ при изменении напряжения на величину ΔU связано с изменением объемного заряда нескомпенсированных ионов акцепторов вблизи левой границы ОПЗ на величину ΔQ и равного ему изменения заряда ионов доноров вблизи правой границы. Эти изменения связаны с подтеканием такого же заряда дырок к левой границе ОПЗ и электронов к правой границе ОПЗ за счет протекания тока проводимости во внешней цепи. Электронно-дырочный переход ведет себя подобно конденсатору. Емкость такого конденсатора называется *барьерной емкостью* p - n перехода, так как она связана с формированием потенциального барьера p - n перехода.

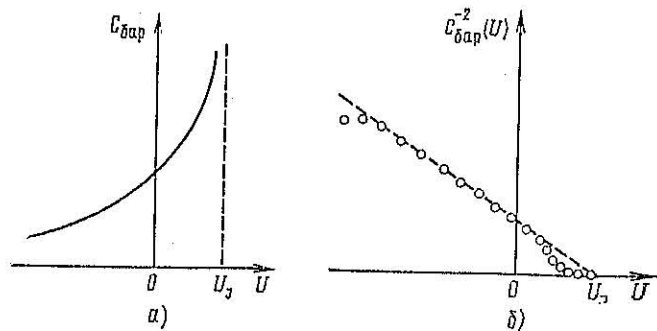


Рис. 1.13. Зависимость барьерной емкости от напряжения (а) и зависимость C^{-2} (б) (точки — эксперимент) (б)

Величину барьерной емкости можно определить следующим образом:

$$C_{\text{бар}} = \frac{dQ}{dU} = \frac{dQ}{d\delta_n} \frac{d\delta_n}{dU} = S \sqrt{\frac{q\epsilon_r \epsilon_0 N^*}{2(U_j - U)}}. \quad (1.76)$$

Отметим, что последнее выражение можно записать в виде $C_{\text{бар}} = S\epsilon_r\epsilon_0/\delta$, что совпадает с формулой для емкости плоского конденсатора, если вместо δ подставить расстояние между обкладками. Величина $C_{\text{бар}}$ нелинейно зависит от приложенного напряжения, как показано на рис. 1.13, а.

В зависимости от концентрации примеси в слаболегированной области толщина ОПЗ составляет 1–10 мкм, а значения удельной барьерной емкости p - n перехода (с площадью 1 см²) лежат в диапазоне 10^{-8} – 10^{-10} Ф/см². При

обратном напряжении $C_{\text{бар}}$ уменьшается, при прямом возрастает, что связано с зависимостью $\delta(U)$. При прямых напряжениях, близких к U_j , формулы для толщины ОПЗ и барьерной емкости дают неопределенно большую погрешность, так как основное допущение теории об обеднении ОПЗ носителями заряда перестает выполняться.

1.3.3. ТОЛЩИНА ОБЛАСТИ ПРОСТРАНСТВЕННОГО ЗАРЯДА И БАРЬЕРНАЯ ЕМКОСТЬ ПЛАВНЫХ p - n ПЕРЕХОДОВ

Большинство p - n переходов создают методом диффузии легирующей примеси в полупроводник. Допустим, что в поверхность полупроводникового кристалла n -типа электропроводности с концентрацией доноров N_D проведена диффузия акцепторной примеси, концентрация которой $N_A(x)$ уменьшается по мере удаления от поверхности полупроводника, как показано на рис. 1.14, а. Если поверхностная концентрация акцепторов $N_{A0} > N_D$, то в некотором сечении x_j кривые пересекаются. Левее этого сечения концентрация акцепторов превышает концентрацию доноров, и полупроводник имеет электропроводность p -типа, а правее x_j — электропроводность n -типа. В сечении x_j электроны атомов донорной примеси заполняют (компенсируют) ненасыщенные связи равного числа атомов акцепторной примеси, и полупроводник имеет собственную проводимость. Таким образом получается p - n переход с плавным распределением примеси, или плавный p - n переход. При рассмотрении большинства физических явлений можно не учитывать, что в некоторой области перехода полупроводник частично или полностью компенсирован, и иметь дело с *разностной (результатирующей) концентрацией* примеси, определяющей электропроводность: $N(x) = N_D - N_A(x)$ (рис. 1.14, б). Следует отметить, что отрицательные значения концентрации примеси лишены физического смысла (концентрация — это число частиц, атомов в единице объема), но можно условиться отличать концентрацию акцепторов от доноров по знаку заряда их ионов.

В плавном p - n переходе ОПЗ возникает по обе стороны от координаты x_j , называемой *металлургической границей* p - n перехода. При соблюдении некоторых условий ОПЗ формируется в интервале значений x , где реальное распределение примеси близко к линейному (рис. 1.14, б).

Рассмотрим задачу о нахождении толщины ОПЗ для p - n перехода с линейным распределением примеси: $N(x) = ax$, где a — градиент кон-