

5.2.2/3
11-19

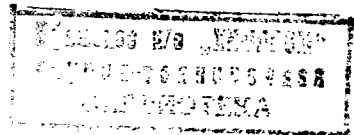
В.В. ПАСЫНКОВ
Л.К. ЧИРКИН
А.Д. ШИНКОВ

Полупроводниковые приборы

ИЗДАНИЕ ТРЕТЬЕ,
ПЕРЕРАБОТАННОЕ
И ДОПОЛНЕННОЕ

Допущено
Министерством высшего и среднего
специального образования СССР
в качестве учебника
для студентов вузов,
обучающихся по специальности
«Полупроводники и диэлектрики»

63046



Москва
«Высшая школа» 1981

ББК 32.852

П19

УДК 621.382

Рецензент — кафедра микроэлектроники
Минского радиотехнического института

Пасынков В. В. и др.

П19 Полупроводниковые приборы: Учебник для вузов. — 3-е изд., перераб. и доп. — М.: Высш. школа, 1981. — 431 с., ил.

Перед загл. авт.: Пасынков В. В., Чиркин Л. К., Шинков А. Д.

В пер.: 1 р. 30 к.

В книге изложена физика полупроводниковых приборов и элементов интегральных микросхем, рассмотрены их основные свойства, характеристики и параметры, в отличие от первых двух изданий включены вопросы конструктивно-технологических особенностей полупроводниковых приборов в интегральном исполнении и общие принципы микроэлектроники.

Предназначается для студентов специальности «Полупроводники и диэлектрики».

П $\frac{30407-132}{001(01)-81}$ 103-81

2403000000

6Ф0.32

ББК 32.852

© Издательство «Высшая школа», 1973

© Издательство «Высшая школа», 1981, с изменениями

ПРЕДИСЛОВИЕ

Содержание учебника соответствует программе, утвержденной МВ и ССО СССР. Книга может быть использована не только в качестве учебника для студентов по специальности «Полупроводники и диэлектрики», но и в качестве учебного пособия для студентов других радиотехнических и электротехнических специальностей. Поэтому в первой главе книги изложены основные сведения по физике полупроводников в объеме, необходимом для понимания принципа действия различных полупроводниковых приборов и элементов интегральных микросхем.

По сравнению со вторым изданием в учебник введены новые главы и расширены главы, в которых рассмотрены наиболее перспективные полупроводниковые приборы. В последнее время полупроводниковые приборы применяют не только как дискретные элементы, но и как элементы интегральных микросхем. Поэтому в учебнике рассмотрены конструктивно-технологические особенности полупроводниковых приборов в интегральном исполнении и общие принципы микроэлектроники. Однако вопросы технологии изготовления полупроводниковых приборов и особенно интегральных микросхем изложены только в объеме, необходимом для понимания принципа действия, свойств и характеристик полупроводниковых приборов, так как подробно изучаются в других последующих курсах.

Учебник написан на основе опыта чтения курсов «Электронные приборы и микроэлектроника», «Физика полупроводниковых приборов» и аналогичных курсов близких названий в Ленинградском ордена Ленина электротехническом институте им. В. И. Ульянова (Ленина).

Авторы весьма признательны рецензентам настоящего учебника — коллективу кафедры микроэлектроники Минского радиотехнического института, возглавляемой чл.-корр. АН БССР В. А. Лабуновым, замечания и советы которых оказали существенное влияние на построение и содержание учебника.

Приносим глубокую благодарность сотрудникам кафедры диэлектриков и полупроводников Ленинградского ордена Ленина электротехнического института им. В. И. Ульянова (Ленина) за детальное обсуждение рукописи учебника.

Замечания и пожелания просим направлять по адресу: 103051, Москва, Неглинная ул., д. 29/14, издательство «Высшая школа».

Авторы

ВВЕДЕНИЕ

В «Основных направлениях экономического и социального развития СССР на 1981—1985 годы и на период до 1990 года» ставится задача обеспечить поступательный рост экономики, настойчиво повышать эффективность производства на основе его всесторонней интенсификации и обеспечить дальнейшее ускорение научно-технического прогресса.

Научно-технический прогресс немыслим без электроники. Интенсивное развитие электроники связано с появлением новых разнообразных полупроводниковых приборов и интегральных микросхем, которые находят широкое применение в вычислительной технике, автоматике, радиотехнике и телевидении, в установках измерительной техники, медицины, биологии и т. д.

Полупроводниковые приборы в виде точечных диодов, или, как их раньше называли, кристаллических детекторов, применяли еще в первых электронных установках. Выпрямительные свойства контактов между металлами и некоторыми сернистыми соединениями были обнаружены в 1874 г. А. С. Поповым в 1895 г. при изобретении радио был применен порошок когерер, в котором использовались нелинейные свойства зернистых систем. В 1922 г. О. В. Лосев использовал отрицательное дифференциальное сопротивление, возникающее при определенных условиях на точечных контактах металла с полупроводником, для генерации и усиления высокочастотных электромагнитных колебаний. Кроме того, им было обнаружено свечение кристаллов карбида кремния при прохождении тока через точечный контакт.

Однако в этот период успешно развивается техника электровакуумных приборов и из-за недостаточного знания строения полупроводников и происходящих в них электрофизических процессов полупроводниковые приборы тогда не получили существенного развития и применения.

В годы Великой Отечественной войны были разработаны точечные высокочастотные и сверхвысокочастотные германиевые и кремниевые диоды. В 1942 г. в СССР был начат выпуск полупроводниковых термоэлектрических генераторов для непосредственного преобразования тепловой энергии в электрическую. Термогенераторы использовались для питания переносных радиостанций партизанских отрядов. Создание и производство этих и многих других приборов стало возможно благодаря фундаментальным теоретическим и экспериментальным исследованиям свойств полупровод-

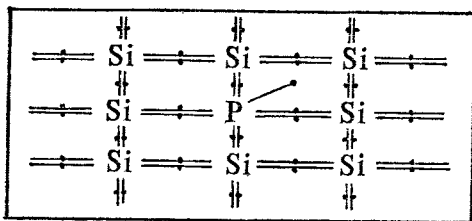
ников, проведенным группой ученых под руководством академика А. Ф. Иоффе.

С 1948 г. с момента создания американскими учеными Дж. Бардиным, В. Браттейном и В. Шокли точечного транзистора начался новый этап полупроводниковой электроники. В пятидесятых годах были разработаны различные типы транзисторов, мощных германиевых и кремниевых выпрямительных диодов, тиристоров, фотодиодов, фототранзисторов, кремниевых фотоэлементов, туннельных диодов и других полупроводниковых приборов. В это же время появились идеи создания полевых транзисторов с управляющим электронно-дырочным переходом и с изолированным затвором (МДП-транзисторов). Но до окончательной разработки полевых транзисторов с изолированным затвором потребовались еще многолетние исследования электрофизических процессов на границе полупроводника с диэлектриком и разработка технологических процессов создания необходимых структур.

Современный этап развития полупроводниковой электроники характеризуется в нашей стране большим объемом научно-исследовательских и технологических работ, направленных на дальнейшее совершенствование имеющихся и создание новых полупроводниковых приборов.

ОСНОВНЫЕ СВЕДЕНИЯ ПО ФИЗИКЕ ПОЛУПРОВОДНИКОВ

1



§ 1.1. ЭНЕРГЕТИЧЕСКИЕ ЗОНЫ ПОЛУПРОВОДНИКОВ

Каждый электрон, входящий в состав атома, обладает определенной полной энергией или занимает определенный энергетический уровень.

В твердом теле благодаря взаимодействию атомов энергетические уровни расщепляются и образуют *энергетические зоны*, состоящие из отдельных близко расположенных по энергии уровней, число которых соответствует числу однородных атомов в данном кристаллическом теле (рис. 1.1). Энергетическую зону или совокупность нескольких перекрывающихся энергетических зон, которые образовались в результате расщепления одного или нескольких энергетических уровней отдельного атома, называют *разрешенной зоной*. Электроны в твердом теле могут иметь энергии, соответствующие разрешенной зоне. Верхний энергетический уровень разрешенной зоны называют *потолком*, нижний — *дном*.

Энергетические уровни валентных электронов при расщеплении образуют *валентную зону*. Разрешенные энергетические уровни, свободные от электронов в невозбужденном состоянии атома, расщепляясь, образуют одну или несколько *свободных зон*. Нижнюю из свободных зон называют *зоной проводимости*.

Наибольший интерес представляют валентная зона и зона проводимости, так как от их взаимного расположения и от степени их заполнения электронами зависят электрические, оптические и другие свойства твердых тел. Между разрешенными зонами находятся *запрещенные зоны*, т. е. области значений энергий, которыми не могут обладать электроны в идеальном кристалле. Для полупроводников (согласно сказанному) наибольшее значение имеет запрещенная зона, разделяющая валентную зону и зону проводимости. Она характеризуется *шириной запрещенной зоны* ΔE , т. е. разностью энергий дна зоны проводимости и потолка валентной зоны.

При температуре 300 К у кремния ширина запрещенной зоны $\Delta\mathcal{E} = 1,12$ эВ; у германия $\Delta\mathcal{E} = 0,75$ эВ; у арсенида галлия $\Delta\mathcal{E} = 1,43$ эВ; у карбида кремния $\Delta\mathcal{E} = 2,4 - 3,4$ эВ (для разных политипов).

Ширина запрещенной зоны изменяется с изменением температуры. Происходит это в результате: 1) изменения амплитуды тепловых колебаний атомов кристаллической решетки; 2) изменений межатомных расстояний, т. е. объема тела. С повышением температуры в первом случае ширина запрещенной зоны уменьшается, во втором случае может быть как уменьшение, так и увеличение ширины запрещенной зоны. У большинства полупроводников ширина запрещенной зоны с повышением температуры уменьшается.

При наличии в полупроводнике электрического поля энергетические диаграммы целесообразно строить, откладывая по вертикальной оси полную энергию электронов $\mathcal{E} - q\varphi$ (с учетом потенциальной энергии электрона в электрическом поле), а по горизонтальной оси — геометрическую координату (рис. 1.2). При таком построении энергетических диаграмм в областях, где существует электрическое поле, энергетические уровни и зоны получают наклонными. Причем угловой коэффициент пропорционален напряженности электрического поля, а относительное смещение энергетических уровней или зон — разности потенциалов между данными точками объема полупроводника. Коэффициент пропорциональности при этом равен элементарному заряду электрона q ; увеличению потенциала соответствует понижение энергетических уровней или зон на энергетической диаграмме.

§ 1.2. ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Образование свободных электронов и дырок — генерация носителей заряда — происходит при воздействии теплового хаотического движения атомов кристаллической решетки (*тепловая генерация*), при воздействии поглощенных полупровод-

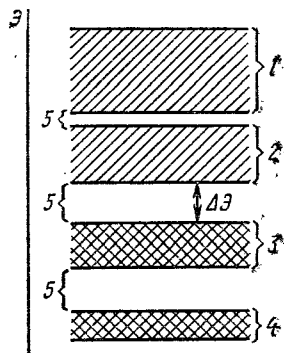


Рис. 1.1. Энергетические зоны полупроводника:

1, 2, 3, 4 — разрешенные зоны; 5 — запрещенные зоны; 1, 2 — свободные зоны; 2 — зона проводимости; 3 — валентная зона; $\Delta\mathcal{E}$ — ширина запрещенной зоны

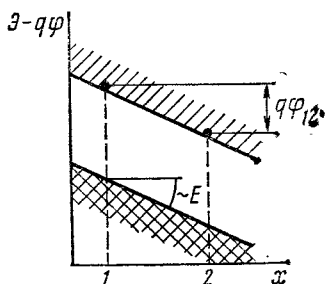


Рис. 1.2. Энергетическая диаграмма полупроводника при наличии в нем электрического поля напряженностью E и разности потенциалов между точками 1 и 2, равной φ_{12}

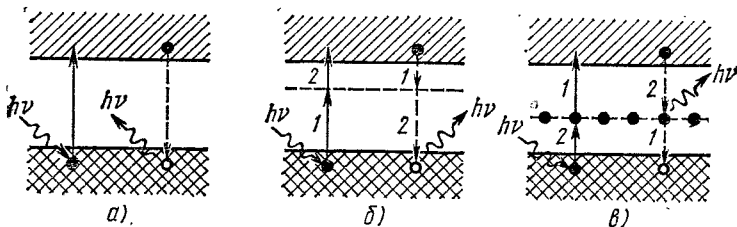


Рис. 1.3. Различные механизмы генерации и рекомбинации носителей заряда:

а — межзонная генерация и рекомбинация; *б* — генерация и рекомбинация с участием пустых рекомбинационных ловушек; *в* — генерация и рекомбинация с участием заполненных электронами рекомбинационных ловушек; — — — — путь генерации; — — — — путь рекомбинации. Цифры означают этапы процессов генерации и рекомбинации

ником квантов света (*световая генерация*) и других энергетических факторов. Так как полупроводник всегда находится под действием всех этих факторов или хотя бы одного ($T \neq 0$), генерация носителей происходит непрерывно.

Одновременно с генерацией в полупроводнике непрерывно происходит и обратный процесс — **рекомбинация носителей заряда**, т. е. возвращение электронов из зоны проводимости в валентную зону, в результате чего исчезает пара носителей заряда. В состоянии термодинамического равновесия процессы генерации и рекомбинации носителей заряда взаимно уравновешены. При этом в полупроводнике существует *равновесная концентрация электронов n_0* и *равновесная концентрация дырок p_0* .

При воздействии на полупроводник нетеплового внешнего энергетического фактора (света, сильного электрического поля и др.) из-за генерации новых носителей заряда их концентрация n и p (*неравновесная концентрация*) будет превышать равновесную концентрацию на величину Δn (или Δp), которую называют *избыточной концентрацией*. Таким образом,

$$\Delta n = n - n_0; \quad \Delta p = p - p_0. \quad (1.1)$$

Избыточная концентрация носителей заряда может быть в отдельных областях полупроводниковой структуры прибора не только в результате внешних энергетических воздействий, но и за счет различных процессов (инжекции, экстракции, аккумуляции и т. д.), которые могут происходить в полупроводниковых приборах (о чем говорится далее).

Механизмы рекомбинации могут быть различны (рис. 1.3).

Межзонная, или непосредственная, рекомбинация происходит при переходе свободного электрона из зоны проводимости в валентную зону на один из свободных энергетических уровней, что соответствует исчезновению пары носителей заряда — свободного электрона и дырки. Однако такой процесс межзонной рекомбинации мало вероятен, так как свободный электрон и дырка должны оказаться одновременно в одном и том же месте кристалла. Кроме того, должен выполняться закон сохранения импульса, т. е. реком-

бинация электрона и дырки возможна только при одинаковых, но противоположно направленных импульсах электрона и дырки. Поэтому, например, в германии на 10 тыс. рекомбинаций лишь одна происходит в результате межзонной рекомбинации.

Рекомбинация с участием рекомбинационных ловушек протекает в два этапа. На первом этапе рекомбинационная ловушка (или энергетический уровень рекомбинационной ловушки) захватывает, например, электрон из зоны проводимости. Таким образом, электрон выбывает из процесса электропроводности. В этом состоянии ловушка будет находиться до тех пор, пока к ней не подойдет дырка, или, другими словами, пока в данном месте кристалла не окажется свободный энергетический уровень валентной зоны. При выполнении этих условий осуществляется второй этап рекомбинации — электрон перейдет на свободный уровень валентной зоны (что эквивалентно захвату дырки из валентной зоны отрицательно заряженной ловушкой).

Двухэтапный процесс рекомбинации более вероятен, так как он не требует одновременного присутствия в данном месте кристалла свободного электрона и дырки. Рекомбинационная ловушка воспринимает количество движения, необходимое для соблюдения закона сохранения импульса, и может забрать часть энергии, освобождаемой в процессе рекомбинации.

Роль рекомбинационных ловушек могут выполнять примесные атомы или ионы, различные включения в кристалле, незаполненные узлы кристаллической решетки, трещины и другие несовершенства объема или поверхности.

В связи с тем, что на поверхности кристалла перечисленных дефектов значительно больше, чем в объеме, процесс рекомбинации на поверхности должен идти значительно интенсивнее. Его рассматривают и оценивают обычно отдельно, считая *поверхностную рекомбинацию* разновидностью рекомбинации с участием рекомбинационных ловушек.

В зависимости от того, как расходуется энергия, освобождающаяся при рекомбинации электрона и дырки, рекомбинацию можно подразделить на два вида.

Излучательной рекомбинацией называют рекомбинацию, при которой энергия, освобождающаяся при переходе электрона на более низкий энергетический уровень, излучается в виде кванта света (фотона).

При *безизлучательной (фононной) рекомбинации* избыточная энергия электрона передается кристаллической решетке полупроводника, т. е. избыточная энергия идет на образование фононов — квантов тепловой энергии.

§ 1.3. КОНЦЕНТРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКЕ ПРИ ТЕРМОДИНАМИЧЕСКОМ РАВНОВЕСИИ

В соответствии со статистикой Ферми—Дирака вероятность заполнения энергетического уровня электроном определяется энергией \mathcal{E} , соответствующей этому уровню, и абсолютной тем-

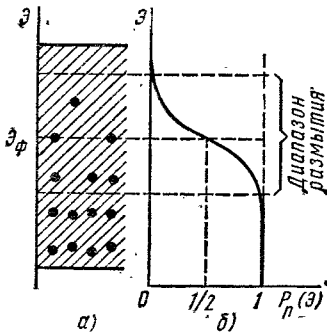


Рис. 1.4. Распределение электронов в частично заполненной зоне (а) и функция вероятности заполнения энергетических уровней (б)

пературой T :

$$P_n(\varepsilon) = \frac{1}{1 + \exp \frac{\varepsilon - \varepsilon_F}{kT}}, \quad (1.2)$$

где ε_F — энергия уровня Ферми, вероятность заполнения которого равна $1/2$ и относительно которого кривая вероятности симметрична (рис. 1.4).

Энергия уровня Ферми соответствует верхней границе электронного распределения при температуре $T = 0$, а также средней энергии «диапазона размытия» при любой другой температуре (рис. 1.4). Симметрия кривой вероятности заполнения относительно уровня Ферми

означает одинаковую вероятность заполнения уровня электроном с энергией, большей на величину $\varepsilon - \varepsilon_F$, и вероятность освобождения уровня от электрона с энергией, на столько же меньшей энергии уровня Ферми.

С помощью соотношения (1.2) можно определять заполнение электронами зоны проводимости или валентной зоны полупроводника. Но для валентной зоны удобнее говорить о дырках — пустых энергетических уровнях в валентной зоне. Любой энергетический уровень может либо быть занят электроном, либо свободен от электрона. Поэтому сумма вероятностей этих двух событий должна быть равна единице:

$$P_n(\varepsilon) + P_p(\varepsilon) = 1.$$

Тогда вероятность заполнения энергетического уровня дыркой

$$P_p(\varepsilon) = \frac{1}{1 + \exp \left(-\frac{\varepsilon - \varepsilon_F}{kT} \right)}. \quad (1.3)$$

В большинстве полупроводников уровень Ферми расположен в запрещенной зоне на больших расстояниях в единицах энергии от зоны проводимости и от валентной зоны по сравнению с энергией kT (при комнатной температуре $kT \approx 0,025$ эВ), т. е.

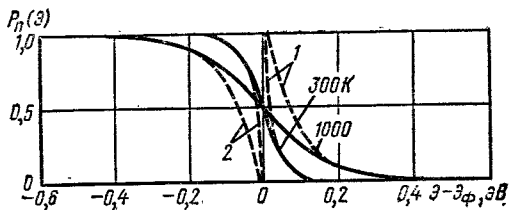
$$|\varepsilon - \varepsilon_F| \gg kT. \quad (1.4)$$

Поэтому, пренебрегая единицей в знаменателе (1.2), вероятность распределения электронов по энергетическим уровням зоны проводимости можно определить с помощью статистики Максвелла—Больцмана:

$$P_n(\varepsilon) \approx \exp \left(-\frac{\varepsilon - \varepsilon_F}{kT} \right). \quad (1.5)$$

Рис. 1.5. Вероятность заполнения электронами энергетических уровней при различных температурах:

— — по статистике Ферми — Дирака; — — — — по статистике Максвелла — Больцмана для электронов в зоне проводимости (1) и в валентной зоне (2)



Аналогично вероятность распределения дырок по энергетическим уровням валентной зоны с учетом (1.3) и (1.4)

$$P_p(\varepsilon) \approx \exp \frac{\varepsilon - \varepsilon_{\Phi}}{kT}. \quad (1.6)$$

Таким образом, для большинства полупроводников (*невырожденных*) можно пользоваться статистикой Максвелла—Больцмана и только в некоторых случаях для полупроводников (*вырожденных*) необходимо использовать статистику Ферми—Дирака. Разница в этих двух функциях распределения электронов по энергиям показана на рис. 1.5.

Для определения концентрации электронов в невырожденном полупроводнике надо проинтегрировать по энергии произведение удвоенной функции распределения плотности энергетических уровней в зоне проводимости $[2N(\varepsilon)]$ и вероятности заполнения этих уровней электронами (1.5). Интегрирование нужно проводить от энергии дна до энергии потолка зоны проводимости. Если же учесть ничтожно малую вероятность заполнения уровней электронами у потолка зоны проводимости, то можно верхний предел интегрирования считать равным бесконечности, т. е.

$$n_0 = \int_{\varepsilon_c}^{\infty} 2N(\varepsilon) \cdot P_n(\varepsilon) d\varepsilon.$$

В результате интегрирования получается

$$n_0 = 2N_c \exp \left(- \frac{\varepsilon_c - \varepsilon_{\Phi}}{kT} \right), \quad (1.7)$$

где $2N_c$ — эффективная плотность состояний в зоне проводимости; N_c — эффективная плотность энергетических уровней в зоне проводимости, энергия которых приведена ко дну зоны проводимости.

Эффективная плотность состояний отличается в два раза от эффективной плотности уровней в связи с тем, что на каждом энергетическом уровне могут находиться два электрона с различными спинами (по принципу Паули).

Аналогично равновесная концентрация дырок в любом невырожденном полупроводнике при термодинамическом равновесии

$$p_0 = 2N_v \exp \frac{\varepsilon_v - \varepsilon_{\Phi}}{kT}, \quad (1.8)$$

где N_v — эффективная плотность энергетических уровней в валентной зоне, энергия которых приведена к потолку валентной зоны; \mathcal{E}_v — энергия потолка валентной зоны.

Таким образом, равновесная концентрация носителей заряда определяется эффективной плотностью уровней в соответствующей зоне, температурой и положением уровня Ферми относительно дна зоны проводимости или потолка валентной зоны.

§ 1.4. СОБСТВЕННЫЕ ПОЛУПРОВОДНИКИ

Собственный полупроводник — это полупроводник без донорных и акцепторных примесей или с концентрацией примеси настолько малой, что она не оказывает существенного влияния на удельную проводимость полупроводника. В собственном полупроводнике при температуре абсолютного нуля отсутствуют носители заряда, так как валентная зона полностью занята электронами (там нет дырок), а в зоне проводимости нет электронов.

При температурах выше абсолютного нуля некоторые электроны валентной зоны могут быть переброшены в зону проводимости — возможна тепловая генерация пар носителей заряда, в зоне проводимости появляются свободные электроны, а в валентной зоне — дырки. Процесс тепловой генерации возможен даже при очень низких температурах из-за значительных флуктуаций (отклонений) энергий тепловых колебаний атомов от средней энергии тепловых колебаний атомов относительно узлов кристаллической решетки.

Как уже отмечалось, кроме тепловой генерации носителей заряда в полупроводнике существует и их рекомбинация, и эти процессы при любой температуре взаимно уравновешены. При этом в собственном полупроводнике существует *собственная концентрация электронов*, которая может быть выражена соотношением, аналогичным (1.7):

$$n_i = \sqrt{4N_c N_v} \exp\left(-\frac{\mathcal{E}_c - \mathcal{E}_F}{kT}\right) \approx 2N_c \exp\left(-\frac{\mathcal{E}_c - \mathcal{E}_F}{kT}\right). \quad (1.9)$$

Собственная концентрация дырок аналогична (1.8):

$$p_i = \sqrt{4N_c N_v} \exp\frac{\mathcal{E}_v - \mathcal{E}_F}{kT} \approx 2N_v \exp\frac{\mathcal{E}_v - \mathcal{E}_F}{kT}. \quad (1.10)$$

Эффективные плотности уровней в зоне проводимости и в валентной зоне определяются соотношениями

$$n \quad \left. \begin{aligned} N_c &= \left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \\ N_v &= \left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2} \end{aligned} \right\} \quad (1.11)$$

где m_n^* и m_p^* — эффективные массы электронов и дырок (смысл этих параметров будет рассмотрен далее); h — постоянная Планка.

Так как в собственном полупроводнике носители заряда образуются за счет ионизации собственных атомов полупроводника,

т. е. благодаря перебросу электронов из валентной зоны в зону проводимости, при этом одновременно возникает два носителя противоположных знаков. Поэтому

$$n_i = p_i. \quad (1.12)$$

Определим положение уровня Ферми для собственного полупроводника. Учитывая соотношения (1.9)—(1.12), можно получить

$$(m_n^*)^{3/2} \exp\left(-\frac{\mathcal{E}_c - \mathcal{E}_\Phi}{kT}\right) = (m_p^*)^{3/2} \exp\frac{\mathcal{E}_v - \mathcal{E}_\Phi}{kT}.$$

Отсюда

$$\mathcal{E}_\Phi = \frac{\mathcal{E}_c + \mathcal{E}_v}{2} + \frac{3}{4} kT \ln \frac{m_p^*}{m_n^*}. \quad (1.13)$$

При небольшом различии эффективных масс электронов и дырок

$$\mathcal{E}_\Phi \approx \frac{\mathcal{E}_c + \mathcal{E}_v}{2} = \mathcal{E}_i, \quad (1.14)$$

т. е. в собственном полупроводнике приближенно можно считать, что уровень Ферми расположен посередине запрещенной зоны.

Приняв в данном случае для простоты начало отсчета энергии от потолка валентной зоны, т. е. $\mathcal{E}_v = 0$, собственные концентрации носителей можно выразить через ширину запрещенной зоны $\Delta\mathcal{E} = \mathcal{E}_c - \mathcal{E}_v$:

$$n_i = 2N_c \exp\left(-\frac{\Delta\mathcal{E}}{2kT}\right); \quad (1.15)$$

$$p_i = 2N_v \exp\left(-\frac{\Delta\mathcal{E}}{2kT}\right). \quad (1.16)$$

Из соотношений (1.15) и (1.16) можно определить собственные концентрации носителей в разных полупроводниках. При комнатной температуре ($T \approx 300$ К) в кремнии $n_i \approx 10^{10}$ см⁻³, в германии $n_i \approx 10^{13}$ см⁻³.

§ 1.5. ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ

Большинство полупроводниковых приборов изготавливают на основе примесных полупроводников, т. е. полупроводников, электрические характеристики которых определяются донорными или акцепторными примесями. Таким образом, в рабочем диапазоне температур полупроводникового прибора поставщиками основного количества носителей заряда в полупроводниковом материале должны быть примеси. Поэтому в практике важное значение имеют такие полупроводниковые материалы, у которых ощутимая собственная концентрация носителей заряда появляется при возможно более высокой температуре, т. е. полупроводники с достаточно большой шириной запрещенной зоны.

В *простых полупроводниках*, состоящих в основном из атомов одного химического элемента, примесями являются чужеродные атомы. В *сложных полупроводниках*, состоящих из атомов двух или

большого числа химических элементов, примесями могут быть не только включения атомов посторонних химических элементов, но и избыточные по отношению к стехиометрическому составу атомы химических элементов, входящих в химическую формулу сложного полупроводника.

Такую же роль, как примеси, могут играть различные дефекты кристаллической решетки: пустые узлы, дислокации или сдвиги, возникающие при пластической деформации кристалла, и т. д.

Рассмотрим роль примесей, атомы которых создают дискретные энергетические уровни в пределах запрещенной зоны полупроводника. При небольшой концентрации примесей их атомы расположены в полупроводнике на таких больших расстояниях друг от друга, что не взаимодействуют между собой. Поэтому нет расщепления примесных уровней. Вероятность непосредственного перехода электронов от одного примесного атома к другому ничтожно мала, т. е. с точки зрения зонной теории ничтожно мала вероятность перехода электрона с одного дискретного примесного уровня на другой.

При большой концентрации примесей в результате взаимодействия примесных атомов между собой примесные уровни одного типа расщепляются в энергетическую *примесную зону*. Электроны, находящиеся в примесной зоне, так же, как в зоне проводимости и в валентной зоне при неполном их заполнении, могут переходить с уровня на уровень, приобретая необходимую для этого энергию за счет ускорения во внешнем электрическом поле на длине свободного пробега.

Примеси могут быть донорного и акцепторного типа.

Д о н о р — это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, занятый в невозбужденном состоянии электроном и способный в возбужденном состоянии отдать электрон в зону проводимости.

А к ц е п т о р — это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, свободный от электрона в невозбужденном состоянии и способный захватить электрон из валентной зоны в возбужденном состоянии.

Доноры и акцепторы в полупроводнике могут быть ионизированы под действием энергии, поступающей в кристалл в виде квантов света, тепла и т. д.

Под *энергией ионизации донора* понимают минимальную энергию, которую необходимо сообщить электрону, находящемуся на донорном уровне, чтобы перевести его в зону проводимости.

Энергия ионизации акцептора — это минимальная энергия, которую необходимо сообщить электрону валентной зоны, чтобы перевести его на акцепторный уровень.

Энергия ионизации примесных атомов значительно меньше энергии ионизации собственных атомов полупроводника или ширины запрещенной зоны. Поэтому в примесных полупроводниках при низких температурах преобладают носители заряда, возникшие из-за ионизации примесей. Если электропроводность полупровод-

ника обусловлена электронами, его называют *полупроводником n-типа*, если электропроводность обусловлена дырками, — *полупроводником p-типа*.

Обычно в полупроводниках присутствуют как доноры, так и акцепторы. Полупроводник, у которого концентрация доноров равна концентрации акцепторов, называют *скомпенсированным*.

Соотношения для концентрации электронов (1.7) и дырок (1.8) можно преобразовать, приняв $\mathcal{E}_в = 0$ и учтя соотношения (1.15) и (1.16):

$$n_0 = 2N_c \exp\left(-\frac{\mathcal{E}_i + \mathcal{E}_i - \mathcal{E}_\Phi}{kT}\right) = n_i \exp\left(-\frac{\mathcal{E}_i - \mathcal{E}_\Phi}{kT}\right); \quad (1.17)$$

$$p_0 = 2N_v \exp\frac{\mathcal{E}_в - \mathcal{E}_\Phi - \mathcal{E}_i + \mathcal{E}_i}{kT} = p_i \exp\frac{\mathcal{E}_i - \mathcal{E}_\Phi}{kT}. \quad (1.18)$$

Отсюда следует, что

$$n_0 p_0 = n_i^2, \quad (1.19)$$

т. е. в невырожденном полупроводнике произведение концентраций свободных электронов и дырок при термодинамическом равновесии есть постоянная величина, равная квадрату собственной концентрации при данной температуре.

Соотношение (1.19) справедливо для невырожденного полупроводника, т. е. для полупроводника, у которого уровень Ферми расположен в запрещенной зоне достаточно далеко (на 2—3 kT) от дна зоны проводимости или от потолка валентной зоны, так как только при этих условиях можно пользоваться функцией распределения Максвелла—Больцмана.

§ 1.6. ВРЕМЯ ЖИЗНИ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

После прекращения энергетического воздействия на полупроводник избыточная концентрация носителей заряда в нем из-за процесса рекомбинации через некоторое время уменьшится до нуля. Количество носителей заряда, рекомбинирующих в единицу времени в единице объема (быстрота изменения концентрации), пропорционально избыточной концентрации и обратно пропорционально некоторому параметру τ , который называют временем жизни:

$$-\frac{dn}{dt} = -\frac{d(\Delta n)}{dt} = \frac{\Delta n}{\tau_n}, \quad -\frac{dp}{dt} = -\frac{d(\Delta p)}{dt} = \frac{\Delta p}{\tau_p}. \quad (1.20)$$

Таким образом, временем жизни неравновесных носителей заряда является отношение избыточной концентрации (Δn или Δp) неравновесных носителей заряда к скорости изменения этой концентрации вследствие рекомбинации:

$$\tau_n = \frac{|\Delta n|}{|d(\Delta n)/dt|} \quad \text{и} \quad \tau_p = \frac{|\Delta p|}{|d(\Delta p)/dt|}. \quad (1.21)$$

Рассмотрим наиболее простой частный случай — постоянство концентрации носителей заряда, с которыми происходит рекомби-

нация неравновесных носителей. Этот случай реализуется, например, в полупроводнике с явно выраженной примесной электропроводностью при введении в него неосновных носителей заряда в небольшом количестве. Тогда появление неравновесных неосновных носителей заряда не вызывает существенного изменения концентрации основных, с которыми происходит рекомбинация. Время жизни при этом оказывается постоянным, а количество носителей заряда, рекомбинировавших в единицу времени в единице объема, пропорционально первой степени избыточной концентрации. Этот случай называют *линейной рекомбинацией*. Уравнения (1.20) решаются очень просто.

Пусть в момент времени $t = 0$ в полупроводнике создана избыточная концентрация $\Delta n(0)$ (или $\Delta p(0)$), которая после окончания действия источника избыточных носителей заряда должна стремиться к нулю при $t \rightarrow \infty$. Решение уравнений (1.20) при этих условиях:

$$\Delta n(t) = \Delta n(0) \exp(-t/\tau_n) \quad \text{и} \quad \Delta p(t) = \Delta p(0) \exp(-t/\tau_p),$$

т. е. при линейной рекомбинации избыточная концентрация носителей за время жизни уменьшается в e раз.

Время жизни неравновесных носителей заряда зависит от температуры полупроводника. Рассмотрим температурную зависимость времени жизни на примере полупроводника с электропроводностью n -типа с рекомбинационными ловушками в верхней половине запрещенной зоны (рис. 1.6). При очень низкой температуре рекомбинационные ловушки заполнены электронами, так как вероятность их заполнения, судя по положению уровня Ферми, намного больше 50 %. При этом первый этап рекомбинации (захват дырки рекомбинационной ловушкой) происходит быстро, и время жизни оказывается небольшим. С повышением температуры уровень Ферми смещается вниз и находится вблизи энергетических уровней рекомбинационных ловушек. Это означает, что теперь не все ловушки заполнены электронами, т. е. не все ловушки могут захватить блуждающие по полупроводнику дырки. Поэтому с повышением температуры время жизни растет.

Необходимо отметить, что рассмотренная температурная зависимость времени жизни справедлива только для полупроводника с рекомбинационными ловушками одного сорта или типа. Если же в полупроводнике будут рекомбинационные ловушки разных типов, создающие в запрещенной зоне несколько различных энергетических уровней, то температурная зависимость времени жизни может быть сложнее представленной на рис. 1.6.

Время жизни носителей заряда в значительной степени зависит от концентрации в полупроводнике рекомбинационных ловушек, так как рекомбинация чаще всего происходит с их помощью. Но кроме рекомбинационных ловушек в запрещенной зоне полупроводника обычно существуют энергетические уровни, которые могут захватывать носители только какого-либо одного типа. Такие уровни называют **уровнями ловушек захвата**,

а дефекты кристаллической решетки, создающие уровни ловушек захвата, — ловушками захвата.

Энергетические уровни ловушек захвата электронов расположены в запрещенной зоне вблизи зоны проводимости, ловушек захвата дырок — вблизи валентной зоны (рис. 1.7).

Через некоторое время после захвата носителя заряда ловушкой захвата может произойти ионизация этой ловушки, т. е. освобождение носителя заряда. Если время нахождения носителя в ловушке захвата велико или велика концентрация ловушек захвата, то эффективное время жизни носителя заряда может оказаться значительно больше действительного времени жизни, так как находящийся в ловушке захвата носитель не может в это время рекомбинировать. Рекомбинация возможна только после ионизации ловушки захвата или после освобождения носителя заряда.

Таким образом, в запрещенной зоне энергетической диаграммы полупроводника может существовать много различных локальных энергетических уровней, связанных с наличием разных примесей. Часть из них может быть уровнями ловушек захвата, часть — уровнями рекомбинационных ловушек. При различной степени отклонения от термодинамического равновесия роль, выполняемая отдельными ловушками, может изменяться, т. е. ловушки захвата могут стать рекомбинационными ловушками и наоборот. Чтобы установить количественный критерий отличия этих энергетических уровней, введены понятия демаркационных уровней, для которых вероятность ионизации с образованием носителя одного знака равна вероятности захвата носителя заряда противоположного знака.

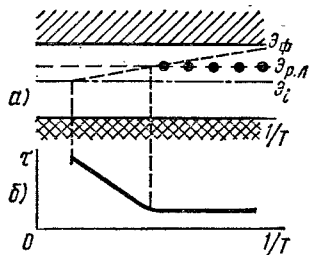


Рис. 1.6. Пояснение температурной зависимости времени жизни:

а — зависимость положения уровня Ферми от температуры и заполнение уровней рекомбинационных ловушек электронами; б — зависимость времени жизни от температуры

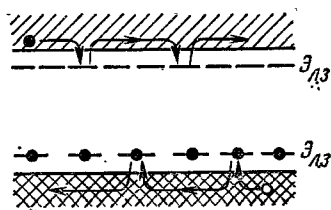


Рис. 1.7. Расположение энергетических уровней ловушек захвата и схематическое изображение процесса захвата электронов и дырок

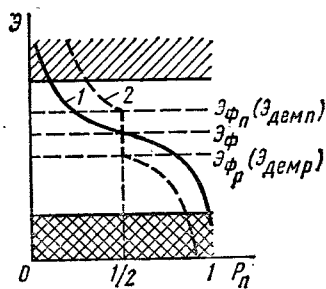


Рис. 1.8. Вероятность заполнения энергетических уровней в состоянии термодинамического равновесия (кривая 1) и в неравновесном состоянии (кривая 2) и расположение квазиуровней Ферми или демаркационных уровней

63046

В неравновесном состоянии при искусственном изменении концентрации носителей заряда распределение свободных электронов и дырок уже не соответствует распределению этих носителей заряда при термодинамическом равновесии. Поэтому в неравновесном состоянии распределение свободных электронов по энергетическим уровням характеризуется своим *квазиуровнем Ферми для электронов*, распределение дырок — своим *квазиуровнем Ферми для дырок*, которые имеют тот же смысл для полупроводника в неравновесном состоянии, что и уровень Ферми в условиях термодинамического равновесия. Чем больше неравновесные концентрации свободных электронов и дырок отклоняются от своих равновесных значений, тем больше отличается положение квазиуровней Ферми для электронов и для дырок от положения уровня Ферми в условиях термодинамического равновесия.

При качественном рассмотрении различных зависимостей параметров полупроводниковых приборов можно в первом приближении считать демаркационные уровни совпадающими с соответствующими квазиуровнями Ферми (рис. 1.8). Уровни, лежащие выше электронного демаркационного уровня или выше квазиуровня Ферми для электронов, являются уровнями захвата электронов. Вероятность их заполнения электронами меньше 50 %, что соответствует большей вероятности (более 50 %) переброса электронов с этих уровней в зону проводимости в результате тепловой генерации. Аналогично уровни, лежащие ниже дырочного демаркационного уровня или квазиуровня Ферми для дырок, являются уровнями захвата дырок.

Для уровней, расположенных между электронным и дырочным демаркационными уровнями или между квазиуровнями Ферми, характерна большая вероятность заполнения с одной стороны электронами, а с другой стороны дырками. В действительности сумма вероятностей заполнения какого-либо уровня электроном и дыркой должна быть равна единице. Поэтому следует считать, что заполнение электронами и дырками всех уровней, расположенных между демаркационными уровнями $\mathcal{E}_{\text{дем } n}$ и $\mathcal{E}_{\text{дем } p}$, одинаково. В связи с этим уровни, расположенные между электронным и дырочным демаркационными уровнями, следует считать уровнями рекомбинационных ловушек.

§ 1.7. ПРОЦЕССЫ ЭЛЕКТРОПРОВОДНОСТИ В ПОЛУПРОВОДНИКАХ

Процесс переноса зарядов — процесс электропроводности может наблюдаться в полупроводниках при наличии электронов в зоне проводимости и условии неполного заполнения электронами валентной зоны. При выполнении этих условий и в отсутствие градиента температуры перенос носителей заряда может происходить либо под действием электрического поля, либо под действием градиента концентрации носителей заряда.

Дрейф носителей заряда

Направленное движение носителей заряда под действием электрического поля называют дрейфом. Электроны, получая ускорение в электрическом поле, приобретают на длине свободного пробега дополнительную энергию около 10^{-8} — 10^{-4} эВ. При этом электроны переходят на более высокие энергетические уровни (разница в энергиях между соседними энергетическими уровнями в разрешенной зоне около 10^{-22} эВ). При очередном соударении электрона с атомом кристаллической решетки электрон отдает кристаллической решетке накопленную на длине свободного пробега энергию, возвращаясь на один из низко лежащих энергетических уровней в разрешенной зоне. Так можно представить процесс электропроводности в результате движения электронов в зоне проводимости и в валентной зоне. Однако учитывая почти полное заполнение электронами энергетических уровней валентной зоны, удобнее обычно рассматривать в валентной зоне движение дырок: дырки, двигаясь по направлению вектора электрического поля и приобретая в этом поле дополнительную энергию, переходят по энергетической диаграмме на более низкие энергетические уровни.

В результате дрейфа электронов в полупроводнике появляется электронная составляющая плотности дрейфового тока, которую можно записать на основании закона Ома в виде

$$\mathbf{J}_{n \text{ др}} = \gamma_n \mathbf{E} = q n \mu_n \mathbf{E}, \quad (1.22)$$

где γ_n — удельная проводимость полупроводника (любого материала) при одном виде имеющих в нем носителей заряда — электронов; μ_n — подвижность электронов, т. е. величина, численно равная средней скорости их направленного движения в электрическом поле с напряженностью, равной единице.

Аналогично дырочная составляющая плотности дрейфового тока

$$\mathbf{J}_{p \text{ др}} = \gamma_p \mathbf{E} = q p \mu_p \mathbf{E}. \quad (1.23)$$

Знаки в правых частях соотношений (1.22) и (1.23) одинаковые, так как заряды электронов и дырок различны по знаку и двигаются в электрическом поле в разные стороны.

Полная плотность дрейфового тока при наличии свободных электронов и дырок равна сумме электронной и дырочной составляющих:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p = q \gamma \mathbf{E} = q (n \mu_n + p \mu_p) \mathbf{E}, \quad (1.24)$$

где γ — удельная проводимость полупроводника (любого материала) с учетом свободных электронов и дырок.

Диффузия носителей заряда

Поведение свободных электронов и дырок в полупроводнике напоминает поведение молекул газа. Эту аналогию можно распространить и на явления, происходящие в результате неравномерного распределения концентрации носителей заряда в объеме полупро-

водника в отсутствие градиента температуры. В этом случае происходит диффузия — движение носителей заряда из-за градиента концентрации, выравнивание концентрации носителей по полупроводнику.

Из молекулярной физики известно, что поток частиц при диффузии (число частиц, пересекающих в единицу времени единичную площадку, перпендикулярную направлению градиента концентрации) пропорционален градиенту концентрации этих частиц:

$$\Phi_m = -D_m \text{grad } m, \quad (1.25)$$

где D_m — коэффициент диффузии, который равен абсолютной величине отношения плотности потока частиц к градиенту их концентрации.

Вектор градиента концентрации направлен в сторону возрастания аргумента, а частицы диффундируют туда, где их меньше, т. е. против градиента концентрации. Поэтому знаки перед правой и левой частями соотношения (1.25) различны.

Остановимся на процессе диффузии электронов и дырок в полупроводниках, т. е. диффузии заряженных частиц (или квазичастиц). Так как всякое направленное движение заряженных частиц это есть электрический ток, то плотность электронной составляющей диффузионного тока может быть получена путем умножения правой части (1.25) на элементарный заряд электрона. Электроны диффундируют против вектора градиента концентрации и имеют отрицательный заряд. Поэтому направление вектора плотности диффузионного тока электронов должно совпадать с направлением вектора градиента концентрации электронов, т. е.

$$\mathbf{J}_{n \text{ диф}} = qD_n \text{grad } n. \quad (1.26)$$

Аналогично плотность дырочной составляющей диффузионного тока

$$\mathbf{J}_{p \text{ диф}} = -qD_p \text{grad } p. \quad (1.27)$$

Заряд дырок положителен, поэтому направление вектора плотности диффузионного тока дырок должно совпадать с направлением их диффузии, т. е. должно быть противоположным направлению вектора градиента концентрации дырок. Следовательно, в правой части соотношения (1.27) должен сохраниться знак «минус».

Одновременно с процессом диффузии неравновесных носителей происходит процесс рекомбинации этих носителей. Поэтому избыточная концентрация уменьшается в направлении от места источника этой избыточной концентрации носителей. Расстояние, на котором при одномерной диффузии в полупроводнике без электрического поля в нем избыточная концентрация носителей заряда уменьшается вследствие рекомбинации в $e = 2,718 \dots$ раз, называют *диффузионной длиной* (L). Иначе — это расстояние, на которое носитель диффундирует за время жизни. Таким образом, диффузионная длина связана со временем жизни носителей соотношениями

$$L_n = \sqrt{D_n \tau_n}; \quad L_p = \sqrt{D_p \tau_p}, \quad (1.28)$$

где D_n и D_p — коэффициенты диффузии электронов и дырок.

В свою очередь коэффициенты диффузии связаны с подвижностью носителей заряда соотношениями Эйнштейна

$$D_n/\mu_n = kT/q; \quad D_p/\mu_p = kT/q. \quad (1.29)$$

Не следует путать диффузионную длину с *длиной свободного пробега* носителей заряда, которая определяется как среднее расстояние, проходимое носителем между двумя последовательными актами рассеяния.

Необходимо отметить, что диффузия носителей заряда может происходить в полупроводнике, имеющем первоначально равномерное распределение концентрации носителей, т. е. равный нулю градиент концентрации, но при наличии в полупроводнике разности температур или градиента температуры. В этом случае носители заряда (например, электроны), находящиеся в местах полупроводника с большей температурой, будут иметь большую энергию, т. е. будут занимать более высокие энергетические уровни в зоне проводимости. В местах полупроводника с меньшей температурой энергетические уровни с аналогичной энергией свободны от электронов. Поэтому возникает диффузия электронов из нагретых мест полупроводника в холодные места. Результатом такой диффузии является возникновение градиента концентрации носителей заряда. Это явление аналогично процессу диффузии в газе — при нагревании давление повышается и частицы газа диффундируют в области с низкой температурой. Процесс, происходящие в полупроводниковых приборах, часто рассматривают, пренебрегая разностью температур между различными областями полупроводниковой структуры прибора. В этом случае для расчета диффузионных токов можно использовать соотношения (1.26) и (1.27).

Уравнения токов

При наличии электрического поля и градиента концентрации носителей заряда в полупроводнике будут существовать дрейфовые и диффузионные токи. Поэтому плотность электронного тока с учетом (1.22) и (1.26):

$$\mathbf{J}_n = \mathbf{J}_{n \text{ др}} + \mathbf{J}_{n \text{ диф}} = qn\mu_n \mathbf{E} + qD_n \text{grad } n. \quad (1.30)$$

Плотность дырочного тока с учетом (1.23) и (1.27)

$$\mathbf{J}_p = \mathbf{J}_{p \text{ др}} + \mathbf{J}_{p \text{ диф}} = qp\mu_p \mathbf{E} - qD_p \text{grad } p. \quad (1.31)$$

Для расчета плотности полного тока следует сложить его электронную и дырочную составляющие, а также добавить к ним плотность тока смещения, равную производной по времени от вектора электрической индукции:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p + \frac{\partial}{\partial t} (\epsilon\epsilon_0 \mathbf{E}). \quad (1.32)$$

**§ 1.8. ТЕМПЕРАТУРНЫЕ ЗАВИСИМОСТИ
КОНЦЕНТРАЦИИ НОСИТЕЛЕЙ ЗАРЯДА И ПОЛОЖЕНИЯ
УРОВНЯ ФЕРМИ**

Температурные зависимости концентрации носителей заряда и положения уровня Ферми в широком диапазоне температур и при различных концентрациях примесей представлены на рис. 1.9.

Рассмотрим характер кривой, соответствующей относительно малой концентрации примесей (доноров) N_d . В области низких температур с увеличением температуры и, следовательно, с увеличением энергии теплового движения концентрация свободных электронов растет за счет ионизации доноров (участок кривой между точками 1 и 2). Угол наклона этого участка кривой характеризует энергию ионизации примесей. В данном диапазоне температур уровень Ферми находится между дном зоны проводимости и энергетическими донорными уровнями. При некоторой температуре, которой соответствует точка 2 кривой, вероятность заполнения донорных уровней оказывается равной 50 %. Поэтому уровень Ферми при температуре T_2 совпадает с энергетическим уровнем донора.

При дальнейшем увеличении температуры концентрация свободных электронов практически не увеличивается (участок кривой между точками 2 и 3), так как все примеси уже ионизированы, а вероятность ионизации собственных атомов полупроводника еще ничтожно мала. Участок кривой, соответствующий постоянной концентрации носителей заряда, называют участком истощения примесей. Первые два участка кривой (1—2 и 2—3) соответствуют примесной электропроводности полупроводника.

При относительно больших температурах (участок кривой за точкой 3) концентрация свободных электронов (носителей заряда) растет с увеличением температуры вследствие перехода электронов через запрещенную зону. Наклон этого участка кривой характеризует ширину запрещенной зоны $\Delta\mathcal{E}$ полупроводника. Уровень Ферми при этих температурах расположен вблизи середины запрещенной зоны, а полупроводник можно считать собственным, так как концентрация носителей заряда определяется ионизацией собственных атомов полупроводника. Температура, при которой наступает собственная электропроводность или при которой полупроводник становится собственным, будет тем меньше, чем меньше

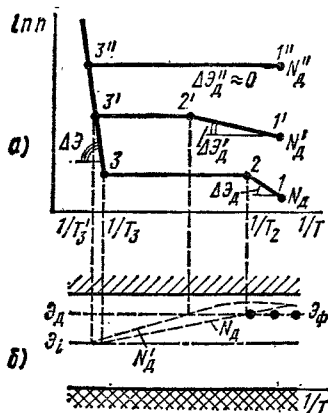


Рис. 1.9. Температурные зависимости концентрации свободных электронов в полупроводнике при различных концентрациях доноров $N_d < N'_d < N''_d$ (а) и соответствующие зависимости положения уровня Ферми (б)

ширина запрещенной зоны полупроводника. Принцип действия большинства полупроводниковых приборов основан на использовании примесной электропроводности, и поэтому появление собственной электропроводности нарушает нормальную работу прибора. Таким образом, температура, которой соответствует точка $З$ кривой, является максимальной рабочей температурой полупроводникового прибора, изготовленного из полупроводника с концентрацией примеси N_d .

Теперь рассмотрим смещение кривых и некоторое изменение их характера при увеличении концентрации примесей ($N_d' > N_d > N_d''$).

С увеличением концентрации примесей участки кривых, соответствующие примесной электропроводности, смещаются вверх, т. е. получается большая концентрация носителей заряда при температурах примесной электропроводности.

Угол наклона первого участка кривой (участка ионизации примесей) с увеличением концентрации примесей уменьшается, так как с увеличением концентрации примесей из-за взаимодействия примесных атомов происходит расщепление примесных энергетических уровней и уменьшение энергии ионизации примесей. Поэтому $\Delta E_d > \Delta E_d' > \Delta E_d''$. При достаточно большой концентрации примесей (N_d'') энергия ионизации примесей стремится к нулю, так как образовавшаяся примесная зона перекрывается с зоной проводимости. Такой полупроводник является *вырожденным* (полуметаллом).

Температура, соответствующая переходу от примесной электропроводности к собственной, увеличивается с увеличением концентрации примесей (например, $T_3' > T_3$). Это значит, что максимальная рабочая температура полупроводникового прибора, созданного на основе полупроводника с большей концентрацией примесей, будет также немного выше максимальной рабочей температуры такого же прибора из того же материала, но с меньшей концентрацией примесей.

§ 1.9. ТЕМПЕРАТУРНЫЕ ЗАВИСИМОСТИ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА И УДЕЛЬНОЙ ПРОВОДИМОСТИ

На величину подвижности носителей заряда в основном влияют два физических фактора: хаотические тепловые колебания атомов кристаллической решетки (рассеяние носителей заряда на тепловых колебаниях атомов кристаллической решетки) и электрические поля ионизированных примесей (рассеяние на ионах примесей). При больших температурах преобладает рассеяние носителей заряда на тепловых колебаниях атомов кристаллической решетки. Поэтому с увеличением температуры в этом диапазоне температур подвижность носителей уменьшается (рис. 1.10).

В диапазоне малых температур с уменьшением температуры уменьшаются тепловые скорости хаотического движения носителей заряда, что приводит к увеличению времени пребывания носителя

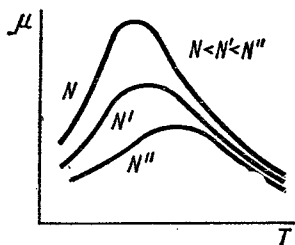


Рис. 1.10. Температурные зависимости подвижности носителей заряда при различных концентрациях примесей

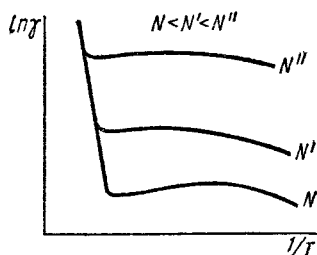


Рис. 1.11. Температурные зависимости удельной проводимости полупроводника при различных концентрациях примесей

(рис. 1.11). Концентрация носителей заряда в полупроводниках очень сильно зависит от температуры — по экспоненциальному закону (1.7), а на подвижность изменение температуры влияет сравнительно слабо — по степенному закону (исключение составляют оксидные полупроводники на основе окислов металлов с переменной валентностью). Поэтому температурная зависимость удельной проводимости похожа на температурную зависимость концентрации носителей при очень малых и при больших температурах. В диапазоне температур, соответствующих истощению примесей, когда концентрация основных носителей заряда остается практически неизменной, температурные изменения удельной проводимости обусловлены температурной зависимостью подвижности.

§ 1.10. ПОЛУПРОВОДНИКИ В СИЛЬНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

В сильных электрических полях в полупроводнике могут происходить физические процессы, приводящие к изменению удельной проводимости полупроводника; вольт-амперная характеристика полупроводника перестает подчиняться закону Ома; может изме-

няться как концентрация носителей заряда, так и их подвижность.

Рассмотрим вначале физические процессы, влияющие на концентрацию носителей заряда.

Ударная ионизация

Свободный электрон (или дырка), разгоняясь под действием большой напряженности электрического поля, может приобрести на длине свободного пробега дополнительную энергию, достаточную для ионизации примеси или собственного атома полупроводника. Процесс ионизации атомов разогнавшимся в поле носителем заряда и называют ударной ионизацией. Ионизацию могут вызывать и дырки, так как движение дырок является лишь способом описания движения совокупности электронов валентной зоны полупроводника.

Количественно процесс ударной ионизации характеризуется коэффициентами ударной ионизации, которые численно равны количеству пар носителей заряда, образуемых первичным носителем на единице пути. По аналогии с теорией электрического разряда в газах коэффициенты ударной ионизации в полупроводниках обозначают α_n и α_p . Коэффициенты ударной ионизации очень сильно зависят от напряженности электрического поля. Для практических расчетов часто пользуются эмпирической аппроксимацией

$$\alpha = A |E|^m, \quad (1.33)$$

где m — довольно большой показатель степени, различный для разных материалов (от 5 до 8).

Туннелирование

Сильному электрическому полю в полупроводнике соответствует большой наклон энергетических зон (рис. 1.12). При этом электроны могут проходить сквозь узкий потенциальный барьер (толщиной Δ) без изменения своей энергии — туннелировать благодаря своим квантовомеханическим свойствам. Так как процесс туннелирования происходит вследствие перехода электронов из валентной зоны в зону проводимости, то этот процесс можно считать аналогичным автоэлектронной эмиссии или холодной эмиссии электронов из металла.

Вероятность перехода электронов из валентной зоны в зону проводимости и, наоборот, из зоны проводимости в валентную зону одна и та же. Но переход электронов из валентной зоны преобладает, поскольку их там значительно больше, чем в зоне проводимости. Поэтому концентрация носителей заряда растет при туннелировании.

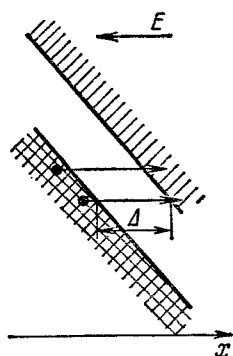


Рис. 1.12. Туннелирование электронов из валентной зоны в зону проводимости при сильном электрическом поле в полупроводнике

Туннельный эффект в полупроводниках проявляется при очень больших напряженностях электрического поля: в кремнии — при $E \approx 10^6$ В/см, в германии — при $E \approx 10^5$ В/см. Напряженности электрического поля, при которых появляется эффект туннелирования, различны для разных материалов, так как толщина потенциального барьера (Δ) зависит от ширины запрещенной зоны полупроводника при неизменной напряженности электрического поля, т. е. при неизменном наклоне энергетических зон.

Теперь рассмотрим влияние сильного электрического поля на подвижность носителей заряда.

Рассеяние носителей заряда в сильных полях

В слабых электрических полях носители заряда на длине свободного пробега приобретают относительно малую энергию. Поэтому их распределение по энергетическим уровням соответствует распределению при данной температуре кристаллической решетки. Дрейфовые скорости движения носителей заряда при этом значительно меньше так называемых тепловых скоростей, т. е. скоростей тепловых хаотических движений.

В сильных электрических полях скорость дрейфа носителей заряда соизмерима с тепловой скоростью; носители заряда на длине свободного пробега приобретают в электрическом поле энергии, соответствующие кинетическим энергиям теплового хаотического движения. При этом распределение носителей заряда по энергетическим уровням соответствует большим температурам, чем температура кристаллической решетки, которая остается практически неизменной. Это явление называют иногда «разогревом» носителей. На подвижность носителей явление разогрева может влиять по-разному.

1. При относительно больших температурах, при которых величина подвижности носителей заряда определяется в основном процес-

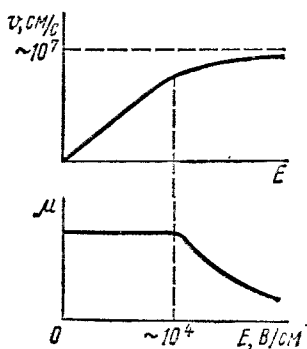


Рис. 1.13. Зависимость дрейфовой скорости и подвижности носителей заряда от напряженности электрического поля

сом рассеяния на тепловых колебаниях атомов кристаллической решетки полупроводника, «разогрев» носителей заряда электрическим полем приводит к увеличению числа столкновений носителей с атомами кристаллической решетки, т. е. к насыщению дрейфовой скорости или к уменьшению подвижности при увеличении напряженности электрического поля (рис. 1.13). Именно это явление надо учитывать в полупроводниковых приборах, если электрические поля превышают значения 10^3 — 10^4 В/см.

2. При относительно малых температурах, при которых величина подвижности носителей заряда определяется в основном процессом рассеяния на

ионизированных примесей, «разогрев» носителей электрическим полем приводит к уменьшению времени нахождения носителя в поле ионизированной примеси, т. е. к уменьшению рассеяния носителя и, следовательно, к увеличению подвижности. Следовательно, увеличение подвижности с увеличением напряженности электрического поля в полупроводниковых приборах может происходить только при очень низких температурах.

Междолинный переход носителей заряда

Рассмотрим вначале зависимость энергии свободного электрона находящегося в вакууме, от величины его импульса P (рис. 1.14). Энергия такого электрона

$$\mathcal{E} = \frac{m_0 v^2}{2} = \frac{m_0^2 v^2}{2m_0} = \frac{P^2}{2m_0},$$

где v — вектор скорости свободного электрона; m_0 — его масса.

Зависимость, представленная на рис. 1.14, является энергетической диаграммой свободных электронов в вакууме, изображенной в пространстве импульсов или в пространстве волновых векторов k ($P = \hbar k = \frac{h}{2\pi} k$).

В полупроводниковом кристалле свободный электрон можно считать свободным только условно, так как на электрон в кристалле действует периодическое потенциальное поле кристаллической решетки. Для того чтобы описать сложные законы движения электрона в кристалле с помощью соотношений, совпадающих по форме с законами классической механики, можно учесть влияние внутренних сил на электрон, изменив соответствующим образом значение его массы, т. е. введя понятие некоторой эффективной массы электрона (или дырки). Таким образом, *эффективная масса* — это коэффициент пропорциональности в законе, связывающем внешнюю силу, действующую на электрон в кристалле, с его ускорением.

Зона проводимости полупроводника может быть образована из нескольких перекрывающихся между собой разрешенных энергетических зон. Структура энергетических зон или энергетическая диаграмма полупроводника в пространстве квазиимпульсов (в k -пространстве) может иметь несколько минимумов

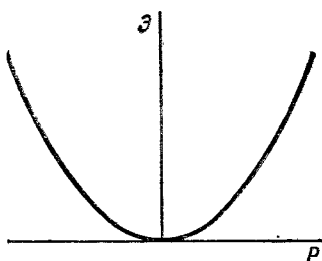


Рис. 1.14. Зависимость энергии свободного электрона, находящегося в вакууме, от величины его импульса

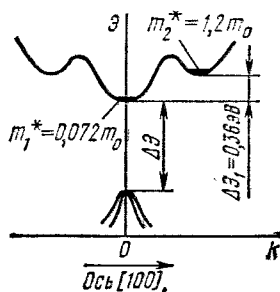


Рис. 1.15. Структура энергетических зон арсенида галлия в кристаллографическом направлении [100]

(рис. 1.15). Так, например, на такой энергетической диаграмме арсенида галлия, зона проводимости кроме центральной долины с минимумом энергии при волновом векторе $\mathbf{k} = 0$ имеет еще боковые долины с минимумом энергии, который отличается от предыдущего на величину $\Delta\mathcal{E}_1$.

По этой энергетической диаграмме можно заключить, что в зоне проводимости арсенида галлия могут существовать электроны, обладающие одной и той же энергией, но имеющие различные квазиимпульсы, а следовательно, и различные эффективные массы:

$$\mathcal{E} = \frac{\mathbf{P}_1^2}{2m_1^*} = \frac{\mathbf{P}_2^2}{2m_2^*}. \quad (1.34)$$

Если $\mathbf{P}_1 < \mathbf{P}_2$, то $m_1^* < m_2^*$, а это значит, что $\mu_1 > \mu_2$. Таким образом, в полупроводнике (арсениде галлия) могут существовать свободные электроны с разными подвижностями: «легкие» электроны с малой эффективной массой и с большой подвижностью в центральной долине и «тяжелые» электроны с большой эффективной массой и с малой подвижностью в боковых долинах.

В слабых электрических полях почти все свободные электроны имеют малые дрейфовые скорости и квазиимпульсы и поэтому находятся в центральной долине. В сильных электрических полях свободные электроны, приобретая дополнительную энергию, превышающую $\Delta\mathcal{E}_1$, получают возможность перейти в боковые долины. Там они характеризуются большей эффективной массой (станут «тяжелыми») и малой подвижностью. По этой причине средняя подвижность всех свободных электронов с увеличением напряженности электрического поля уменьшается. Подвижности «легких» и «тяжелых» электронов могут отличаться в десятки раз.

§ 1.11. ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

Поглощение света

Существуют различные виды поглощения света. При поглощении полупроводником квантов света — фотонов их энергия может быть передана электронам валентной зоны с переводом этих электронов в зону проводимости, т. е. энергия квантов света идет на ионизацию атомов полупроводника. Этот процесс называют собственным поглощением. Существует поглощение энергии квантов света свободными электронами зоны проводимости или дырками валентной зоны, т. е. поглощение носителями заряда. При этом энергия квантов света расходуется также на перевод носителей на более высокие для них энергетические уровни, но в пределах соответствующей разрешенной зоны. Возможно примесное поглощение, при котором энергия фотонов идет на ионизацию или возбуждение примесных атомов. Кроме того, в полупроводниках может происходить поглощение фотонов кристаллической решеткой, поглощение с переходом электронов с акцепторного на донорный энергетический уровень и

некоторые другие виды поглощения. Процессы поглощения фотонов не следует смешивать с процессами рассеяния, которые также приводят к уменьшению плотности потока фотонов.

При собственном поглощении фотонов переход электронов из валентной зоны в зону проводимости полупроводника может происходить без изменения квазиимпульса или волнового вектора электрона, т. е. возможны *прямые переходы* (рис. 1.16). Может происходить также переброс электронов из валентной зоны в зону проводимости и с изменением волнового вектора — *непрямые переходы* (рис. 1.17). При непрямах переходах в процессе поглощения кроме фотона и электрона должна участвовать еще третья квазичастица, которая заберет часть квазиимпульса на себя, т. е. обеспечит выполнение закона сохранения импульса. Таким образом, не прямые переходы — это переходы с участием третьей квазичастицы. Третьей квазичастицей обычно является фотон — квант тепловой энергии кристаллической решетки полупроводника.

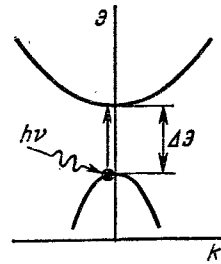


Рис. 1.16. Прямой переход электрона из валентной зоны в валентной зоны в зону проводимости полупроводника

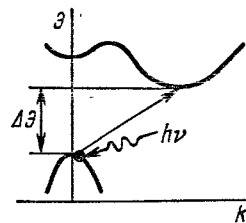


Рис. 1.17. Непрямой переход электрона из валентной зоны в зону проводимости полупроводника

Поглощение света или вообще фотонов характеризуют показателем поглощения α , который равен относительному изменению светового потока (потока фотонов) в слое полупроводника единичной толщины (рис. 1.18):

$$\alpha = -\frac{1}{\Phi} \frac{d\Phi}{dx}.$$

Это соотношение представляет собой дифференциальное уравнение с разделяющимися переменными. Поэтому

$$\int_{\Phi_0}^{\Phi} \frac{d\Phi}{\Phi} = -\alpha \int_0^x dx; \quad \ln \frac{\Phi}{\Phi_0} = -\alpha x \quad \text{и} \quad \Phi = \Phi_0 \exp(-\alpha x).$$

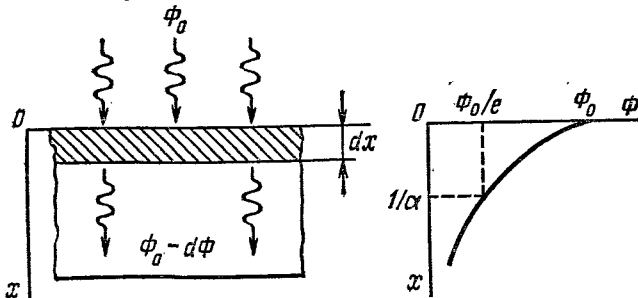


Рис. 1.18. Поглощение света в полупроводнике

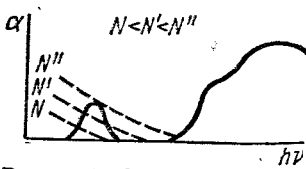


Рис. 1.19. Спектр поглощения полупроводника

Таким образом, показатель поглощения α можно определить как величину, обратную толщине слоя полупроводника, после прохождения которого световой поток (поток фотонов) уменьшится в $e = 2,718 \dots$ раз.

Зависимость показателя поглощения от энергии фотонов называют спектром поглощения полупроводника (рис. 1.19).

При больших энергиях фотонов происходит собственное поглощение с образованием пар носителей электрон—дырка. Показатель поглощения при этом велик. При малой энергии фотонов (меньше ширины запрещенной зоны полупроводника) показатель поглощения уменьшается.

При еще меньших энергиях квантов света может происходить примесное поглощение, если не все примеси ионизированы при данной температуре. Примесному поглощению соответствует один или несколько максимумов в спектре поглощения при энергиях квантов света, равных энергиям ионизации примесей.

При малых энергиях фотонов основным процессом поглощения является поглощение носителями заряда. Поэтому показатель поглощения в этом диапазоне частот зависит от концентрации носителей или от концентрации примесей.

Фоторезистивный эффект

Фоторезистивный эффект (внутренний фотоэлектрический эффект) — это изменение удельного сопротивления полупроводника, обусловленное исключительно действием электромагнитного излучения (квантов света) и не связанное с нагреванием полупроводника. Сущность этого явления состоит в том, что при поглощении квантов света с энергией, достаточной для ионизации собственных атомов полупроводника или ионизации примесей, происходит увеличение концентрации носителей заряда. В результате увеличения концентрации носителей уменьшается удельное сопротивление полупроводника.

Для существования фоторезистивного эффекта необходимо, чтобы в полупроводнике происходило либо собственное поглощение квантов света с образованием новых пар носителей заряда, либо примесное поглощение с образованием носителей одного знака.

В области собственного поглощения избыточные концентрации электронов и дырок, равные между собой, должны быть пропорциональны показателю поглощения, интенсивности света и времени жизни:

$$\Delta n = \Delta p = \beta \alpha I \tau.$$

Коэффициент пропорциональности β называют *внутренним квантовым выходом*, так как он определяет число пар носителей заряда (или число носителей заряда при примесном поглощении),

образуемых одним поглощенным квантом света, если интенсивность света I измерять числом квантов в секунду. Внутренний квантовый выход будет меньше единицы, если существуют условия для поглощения носителями заряда. В этом случае энергия некоторых квантов будет расходоваться на повышение энергии носителей заряда, а не на образование новых носителей заряда.

При освещении полупроводника наряду с генерацией новых носителей заряда происходит и обратный процесс — рекомбинация. При непрерывном освещении между этими двумя противоположными процессами генерации и рекомбинации устанавливается некоторое динамическое равновесие.

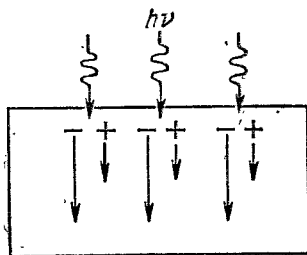


Рис. 1.20. Возникновение э. д. с. в полупроводнике при его освещении

Э. д. с., возникающая в полупроводнике при его освещении

Если однородный полупроводник осветить сильно поглощаемым светом, то в его поверхностном слое, где происходит поглощение света, возникнет избыточная концентрация электронов и дырок, которые будут диффундировать в глубь полупроводника (рис. 1.20). Коэффициент диффузии электронов обычно значительно больше коэффициента диффузии дырок. Поэтому при диффузии электроны опережают дырки, происходит некоторое разделение зарядов — поверхность полупроводника приобретает положительный заряд по сравнению с объемом. Таким образом, в полупроводнике при его освещении возникает электрическое поле или э. д. с., которую иногда называют *э. д. с. Дембера*.

Величину напряженности электрического поля, возникающую в полупроводнике при его освещении, можно найти, используя уравнения (1.30)—(1.32) с учетом того, что в установившемся состоянии тока через полупроводник нет. Тогда

$$0 = qn\mu_n E + qD_n \text{grad } n + qp\mu_p E - qD_p \text{grad } p.$$

Если же считать, что $\text{grad } n \approx \text{grad } p$, то

$$E = - \frac{D_n - D_p}{\mu_n n - \mu_p p} \text{grad } p, \quad (1.35)$$

т. е. напряженность электрического поля пропорциональна возникающему при освещении полупроводника градиенту концентрации носителей заряда.

Люминесценция полупроводников

Люминесценцией называют электромагнитное нетепловое излучение, обладающее длительностью, значительно превышающей период световых колебаний.

Для возникновения люминесценции в полупроводнике атомы полупроводника должны быть выведены из состояния термодинамического равновесия, т. е. возбуждены. Они могут быть переведены в возбужденное состояние электрическим полем (*электролюминесценция*), бомбардировкой полупроводника электронами (*катодолюминесценция*), освещением (*фотолюминесценция*) и с помощью других энергетических воздействий. При люминесценции акты поглощения энергии полупроводником и излучения квантов света разделены во времени (а может быть, и в пространстве) промежуточными процессами, что приводит к относительно длительному существованию свечения полупроводника после прекращения возбуждения.

Излучение квантов света из полупроводника может происходить в результате перехода электрона на более низкий энергетический уровень при межзонной рекомбинации или при рекомбинации с участием рекомбинационных ловушек (см. рис. 1.3).

Излучательная рекомбинация носителей заряда может произойти без электромагнитного воздействия, т. е. самопроизвольно. Такую рекомбинацию называют *самопроизвольной* или *спонтанной*. Так как спонтанное излучение случайно, имеет статистический характер, то оно некогерентно. Это обусловлено тем, что акты спонтанного излучения происходят независимо друг от друга в разные моменты времени.

Переход электрона на более низкий энергетический уровень с излучением кванта света может произойти с помощью электромагнитного воздействия. Такую рекомбинацию (такой переход) называют *вынужденной*, *индуцированной* или *стимулированной*. Следовательно, квант света с определенной частотой может не только поглощаться полупроводником, но и вызывать добавочное индуцированное излучение. Индуцированное излучение происходит в том же направлении, что и вызвавшие его излучение, в одной и той же фазе и с одинаковой поляризацией, т. е. индуцированное излучение является когерентным.

§ 1.12. ОБЕДНЕННЫЕ, ИНВЕРСНЫЕ И ОБОГАЩЕННЫЕ ПОВЕРХНОСТНЫЕ СЛОИ

Поверхность полупроводника представляет собой нарушение периодичности кристаллической решетки. Из-за этого возникают дополнительные энергетические уровни, расположенные в запрещенной зоне энергетической диаграммы полупроводника. Эти уровни, теоретически предсказанные И. Е. Таммом, называют уровнями Тамма. Уровни Тамма являются акцепторными, так как у атомов полупроводника (например, кремния), находящихся у поверхности кристалла, оказывается всего по три соседних атома вместо четырех и, следовательно, отсутствует одна электронная связь. Плотность поверхностных уровней Тамма или поверхностных состояний должна быть того же порядка, что и число атомов на единице поверхности кристалла, т. е. около 10^{15} см⁻².

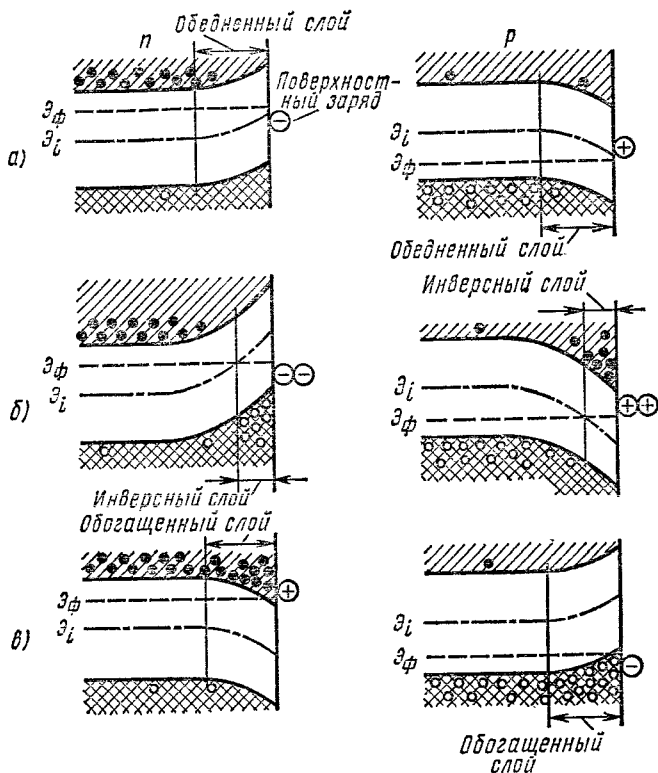


Рис. 1.21. Искажение энергетических зон у поверхности полупроводника и образование поверхностных слоев

На реальной поверхности полупроводника поверхностные состояния возникают также вследствие адсорбции различных примесей (кислорода, воды и других атомов, ионов и молекул). Очевидно, что такая сложная структура поверхности реального полупроводника характеризуется дополнительными энергетическими уровнями в запрещенной зоне донорного, акцепторного типа или типа ловушек.

Локальные энергетические уровни, обусловленные нарушением периодичности кристалла у поверхности полупроводника или примесями на поверхности, называют **поверхностными уровнями**.

При комнатной температуре большинство примесей обычно ионизировано, т. е. на поверхностных уровнях находятся заряды. Для компенсации этих зарядов в соответствии с условием электрической нейтральности должен существовать объемный заряд в полупроводнике, что соответствует существованию электрического поля и изгибу энергетических зон вблизи поверхности полупроводника.

На рис. 1.21 показано образование трех возможных вариантов поверхностных слоев в полупроводниках *n*- и *p*-типа при наличии положительных или отрицательных поверхностных состояний.

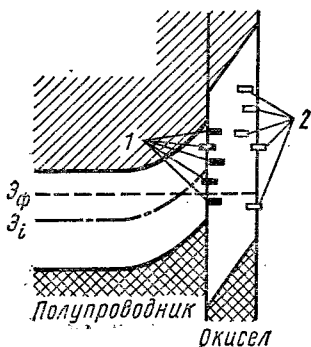


Рис. 1.22. Быстрые (1) и медленные (2) поверхностные состояния на реальной поверхности полупроводника

При малой плотности отрицательных поверхностных состояний на полупроводнике n -типа образуется обедненный слой (рис. 1.21, а), так как основные носители заряда — электроны отталкиваются отрицательным поверхностным зарядом в глубь полупроводника. Электростатическое поле поверхностных зарядов проникает в полупроводник на определенную глубину, которая зависит от удельного сопротивления этого полупроводника.

При большой плотности отрицательных поверхностных состояний у поверхности полупроводника n -типа образуется слой с противоположным типом электропроводности — инверсный слой. Граница инверсного слоя в глубине полупроводника расположена там, где уровень Ферми пересекает середину запрещенной зоны. Под инверсным слоем в полупроводнике находится обедненный слой.

Если на поверхности полупроводника n -типа преобладают положительные поверхностные состояния, то поверхность полупроводника обогащается электронами — основными носителями заряда, т. е. образуется обогащенный слой. Граница обогащенного слоя в глубине полупроводника находится там, где начинается изгиб энергетических уровней, т. е. определяется глубиной проникновения электростатического поля поверхностных зарядов.

Таким образом, толщина области объемного заряда зависит от плотности поверхностных состояний, от удельного сопротивления полупроводника или от концентрации примеси и составляет обычно 10^{-2} — 10 мкм.

На поверхности полупроводника p -типа могут образовываться аналогичные поверхностные слои, но при других знаках поверхностных зарядов (рис. 1.21).

На реальном полупроводнике всегда имеется слой окисла. Поэтому поверхностные состояния могут находиться не только непосредственно на полупроводнике, но также в слое окисла и на его поверхности. При изменении внешнего электрического поля и при соответствующем изменении энергетической диаграммы вблизи поверхности полупроводника должно происходить заполнение или опустошение электронами, по крайней мере, некоторых поверхностных состояний. Поверхностные состояния, расположенные вблизи границы раздела полупроводник — окисел, заполняются или опустошаются относительно быстро, так как в полупроводнике есть достаточное количество свободных электронов в зоне проводимости и дырок (пустых энергетических уровней) в валентной зоне. Из-за малого времени релаксации таких поверхностных состояний (10^{-8} — 10^{-4} с) их называют *быстрыми поверхностными состояниями* (рис. 1.22).

В слое окисла концентрация носителей заряда ничтожно мала. Поэтому с изменением внешнего электрического поля поверхностные состояния, расположенные в слое окисла или на его поверхности, перезаряжаются медленно (время их релаксации велико). Такие поверхностные состояния называют *медленными поверхностными состояниями*. Время релаксации медленных поверхностных состояний обычно от миллисекунд до нескольких часов.

§ 1.13. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

Некоторые поверхностные состояния создают вблизи середины запрещенной зоны энергетические уровни, которые являются уровнями рекомбинационных ловушек. Рекомбинационными ловушками могут быть только быстрые поверхностные состояния, так как время перехода носителей на медленные поверхностные состояния очень велико. Явление поверхностной рекомбинации принято характеризовать *скоростью поверхностной рекомбинации* носителей заряда, которая определяется как отношение плотности потока носителей заряда у поверхности полупроводника к избыточной концентрации этих носителей у поверхности, т. е.

$$s_p = \frac{\Phi_p}{\Delta p} = \frac{J_p}{q \Delta p}. \quad (1.36)$$

Таким образом, скорость поверхностной рекомбинации показывает, какое количество носителей заряда рекомбинирует за 1 с на поверхности полупроводника площадью в 1 см². Размерность скорости поверхностной рекомбинации — скорость.

Плотность поверхностных состояний вообще и поверхностных состояний, являющихся рекомбинационными ловушками, в частности зависит от обработки поверхности полупроводника и от свойств внешней среды, с которой он соприкасается. Так как плотность поверхностных состояний обычно велика, то рекомбинация носителей на поверхности идет значительно быстрее (интенсивнее), чем в объеме полупроводника. При малых размерах кристалла полупроводника явление поверхностной рекомбинации будет существенно уменьшать эффективное время жизни носителей заряда, так как

$$1/\tau_{эф} = 1/\tau_v + 1/\tau_s,$$

где $\tau_{эф}$ — эффективное время жизни; τ_v — время жизни в объеме полупроводника; τ_s — время жизни на поверхности полупроводника.

Кроме того, плотность поверхностных состояний может изменяться со временем из-за испарения или конденсации влаги на поверхности кристалла, из-за возможных миграций адсорбированных примесей на поверхности и т. п. Эти процессы, приводя к изменению эффективного времени жизни носителей заряда, могут являться причиной нестабильности параметров и характеристик полупроводниковых приборов.

Таким образом, при изготовлении полупроводниковых приборов необходимо, во-первых, выбирать метод обработки поверхности кри-

сталлов полупроводника, при котором скорость поверхностной рекомбинации минимальная, и, во-вторых, находить способ длительного сохранения достигнутых значений скорости поверхностной рекомбинации. Последнюю задачу обычно решают посредством нанесения на поверхность кристалла специальных покрытий и герметизации прибора в корпус.

§ 1.14. ПРОВОДИМОСТЬ КАНАЛА ПОВЕРХНОСТНОЙ ЭЛЕКТРОПРОВОДНОСТИ

При наличии на поверхности полупроводника инверсного слоя по нему вдоль поверхности могут проходить токи, т. е. могут существовать каналы поверхностной электропроводности. При оценке проводимости канала поверхностной электропроводности необходимо учитывать, что концентрация носителей заряда является величиной переменной по глубине канала и подвижность носителей заряда в канале может значительно отличаться от подвижности тех же носителей в объеме полупроводника из-за дополнительного рассеяния носителей на поверхности кристалла.

Проводимость тонкого слоя (толщиной dz) канала, например, с электропроводностью p -типа

$$d\sigma = q\rho\mu_p dz \frac{b}{l},$$

где b — ширина канала; l — длина канала.

Проводимость канала толщиной W

$$\sigma = q \frac{b}{l} \int_0^W \rho\mu_p dz,$$

его сопротивление

$$R = \frac{l}{\sigma} = \frac{l}{q \int_0^W \rho\mu_p dz} \cdot \frac{l}{b}.$$

Вместо удельного сопротивления при оценке токов в канале пользуются другим параметром — удельным сопротивлением слоя (ρ_s), аналогичным удельному поверхностному сопротивлению диэлектриков, т. е. *удельным сопротивлением слоя* называют сопротивление квадрата этого слоя току, проходящему между двумя противоположными сторонами квадрата.

Тогда, используя удельное сопротивление слоя, можно записать, что

$$R = \rho_s \frac{l}{b},$$

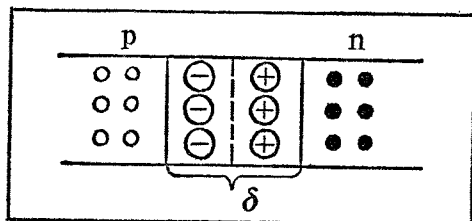
где $\rho_s = \frac{1}{q \int_0^W \rho\mu_p dz}$ — удельное сопротивление слоя или канала поверхностной электропроводности, Ом.

Контрольные вопросы

1. Что такое разрешенные и запрещенные энергетические зоны?
2. Что такое ширина запрещенной зоны?
3. Каковы правила наиболее целесообразного построения энергетических диаграмм?
4. Чем определяется наклон (угловой коэффициент) энергетических уровней и зон на энергетической диаграмме полупроводника?
5. Что такое уровень Ферми?
6. В каких случаях и почему надо пользоваться различными статистиками распределения электронов по энергиям?
7. Что такое собственный полупроводник?
8. Чему равно произведение концентрации электронов и дырок в невырожденном полупроводнике при термодинамическом равновесии?
9. Что такое диффузия и дрейф носителей заряда?
10. Что такое подвижность носителей заряда?
11. Что такое диффузионная длина и длина свободного пробега носителей заряда?
12. Как объяснить температурную зависимость концентрации носителей заряда в полупроводнике?
13. Какими физическими факторами объясняется температурная зависимость подвижности носителей заряда?
14. Как и почему изменяется рассеяние носителей заряда в сильных электрических полях?
15. Что такое прямые и непрямые переходы носителей заряда между разрешенными зонами полупроводника?
16. Что такое показатель поглощения света полупроводником?
17. Что такое фоторезистивный эффект?
18. Какие разновидности поверхностных слоев могут возникать на полупроводнике при наличии различных поверхностных состояний?
19. Что такое скорость поверхностной рекомбинации?
20. Что такое удельное сопротивление слоя и какова размерность этого параметра?

КОНТАКТНЫЕ ЯВЛЕНИЯ

2



§ 2.1. ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

Образование электронно-дырочного перехода

При идеальном контакте двух полупроводников с различным типом электропроводности из-за градиента концентрации носителей заряда возникает их диффузия в области с противоположным типом электропроводности через плоскость *металлургического контакта* (плоскость, где изменяется тип примесей, преобладающих в полупроводнике). В результате диффузии носителей заряда нарушается электрическая нейтральность примыкающих к металлургическому контакту частей монокристалла полупроводника. В *p*-области вблизи металлургического контакта после диффузии из нее дырок остаются нескомпенсированные ионизированные акцепторы (отрицательные неподвижные заряды), а в *n*-области — нескомпенсированные ионизированные доноры (положительные неподвижные заряды). Образуется область объемного заряда, состоящая из двух разноименно заряженных слоев. Между нескомпенсированными разноименными зарядами ионизированных примесей возникает электрическое поле, направленное от *n*-области к *p*-области и называемое *диффузионным электрическим полем* (рис. 2.1, а). Возникшее диффузионное электрическое поле препятствует дальнейшей диффузии основных носителей через металлургический контакт — устанавливается равновесное состояние. Между *n*- и *p*-областями при этом существует разность потенциалов, называемая *контактной разностью потенциалов*. Потенциал *n*-области положителен по отношению к потенциалу *p*-области.

Таким образом, *электронно-дырочный переход* (*p-n*-переход) — это переходный слой между двумя областями полупроводника с разной электропроводностью, в котором существует диффузионное электрическое поле.

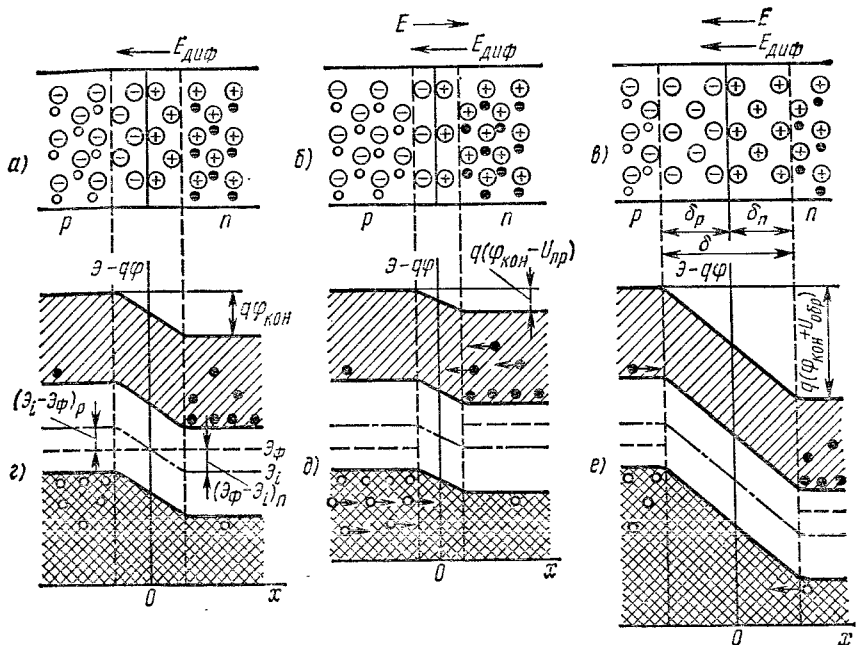


Рис. 2.1. Пространственное распределение зарядов (а, б, в) и энергетические диаграммы электронно-дырочного перехода (г, д, е):

а, г — внешнее напряжение отсутствует ($U = 0$); б, д — внешнее напряжение прямое ($U > 0$); в, е — внешнее напряжение обратное ($U < 0$)

Энергетическая диаграмма электронно-дырочного перехода

Согласно сказанному в § 1.1 энергетическую диаграмму электронно-дырочного перехода можно изобразить, как показано на рис. 2.1, г. Вдали от контакта двух областей электрическое поле отсутствует (если соответствующие области легированы равномерно) или относительно мало по сравнению с полем в p - n -переходе. Поэтому энергетические зоны в этих областях изображены горизонтальными. Взаимное расположение разрешенных зон и уровня Ферми за пределами n - p -перехода остается таким же, каким было в соответствующих полупроводниках.

Так как напряженность диффузионного электрического поля в p - n -переходе направлена от электронного полупроводника к дырочному, на диаграмме соответствующие энергетические зоны для n -области должны быть ниже, чем для p -области. Величина такого сдвига зон определяется тем, что в равновесном состоянии уровень Ферми должен быть расположен на одной высоте энергетической диаграммы всей системы. Величина этого сдвига зон соответствует также контактной разности потенциалов ($\varphi_{\text{кон}}$) или высоте потенциального барьера ($q\varphi_{\text{кон}}$) электронно-дырочного перехода.

Высота потенциального барьера и контактная разность потенциалов

Как видно из рис. 2.1, z , высота потенциального барьера p - n -перехода

$$q\varphi_{\text{кон}} = (\mathcal{E}_\Phi - \mathcal{E}_i)_n + (\mathcal{E}_i - \mathcal{E}_\Phi)_p. \quad (2.1)$$

На основе соотношений (1.17) и (1.18) можно записать:

$$\ln \frac{n_{n0}}{n_i} = \frac{\mathcal{E}_\Phi - \mathcal{E}_i}{kT} \quad \text{и} \quad \ln \frac{p_{p0}}{n_i} = \frac{\mathcal{E}_i - \mathcal{E}_\Phi}{kT}.$$

Следовательно, высота потенциального барьера

$$q\varphi_{\text{кон}} = kT \ln \frac{n_{n0} p_{p0}}{n_i^2} \quad (2.2)$$

или, при использовании соотношения (1.19),

$$q\varphi_{\text{кон}} = kT \ln \frac{p_{p0}}{p_{n0}} = kT \ln \frac{n_{n0}}{n_{p0}}. \quad (2.3)$$

Все приведенные выражения для высоты потенциального барьера равносильны, но удобнее пользоваться (2.2), так как концентрация основных носителей заряда практически равна концентрации легирующей примеси либо легко может быть определена из удельной проводимости прилегающих к p - n -переходу областей. Значение собственной концентрации носителей n_i для одной температуры дается в литературе, а температурная зависимость собственной концентрации определяется шириной запрещенной зоны полупроводника [см. (1.15)].

Для удобства анализа подставим (1.9) или (1.15) в (2.2) и учтем, что концентрация основных носителей заряда значительно меньше плотности возможных состояний в соответствующих зонах, т. е. $n_{n0} p_{p0} \ll 4N_c N_v$. Тогда

$$q\varphi_{\text{кон}} = \Delta\mathcal{E} - kT \ln \frac{4N_c N_v}{n_{n0} p_{p0}}. \quad (2.4)$$

Из соотношения (2.4) для высоты потенциального барьера можно сделать следующие выводы:

1) при одних и тех же концентрациях примесей высота потенциального барьера больше в p - n -переходах, созданных в полупроводниках с большей шириной запрещенной зоны;

2) высота потенциального барьера возрастает при увеличении концентрации примесей в соответствующих областях;

3) с увеличением температуры высота потенциального барьера уменьшается.

§ 2.2. ТОКИ ЧЕРЕЗ ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

При нарушении равновесия внешним электрическим полем через p - n -переход начинает проходить ток. Если внешнее напряжение приложено так, что создаваемая им напряженность электрического поля противоположна по направлению диффузионной напряженно-

сти (рис. 2.1, б), то суммарная напряженность поля в p - n -переходе падает, высота потенциального барьера уменьшается (рис. 2.1, д). Часть основных носителей, имеющих наибольшие значения энергии, может теперь преодолевать понизившийся потенциальный барьер, переходя через p - n -переход. Это приводит к появлению сравнительно большого тока через p - n -переход. Напряжение рассмотренной полярности называют прямым и считают положительным.

Преодолевшие потенциальный барьер носители заряда оказываются в соседней области неосновными. Таким образом, через p - n -переход происходит инжекция неосновных носителей заряда в область, примыкающую к p - n -переходу. Ту область, в которую происходит инжекция неосновных носителей, называют базой полупроводникового прибора.

С увеличением внешнего прямого напряжения уменьшается суммарная напряженность электрического поля в p - n -переходе. С уменьшением напряженности электрического поля уменьшается глубина проникновения этого поля в области полупроводника, прилегающие к контакту. Поэтому уменьшается толщина p - n -перехода или ширина области объемного заряда.

Если созданное внешним источником электрическое поле в p - n -переходе совпадает по направлению с диффузионным (рис. 2.1, в), то высота потенциального барьера для основных носителей увеличивается (рис. 2.1, е). Однако для неосновных носителей, т. е. для дырок в n -области и для электронов в p -области, потенциальный барьер в p - n -переходе вообще отсутствует. Неосновные носители заряда будут втягиваться электрическим полем в p - n -переход и проходить через него в соседнюю область — будет происходить так называемая экстракция. При этом через p - n -переход будет идти обратный ток, который относительно мал из-за малой концентрации неосновных носителей заряда в прилегающих к p - n -переходу областях.

Напряжение, имеющее рассмотренную полярность, называют обратным и считают отрицательным. Толщина p - n -перехода с увеличением обратного напряжения по абсолютному значению увеличивается, так как при этом увеличивается суммарная напряженность электрического поля в p - n -переходе и увеличивается глубина проникновения этого поля в прилегающие к контакту области.

§ 2.3. КОНЦЕНТРАЦИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА У ГРАНИЦ ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

Рассмотрим зависимость концентрации неосновных носителей заряда у границ p - n -перехода от внешнего напряжения, приложенного к электронно-дырочному переходу, для частных случаев.

Малые токи

Как известно, плотности электронного и дырочного токов определяются алгебраической суммой дрейфовых и диффузионных составляющих [см. (1.30) и (1.31)]. При движении носителей заряда

только в одном направлении x , параллельном вектору электрического поля,

$$\begin{aligned} J_p &= -q p \mu_p \frac{d\varphi}{dx} - q D_p \frac{dp}{dx}; \\ J_n &= -q n \mu_n \frac{d\varphi}{dx} + q D_n \frac{dn}{dx}. \end{aligned}$$

Учитывая малость токов и воспользовавшись соотношением Эйнштейна (1.29), можно считать, что при $J_p \approx 0$ и $J_n \approx 0$

$$\begin{aligned} -\frac{q}{kT} \frac{d\varphi}{dx} = \frac{1}{p} \frac{dp}{dx} \quad \text{или} \quad \frac{q}{kT} \int_0^{\varphi_{\text{кон}} - u} d\varphi &= - \int_{p_p}^{p_{n, \text{гр}}} \frac{dp}{p}; \\ \frac{q}{kT} \frac{d\varphi}{dx} = \frac{1}{n} \frac{dn}{dx} \quad \text{или} \quad \frac{q}{kT} \int_0^{\varphi_{\text{кон}} - u} d\varphi &= \int_{n_{p, \text{гр}}}^{n_n} \frac{dn}{n}. \end{aligned}$$

При малых токах концентрации основных носителей за пределами p - n -перехода p_p и n_n практически равны равновесным концентрациям p_{p0} и n_{n0} . Тогда с учетом (2.3)

$$p_{n, \text{гр}} = p_{n0} \exp \frac{qu}{kT}; \quad n_{p, \text{гр}} = n_{p0} \exp \frac{qu}{kT}. \quad (2.5)$$

Полученные соотношения имеют довольно простой физический смысл. В невырожденном полупроводнике носители заряда подчиняются статистике Максвелла—Больцмана, т. е. число их с энергией выше некоторого значения экспоненциально падает с увеличением этой энергии. В состоянии равновесия концентрация неосновных носителей заряда по одну сторону p - n -перехода равна концентрации основных носителей по другую сторону p - n -перехода, имеющих энергию, большую на $q\varphi_{\text{кон}}$. При изменении высоты потенциального барьера на qu количество носителей, имеющих энергию, достаточную для его преодоления, изменяется в $\exp \frac{qu}{kT}$ раз, что и характеризуется формулами (2.5).

При напряжении на p - n -переходе, равном нулю, граничная концентрация неосновных носителей заряда равна равновесной. С ростом прямого напряжения ($u > 0$) граничная концентрация неосновных носителей заряда растет, что соответствует явлению инжекции. При обратном напряжении ($u < 0$) граничная концентрация неосновных носителей падает, что соответствует явлению экстракции.

Большие прямые токи

При увеличении прямого напряжения на p - n -переходе высота его потенциального барьера уменьшается, что ведет к выравниванию концентраций носителей по обе стороны от p - n -перехода. В пределе, если контактная разность потенциалов могла бы быть полностью скомпенсирована внешним напряжением (практически это

недостижимо из-за падения части внешнего напряжения в объеме полупроводника), выполнялись бы условия

$$p_{n, \text{гр max}} = p_p; \quad n_{p, \text{гр max}} = n_n. \quad (2.6)$$

Соотношения (2.6) можно рассматривать как верхние пределы концентрации инжектированных через p - n -переход носителей.

Большие обратные токи

Соотношения (2.5) показывают, что концентрация неосновных носителей заряда на границе p - n -перехода, смещенного в обратном направлении, с ростом абсолютного значения обратного напряжения должна очень быстро падать. Однако такое падение концентрации ограничено тем, что скорость движения носителей заряда в электрическом поле растет до определенного предела v_{max} (см. § 1.10). Так как плотность тока, например дырочного, связана со скоростью движения носителей соотношением

$$J_p = q p v_p, \quad (2.7)$$

то минимальное значение концентрации неосновных носителей заряда, которое может получиться на границе p - n -перехода,

$$p_{n, \text{гр min}} = \frac{J_p}{q v_{p \text{ max}}}; \quad n_{p, \text{гр min}} = \frac{J_n}{q v_{n \text{ max}}}. \quad (2.8)$$

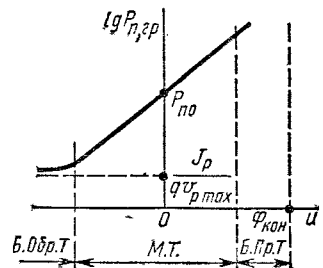
Выражения (2.8) можно рассматривать как нижние пределы концентрации неосновных носителей на границе p - n -перехода.

Зависимость граничной концентрации неосновных носителей заряда от напряжения

Для того чтобы подвести итог, построим в полулוגарифмическом масштабе зависимость граничной концентрации неосновных носителей от напряжения на p - n -переходе (рис. 2.2). При малых токах эта зависимость экспоненциальна, а в выбранном масштабе — прямая, проходящая через значение равновесной концентрации неосновных носителей заряда при $u = 0$. При больших прямых токах, т. е. при напряжениях на p - n -переходе, стремящихся к контактной разности потенциалов, изображать зависимость не имеет смысла. При больших обратных токах граничная концентрация неосновных носителей заряда стремится к значению $J_p/qv_{p \text{ max}}$.

Рис. 2.2. Зависимость граничной концентрации неосновных носителей заряда около p - n -перехода от напряжения, приложенного к p - n -переходу:

Б. Обр. Т. — диапазон больших обратных токов;
М. Т. — диапазон малых токов; Б. Пр. Т. —
диапазон больших прямых токов



§ 2.4. МЕТОДЫ ФОРМИРОВАНИЯ И КЛАССИФИКАЦИЯ ЭЛЕКТРОННО-ДЫРОЧНЫХ ПЕРЕХОДОВ

Среди разнообразных методов формирования p - n -переходов наибольшее значение имеют два: метод в плавления и метод диффузии примесей. Электронно-дырочный переход, полученный методом в плавления в полупроводник (с последующей рекристаллизацией полупроводника) металла или сплава, содержащего донорные или акцепторные примеси, называют *сплавным* переходом, а переход, полученный в результате диффузии примеси в полупроводник, — *диффузионным*.

При в плавлении в полупроводник металла или сплава, содержащего донорные или акцепторные примеси, полупроводник с навеской в плавления материала нагревают до расплавления навески, в результате чего часть кристалла полупроводника растворяется в расплаве навески. При последующем охлаждении происходит рекристаллизация полупроводникового кристалла с примесью в плавления материала. Если рекристаллизованный слой получился с другим типом электропроводности по сравнению с электропроводностью исходного полупроводника, то на границе их раздела возникает p - n -переход (рис. 2.3).

При создании *диффузионного p - n -перехода* используют диффузию в полупроводник примеси, находящейся в газообразной, жидкой или твердой фазе (рис. 2.4). Диффузионные p - n -переходы в свою очередь могут быть нескольких разновидностей. Так, диффузионный p - n -переход, образованный в результате диффузии примеси сквозь отверстие в защитном слое, нанесенном на поверхность полупроводника, называют *планарным p - n -переходом*. В качестве защитного слоя на кремнии обычно используют двуокись того же самого кремния.

На рис. 2.5 показаны этапы технологического процесса формирования планарного p - n -перехода. Основой такой технологии является фотолитография. На исходную окисленную пластину монокристалла кремния наносят слой фоточувствительного вещества — фоторезиста ΦP (рис. 2.5, а). Пленку фоторезиста освещают через маску ультрафиолетовым светом (рис. 2.5, б). Экспонированные места фоторезиста полимеризуются и становятся нерастворимыми. После этого незаполимеризованные части фото-

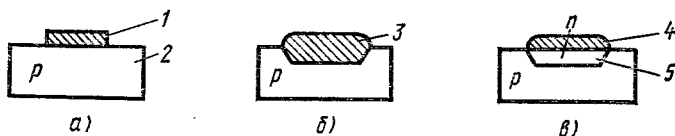


Рис. 2.3. Технологические стадии в плавления примесей в монокристалл полупроводника:

а — навеска в плавления материала 1, содержащая донорную примесь, на поверхности монокристалла полупроводника 2 перед в плавлением; б — расплав материала навески с полупроводником 3 при высокой температуре в плавления; в — избыточный слой материала навески 4 на поверхности рекристаллизованного полупроводника 5 после охлаждения

резиста смывают, так что он остается только на облученных местах (рис. 2.5, в). Затем производят травление пленки окисла, которая остается только в тех местах, где она была защищена фоторезистом (рис. 2.5, г). В дальнейшем проводят диффузию примеси в пластину исходного кремния. Диффузия примеси происходит селективно — только через отверстия в слое окиси кремния (рис. 2.5, д).

Для создания более сложных структур, например транзисторных с близко расположенными двумя p - n -переходами, необходимо еще раз повторить все рассмотренные этапы технологического процесса, т. е. заново окислить пластину кремния, нанести слой фоторезиста, засветить его определенные части поверхности, провести травление и диффузию через образовавшиеся отверстия — окошки.

Диффузионный p - n -переход, образованный в результате обратной диффузии примеси из полупроводника в соседнюю область, которая может быть металлом или полупроводником, называют *конверсионным p - n -переходом*. Так, например, для создания конверсионного p - n -перехода в германиевом монокристалле применяют германий, содержащий два типа примесей — донорную и акцепторную (медь). Концентрация меди в германии должна быть больше концентрации донорной примеси, т. е. исходный германий имеет электропроводность p -типа. В исходный германий проводят вплавление навески металла или сплава. Медь, отличаясь высоким коэффициентом диффузии в германии, при вплавлении диффундирует из германия в навеску. В результате из слоя

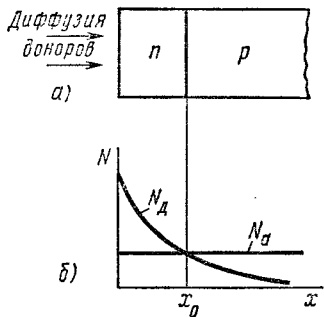


Рис. 2.4. Диффузионный p - n -переход (а) и распределение примесей в полупроводнике после диффузии в него доноров (б):

N_a — исходная концентрация акцепторов; N_d — концентрация доноров после диффузии; x_0 — глубина залегания сформированного p - n -перехода

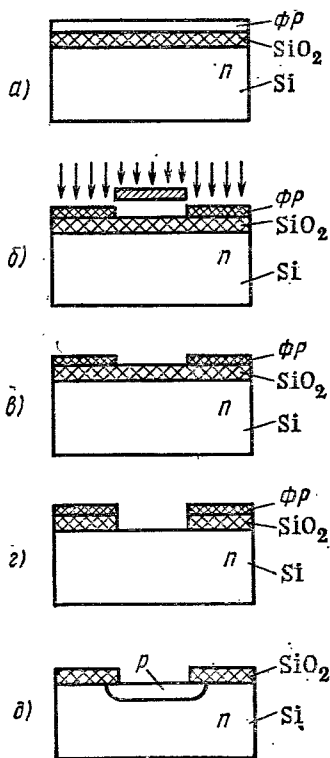


Рис. 2.5. Схема технологического процесса формирования планарного p - n -перехода:

а — окисленная пластинка с нанесенным фоторезистом; б — экспонирование фоторезиста через маску; в — пластинка после растворения незаполнеризованного фоторезиста; г — травление пленки окисла; д — пластинка после снятия фоторезиста и проведения диффузии

германия, примыкающего к навеске, удаляется акцепторная примесь меди, изменяется тип электропроводности этого слоя, т. е. происходит так называемая *конверсия*.

Кроме основных методов формирования *p-n*-переходов (вплавление и диффузия примесей) отметим еще метод *эпитаксиального наращивания*. Сущность эпитаксиального наращивания состоит в том, что какое-либо химическое соединение полупроводникового элемента, разлагаясь у поверхности полупроводника, образует на нем слой, структура которого является продолжением монокристаллической структуры исходного полупроводника — подложки. Чаще всего в качестве соединения применяют хлориды или иодиды германия и кремния с примесью веществ, необходимых для легирования эпитаксиального слоя. Таким образом, эпитаксиальный слой можно получить с тем же типом электропроводности, что и исходный полупроводник — подложка, но с другим удельным сопротивлением; а можно получить эпитаксиальный слой с другим типом электропроводности, т. е. создать *эпитаксиальный p-n-переход*.

По характеру распределения концентрации примеси различают резкие и плавные *p-n*-переходы. Переход, в котором толщина области изменения концентрации примеси значительно меньше толщины *p-n*-перехода, называют *резким p-n-переходом*. Резкий *p-n*-переход получается обычно при методе вплавления примеси. Переход, в котором толщина области изменения концентрации примеси сравнима или больше толщины *p-n*-перехода, называют *плавным p-n-переходом*. Плавный переход получают обычно при изготовлении методом диффузии примеси.

По соотношению концентраций основных носителей заряда или соответствующих примесей в *p*- и *n*-областях различают симметричные и несимметричные *p-n*-переходы. У *симметричных p-n-переходов* концентрации основных носителей заряда в прилегающих к переходу *p*- и *n*-областях приблизительно равны ($p_{p0} \approx n_{n0}$). Для *несимметричных p-n-переходов* справедливо неравенство $p_{p0} \gg n_{n0}$ (или $n_{n0} \gg p_{p0}$). В полупроводниковых приборах обычно существуют несимметричные *p-n*-переходы. Для обозначения несимметричных *p-n*-переходов пользуются следующими символами: p^+-n (или n^+-p).

§ 2.5. РАСПРЕДЕЛЕНИЕ НАПРЯЖЕННОСТИ ЭЛЕКТРИЧЕСКОГО ПОЛЯ И ПОТЕНЦИАЛА В ЭЛЕКТРОННО-ДЫРОЧНОМ ПЕРЕХОДЕ

При отсутствии токов распределение напряженности электрического поля и потенциала в *p-n*-переходе можно найти путем решения уравнения Пуассона

$$\nabla^2 \varphi = - \frac{q}{\epsilon \epsilon_0} (p - n + N_d - N_a). \quad (2.9)$$

Однако даже для одномерной системы решение такого уравнения с учетом носителей заряда затруднено. Поэтому обычно

влиянием носителей заряда на распределение потенциала пренебрегают, считая, что в p - n -переходе они отсутствуют, в то время как прилегающие к p - n -переходу области полупроводника остаются электрически нейтральными. Такие допущения соответствуют обратному и малому прямому напряжениям на p - n -переходе при малых токах через него.

Тогда для плоского одномерного p - n -перехода распределение потенциала определяется уравнением

$$\frac{d^2\varphi}{dx^2} = -\frac{q}{\epsilon\epsilon_0} N(x), \quad (2.10)$$

где $N(x) = N_d - N_a$ — разность концентраций доноров и акцепторов.

Необходимые для решения уравнения граничные условия можно записать с учетом следующих соображений. На одной границе p - n -перехода, где $x = -\delta_p$ (см. рис. 2.1) потенциал p -области примем за нуль, так как одно значение потенциала может быть взято произвольно. В связи с тем, что напряженность электрического поля при отсутствии заряженных поверхностей изменяется непрерывно, то на границе p - n -перехода она должна быть такой же, как и вне его. Практически в нейтральных областях полупроводника напряженность поля всегда значительно меньше, чем в p - n -переходе, и может быть принята равной нулю. На другой границе p - n -перехода при $x = \delta_n$ значение потенциала представляет собой алгебраическую сумму контактной разности потенциалов и внешнего напряжения u . Так как прямое напряжение вычитается из контактной разности потенциалов, а обратное — складывается, то $\varphi(\delta_n) = \varphi_{\text{кон}} - u$. Таким образом, граничные условия для уравнения (2.10) примут следующий вид:

$$\left. \begin{array}{l} \text{при } x = -\delta_p \quad \varphi = 0, \quad \frac{d\varphi}{dx} = 0; \\ \text{при } x = \delta_n \quad \varphi = \varphi_{\text{кон}} - u, \quad \frac{d\varphi}{dx} = 0. \end{array} \right\} \quad (2.11)$$

Количество граничных условий оказалось больше, чем порядок уравнения (4 и 2). Это связано с тем, что координаты перехода $-\delta_p$ и δ_n неизвестны и для их нахождения потребуются дополнительные условия.

Распределение напряженности

Перепишем уравнение (2.10) в следующем виде:

$$\frac{d}{dx} \left(\frac{d\varphi}{dx} \right) = -\frac{q}{\epsilon\epsilon_0} N(x).$$

Тогда

$$d \left(\frac{d\varphi}{dx} \right) = -\frac{q}{\epsilon\epsilon_0} N(x) dx.$$

Для упрощения записи в дальнейшем будем пользоваться определенными интегралами с переменными пределами. Это позволяет

не применять произвольные постоянные, если пределы интегрирования согласованы. Тогда

$$\int_0^{d\varphi/dx} d\left(\frac{d\varphi}{dx}\right) = -\frac{q}{\varepsilon\varepsilon_0} \int_{-\delta_p}^x N(x) dx,$$

или

$$E = -\frac{d\varphi}{dx} = \frac{q}{\varepsilon\varepsilon_0} \int_{-\delta_p}^x N(x) dx. \quad (2.12)$$

Соотношение (2.12) дает возможность найти распределение напряженности электрического поля или градиента потенциала в p - n -переходе с любым характером изменения в нем разности концентраций доноров и акцепторов. Если для этого уравнения учесть граничные условия (2.11), то

$$\int_{-\delta_p}^{\delta_n} N(x) dx = 0. \quad (2.13)$$

Уравнение (2.13) является условием электрической нейтральности p - n -перехода, т. е. свидетельствует о том, что суммарный неподвижный заряд ионизированных примесей с одной стороны металлургического контакта равен суммарному неподвижному заряду ионизированных примесей с другой стороны металлургического контакта.

Распределение потенциала

Для расчета распределения потенциала надо проинтегрировать градиент потенциала, т. е.

$$\varphi = \int_{-\delta_p}^x \frac{d\varphi}{dx} dx. \quad (2.14)$$

После подстановки (2.12) в (2.14)

$$\varphi = -\frac{q}{\varepsilon\varepsilon_0} \int_{-\delta_p}^x \left[\int_{-\delta_p}^x N(x) dx \right] dx.$$

Правую часть полученного выражения можно проинтегрировать по частям $\left(\int y dx = yx - \int x dy, \text{ где } y = \int_{-\delta_p}^x N(x) dx \text{ и } dy = \right.$
 $\left. = d \left[\int_{-\delta_p}^x N(x) dx \right] = N(x) dx \Big|_{-\delta_p}^x \right)$. Тогда

$$\varphi = -\frac{q}{\varepsilon\varepsilon_0} \left[x \int_{-\delta_p}^x N(x) dx - \int_{-\delta_p}^x xN(x) dx \right]. \quad (2.15)$$

Если для этого уравнения, показывающего распределение потенциала в p - n -переходе, учесть граничные условия (2.11), то полное падение потенциала на p - n -переходе

$$\Phi_{\text{кон}} - u = \frac{q}{\epsilon\epsilon_0} \int_{-\delta_p}^{\delta_n} xN(x) dx. \quad (2.16)$$

Уравнение электрической нейтральности p - n -перехода (2.13) и уравнение для полного падения напряжения на p - n -переходе (2.16) представляют собой систему двух уравнений с двумя неизвестными — пределами интегрирования. Решение этой системы уравнений дает положение границ p - n -перехода (δ_p и δ_n), которые до сих пор оставались неопределенными.

Методы решения полученных уравнений

Численное решение. Численные методы решения уравнений для распределения напряженности электрического поля и потенциала в p - n -переходе целесообразно применять при сложном распределении концентрации примеси. Для такого решения используют численный метод интегрирования. Решение начинают при произвольно взятой координате $x = -\delta_p$. Выполняют численное определение градиента потенциала $d\phi/dx$. Этот интеграл сначала возрастает по абсолютному значению, а затем падает, проходит через нуль и изменяет знак. Значение координаты, при которой выполняется условие электрической нейтральности, определяет вторую границу p - n -перехода.

Однако при этом остается неизвестным значение напряжения. Для его определения проводят интегрирование градиента потенциала $d\phi/dx$. Значение этого интеграла при $x = \delta_n$ и дает значение напряжения на p - n -переходе $\Phi_{\text{кон}} - u$. Таким образом, напряжение на p - n -переходе при численном методе решения получается не аргументом, а функцией. Если же требуется выполнить расчет для заданного напряжения, то производят корректировку координаты $-\delta_p$.

В ходе такого решения распределения напряженности электрического поля и потенциала в p - n -переходе получают как промежуточные значения интегралов.

Аналитическое решение. Аналитическое решение возможно в тех случаях, когда распределение концентрации примесей является интегрируемой аналитически функцией и полученная система уравнений имеет аналитическое решение. В частности, аналитическое решение возможно для резкого p - n -перехода, для плавного p - n -перехода с линейным распределением концентрации примесей и других плавных p - n -переходов.

§ 2.6. АНАЛИТИЧЕСКИЙ РАСЧЕТ РЕЗКОГО ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

В соответствии с определением резкого p - n -перехода (см. § 2.4) можно считать, что концентрация примесей изменяется скачком на металлургическом контакте, а в остальных частях перехода остается постоянной, т. е.

$$\text{и} \quad \left. \begin{aligned} N(x) &= -N_{ap} & \text{при} & \quad -\delta_p < x < 0 \\ N(x) &= N_{dn} & \text{при} & \quad \delta_n > x > 0. \end{aligned} \right\} \quad (2.17)$$

Распределение напряженности

При $x < 0$ в соответствии с уравнением (2.12) и с условиями (2.17) градиент потенциала

$$\left(\frac{d\varphi}{dx}\right)_p = \frac{q}{\varepsilon\varepsilon_0} N_{ap} (x + \delta_p). \quad (2.18)$$

При $x > 0$ следует учесть, что подынтегральная функция в (2.12) имеет разрыв. Поэтому целесообразно записать

$$\left(\frac{d\varphi}{dx}\right)_n = -\frac{q}{\varepsilon\varepsilon_0} \left[\int_{-\delta_p}^0 N(x) dx + \int_0^x N(x) dx \right] = \frac{q}{\varepsilon\varepsilon_0} (N_{ap}\delta_p - N_{dn}x). \quad (2.19)$$

Таким образом, для резкого p - n -перехода получается линейная зависимость градиента потенциала или напряженности электрического поля в p - n -переходе, наибольшее значение градиента получается на металлургическом контакте (рис. 2.6).

Распределение потенциала

Для расчета распределения потенциала надо проинтегрировать выражения для его градиента, т. е. (2.18) и (2.19). Тогда

$$\varphi_p = \frac{q}{\varepsilon\varepsilon_0} N_{ap} \int_{-\delta_p}^x (x + \delta_p) dx = \frac{q}{2\varepsilon\varepsilon_0} N_{ap} (x + \delta_p)^2 \quad (2.20)$$

и

$$\begin{aligned} \varphi_n &= \int_{-\delta_p}^0 \left(\frac{d\varphi}{dx}\right)_p dx + \int_0^x \left(\frac{d\varphi}{dx}\right)_n dx = \\ &= \frac{q}{2\varepsilon\varepsilon_0} N_{ap}\delta_p^2 + \frac{q}{\varepsilon\varepsilon_0} N_{ap}\delta_p x - \frac{q}{2\varepsilon\varepsilon_0} N_{dn}x^2. \end{aligned} \quad (2.21)$$

Таким образом, потенциал в резком p - n -переходе изменяется с координатой по квадратичной параболе (рис. 2.6).

Отношение δ_p/δ_n

Полученные выражения (2.18) — (2.21) содержат пока неизвестные значения границ p - n -перехода. Воспользовавшись условием электрической нейтральности (2.13), для резкого p - n -перехода получим

$$\int_{-\delta_p}^0 (-N_{ap}) dx + \int_0^{\delta_n} N_{dn} dx = 0;$$

$$\delta_p/\delta_n = N_{dn}/N_{ap}, \quad (2.22)$$

т. е. металлургический контакт делит резкий p - n -переход на части, толщины которых обратно пропорциональны концентрациям примесей в соответствующих областях. Следовательно, при несимметричном легировании областей (что чаще всего и бывает) p - n -переход расположен преимущественно в области с меньшей концентрацией примесей.

Определение границ и толщины p - n -перехода

Для определения границ резкого p - n -перехода целесообразно ввести полную толщину перехода

$$\delta = \delta_p + \delta_n.$$

Полную толщину p - n -перехода можно выразить через толщину одной его части, учитывая (2.22). Таким образом,

$$\delta = \delta_p + \delta_p \frac{N_{ap}}{N_{dn}} = \delta_p \frac{N_{ap} + N_{dn}}{N_{dn}}; \quad \delta = \delta_n \frac{N_{dn}}{N_{ap}} + \delta_n = \delta_n \frac{N_{ap} + N_{dn}}{N_{ap}}.$$

Отсюда

$$\delta_p = \delta \frac{N_{dn}}{N_{ap} + N_{dn}} \quad \text{и} \quad \delta_n = \delta \frac{N_{ap}}{N_{ap} + N_{dn}}. \quad (2.23)$$

Найдя полную толщину p - n -перехода, можно с помощью соотношений (2.23) определить координаты его границ. Для получения выражения полной толщины p - n -перехода запишем полное падение потенциала после интегрирования (2.16) с учетом условий (2.17):

$$\varphi_{\text{кон}} - u = \frac{q}{2\epsilon\epsilon_0} (N_{ap}\delta_p^2 + N_{dn}\delta_n^2). \quad (2.24)$$

Подставив в (2.24) соотношения (2.23), получим

$$\varphi_{\text{кон}} - u = \frac{qN_{ap}N_{dn}\delta^2}{2\epsilon\epsilon_0(N_{ap} + N_{dn})}. \quad (2.25)$$

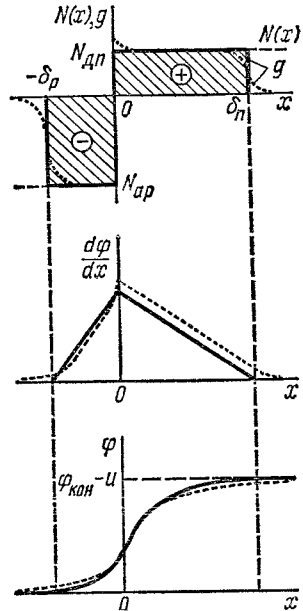


Рис. 2.6. Распределение концентрации примесей $N(x)$, плотности объемного заряда g , градиента потенциала $d\varphi/dx$ и потенциала φ в резком p - n -переходе без учета носителей заряда (—) и с учетом носителей заряда (...)

Отсюда полная толщина резкого p - n -перехода

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0}{q} \frac{N_{ap} + N_{dn}}{N_{ap}N_{dn}}} (\Phi_{\text{кон}} - u). \quad (2.26)$$

Реальные резкие p - n -переходы (например, сплавные) обычно бывают несимметричными, т. е.

$$N_{ap} \gg N_{dn} \text{ или } N_{dn} \gg N_{ap}.$$

Поэтому толщина резкого несимметричного p - n -перехода

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0}{q} \frac{\Phi_{\text{кон}} - u}{N}}, \quad (2.27)$$

где N — концентрация примесей в слаболегированной области.

Хотя в подкоренном выражении и стоит разность $\Phi_{\text{кон}} - u$, отрицательные значения этой разности не имеют физического смысла. При обратных напряжениях $u < 0$ и подкоренное выражение всегда положительно. При прямых напряжениях $u > 0$, но оно не может превзойти $\Phi_{\text{кон}}$, так как при $u \rightarrow \Phi_{\text{кон}}$ p - n -переход исчезает.

Влияние носителей заряда

Подвижные носители заряда могут оказывать влияние на распределение градиента потенциала и потенциала в резком p - n -переходе вблизи его границ и вблизи металлургического контакта.

Вблизи границ p - n -перехода с учетом предположений, принятых для расчетов, плотность объемного заряда некомпенсированных примесей изменяется скачком от нуля за пределами p - n -перехода до значения, отличающегося только множителем q от $N(x)$, т. е. в пределах p - n -перехода график распределения плотности объемного заряда отличается от графика распределения концентрации примесей только масштабом (см. рис. 2.6). В действительности же плотность объемного заряда не может изменяться скачком, так как это соответствовало бы бесконечному градиенту концентрации носителей заряда и наличию бесконечно больших диффузионных токов. Поэтому границы p - n -перехода должны быть размытыми, а в распределении градиента потенциала и потенциала должны быть «хвосты» (рис. 2.6).

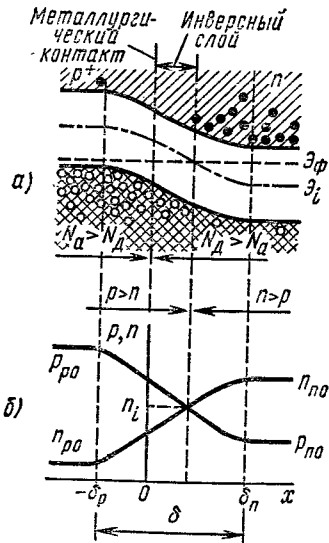


Рис. 2.7. Энергетическая диаграмма несимметричного резкого p - n -перехода в состоянии равновесия с инверсным слоем в слаболегированной области перехода (а) и распределение носителей заряда в таком переходе (б)

Вблизи металлургического контакта несимметричного резкого p - n -перехода концентрация носителей заряда также может существенно отличаться от собственной, что иллюстрируется энергетической диаграммой p - n -перехода в состоянии равновесия (рис. 2.7). При несимметрии p - n -перехода металлургический контакт оказывается ближе к сильнолегированной области, а падение потенциала там меньше. Уровень Ферми пересекает середину запрещенной зоны не на металлургическом контакте, а в слаболегированной части p - n -перехода. Тип преобладающих носителей заряда определяется взаимным положением уровня Ферми и середины запрещенной зоны: где уровень Ферми выше середины запрещенной зоны — преобладают электроны, где ниже — дырки. Следовательно, граница смены типа носителей заряда не совпадает с металлургическим контактом, т. е. в некотором слое слаболегированной части p - n -перехода тип преобладающих (основных) носителей заряда не соответствует типу примесей. Такой слой, как указывалось (см. § 1.12), называют инверсным.

§ 2.7. АНАЛИТИЧЕСКИЙ РАСЧЕТ ПЛАВНОГО ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА С ЛИНЕЙНЫМ РАСПРЕДЕЛЕНИЕМ КОНЦЕНТРАЦИИ ПРИМЕСЕЙ

Линейный закон распределения примесей, т. е.

$$N(x) = ax \quad (2.28)$$

(где a — градиент концентрации примесей), может быть принят при очень малой толщине p - n -перехода по сравнению с областью, где концентрация примесей переменна.

Распределение напряженности

Решение уравнения (2.12) с учетом (2.28) дает зависимость градиента потенциала в плавном p - n -переходе с линейным распределением концентрации примесей:

$$\frac{d\varphi}{dx} = - \frac{q}{2\epsilon\epsilon_0} ax^2 \Big|_{-\delta_p}^x = - \frac{qa}{2\epsilon\epsilon_0} (x^2 - \delta_p^2). \quad (2.29)$$

В данном случае градиент потенциала изменяется с изменением координаты по квадратичной параболе (рис. 2.8).

Распределение потенциала

Для определения потенциала надо проинтегрировать градиент потенциала (2.29):

$$\varphi = - \frac{qa}{2\epsilon\epsilon_0} \int_{-\delta_p}^x (x^2 - \delta_p^2) dx = - \frac{qa}{\epsilon\epsilon_0} \left(\frac{x^3}{6} - \frac{\delta_p^2 x}{2} - \frac{\delta_p^3}{3} \right). \quad (2.30)$$

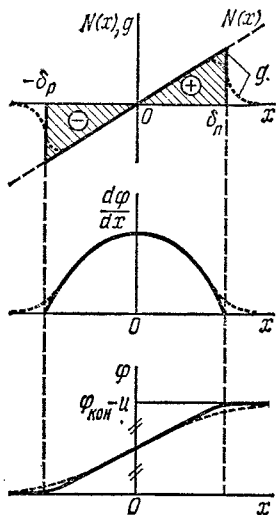


Рис. 2.8. Распределение концентрации примесей $N(x)$, плотности объемного заряда g , градиента потенциала $d\varphi/dx$ и потенциала φ в плавном электронно-дырочном переходе без учета носителей заряда (—) и с учетом носителей заряда (...)

Таким образом, потенциал в плавном p - n -переходе с линейным распределением примесей изменяется с изменением координаты по кубической параболе (рис. 2.8).

Определение границ и толщины p - n -перехода

Решение уравнения электрической нейтральности (2.13) в данном случае приводит к соотношению

$$\delta_p = \delta_n = \delta/2, \quad (2.31)$$

т. е. рассматриваемый переход является симметричным.

Из уравнения (2.16) с учетом (2.28) можно получить полное падение потенциала на p - n -переходе:

$$\begin{aligned} \varphi_{\text{кон}} - u &= \varphi(\delta_n) - \varphi(\delta_p) = \frac{2qa}{3\epsilon\epsilon_0} \delta_p^3 = \\ &= \frac{qa}{12\epsilon\epsilon_0} \delta^3. \end{aligned} \quad (2.32)$$

Отсюда толщина плавного p - n -перехода с линейным распределением примесей

$$\delta = \sqrt[3]{\frac{12\epsilon\epsilon_0(\varphi_{\text{кон}} - u)}{qa}}. \quad (2.33)$$

Влияние носителей заряда

В плавном p - n -переходе с линейным распределением концентрации примесей из-за симметрии перехода инверсный слой не образуется, если нет существенной несимметрии плотности энергетических состояний в разрешенных энергетических зонах, т. е. если нет резкого отличия эффективных масс носителей заряда разных знаков.

У границ плавного p - n -перехода влияние носителей заряда аналогично влиянию носителей у резкого p - n -перехода.

§ 2.8. БАРЬЕРНАЯ ЕМКОСТЬ ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА Барьерная емкость как проявление токов смещения

Барьерная емкость p - n -перехода проявляется при приложении к p - n -переходу изменяющегося во времени напряжения. При этом через p - n -переход проходит ток. Та доля тока, которая не связана с движением носителей заряда через p - n -переход, и определяет барьерную емкость.

Поэтому барьерная емкость должна быть связана с прохождением токов смещения. Для одномерного плоского p - n -перехода ток смещения, одинаковый во всех его сечениях,

$$i = SJ_{\text{см}} = S \frac{d}{dt} (\epsilon \epsilon_0 E),$$

где S — площадь p - n -перехода.

Значение тока смещения можно связать с изменением объемного заряда. Для этого выделим мысленно в p - n -переходе объем в виде цилиндра (или призмы), образующие которого параллельны оси x — направлению электрического поля (рис. 2.9). Одно основание цилиндра пусть лежит за пределами p - n -перехода, а другое — внутри его. Тогда согласно теореме Остроградского — Гаусса можно определить поток вектора электрической индукции через поверхность, ограничивающую выделенный объем. Этот поток проходит только через одно основание цилиндра, так как боковые его поверхности параллельны электрическому полю, а второе основание лежит в области, где поле отсутствует. Тогда

$$\epsilon \epsilon_0 ES = Q,$$

где Q — заряд ионизированных примесей.

Ток смещения можно записать теперь таким образом:

$$i = \frac{dQ}{dt} = \frac{dQ}{dU} \frac{dU}{dt}.$$

Сравнив последнее выражение с обычным выражением для тока через емкость, т. е. с

$$i_C = C \frac{dU}{dt},$$

получаем, что в качестве барьерной емкости следует взять величину

$$C_{\text{бар}} = \left| \frac{dQ}{dU} \right|. \quad (2.34)$$

Абсолютное значение этого отношения взято потому, что объемный заряд в p - n -переходе может быть положительным и отрицательным, а правило знаков для напряжения выбрано произвольно.

Таким образом, барьерная емкость связана с током смещения (как и обычная емкость).

Общее соотношение для барьерной емкости электронно-дырочного перехода

Исходя из определения барьерной емкости (2.34), можно получить общую формулу для барьерной емкости плоского p - n -перехода. Объемный заряд ионизированных примесей в цилиндре,

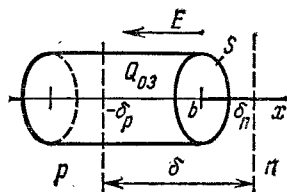


Рис. 2.9. К выводу выражения для барьерной емкости p - n -перехода

выделенном в p - n -переходе,

$$Q = qS \int_{-\delta_p}^b N(x) dx.$$

Дифференциал этого объемного заряда можно определить путем дифференцирования по единственной переменной — нижнему пределу интегрирования:

$$dQ = -qSN(-\delta_p) d(-\delta_p). \quad (2.35)$$

Приращение напряжения или дифференциал падения напряжения на p - n -переходе можно найти, продифференцировав выражение (2.16). Но при этом следует учесть, что переменными являются оба предела интегрирования в выражении (2.16). Поэтому целесообразно разбить интеграл на два, тогда каждый из них будет иметь по одному переменному пределу:

$$\varphi_{\text{кон}} - U = \frac{q}{\varepsilon\varepsilon_0} \int_{-\delta_p}^{\delta_n} xN(x) dx = \frac{q}{\varepsilon\varepsilon_0} \left[\int_{-\delta_p}^0 xN(x) dx + \int_0^{\delta_n} xN(x) dx \right]$$

и

$$dU = \frac{q}{\varepsilon\varepsilon_0} [\delta_p N(-\delta_p) d(-\delta_p) + \delta_n N(\delta_n) d\delta_n]. \quad (2.36)$$

Теперь, взяв отношение дифференциалов (2.35) и (2.36), получим

$$C_{\text{бар}} = \left| \frac{qSN(-\delta_p) d\delta_p}{\frac{q}{\varepsilon\varepsilon_0} [\delta_p N(-\delta_p) d(-\delta_p) + \delta_n N(\delta_n) d\delta_n]} \right|.$$

Для преобразования последнего выражения к более наглядному виду продифференцируем условие электрической нейтральности p - n -перехода (2.13):

$$N(\delta_n) d\delta_n - N(-\delta_p) d(-\delta_p) = 0,$$

т. е.

$$N(\delta_n) d\delta_n = N(-\delta_p) d(-\delta_p).$$

Учтя, что $\delta = \delta_p + \delta_n$, получим

$$C_{\text{бар}} = \frac{\varepsilon\varepsilon_0 S}{\delta}. \quad (2.37)$$

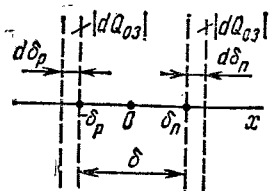


Рис. 2.10. Изменение толщины p - n -перехода и объемного заряда при изменении напряжения на p - n -переходе

Таким образом, барьерная емкость плоского одномерного p - n -перехода может быть рассчитана по формуле плоского конденсатора. Такой результат не является

очевидным, так как распределение зарядов в плоском конденсаторе и в электронно-дырочном переходе не одинаково. Причина совпадения формул — в характере изменения заряда p - n -перехода: при изменении напряжения на p - n -переходе заряд изменяется потому, что сдвигаются границы p - n -перехода. Заряды, обуславливающие барьерную емкость, сосредоточены в двух тонких слоях, расположенных на расстоянии δ один от другого (рис. 2.10), что очень напоминает поверхностные заряды на металлических обкладках конденсатора.

Частные соотношения для барьерной емкости различных электронно-дырочных переходов

С помощью выражения (2.37) можно определить барьерную емкость на основе результатов расчета толщины p - n -перехода. Поэтому для резкого p - n -перехода, если учесть (2.26),

$$C_{\text{бар}} = S \sqrt{\frac{q\epsilon\epsilon_0 N_{ap} N_{dn}}{2(N_{ap} + N_{dn})(\varphi_{\text{кон}} - U)}}. \quad (2.33)$$

Для резкого несимметричного p - n -перехода с учетом (2.27)

$$C_{\text{бар}} = S \sqrt{\frac{q\epsilon\epsilon_0 N}{2(\varphi_{\text{кон}} - U)}}, \quad (2.39)$$

где N — концентрация примесей в слаболегированной области.

Для плавного p - n -перехода с линейным распределением концентрации примесей с учетом (2.33)

$$C_{\text{бар}} = S \sqrt[3]{\frac{(\epsilon\epsilon_0)^2 qa}{12(\varphi_{\text{кон}} - U)}}. \quad (2.40)$$

Как видно из данных результатов, при разных распределениях примесей получаются разные вольт-фарадные характеристики электронно-дырочных переходов. Это дает возможность оценивать характер распределения примесей в различных p - n -переходах. Часто применяют также графический метод. Для резкого p - n -перехода вольт-фарадная характеристика оказывается прямой в координатах $1/C_{\text{бар}}^2$ от U , а для плавного с линейным распределением примесей — в координатах $1/C_{\text{бар}}^3$ от U (рис. 2.11).

Если экспериментальные точки ложатся на прямые в указанных системах координат, то это служит подтверждением (но не доказательством) принятого при построении характера распределения примесей. Однако вольт-фарадные характеристики неоднозначно связаны с распределением примесей в p - n -переходе, т. е. одинаковым вольт-фарадным характеристикам могут соответствовать разные распределения примесей. Например, распределения, показанные на рис. 2.12, дают прямые вольт-фарадные характеристики в координатах $1/C_{\text{бар}}^3$ от U . Поэтому при анализе вольт-фарадных характеристик необходимо принимать во внимание дополнительные сведения о технологии и т. п.

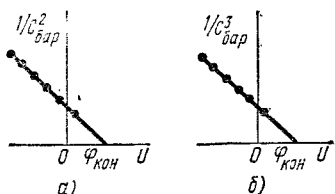


Рис. 2.11. Зависимость барьерной емкости резкого (а) и плавного (б) p - n -переходов от постоянного смещения на переходе

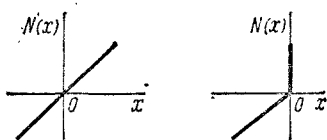


Рис. 2.12. Различные распределения примесей, обеспечивающие одинаковые вольт-фарадные характеристики p - n -переходов

Кроме того, вольт-фарадные характеристики дают возможность определять величину контактной разности потенциалов на p - n -переходе (или высоту потенциального барьера). Если при построении вольт-фарадной характеристики получилась прямая линия, то отрезок, отсекаемый ею на горизонтальной оси, дает значение контактной разности потенциалов (см. рис. 2.11).

§ 2.9. КОНТАКТ ПОЛУПРОВОДНИКОВ С ОДНИМ ТИПОМ ЭЛЕКТРОПРОВОДНОСТИ

Контакт полупроводников с одной электропроводностью, но с разной удельной проводимостью обозначают как p^+ - p - или n^+ - n -переход, причем значком «+» отмечают область полупроводника с большей удельной проводимостью, т. е. с большей концентрацией примесей (рис. 2.13).

Вблизи такого контакта происходят процессы, аналогичные процессам в электронно-дырочном переходе, т. е. носители из области с большей концентрацией диффундируют в область с меньшей концентрацией. При диффузии в сильнолегированной области нарушается компенсация заряда ионизированных примесей, а в слаболегированной — создается избыток основных носителей заряда (рис. 2.13, в). Таким образом, на контакте двух полупроводников с одним типом электропроводности, но с разной удельной проводимостью также образуется область объемного заряда, диффузионное поле и контактная разность потенциалов. В отличие от p - n -перехода в данном случае в слаболегированной области объемный заряд образован избыточной концентрацией основных носителей заряда.

Контактная разность потенциалов для таких переходов определяется соотношениями, аналогичными (2.3), т. е. для контакта типа p^+ - p

$$\Phi_{\text{кон}} = \frac{kT}{q} \ln \frac{p_{p^+0}}{p_{p0}}; \quad (2.41)$$

для контакта типа n^+ - n

$$\Phi_{\text{кон}} = \frac{kT}{q} \ln \frac{n_{n^+0}}{n_{n0}}. \quad (2.42)$$

Распределение носителей заряда вблизи контакта полупроводников с одним типом электропроводности (рис. 2.13, в) показывает, что в данном случае отсутствует слой с меньшей концентрацией по сравнению с концентрацией носителей в слаболегированной области, т. е. отсутствует высокоомный слой. Из-за этого при прохождении тока на таком контакте падает сравнительно небольшое напряжение, высота его потенциального барьера практически не изменяется от величины и направления тока. Следовательно, контакт двух полупроводников с одной электропроводностью не обладает выпрямляющими свойствами и имеет малое сопротивление по сравнению с сопротивлением слаболегированной области.

Другой особенностью контакта двух полупроводников с одной электропроводностью является отсутствие инжекции неосновных носителей заряда в слаболегированную высокоомную область. Действительно, если внешнее напряжение приложено плюсом к высокоомной n -области $n^+ - n$ -перехода, что аналогично прямому включению $p - n$ -перехода, то при этом из n^+ -области в n -область вводятся электроны, которые являются основными носителями. При противоположной полярности дырочный ток из n^+ -области в n -область аналогичен обратному току через $p - n$ -переход. Однако из-за ничтожно малой концентрации неосновных носителей заряда в сильнолегированной n^+ -области [см. (1.19)] инжекция дырок в высокоомную n -область также оказывается ничтожно малой.

Невыпрямляющие и неинжектирующие контакты широко используются в полупроводниковых приборах наряду с выпрямляющими и инжектирующими.

Однако из-за существования на контакте двух полупроводников с одной электропроводностью потенциального барьера для неосновных носителей заряда,двигающихся из слаболегированной области к контакту, может проис-

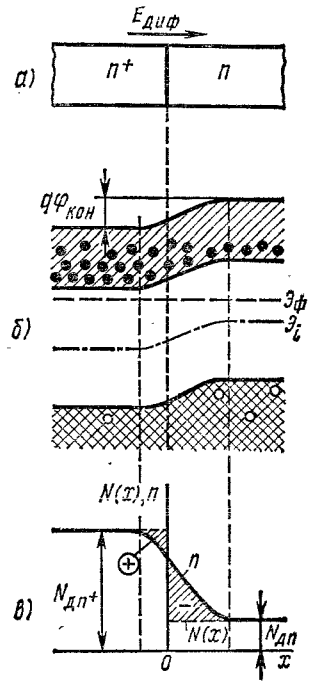


Рис. 2.13. Контакт полупроводников с одним типом электропроводности:

а — структура контакта; б — энергетическая диаграмма контакта; в — зависимости концентрации примесей и концентрации носителей заряда от координаты

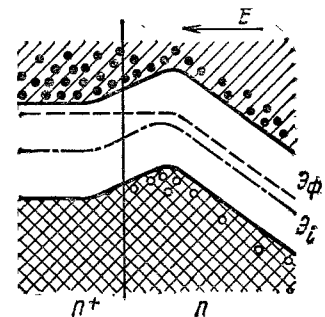


Рис. 2.14. Накопление неосновных носителей заряда (дырок) вблизи контакта двух полупроводников с электропроводностью n -типа при наличии внешнего электрического поля

ходить накопление этих неосновных носителей вблизи контакта при определенной полярности внешнего напряжения (рис. 2.14). При приложении внешнего напряжения оно падает в основном в объеме слабоэнергетических уровней и зон, приводящий к образованию потенциальных ям для неосновных носителей заряда. Эффект накопления неосновных носителей заряда и последующего их рассасывания — эффект инерционный. Поэтому он может ухудшать быстродействие полупроводниковых приборов.

§ 2.10. КОНТАКТ МЕТАЛЛ — ПОЛУПРОВОДНИК

При идеальном контакте металла с полупроводником (т. е. при отсутствии каких-либо промежуточных слоев, отличающихся химическим составом) происходят явления, связанные с разницей в работе выхода электронов в металле A_m и в полупроводнике A_n . Под *работой выхода электронов* будем понимать энергию, необходимую для перевода электрона с уровня Ферми на потолок верхней свободной зоны (без удаления электрона в вакуум на бесконечное расстояние от поверхности полупроводника).

При образовании идеального контакта металла с полупроводником возникает диффузия электронов из материала с меньшей работой выхода в материал с большей работой выхода. Перераспределение зарядов вызывает появление электрического поля и контактной разности потенциалов:

$$\Phi_{\text{кон}} = (A_m - A_n)/q. \quad (2.43)$$

Это электрическое поле сосредоточено практически только в полупроводнике, так как из-за большой концентрации носителей заряда в металле их перераспределение происходит в очень тонком слое, сравнимым с межатомным расстоянием.

В зависимости от типа электропроводности полупроводника и от соотношения работ выхода в полупроводнике может возникать обедненный, инверсный или обогащенный слой (рис. 2.15). Если работа выхода в металле меньше работы выхода в полупроводнике ($A_m < A_n$), то электроны с большей вероятностью будут переходить из металла в полупроводник. Это приводит к образованию в полупроводнике обедненного слоя, если полупроводник p -типа (рис. 2.15, а), или даже инверсного слоя, если $A_m \ll A_n$ (рис. 2.15, б). Если полупроводник n -типа, то образуется обогащенный слой (рис. 2.15, в).

При противоположном соотношении работ выхода ($A_m > A_n$) в полупроводнике n -типа получается обедненный или инверсный слой, а в дырочном — обогащенный.

В обедненных слоях объемный заряд формируется в результате нарушения компенсации заряда ионизированных примесей основными носителями, а в обогащенных — из-за накопления основных носителей заряда. Обогащенный слой обуславливает малое сопротивление приконтактной области полупроводника по сравнению

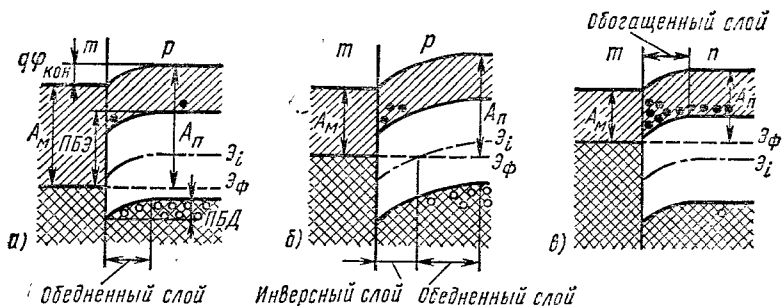


Рис. 2.15. Образование обедненного (а), инверсного (б) и обогащенного (в) слоев в полупроводнике вблизи контакта с металлом при работе выхода в металле меньше, чем в полупроводнике

с сопротивлением объема полупроводника. Поэтому такой контакт не обладает выпрямляющими свойствами. При наличии обедненного или инверсного слоя контакт металл — полупроводник обладает выпрямляющими свойствами, так как внешнее напряжение, падая в основном на высокоомном контакте, будет изменять высоту его потенциального барьера, т. е. условия для прохождения носителей заряда через контакт.

Характерной особенностью контакта металл — полупроводник в отличие от p - n -перехода является разная высота потенциальных барьеров для электронов и дырок. В результате контакты металл — полупроводник могут быть неинжектирующими. Рассмотрим рис. 2.16. При включении такого контакта в прямом направлении (рис. 2.16, б) высота потенциального барьера для дырок (ПБД) в полупроводнике понижается, в результате чего дырки будут переходить из полупроводника в металл. Чем больше прямое напряжение, тем больше вероятность такого перехода дырок. Однако при этом высота потенциального барьера для электронов (ПБЭ), которые могут двигаться из металла в полупроводник, изменяется относительно мало, оставаясь еще большой. Поэтому переход электронов из металла в полупроводник маловероятен, соответственно отсутствует и инжекция неосновных носителей заряда в полупроводник.

При другой полярности внешнего напряжения (при обратном напряжении) потенциальный барьер для дырок в полупроводнике повышается (рис. 2.16, в) и их движение через контакт прекращается. Для неосновных носителей заряда (для электронов в данном случае) поле вблизи контакта оказывается ускоряющим. Поэтому, проходя через контакт, неосновные носители заряда образуют обратный ток, который будет мал из-за малой концентрации неосновных носителей.

Если разница в работах выхода велика, то в полупроводнике образуется инверсный слой (см. рис. 2.15, б). В этом случае при прямых напряжениях через такой контакт будет происходить инжекция неосновных носителей заряда в полупроводник.

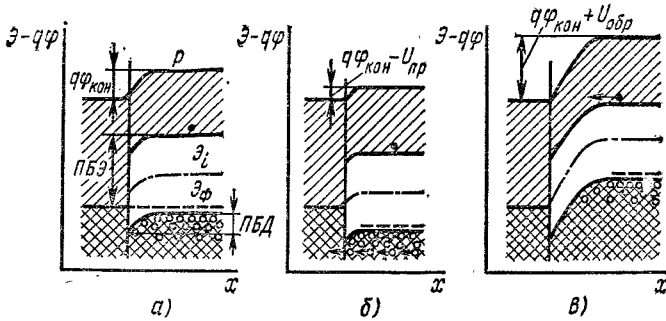


Рис. 2.16. Изменение высоты потенциального барьера на выпрямляющем неинжектирующем контакте металл — полупроводник при изменении внешнего напряжения:

а — внешнее напряжение отсутствует; *б* — внешнее напряжение прямое; *в* — внешнее напряжение обратное

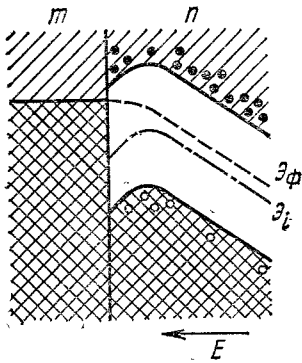


Рис. 2.17. Накопление неосновных носителей (дырок) вблизи невыпрямляющего контакта металл — полупроводник при наличии внешнего электрического поля

В невыпрямляющих контактах металл — полупроводник с обогащенным слоем может происходить накопление неосновных носителей заряда из-за образования потенциальных ям для этих неосновных носителей в приконтактной области полупроводника (рис. 2.17). Такое явление, как уже отмечалось в § 2.9, может влиять на быстродействие полупроводниковых приборов. Для устранения этого явления необходимо устранить потенциальный барьер на контакте металл — полупроводник путем подбора контактирующих пар материалов с одинаковыми работами выхода. Однако практически это неосуществимо из-за ограниченного набора материалов и необходимости подбирать металл заново для каждой концентрации примесей в полупроводнике и для каждой температуры.

Для исключения эффекта накопления неосновных носителей заряда в полупроводнике около контакта можно провести дополнительное легирование приконтактной области полупроводника. Потенциальный барьер при этом остается, но его толщина будет очень малой из-за сильного легирования приконтактной области полупроводника. Малая толщина потенциального барьера обеспечивает возможность туннелирования неосновных носителей заряда из потенциальной ямы в полупроводнике в металл.

§ 2.11. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называют переходный слой с существующим там диффузионным электрическим полем между двумя

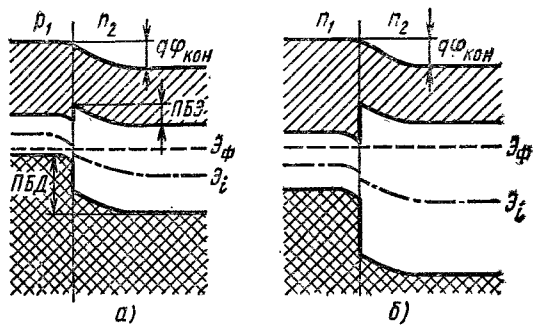


Рис. 2.18. Энергетические диаграммы гетеропереходов:

а — выпрямляющий контакт между полупроводниками *p*- и *n*-типов с преимущественной инжекцией электронов в узкозонный полупроводник; *б* — выпрямляющий контакт между полупроводниками *n*-типа — контакт без инжекции неосновных носителей заряда

различными полупроводниками, которые отличаются друг от друга шириной энергетических зон. При образовании таких контактов происходит перераспределение носителей заряда, что приводит к появлению контактной разности потенциалов и к выравниванию уровней Ферми (рис. 2.18). Все остальные энергетические уровни после возникновения контактной разности потенциалов должны соответственно изогнуться. При этом энергетический уровень потолка верхней свободной зоны должен быть непрерывным. Обычно энергетический уровень потолка верхней свободной зоны является энергетическим уровнем потолка зоны проводимости, так как свободные энергетические зоны перекрываются друг с другом.

Так как ширина энергетических зон различных полупроводников различна, то на металлургическом контакте гетероперехода получается обычно разрыв дна зоны проводимости. Величина разрыва дна зоны проводимости определяется различием энергий сродства к электрону двух различных полупроводников (*энергия сродства к электрону* есть разница энергий потолка верхней свободной зоны и дна зоны проводимости).

Величина разрыва потолка валентной зоны зависит как от разницы энергий сродства, так и от различия ширины запрещенных зон контактирующих полупроводников.

В результате разрывов дна зоны проводимости и потолка валентной зоны высота потенциальных барьеров для электронов и дырок в гетеропереходе оказывается различной. Это является особенностью гетеропереходов, обуславливающей специфические свойства гетеропереходов в отличие от гомопереходов.

Каждый из полупроводников, образующих гетеропереход, может иметь различный тип электропроводности. Поэтому для каждой пары полупроводников в принципе можно осуществить четыре комбинации структур: p_1-n_2 ; n_1-n_2 ; n_1-p_2 и p_1-p_2 .

Из-за различия по высоте потенциальных барьеров для электронов и для дырок прямой ток через гетеропереход связан обычно с движением носителей заряда только одного типа. Поэтому гетеропереходы могут быть как инжектирующими (рис. 2.18, *а*), так и неинжектирующими (рис. 2.18, *б*).

Однако для того чтобы получить гетеропереходы с достаточно хорошими характеристиками, кристаллическая решетка одного

полупроводника должна с минимальным количеством нарушений переходить в кристаллическую решетку другого полупроводника, т. е. должны быть близки параметры кристаллических решеток полупроводников, образующих гетеропереход. Дефекты, возникающие на границе раздела двух полупроводников, создают условия для рекомбинации и генерации носителей заряда в гетеропереходе — рекомбинационные ловушки. В результате механизмы прохождения тока через реальный и идеальный гетеропереходы могут отличаться, что не даст возможности использовать специфические свойства гетероперехода.

Выращивание слоев одного полупроводника на пластинах другого осуществляется при относительно высоких температурах. При этом может происходить диффузия примесей из одного полупроводника в другой, изменение исходных типов электропроводности, искажение теоретических энергетических диаграмм и соответственно изменение свойств гетероперехода.

Все рассмотренные ранее виды электрических переходов (электронно-дырочный переход, контакт полупроводников с одинаковыми типами электропроводности, контакт металл — полупроводник) можно считать частными случаями более общего вида контактов — гетеропереходов.

§ 2.12. СВОЙСТВА И ПАРАМЕТРЫ НЕВЫПРЯМЛЯЮЩИХ КОНТАКТОВ

Требования к невыпрямляющим контактам

Невыпрямляющие контакты имеют очень большое значение в полупроводниковых приборах и при проведении исследований полупроводников. Основное назначение невыпрямляющих контактов — электрическое соединение полупроводника с металлическими токопроводящими частями полупроводникового прибора. Невыпрямляющих контактов в полупроводниковых приборах больше, чем выпрямляющих. Случай отказов и производственного брака полупроводниковых приборов из-за невыпрямляющих контактов довольно часты. При разработке технологии полупроводниковых приборов создание совершенных невыпрямляющих контактов нередко требует больших усилий, чем создание *p-n*-переходов.

Несмотря на это, теория невыпрямляющих контактов разработана слабее, чем теория *p-n*-переходов; а создание невыпрямляющих контактов часто основано на эксперименте.

Невыпрямляющий контакт меньше влияет на свойства и параметры полупроводникового прибора при следующих условиях:

- 1) если отсутствует инжекция неосновных носителей заряда через невыпрямляющий контакт в полупроводник;
- 2) при минимально возможном падении напряжения на невыпрямляющем контакте, т. е. при минимальном сопротивлении;
- 3) если вольт-амперная характеристика невыпрямляющего контакта линейна (это требование не всегда существенно).

Параметры невыпрямляющих контактов

Для того чтобы можно было установить, насколько контакт удовлетворяет предъявленным к нему требованиям, и чтобы можно было сравнить разные контакты, надо ввести количественные параметры, характеризующие невыпрямляющий контакт.

Скорость рекомбинации на контакте. Скорость рекомбинации на контакте показывает, насколько концентрация носителей заряда вблизи контакта может отклоняться от равновесной концентрации. Вводится скорость рекомбинации на контакте аналогично скорости поверхностной рекомбинации [см. (1.36)] — как отношение плотности потока носителей заряда через контакт к избыточной концентрации этих носителей у контакта, т. е.

$$s_p = \Phi_p / (p_{гр} - p_0).$$

Размерность скорости рекомбинации на контакте — скорость, так как плотность потока — это произведение концентрации на скорость ($\Phi_p = p v_p$). Очевидно, что чем больше скорость рекомбинации, тем меньше при данном потоке носителей заряда отклонение концентрации носителей от равновесной, тем выше качество невыпрямляющего контакта.

С увеличением плотности потока носителей заряда к реальному контакту неизбежно будет возрастать и граничная концентрация этих носителей у контакта. При большой плотности потока граничная концентрация носителей будет существенно превышать равновесную концентрацию тех же носителей ($p_{гр} \gg p_0$). Тогда

$$s_p = \frac{\Phi_p}{p_{гр} - p_0} = \frac{p_{гр} v_p}{p_{гр} - p_0} \approx v_p.$$

Таким образом, скорость рекомбинации на контакте не превосходит скорости движения носителей заряда. Если перенос носителей заряда к контакту связан с диффузией, то скорость движения, в свою очередь, не может превзойти тепловой скорости. Если движение носителей заряда связано с дрейфом, то их скорость тоже ограничена величиной v_{\max} , имеющей порядок тепловой скорости (см. § 1.10.)

Итак, скорость рекомбинации на контакте не превосходит тепловой скорости движения носителей заряда. Это имеет практическое значение, так как нет смысла пытаться совершенствовать контакты, скорость рекомбинации на которых приближается к максимально возможной.

На рис. 2.19 показаны зависимости граничной концентрации носителей заряда около идеального контакта с бесконечно большой скоростью рекомбинации (1), для реального контакта со скоростью рекомбинации, равной максимально возможной скорости движения носителей (2), и для реального контакта со скоростью рекомбинации, не превышающей максимальной скорости движения носителей (3).

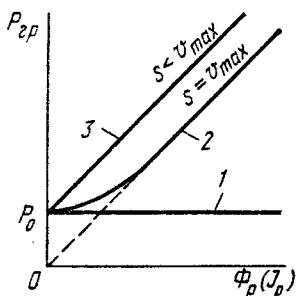


Рис. 2.19. Зависимость граничной концентрации носителей заряда около невыпрямляющего контакта от погока или от плотности тока этих носителей

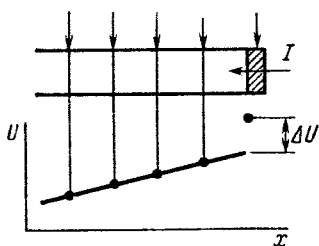


Рис. 2.20. Определение падения напряжения на контакте путем экстраполяции распределения потенциала в полупроводнике

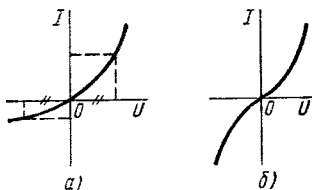


Рис. 2.21. Вольт-амперные характеристики контактов:

a — контакта, обладающего некоторым эффектом выпрямления; *b* — контакта без эффекта выпрямления. Но с сопротивлением, зависящим от приложенного напряжения

Сопротивление контакта. Сопротивление контакта — это отношение падения напряжения на контакте к проходящему через контакт току:

$$R_{\text{кон}} = \Delta U / I.$$

Чем меньше сопротивление контакта, тем он лучше. Для выяснения смысла падения напряжения на контакте сделаем мысленно следующий опыт. Пусть через полупроводник в виде призмы или цилиндра с контактом протекает ток (рис. 2.20). Потенциал в полупроводнике должен распределиться линейно. В этом можно убедиться, выполнив измерения с помощью зондов. Если один из зондов поставить на металлический электрод контакта, то полученное значение потенциала не попадет на линейную зависимость. Под падением напряжения на контакте следует понимать отклонение напряжения на металлическом электроде контакта от значения экстраполированного по распределению потенциала в полупроводнике. При практическом определении падения напряжения на контакте не обязательно иметь полупроводник в форме призмы или цилиндра, необходимо установить реальное распределение потенциала в нем для возможности экстраполяции.

Сопротивление контакта зависит от его площади. Поэтому можно сравнивать контакты с одинаковой площадью или ввести *удельное контактное сопротивление*, определив его как отношение падения напряжения на контакте к плотности проходящего тока:

$$\rho_{\text{кон}} = \Delta U / J, \text{ Ом} \cdot \text{см}^2. \quad (2.44)$$

Сопротивление контакта связано с удельным контактным сопротивлением:

$$R_{\text{кон}} = \rho_{\text{кон}} / S.$$

Параметры линейности. В зависимости от того, какие требования предъявляются к линейности вольт-амперной характеристики

контакта и каков вид нелинейности, ее можно характеризовать и оценивать по-разному. Если контакт обладает эффектом выпрямления (рис. 2.21, а), то нелинейность вольт-амперной характеристики контакта можно характеризовать коэффициентом выпрямления, т. е. отношением прямого тока к обратному при равных значениях приложенного напряжения. Идеальный невыпрямляющий контакт должен иметь коэффициент выпрямления, равный единице.

Невыпрямляющие контакты с симметричной вольт-амперной характеристикой (рис. 2.21, б) можно характеризовать коэффициентом нелинейности — отношением статического сопротивления к дифференциальному. Идеальный невыпрямляющий контакт с линейной вольт-амперной характеристикой будет иметь коэффициент нелинейности, равный единице.

Структура реального невыпрямляющего контакта

Невыпрямляющий контакт, удовлетворяющий перечисленным требованиям, имеет сравнительно сложную структуру. Он состоит из нескольких последовательно соединенных контактов (рис. 2.22).

Для меньшей вероятности накопления неосновных носителей заряда около контакта металл — полупроводник потенциальный барьер здесь должен быть по возможности меньшим. Поэтому металл желательно подобрать с работой выхода, которая мало бы отличалась от работы выхода электронов, в полупроводнике (см. § 2.10). Так как это обеспечить трудно, то поверхностный слой полупроводника должен быть сильно легирован для обеспечения возможности туннелирования носителей заряда сквозь потенциальный барьер.

На контакте двух полупроводников с разной степенью легирования также может происходить накопление неосновных носителей заряда (см. § 2.9). Для меньшего влияния этого явления на характеристики и параметры полупроводникового прибора целесообразно в поверхностный слой полупроводника вводить примеси рекомбинационных ловушек (например, золото), что уменьшит время жизни носителей заряда в этой части структуры. При малом времени жизни накопленные носители заряда будут быстро рекомбинировать, т. е. процесс их рассасывания меньше будет сказываться на параметрах прибора.

Однако в реальных невыпрямляющих контактах часто образуются различные промежуточные слои, ухудшающие свойства контактов. Поэтому окончательную доработку технологии невыпрямляющих контактов проводят экспериментально.

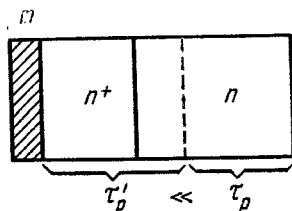


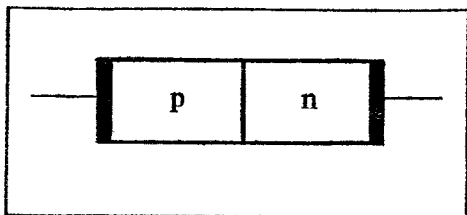
Рис. 2.22. Структура реального невыпрямляющего контакта

Контрольные вопросы

1. Что такое электронно-дырочный переход?
2. Как и почему изменяется высота потенциального барьера p - n -перехода с изменением температуры и с изменением концентрации примесей в прилегающих к переходу областях?
3. Что такое инжекция и экстракция неосновных носителей заряда?
4. От чего зависит и чем определяется концентрация неосновных носителей заряда на границах p - n -перехода?
5. Как можно записать условие электрической нейтральности p - n -перехода и каков физический смысл этого условия?
6. Как распределяются напряженности электрического поля и потенциалы в резком и в плавном p - n -переходах?
7. Как изменяется толщина резкого и плавного p - n -переходов при изменении приложенного напряжения?
8. Что такое барьерная емкость p - n -перехода?
9. Как можно определить величину контактной разности потенциалов на p - n -переходе с помощью экспериментальных вольт-фарадных характеристик этого перехода?
10. Почему контакт двух полупроводников с одним типом электропроводности является невыпрямляющим и неинжектирующим неосновные носители заряда в высокоомную область?
11. При каких условиях контакт металл — полупроводник является невыпрямляющим?
12. При каких условиях контакт металл — полупроводник будет выпрямляющим без инжекции неосновных носителей заряда в полупроводник?
13. В каком случае может происходить накопление неосновных носителей заряда вблизи невыпрямляющего контакта металл—полупроводник?
14. Каковы правила построения энергетических диаграмм гетеропереходов?
15. Почему и при каких условиях на гетеропереходах может происходить выпрямление без инжекции неосновных носителей заряда?
16. Почему на гетеропереходе между двумя полупроводниками с одним типом электропроводности может наблюдаться эффект выпрямления?
17. Каким требованиям должны удовлетворять невыпрямляющие контакты?
18. Что такое скорость рекомбинации на контакте?
19. Что такое сопротивление контакта и как его можно определить экспериментально?
20. Что такое коэффициент выпрямления и почему этот параметр применим только при наличии слабого эффекта выпрямления?
21. Что такое коэффициент нелинейности?

ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

3



§ 3.1. СТРУКТУРА И ОСНОВНЫЕ ЭЛЕМЕНТЫ

Полупроводниковый диод — это полупроводниковый прибор с одним выпрямляющим электрическим переходом и двумя внешними выводами, в котором используется то или иное свойство выпрямляющего перехода.

В качестве выпрямляющего электрического перехода в полупроводниковых диодах может быть электронно-дырочный переход, гетеропереход или контакт металл — полупроводник.

В диоде с электронно-дырочным переходом кроме выпрямляющего электрического перехода должно быть два невыпрямляющих перехода, через которые p - и n -области диода соединены с выводами (рис. 3.1, *а*). В диоде с выпрямляющим электрическим переходом в виде контакта металл — полупроводник всего один невыпрямляющий переход (рис. 3.1, *б*).

Обычно полупроводниковые диоды имеют несимметричные электронно-дырочные переходы. Поэтому при прямом включении диода количество неосновных носителей, инжектированных из сильнолегированной области в слаболегированную область, значительно больше, чем количество неосновных носителей, проходящих в противоположном направлении. В соответствии с общим определением (см. § 2.2) область полупроводникового диода, в которую происходит инжекция неосновных для этой области носителей, называют

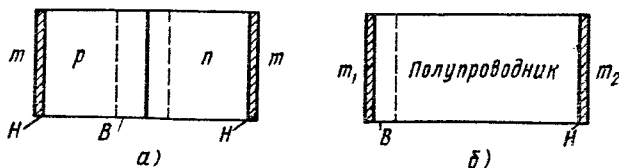


Рис. 3.1. Структура полупроводниковых диодов:

а — с электронно-дырочным переходом; *б* — с выпрямляющим контактом металл — полупроводник; *В* — выпрямляющие контакты; *Н* — невыпрямляющие контакты

базой днода. Таким образом, в диоде базовой областью является слаболегированная область.

В зависимости от соотношения линейных размеров выпрямляющего перехода и характеристической длины различают плоскостные и точечные диоды. Характеристической длиной для диода является наименьшая из двух величин, определяющая свойства и характеристики диода: диффузионная длина неосновных носителей в базе или толщина базы.

Плоскостным называют диод, у которого линейные размеры, определяющие площадь выпрямляющего электрического перехода, значительно больше характеристической длины.

Точечным называют диод, у которого линейные размеры, определяющие площадь выпрямляющего электрического перехода, значительно меньше характеристической длины.

Выпрямляющий переход кроме эффекта выпрямления обладает и другими свойствами: нелинейностью вольт-амперной характеристики; явлением ударной ионизации атомов полупроводника при относительно больших для данного перехода напряжениях; явлением туннелирования носителей сквозь потенциальный барьер перехода как при обратном, так в определенных условиях и при прямом напряжении; барьерной емкостью. Эти свойства выпрямляющего перехода используют для создания различных видов полупроводниковых диодов: выпрямительных диодов, смесителей, умножителей, модуляторов, стабилитронов, стабилитронов, лавинно-пролетных диодов, туннельных и обращенных диодов, варикапов.

§ 3.2. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ДИОДА ПРИ ИНЖЕКЦИИ И ЭКСТРАКЦИИ НОСИТЕЛЕЙ ЗАРЯДА

Прямое включение диода

При прямом напряжении на диоде внешнее напряжение частично компенсирует контактную разность потенциалов на электронно-дырочном переходе, так как внешнее электрическое поле при прямом включении диода направлено в противоположную сторону диффузионному полю. Поэтому высота потенциального барьера перехода уменьшается пропорционально приложенному к диоду напряжению. Пренебрегая падением напряжения на базе диода, рассмотрим диод при малых прямых токах.

С уменьшением высоты потенциального барьера увеличивается количество носителей заряда, которые могут преодолеть потенциальный барьер и перейти в соседнюю область днода, где они окажутся неосновными носителями (см. рис. 2.1). Этот процесс, как было отмечено в § 2.2, называют инжекцией неосновных носителей через электронно-дырочный переход.

Так как высота потенциального барьера уменьшается пропорционально приложенному напряжению, а носители заряда распределены по энергиям по экспоненциальному закону в соответствии

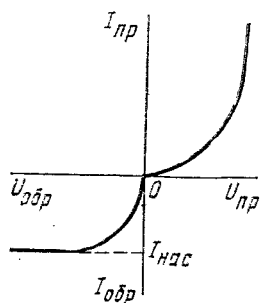
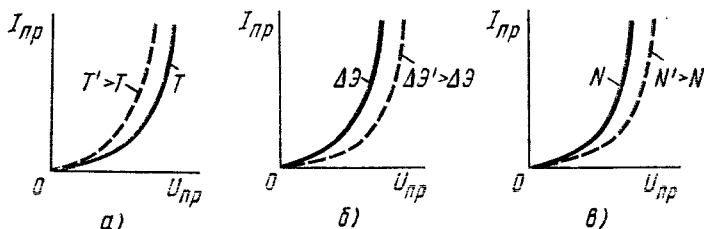


Рис. 3.2. Вольт-амперная характеристика полупроводникового диода

Рис. 3.3. Прямые ветви вольт-амперных характеристик диода при разных температурах (а), при разной ширине запрещенной зоны исходного материала (б) и при разной концентрации примесей в прилегающих к *p-n*-переходу областях (в)



со статистикой Ферми — Дирака или Максвелла — Больцмана, то прямая ветвь вольт-амперной характеристики диода должна быть похожа на экспоненту (рис. 3.2).

Рассмотрим влияние некоторых факторов на прямую ветвь вольт-амперной характеристики диода.

При увеличении температуры диода уменьшается высота потенциального барьера (см. § 2.1) и изменяется распределение носителей заряда по энергиям — электроны, например, занимают более высокие энергетические уровни в зоне проводимости. Из-за этих двух причин прямой ток через диод увеличивается с ростом температуры при неизменном приложенном напряжении (рис. 3.3, а).

Если сравнить прямые ветви двух диодов, изготовленных из разных материалов, с разной шириной запрещенной зоны, то у диода из материала с большей шириной запрещенной зоны будет больше высота потенциального барьера (см. § 2.1). Следовательно, прямой ток через диод из материала с большей шириной запрещенной зоны будет меньше при том же прямом напряжении (рис. 3.3, б).

С увеличением концентрации примесей в прилегающих к электронно-дырочному переходу областях будет увеличиваться высота потенциального барьера перехода (см. § 2.1), а значит, будет меньше прямой ток при неизменном напряжении (рис. 3.3, в).

Обратное включение диода

При обратном включении диода внешнее электрическое поле и диффузионное поле в электронно-дырочном переходе совпадают по направлению, происходит экстракция неосновных носителей из прилегающих к переходу областей (см. § 2.2 и рис. 2.1, е). Это

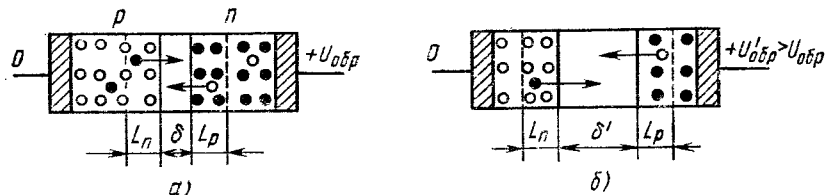


Рис. 3.4. Экстракция неосновных носителей из прилегающих к p - n -переходу областей при разных обратных напряжениях на диоде

приводит к уменьшению граничной концентрации неосновных носителей около электронно-дырочного перехода и к появлению диффузии неосновных носителей к переходу — идет диффузионный ток неосновных носителей, возникающих в результате тепловой генерации в объеме n - и p -областей диода, а также на невыпрямляющих контактах.

За время жизни до электронно-дырочного перехода могут продиффундировать неосновные носители, возникшие в n - и p -областях на расстоянии, не превышающем соответствующую диффузионную длину (рис. 3.4). Остальные неосновные носители, не успев дойти до перехода, рекомбинируют в объеме. Это справедливо для разных обратных напряжений на диоде, если толщины прилегающих к переходу областей превышают диффузионные длины неосновных носителей. Поэтому обратный ток, начиная с очень малых значений обратного напряжения, не будет изменяться с изменением напряжения (см. рис. 3.2). Этот неизменный ток с изменением напряжения обратный ток через диод называют током насыщения $J_{нас}$.

Рассмотрев физические процессы в диоде при обратном напряжении, можно выразить величину плотности тока насыщения через параметры полупроводникового материала. Для этого надо вспомнить общее соотношение для плотности тока при наличии носителей двух типов:

$$J = q(pv_p + nv_n),$$

где v_p и v_n — скорости либо диффузии, либо дрейфа дырок и электронов.

В диоде неосновные носители диффундируют к переходу, поэтому их скорости можно представить как диффузионные длины, деленные на соответствующие времена жизни. Вместо общих концентраций p и n подставим концентрации неосновных носителей, так как с движением именно неосновных носителей связан обратный ток диода. Тогда

$$J_{нас} = q \left(p_{n0} \frac{L_p}{\tau_p} + n_{p0} \frac{L_n}{\tau_n} \right). \quad (3.1)$$

Если учесть соотношения (1.19), а также практически полную ионизацию примесей при комнатных температурах, то (3.1) можно привести к виду:

$$J_{нас} = qn_i^2 \left(\frac{L_p}{N_{дп}\tau_p} + \frac{L_n}{N_{ап}\tau_n} \right). \quad (3.2)$$

При увеличении температуры диода плотность тока насыщения увеличивается, так как с температурой экспоненциально растет собственная концентрация носителей заряда [см. (1.15)].

В диодах на основе материала с большей шириной запрещенной зоны плотность тока насыщения должна быть значительно меньше, так как собственная концентрация экспоненциально уменьшается с увеличением ширины запрещенной зоны [см. (1.15)]. Сравнивая германиевые и кремниевые диоды и учитывая разницу в собственных концентрациях носителей в германии и кремнии, которая составляет три порядка (см. § 1.4), следует заключить, что плотность тока насыщения в кремниевых диодах должна быть меньше на шесть порядков.

С увеличением концентрации примесей в прилегающих к переходу областях плотность тока насыщения в соответствии с (3.2) должна уменьшиться.

§ 3.3. РАСЧЕТ РАСПРЕДЕЛЕНИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА В БАЗЕ ДИОДА

В качестве примера проведем расчет для диода с несимметричным электронно-дырочным переходом p^+-n (рис. 3.5) при приложении к нему напряжения, имеющего постоянную и малую переменную составляющие:

$$u = U + U_m \exp j\omega t. \quad (3.3)$$

Здесь переменная составляющая записана в виде комплексной величины в показательной форме. Так как в (3.3) складывается постоянная и переменная составляющие, то физический смысл имеет только проекция вектора переменного напряжения на действительную ось.

Условием малости переменной составляющей напряжения для удобства математических преобразований выберем следующее:

$$U_m \ll kT/q, \quad (3.4)$$

т. е. амплитуда переменной составляющей напряжения не должна превосходить ~ 10 мВ.

Допущения при расчете

Для облегчения расчетов обычно выбирают упрощенную модель структуры того или иного прибора. В данном случае допустим, что:

1) электронно-дырочный переход диода плоский, т. е. будем рассматривать одномерную модель диода (рис. 3.5);

2) токи малы и не вызывают существенного падения напряжения на сопротивлении базы диода; таким образом, электрическое поле сосредоточено только в электронно-дырочном переходе;

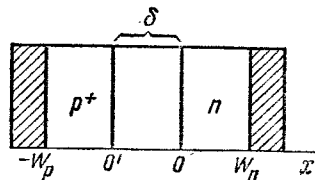


Рис. 3.5. Одномерная модель диода, принятая для расчета

3) невыпрямляющие контакты идеальны, т. е. около невыпрямляющих контактов в полупроводнике существует всегда только равновесная концентрация носителей заряда;

4) поверхностные явления несущественны;

5) в электронно-дырочном переходе не происходят процессы генерации или рекомбинации носителей;

6) рекомбинация неосновных носителей в объеме базы линейная, т. е. количество носителей, рекомбинирующих в единице объема за единицу времени, прямо пропорционально избыточной концентрации:

$$R_p = \frac{\Delta p_n}{\tau_p} = \frac{p_n - p_{n0}}{\tau_p}.$$

Дифференциальные уравнения

Основным для решения поставленной задачи является уравнение непрерывности, например, для дырок в базе диода с электропроводностью n -типа:

$$\frac{\partial p_n}{\partial t} = -\frac{1}{q} \operatorname{div} \mathbf{J}_p - R_p + G_p. \quad (3.5)$$

Это уравнение показывает, как и по каким причинам изменяется концентрация дырок со временем. Во-первых, концентрация дырок может изменяться из-за существования тока дырок, что учитывает первое слагаемое. Во-вторых, концентрация дырок может изменяться из-за их рекомбинации, что учитывает второе слагаемое (R_p — скорость рекомбинации). Это же слагаемое может учитывать изменение концентрации дырок из-за тепловой генерации. В-третьих, концентрация дырок может изменяться из-за нетепловой генерации (ударная ионизация, ионизация под действием света и т. д.). В данном случае $G_p = 0$.

С учетом принятых в начале параграфа допущений уравнение непрерывности (3.5) перепишем следующим образом:

$$\frac{\partial p_n}{\partial t} = -\frac{1}{q} \cdot \frac{\partial J_p}{\partial x} - \frac{p_n - p_{n0}}{\tau_p}. \quad (3.6)$$

Воспользуемся уравнением (1.31) для плотности тока дырок, также упростив его с учетом принятых допущений:

$$\mathbf{J}_p = -qD_p \frac{\partial p_n}{\partial x}. \quad (3.7)$$

После подстановки (3.7) в (3.6):

$$\frac{\partial p_n}{\partial t} - D_p \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p}. \quad (3.8)$$

Получим дифференциальное уравнение второго порядка в частных производных. Для его решения необходимы граничные и начальные условия.

Граничные условия

При малых токах концентрация неосновных носителей в базе около p - n -перехода определяется соотношением (2.5):

$$p_n(0) = p_{n0} \exp \frac{qu}{kT}.$$

Подставив сюда значение напряжения (3.3), получим при $x = 0$

$$\begin{aligned} p_n(0) &= p_{n0} \exp \frac{q}{kT} (U + U_m \exp j\omega t) = \\ &= p_{n0} \left(\exp \frac{qU}{kT} \right) \exp \left(\frac{qU_m}{kT} \exp j\omega t \right). \end{aligned} \quad (3.9)$$

Аргумент второй экспоненты мал. Поэтому ее можно разложить в ряд, ограничившись двумя первыми его членами:

$$\exp y \approx 1 + y + \dots$$

Тогда (3.9) с учетом разложения в ряд

$$p_n(0) = p_{n0} \exp \frac{qU}{kT} + p_{n0} \left(\exp \frac{qU}{kT} \right) \frac{qU_m}{kT} \exp j\omega t. \quad (3.10)$$

Таким образом, концентрация неосновных носителей в базе около электронно-дырочного перехода имеет постоянную и переменную составляющие. Частота изменения переменной составляющей та же, что и частота приложенного переменного напряжения. Если бы переменное напряжение не было мало, то у переменной составляющей концентрации носителей появились бы гармоники высшего порядка.

Второе граничное условие следует из идеальности невыпрямляющего контакта, т. е. при $x = W_n$

$$p_n(W_n) = p_{n0}. \quad (3.11)$$

Форма решения

Для преобразования дифференциального уравнения (3.8) выберем форму решения в виде суммы постоянной и переменной составляющих концентрации, т. е. форму, аналогичную граничному условию (3.10). Это обусловлено линейностью уравнения для концентрации носителей, так что новые гармоники появиться не могут. Тогда решение дифференциального уравнения (3.8) должно иметь следующий вид:

$$p_n(x, t) = p_{n0} + \Delta p_n(x) + \dot{p}_n(x) \exp j\omega t, \quad (3.12)$$

где $p_n(x, t)$ — полная концентрация неосновных носителей в базе; $\Delta p_n(x)$ — постоянная составляющая избыточной концентрации неосновных носителей, зависящая только от координаты; $\dot{p}_n(x) \exp j\omega t$ — переменная составляющая избыточной концентрации неосновных носителей, зависящая как от координаты, так и от времени.

Таким образом, постоянная составляющая в (3.12) представлена в виде суммы равновесной и избыточной концентраций.

Преобразование общего дифференциального уравнения

После подстановки выбранной формы решения (3.12) в (3.8) получим

$$\begin{aligned} \dot{p}_n(x) j\omega \exp j\omega t = D_p \frac{\partial^2 [\Delta p_n(x)]}{\partial x^2} + D_p \frac{\partial^2 \dot{p}_n(x)}{\partial x^2} \exp j\omega t - \\ - \frac{\Delta p_n(x)}{\tau_p} - \frac{\dot{p}_n(x)}{\tau_p} \exp j\omega t. \end{aligned} \quad (3.13)$$

В этом уравнении есть слагаемые, зависящие от времени и не зависящие от него. Уравнение справедливо только в том случае, если алгебраические суммы не зависящих от времени и зависящих от него составляющих отдельно равны нулю. Поэтому для постоянной составляющей избыточной концентрации из (3.13) получим

$$\frac{d^2 [\Delta p_n(x)]}{dx^2} = \frac{\Delta p_n(x)}{L_p^2}, \quad (3.14)$$

где

$$L_p = \sqrt{D_p \tau_p}.$$

Граничные условия для постоянной составляющей избыточной концентрации будут

$$\text{при } x=0 \quad \Delta p_n(0) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right), \quad (3.15)$$

$$\text{при } x=W_n \quad \Delta p_n(W_n) = 0. \quad (3.16)$$

Для переменной составляющей избыточной концентрации

$$\frac{d^2 \dot{p}_n(x)}{dx^2} = \frac{\dot{p}_n(x)}{\Lambda_p^2}. \quad (3.17)$$

где

$$\Lambda_p = \frac{L_p}{\sqrt{1 + j\omega \tau_p}}.$$

Граничные условия для переменной составляющей избыточной концентрации запишем, исходя из общих граничных условий (3.10) и (3.11):

$$\text{при } x=0 \quad \dot{p}_n(0) \exp j\omega t = p_{n0} \left(\exp \frac{qU}{kT} \right) \frac{q\dot{U}_m}{kT} \exp j\omega t,$$

или окончательно

$$\dot{p}_n(0) = p_{n0} \left(\exp \frac{qU}{kT} \right) \frac{q\dot{U}_m}{kT}, \quad (3.18)$$

$$\text{при } x=W_n \quad \dot{p}_n(W_n) = 0. \quad (3.19)$$

Полученные уравнения для постоянной и переменной составляющих аналогичны, так что решать можно только одно из них.

Решение для постоянной составляющей избыточной концентрации неосновных носителей в базе

Решение дифференциального уравнения (3.14) удобнее искать в виде:

$$\Delta p_n(x) = A_1 \operatorname{ch} \frac{x}{L_p} + A_2 \operatorname{sh} \frac{x}{L_p}. \quad (3.20)$$

Такой вид решения упрощает поиск произвольных постоянных, если заданы условия на границах и при этом на одной границе — нуль.

Подставив в (3.20) $x = 0$ и учтя граничное условие (3.15), получим

$$\Delta p_n(0) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) = A_1.$$

С учетом граничного условия (3.16) и значения постоянной интегрирования A_1 при $x = W_n$

$$\Delta p_n(W_n) = 0 = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \operatorname{ch} \frac{W_n}{L_p} + A_2 \operatorname{sh} \frac{W_n}{L_p}.$$

Следовательно,

$$A_2 = - p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \operatorname{cth} \frac{W_n}{L_p}.$$

Подставив в (3.20) значения постоянных интегрирования, окончательно получим

$$\Delta p_n(x) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \left(\operatorname{ch} \frac{x}{L_p} - \operatorname{cth} \frac{W_n}{L_p} \operatorname{sh} \frac{x}{L_p} \right). \quad (3.21)$$

Таково распределение постоянной составляющей избыточной концентрации неосновных носителей в базе диода при разных напряжениях.

§ 3.4. РАСЧЕТ ПОСТОЯННЫХ ТОКОВ, ПРОХОДЯЩИХ ЧЕРЕЗ ДИОД И СВЯЗАННЫХ С ИНЖЕКЦИЕЙ И ЭКСТРАКЦИЕЙ НОСИТЕЛЕЙ ЗАРЯДА

Для определения дырочной составляющей плотности тока в произвольном сечении базы диода воспользуемся соотношениями (1.31) и (3.21), учтя допущение, принятое в предыдущем параграфе, а именно отсутствие электрического поля в базе диода. Тогда после дифференцирования

$$J_p(x) = - \frac{qD_p p_{n0}}{L_p} \left(\operatorname{sh} \frac{x}{L_p} - \operatorname{cth} \frac{W_n}{L_p} \operatorname{ch} \frac{x}{L_p} \right) \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.22)$$

Такова плотность постоянного тока в различных частях базы диода.

При $x = 0$, т. е. на границе базы с p - n -переходом, дырочная составляющая плотности тока через переход

$$J_p(0) = \frac{qD_p p_{n0}}{L_p} \operatorname{cth} \frac{W_n}{L_p} \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.23)$$

Аналогично можно записать и электронную составляющую плотности тока через переход диода, т. е. через границу между p -областью и p - n -переходом:

$$J_n(0') = \frac{qD_n n_{p0}}{L_n} \operatorname{cth} \frac{W_p}{L_n} \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.24)$$

Для практики бóльший интерес представляет знание полной плотности тока, т. е. суммы электронной и дырочной составляющих. Суммировать надо плотности токов в одном и том же сечении. Однако расчет тока основных носителей в базе потребовал бы использования другой методки по сравнению с методикой расчета тока неосновных носителей. Поэтому используем допущение, что в электронно-дырочном переходе не происходят процессы генерации и рекомбинации носителей. Следовательно, составляющие плотностей тока (электронного и дырочного) одинаковы по обе стороны перехода. Теперь можно определить плотности тока неосновных носителей с двух сторон электронно-дырочного перехода и сложить их:

$$J = J_p(0) + J_n(0') = q \left(\frac{p_{n0} D_p}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{n_{p0} D_n}{L_n} \operatorname{cth} \frac{W_p}{L_n} \right) \times \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.25)$$

Первые два множителя в (3.25) определяются параметрами полупроводниковых областей с двух сторон p - n -перехода и толщиной этих областей. Произведение этих множителей, не зависящее от напряжения, называют плотностью тока насыщения:

$$J_{\text{нас}} = q \left(\frac{p_{n0} D_p}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{n_{p0} D_n}{L_n} \operatorname{cth} \frac{W_p}{L_n} \right). \quad (3.26)$$

Строго говоря, этот ток не совсем и не всегда насыщенный, так как толщина базы диода зависит от напряжения в связи с изменением толщины p - n -перехода при изменении приложенного напряжения.

Таким образом, вольт-амперную характеристику диода обычно записывают в виде

$$J = J_{\text{нас}} \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.27)$$

Графическое изображение вольт-амперной характеристики показано на рис. 3.2. Для удобства масштабы прямых и обратных напряжений, а также прямых и обратных токов выбирают разными.

**§ 3.5. ЧАСТНЫЕ СЛУЧАИ РАСЧЕТА
РАСПРЕДЕЛЕНИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА
И ТОКА НАСЫЩЕНИЯ**

Чтобы получить еще более простые выражения и лучше разобраться в смысле результата, рассмотрим предельные частные случаи диодов с толстой и тонкой базой.

Диод с толстой базой — это диод, толщина базы которого значительно превышает диффузионную длину неосновных носителей ($W_n \gg L_p$). Для другой области диода из-за сильного ее легирования тем более справедливо подобное неравенство, т. е. $W_p \gg L_n$. Тогда аргументы гиперболических котангенсов в соотношении (3.21) значительно превышают единицу, а сами гиперболические котангенсы близки к единице:

$$\operatorname{cth} \frac{W_n}{L_p} \approx 1 \quad \text{и} \quad \operatorname{cth} \frac{W_p}{L_n} \approx 1.$$

Воспользовавшись формулами Эйлера $\left(\operatorname{sh} y = \frac{e^y - e^{-y}}{2}; \operatorname{ch} y = \frac{e^y + e^{-y}}{2} \right)$, получим из соотношения (3.21) распределение избыточной концентрации неосновных носителей в базе диода с толстой базой

$$\Delta p_n(x) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \exp \left(-\frac{x}{L_p} \right). \quad (3.28)$$

Следовательно, в диоде с толстой базой абсолютное значение избыточной концентрации неосновных носителей в базе ($\Delta p_n = p_n - p_{n0}$) экспоненциально уменьшается с увеличением расстояния от p - n -перехода (рис. 3.6). При прямых напряжениях избыточная концентрация положительна, что соответствует инжекции неосновных носителей в базу. При обратных напряжениях избыточная концентрация отрицательна, т. е. $p_n < p_{n0}$, что соответствует экстракции неосновных носителей из базы.

Выражение для плотности тока насыщения в диоде с толстой базой можно найти из соотношения (3.26) с учетом того, что $\operatorname{cth} \frac{W}{L} \approx 1$:

$$J_{\text{нас}} = q \left(\frac{p_{n0} D_p}{L_p} + \frac{n_{p0} D_n}{L_n} \right). \quad (3.29)$$

Это выражение для плотности тока насыщения в диоде с толстой базой можно получить также исходя из того, что обратный ток обусловлен только диффузией неосновных носителей в прилегающих к электронно-дырочному переходу областях. Поэтому для вычисления плотности тока насыщения надо воспользоваться вторыми слагаемыми в (1.30) и (1.31) или соотношениями (1.26) и (1.27). При этом градиенты концентраций неосновных носителей в n - и p -областях около электронно-дырочного перехода можно определить как p_{n0}/L_p и n_{p0}/L_n (см. рис. 3.6).

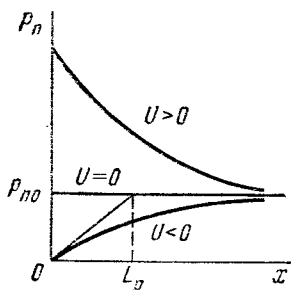


Рис. 3.6. Распределение концентрации неосновных носителей в базе диода с толстой базой при разных напряжениях

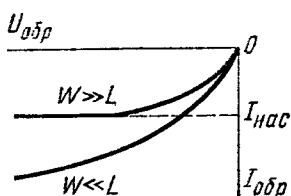


Рис. 3.7. Обратные ветви вольт-амперных характеристик диодов с толстой и с тонкой базой с учетом экстракции неосновных носителей заряда из прилегающих к $p-n$ -переходу областей

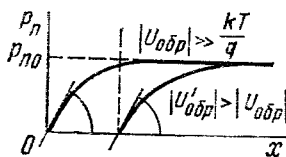


Рис. 3.8. Распределение концентрации неосновных носителей заряда в базе диода при разных обратных напряжениях, поясняющее неизменность обратного тока (тока насыщения) в диоде с толстой базой при экстракции неосновных носителей

Таким образом, в диоде с толстой базой плотность тока насыщения не зависит от напряжения. Поэтому обратный ток через диод с учетом только экстракции неосновных носителей заряда, начиная с обратных напряжений $|U_{обр}| \gg kT/q$ или $|U_{обр}| > 0,1В$, не изменяется с напряжением (рис. 3.7). Тот же вывод можно сделать, если представить распределение неосновных носителей в базе диода с толстой базой при разных обратных напряжениях (рис. 3.8). С увеличением обратного напряжения по абсолютному значению и соответственно с увеличением толщины $p-n$ -перехода за счет толщины базы происходит смещение кривых распределения концентрации неосновных носителей заряда в глубь базы при неизменном градиенте концентрации около $p-n$ -перехода, что по соотношению (1.27) соответствует неизменному току.

Выражение (3.29) аналогично выражению (3.1), полученному из чисто физических соображений, так как $L = \sqrt{D\tau}$.

В соответствии с (3.29) характеристической длиной, определяющей свойства и многие параметры диода с толстой базой, является диффузионная длина неосновных носителей заряда в базе диода.

Диод с тонкой базой — это диод, толщина базы которого значительно меньше диффузионной длины неосновных носителей ($W_n \ll L_p$). В этом случае аргументы всех гиперболических функций в соотношении (3.21) будут малы (меньше единицы). Поэтому, раскладывая гиперболические функции в ряд, можно ограничиться всего одним членом разложения ($\text{cth } y \approx 1/y$; $\text{sh } y \approx y$; $\text{ch } y \approx 1$). Тогда для распределения концентрации неосновных носителей в диоде с тонкой базой из (3.21) получим

$$\Delta p_n(x) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \left(1 - \frac{x}{W_n} \right). \quad (3.30)$$

Следовательно, в диоде с тонкой базой концентрация неосновных носителей уменьшается по мере удаления от $p-n$ -перехода

линейно (рис. 3.9). Другими словами, например, для диода с несимметричным p^+n -переходом плотность тока в любом сечении базы неизменна в соответствии с (3.7) и с учетом (3.30), т. е. рекомбинация неосновных носителей заряда в базе несущественна. Все неосновные носители заряда, инжектированные в базу при прямом напряжении, доходят до невыпрямляющего контакта, где и рекомбинируют. При обратном напряжении через p -переход диода с тонкой базой происходит экстракция всех носителей заряда, поставляемых в базу невыпрямляющим контактом.

Плотность тока насыщения в диоде с тонкой базой из общего соотношения (3.26) с учетом разложения в ряд гиперболических функций

$$J_{\text{нас}} = q \left(\frac{p_{n0} D_p}{W_n} + \frac{n_{p0} D_n}{W_p} \right). \quad (3.31)$$

Это выражение для плотности тока насыщения можно также получить исходя из того, что обратный ток обусловлен только диффузией неосновных носителей заряда от невыпрямляющих контактов к электронно-дырочному переходу по прилегающим к переходу областям. Поэтому для вычисления плотности тока насыщения надо воспользоваться вторыми слагаемыми в (1.30) и (1.31) или соотношениями (1.26) и (1.27). При этом градиенты концентраций неосновных носителей в n - и p -областях можно определить как p_{n0}/W_n и n_{p0}/W_p (см. рис. 3.9).

Таким образом, в диоде с тонкой базой плотность тока насыщения зависит от обратного напряжения, так как с изменением обратного напряжения изменяется и толщина базы (W_n или W_p) в связи с изменением толщины p - n -перехода (см. рис. 3.7). Тот же вывод можно сделать, если представить распределение неосновных носителей заряда в базе диода с тонкой базой при разных обратных напряжениях (рис. 3.10). С увеличением обратного напряжения по абсолютному значению происходит изменение градиента концентрации неосновных носителей заряда в базе, т. е. невыпрямляющий контакт (источник неосновных носителей заряда в данном случае) тем сильнее влияет на обратный ток, чем он ближе оказывается к p - n -переходу.

В соответствии с (3.31) характеристической длиной, определяющей свойства и многие параметры диода с тонкой базой, является толщина базы диода.

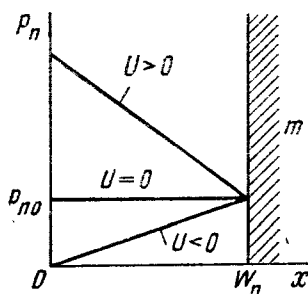


Рис. 3.9. Распределение концентраций неосновных носителей в базе диода с тонкой базой

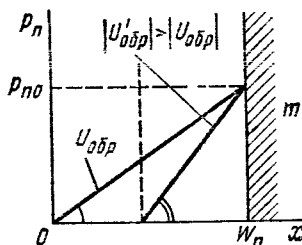


Рис. 3.10. Распределение концентраций неосновных носителей в базе диода при разных обратных напряжениях, поясняющее увеличение обратного тока в диоде с тонкой базой

§ 3.6. РАСЧЕТ ПЕРЕМЕННЫХ ТОКОВ И ПОЛНОЙ ПРОВОДИМОСТИ ДИОДА

В связи с аналогией дифференциальных уравнений для распределения постоянной и переменной составляющих концентрации неосновных носителей (3.14) и (3.17), а также в связи с аналогией граничных условий можно сразу записать выражение для переменной составляющей плотности тока через диод. Для этого достаточно сделать следующие замены: вместо L подставить

$$\Lambda = \frac{L}{\sqrt{1 + j\omega\tau}}; \quad (3.32)$$

вместо $(\exp \frac{qU}{kT} - 1)$ подставить $(\exp \frac{qU}{kT}) \frac{q\dot{U}_m}{kT}$.

В результате получим выражение для переменной составляющей плотности тока, аналогичное выражению (3.25):

$$j_m = \frac{q^2 \dot{U}_m}{kT} \left(\frac{p_{n0} D_p}{\Lambda_p} \operatorname{cth} \frac{W_n}{\Lambda_p} + \frac{n_{p0} D_n}{\Lambda_n} \operatorname{cth} \frac{W_p}{\Lambda_n} \right) \exp \frac{qU}{kT}. \quad (3.33)$$

Видно, что переменная составляющая плотности тока через диод линейно связана с переменным напряжением, если это напряжение мало (см. § 3.3). Отсюда следует, что для описания свойств диода целесообразно воспользоваться обычным приемом электротехники — ввести полную проводимость (либо полное сопротивление) диода для переменного тока:

$$\dot{Y} = j_m / \dot{U}_m = S j_m / \dot{U}_m.$$

Тогда для диода, используя (3.33):

$$\dot{Y} = \frac{q^2 S}{kT} \left(\frac{p_{n0} D_p}{\Lambda_p} \operatorname{cth} \frac{W_n}{\Lambda_p} + \frac{n_{p0} D_n}{\Lambda_n} \operatorname{cth} \frac{W_p}{\Lambda_n} \right) \exp \frac{qU}{kT}. \quad (3.34)$$

Полученное выражение позволяет выполнить расчет проводимости полупроводникового диода на любой частоте и при любых соотношениях размеров диода и диффузионной длины.

Как видно из (3.4), проводимость диода для переменного тока оказывается комплексной. Запишем проводимость диода в виде:

$$\dot{Y} = \frac{1}{r} + j\omega C_{\text{диф}},$$

что соответствует параллельной схеме замещения.

Значения диффузионной емкости $C_{\text{диф}}$ и дифференциального сопротивления r могут быть вычислены для любого случая из общего выражения полной проводимости диода (3.34) путем соответствующих преобразований — извлечения корней из комплексных чисел, взятия гиперболического котангенса сумм и т. д. Однако получающиеся выражения оказываются довольно сложными, поэтому целесообразно ограничиться частными случаями. Таких частных случаев можно выделить четыре — два соотношения размеров ($W \gg L$ и $W \ll L$) для двух диапазонов частот (высоких и низких).

1) $W \gg L$, **низкие частоты**. Условие $W \gg L$ означает, что аргументы гиперболических котангенсов в (3.34) являются большими. Следовательно,

$$\operatorname{cth} \left(\frac{W}{L} \sqrt{1 + j\omega\tau} \right) \approx 1.$$

Теперь остается преобразовать соотношение (3.32). Преобразование этого соотношения оказывается простым, если $\omega\tau \ll 1$. Такое условие и будем считать критерием низкой частоты. Тогда, раскладывая в биномиальный ряд $\sqrt{1 + j\omega\tau}$, получим

$$\frac{1}{\Lambda} \approx \frac{1}{L} \left(1 + j \frac{\omega\tau}{2} \right) = \frac{1}{L} + j\omega \frac{\tau}{2L}.$$

Подставив полученное выражение в формулу для полной проводимости (3.34), получим после группирования слагаемых

$$\dot{Y} = \frac{q^2 S}{kT} \left[\left(\frac{p_{n0} D_p}{L_p} + \frac{n_{p0} D_n}{L_n} \right) + \frac{j\omega}{2} (p_{n0} L_p + n_{p0} L_n) \right] \exp \frac{qU}{kT}. \quad (3.35)$$

Следовательно,

$$\frac{1}{r} = \frac{q^2 S}{kT} \left(\frac{p_{n0} D_p}{L_p} + \frac{n_{p0} D_n}{L_n} \right) \exp \frac{qU}{kT},$$

или, учтя (3.29), $\frac{1}{r} = \frac{qS}{kT} J_{\text{нас}} \exp \frac{dU}{kT}$.

Воспользовавшись выражением (3.27) для вольт-амперной характеристики в виде

$$I = I_{\text{нас}} \left(\exp \frac{qU}{kT} - 1 \right) \quad \text{или} \quad I_{\text{нас}} \exp \frac{qU}{kT} = I + I_{\text{нас}}, \quad (3.36)$$

получим окончательно

$$\frac{1}{r} = \frac{q}{kT} (I + I_{\text{нас}}). \quad (3.37)$$

Диффузионную емкость получаем из (3.35):

$$C_{\text{диф}} = \frac{q^2 S}{2kT} (p_{n0} L_p + n_{p0} L_n) \exp \frac{qU}{kT}. \quad (3.38)$$

Если $p_{n0} \gg p_{p0}$, либо $n_{p0} \gg n_{n0}$, либо $\tau_p = \tau_n = \tau$, то выражение для диффузионной емкости с учетом (3.38) и (3.1) или (3.29) приводим к виду

$$C_{\text{диф}} = \frac{qS}{2kT} J_{\text{нас}} \left(\exp \frac{qU}{kT} \right) \tau,$$

или, учитывая (3.36),

$$C_{\text{диф}} = \frac{q}{2kT} (I + I_{\text{нас}}) \tau. \quad (3.39)$$

Для диодов с несимметричным p - n -переходом τ в этой формуле — время жизни неосновных носителей в слаболегированной области, т. е. в базе диода.

Для такого значения диффузионной емкости можно легко найти постоянную времени:

$$rC_{\text{диф}} = \tau/2. \quad (3.40)$$

2) $W \gg L$, высокие частоты. Критерий высокой частоты $\omega\tau \gg 1$, т. е. нарушение условия низкой частоты. При этом, как и ранее,

$$\text{cth} \left(\frac{W}{L} \sqrt{1 + j\omega\tau} \right) \approx 1.$$

Теперь надо преобразовать (3.32) с учетом высокой частоты. Чтобы извлечь корень, удобно воспользоваться условием $\omega\tau \gg 1$. Тогда, применяя формулу Муавра, получим:

$$\begin{aligned} \sqrt{1 + j\omega\tau} &\approx \sqrt{j\omega\tau} = \sqrt{\cos 90^\circ + j\omega\tau \sin 90^\circ} = \\ &= \sqrt{\omega\tau} (\cos 45^\circ + j \sin 45^\circ) = \sqrt{\omega\tau} \left(\frac{1}{\sqrt{2}} + j \frac{1}{\sqrt{2}} \right), \end{aligned}$$

т. е.

$$\frac{1}{\Lambda} \approx \frac{1}{L} \sqrt{\frac{\omega\tau}{2}} + j \frac{1}{L} \sqrt{\frac{\omega\tau}{2}}.$$

Поэтому полная проводимость днада

$$\begin{aligned} \dot{Y} &= \frac{q^2 S}{kT} \left[\sqrt{\frac{\omega}{2}} \left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n} \right) + \right. \\ &\left. + j \sqrt{\frac{\omega}{2}} \left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n} \right) \right] \exp \frac{qU}{kT}. \end{aligned} \quad (3.41)$$

Тогда

$$\frac{1}{r} = \frac{q^2 S}{kT} \sqrt{\frac{\omega}{2}} \left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n} \right) \exp \frac{qU}{kT}. \quad (3.42)$$

Если $p_{n0} \gg n_{p0}$, либо $n_{p0} \gg p_{n0}$, либо $\tau_p = \tau_n = \tau$, то выражение для активной проводимости

$$\frac{1}{r} = \frac{qS}{kT} \sqrt{\frac{\omega\tau}{2}} J_{\text{нас}} \exp \frac{qU}{kT},$$

т. е.

$$\frac{1}{r} = \frac{q}{kT} (I + I_{\text{нас}}) \sqrt{\frac{\omega\tau}{2}}. \quad (3.43)$$

Сравним (3.43) и (3.37), получим

$$\left(\frac{1}{r} \right)_{\text{вч}} = \left(\frac{1}{r} \right)_{\text{вч}} \sqrt{\frac{\omega\tau}{2}}. \quad (3.44)$$

Аналогично, сравним мнимые части (3.41) и (3.35):

$$(C_{\text{диф}})_{\text{вч}} = \frac{q}{kT} (I + I_{\text{нас}}) \sqrt{\frac{\tau}{2\omega}} = (C_{\text{диф}})_{\text{вч}} \sqrt{\frac{2}{\omega\tau}}. \quad (3.45)$$

Постоянная времени в данном случае

$$rC_{\text{диф}} = 1/\omega \quad (3.46)$$

подтверждает, что фазовый сдвиг составляет $\pi/4$. Это же можно заключить из равенства действительной и мнимой частей полной проводимости диода (3.41), т. е.

$$\varphi = \operatorname{arctg} \frac{x_c}{r} = \operatorname{arctg} 1 = 45^\circ.$$

3) $W \ll L$, низкие частоты. В данном случае критерий низких частот отличается от того, который применялся для диода с толстой базой. Действительно, условие тонкой базы диода ($W \ll L$) при низких частотах ($\omega \rightarrow 0$) приводит к тому, что аргументы гиперболических котангенсов в (3.34) будут малыми. Именно малость аргумента гиперболического котангенса будем считать условием низкой частоты, т. е. критерий низких частот в данном случае

$$\left| \frac{W}{L} \sqrt{1 + j\omega\tau} \right| \ll 1.$$

Воспользовавшись этим условием, разложим гиперболический котангенс в ряд. Если ограничиться одним членом ряда, как это было сделано в § 3.5, то в выражении для проводимости исчезнет мнимая составляющая. Поэтому возьмем два члена ряда, чтобы не потерять интересующий нас эффект. Тогда

$$\operatorname{cth} \frac{W}{\Lambda} \approx \frac{\Lambda}{W} + \frac{W}{3\Lambda} = \frac{\Lambda}{W} \left(1 + \frac{W^2}{3\Lambda^2} \right)$$

или

$$\operatorname{cth} \frac{W}{\Lambda} \approx \frac{\Lambda}{W} \left(1 + \frac{W^2}{3L^2} + j\omega\tau \frac{W^2}{3L^2} \right) \approx \frac{\Lambda}{W} \left(1 + j\omega\tau \frac{W^2}{3L^2} \right). \quad (3.47)$$

В полученном выражении слагаемое $W^2/3L^2$ пренебрежимо мало по сравнению с единицей. Хотя мнимое слагаемое тоже мало по сравнению с единицей, но пренебрегать им нельзя, опять-таки чтобы не потерять интересующую нас величину мнимой составляющей проводимости. Такой подход приводит к ошибке в действительной составляющей проводимости. Ошибка получается порядка W^2/L^2 , что, конечно, несущественно, так как $W \ll L$.

Подставив (3.47) в формулу полной проводимости диода (3.34), после сокращения получим

$$\dot{Y} = \frac{q^2 S}{kT} \left[\frac{p_{n0} D_p}{W_n} + \frac{n_{p0} D_n}{W_p} \right] + \frac{i\omega}{3} (p_{n0} W_n + n_{p0} W_p) \exp \frac{qU}{kT}. \quad (3.48)$$

Тогда с учетом (3.31)

$$\frac{1}{r} = \frac{qS}{kT} J_{\text{нас}} \exp \frac{qU}{kT}$$

или

$$\frac{1}{r} = \frac{q}{kT} (I + I_{\text{нас}}). \quad (3.49)$$

Диффузионная емкость диода с тонкой базой

$$C_{\text{диф}} = \frac{q^2 S}{3kT} (p_{n0} W_n + n_{p0} W_p) \exp \frac{qU}{kT}. \quad (3.50)$$

Если $p_{n0} \gg n_{p0}$ или $n_{p0} \gg p_{r0}$, то диффузионная емкость

$$C_{\text{диф}} = \frac{q}{kT} (I + I_{\text{нас}}) \frac{W^2}{3D}, \quad (3.51)$$

где W и D относятся к слаболегированной области, т. е. к базе диода.

Постоянную времени в данном случае выразим, используя формулы (3.50) и (3.51):

$$rC_{\text{диф}} = W^2/(3D). \quad (3.52)$$

4) $W \ll L$, высокие частоты. Критерием высокой частоты для диода с тонкой базой оказывается большое значение аргументов гиперболических котангенсов в (3.34), т. е.

$$\left| \frac{W}{\Lambda} \right| = \left| \frac{W}{L} \sqrt{1 + j\omega\tau} \right| \gg 1$$

или иначе, пренебрегая единицей по сравнению с $\omega\tau$,

$$\frac{W^2}{L^2} \omega\tau \gg 1.$$

Так как аргументы гиперболических котангенсов оказываются большими, расчет параметров диода в данном случае аналогичен расчету при высоких частотах диода с толстой базой. Это относится к тем формулам, в которые еще не введен ток насыщения.

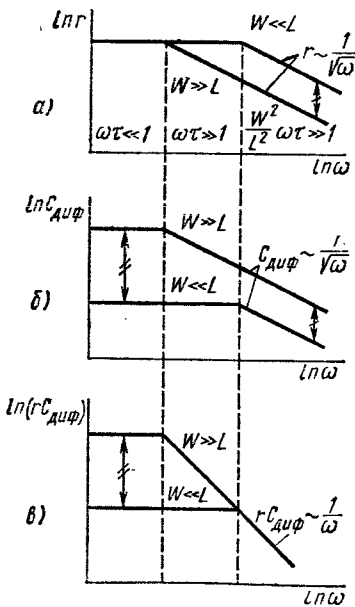


Рис. 3.11. Частотные зависимости активного сопротивления (а), диффузионной емкости (б) и постоянной времени (в) диодов с тонкой и толстой базами

§ 3.7. ГРАФИКИ ЧАСТОТНЫХ ЗАВИСИМОСТЕЙ ПАРАМЕТРОВ ДИОДА

При построении графиков примем одинаковыми плотности тока для диодов с разной толщиной базы. Одинаковыми будем считать и электрофизические параметры материала диодов. Масштаб по осям удобно взять логарифмический (рис. 3.11).

1) Активное сопротивление r . Для диода с толстой базой в области низких частот ($\omega\tau \ll 1$) сопротивление не зависит от частоты в соответствии с (3.37). При высоких частотах ($\omega\tau \gg 1$) сопротивление падает обратно пропорционально корню квадратному из частоты (3.43), что в логарифмическом

масштабе дает прямую линию с угловым коэффициентом $1/2$ (рис. 3.11, а).

Для диода с тонкой базой низкочастотное значение сопротивления оказывается тем же [ср. (3.49) и (3.37)] и сохраняется постоянным до тех пор, пока не нарушится условие малости аргументов гиперболических котангенсов. При $\frac{W^2}{L^2} \omega \tau \gg 1$ сопротивление падает обратно пропорционально корню квадратному из частоты.

2) Диффузионная емкость $C_{диф}$. Диффузионная емкость диода с толстой базой не зависит от частоты при $\omega \tau \ll 1$ (3.39) и обратно пропорциональна корню квадратному из частоты при $\omega \tau \gg 1$ (3.45).

Низкочастотное значение диффузионной емкости диода с тонкой базой значительно меньше, чем диода с толстой базой [ср. (3.51) и (3.39)], так как

$$\frac{W^2}{3D} = \frac{W^2}{3L^2} \tau \ll \frac{\tau}{2} \quad \text{при } W \ll L.$$

В диапазоне частот, когда $\frac{W^2}{L^2} \omega \tau \gg 1$, диффузионная емкость обратно пропорциональна корню квадратному из частоты.

При высоких частотах постоянные времени любого диода должны быть равны между собой и равны $1/\omega$ в соответствии с (3.46), поэтому на рис. 3.11, а и б отмечены равные отрезки в диапазоне высоких частот.

3) Постоянная времени $rC_{диф}$. Построение графиков для постоянной времени следует из двух предыдущих. При низких частотах постоянная времени диода с толстой базой много больше постоянной времени диода с тонкой базой [ср. (3.40) и (3.52)]. При высоких частотах постоянные времени этих диодов одинаковы (рис. 3.11, в).

§ 3.8. ФИЗИЧЕСКИЙ СМЫСЛ ПАРАМЕТРОВ ДИОДА Низкочастотные значения

1) *Сопротивление диода r* представляет собой просто дифференциальное сопротивление диода, т. е. сопротивление диода малому переменному току при постоянном смещении. Чтобы в этом убедиться, надо продифференцировать выражение вольт-амперной характеристики диода (3.36) и сравнить с (3.37):

$$\frac{dI}{dU} = \frac{q}{kT} I_{нас} \exp \frac{qU}{kT} = \frac{q}{kT} (I + I_{нас}) = \frac{1}{r}.$$

2) *Диффузионную емкость* обычно связывают с изменением заряда инжектированных носителей при изменении напряжения на диоде. Действительно, инжектированные носители в течение некоторого времени существуют в областях диода, примыкающих к p - n -переходу. При изменении напряжения часть накопленных неосновных носителей может возвратиться в p - n -переход и пройти

через него в соседнюю область. Образующийся при этом ток аналогичен емкостному току.

Однако к этому требуются некоторые пояснения и дополнения. Дело в том, что при инжекции области, примыкающие к p - n -переходу, остаются нейтральными, т. е. никакой суммарный заряд в них не появляется. Нейтрализация заряда происходит из-за подхода основных носителей в те области, куда произошла инжекция неосновных носителей. Нейтрализация устанавливается за очень малый промежуток времени — порядка времени максвелловской или диэлектрической релаксации (обычно 10^{-11} — 10^{-12} с). Так как концентрация основных носителей относительно велика и необходимое их количество пополняется невыпрямляющим контактом, нейтрализация получается практически полной. Следует заметить, что нейтрализуется не только заряд в среднем по всей области, но и заряд в каждой точке, т. е. выполняется условие локальной электрической нейтральности.

Несмотря на то что при инжекции примыкающие к p - n -переходу области не заряжаются, диффузионную емкость можно связать с зарядом инжектированных носителей, так как и инжектированные неосновные носители и нейтрализующие их основные носители не исчезают. Для сравнения вспомним, что и обычный конденсатор в целом электрически нейтрален. Но в обычном конденсаторе положительный и отрицательный заряды пространственно разделены (то же самое можно сказать и об электронно-дырочном переходе при рассмотрении его барьерной емкости), в то время как при инжекции через p - n -переход и положительный, и отрицательный заряды оказываются в одной и той же области и пространственно не разделяются. В результате чего невозможно обнаружить область, где проходят токи смещения.

Следовательно, диффузионную емкость можно связать с изменением заряда инжектированных неосновных носителей, но нельзя связать с прохождением токов смещения. В этом существенное физическое отличие диффузионной емкости от барьерной емкости p - n -перехода и от емкости обычного конденсатора.

Диффузионную емкость можно представить следующим образом:

$$C_{\text{диф}} = \left| \frac{dQ_{\text{инж.эф}}}{dU} \right|, \quad (3.53)$$

где $dQ_{\text{инж.эф}}$ — эффективное значение инжектированного заряда.

Здесь абсолютная величина отношения взята для того, чтобы не возникала путаница из-за правила знаков для напряжения, а также из-за того, что инжектированный заряд может быть как положительным, так и отрицательным. Эффективное значение инжектированного заряда следует брать потому, что из-за распределенного характера этого заряда он не весь одинаково участвует в образовании емкости. Поэтому приходится проводить какое-то усреднение.

Чтобы убедиться, что данное определение не противоречит полученным ранее значениям диффузионной емкости, можно найти

значение

$$|dQ_{\text{инж}}/dU|,$$

где $Q_{\text{инж}}$ — полный инжектированный заряд.

Полученный при дифференцировании $Q_{\text{инж}}$ по U результат, по-видимому, должен несущественно отличаться от значения диффузионной емкости (3.53). Сравнение $dQ_{\text{инж}}/dU$ с $C_{\text{диф}}$ дает возможность найти и способ выражения эффективного инжектированного заряда через полный заряд инжектированных носителей.

Запишем выражение для полного инжектированного заряда, учитывая, что он складывается в общем случае из заряда дырок, инжектированных в n -область, и электронов, инжектированных в p -область. Тогда в системе координат, принятой на рис. 3.5,

$$Q_{\text{инж}} = qS \left(\int_0^{W_n} \Delta p_n dx + \int_{-W_p}^{0'} \Delta n_p dx \right). \quad (3.54)$$

Применим это соотношение к частным случаям.

а) **Диод с толстой базой.** Для диода с толстой базой распределение избыточной концентрации дырок в n -области соответствует выражению (3.28). Аналогичное выражение будет и для распределения избыточной концентрации инжектированных в p -область электронов. При интегрировании положим $W_n = \infty$ и $W_p = \infty$. Тогда

$$Q_{\text{инж}} = qS (p_{n0}L_p + n_{p0}L_n) \left(\exp \frac{qU}{kT} - 1 \right)$$

и

$$\left| \frac{dQ_{\text{инж}}}{dU} \right| = \frac{q^2 S}{kT} (p_{n0}L_p + n_{p0}L_n) \exp \frac{qU}{kT}. \quad (3.55)$$

Полученный результат отличается от выражения для диффузионной емкости диода с толстой базой (3.38) сравнительно небольшим множителем 2, который можно отнести за счет усреднения. Следовательно, для диода с толстой базой $Q_{\text{инж.эф}} = \frac{1}{2} Q_{\text{инж}}$.

б) **Диод с тонкой базой.** Для диода с тонкой базой распределение избыточной концентрации дырок в n -области соответствует выражению (3.30). Аналогичное выражение будет и для распределения избыточной концентрации инжектированных в p -область электронов. После интегрирования получим

$$Q_{\text{инж}} = \frac{1}{2} qS (p_{n0}W_n + n_{p0}W_p) \left(\exp \frac{qU}{kT} - 1 \right)$$

и

$$\left| \frac{dQ_{\text{инж}}}{dU} \right| = \frac{q^2 S}{2kT} (p_{n0}W_n + n_{p0}W_p) \exp \frac{qU}{kT}. \quad (3.56)$$

Этот результат всего в 1,5 раза отличается от выражения для диффузионной емкости диода с тонкой базой (3.50), что также можно отнести за счет усреднения.

Приведенные примеры показывают, что диффузионную емкость диода можно связать с изменением заряда инжектированных носителей, но надо учитывать эффективный инжектированный заряд.

Иногда в качестве определения диффузионной емкости берут

$$C_{\text{диф}} = \left| \frac{dQ_{\text{инж}}}{dU} \right|.$$

Такой способ вычисления диффузионной емкости неудобен, так как для расчета полной проводимости диода в его эквивалентную схему придется подставлять не полное значение $C_{\text{диф}}$, а некоторую ее часть, зависящую к тому же от конструкции диода ($1/2$ для диода с толстой базой и $1/1.5$ для диода с тонкой базой).

3) *Постоянная времени.* Если продолжить аналогию полупроводникового диода с конденсатором, то можно выяснить физический смысл постоянной времени $\tau C_{\text{диф}}$. Для конденсатора постоянная времени показывает, за какое время его заряд уменьшится в e раз, т. е. постоянная времени характеризует время исчезновения заряда конденсатора.

Постоянная времени диода с толстой базой при низкой частоте $\tau C_{\text{диф}} = \tau/2$ (3.40) тоже характеризует время исчезновения заряда. Действительно, τ — время жизни неосновных носителей как раз и показывает, в течение какого времени концентрация неосновных носителей изменится в e раз из-за рекомбинации.

Для диода с тонкой базой постоянная времени при низкой частоте в соответствии с (3.52) $\tau C_{\text{диф}} = W^2/(3D)$.

Чтобы разобраться в физическом смысле этого результата, вспомним, что в данном случае из-за малой толщины базы процесс рекомбинации носителей в ее объеме уже не является определяющим. Главным теперь оказывается уход неосновных носителей заряда на невыпрямляющий контакт с последующей рекомбинацией. Отсюда следует, что время существования заряда в базе теперь будет определяться временем пролета носителей заряда через базу диода.

Чтобы убедиться в этом, выполним расчет времени пролета неосновных носителей через базу диода. Плотность дырочного тока в общем случае

$$J_p = qp v_p = qp \frac{dx}{dt},$$

отсюда

$$dt = \frac{qp}{J_p} dx.$$

Проинтегрировав последнее выражение в пределах от 0 до $t_{\text{прол}}$ по времени и соответственно от 0 до W_n по координате x , получим время пролета:

$$t_{\text{прол}} = q \int_0^{W_n} \frac{p}{J_p} dx.$$

Для диода с тонкой базой, с учетом (3.30),

$$p_n(x) = p_{n0} + \Delta p_n(x) = p_{n0} + p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \left(1 - \frac{x}{W_n} \right),$$

а плотность дырочного тока одинаковая во всех сечениях, с учетом (3.27) и (3.31),

$$J_p = \frac{q p_{n0} D_p}{W_n} \left(\exp \frac{qU}{kT} - 1 \right).$$

При расчете для упрощения пренебрежем равновесной концентрацией неосновных носителей, что соответствует большим прямым напряжениям на p - n -переходе. Тогда

$$t_{\text{прол}} = q \int_0^{W_n} \frac{W_n - x}{q D_p} dx = \frac{W_n^2}{2 D_p}. \quad (3.57)$$

Сравнение (3.57) с постоянной времени диода с тонкой базой (3.52) подтверждает, что она определяется временем пролета неосновных носителей через базу диода.

Высокочастотные значения

Для высокочастотных значений сопротивления и диффузионной емкости диода характерна их частотная зависимость. Это сильно ограничивает использование таких параметров диода, так как затрудняет расчет частотных характеристик схем с использованием полупроводниковых диодов.

С точки зрения причин такой частотной зависимости можно сказать, что она появилась из-за того, что система с распределенными параметрами (диод на высокой частоте) была представлена моделью диода с сосредоточенными параметрами. Такую модель нельзя считать удачной для высокой частоты. Отсюда следует, что искать физический смысл высокочастотных параметров полупроводникового диода нет смысла — их надо рассматривать как формальные.

§ 3.9. ПРЕДЕЛЫ ПРИМЕНИМОСТИ ЧАСТНЫХ СЛУЧАЕВ РАСЧЕТА ПАРАМЕТРОВ ДИОДА

Предыдущие расчеты были приведены для частных случаев диодов с толстой и тонкой базами, при низкой и высокой частотах. Однако полученные выражения со сравнительно небольшими погрешностями могут применяться практически всегда. Для того чтобы в этом убедиться, рассмотрим ошибки, которые могут возникать по разным причинам.

1. Размеры базы диода. Частные случаи толщины базы ($W \gg L$ и $W \ll L$) использовались для того, чтобы избавиться от гиперболических функций в формулах. По-видимому, наибольшую погрешность полученные соотношения будут иметь при $W \approx L$.

Пусть $W = L$, тогда

$$\text{cth} \frac{W}{L} = \text{cth} 1 = 1,31.$$

Если рассматривать такой диод как диод с толстой базой, то $\text{cth} \frac{W}{L} \approx 1$, что дает погрешность 24 % от точного значения.

Если же рассматривать такой диод как диод с тонкой базой, то при разложении гиперболического котангенса в ряд $\text{cth} \frac{W}{L} \approx \approx \frac{L}{W} = 1$, что дает ту же погрешность 24 %. Для расчета токов в полупроводниковом диоде эту погрешность можно считать умеренной.

2. Частотные диапазоны. Аналогично сказанному, можно считать, что наибольшая погрешность приближенных расчетов будет при $\omega\tau \approx 1$. Пусть $\omega\tau = 1$. Точное выражение для

$$\sqrt{1+j\omega\tau} = \sqrt{1+j} = \sqrt[4]{2} \left(\cos \frac{\pi}{8} + j \sin \frac{\pi}{8} \right) = 1,1 + j0,45.$$

Если принять приближение $\omega\tau \ll 1$, то раскладывая $\sqrt{1+j}$ в биномиальный ряд, получим $\sqrt{1+j} \approx 1,0 + j0,5$, что дает погрешность около 10 % для действительной и мнимой частей.

Если же использовать методику расчета параметров такого диода, соответствующую высокой частоте, то

$$\begin{aligned} \sqrt{1+j} &\approx \sqrt{j} = \sqrt{\cos 90^\circ + j \sin 90^\circ} = \\ &= \cos 45^\circ + j \sin 45^\circ = 0,71 + j0,71, \end{aligned}$$

что дает погрешность действительной части около 35 %, а мнимой — около 60 %, т. е. получается правильный порядок.

Таким образом, судя по максимально возможным погрешностям, для практических расчетов можно ограничиться только рассмотренными частными случаями.

§ 3.10. ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА В ЭЛЕКТРОННО-ДЫРОЧНОМ ПЕРЕХОДЕ

Толщина электронно-дырочного перехода в диодах значительно меньше диффузионной длины носителей заряда. Однако процессы генерации и рекомбинации носителей в электронно-дырочном переходе идут более интенсивно, чем в прилегающих к *p-n*-переходу областях. Поэтому составляющие тока полупроводникового диода, связанные с генерацией и рекомбинацией носителей в электронно-дырочном переходе, могут быть существенными.

При обратном напряжении на диоде образующиеся из-за тепловой генерации носители разного знака растаскиваются электрическим полем перехода в разные стороны и уже не могут рекомбинировать (рис. 3.12, *a*). Именно поэтому процесс генерации носителей в переходе не уравнивается процессом рекомбинации. Обратный ток, вызванный генерацией носителей в электронно-дырочном переходе, называют генерационным током ($I_{ген}$). Следует, однако, помнить, что ток насыщения, связанный

с экстракцией неосновных носителей и рассмотренный в § 3.2 и 3.4, также вызван генерацией неосновных носителей заряда, но генерацией в прилегающих к переходу областях.

Чтобы выяснить влияние генерационного тока на обратный ток через диод, проведем качественное сравнение плотности тока насыщения и плотности тока генерации. Предположим, что диод имеет симметричный p - n -переход и все электрофизические параметры прилегающих к переходу областей также одинаковы, т. е. $p_{n0} = n_{p0}$, $L_p = L_n = L$, $\tau_p = \tau_n = \tau$. Тогда в соответствии с (3.1) плотность тока насыщения в диоде с $W \gg L$

$$J_{\text{нас}} = \frac{2q p_{n0} L}{\tau}.$$

Исходя из физики процессов образования тока насыщения и тока генерации можно по аналогии с предыдущим выражением записать соотношение, показывающее, чему пропорционален ток генерации. При этом следует учесть, что если в образовании тока насыщения принимали участие неосновные носители с концентрацией p_{n0} , возникающие в объеме толщиной $2L$, то в образовании тока генерации принимают участие носители с концентрацией n_i , возникающие в p - n -переходе толщиной δ . Поэтому

$$J_{\text{ген}} \sim \frac{q n_i \delta}{\tau}.$$

Тогда отношение плотности тока генерации к плотности тока насыщения с учетом (1.19) пропорционально следующим величинам:

$$\frac{J_{\text{ген}}}{J_{\text{нас}}} \sim \frac{n_i \delta}{p_{n0} L} = \frac{n_{n0} \delta}{n_i L}.$$

Таким образом, во-первых, процесс генерации носителей в p - n -переходе следует учитывать для диодов, изготовленных из полупроводника с большой шириной запрещенной зоны, т. е. с малой собственной концентрацией носителей n_i . Чтобы лучше осознать этот вывод, надо учесть, что плотность тока насыщения в соответствии с (3.2) пропорциональна n_i^2 . Так, например, в кремние-

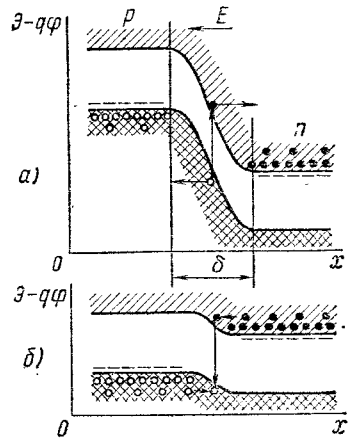


Рис. 3.12. Генерация (а) и рекомбинация (б) носителей в p - n -переходе

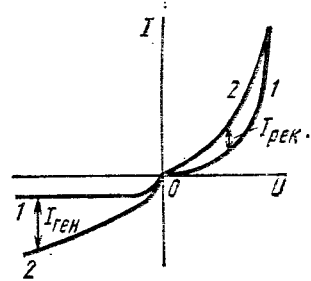


Рис. 3.13. Вольт-амперные характеристики диода без учета генерации и рекомбинации носителей заряда в p - n -переходе (1) и с учетом этих процессов (2)

вых диодах обратный ток определяется током генерации, а в германиевых диодах — током насыщения.

Во-вторых, с увеличением обратного напряжения по абсолютному значению величина тока генерации также растет в связи с увеличением толщины p - n -перехода δ . Поэтому в диодах с преобладающим током генерации нет участка постоянного тока на обратной ветви вольт-амперной характеристики (рис. 3.13).

В-третьих, соотношение между током генерации и током насыщения изменяется при изменении температуры — с повышением температуры оба тока растут, но ток насыщения растет быстрее (как n_i^2), а ток генерации медленнее (как n_i). Значит, с повышением температуры относительная роль тока генерации уменьшается.

В-четвертых, соотношение $J_{ген}/J_{нас}$ изменяется при изменении концентрации примесей ($n_{i,0} \approx N_{эп}$) — с увеличением концентрации примеси в базе диода относительная роль тока генерации также увеличивается.

При прямом напряжении на диоде носители разных знаков подходят к электронно-дырочному переходу. Если прямое напряжение мало, то высота потенциального барьера на переходе велика и основная часть носителей не может преодолеть потенциальный барьер перехода, но вблизи середины p - n -перехода может происходить их рекомбинация (рис. 3.12, б). Составляющую прямого тока, связанную с процессом рекомбинации носителей в электронно-дырочном переходе, называют **рекомбинационным током**. Надо понимать условность этого термина, так как прямой ток, связанный с инжекцией неосновных носителей в прилегающие к переходу области (см. § 3.2), также сопровождается рекомбинацией инжектированных носителей либо в базе диода, либо на невыпрямляющем контакте диода.

При больших для диода прямых напряжениях высота потенциального барьера на переходе оказывается малой. Поэтому прямой ток при больших прямых напряжениях будет вызван в основном инжекцией носителей через уменьшенный потенциальный барьер перехода. Следовательно, рекомбинационный ток может сказываться на величине прямого тока через диод только при малых прямых напряжениях (рис. 3.13).

Плотность рекомбинационного тока в диоде с симметричным p - n -переходом

$$J_{рек} = \frac{qn_i\delta}{\tau} \cdot \frac{kT}{q(\varphi_{кон} - U)} \left[\exp \frac{qU}{2kT} - \exp \left(-\frac{qU}{2kT} \right) \right]. \quad (3.58)$$

Тогда для диода с $W \gg L$, учитывая (3.2) и (3.27), получим

$$\frac{J_{рек}}{J_{инж}} \approx \frac{N\delta}{2n_iL} \cdot \frac{kT}{q(\varphi_{кон} - U)} \exp \left(-\frac{qU}{2kT} \right).$$

Отношение величин рекомбинационного тока к инжекционному зависит от ширины запрещенной зоны полупроводника. В диодах из материала с большой шириной запрещенной зоны (кремниевый

диод) из-за большой высоты потенциального барьера инжекция носителей затруднена. Поэтому прямой ток при малых прямых напряжениях будет определяться рекомбинационным током.

§ 3.11. ЛАВИННЫЙ ПРОБОЙ

Механизм лавинного пробоя и основные определения

Пробой диода — это явление резкого увеличения обратного тока через диод при достижении обратным напряжением некоторого критического для данного диода значения. В зависимости от физических явлений, приводящих к пробую, различают лавинный, туннельный и тепловой пробой.

Лавинный пробой связан с образованием лавины носителей заряда под действием сильного электрического поля, при котором носители приобретают энергии, достаточные для образования новых электронно-дырочных пар в результате ударной ионизации атомов полупроводника.

Процесс ударной ионизации характеризуют коэффициентами ударной ионизации α_n и α_p (см. § 1.10). Чтобы количественно охарактеризовать увеличение тока из-за процесса ударной ионизации в p - n -переходе диода вводят коэффициенты лавинного умножения M_n и M_p , показывающие, во сколько раз увеличивается ток данных носителей в результате ударной ионизации. В связи с предположением о равенстве коэффициентов ионизации автоматически получается равенство коэффициентов лавинного умножения: $M_n = M_p = M$.

Наконец, параметром диода или его выпрямляющего перехода, характеризующим явление пробоя, является пробивное напряжение — это напряжение, при котором происходит неограниченное возрастание тока (рис. 3.14). Формально при пробивном напряжении $M \rightarrow \infty$. В производственных условиях пробивное напряжение определяют при достижении обратным током заданного значения.

Связь коэффициента лавинного умножения с коэффициентом ударной ионизации и вывод условия лавинного пробоя

Для расчета связи параметра, характеризующего физический процесс — коэффициента ударной ионизации, с параметром, характеризующим электронно-дырочный переход — коэффициентом лавинного умножения, воспользуемся уравнением непрерывности, например для электронов, которое имеет вид, аналогичный виду уравнения непрерывности для дырок (3.5):

$$\frac{\partial n}{\partial t} = \frac{1}{q} \operatorname{div} \mathbf{J}_n - R_n + G_n. \quad (3.59)$$

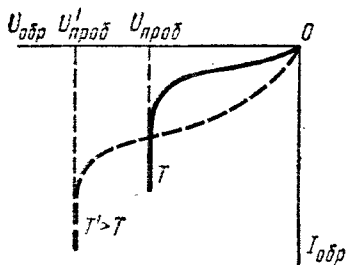


Рис. 3.14. Вольт-амперные характеристики диода при лавинном пробое

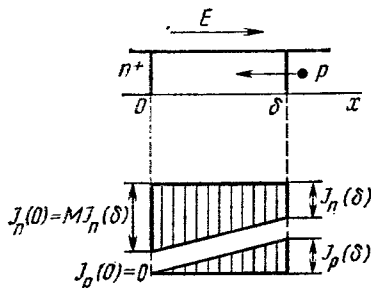


Рис. 3.15. Плотности электронного и дырочного токов при лавинном умножении в несимметричном n - p -переходе

Если при рассмотрении установившегося процесса ударной ионизации ($\partial n/\partial t = 0$) пренебречь рекомбинацией в p - n -переходе при большом обратном напряжении на диоде ($R_n = 0$), то уравнение непрерывности для одномерной модели диода примет следующий вид:

$$-\frac{1}{q} \cdot \frac{dJ_n}{dx} = G_n. \quad (3.60)$$

В данном случае скорость тепловой генерации электронов (G_n) учитывает их генерацию под действием сильного электрического поля. Скорость нетепловой генерации численно равна числу носителей, образующихся в единицу времени в единице объема. Чтобы выразить скорость нетепловой генерации через коэффициенты ударной ионизации, рассмотрим единичный объем, имеющий длину в направлении прохождения носителей и площадь поперечного сечения, равные единице. Тогда каждый проходящий через этот объем носитель образует в единичном объеме количество носителей, равное коэффициенту

ионизации α . В единицу времени через рассматриваемый объем проходит число носителей, равное J/q . Следовательно, если учесть генерацию электронов (или дырок) в результате ионизации атомов электронами и дырками, получим

$$G_n = G_p = \alpha \frac{J_n}{q} + \alpha \frac{J_p}{q}. \quad (3.61)$$

Подставим (3.61) в (3.60). Тогда

$$-\frac{dJ_n}{dx} = \alpha (J_n + J_p) = \alpha J. \quad (3.62)$$

Для решения дифференциального уравнения (3.62) выясним граничные условия. Для несимметричного p - n -перехода (например, n^+ - p) можно считать, что в переход входит только поток электронов, соответствующий плотности электронного тока $J_n(\delta)$ (рис. 3.15). В соответствии с определением коэффициента лавинного умножения выходящая из перехода плотность электронного тока равна $MJ_n(\delta)$. Эта же плотность тока является полной плотностью тока через рассматриваемый переход, так как через левую границу несимметричного n^+ - p -перехода дырочная составляющая тока не про-

ходит. Таким образом, граничные условия принимают следующий вид:

$$\left. \begin{array}{l} \text{при } x=0 \quad J_n(0) = MJ_n(\delta) = J; \quad J_p = J_p(0) = 0; \\ \text{при } x=\delta \quad J_n = J_n(\delta); \quad J_p = J_p(\delta). \end{array} \right\} \quad (3.63)$$

С учетом проведенных рассуждений уравнение (3.62) примет вид

$$- \frac{dJ_n}{dx} = \alpha MJ_n(\delta)$$

или

$$- dJ_n = \alpha MJ_n(\delta) dx. \quad (3.64)$$

При интегрировании дифференциального уравнения (3.64) надо согласовать пределы интегрирования, используя граничные условия (3.63), которые для наглядности отражены на рис. 3.15. Тогда

$$- \int_{J_n(\delta)}^{MJ_n(\delta)} dJ_n = MJ_n(\delta) \int_{\delta}^0 \alpha dx$$

или

$$- MJ_n(\delta) + J_n(\delta) = MJ_n(\delta) \int_{\delta}^0 \alpha dx.$$

В результате

$$1 - \frac{1}{M} = \int_{\delta}^0 \alpha dx. \quad (3.65)$$

Соотношение (3.65) отражает связь коэффициента лавинного умножения с коэффициентом ударной ионизации. При пробое электронно-дырочного перехода коэффициент лавинного умножения $M \rightarrow \infty$. Тогда

$$\int_{\delta}^0 \alpha dx = 1. \quad (3.66)$$

Уравнение (3.66) является *условием лавинного пробоя электронно-дырочного перехода*.

Расчет коэффициента лавинного умножения и пробивного напряжения при лавинном пробое резкого электронно-дырочного перехода

Распределение напряженности электрического поля в резком несимметричном n^+p -переходе линейное (2.18), и практически вся область объемного заряда расположена в слаболегированной области полупроводника (2.22). Тогда в соответствии с (2.18) и с учетом изменения принятого начала координат ($\delta \approx \delta_p$)

$$|E| = \frac{qN_{ap}}{\epsilon\epsilon_0} (\delta - x).$$

Учитывая соотношение между коэффициентами лавинного умножения и ударной ионизации (3.65) и принимая аппроксимацию зависимости коэффициента ударной ионизации от напряженности электрического поля (1.33), получим

$$1 - \frac{1}{M} = \int_0^{\delta} \alpha dx = A \left(\frac{qN_{ap}}{\varepsilon\varepsilon_0} \right)^m \int_0^{\delta} (\delta - x)^m dx = A \left(\frac{qN_{ap}}{\varepsilon\varepsilon_0} \right)^m \frac{\delta^{m+1}}{m+1}. \quad (3.67)$$

Подставим в последнее выражение соотношение для толщины резкого несимметричного электронно-дырочного перехода (2.27), в котором можно пренебречь значением контактной разности потенциалов по сравнению с обратным напряжением ($\varphi_{кон} \ll |U|$). Тогда

$$1 - \frac{1}{M} = \frac{A}{m+1} \left(\frac{qN_{ap}}{\varepsilon\varepsilon_0} \right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_{ap}} \right)^{\frac{m+1}{2}} |U|^{\frac{m+1}{2}}. \quad (3.68)$$

С помощью этого соотношения можно рассчитать коэффициент лавинного умножения, зная концентрацию примесей в базе диода и другие параметры материала базы.

При лавинном пробое (при $U_{проб}$) коэффициент лавинного умножения $M \rightarrow \infty$. Поэтому

$$1 = \frac{A}{m+1} \left(\frac{qN_{ap}}{\varepsilon\varepsilon_0} \right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_{ap}} \right)^{\frac{m+1}{2}} U_{проб}^{\frac{m+1}{2}}. \quad (3.69)$$

Разделив (3.68) на (3.69), получим

$$1 - 1/M = (U/U_{проб})^{\frac{m+1}{2}},$$

т. е.

$$M = \frac{1}{1 - (U/U_{проб})^{\frac{m+1}{2}}}$$

или

$$M = \frac{1}{1 - (U/U_{проб})^b}, \quad (3.70)$$

где $b = (m+1)/2$ — коэффициент, различный для диодов из разных материалов ($b = 2 \div 6$).

При $U \rightarrow 0$ $M \rightarrow 1$, при $U \rightarrow U_{проб}$ $M \rightarrow \infty$ (рис. 3.16).

Из выражения (3.69) пробивное напряжение при лавинном пробое

$$U_{проб}^{\frac{m+1}{2}} = \frac{m+1}{A} \left(\frac{1}{2} \right)^{\frac{m+1}{2}} \left(\frac{\varepsilon\varepsilon_0}{q} \right)^{m - \frac{m+1}{2}} \left(\frac{1}{N_{ap}} \right)^{m - \frac{m+1}{2}}$$

или

$$U_{проб} = \left[\frac{1}{2} \left(\frac{m+1}{A} \right)^{\frac{2}{m+1}} \left(\frac{\varepsilon\varepsilon_0}{q} \right)^{\frac{m-1}{m+1}} \right] \frac{1}{N_{ap}^{\frac{m-1}{m+1}}}. \quad (3.71)$$

Рис. 3.16. Зависимость коэффициента лавинного умножения от напряжения на электронно-дырочном переходе

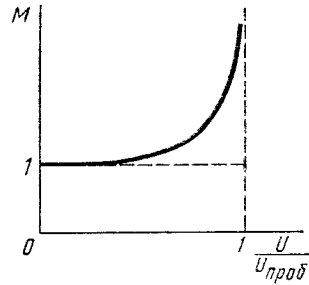
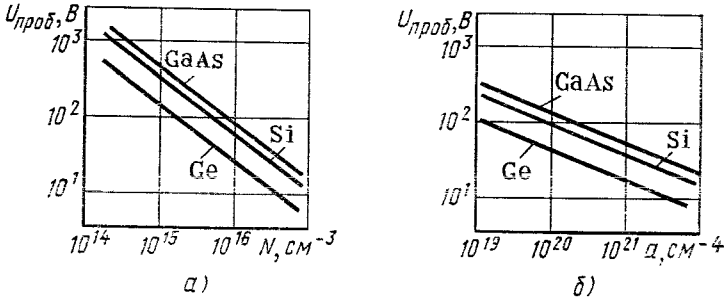


Рис. 3.17. Зависимости пробивного напряжения при лавинном пробое от концентрации примесей в базе диода с несимметричным резким p - n -переходом (а) и от градиента концентрации примесей в плавном p - n -переходе (б)



Здесь выражение в квадратных скобках включает в себя только величины, которые можно считать постоянными для данного материала. Тогда, вводя новые обозначения, получим

$$U_{\text{проб}} = B_1 \frac{1}{N^k}, \quad (3.72)$$

где N — концентрация примесей в слаболегированной области, т. е. в базе диода; $k = (m - 1) / (m + 1)$.

Эксперименты подтверждают соотношение (3.72). В системе координат $\lg U_{\text{проб}}$ от $\lg N$ экспериментальные точки ложатся на прямую линию (рис. 3.17, а). При этом для несимметричных электронно-дырочных переходов p^+n и n^+p , сформированных в одном и том же полупроводнике, зависимости $\lg U_{\text{проб}} = \lg N$ совпадают. Этот экспериментальный факт доказывает правомерность допущения равенства коэффициентов ударной ионизации для электронов и дырок, что было принято в начале этого параграфа.

Однако чаще бывает известна не концентрация примесей в базе диода, а ее удельное сопротивление, т. е. удельное сопротивление исходного полупроводника. Поэтому, учтя связь концентрации примеси или концентрации основных носителей с удельным сопротивлением, соотношение (3.72) удобнее записать в следующем виде:

$$U_{\text{проб}} = B\rho^a. \quad (3.73)$$

Для кремниевых p^+n -переходов $U_{\text{проб}} = 96\rho^{0,78}$; n^+p -переходов $U_{\text{проб}} = 48\rho^{0,78}$; для германиевых p^+n -переходов $U_{\text{проб}} = 100\rho^{0,8}$; n^+p -переходов $U_{\text{проб}} = 55\rho^{0,8}$, где ρ — удельное сопротивление базы, Ом·см.

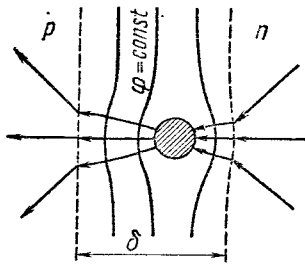


Рис. 3.18. Искажения линий тока и эквипотенциалей в p - n -переходе и в прилегающих к нему областях, связанные с проводящим дефектом в p - n -переходе

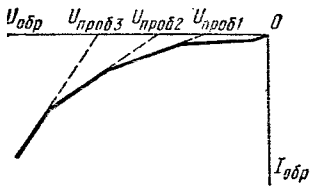


Рис. 3.19. Вольт-амперная характеристика диода при пробое по дефектам

рости на длине свободного пробега в электрическом поле. Следовательно, повышение температуры приводит к увеличению пробивного напряжения при лавинном пробое (см. рис. 3.14).

Особенности лавинного пробоя в электронно-дырочных переходах с дефектами

Электронно-дырочные переходы реальных диодов часто имеют различные дефекты: нарушения кристаллической решетки, инородные включения и т. п. Такие неоднородности, независимо от их характера, приводят к нарушениям картины электрического поля, к появлению областей с повышенной напряженностью. В тех местах, где напряженность электрического поля оказывается больше, развивается пробой. После развития пробоя в районе дефекта образуется шнур с повышенной удельной проводимостью. Однако из-за малого сечения шнура последовательно с ним оказывается включенным большое так называемое сопротивление растекания, сосредоточенное в примыкающих к переходу областях (рис. 3.18). Поэтому ток при лавинном пробое по дефектам ограничен большим сопротивлением растекания.

В связи с тем, что количество различных дефектов в p - n -переходе может быть большим, и они по своим свойствам как-то рас-

Эмпирические коэффициенты B и a различны не только для диодов из разных материалов, но и для диодов из одного и того же материала с разными резкими электронно-дырочными переходами (p^+-n и n^+-p). Связано это различие в коэффициентах с тем, что подвижность электронов отличается от подвижности дырок в одном и том же материале.

Таким образом, пробивное напряжение резких несимметричных электронно-дырочных переходов определяется концентрацией примеси в слаболегированной области или ее удельным сопротивлением, так как от этих величин зависит толщина p - n -перехода.

Толщина плавных p - n -переходов зависит от градиента концентрации примесей (2.33). Поэтому пробивное напряжение плавных p - n -переходов определяется градиентом концентрации примесей (рис. 3.17, б).

С повышением температуры уменьшается длина свободного пробега носителей заряда, а значит, и энергия, которую носитель заряда может приобре-

пределены статистически, пробой по каждому из них приводит к некоторому росту тока через диод. Таким образом, при лавинном пробое *p-n*-перехода по дефектам получается «мягкая» обратная ветвь вольт-амперной характеристики (рис. 3.19).

С явлениями лавинного пробоя по дефектам связано также, по-видимому, то, что экспериментальные данные пробивных напряжений и коэффициентов лавинного умножения получаются с большим разбросом.

§ 3.12. ТУННЕЛЬНЫЙ ПРОБОЙ

Туннельным пробоем электронно-дырочного перехода называют электрический пробой перехода, вызванный квантово-механическим туннелированием носителей заряда сквозь запрещенную зону полупроводника без изменения их энергии (см. рис. 1.12). Туннелирование электронов возможно при условии, если ширина потенциального барьера Δ , который необходимо преодолеть электронам, достаточно мала. При одной и той же ширине запрещенной зоны (для одного и того же материала) ширина потенциального барьера определяется напряженностью электрического поля, т. е. наклоном энергетических уровней и зон (см. рис. 1.2). Следовательно, условия для туннелирования возникают только при определенной напряженности электрического поля или при определенном напряжении на электронно-дырочном переходе — при пробивном напряжении. Значение этой критической напряженности электрического поля составляет примерно $8 \cdot 10^5$ В/см для кремниевых переходов и $3 \cdot 10^5$ В/см — для германиевых. Так как вероятность туннелирования очень сильно зависит от напряженности электрического поля, то внешне туннельный эффект проявляется как пробой диода (рис. 3.20).

Рассмотрим в качестве примера расчет пробивного напряжения при туннельном пробое резкого близкого к симметричному электронно-дырочного перехода.

Наибольшее значение напряженности электрического поля в таком переходе существует на металлургическом контакте [см. § 2.6 и, в частности, формулу (2.18)], т. е. при $x = 0$

$$|E|_{\max} = \frac{qN_{ap}}{\epsilon\epsilon_0} \delta_p$$

Воспользовавшись уравнением (2.23), получим

$$|E|_{\max} = \frac{qN_{ap}}{\epsilon\epsilon_0} \cdot \frac{N_{дп}}{N_{ap} + N_{дп}} \delta.$$

С учетом соотношения для толщины резкого электронно-дырочного перехода (2.26) получим

$$E_{\max}^2 = \frac{2q}{\epsilon\epsilon_0} \cdot \frac{N_{ap}N_{дп}}{N_{ap} + N_{дп}} (\varphi_{\text{кон}} - u).$$

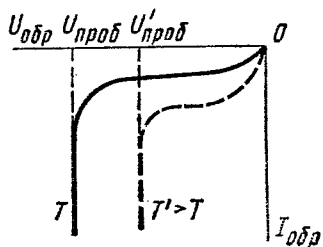


Рис. 3.20. Вольт-амперные характеристики диода при туннельном пробое

При $E_{\text{max}} = E_{\text{кр}}$ напряжение на электронно-дырочном переходе будет пробивным. Поэтому

$$U_{\text{проб}} = \frac{\varepsilon\varepsilon_0 E_{\text{кр}}^2}{2q} \cdot \frac{N_{ap} + N_{дп}}{N_{ap}N_{дп}} - \varphi_{\text{кон}}$$

или

$$U_{\text{проб}} = \frac{\varepsilon\varepsilon_0 E_{\text{кр}}^2}{2q} \left(\frac{1}{N_{ap}} + \frac{1}{N_{дп}} \right) - \varphi_{\text{кон}}. \quad (3.74)$$

Следовательно, при туннельном пробое пробивное напряжение оказывается обратно пропорциональным концентрации примесей в первой степени.

Туннельный пробой может происходить только в электронно-дырочных переходах, изготовленных в полупроводниках с большой концентрацией примесей, так как для туннелирования необходима малая толщина потенциального барьера и, следовательно, малая ширина перехода. При малой ширине перехода пробивные напряжения его оказываются также малыми. Поэтому пренебрегать величиной контактной разности потенциалов в выражении (3.74) не следует — она может быть сравнима с пробивным напряжением.

Если в (3.74) от концентраций примесей перейти к удельным сопротивлениям и все-таки пренебречь контактной разностью потенциалов, то получим широко известное соотношение

$$U_{\text{проб}} = C_p \rho_p + C_n \rho_n. \quad (3.75)$$

Для кремневых электронно-дырочных переходов после подстановки численных значений параметров пробивное напряжение при туннельном пробое

$$U_{\text{проб}} = 200\rho_n + 73\rho_p;$$

для германиевых переходов

$$U_{\text{проб}} = 190\rho_n + 94\rho_p,$$

где ρ_n и ρ_p — удельные сопротивления n - и p -областей, прилегающих к переходу, Ом·см.

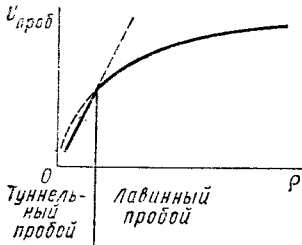


Рис. 3.21. Зависимость пробивного напряжения при туннельном и лавинном пробое от удельного сопротивления базы диода

С повышением температуры у большинства полупроводников ширина запрещенной зоны уменьшается. Следовательно, при этом уменьшается и толщина потенциального барьера при той же напряженности электрического поля, что приводит к увеличению вероятности туннелирования носителей сквозь потенциальный барьер. Поэтому пробивное напряжение при туннельном пробое уменьшается с увеличением температуры (см. рис. 3.20).

Так как для туннельного пробоя необходима малая толщина p - n -перехода,

он наблюдается в диодах, изготовленных на основе полупроводников с большой концентрацией примесей. Напряжения туннельного пробоя поэтому не превосходят нескольких вольт. В диодах из полупроводников с меньшей концентрацией примесей и, значит, с большей толщиной p - n -перехода пробой будет происходить при больших напряжениях и будет иметь лавинный характер (рис. 3.21).

§ 3.13. ТЕПЛОВОЙ ПРОБОЙ

Вольт-амперная характеристика с учетом тепловыделения в диоде

Тепловой пробой диода — это пробой, развитие которого обусловлено выделением в выпрямляющем электрическом переходе тепла вследствие прохождения тока через переход. При подаче на диод обратного напряжения практически все оно падает на p - n -переходе, через который идет, хотя и небольшой, обратный ток. Выделяющаяся на p - n -переходе мощность

$$P_{\text{выт}} = U_{\text{обр}} I_{\text{обр}} \quad (3.76)$$

вызывает разогрев p - n -перехода и прилегающих к нему областей полупроводника.

Отводимая от p - n -перехода мощность в результате теплопроводности и дальнейшего рассеяния тепла в окружающую среду пропорциональна перегреву p - n -перехода ($T - T_{\text{окр}}$) и обратно пропорциональна тепловому сопротивлению конструкции диода $R_{T\text{пс}}$:

$$P_{\text{отв}} = \frac{T - T_{\text{окр}}}{R_{T\text{пс}}}. \quad (3.77)$$

Таким образом, тепловое сопротивление диода между p - n -переходом и окружающей средой определяется мощностью, отводимой от перехода в окружающую среду при разности температур между ними в один градус.

Через некоторое время после подачи на диод обратного напряжения устанавливается тепловое равновесие между выделяемым и отводимым теплом. При этом $P_{\text{выт}} = P_{\text{отв}}$ или

$$U_{\text{обр}} I_{\text{обр}} = \frac{T - T_{\text{окр}}}{R_{T\text{пс}}}. \quad (3.78)$$

Учтем теперь зависимость обратного тока от температуры. Для простейшего случая, когда обратный ток определяется экстракцией неосновных носителей в диоде с толстой базой (3.2) или (3.29),

$$I_{\text{обр}} = J_{\text{нас}} S = q S n_i^2 \left(\frac{L_p}{N_{\text{дн}} \tau_p} + \frac{L_n}{N_{\text{ап}} \tau_n} \right).$$

С учетом (1.15), пренебрегая слабой зависимостью от температуры концентрации примесей, диффузионной длины и времени жизни,

$$I_{\text{обр}} = I_{\infty} \exp \left(- \frac{\Delta \mathcal{E}}{kT} \right), \quad (3.79)$$

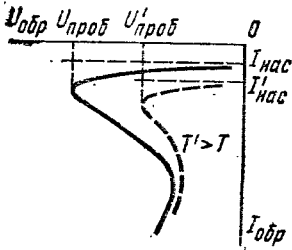


Рис. 3.22. Вольт-амперная характеристика диода с учетом тепловыделения в $p-n$ -переходе при разных температурах окружающей среды

ловыделения в $p-n$ -переходе. Общий вид такой кривой показан на рис. 3.22. Кривая идет не из начала координат, так как обратный ток был принят равным току насыщения, т. е. нас не интересовала начальная часть характеристики.

На кривой могут быть два экстремума — максимум и минимум напряжения. Между этими экстремумами находится участок вольт-амперной характеристики с отрицательным дифференциальным сопротивлением. Возможность существования отрицательного дифференциального сопротивления объясняется тем, что с увеличением выделяемой мощности растет температура, а следовательно, увеличиваются ток и выделяемая мощность. Таким образом, в диоде возникает внутренняя положительная обратная связь, которая и может привести к появлению теплового пробоя и отрицательного дифференциального сопротивления.

Проведем исследование экстремумов вольт-амперной характеристики. Для этого надо продифференцировать уравнение (3.80) по температуре и приравнять полученное выражение нулю:

$$\frac{dU_{обр}}{dT} = \frac{1}{R_{T\text{ пс}} I_{\infty}} \left[\exp \frac{\Delta \mathcal{E}}{kT} + (T - T_{\text{окр}}) \left(\exp \frac{\Delta \mathcal{E}}{kT} \right) \frac{\Delta \mathcal{E}}{k} \left(-\frac{1}{T^2} \right) \right] = 0$$

или

$$T^2 - \frac{\Delta \mathcal{E}}{k} T + \frac{\Delta \mathcal{E}}{k} T_{\text{окр}} = 0.$$

Отсюда температура $p-n$ -перехода диода, соответствующая экстремальным напряжениям на диоде,

$$T = \frac{\Delta \mathcal{E}}{2k} \pm \sqrt{\left(\frac{\Delta \mathcal{E}}{2k} \right)^2 - \frac{\Delta \mathcal{E}}{k} T_{\text{окр}}}. \quad (3.81)$$

Диод будет обладать отрицательным дифференциальным сопротивлением при обратном включении, а обратная ветвь вольт-амперной характеристики будет иметь два экстремальных напряжения

при условии, если корни последнего уравнения будут действительными и разными, т. е. если

$$\left(\frac{\Delta\theta}{2k}\right)^2 > \frac{\Delta\theta}{k} T_{\text{окр}}$$

или

$$\Delta\theta > 4kT_{\text{окр}}. \quad (3.82)$$

Последнее неравенство справедливо для всех полупроводников, применяемых для изготовления полупроводниковых диодов.

Расчет пробивного напряжения при тепловом пробое

Чтобы оценить влияние разогрева на обратный ток диода, введем понятие температурного коэффициента обратного тока (по аналогии с другими температурными коэффициентами — ТК R , ТК ϵ и т. д.):

$$\alpha = \text{ТК } I = \frac{1}{I} \cdot \frac{dI}{dT}. \quad (3.83)$$

Подставим в (3.83) аппроксимацию обратного тока (3.79). Тогда

$$\alpha = \frac{1}{I_{\infty} \exp\left(-\frac{\Delta\theta}{kT}\right)} \frac{I_{\infty} \left[\exp\left(-\frac{\Delta\theta}{kT}\right) \right] \Delta\theta}{kT^2} = \frac{\Delta\theta}{kT^2}. \quad (3.84)$$

В дифференциальном уравнении (3.83) можно разделить переменные, а потом провести интегрирование, согласовав пределы интегрирования:

$$\int_{I_{\text{окр}}}^I \frac{dI}{I} = \alpha \int_{T_{\text{окр}}}^T dT,$$

где $I_{\text{окр}}$ — обратный ток через диод при начальной температуре $T_{\text{окр}}$.

При интегрировании пренебрежем температурной зависимостью α , вынося этот параметр за знак интеграла. В таком случае окончательный результат можно будет считать справедливым в первом приближении только при небольших перегревах p - n -перехода относительно среды, окружающей диод. Тогда в результате интегрирования получим

$$\alpha (T - T_{\text{окр}}) = \ln \frac{I}{I_{\text{окр}}}$$

или

$$\exp[\alpha (T - T_{\text{окр}})] = \frac{I}{I_{\text{окр}}},$$

т. е.

$$I = I_{\text{окр}} \exp[\alpha (T - T_{\text{окр}})]. \quad (3.85)$$

Если теперь подставить выражение обратного тока через температурный коэффициент этого тока (3.85) в уравнение теплового баланса (3.78), то получим

$$U = \frac{T - T_{\text{окр}}}{R_T \text{ пс } I_{\text{окр}} \exp[\alpha(T - T_{\text{окр}})]}. \quad (3.86)$$

При тепловом пробое $dU/dT = 0$. Поэтому после дифференцирования формулы (3.86) и сокращений получим

$$1 - \alpha(T - T_{\text{окр}}) = 0.$$

Отсюда температура p - n -перехода при тепловом пробое

$$T = T_{\text{окр}} + \frac{1}{\alpha}. \quad (3.87)$$

Для полупроводниковых диодов значение температурного коэффициента обратного тока обычно около $0,1 \text{ K}^{-1}$, т. е. при тепловом пробое температура p - n -перехода превышает температуру окружающей среды всего примерно на 10 K . Именно из-за малого перегрева p - n -перехода в начале развития теплового пробоя можно считать температурный коэффициент обратного тока величиной неизменной при изменении температуры и вынести его за знак интеграла. Конечно, при развитии теплового пробоя с дальнейшим увеличением обратного тока температура p - n -перехода может значительно повыситься — вплоть до плавления полупроводникового материала.

Воспользовавшись полученным значением температуры p - n -перехода при тепловом пробое (3.87) и подставив его в (3.86), получим

$$U_{\text{проб}} = \frac{1}{\alpha R_T \text{ пс } I_{\text{окр}} e} \quad (3.88)$$

или, с учетом (3.84),

$$U_{\text{проб}} = \frac{kT^2}{\Delta \mathcal{E} R_T \text{ пс } I_{\text{окр}} e}.$$

Следовательно, пробивное напряжение при тепловом пробое диода определяется его обратным током, температурным коэффициентом обратного тока и тепловым сопротивлением. Особое внимание следует обратить на сильную зависимость напряжения теплового пробоя от температуры окружающей среды. С увеличением температуры окружающей среды пробивное напряжение при тепловом пробое в соответствии с (3.88) и (3.79) уменьшается (см. рис. 3.22). Пробивное напряжение уменьшается, во-первых, в связи с увеличением выделяющейся мощности при тех же обратных напряжениях и, во-вторых, из-за ухудшения теплоотвода от p - n -перехода.

Так как пробивное напряжение при тепловом пробое зависит от обратного тока через диод при температуре окружающей среды, то в диодах с большими обратными токами даже при комнатных

температурах создаются условия для теплового пробоя, и он наступает раньше, чем лавинный пробой. Это справедливо, в частности, для германиевых диодов. И наоборот, в кремниевых диодах из-за значительно меньших обратных токов напряжение теплового пробоя получается настолько большим, что раньше наступает лавинный пробой. Однако это не означает, что в кремниевых диодах не может быть теплового пробоя. Он может происходить при высоких температурах окружающей среды. Кроме того, пробой может начаться как лавинный, а затем, по мере увеличения обратного тока, перейти в тепловой.

В связи с тем, что пробивное напряжение при тепловом пробое уменьшается с увеличением теплового сопротивления, на совершенство конструкции диода с точки зрения уменьшения его теплового сопротивления следует обратить особое внимание. Необходимо также отметить, что тепловое сопротивление может увеличиться из-за неправильной установки диода, когда он оказывается теплоизолированным. Напряжение теплового пробоя при этом может существенно уменьшиться. То же самое может произойти при изменении условий в окружающей среде (например, при повышении давления воздуха в связи с подъемом на большую высоту).

Особенности теплового пробоя в реальных диодах

Тепловой пробой в реальных диодах всегда происходит с образованием так называемого «шнура» или канала высокой проводимости, температура в котором превышает среднюю температуру остальной части $p-n$ -перехода. В свою очередь, образование шнура может быть вызвано либо дефектами в реальном $p-n$ -переходе, либо статистической флуктуацией плотности обратного тока по площади $p-n$ -перехода. Действительно, если в каком-нибудь месте $p-n$ -перехода в некоторый момент времени плотность обратного тока оказалась несколько выше плотности обратного тока в остальной части $p-n$ -перехода, то температура этого места $p-n$ -перехода будет выше из-за большей выделяющейся там удельной мощности. Увеличение температуры приведет к росту плотности обратного тока в данном месте $p-n$ -перехода в связи с увеличением тепловой генерации носителей либо в самом переходе (см. § 3.10), либо в областях полупроводника, прилегающих к $p-n$ -переходу (см. § 3.2, 3.4 и 3.5). Локальное увеличение плотности тока вызовет локальное увеличение температуры, увеличение температуры вызовет увеличение плотности тока и т. д.

Диаметр шнура, возникшего при тепловом пробое, может составлять всего несколько микрометров. Длина его определяется толщиной $p-n$ -перехода, т. е. может составлять десятки микрометров. Поэтому, учитывая малый объем шнура, следует заключить, что для развития теплового пробоя в реальных диодах при шнуровании тока требуется очень малая мощность, т. е. тепловой пробой может происходить при малых обратных токах и малых обратных

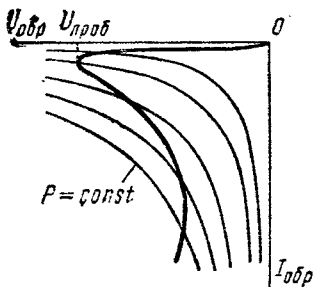


Рис. 3.23. Вольт-амперная характеристика при тепловом пробое диода и гиперболы равной мощности

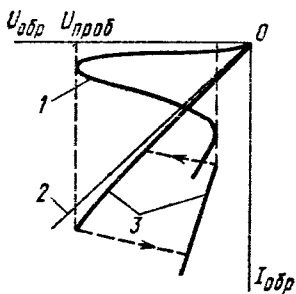


Рис. 3.24. Вольт-амперные характеристики шнура (1) и шунтирующего сопротивления остальной части $p-n$ -перехода (2), которые в сумме могут дать у-образную вольт-амперную характеристику диода (3) при тепловом пробое

напряжениях. Удельная мощность, выделяемая в единичном объеме шнура, даже при очень малых обратных токах через диод оказывается довольно большой.

Для возникновения теплового пробоя в соответствии с (3.87) необходим перегрев шнура примерно на 10 К. Это свидетельствует, во-первых, о том, что опять необходима малая мощность для развития теплового пробоя, и, во-вторых, о том, что тепловой пробой — это явление обратимое, если, конечно, ограничить обратный ток при пробое, не доводя разогретый шнур до появления необратимых физико-химических процессов в полупроводнике.

Следствием малого объема шнура, по которому происходит тепловой пробой, является также малая инерционность процесса теплового пробоя реальных диодов. Тепловые постоянные времени нагрева и охлаждения шнура могут быть порядка 10^{-6} — 10^{-7} с. При этом еще надо учесть, что при тепловом пробое напряжение на диоде уменьшается и барьерная емкость $p-n$ -перехода диода разряжается через сопротивление шнура с выделением добавочной мощности в шнур. Это явление способствует ускорению разогрева шнура и уменьшению инерционности процесса теплового пробоя.

Еще одним следствием шнурования тока при тепловом пробое диода является возможность получения своеобразной вольт-амперной характеристики — так называемой у-образной вольт-амперной характеристики, которая на первый взгляд противоречит тепловому механизму пробоя диода. Действительно, с увеличением тока через диод температура $p-n$ -перехода все время должна возрастать, что видно как из уравнения теплового баланса (3.78), так и из соотношения (3.79). Таким образом, кривые (гиперболы) равной мощности должны пересекать вольт-амперную характеристику диода с учетом тепловыделения в $p-n$ -переходе только в одной точке (рис. 3.23). Однако тепловой пробой происходит по шнуру с очень малым поперечным сечением. Если бы удалось из всего $p-n$ -перехода выделить шнур, то его вольт-амперная характеристика соответствовала бы всем условиям теплового пробоя (кривая 1 на рис. 3.24). Через остальную часть $p-n$ -перехода, площадь кото-

рой на несколько порядков больше поперечного сечения шнура, идет обратный ток, т. е. ее можно характеризовать некоторым сопротивлением R . Для простоты будем считать сопротивление R постоянным — с линейной вольт-амперной характеристикой (кривая 2). Если сопротивление R окажется меньше абсолютного значения отрицательного дифференциального сопротивления шнура на падающем участке его вольт-амперной характеристики, то суммарная вольт-амперная характеристика диода (кривая 3) будет иметь у-образный вид. Следовательно, у-образные вольт-амперные характеристики не противоречат тепловому механизму пробоя диода.

Как было отмечено во второй части этого параграфа, пробивное напряжение при тепловом пробое уменьшается с повышением температуры окружающей среды. Однако тепловому пробую может предшествовать лавинный пробой, для которого характерен положительный температурный коэффициент пробивного напряжения. Поэтому температурная зависимость пробивного напряжения для диода при наличии теплового и лавинного пробоя может быть сложной и даже немонотонной, так как при больших температурах тепловой пробой может наступать без предшествующего лавинного пробоя.

§ 3.14. ВЛИЯНИЕ ПОВЕРХНОСТНЫХ СОСТОЯНИЙ НА ВОЛЬТ-АМПЕРНУЮ ХАРАКТЕРИСТИКУ ДИОДА

В реальном полупроводниковом диоде выпрямляющий электрический переход обязательно выходит на поверхность полупроводника. В связи с этим состояние поверхности влияет на вольт-амперную характеристику диода. Это влияние значительно сильнее сказывается на обратной ветви вольт-амперной характеристики, так как обратные токи обычно малы.

Характер влияния поверхностных состояний зависит от знака и величины поверхностного заряда. Рассмотрим три возможных варианта влияния поверхностных состояний.

Генерация носителей заряда на поверхности полупроводника

Генерация и рекомбинация носителей заряда на поверхности полупроводника, как правило, проходит более интенсивно, чем в его объеме (см. § 1.2). На обратные токи полупроводникового диода генерация носителей на поверхности оказывает такое же влияние, как и генерация носителей в объеме. Однако

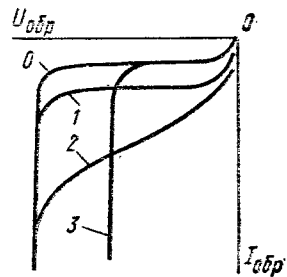


Рис. 3.25. Влияние поверхностных состояний на обратную ветвь вольт-амперной характеристики диода:

0 — без учета поверхностных состояний; 1 — с учетом генерации носителей заряда на поверхности; 2 — при наличии канала поверхностной электропроводности; 3 — при обогащенном слое на поверхности базы

обратные токи диода зависят от скорости поверхностной генерации, а скорость поверхностной генерации может изменяться со временем из-за изменения поверхностного заряда.

В любом случае обратные токи диода с учетом генерации носителей на поверхности должны быть больше по сравнению с токами без генерации носителей на поверхности (кривая 1 на рис. 3.25).

Каналы поверхностной электропроводности

Если на поверхности полупроводника существует большой поверхностный заряд, совпадающий по знаку с зарядом основных носителей в базе диода, то такой поверхностный заряд отталкивает основные носители от поверхности базы и притягивает к поверхности неосновные носители, что приводит к образованию инверсного слоя по всей поверхности базы (рис. 3.26, а). При возникновении инверсного слоя увеличивается площадь электронно-дырочного перехода. Теперь экстракция неосновных носителей из базы будет происходить не только из слоя толщины L_p , примыкающего к электронно-дырочному переходу в объеме, но и из такого же слоя, примыкающего к поверхности базы (рис. 3.27, а). Таким образом, инверсный слой тоже может принимать участие в сборе неосновных носителей из базы диода.

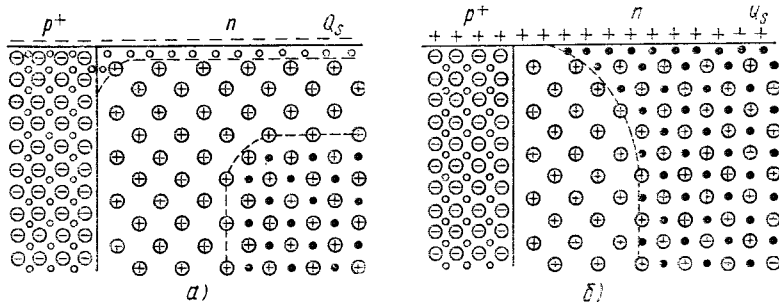


Рис. 3.26. Искажения границ p - n -перехода под влиянием поверхностного заряда:

а — при образовании инверсного слоя на поверхности базы диода; б — при образовании обогащенного слоя на поверхности базы диода

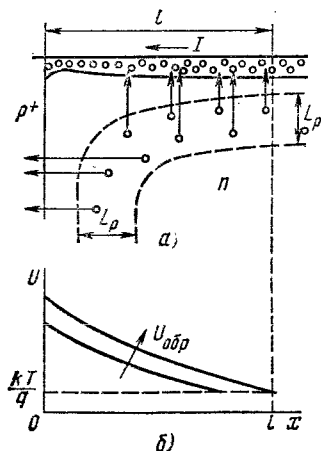


Рис. 3.27. Пути движения неосновных носителей при наличии канала поверхностной электропроводности (а) и распределение напряжения вдоль канала (б)

В связи с тем, что вдоль инверсного слоя течет ток, напряжение между инверсным слоем и объемом базы не остается постоянным — по мере удаления от металлургической границы электронно-дырочного перехода оно уменьшается (рис. 3.27, б). При этом надо учитывать большое удельное сопротивление инверсного слоя из-за малой его толщины. Тогда на некотором расстоянии l от металлургической границы электронно-дырочного перехода это напряжение примерно равно kT/q . Следовательно, не весь инверсный слой участвует в сборе неосновных носителей, а только его часть протяженностью l . Та часть инверсного слоя, которая принимает участие в экстракции неосновных носителей, называется каналом поверхностной электропроводности.

При увеличении обратного напряжения на диоде будет увеличиваться и длина канала поверхностной электропроводности (рис. 3.27, б). Таким образом, обратная ветвь вольт-амперной характеристики диода при наличии канала поверхностной электропроводности в базовой области не будет иметь участка насыщения (см. рис. 3.25, кривая 2).

Поверхностный пробой

Поверхностным пробоем электронно-дырочного перехода называют пробой перехода, который происходит в месте выхода перехода на поверхность кристалла и на пробивное напряжение которого оказывают влияние поверхностные состояния.

Если поверхностный заряд — заряд поверхностных состояний имеет знак, противоположный знаку основных носителей в базе диода, то на поверхности базы образуется обогащенный слой (см. § 1.12). Из-за возникновения обогащенного слоя толщина электронно-дырочного перехода у поверхности базы уменьшается, так как диффузионное электрическое поле перехода проникает в обогащенный слой на меньшую глубину (см. рис. 3.26, б). Влияние поверхностного заряда сказывается в основном на области базы, потому что ее удельное сопротивление велико.

В связи с меньшей толщиной электронно-дырочного перехода у поверхности пробой диода будет происходить именно там, а пробивное напряжение окажется тем меньше, чем будет большее сужение перехода у поверхности (см. рис. 3.25, кривая 3). Таким образом, величина пробивного напряжения зависит в данном случае от плотности поверхностных состояний или от величины поверхностного заряда, создающего обогащенный слой у поверхности базы (рис. 3.28). По своей природе поверхностный пробой может быть лавинным, туннельным или тепловым.

§ 3.15. ПРОЦЕССЫ В ДИОДАХ ПРИ БОЛЬШИХ ПРЯМЫХ ТОКАХ

Прежде чем рассматривать явления в полупроводниковых диодах при больших прямых токах, установим понятия об уровне инжекции. Под уровнем инжекции будем понимать соотношение

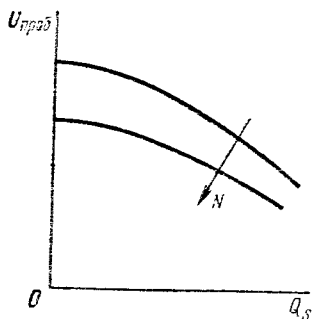


Рис. 3.28. Зависимости пробивного напряжения диода от величины поверхностного заряда, создающего обогащенный слой у поверхности базы, при разных концентрациях примесей в базе

разница между концентрацией инжектированных неосновных носителей и концентрацией основных носителей в состоянии равновесия мала, т. е.

$$\Delta p_n \approx p_n \approx n_{n0}.$$

Высоким уровнем инжекции будем считать такой, при котором

$$\Delta p_n \approx p_n \gg n_{n0}.$$

Прохождение через полупроводниковый диод больших токов, как правило, сопровождается повышением уровня инжекции, по крайней мере, в некоторых частях базы диода.

При больших прямых токах в базе полупроводникового диода изменяется концентрация как неосновных, так и основных носителей заряда, появляется электрическое поле, изменяются электрофизические параметры материала базы (время жизни, подвижность или коэффициент диффузии, удельное сопротивление), изменяются условия как на выпрямляющем, так и на невыпрямляющем контактах. При больших токах необходимо также учитывать сопротивление базы диода.

Изменение концентрации основных носителей

При инжекции неосновных носителей заряда в базе диода возникает избыточная концентрация этих носителей и нарушается электрическая нейтральность базы. Из-за диффузии в базе и рекомбинации получается неравномерное распределение неосновных носителей в базе (рис. 3.29). Для компенсации инжектированных неосновных носителей заряда через невыпрямляющий контакт в базу входят основные носители и также распределяются в базе неравномерно.

между концентрацией неосновных носителей и концентрацией основных носителей в состоянии равновесия (или, что почти то же самое, концентрацией примесей). Низким уровнем инжекции будем считать такой, при котором концентрация инжектированных неосновных носителей значительно меньше концентрации основных носителей в состоянии равновесия, т. е. для полупроводника *n*-типа:

$$p_n - p_{n0} = \Delta p_n \ll n_{n0},$$

для полупроводника *p*-типа:

$$n_p - n_{p0} = \Delta n_p \ll p_{p0}.$$

Средним уровнем инжекции будем считать такой, при котором

С достаточной для практических расчетов точностью можно считать, что при этом восстанавливается электрическая нейтральность базы, т. е. избыточная концентрация основных носителей в базе диода равна избыточной концентрации неосновных носителей

$$p_n - p_{n0} \approx n_n - n_{n0} \quad (3.89)$$

или

$$p_n \approx n_n - N_A.$$

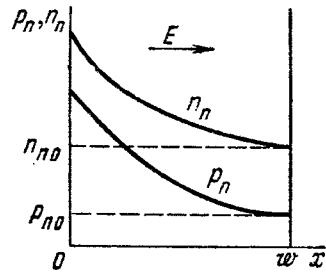


Рис. 3.29. Распределение концентрации неосновных и основных носителей в базе диода при прямом его включении

Появление электрического поля в базе

Равенство (3.89) является приближительным, так как для поддержания неравномерного распределения основных носителей в базе необходимо электрическое поле, которое и возникает из-за неполной компенсации заряда неосновных носителей вблизи электронно-дырочного перехода.

В принципе нейтрализация заряда инжектированных носителей не может быть абсолютно полной. Действительно, если предположить полную нейтрализацию, то, во-первых, тогда не будет причины, вызывающей неравномерное распределение основных носителей. Во-вторых, неравномерное распределение основных носителей без электрического поля приведет к возникновению диффузионного тока электронов в направлении, противоположном истинному направлению движения электронов в базе диода при его прямом включении. Таким образом, предположение об абсолютной нейтрализации заряда инжектированных носителей ведет к противоречиям.

Остается принять, что нейтрализация заряда оказывается неполной — избыточная концентрация основных носителей около электронно-дырочного перехода остается несколько меньшей избыточной концентрации неосновных носителей. Возникающая из-за нескомпенсированного заряда напряженность электрического поля поддерживает неравномерное распределение основных носителей в базе.

Другая составляющая электрического поля в базе диода является следствием прохождения тока через объемное сопротивление базы.

Чтобы определить суммарное значение электрического поля в базе, воспользуемся выражениями для плотности электронного (1.30) и дырочного (1.31) токов, а также выражением для полной плотности тока (1.32) с учетом того, что в установившемся состоянии $\frac{\partial(\epsilon\epsilon_0 E)}{\partial t} = 0$.

Из условия электрической нейтральности (3.89) для днода с равномерно легированной базой, т. е. при $N_d(x) = \text{const}$, следует, что

$$\text{grad } n = \text{grad } p.$$

Тогда

$$\mathbf{J} = q(p\mu_p + n\mu_n) \mathbf{E} + q(D_n - D_p) \text{grad } p$$

и

$$\mathbf{E} = \frac{\mathbf{J}}{q(p\mu_p + n\mu_n)} - \frac{D_n - D_p}{p\mu_p + n\mu_n} \text{grad } p. \quad (3.90)$$

Полученный результат показывает, что напряженность электрического поля действительно состоит из двух составляющих. Первая из них равна отношению плотности полного тока к удельной проводимости базы, т. е. представляет собой в соответствии с законом Ома в дифференциальной форме напряженность электрического поля, возникающую из-за прохождения тока через материал базы с конечной проводимостью. Следует отметить, что удельная проводимость в данном случае не является постоянной, она зависит от концентрации носителей n , значит, от плотности тока через днод.

Вторая составляющая напряженности электрического поля определяется градиентом концентрации носителей заряда. Эта составляющая и есть напряженность электрического поля, которое поддерживает неравномерное распределение концентрации основных носителей для выполнения почти полной электрической нейтральности базы.

Если проинтегрировать (3.90) по координатам, то получим разность потенциалов, которая тоже будет иметь две составляющие. Первая из них — это падение напряжения на объемном сопротивлении базы. Вторая составляющая представляет собой так называемую э. д. с. Дембера, которая была рассмотрена в § 1.11. Эта э. д. с. может возникнуть в полупроводнике при неравномерном распределении носителей даже при отсутствии тока через полупроводник, как видно из соотношения (3.90). Так, например, при освещении полупроводника между освещенными и неосвещенными его частями возникает разность потенциалов. Аналогичную природу имеет посленижекционная э. д. с.

Таким образом, при больших токах в базе днода помимо диффузии носителей заряда необходимо учитывать их дрейф. Кроме того, из-за наличия в базе при больших токах через днод электрического поля напряжение на электроно-дырочном переходе отличается от напряжения, приложенного к выводам днода.

Изменение электрических параметров

Время жизни носителей заряда. При высоком уровне инжекции изменяется ход процесса рекомбинации. Это справедливо для любой модели процесса рекомбинации. Так, при увеличении концентрации неосновных и основных носителей изменяются положения

демаркационных уровней, и рекомбинация может происходить через такие ловушки, через которые она не происходила при низком уровне инжекции.

Вообще зависимость времени жизни от уровня инжекции определяется многими факторами — типом и расположением энергетических уровней ловушек в запрещенной зоне полупроводника, температурой и т. д. С увеличением уровня инжекции время жизни может уменьшаться, увеличиваться, может наблюдаться и немонотонная зависимость.

Подвижность (коэффициент диффузии). При высоком уровне инжекции может проявляться еще один механизм рассеяния носителей заряда — рассеяние носителей на носителях, что приводит к уменьшению подвижности и коэффициента диффузии носителей заряда. Однако это явление часто можно не принимать во внимание, так как оно начинает сказываться при сравнительно больших концентрациях инжектированных носителей.

Изменение условий на контактах (на выпрямляющем переходе и невыпрямляющих контактах). Граничная концентрация неосновных носителей около электронно-дырочного перехода при больших прямых напряжениях перестает экспоненциально зависеть от напряжения на переходе (см. § 2.3). Можно считать, что при очень больших прямых напряжениях электронно-дырочный переход полупроводникового диода как бы исчезает и диод становится похожим на резистор с переменным сопротивлением.

В то же время при больших токах нарушается равновесие у невыпрямляющего контакта. Здесь из-за ограниченной скорости движения начинает расти концентрация неосновных носителей, соответственно растет и концентрация основных носителей (см. § 2.12). База диода при этом заполняется дополнительным количеством электронов и дырок, что приводит к уменьшению ее сопротивления.

§ 3.16. РАСЧЕТ ВОЛЬТ-АМПЕРНОЙ ХАРАКТЕРИСТИКИ ДИОДА ПРИ БОЛЬШИХ ПРЯМЫХ ТОКАХ

Выбор модели диода

Большое количество факторов, которые следует принимать во внимание при расчете вольт-амперной характеристики диода при больших прямых токах, приводит к тому, что в общем виде решить эту задачу не удастся.

В связи с этим для выполнения аналитического расчета целесообразно выбрать такую структуру модели полупроводникового

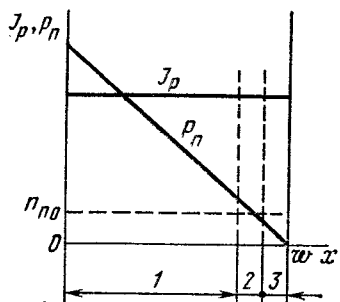


Рис. 3.30. Распределение концентрации неосновных носителей в базе диода с тонкой базой:

1 — слой с высоким уровнем инжекции; 2 — слой со средним уровнем инжекции; 3 — слой с низким уровнем инжекции

диода, которая позволила бы выполнить расчет и в то же время по возможности больше соответствовала бы реальной структуре диода.

Одной из таких структур является полупроводниковый диод с несимметричным электронно-дырочным p^+-n -переходом, с тонкой базой и с идеальным невыпрямляющим контактом к базе. В базе такого диода можно выделить слои с высоким уровнем инжекции — у электронно-дырочного перехода, средним — там, где концентрация неосновных носителей сравнима с концентрацией основных носителей в равновесном состоянии, и с низким уровнем инжекции вблизи невыпрямляющего контакта (рис. 3.30).

Очевидно, что с увеличением тока через диод будет увеличиваться та часть базы, где существует высокий уровень инжекции. При достаточно больших токах почти во всей базе будет высокий уровень инжекции. Поэтому весь расчет можно выполнять только для высокого уровня инжекции. Это первое преимущество выбранной модели для расчета.

Второе преимущество диода с тонкой базой как модели для расчета состоит в том, что ток одного вида носителей во всей базе такого диода остается постоянным. Это следует из того, что во всей тонкой базе практически отсутствует рекомбинация носителей — она происходит только на невыпрямляющем контакте.

Третье преимущество рассматриваемой модели следует из несимметрии ее перехода. Если концентрация инжектированных неосновных носителей у границы электронно-дырочного перехода еще не равна полной концентрации основных носителей, что может быть только при очень больших прямых токах, то через несимметричный переход происходит инжекция только одного вида носителей — из сильнолегированной области в слаболегированную. В этом случае весь ток связан с движением носителей одного вида — дырок. Значит, можно рассматривать движение только одного вида носителей во всей базе диода.

Четвертое преимущество рассматриваемой модели заключается в том, что она соответствует реальной структуре большинства полупроводниковых диодов.

Расчет вольт-амперной характеристики диода с тонкой базой при больших прямых токах

Исходя из перечисленных допущений можно считать, что в базе рассматриваемого диода электронный ток равен нулю. Тогда напряженность электрического поля в базе диода может быть найдена из уравнения для плотности электронного тока [см. (1.30)]:

$$J_n = 0 = qn\mu_n E + qD_n \frac{dn_n}{dx}.$$

С учетом соотношения Эйнштейна (1.29)

$$E = -\frac{kT}{q} \cdot \frac{1}{n_n} \cdot \frac{dn_n}{dx}. \quad (3.91)$$

Используя условие электрической нейтральности базы (3.89), получим

$$E = -\frac{kT}{q} \cdot \frac{1}{p_n + N_{дн}} \cdot \frac{dp_n}{dx}. \quad (3.92)$$

Соотношение (3.92) представляет собой полное значение напряженности электрического поля, включающее поле, связанное с падением напряжения на объемном сопротивлении базы, и поле, обусловленное э. д. с. Дембера.

Тогда, подставив (3.92) в (1.31), получим дырочную составляющую плотности тока в базе диода, а следовательно, и ток через диод:

$$J_p = -qD_p \frac{p_n}{p_n + N_{дн}} \cdot \frac{dp_n}{dx} - qD_p \frac{dp_n}{dx}$$

или

$$J_p = -qD_p \left(1 + \frac{p_n}{p_n + N_{дн}} \right) \frac{dp_n}{dx}. \quad (3.93)$$

Выражение (3.93) напоминает формулу для расчета диффузионной составляющей плотности тока (1.27), так как плотность тока получилась пропорциональной градиенту концентрации носителей. Однако такое сходство формально, так как на самом деле в расчете была принята во внимание и напряженность электрического поля в базе диода.

Особенно простой вид принимает выражение для плотности дырочного тока через диод при высоком уровне инжекции, т. е. когда $p_n \gg N_{дн}$:

$$J_p = -q2D_p \frac{dp_n}{dx}. \quad (3.94)$$

Таким образом, дырочный ток через диод прямо пропорционален градиенту концентрации носителей заряда в базе, причем коэффициент пропорциональности не зависит от концентрации носителей. В выражении (3.94) коэффициент диффузии удвоен, что отражает влияние электрического поля в базе диода.

Формальное сходство выражения для плотности тока при высоком уровне инжекции (3.94) с выражением для чисто диффузионного тока (1.27) дает возможность упростить расчет вольт-амперной характеристики диода при высоком уровне инжекции, так как все отличие от рассмотренного ранее расчета при низком уровне инжекции заключается только в удвоенном значении коэффициента диффузии, т. е.

$$J = J_p = \frac{q2D_p p_{n0}}{W_n} \left(\exp \frac{qU_p}{kT} - 1 \right). \quad (3.95)$$

В уравнении для вольт-амперной характеристики диода при низком уровне инжекции [см. (3.27)] напряжение U имело смысл напряжения на электронно-дырочном переходе. Оно же считалось равным напряжению на диоде, так как падением напряжения на базе диода при низком уровне инжекции можно было пренебречь. При высоком уровне инжекции следует учесть разницу в падениях

напряжения на переходе и на диоде, пометив индексом $p-n$ напряжение в формуле (3.95). Теперь понятен физический смысл полученной формулы, но необходимо еще найти распределение падений напряжения на диоде.

Расчет падения напряжения на базе в диоде с тонкой базой

Полное падение напряжения на диоде складывается из падения напряжения на электронно-дырочном переходе и на базе диода. Падение напряжения на базе диода может быть найдено с использованием соотношения для напряженности электрического поля в базе (3.91) и с преобразованием интеграла по координате в интеграл по концентрации носителей:

$$\begin{aligned}
 U_B &= \int_0^{U_B} dU_B = - \int_{W_n}^0 E dx = - \frac{kT}{q} \int_{W_n}^0 \frac{1}{n_n} \frac{dn_n}{dx} dx = \\
 &= \frac{kT}{q} \int_{n_{n0}}^{n_{n\Gamma p}} \frac{dn_n}{n_n} = \frac{kT}{q} \ln \frac{n_{n\Gamma p}}{n_{n0}}.
 \end{aligned} \tag{3.96}$$

Нижний предел у последнего интеграла соответствует концентрации основных носителей в базе в состоянии равновесия. Это можно считать справедливым даже при токах, близких к предельно допустимым для полупроводниковых диодов, при хорошем невыпрямляющем контакте с базой. В этом случае концентрация основных носителей у невыпрямляющего контакта изменяется с изменением тока относительно мало, или, другими словами, мало падение напряжения на невыпрямляющем контакте.

Концентрацию основных носителей около электронно-дырочного перехода можно определить из условия электрической нейтральности ($p_{n\Gamma p} \approx n_{n\Gamma p}$) и выражения для граничной концентрации неосновных носителей (2.5), т. е. при $x = 0$

$$n_{n\Gamma p} = p_{n\Gamma p} = p_{n0} \exp \frac{qU_{p-n}}{kT}. \tag{3.97}$$

После подстановки (3.97) в (3.96)

$$U_B = \frac{kT}{q} \ln \left(\frac{p_{n0}}{n_{n0}} \exp \frac{qU_{p-n}}{kT} \right)$$

или с учетом (1.19)

$$U_B = \frac{kT}{q} \ln \left(\frac{n_i}{n_{nc}} \right)^2 + U_{p-n}.$$

Тогда полное падение напряжения на диоде

$$U = U_B + U_{p-n} = 2 \frac{kT}{q} \ln \frac{n_i}{n_{nc}} + 2U_{p-n},$$

отсюда

$$U_{p-n} = \frac{U}{2} - \frac{kT}{q} \ln \frac{n_i}{n_{i0}} = \frac{U}{2} + \frac{kT}{q} \ln \frac{n_{n0}}{n_i}. \quad (3.98)$$

Подставив последнее выражение (3.98) в уравнение вольт-амперной характеристики (3.95), получим

$$J = \frac{q2D_p p_{n0}}{W_n} \left[\left(\exp \frac{qU}{2kT} \right) \frac{n_{n0}}{n_i} - 1 \right].$$

Единицей в квадратных скобках можно пренебречь, так как $n_{n0} \gg n_i$ и $\exp \frac{qU}{2kT} \gg 1$. Тогда

$$J = \frac{q2D_n n_i}{W} \exp \frac{qU}{2kT}. \quad (3.99)$$

Здесь, как и ранее, опущены индексы у параметров материала базы диода — все входящие в формулу величины относятся к слаболегированной области, т. е. к базе диода.

Полученное выражение напоминает уравнение вольт-амперной характеристики диода с тонкой базой при низком уровне инжекции (3.27) с учетом (3.31), но есть и некоторые отличия:

1) вместо обычного коэффициента диффузии стоит его удвоенное значение, что отражает влияние электрического поля в базе;

2) в показателе экспоненты стоит уменьшенное вдвое напряжение, что связано с распределением приложенного к диоду напряжения приблизительно пополам между переходом и базой диода;

3) вместо концентрации неосновных носителей в последней формуле стоит собственная концентрация. Значит вид вольт-амперной характеристики при высоком уровне инжекции перестает зависеть от концентрации примесей в слаболегированной области (в базе диода). Объясняется это тем, что при высоком уровне инжекции свойства полупроводника базы определяются в основном концентрацией инжектированных носителей. Полупроводник базы ведет себя при этом аналогично собственному полупроводнику.

Диапазон применимости полученного соотношения для вольт-амперной характеристики диода при больших токах

Диапазон применимости соотношения (3.99) ограничен снизу тем, что при малых токах через диод не выполняется условие высокого уровня инжекции; сверху — тем, что при очень больших токах становятся несправедливыми принятые условия на границах электронно-дырочного перехода и на невыпрямляющем контакте с базой.

Нижняя граница диапазона применимости. Строго нижнюю границу диапазона применимости соотношения (3.99) следовало бы определить по концентрации носителей заряда в соответствии с определением уровня инжекции. Вольт-амперная характеристика, рас-

считанная по приближенным формулам для низкого и высокого уровней инжекции, должна быть непрерывной. Тогда в качестве предельного тока, являющегося границей диапазонов применимости приближений, можно взять такой, при котором вычисленные токи в том и другом приближении окажутся одинаковыми, т. е. соотношения (3.27) с учетом (3.31) и (3.99) имеют вид

$$\frac{qD_p p_{n0}}{W_n} \left(\exp \frac{qU}{kT} - 1 \right) = \frac{q2D_p n_i}{W_n} \exp \frac{qU}{2kT}. \quad (3.100)$$

Если пренебречь единицей в формуле тока при низком уровне инжекции, то

$$\exp \frac{qU}{2kT} = \frac{2n_i}{p_{n0}}.$$

Подставив этот результат в любую часть выражения (3.100), получим для граничной плотности тока

$$J_n = \frac{q2D_p n_i 2n_i}{W_n p_{n0}}$$

или после преобразований с помощью соотношения (1.19)

$$J_n = \frac{4qDN}{W}. \quad (3.101)$$

Здесь опять опущены индексы, которые должны соответствовать слаболегированной области диода, т. е. его базе.

Оценка граничной плотности тока с помощью соотношения (3.101) показывает, что $J_n \approx 0,1$ А/см², что примерно на три порядка меньше, чем предельно допустимая плотность тока диода. Следовательно, практически на всей прямой ветви вольт-амперной характеристики полупроводниковый диод работает в условиях больших токов.

Верхняя граница диапазона применимости. Верхняя граница диапазона применимости соотношения (3.99) определяется неравенством $U_{p-n} < \varphi_{кон}$.

Подставив в это неравенство выражения (3.98) и (2.2), получим

$$\frac{U}{2} - \frac{kT}{q} \ln \frac{n_i}{n_{n0}} < \frac{kT}{q} \ln \frac{p_{p0} n_{n0}}{n_i^2}$$

или

$$\frac{U}{2} < \frac{kT}{q} \ln \frac{p_{p0}}{n_i}.$$

Следовательно,

$$J < \frac{2qD_p n_i}{W_n} \cdot \frac{p_{p0}}{n_i}$$

или

$$J_B = \frac{2qDN^+}{W}, \quad (3.102)$$

где N^+ — концентрация примесей в сильнолегированной области, а остальные величины следует брать для слаболегированной области диода.

Порядок плотности тока, соответствующий верхней границе диапазона применимости, вычисленной по соотношению (3.102), около 100 А/см^2 , что совпадает с предельно допустимой плотностью тока полупроводниковых диодов.

Вольт-амперная характеристика диода с толстой базой при больших токах

В диодах с толстой базой, когда толщина базы во много раз превышает диффузионную длину неосновных носителей (например, в диодах на основе материала с большой шириной запрещенной зоны), сопротивление базы можно считать не зависящим от тока, т. е. можно пренебречь модуляцией сопротивления базы. Тогда, при условии экспоненциальной зависимости граничной концентрации неосновных носителей около перехода от напряжения на нем, уравнение вольт-амперной характеристики диода с толстой базой может быть представлено в виде

$$I = I_{\text{нас}} \left[\exp \frac{q(U - IR_B)}{kT} - 1 \right],$$

где R_B — сопротивление базы диода.

Если сопротивление базы известно, то эта формула позволяет рассчитать вольт-амперную характеристику, принимая в качестве независимой переменной значение тока.

При больших токах, когда напряжение на электронно-дырочном переходе приближается к контактной разности потенциалов, это напряжение как бы стабилизируется. Тогда вольт-амперная характеристика диода с толстой базой принимает вид

$$U = \varphi_{\text{кон}} + IR_B.$$

Таким образом, при больших токах на вольт-амперной характеристике диода с толстой базой может наблюдаться прямолинейный участок (рис. 3.31). Это дает возможность оценить контактную разность потенциалов на электронно-дырочном переходе диода.

Однако в большинстве случаев сопротивление базы не является постоянным при изменении тока, и расчет вольт-амперной характеристики сильно усложняется. Но заранее можно сказать, что при напряжениях на переходе, близких к контактной разности потенциалов, вольт-амперная характеристика вернее всего не будет экспоненциальной, так как экспоненциальный ее вид прежде всего связан с экспо-

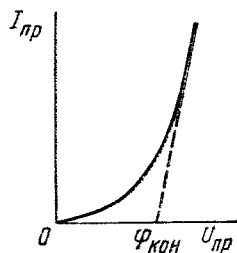


Рис. 3.31. Прямая ветвь вольт-амперной характеристики диода с толстой базой

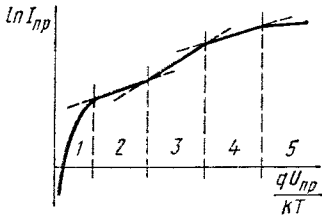


Рис. 3.32. Прямая ветвь вольт-амперной характеристики диода:

1 — участок малых токов; 2 — участок с преобладанием рекомбинации в $p-n$ -переходе; 3 — участок с преобладанием процесса инжекции; 4 — участок, связанный с модуляцией сопротивления базы при высоком уровне инжекции; 5 — участок, связанный с нарушением условий на границах $p-n$ -перехода и на невыпрямляющем контакте с базой

ненциальной зависимостью концентрации инжектированных носителей от напряжения. А эта зависимость при больших прямых токах нарушается.

§ 3.17. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ДИОДА В ПОЛУЛОГАРИФИЧЕСКИХ КООРДИНАТАХ

На прямой ветви вольт-амперной характеристики диода, построенной в полулогарифмических координатах, могут существовать участки, связанные с различными механизмами образования тока. Там, где вольт-амперная характеристика экспоненциальна, в данной системе координат получается прямолинейный отрезок (рис. 3.32).

При напряжении на диоде, равном нулю, ток равен нулю. Следовательно, $\lg I \rightarrow -\infty$.

При малых напряжениях на диоде ток может определяться рекомбинацией в электронно-дырочном переходе. Тогда он зависит от напряжения как $\exp \frac{qU}{2kT}$ [см. (3.58)], что дает на графике отрезок прямой линии с угловым коэффициентом $1/2$.

При больших напряжениях преобладает ток, связанный с инжекцией носителей заряда и зависящий от напряжения как $\exp \frac{qU}{kT}$ [см. (3.27)], что дает на графике отрезок прямой линии с угловым коэффициентом 1.

При дальнейшем увеличении тока наряду с инжекцией носителей заряда начинает сказываться модуляция сопротивления базы, что опять дает зависимость $\exp \frac{qU}{2kT}$ [см. (3.99)] или на графике — отрезок прямой с угловым коэффициентом $1/2$.

Наконец, при очень больших прямых токах вольт-амперная характеристика диода становится неэкспоненциальной из-за нарушения условий на электронно-дырочном переходе и на невыпрямляющем контакте. Соответствующий отрезок на графике не будет прямой линией.

Таким образом, построение вольт-амперной характеристики диода в полулогарифмическом масштабе может дать информацию относительно возможных механизмов прохождения тока. Надо учесть, что прямолинейные участки на этой кривой могут плавно переходить один в другой, так что определить их границы точно не представляется возможным. Иногда некоторые участки отсутствуют. Так, для некоторых кремниевых диодов участок, соответствующий рекомбинационному току, может сразу перейти в участок, соответствующий высокому уровню инжекции. Их границу обнаружить

нельзя. Искажает вид характеристик также изменение времени жизни носителей, изменение температуры, наличие каналов поверхностной электропроводности и т. п.

§ 3.18. ПЕРЕХОДНЫЕ ПРОЦЕССЫ В ДИОДАХ

Переходные процессы в полупроводниковых диодах связаны в основном с двумя явлениями, происходящими при быстром изменении напряжения на диоде или тока через диод.

Первое из них — это накопление неосновных носителей в базе диода при его прямом включении и их рассасывание при уменьшении напряжения. Так как электрическое поле в базе диода обычно невелико, то движение неосновных носителей в базе определяется законами диффузии и происходит относительно медленно. Поэтому накопление носителей в базе и их рассасывание могут влиять на свойства диодов в режиме переключения.

Второе явление, происходящее в диодах при их переключении, это заряд и разряд барьерной емкости, что также происходит не мгновенно и поэтому может влиять на свойства диодов.

При сравнительно больших плотностях прямого тока через диод существенно накопление неосновных носителей в базе диода, а заряд и разряд барьерной емкости диода является второстепенным процессом. При малых плотностях тока переходные процессы в диоде определяются зарядом и разрядом барьерной емкости диода, накопление же неосновных носителей в базе практически не сказывается.

Временные зависимости напряжения и тока, характеризующие переходные процессы в полупроводниковом диоде, зависят также от сопротивления внешней цепи, в которую включен диод. Поэтому рассмотрим четыре предельных случая переходных процессов в полупроводниковом диоде с несимметричным p^+n -переходом.

Процессы при больших напряжениях и токах

Работа диода в схеме с генератором напряжения. Рассмотрим процессы, происходящие в полупроводниковом диоде при его включении на генератор напряжения, т. е. при включении диода в цепь с малым полным сопротивлением (в том числе и с малым сопротивлением источника питания) по сравнению с сопротивлением диода.

При подаче на диод прямого напряжения ток через диод устанавливается не сразу, так как с течением времени происходит накопление в базе инжектированных через p - n -переход неосновных носителей (дырок) и уменьшение в связи с этим сопротивления базы (рис. 3.33). Этот процесс *модуляции сопротивления базы* происходит не мгновенно, так как накопление дырок в базе диода связано с относительно медленным процессом диффузии их от p - n -перехода в глубь базы.

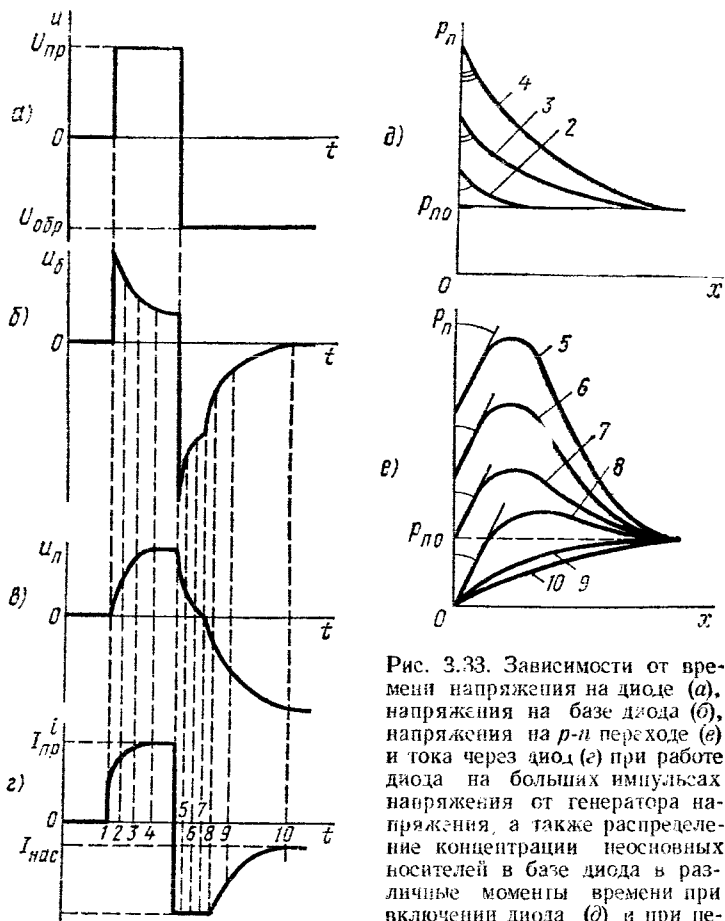


Рис. 3.33. Зависимости от времени напряжения на диоде (а), напряжения на базе диода (б), напряжения на p - n переходе (в) и тока через диод (г) при работе диода на больших импульсах напряжения от генератора напряжения, а также распределение концентрации неосновных носителей в базе диода в различные моменты времени при включении диода (д) и при переключении диода на обратное напряжение (е)

По мере накопления дырок и уменьшения сопротивления базы происходит перераспределение всего внешнего напряжения между сопротивлением базы и p - n -переходом; падение напряжения на базе диода уменьшается (рис. 3.33, б), а на p - n -переходе увеличивается (рис. 3.33, в), вызывая увеличение уровня инжекции (рис. 3.33, д). При длительном прохождении прямого тока процесс инжекции дырок уравнивается процессом их рекомбинации. Возникает некоторое установившееся состояние, характеризующееся таким распределением дырок в базе, при котором их концентрация превышает равновесную вблизи p - n -перехода и снижается, стремясь к равновесной, при удалении от него в глубь базы (кривая 4 на рис. 3.33, д).

О величине тока через p - n -переход можно судить по градиенту концентрации дырок в базе диода около p - n -перехода (см. § 3.4). Градиент концентрации дырок около p - n -перехода возрастает

с увеличением напряжения на p - n -переходе при постоянном напряжении на диоде и при уменьшении напряжения на базе диода (рис. 3.33, δ). Следует отметить, что сопротивление базы диода уменьшается не только из-за увеличения концентрации неосновных носителей (дырок), но и из-за увеличения концентрации основных носителей (электронов). Концентрация электронов около p - n -перехода возрастает в соответствии с принципом электрической нейтральности, согласно которому в любой части базовой области сумма всех зарядов должна быть равна нулю. Очевидно, что число накопленных дырок в базе тем больше, чем больше ток через диод и чем больше время жизни дырок. Кроме того, число накопленных дырок зависит от геометрии базы. При переключении диода с прямого напряжения на обратное в начальный момент наблюдается большой обратный ток, ограниченный в основном последовательным сопротивлением базы диода. Источник питания вместе с сопротивлением базы в это время можно считать генератором тока для p - n -перехода.

После переключения диода на обратное напряжение начинается процесс рассасывания неосновных носителей, накопленных в базе. Из-за ограничения обратного тока концентрация дырок в базе около p - n -перехода не может мгновенно уменьшиться до равновесного значения. Пока концентрация дырок в базе около p - n -перехода превышает равновесное значение, на p - n -переходе сохраняется прямое падение напряжения (рис. 3.33, ϵ).

После уменьшения концентрации дырок в базе около p - n -перехода до нуля величина обратного тока начинает уменьшаться со временем, о чем свидетельствует уменьшение градиента концентрации дырок около p - n -перехода (рис. 3.33, ϵ). С течением времени все накопленные в базе дырки

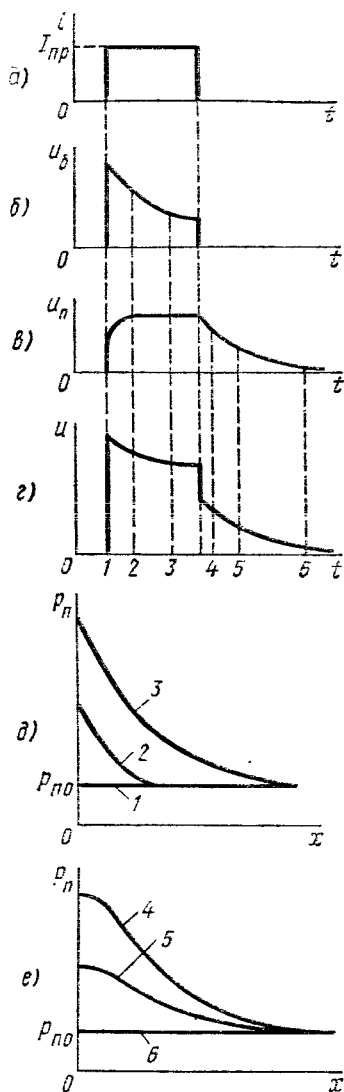


Рис. 3.34. Зависимости от времени тока через диод (a), напряжения на базе (b), напряжения на p - n -переходе (c) и напряжения на диоде (ϵ) при работе диода на больших импульсах тока в схеме с генератором тока, а также распределение концентрации неосновных носителей в базе диода в различные моменты времени при включении диода (d) и при выключении диода (e)

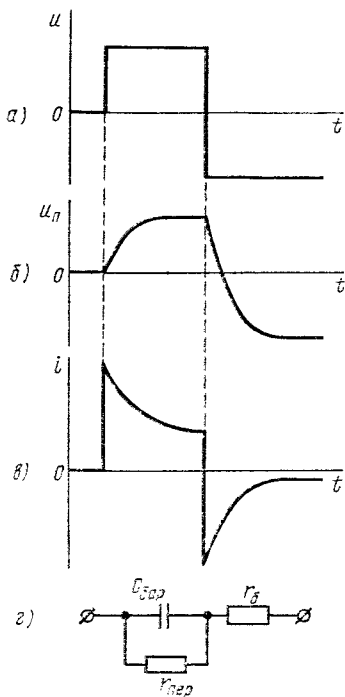


Рис. 3.35. Зависимости от времени напряжения на диоде (а), напряжения на p - n -переходе (б) и тока через диод (в) при малых импульсных напряжениях в схеме с генератором напряжения, а также эквивалентная схема диода для малых сигналов (г)

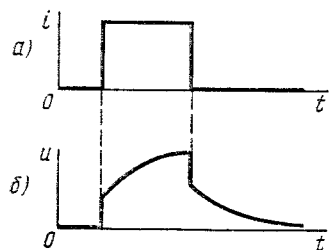


Рис. 3.36. Зависимости тока через диод (а) и напряжения на диоде (б) при работе диода на малых импульсах тока от генератора тока

уходят через p - n -переход или рекомбинируют в базе диода, в результате чего обратный ток уменьшается до стационарного значения тока насыщения (рис. 3.33, з). В это время заканчивается восстановление обратного сопротивления диода.

Процесс рассасывания накопленных носителей происходит значительно медленнее процесса их накопления, поэтому именно процесс рассасывания и определяет частотные свойства большинства диодов.

Работа диода в схеме с генератором тока соответствует включению диода в схему с большим сопротивлением, которое и определяет величину тока в цепи с диодом. Рассмотрим процессы, происходящие в диоде, при прохождении через диод импульса прямого тока прямоугольной формы (рис. 3.34).

В первый момент прохождения через диод импульса прямого тока на диоде падает относительно большое напряжение, которое в дальнейшем уменьшается из-за уменьшения сопротивления базовой области диода, связанного с накоплением неравновесных носителей в базе.

После окончания процесса накопления несбалансированных носителей в базе величины напряжений на диоде, на базе диода и на p - n -переходе достигают установившихся значений. Распределение дырок в базе в это время также соответствует некоторому установившемуся состоянию (кривая 3 на рис. 3.34, д).

В момент окончания импульса тока через диод, т. е. в момент разрыва цепи с диодом исчезает падение напряжения на объемном сопротивлении базы диода (рис. 3.34, б). Концентрация дырок в базе около p - n -перехода мгновенно измениться не может. Поэтому напряжение на p - n -переходе и соответственно на диоде после выключения тока уменьшается

замедленно по мере рекомбинации неравновесных носителей в базе. Изменения в распределении концентрации дырок в базе диода со временем показаны на рис. 3.34, *е*. Остаточное напряжение на диоде уменьшится до нуля после рекомбинации всех неравновесных носителей в базовой области диода.

Процессы при малых напряжениях и токах

Работа диода в схеме с генератором напряжения. При приложении к диоду малого прямого напряжения (рис. 3.35) эффект модуляции сопротивления базы диода из-за малого уровня инжекции пренебрежимо мал. Поэтому сопротивление диода в данном случае имеет емкостной характер. В первый момент напряжение на *p-n*-переходе близко к нулю, а ток через диод ограничен только сопротивлением базы диода (рис. 3.35). По мере заряда барьерной емкости напряжение на *p-n*-переходе и ток через диод стремятся к некоторым установившимся значениям, которые определяются активной составляющей сопротивления *p-n*-перехода.

В момент переключения диода напряжение на барьерной емкости не может измениться мгновенно, оно достигает установившегося значения через некоторое время. Ток через диод также зависит от времени, что характерно для емкостного сопротивления.

Работа диода в схеме с генератором тока. Осциллограммы импульса прямого тока малой амплитуды и падения напряжения на диоде при питании диода от генератора тока представлены на рис. 3.36. В первый момент пропускания импульса через диод весь ток состоит из емкостной составляющей. Поэтому напряжение на диоде в первый момент определяется падением напряжения на сопротивлении базы диода. По мере заряда барьерной емкости увеличивается и напряжение на диоде.

При выключении диода на нем некоторое время сохраняется остаточное напряжение, уменьшающееся со временем. Остаточное напряжение в данном случае связано с тем, что барьерная емкость остается еще заряженной. По мере разряда этой емкости через активное сопротивление *p-n*-перехода диода уменьшается напряжение на емкости и остаточное напряжение на диоде.

§ 3.19. ВЫПРЯМИТЕЛЬНЫЕ ПЛОСКОСТНЫЕ ДИОДЫ

Выпрямительный полупроводниковый диод — это полупроводниковый диод, предназначенный для выпрямления переменного тока.

Германиевые диоды

Технология изготовления и конструкция. Электронно-дырочные переходы большинства германиевых плоскостных диодов формируют путем сплавления индия в германий с электропроводностью *n*-типа (см. § 2.4). Исходные пластинки толщиной 0,3—0,4 мм вырезают из монокристаллического германия с удельным сопротивлением около 30 Ом·см. От площади *p-n*-перехода зависит допустимое

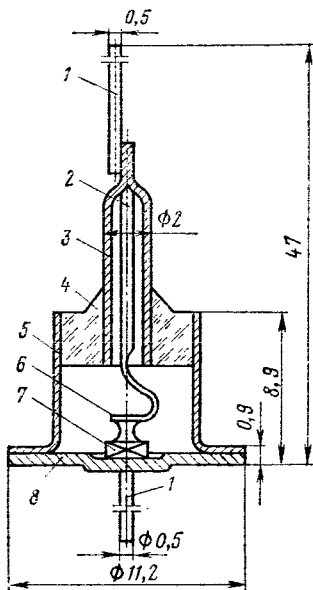


Рис. 3.37. Конструкция германиевого плоскостного диода Д7:

1 — наружные выводы; 2 — внутренний вывод; 3 — коваровая трубка; 4 — стеклянный проходной изолятор; 5 — коваровый баллон (корпус); 6 — электрод; 7 — кристалл германия; 8 — кристаллодержатель

примесью сурьмы, которая является донором для германия и обеспечивает невыпрямляющий контакт.

В мощных германиевых диодах полупроводниковая пластинка припаивается на медное основание с использованием молибденовой прокладки. Назначение последней — уменьшать механические напряжения, возникающие при работе диода из-за разных температурных коэффициентов линейного расширения меди и германия.

Для защиты $p-n$ -перехода от внешних воздействий кристалл германия помещают в корпус (рис. 3.37). В корпус сварен стеклянный проходной изолятор, через который проходит трубка для вывода. Наружная поверхность стеклянного проходного изолятора покрывается светонепроницаемым лаком, чтобы свет не проник в кристалл германия. Кванты света могли бы создать вблизи $p-n$ -перехода пары электрон — дырка и тем самым увеличить обратный ток диода.

Конструкция мощных германиевых диодов предусматривает их работу с естественным либо искусственным воздушным или водяным охлаждением.

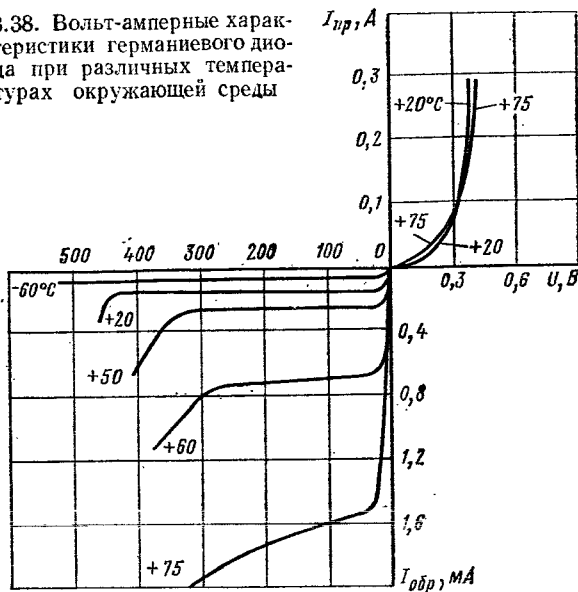
Электрические свойства. Германиевые плоскостные диоды различных типов рассчитаны на различные допустимые прямые токи —

значение прямого тока через диод. Поэтому площадь $p-n$ -перехода рассчитывают, исходя из допустимой плотности прямого тока, которая для германиевых диодов равна 100 A/cm^2 .

Пластинку германия с $p-n$ -переходом припаивают к металлическому кристаллодержателю. При пайке важно получить невыпрямляющий контакт. Во-первых, это необходимо для уменьшения падения напряжения при прохождении прямого тока через диод и, во-вторых, для устранения возможности инжекции неосновных носителей контактом кристалл германия — кристаллодержатель. В случае инжекции неосновных носителей в кристалл n -типа со стороны этого контакта при приложении к диоду обратного напряжения дырки из-за большой диффузионной длины в германии могут доходить до $p-n$ -перехода, увеличивая обратный ток диода и снижая пробивное напряжение.

В связи с изложенным пайку кристалла n -германия к кристаллодержателю производят припоем с

3.38. Вольт-амперные характеристики германиевого диода при различных температурах окружающей среды



от 0,3 до 1000 А. Падение напряжения на диодах при этих токах не превышает 0,5 В.

Температурная зависимость прямого падения напряжения на германиевых диодах может быть различной при малых и при больших прямых токах (рис. 3.38). При малых прямых токах, когда почти все внешнее напряжение приложено к *p-n*-переходу, прямое падение напряжения уменьшается с увеличением температуры в связи с уменьшением высоты потенциального барьера и с перераспределением носителей по энергиям (см. § 3.2). Поэтому прямое падение напряжения уменьшается при увеличении температуры.

При больших для данного диода прямых токах прямое падение напряжения может зависеть от сопротивления базы, которое увеличивается с увеличением температуры из-за уменьшения подвижности носителей заряда. Поэтому при больших прямых токах прямое падение напряжения на диоде может возрастать.

Обратная ветвь вольт-амперной характеристики германиевых плоскостных диодов имеет участок насыщения, так как обратные токи германиевых диодов связаны в основном с процессом экстракции неосновных носителей из базы. Допустимое обратное напряжение не превышает 400 В. Пробой германиевых диодов имеет тепловой характер. Поэтому пробивное напряжение уменьшается с повышением температуры.

Верхний предел диапазона рабочих температур германиевых диодов (75—85 °С) определяется резким ухудшением выпрямления из-за роста обратного тока — сказывается тепловая ионизация атомов германия. Нижний предел диапазона рабочих температур (—60 °С) обусловлен различием температурных коэффициентов

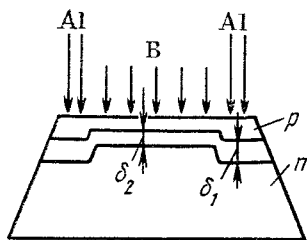


Рис. 3.39. Структура плоскостного кремниевого диода с контролируемым лавинообразованием, p - n -переход которого сформирован диффузией бора в центральной части пластинки и диффузией алюминия в кольцевой части пластинки

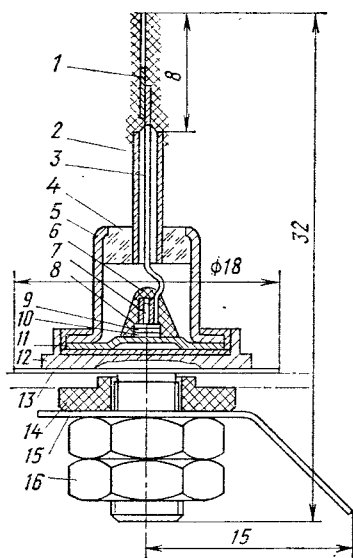


Рис. 3.40. Конструкция кремниевого плоскостного диода (D202—D205):

1 — наружный вывод; 2 — трубка; 3 — внутренний вывод; 4 — проходной изолятор; 5 — корпус; 6 — лаковое защитное покрытие; 7 — алюминий; 8 — кристалл кремния; 9 — сплав золото — сурьма; 10 — слой олова; 11 — кристаллодержатель; 12 — теплоотводящее основание; 13 — слюдяная изолирующая шайба; 14 — изоляционная втулка; 15 — токоотводящий элемент; 16 — крепежные гайки

расширения германия и индия: при низких температурах возникают механические напряжения, которые могут привести к растрескиванию монокристалла германия.

Германиевые плоскостные диоды могут быть использованы для выпрямления переменного тока частотой не более 50 кГц, что связано с инерционностью процессов накопления и высвобождения неосновных носителей заряда в базе (см. § 3.18).

Выпускаемые промышленностью германиевые плоскостные диоды могут быть классифицированы по величине допустимого прямого тока на диоды малой мощности (допустимый прямой ток не превышает 0,3 А), диоды средней мощности (допустимый прямой ток от 0,3 до 10 А) и диоды большой мощности (допустимый прямой ток свыше 10 А).

Для выпрямления больших напряжений при отсутствии высоковольтных диодов может быть применено последовательное включение низковольтных диодов. При этом должны быть учтены специфические особенности полупроводниковых диодов. Диоды одного типа имеют большой разброс обратных сопротивлений и пробивных напряжений. Кроме того, обратные вольт-амперные характеристики диодов по-разному могут изменяться при изменении температуры и со временем. Все это приводит к тому, что приложенное к последовательно соединенным диодам напряженно падает в основном на диоде с большим обратным сопротивлением. В результате возможен необратимый пробой. Выход из строя одного из диодов в последовательной цепочке ведет к пробую остальных.

Для надежной работы последовательно включенных германиевых диодов рекомендуется каждый из них шунтировать резистором постоянного сопротивления с таким расчетом, чтобы распределение напряже-

ний на диодах определялось в основном сопротивлениями шунтирующих резисторов.

Существенным недостатком германиевых диодов является то, что они плохо выдерживают даже очень кратковременные импульсные перегрузки. Определяется это механизмом пробоя германиевых диодов — тепловым пробоем, происходящим при шнуровании тока с выделением большой удельной мощности в месте пробоя.

Кремниевые диоды

Технология изготовления и конструкция. Разнообразные типы кремниевых выпрямительных плоскостных диодов с технологической точки зрения являются либо сплавными, либо диффузионными. Сплавные *p-n*-переходы кремниевых диодов формируют в монокристаллических пластинках кремния с электропроводностью *n*-типа вплавлением алюминия; в пластинках кремния с электропроводностью *p*-типа — вплавлением сплава олова с фосфором или золота с сурьмой. Диффузионные *p-n*-переходы формируют диффузией алюминия или бора в пластинки кремния с электропроводностью *n*-типа и диффузией фосфора в пластинки кремния с электропроводностью *p*-типа. Относительно небольшие по площади *p-n*-переходы кремниевых диодов малой мощности создают обычно методом вплавления. Электронно-дырочные переходы большой площади для мощных кремниевых диодов удобнее делать методом диффузии примесей.

Площадь *p-n*-перехода рассчитывают, исходя из необходимого значения допустимого прямого тока и учитывая допустимую плотность прямого тока, которая для кремниевых *p-n*-переходов равна 200 А/см². Толщина исходных пластинок кремния обычно 0,3—0,6 мм.

Пробой реальных *p-n*-переходов часто происходит вблизи поверхности полупроводника, в местах выхода *p-n*-перехода на поверхность монокристалла (см. § 3.14). Для исключения возможности поверхностного пробоя толщину диффузионного *p-n*-перехода (δ_1) в местах выхода его на поверхность пластинки кремния, т. е. в кольцевой части по периферии пластинки, делают большей по сравнению с толщиной *p-n*-перехода (δ_2) в центральной части пластинки (рис. 3.39). Кольцевая часть *p-n*-перехода формируется диффузией алюминия, центральная часть — диффузией бора. Коэффициент диффузии алюминия больше коэффициента диффузии бора. Именно поэтому градиент концентрации акцепторной примеси в кольцевой части оказывается меньше, чем в центральной части *p-n*-перехода, а толщина диффузионного *p-n*-перехода зависит от градиента концентрации примесей (см. рис. 3.17, б). Пробивное напряжение центральной части *p-n*-перехода оказывается меньше пробивного напряжения кольцевой его части. Пробивное напряжение всего диода не зависит от состояния поверхности; создаются условия для равномерного распределения плотности обратного тока при лавинном пробое. Поэтому такие диоды, называемые диодами

с контролируемым лавинообразованием или лавинными диодами, могут выдерживать значительные импульсные перегрузки по обратному току. Диоды с контролируемым лавинообразованием можно соединять последовательно без шунтирующих резисторов, выравнивающих падение напряжения на каждом диоде.

Другой технологической операцией, также направленной на уменьшение возможности поверхностного пробоя, является снятие фаски по периметру пластинки кремния после формирования в ней $p-n$ -перехода (см. рис. 3.39).

Частотные свойства выпрямительных диодов, работающих обычно при большом уровне инжекции, определяются процессами накопления и рассасывания неосновных носителей заряда в базе (см. § 3.18). Поэтому для улучшения частотных свойств кремниевых диодов в исходную пластинку кремния проводят диффузию золота, примесь которого создает энергетические уровни рекомбинационных ловушек и уменьшает время жизни неосновных носителей. Такие кремниевые диоды часто называют частотными, подчеркивая тем самым, что они способны работать при повышенных частотах (до 100 кГц).

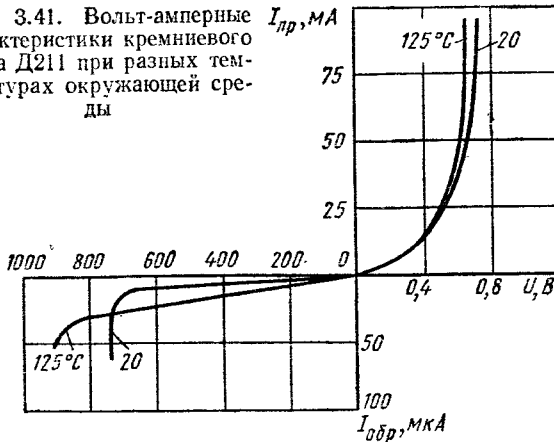
Для защиты от внешних воздействий и обеспечения хорошего теплоотвода пластинку кремния с электронно-дырочным переходом монтируют в металлическом корпусе (рис. 3.40), припаявая ее к медному основанию серебряным сплавом или сплавом золота с сурьмой (1% сурьмы), чтобы получить невыпрямляющий контакт. Иногда (для мощных диодов) между медным основанием и пластинкой кремния помещают металлическую прокладку из вольфрама или кобальта, имеющую примерно тот же коэффициент линейного расширения, что и кремний. Таким образом, устраняют механические напряжения в кристалле кремния при изменении температуры.

Электрические свойства. Вольт-амперные характеристики одного из кремниевых плоскостных диодов (Д211) приведены на рис. 3.41. Кремниевые плоскостные диоды различных типов рассчитаны на различные допустимые прямые токи — от 0,5 до 1600 А. Падение напряжения на диодах при этих токах не превышает обычно 1,5 В. Большие падения напряжения при прохождении прямого тока через кремниевые диоды по сравнению с прямым падением напряжения на германиевых диодах связаны с большей высотой потенциального барьера $p-n$ -переходов, сформированных в кремнии (см. § 2.1).

С увеличением температуры кремниевого диода прямое падение напряжения уменьшается, что связано с уменьшением высоты потенциального барьера и с перераспределением носителей заряда по энергиям (см. § 3.2).

Обратная ветвь вольт-амперной характеристики кремниевых диодов не имеет участка насыщения обратного тока, так как обратный ток в кремниевых диодах вызван в основном процессом генерации носителей в $p-n$ -переходе (см. § 3.10). Допустимое обратное напряжение кремниевых диодов (до 1600 В) значительно превосходит аналогичный параметр германиевых диодов. Пробой кремние-

Рис. 3.41. Вольт-амперные характеристики кремниевого диода Д211 при разных температурах окружающей среды



вых диодов имеет лавинный характер. Поэтому с увеличением температуры пробивное напряжение увеличивается (см. § 3.11).

Верхний предел диапазона рабочих температур кремниевых диодов (приблизительно 125 °С) оказывается большим аналогичного параметра германиевых диодов, так как ширина запрещенной зоны кремния превышает ширину запрещенной зоны германия. Поэтому тепловая генерация носителей заряда в результате ионизации собственных атомов полупроводника (собственная электропроводность) в кремниевых диодах начинает сказываться при больших температурах.

Промышленность выпускает кремниевые выпрямительные диоды малой, средней и большой мощности, рассчитанные на допустимый прямой ток до нескольких сотен ампер.

Для увеличения допустимого обратного напряжения кремниевые диоды соединяют последовательно. Группа последовательно соединенных диодов, оформленных в общем металлическом или пластмассовом корпусе, называется выпрямительным столбом. Некоторые кремниевые выпрямительные столбы рассчитаны на допустимое обратное напряжение до 10 кВ.

Помимо выпрямительных столбов выпускают также и кремниевые выпрямительные блоки — группы диодов, соединенные по различным выпрямительным схемам (мостовым, схемам удвоителя напряжения и т. д.).

§ 3.20. СЕЛЕНОВЫЕ ВЫПРЯМИТЕЛИ

Технология изготовления и конструкция. Селеновые выпрямительные пластины (рис. 3.42) изготавливают на стальной или алюминиевой подложке (основании). В целях обеспечения возможно меньшего переходного сопротивления на подложку наносят дополнительное покрытие. На стальные основания гальваническим методом осаждают никель, а на алюминиевые — тонкий слой (1—2 мкм) висмута или никеля (висмут наносят испарением в вакууме).

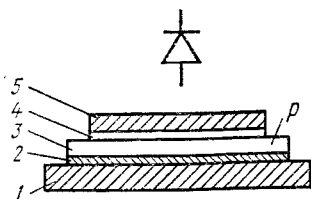


Рис. 3.42. Структура селеновой выпрямительной пластины:

1 — стальная или алюминиевая подложка; 2 — слой висмута или никеля; 3 — слой селена; 4 — слой селенида кадмия; 5 — электрод из сплава висмута, кадмия и олова

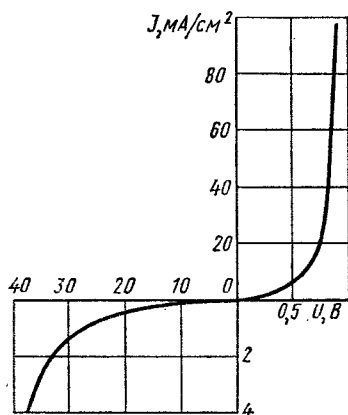


Рис. 3.43. Вольт-амперная характеристика селеновой выпрямительной пластины

Затем на подложку наносят аморфный селен. Для получения большей удельной проводимости селена в него вводят акцепторную примесь (хлор).

Следующей операцией является термообработка при температуре 215°C , близкой к температуре плавления селена. При этом происходит кристаллизация селена и резко уменьшается его удельное сопротивление. Для создания выпрямляющего электрического перехода на поверхность селена наносят легкоплавкий сплав (Bi, Cd, Sn). При нанесении верхнего электрода кадмий реагирует с селеном и образуется тонкий слой селенида кадмия. Таким образом, выпрямляющим контактом в селеновых выпрямителях является гетеропереход, возникающий между селеном и селенидом кадмия.

Для улучшения выпрямляющих свойств селеновые пластины подвергают электрической формовке, которую производят путем длительного приложения постоянного напряжения в обратном направлении. При этом создаются благоприятные условия для диффузии кадмия в селен и наращивания слоя селенида кадмия. Прохождение электрического тока способствует образованию равномерного по толщине гетероперехода, так как в наиболее слабых местах увеличивается плотность тока, что приводит к локальному повышению температуры и более интенсивному соединению кадмия с селеном.

Практически в установках для выпрямления переменного тока в большинстве случаев соединяют ряд селеновых пластин последовательно для получения более высокого выпрямленного напряжения и параллельно для выпрямления больших токов. При этом в один конструктивный узел входит несколько отдельных пластин, образуя выпрямительные столбы и блоки.

Электрические свойства. Вольт-амперная характеристика селеновой выпрямительной пластины представлена на рис. 3.43.

В селеновых выпрямителях допускается плотность прямого тока до $0,1\text{ A}/\text{cm}^2$, допустимые обратные напряжения до 60 В . Большая барьерная емкость при относительно малой допустимой плотности прямого тока ограничивает применение селеновых выпрямителей при повышенных частотах.

Параметры селеновых выпрямителей изменяются со временем как при хранении, так и в работе. Длительное хранение приводит к возрастанию обратного тока в 5—10 раз. Этот процесс, известный под названием расформовки, заключается в увеличении обратного тока выпрямляющего гетероперехода вследствие диффузии в него галоидов и из-за некоторого изменения химического состава полупроводника. Расформовка, как правило, — процесс обратимый. После подачи на выпрямитель переменного напряжения обратный ток уменьшается, достигая номинальной величины в течение 2—3 мин.

Прохождение через селеновый элемент только прямого тока сильно увеличивает расформовку. Поэтому нельзя применять селеновые выпрямители в тех цепях, где через диод приходит только прямой ток.

Необратимое возрастание прямого падения напряжения селеновых выпрямителей, называемое старением, при хранении незначительно, но ускоряется при эксплуатации выпрямителей. Процесс старения связан с увеличением сопротивления слоя селена вследствие обеднения его галогенными примесями, а также с увеличением сопротивления контакта селена с электродом. Интенсивность этих процессов возрастает с повышением температуры. Старение становится особенно активным, когда температура селеновых элементов превышает 75 °С. Эта температура считается максимальной рабочей температурой селеновых выпрямителей.

Срок службы селеновых выпрямителей ~ 10 000 ч. При этом под сроком службы понимают время, по истечении которого выпрямленное напряжение снижается на 10 %.

Важнейшей особенностью селеновых выпрямителей является их способность выдерживать значительные кратковременные перегрузки и быстро восстанавливаться при пробое. Пробой возникает не по всей площади гетероперехода, а только в отдельных «слабых» местах. При этом из-за резкого повышения плотности тока происходит местное расплавление селена, который затем при быстром охлаждении переходит в аморфное состояние. Удельное сопротивление аморфного селена на несколько порядков выше, чем кристаллического. Таким образом, пробитое место выпрямительной пластины оказывается изолированным, происходит «самозалечивание» места пробоя.

В процессе изготовления селеновых выпрямителей их подвергают воздействию высокого импульсного напряжения для пробоя «слабых» мест.

§ 3.21. ВЫПРЯМИТЕЛЬНЫЕ ТОЧЕЧНЫЕ ВЫСОКОЧАСТОТНЫЕ ДИОДЫ

Конструкция германиевого точечного диода показана на рис. 3.44. Диод состоит из кристалла германия, припаянного к кристаллодержателю, контактного электрода в виде тонкой проволоочки и стеклянного баллона.

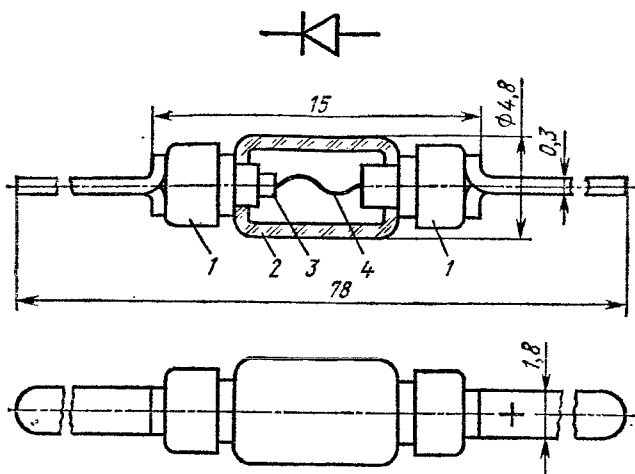


Рис. 3.44. Конструкция германиевого точечного диода Д2:
1 — выводы; 2 — стеклянный баллон; 3 — кристалл германия;
4 — электрод из вольфрамовой проволоки

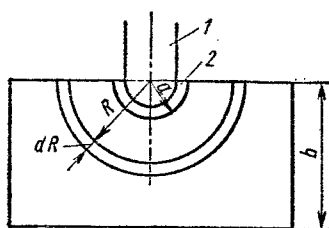


Рис. 3.45. Структура точечного $p-n$ -перехода:

1 — металлический проволочный электрод; 2 — слой полупроводника с электропроводностью p -типа в полупроводнике n -типа

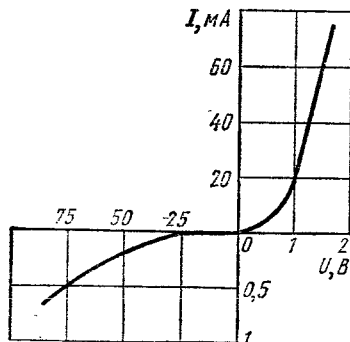


Рис. 3.46. Вольт-амперная характеристика германиевого точечного диода

Германиевые точечные выпрямительные диоды обычно изготавливают на основе высокоомного германия n -типа с удельным сопротивлением 10—30 Ом·см. Пластины германия шлифуют, полируют и протравливают в кислотах или очищают электролитическим методом. Электролитически заостренную тонкую проволоку приваривают к германию импульсом тока до 1 А. В момент приварки в германии под проволочным электродом образуется область с электропроводностью p -типа. Существенное влияние на результаты формовки оказывает состав материала контактной пружины. Поэтому контактную пружину иногда покрывают индием или другими материалами, чтобы получить высококачественный $p-n$ -переход. Таким образом, в германиевых точечных высокочастотных диодах выпрямление происходит на $p-n$ -переходах, образованных в результате формовки.

Сопротивление базы точечного диода определяется сопротивлением растекания r_s , т. е. сопротивлением области полупроводника, расположенной вблизи точечного контакта,

где сгущаются линии тока. Поэтому сопротивление растекания зависит от площади точечного контакта: чем меньше эта площадь, тем больше величина r_s .

Определим r_s , если контакт имеет полусферическую форму (рис. 3.45). Здесь a — радиус p - n -перехода под точечным контактом; b — толщина кристалла полупроводника. Сопротивление полусферического слоя толщиной dR на расстоянии R от центра сферы

$$dr = \rho \frac{dR}{2\pi R^2}.$$

Полное сопротивление растекания определяется суммой сопротивлений полусферических слоев. Если удельное сопротивление ρ считать постоянным, то

$$r_s = \int_0^b dr = \frac{\rho}{2\pi} \int_a^b \frac{dR}{R^2} = \frac{\rho}{2\pi} \left(\frac{1}{a} - \frac{1}{b} \right).$$

Значение a мало по сравнению с b . Обычно $a \approx 2 \cdot 10^{-3}$ см, $b \approx 2 \cdot 10^{-2}$ см. В этих условиях величиной $1/b$ можно пренебречь, тогда

$$r_s \approx \frac{\rho}{2\pi a}. \quad (3.103)$$

Действительная величина сопротивления растекания может быть несколько меньше вычисленной из-за инжекции неосновных носителей через переход в объем кристалла.

На рис. 3.46 изображена вольт-амперная характеристика германиевого точечного диода. При достаточно больших прямых токах сопротивление перехода мало по сравнению с сопротивлением растекания, и ток через точечный диод определяется величиной r_s .

Германиевые точечные высокочастотные диоды могут иметь обратное напряжение до 350 В и прямой ток до 100 мА при $U_{пр} = 1-2$ В. Барьерная емкость точечных германиевых диодов мала (около 1 пФ), но при СВЧ они применяться не могут из-за эффекта накопления. При частоте выше 150 МГц инжектированные носители заряда за время действия обратного напряжения не успевают рекомбинировать и уйти из базы.

Кремниевые точечные диоды в конструктивном отношении почти не отличаются от германиевых. Для их изготовления используют кремний с электропроводностью n -типа с удельным сопротивлением 3—8 Ом·см. Контактная пружина покрывается алюминием, что обеспечивает после формовки образование высококачественного электронно-дырочного перехода.

§ 3.22. ИМПУЛЬСНЫЕ ДИОДЫ

Импульсный полупроводниковый диод — это полупроводниковый диод, предназначенный для применения в импульсных режимах.

Импульсные диоды чаще всего работают при больших сигналах. Поэтому свойства и параметры импульсных диодов определяются процессами, рассмотренными в первой части § 3.18.

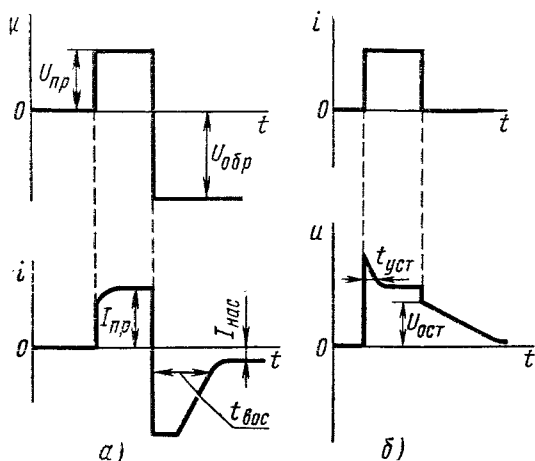


Рис. 3.47. Осциллограммы токов и напряжений для импульсных диодов при его работе в схемах с генератором напряжения (а) и с генератором тока (б)

При переключении диода с прямого напряжения на обратное в начальный момент наблюдается большей обратный ток, ограниченный в основном последовательным сопротивлением r_s . Все инжектированные неосновные носители рекомбинируют или уходят через переход, после чего обратный ток уменьшается до своего стационарного значения $I_{нас}$ (рис. 3.47, а). В связи с этим явлением введен параметр — время восстановления обратного сопротивления или тока $t_{1,0.}$, равное

интервалу времени между моментом переключения напряжения на диод с прямого на обратное и моментом, когда обратный ток достигает заданного значения.

По этому параметру импульсные диоды можно подразделить на две группы:

- 1) скоростные, или микросекундные, с $0,1 \text{ мкс} < t_{вос} < 0,1 \text{ мс}$;
- 2) сверхскоростные, или наносекундные, с $t_{вос} < 0,1 \text{ мкс}$.

Эффект накопления проявляется тем сильнее, чем больший прямой ток протекает через диод перед переключением. Чем выше значение прямого тока, тем большая концентрация неосновных носителей создается в базе диода и тем больше время восстановления. При подаче на диод прямого напряжения ток устанавливается не сразу, так как с течением времени происходит накопление инжектированных неосновных носителей и снижение сопротивления базы. Передний фронт импульса тока получается искаженным (рис. 3.47, а). Однако этот процесс протекает быстро и не имеет практически столь большого значения, как переходные процессы при восстановлении обратного тока.

При пропуске импульса прямого тока через диод наблюдается выброс напряжения в первый момент после включения (рис. 3.47, б). Это вызвано повышенным падением напряжения до тех пор, пока инжектированные носители не понизят сопротивления базы диода. В связи с этим для импульсных диодов вводят параметр — время установления прямого сопротивления или напряжения $t_{уст}$, равное интервалу времени между началом протекания через диод прямого тока заданной величины и моментом, когда напряжение на диоде достигает 1,2 установившегося значения.

К специфическим параметрам импульсных диодов кроме $t_{\text{нос}}$ и $t_{\text{уст}}$ можно отнести еще максимальный ток восстановления $I_{\text{обр. н макс}}$, равный наибольшему значению обратного тока через диод после переключения напряжения на диоде с прямого на обратное, а также максимальное обратное сопротивление $r_{\text{пр. н макс}}$, равное отношению наибольшего значения амплитуды импульса прямого напряжения на диоде к току, при котором производилось измерение.

В настоящее время в качестве импульсных диодов используют точечные и плоскостные диоды. Технология изготовления импульсных точечных диодов аналогична технологии изготовления выпрямительных точечных высокочастотных диодов. Плоскостные импульсные диоды изготавливают либо методом сплавления примесей, либо диффузией донорных или акцепторных примесей в полупроводниковый кристалл. Практически технология изготовления импульсных плоскостных диодов почти не отличается от технологии обычных выпрямительных диодов. Разница заключается только в выборе исходного материала и в площади p - n -перехода. Импульсные диоды должны иметь как можно меньшую площадь перехода для получения наименьшей емкости. Исходным для импульсных диодов выбирают полупроводниковый материал с малым временем жизни носителей заряда.

§ 3.23. ПЛОСКОСТНЫЕ ДИОДЫ С ВЫПРЯМЛЕНИЕМ НА КОНТАКТЕ МЕТАЛЛ — ПОЛУПРОВОДНИК

Первыми приборами с использованием выпрямления на контакте металл — полупроводник были точечные диоды с прижимным металлическим контактом (см. § 3.24). Основные математические соотношения для электрических характеристик таких контактов получил немецкий ученый В. Шоттки, вследствие чего потенциальный барьер на контакте металл — полупроводник часто называют барьером Шоттки. Однако точечные диоды с прижимным металлическим контактом имеют свойства, отличные от свойств идеализированной модели из-за значительных механических напряжений в приконтактной области, наличия промежуточных окисных слоев, мультikonтактности и т. п. Лишь в последние годы удалось получить совершенные контакты металл — полупроводник с характеристиками идеального барьера Шоттки.

Технология изготовления и конструкция. Обычно для изготовления диодов с барьером Шоттки используют подложки из низкогомоного n -кремния с высокоомным тонким эпитаксиальным слоем того же кремния (рис. 3.48).

На поверхность эпитаксиального слоя наносят металлический электрод. Металл выбирают таким образом, чтобы получить выпрямляющий контакт, но не инжектирующий неосновные носители в базовую область (см. § 2.10). Основным методом нанесения металла является его осаждение на монокристалл полупроводника при вакуумном испарении. Для лучшей адгезии металл осаждают обычно на подогретый монокристалл (например, оптимальная температура

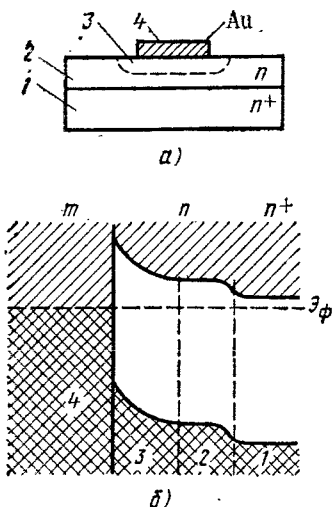


Рис. 3.48. Структура диода с выпрямлением на контакте металл — полупроводник (а) и его энергетическая диаграмма (б):

1 — низкоомный исходный кристалл кремния; 2 и 3 — эпитаксиальный слой высокоомного кремния; 3 — область объемного заряда; 4 — металлический (золотой) контакт

основных носителей в базе. В результате инерционность диодов с выпрямлением на контакте металл — полупроводник определяется величиной барьерной емкости выпрямляющего контакта. При малой площади нанесенного электрода (диаметр 20—30 мкм) и при относительно широком потенциальном барьере из-за высокого удельного сопротивления эпитаксиального слоя барьерная емкость диода может иметь величину около 1 пФ.

Важным преимуществом диодов с выпрямлением на контакте металл — полупроводник по сравнению с диодами на $p-n$ -переходах является также незначительность последовательных электрических и тепловых сопротивлений, так как металлический слой по этим свойствам превосходит любой сильнолегированный слой полупроводника.

В связи с тем, что барьерная емкость и последовательное сопротивление в таких диодах небольшие, соответственно мало и время перезаряда емкости; это дает возможность использовать диоды с выпрямлением на контакте металл — полупроводник в качестве сверхскоростных импульсных диодов.

Рассматриваемые диоды обладают почти идеальной прямой ветвью вольт-амперной характеристики, т. е. вольт-амперная характеристика соответствует выражению (3.27). При этом с изменением прямого тока в пределах нескольких порядков зависимость $\lg I_{пр} = f(U_{пр})$ близка к линейной, или в показателе экспоненты

при нанесении золота на кремний близка к 200 °С). Для очистки поверхности полупроводника перед нанесением металла производят химическое травление. Лучшее результаты дает ионно-плазменная бомбардировка кремния непосредственно в рабочей камере перед испарением металла.

Иногда для нанесения металлического контакта применяют метод химического осаждения путем восстановления водородом газообразного соединения металла при повышенной температуре. Таким методом наносят металлы, обладающие высокой температурой плавления (вольфрам, молибден, ванадий и др.). Электрические свойства выпрямляющего контакта с такими металлами более стабильны.

Электрические свойства. Поскольку в рассматриваемых диодах из-за разной высоты потенциальных барьеров для электронов и дырок нет инжекции неосновных носителей заряда, то нет и таких медленных процессов, как накопление и рассасывание не-

при изменении тока не появляются дополнительных множителей (см. § 3.17 и рис. 3.32). Учитывая эту особенность, диоды с выпрямлением на контакте металл — полупроводник можно использовать в некоторых схемах в качестве быстродействующих логарифмических элементов.

Кроме того, важным преимуществом рассматриваемых диодов может быть их низкая стоимость по сравнению со стоимостью диодов на *p-n*-переходах.

При использовании выпрямления на контакте металл — полупроводник можно создавать мощные высокочастотные выпрямительные диоды. Лабораторные образцы таких диодов способны работать на частотах до 1 МГц при $U_{обр} = 50$ В и $I_{пр} = 10$ А.

§ 3.24. СВЧ-ДИОДЫ

Большинство СВЧ-диодов представляют собой точечные диоды, выпрямление в которых происходит на контакте металл — полупроводник. Особенностью таких контактов является возможность выпрямления без инжекции неосновных носителей в кристалл полупроводника, т. е. в базу диода. Отсутствие инжекции неосновных носителей приводит, как отмечалось, к отсутствию процессов накопления и рассасывания неосновных носителей в базе диода. Именно эти процессы ограничивают частотный диапазон работы большинства полупроводниковых диодов.

Выпрямительные диоды

Конструкция СВЧ-диодов показана на рис. 3.49. Корпус диода патронной конструкции (рис. 3.49, *а*) состоит из двух металлических фланцев, разделенных керамической втулкой. На нижнем фланце крепится кристаллодержатель с полупроводниковым кристаллом, а в верхнем — настроечный штифт с вольфрамовой контактной пружинкой, которая имеет изгибы в обе стороны от оси. Двусторонний симметричный изгиб уменьшает тангенциальную составляющую силы давления в точке контакта и предотвращает скольжение пружинки по кристаллу. Для повышения поверхностной проводимости, улучшения контакта и коррозионной стойкости металлические фланцы серебрят или золотят.

В целях уменьшения собственной емкости и индуктивности СВЧ-диода в качестве корпуса для него используют непосредственно элементы волноводных систем. Для работы в коаксиальных фидерных линиях применяют коаксиальную конструкцию СВЧ-диода (рис. 3.49, *б*). Конструкцию диода, в которой полупроводниковый кристалл и контактную пружинку монтируют непосредственно в отрезке волновода, называют волноводной (рис. 3.49, *в*).

Эквивалентная схема СВЧ-диода кроме сопротивления растекания r_6 , емкости перехода $C_{бар}$ и сопротивления перехода $r_{пер}$ включает в себя индуктивность контактной пружинки L и емкость патрона $C_{корп}$ (рис. 3.50).

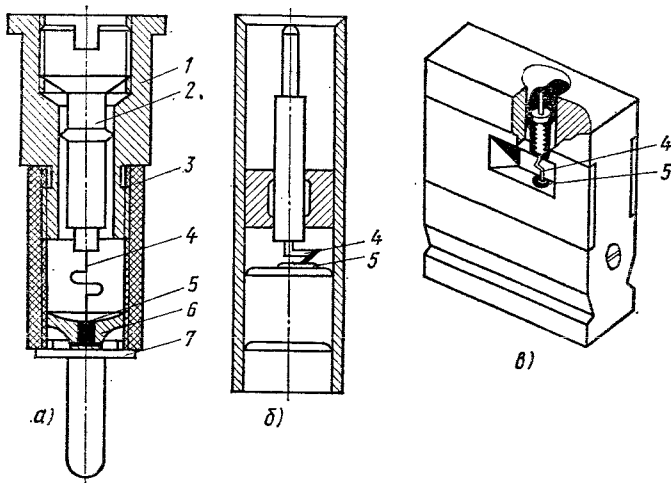


Рис. 3.49. Конструкция СВЧ-диодов:

a — патронная; *б* — коаксиальная; *в* — волноводная;
 1 — верхний фланец; 2 — настроечный штифт; 3 — керамическая втулка; 4 — контактная пружинка; 5 — кристалл полупроводника; 6 — кристаллодержатель; 7 — нижний фланец

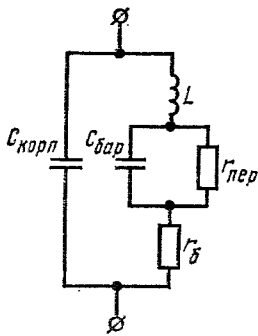


Рис. 3.50. Эквивалентная схема СВЧ-диода:

L — индуктивность контактной пружины; $C_{бар}$ — барьерная емкость $p-n$ -перехода; $r_{пер}$ — сопротивление $p-n$ -перехода; $C_{корп}$ — емкость корпуса

Представить диод в виде системы с сосредоточенными параметрами можно, только если линейные размеры (толщина перехода, радиус перехода) малы по сравнению с длиной волны. Толщина выпрямляющего слоя в СВЧ-диодах имеет величину порядка долей микронметра и не определяет частоту, до которой можно пользоваться данной эквивалентной схемой. Полусферическая область кристалла, определяющая сопротивление $r_б$, примыкает к точечному контакту и имеет радиус в несколько десятков микронметров при радиусе контактного острья в несколько микронметров. Следовательно, сопротивление растекания можно представить в эквивалентной схеме в виде сосредоточенного параметра даже до длин волн порядка миллиметра. Размеры конструктивных элементов СВЧ-диода уже в диапазоне волн менее 3 см становятся соизмеримыми с длиной волны. При этом индуктивность контактной пружины и емкость патрона можно рассматривать как сосредоточенные лишь с известными допущениями.

Важной характеристикой СВЧ-диодов (особенно если они предназначены для применения во входных цепях приемных устройств)

является коэффициент шума, который можно определить как изменение отношения мощности сигнала к мощности шума на входе и выходе рассматриваемой системы:

$$F = (P_{\text{сигн}}/P_{\text{ш}})_{\text{вх}} / (P_{\text{сигн}}/P_{\text{ш}})_{\text{вых}}$$

Иногда коэффициент шума определяют как отношение мощности шумов на выходе рассматриваемого четырехполюсника к мощности шумов на входе идеального (нешумящего) четырехполюсника с таким же коэффициентом передачи мощности. Нетрудно показать, что эти определения коэффициента шума идентичны.

Часто используют также такой параметр, как относительная температура шума. Параметр этот не имеет размерности, и название «температура» является в определенной степени условным. Собственные шумы детектора превосходят шумы чисто активного линейного сопротивления того же значения. Мощность флуктуационных шумов активного сопротивления в диапазоне частот Δf определяется соотношением

$$P_{\text{ш}R} = kT_0 \Delta f,$$

где T_0 — температура, при которой производится сравнение диода с эквивалентным сопротивлением.

При некоторой более высокой температуре T_1 шумы эквивалентного сопротивления равны шумам диода (имеется в виду диод в рабочем режиме). Отношение этих температур и принято называть относительной температурой шума:

$$t_{\text{ш}} = T_1/T_0.$$

За стандартную температуру принимают $T_0 = 290$ К.

Основным параметром одного из видов СВЧ-диода — в и д е о - д е т е к т о р а является добротность

$$Q = \frac{\beta_I r}{\sqrt{rt_{\text{ш}} + R_{\text{ш}}}},$$

где β_I — чувствительность по току, определяемая как отношение выпрямленного тока к поданной на диод мощности СВЧ в режиме короткого замыкания цепи выпрямленного тока: $\beta_I = I_{\text{пр}}/P_{\text{вх}}$; r — сопротивление диода в рабочей точке; $R_{\text{ш}}$ — эквивалентное шумовое сопротивление на входе видеоусилителя (обычно 1000 Ом).

Кроме добротности и чувствительности по току важным параметром видеодетектора является коэффициент стоячей волны напряжения (КСВН) — отношение максимальной напряженности к минимальной напряженности электрического поля в волноводе, нагрузкой которого является камера с диодом.

Статические вольт-амперные характеристики СВЧ-диодов отличаются от статических характеристик точечных высокочастотных диодов: у СВЧ-диодов пробивное напряжение значительно ниже. Низкое пробивное напряжение объясняется тем, что для изготовления СВЧ-диодов применяют сильнолегированные полупровод-

никовые материалы. Низкое удельное сопротивление исходного материала необходимо для уменьшения сопротивления растекания диода, а также для уменьшения времени жизни неосновных носителей.

К вольт-амперным характеристикам *измерительных диодов* предъявляются особые требования. Прямая ветвь их вольт-амперной характеристики примерно квадратична, т. е. ток пропорционален квадрату приложенного напряжения. Это дает возможность, измеряя выпрямленный ток, определять мощность, которая пропорциональна квадрату напряженности электрического поля в волноводе.

Электрические данные германневого видеодетектора Д602А	
Номинальная рабочая длина волны, см	3,2
Чувствительность по току при подводимой мощности не более 0,02 мВт, А/Вт	1,5
Добротность, Вт ^{-1/2}	15
КСВН	3,2

Электрические данные кремниевого измерительного СВЧ-диода ДК-И1

Номинальная рабочая длина волны, см	10
Чувствительность по току при подводимой мощности не более 0,02 мВт, А/Вт	0,5
Выпрямленный ток при подводимой мощности 0,5 мВт, мА	0,4

Переключательные диоды

В переключательных диодах используется резкое изменение сопротивления диода переменному току при подаче на него прямого и обратного постоянного смещения. Переключательные диоды можно подразделить на *резонансные* и *нерезонансные*.

В *резонансных переключательных диодах* при прямом смещении малое активное сопротивление *p-n*-перехода $r_{пер}$ шунтирует емкость $C_{бар}$ (см. рис. 3.50). В этом случае диод можно представить как параллельный колебательный контур, состоящий из емкости патрона диода $C_{корп}$, индуктивности контактной пружинки L и сопротивления потерь $(r_б + r_{пер})$. Таким образом, при прямом смещении диод будет обладать большим сопротивлением для СВЧ-сигнала определенной (резонансной) частоты, т. е. для сигнала с частотой $\omega = 1/\sqrt{LC_{корп}}$.

При обратном смещении диод можно представить в виде последовательного контура, состоящего из индуктивности контактной пружинки L и емкости *p-n*-перехода $C_{бар}$ с последовательно включенным сопротивлением потерь $r_б$. Кроме того, этот последовательный контур зашунтирован емкостью патрона диода $C_{корп}$. Таким образом, при обратном смещении диод будет оказывать малое сопротивление СВЧ-сигналу определенной (резонансной) частоты.

Переключая диод из состояния параллельного резонанса в состояние последовательного резонанса путем изменения полярности постоянного смещения, можно резко изменять полное сопротивление

ние в СВЧ-тракте и коммутировать (закрывать или открывать) тот или иной тракт для СВЧ-мощности. Для идеального резонансного переключающего диода необходимо иметь большие добротности параллельного и последовательного контуров, т. е. необходимо, чтобы активная составляющая дифференциального сопротивления p - n -перехода при прямом смещении диода была мала. Малым должно быть и сопротивление базы диода r_b .

В качестве переключаемых используют точечные и плоскостные диоды, параметры которых обеспечивают резонанс на выбранной частоте. Для этого диоды должны обладать малой собственной емкостью. Снижение емкости может быть достигнуто за счет расширения области объемного заряда, т. е., например, использования диодов со структурой p - i - n .

В нерезонансных переключаемых диодах используется изменение активной составляющей динамического сопротивления диода при изменении полярности постоянного смещения. Нерезонансный переключаемый диод обладает малым сопротивлением для СВЧ-сигнала при прямом смещении и большим сопротивлением при обратном смещении.

В конструкции нерезонансного переключающего диода должны быть предусмотрены уменьшенные до возможного минимума индуктивность, емкость p - n -перехода и емкость патрона. В связи с этим требованием нерезонансные переключаемые диоды делают без контактной проволоки и без специального корпуса, а емкость p - n -перехода часто компенсируют путем параллельного подключения внешней индуктивности. При этом сопротивление всей системы (диод с внешней индуктивностью) становится практически активным.

Поскольку переключаемые диоды предназначены для работы в двух фиксированных рабочих точках, форма их вольт-амперных характеристик не имеет существенного значения, особенно для резонансных переключаемых диодов.

Преобразовательные диоды

Преобразовательные СВЧ-диоды применяют в качестве *смесителей, умножителей и модуляторов*. Все эти функции диоды могут выполнять благодаря нелинейности вольт-амперной характеристики.

Смесительные СВЧ-диоды характеризуются параметрами, определяющими работоспособность этих приборов в схемах, предназначенных для преобразования высокочастотного сигнала и напряжения гетеродина в сигнал промежуточной частоты. В связи с этим основными параметрами смесительных СВЧ-диодов являются потери преобразования, коэффициент шума и выпрямленный ток.

Потери преобразования определяются как отношение мощности СВЧ-сигнала, подаваемого на смеситель, к мощности промежуточной частоты:

$$L_{прб} = P_{вх} / P_{пч}$$

Выпрямленный ток определяет выпрямляющие свойства СВЧ-диода на рабочей частоте и до известной степени потери преобразования (чем больше выпрямленный ток, тем меньше потери преобразования). Величину выпрямленного тока измеряют обычно при подаче на диод некоторого фиксированного уровня мощности.

Кроме перечисленных параметров, смесительные диоды могут характеризоваться полным входным (по отношению к входному сигналу) и выходным (по отношению к тракту промежуточной частоты) сопротивлениями. Качество смесительного диода в значительной степени определяется свойствами полупроводника, из которого он изготовлен.

У м н о ж и т е л ь н ы е СВЧ-диоды характеризуются мощностью какой-то определенной гармоники при заданном значении мощности основной гармоники, подаваемой на диод; м о д у л ь а т о р н ы е СВЧ-диоды — потерями преобразования по модулирующей частоте (отношением мощности модулируемых колебаний к выходной мощности).

Электрические данные германиевого смесительного диода Д405

Номинальная рабочая длина волны, см	3,2
Потери преобразования, дБ	7,0
Относительная температура шумов при 20 °С	2,2
Выпрямленный ток при подводимой мощности 1 мВт, мА	1,0
Выходное сопротивление, Ом	350—550
КСВН	2,0

Электрические данные кремниевого умножительного диода Д501

Номинальная длина волны основной частоты, см	25,6
Мощность 8-й гармоники при подводимой мощности основной частоты 130 мВт и сопротивлении смещения в цепи умножительного диода 100 Ом, мкВт	300

Электрические данные германиевого модуляторного диода Д401

Рабочий диапазон длин волн, см	7—10
Частота модулирующих колебаний, МГц	70—150
Выходная мощность, мВт	15
Наибольшие потери преобразования по несущей и модулирующей частотам, дБ	13

§ 3.25. СТАБИЛИТРОНЫ

При лавинном и туннельном пробое диодов происходит резкое увеличение обратного тока, а обратное напряжение изменяется очень мало (см. § 3.11 и 3.12). Это свойство использовано для создания стабилизаторов напряжения — полупроводниковых стабилитронов. Таким образом, п о л у п р о в о д н и к о в ы й с т а б и л и т р о н — это полупроводниковый диод, напряжение на котором в области электрического пробоя слабо зависит от тока и который используется для стабилизации напряжения. В связи с тем, что лавинный пробой характерен для диодов, изготовленных на

основе полупроводника с большой шириной запрещенной зоны, исходным материалом для стабилитронов служит кремний.

Пробивное напряжение диода — напряжение стабилизации стабилитрона зависит от толщины $p-n$ -перехода или от удельного сопротивления базы диода (см. рис. 3.21). Поэтому разные типы стабилитронов имеют различные напряжения стабилизации $U_{ст}$ (от 3 до 400 В).

Низковольтные стабилитроны изготавливают на основе сильнолегированного кремния. В связи с этим в низковольтных стабилитронах с напряжением стабилизации менее 6 В происходит туннельный пробой. *Высковольтные стабилитроны* изготавливают на основе слаболегированного кремния. Поэтому их принцип действия связан с лавинным пробоем. Среди различных типов стабилитронов — большинство с лавинным пробоем (рис. 3.51).

Качество стабилитрона, т. е. его способность стабилизировать напряжение при изменении проходящего тока характеризуется дифференциальным сопротивлением стабилитрона $r_{ст}$ — величиной, определяемой отношением приращения напряжения на стабилитроне к вызвавшему его малому приращению тока. Так как максимальным изменениям тока для лучшей стабилизации должны соответствовать минимальные изменения напряжения, то качество стабилитрона выше, если он имеет меньшее дифференциальное сопротивление.

Важным параметром стабилитрона является температурный коэффициент напряжения стабилизации $\alpha_{ст}$. Значения этого параметра у разных стабилитронов различны. Обобщенная зависимость температурного коэффициента напряжения стабилизации от напряжения стабилизации приведена на рис. 3.52. Как видно из рисунка, $\alpha_{ст}$ может иметь положительные значения для высоковольтных

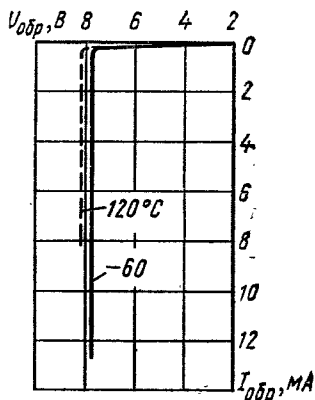


Рис. 3.51. Вольт-амперные характеристики кремниевого стабилитрона Д808 при разных температурах окружающей среды

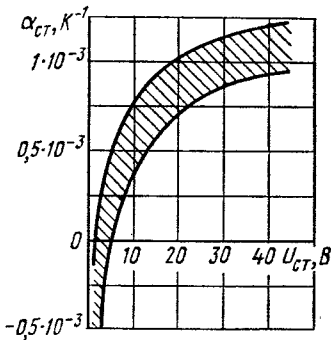


Рис. 3.52. Зависимость температурного коэффициента напряжения стабилизации стабилитронов различных марок от напряжения стабилизации

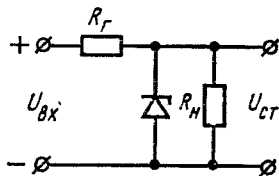


Рис. 3.53. Схема стабилизатора постоянного напряжения с использованием стабилитрона

и отрицательные для низковольтных стабилитронов, что связано с различной температурной зависимостью пробивного напряжения при лавинном и туннельном пробое (см. § 3.11 и 3.12). Изменение знака $\alpha_{ст}$ соответствует напряжению стабилизации $U_{ст} \approx 6$ В. У стабилитронов с таким напряжением стабилизации температурные изменения напряжения стабилизации минимальны.

Один из способов уменьшения температурного коэффициента напряжения стабилизации, который используют для создания *термокомпенсированных стабилитронов*, заключается в последовательном соединении стабилитрона и *p-n*-перехода, включенного в прямом направлении. С повышением температуры падение напряжения на *p-n*-переходе, включенном в прямом направлении, уменьшается. Одновременно падение напряжения на обратном смещенном *p-n*-переходе растет при лавинном пробое. Таким образом, у термокомпенсированных стабилитронов удается получить ничтожно малый температурный коэффициент напряжения стабилизации. Например, у термокомпенсированного стабилитрона СКТ-15 с напряжением стабилизации 8—9 В температурный коэффициент напряжения стабилизации $\alpha_{ст} = (-1 \div 20) \cdot 10^{-5}$ 1/К. Такие стабилитроны можно применять в источниках эталонного напряжения вместо нормальных элементов.

При возникновении лавинного пробоя в стабилитроне появляются шумы. Эти шумы являются отличительной особенностью процесса ударной ионизации. Вначале этот процесс неустойчив: он возникает, срывается, возникает снова. С увеличением тока процесс ударной ионизации становится устойчивым и шумы исчезают. В стабилитронах с напряжением стабилизации менее 6 В шумы практически отсутствуют, так как пробой в них имеет туннельный характер.

Диапазон токов, в котором может работать стабилитрон, устанавливают из следующих соображений. *Минимально допустимый ток стабилизации* определяется тем, что при малых токах, во-первых, дифференциальное сопротивление оказывается еще большим и, во-вторых, могут быть значительные шумы. *Максимально допустимый ток стабилизации* определяется допустимой для данного прибора мощностью рассеяния.

Методом сплавления или диффузии примесей одновременно с двух сторон пластинки кремния можно сформировать два *p-n*-перехода, которые при подаче напряжения на крайние области структуры оказываются включенными встречно. Так изготавливают стабилитроны с симметричной вольт-амперной характеристикой — *симметричные стабилитроны*, которые могут стабилизировать напряжение разной полярности и применяются для защиты различных элементов электрических схем от перенапряжений обеих полярностей.

Конструктивно стабилитроны оформляют аналогично выпрямительным диодам. В зависимости от площади *p-n*-перехода и конструктивного оформления стабилитроны могут иметь различные допустимые мощности рассеяния (от 0,25 до 50 Вт).

На рис. 3.53 изображена простейшая схема использования стабилитрона для стабилизации постоянного напряжения и для защиты различных приборов и элементов электрических схем от перенапряжений. При увеличении входного напряжения резко уменьшается сопротивление стабилитрона, вследствие чего избыточное напряжение падает на сопротивлении гасящего резистора R_r , а напряжение на стабилитроне и на параллельно включенной нагрузке практически остается неизменным.

§ 3.26. СТАБИСТОРЫ

Стабилизацию постоянного напряжения можно также получить при использовании диода, включенного в прямом направлении. Диод, в котором для стабилизации напряжения используется прямая ветвь вольт-амперной характеристики, называют с т а б и с т о р о м. Для изготовления стабисторов применяют кремний с относительно большой концентрацией примесей, что необходимо для получения меньшего сопротивления базы диода и соответственно меньшего дифференциального сопротивления при прямом включении.

Отличительной особенностью стабисторов по сравнению со стабилитронами является меньшее напряжение стабилизации, которое определяется прямым падением напряжения на диоде и для кремниевых стабисторов составляет примерно 0,7 В. Последовательное соединение двух или трех стабисторов дает возможность получить удвоенное или утроенное значение напряжения стабилизации (1,4—2,1 В). Такие цепочки из последовательно соединенных диодов монтируют в одном корпусе.

С увеличением температуры высота потенциального барьера p - n -перехода уменьшается (см. § 2.1), поэтому стабисторы имеют отрицательный температурный коэффициент напряжения стабилизации.

Для использования в качестве стабисторов предназначены кремниевые диоды Д219С, Д220С и Д223С, а также селеновые выпрямители 7ГЕ1А-С и 7ГЕ2А-С.

§ 3.27. ЛАВИННО-ПРОЛЕТНЫЕ ДИОДЫ

Лавинно-пролетный диод — это полупроводниковый диод, работающий в режиме лавинного умножения носителей заряда при обратном смещении электронно-дырочного перехода и предназначенный для генерации сверхвысокочастотных электромагнитных колебаний.

На рис. 3.54, а изображена структура диода p^+n , на рис. 3.54, б — изменение напряженности электрического поля по структуре при обратном напряжении, соответствующем лавинному пробое. Пары электрон — дырка, генерируемые в узкой части области объемного заряда, где напряженность поля достаточна для ударной ионизации, разделяются полем перехода. Ток через диод идет до тех пор, пока носители двигаются по области объемного заряда. Из-за не-

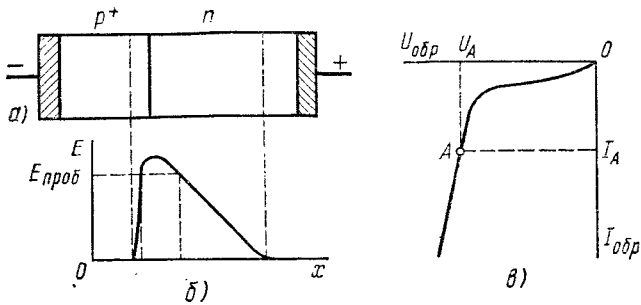


Рис. 3.54. Структура лавинно-пролетного диода (а), распределение напряженности электрического поля по структуре (б) и положение рабочей точки (постоянного смещения) на вольт-амперной характеристике (в)

симметричности p - n -перехода время дрейфа электронов через область объемного заряда до относительно низкоомной n -области получается больше времени дрейфа дырок. Это время дрейфа электронов (время пролета) определяет на высоких частотах фазовый сдвиг между приложенным к диоду напряжением и проходящим током.

Фазовый сдвиг между напряжением и током определяется не только пролетным временем, но и инерционностью лавинного процесса ударной ионизации. Только из-за инерционности развития лавины фазовый сдвиг между напряжением и током при высоких частотах может иметь величину, близкую к 90° .

Допустим, что пролетное время вместе со временем, определяемым инерционностью лавинного процесса, равно половине периода колебаний некоторой частоты (рис. 3.55, а). В этом случае ток через диод будет отставать на полпериода от вызвавшего его напряжения. При этом нарастание напряжения все время будет сопровождаться уменьшением тока, а уменьшение напряжения, наоборот, — ростом тока. Это свидетельствует о том, что для данной частоты переменного напряжения в течение всего периода колебаний выполняется условие отрицательного дифференциального сопротивления.

При понижении частоты приложенного напряжения (при увеличении периода колебаний) и при той же величине времени пролета ток будет отставать от напряжения на угол, меньший 180° . Большую часть периода колебаний будут сохраняться условия отрицательного дифференциального сопротивления, меньшую часть периода эти условия выполняться не будут. Когда время пролета с понижением частоты приложенного напряжения будет составлять четверть периода, условия отрицательного дифференциального сопротивления будут выполняться на протяжении половины периода, чередуясь через каждые четверть периода с условиями положительного дифференциального сопротивления (рис. 3.55, б). В этом предельном случае в среднем за период не будет наблюдаться отрицательное дифференциальное сопротивление лавинно-пролетного диода.

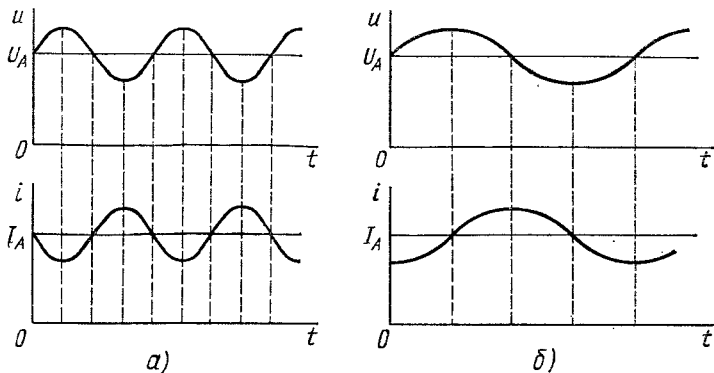


Рис. 3.55. Зависимости напряжений и токов, иллюстрирующие появление отрицательного дифференциального сопротивления лавинно-пролетного диода:

а — при сдвиге фаз 180° , возникающем на больших частотах переменной составляющей напряжения; *б* — при сдвиге фаз 90° , соответствующем меньшим частотам переменной составляющей напряжения и отсутствию отрицательного дифференциального сопротивления в среднем за период

Аналогично с повышением частоты пропадает отрицательное дифференциальное сопротивление при фазовом сдвиге между током и напряжением, достигающем 270° . Таким образом, отрицательное дифференциальное сопротивление лавинно-пролетных диодов проявляется только в диапазоне достаточно высоких частот и не наблюдается в статическом режиме.

Лавинно-пролетные диоды изготавливают на основе германия, кремния и арсенида галлия. На кремниевых лавинно-пролетных диодах миллиметрового диапазона (50 ГГц) в импульсном режиме получена мощность 0,35 Вт при к. п. д. 0,5 %. При несколько меньших частотах удается получать большие колебательные мощности (до нескольких сотен ватт при частоте в 1 ГГц) и большие к. п. д. (до 40 %).

Повышение к. п. д. лавинно-пролетных диодов связано с определенными трудностями. Объясняется это тем, что напряжение на диоде должно иметь величину, не намного превышающую критическое напряжение лавинного пробоя.

Кроме того, время пролета электронов через слой пространственного заряда, определяемое скоростью дрейфа и толщиной этого слоя, должно оставаться постоянным при изменении напряжения на диоде. Скорость дрейфа возрастает с напряженностью электрического поля лишь до определенной величины. Начиная с некоторого значения напряженности поля, наступает насыщение скорости дрейфа. Таким образом, допустимо только такое снижение напряжения на диоде, при котором скорость дрейфа электронов оставалась бы равной скорости насыщения. Это ограничивает допустимую амплитуду колебаний относительно небольшой долей постоянного напряжения.

§ 3.28. ТУННЕЛЬНЫЕ ДИОДЫ

Структура и принцип действия

Туннельный диод — это полупроводниковый диод на основе вырожденного полупроводника, в котором туннельный эффект приводит к появлению на вольт-амперной характеристике при прямом напряжении участка отрицательной дифференциальной проводимости.

В отличие от всех остальных полупроводниковых диодов для изготовления туннельных диодов используют полупроводниковый материал с очень большой концентрацией примесей (10^{18} — 10^{20} см⁻³). Следствием большой концентрации примесей в прилегающих к электронно-дырочному переходу областях является, во-первых, малая толщина перехода (около 10^{-2} мкм), т. е. на два порядка меньше, чем в других полупроводниковых диодах. Сквозь такие тонкие потенциальные барьеры возможно туннелирование носителей заряда (см. § 1.10).

Другим следствием большой концентрации примесей является расщепление примесных энергетических уровней с образованием примесных энергетических зон, которые примыкают к зоне проводимости в *n*-области и к валентной зоне в *p*-области. Уровни Ферми при этом оказываются расположенными в разрешенных зонах (рис. 3.56).

В диоде без внешнего напряжения существует туннелирование электронов из *n*-области в *p*-область и обратно. Встречные потоки электронов равны, поэтому суммарный ток через диод равен нулю (рис. 3.56, а).

При небольшом прямом напряжении на туннельном диоде происходит уменьшение высоты потенциального барьера электронно-дырочного перехода или смещение энергетической диаграммы *n*-области относительно энергетической диаграммы *p*-области. Свободные энергетические уровни *p*-области (занятые дырками), расположенные непосредственно над уровнем Ферми, оказываются на одной высоте по энергетической диаграмме или при одних и тех же значениях с энергетическими уровнями *n*-области, занятыми электронами (рис. 3.56, б). Поэтому будет происходить преимущественное туннелирование электронов из *n*-области в *p*-область.

При прямом напряжении на диоде, когда свободные энергетические уровни валентной и примесной зон *p*-области окажутся на одной высоте с занятыми электронами энергетическими уровнями зоны проводимости и примесной зоны *n*-области, туннельный ток через диод будет максимальным (рис. 3.56, в).

При дальнейшем увеличении прямого напряжения на диоде туннельный ток через диод будет уменьшаться, так как из-за смещения энергетических диаграмм будет уменьшаться количество электронов, способных туннелировать из *n*-области в *p*-область (рис. 3.56, г).

Туннельный ток через диод окажется равным нулю при некотором еще большем прямом напряжении, когда из-за относительного смещения энергетических диаграмм *n*- и *p*-областей для сво-

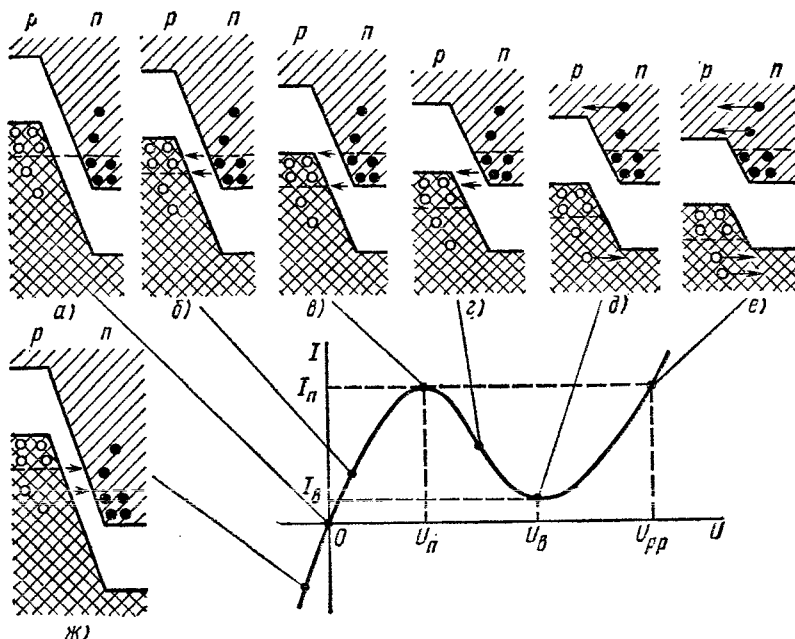


Рис. 3.56. Вольт-амперная характеристика и энергетические диаграммы туннельного диода при:

а — отсутствии напряжения; б — небольшом прямом напряжении; в — напряжении пика; г — напряжении, соответствующем отрицательному дифференциальному сопротивлению; д — напряжении впадины; е — напряжении, вызывающем значительный инжекционный ток; ж — обратном напряжении

бодных электронов n -области не будет свободных энергетических уровней в p -области (рис. 3.56, д). Однако при этом через диод будет проходить прямой ток, обусловленный переходом носителей через понизившийся потенциальный барьер электронно-дырочного перехода, т. е. ток, связанный с инжекцией.

С дальнейшим увеличением прямого напряжения в связи с уменьшением высоты потенциального барьера прямой ток через туннельный диод будет возрастать, как и в обычных выпрямительных диодах (рис. 3.56, е).

При обратном напряжении на туннельном диоде опять возникают условия для туннелирования электронов (рис. 3.56, ж). Только теперь электроны туннелируют из валентной зоны p -области в зону проводимости n -области. Возникающий при этом обратный ток будет расти с увеличением обратного напряжения по абсолютному значению. Таким образом, туннельный диод обладает относительно высокой проводимостью при обратном напряжении. Можно считать, что у туннельного диода при ничтожно малых обратных напряжениях происходит туннельный пробой.

Таким образом, туннельный диод обладает отрицательным дифференциальным сопротивлением в некотором диапазоне прямых

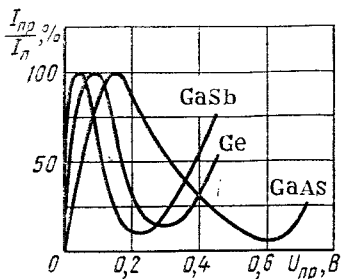


Рис. 3.57. Вольт-амперные характеристики туннельных диодов на основе различных полупроводниковых материалов

напряжений. Это и является самым интересным свойством туннельного диода, так как всякий прибор с отрицательным дифференциальным сопротивлением может быть использован для генерации и усиления электромагнитных колебаний, а также в переключающих схемах.

Параметры

Промышленность выпускает туннельные диоды из арсенида галлия и из германия. Их вольт-амперные характеристики показаны на рис. 3.57.

Из рисунка видно, что чем больше ширина запрещенной зоны исходного полупроводника, тем при больших напряжениях наблюдается отрицательное дифференциальное сопротивление.

Туннельные диоды характеризуются специфическими параметрами (см. рис. 3.56):

1. *Пиковый ток* I_{Π} — прямой ток в точке максимума вольт-амперной характеристики, при котором значение di/du равно нулю. Этот ток различен для туннельных диодов разного назначения. Величина его может быть от десятых долей миллиампера до сотен миллиампер.

2. *Ток впадины* $I_{В}$ — прямой ток в точке минимума вольт-амперной характеристики, при котором значение di/du равно нулю.

3. *Отношение токов туннельного диода* $I_{\Pi}/I_{В}$ — отношение пикового тока к току впадины. Для туннельных диодов из арсенида галлия $I_{\Pi}/I_{В} \geq 10$, для германиевых туннельных диодов $I_{\Pi}/I_{В} = 3 \div 6$.

4. *Напряжение пика* U_{Π} — прямое напряжение, соответствующее пиковому току. Для туннельных диодов из арсенида галлия $U_{\Pi} = 100 - 150$ мВ, для германиевых $U_{\Pi} = 40 - 60$ мВ.

5. *Напряжение впадины* $U_{В}$ — прямое напряжение, соответствующее току впадины. У туннельных диодов из арсенида галлия $U_{В} = 400 - 500$ мВ, у германиевых $U_{В} = 250 - 350$ мВ.

6. *Напряжение раствора* $U_{рр}$ — прямое напряжение, большее напряжение впадины, при котором ток равен пиковому.

7. *Удельная емкость туннельного диода* C_d/I_{Π} — отношение емкости туннельного диода к пиковому току.

8. *Предельная резистивная частота* f_R — расчетная частота, на которой активная составляющая полного сопротивления последовательной цепи, состоящей из $p-n$ -перехода и сопротивления потерь, обращается в нуль.

9. *Резонансная частота туннельного диода* f_0 — расчетная частота, на которой общее реактивное сопротивление $p-n$ -перехода и индуктивности корпуса туннельного диода обращается в нуль.

Температурные зависимости параметров

Изменение температуры туннельного диода может по-разному влиять на туннельную составляющую тока i на составляющую, связанную с инжекцией.

На температурную зависимость туннельной составляющей тока могут влиять следующие физические факторы.

1. С повышением температуры уменьшается ширина запрещенной зоны арсенида галлия и германия — основных исходных полупроводниковых материалов для туннельных диодов. Уменьшение ширины запрещенной зоны приводит к уменьшению толщины потенциального барьера, сквозь который туннелируют электроны, при этом вероятность туннелирования растет. Туннельная составляющая тока i , в частности, пиковый ток увеличиваются.

2. При увеличении температуры изменяется распределение электронов по энергетическим уровням — количество электронов под уровнем Ферми в зоне проводимости n -области уменьшается, так как часть свободных электронов переходит на более высокие энергетические уровни, а уровень Ферми смещается вниз. Поэтому уменьшается число электронов, которые могут туннелировать из n -области в p -область. Туннельная составляющая прямого тока уменьшается.

Так как эти факторы действуют, так сказать, в разные стороны, то суммарное их влияние, во-первых, должно быть малым, а во-вторых, оно может привести как к увеличению, так и к уменьшению пикового тока туннельного диода с увеличением температуры.

Инжекционная составляющая тока туннельного диода растет с увеличением температуры по двум причинам, имеющим место и в выпрямительных диодах (см. § 3.2), — уменьшение высоты потенциального барьера и перераспределение носителей заряда по энергетическим уровням. Поэтому у туннельного диода ток впадины растет с увеличением температуры.

Частотные свойства туннельных диодов

Механизм действия туннельных диодов связан с туннелированием электронов сквозь потенциальный барьер. Время, необходимое для завершения этого процесса, составляет 10^{-13} — 10^{-14} с. Эффекта накопления неосновных носителей в базе туннельных диодов практически нет, так как они используются при малых напряжениях, соответствующих падающему участку вольт-амперной характеристики (с отрицательным дифференциальным сопротивлением). Поэтому туннельные диоды способны работать на частотах до сотен гигагерц, что соответствует миллиметровому диапазону радиоволн. Верхний предел частотного диапазона работы туннельных диодов ограничен лишь паразитными реактивностями — собственной емкостью, основой которой составляет барьерная емкость электронно-дырочного перехода, и индуктивностью выводов и корпуса.

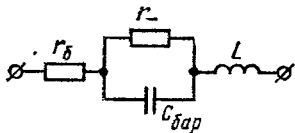


Рис. 3.58. Эквивалентная схема туннельного диода

Для анализа и расчета параметров, характеризующих частотные свойства туннельных диодов, воспользуемся эквивалентной схемой туннельного диода для малого переменного сигнала при наличии постоянного напряжения, которое смещает рабочую точку на падающий участок вольт-амперной характеристики.

Эквивалентная схема туннельного диода отличается от эквивалентной схемы обычного диода только тем, что здесь вместо активного сопротивления перехода введено отрицательное дифференциальное сопротивление r_- и учитывается индуктивность выводов L (рис. 3.58).

Полное сопротивление туннельного диода при малом синусоидальном напряжении

$$\dot{z} = r_б + j\omega L + \frac{1}{g + j\omega C_{бap}} = r_б + \frac{g}{g^2 + \omega^2 C_{бap}^2} + j\left(\omega L - \frac{\omega C_{бap}}{g^2 + \omega^2 C_{бap}^2}\right), \quad (3.104)$$

где $g = 1/r_-$ — отрицательная дифференциальная проводимость туннельного диода.

Отрицательное дифференциальное сопротивление у туннельного диода будет существовать, если вещественная часть полного сопротивления будет меньше нуля, т. е. при

$$r_б < \left| \frac{g}{g^2 + \omega^2 C_{бap}^2} \right|. \quad (3.105)$$

При большой частоте второе слагаемое в вещественной части (3.104) окажется меньше (по абсолютному значению) первого слагаемого. Это значит, что туннельный диод при больших частотах не будет обладать отрицательным дифференциальным сопротивлением. Определим *предельную резистивную частоту*, на которой активная составляющая полного сопротивления туннельного диода обращается в нуль, т. е. неравенство (3.105) обращается в равенство:

$$r_б = \frac{g}{g^2 + \omega_R^2 C_{бap}^2}.$$

Отсюда

$$\omega_R = \frac{1}{|r_-| C_{бap}} \sqrt{\frac{|r_-|}{r_б} - 1}. \quad (3.106)$$

Из соотношения (3.106) видно, что предельная резистивная частота туннельного диода зависит от соотношения сопротивления базы диода и абсолютного значения отрицательного дифференциального сопротивления. Для нахождения экстремальных значений предельной резистивной частоты надо продифференцировать (3.106) по $|r_-|$ и приравнять полученное выражение нулю. Тогда получаем условие

$$|r_-| = 2r_б, \quad (3.107)$$

при выполнении которого предельная резистивная частота будет максимальной. Связь ее с параметрами эквивалентной схемы можно получить, подставив (3.107) в (3.106). Тогда

$$f_R = \frac{1}{2\pi |r_-| C_{бар}}. \quad (3.108)$$

Усредненное отрицательное дифференциальное сопротивление можно представить в виде:

$$|r_-| = \frac{\Delta U}{\Delta I} = \frac{U_B - U_{\Pi}}{I_{\Pi} - I_B} \approx \frac{\text{const}}{I_{\Pi}}.$$

Если полученное соотношение подставить в (3.108), то

$$f_R \sim I_{\Pi} / C_{бар}.$$

Поэтому отношение емкости туннельного диода к пиковому току является одним из основных параметров туннельного диода, который характеризует его частотные свойства.

В туннельном диоде, имеющем паразитную индуктивность и емкость, при некоторой частоте могут возникнуть условия для резонанса, что нарушит нормальную работу схемы с туннельным диодом. Поэтому резонансная частота туннельного диода f_0 является еще одним параметром, характеризующим частотные свойства туннельного диода.

Резонансную частоту туннельного диода можно определить из соотношения (3.104) при условии равенства нулю мнимой части полного сопротивления диода:

$$\omega_0 L = \frac{\omega_0 C_{бар}}{g^2 + \omega_0^2 C_{бар}^2},$$

тогда

$$\omega_0 = \frac{1}{|r_-| C_{бар}} \sqrt{\frac{|r_-|^2 C_{бар}}{L} - 1}. \quad (3.109)$$

Туннельные диоды должны быть сконструированы так, чтобы резонансная частота была выше предельной резистивной частоты ($\omega_0 > \omega_R$), тогда возможные нежелательные резонансные явления возникают только на частотах, на которых туннельный диод уже не будет обладать отрицательным дифференциальным сопротивлением. Неравенство $\omega_0 > \omega_R$ с учетом соотношений (3.109) и (3.106) преобразуем следующим образом: $L < r_0 |r_-| C_{бар}$.

Следовательно, индуктивность эквивалентной схемы туннельного диода, которая определяется в основном индуктивностью внутренних и внешних выводов диода, должна быть по возможности малой. Поэтому внутренние и внешние выводы у туннельных диодов делают не из тонкой проволоки, а из плющенко или в виде мембран, пластин и т. п.

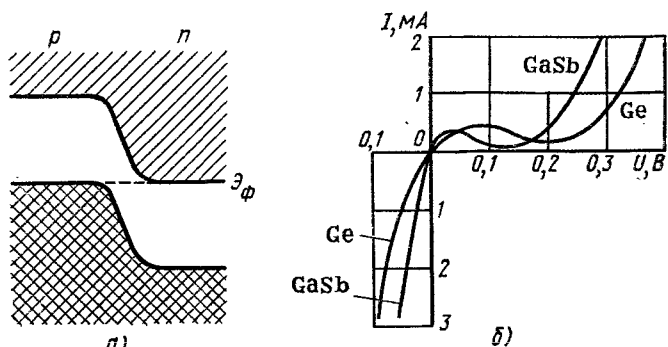


Рис. 3.59. Энергетическая диаграмма (а) и вольт-амперные характеристики (б) обращенных диодов

§ 3.29. ОБРАЩЕННЫЕ ДИОДЫ

Обращенным называют диод на основе полупроводника с критической концентрацией примесей, в котором проводимость при обратном напряжении вследствие туннельного эффекта значительно больше, чем при прямом напряжении.

При концентрациях примесей в p - и n -областях диода, меньших, чем в туннельных диодах, но больших, чем в обычных выпрямительных диодах, можно получить диод, энергетическая диаграмма которого показана на рис. 3.59, а. Уровень Ферми при такой средней концентрации примесей может быть расположен на потолке валентной зоны p -области и на дне зоны проводимости n -области диода, т. е. потолок валентной зоны p -области и дно зоны проводимости n -области при нулевом смещении на диоде находятся на одной высоте по энергетической диаграмме.

Обратная ветвь вольт-амперной характеристики обращенного диода аналогична обратной ветви вольт-амперной характеристики туннельного диода, так как при обратных напряжениях происходит туннелирование электронов из валентной зоны p -области в зону проводимости n -области. Поэтому обратные токи в обращенных диодах оказываются большими при ничтожно малых обратных напряжениях (десятки милливольт).

Прямая ветвь вольт-амперной характеристики обращенного диода аналогична прямой ветви вольт-амперной характеристики обычного выпрямительного диода, так как при прямых напряжениях на обращенном диоде прямой ток может быть образован только в результате инжекции носителей заряда через потенциальный барьер электронно-дырочного перехода. Но заметная инжекция может наблюдаться только при прямых напряжениях в несколько десятых долей вольта. При меньших напряжениях прямые токи в обращенных диодах оказываются меньше обратных (рис. 3.59, б).

Таким образом, обращенные диоды обладают выпрямляющим эффектом, но пропускное (проводящее) направление у них соответ-

ствует обратному включению, а запирающее (непроводящее) — прямому включению.

Из принципа действия обращенных диодов ясно, что они, во-первых, способны работать на очень малых сигналах. Во-вторых, должны обладать хорошими частотными свойствами, так как туннелирование — процесс малоинерционный, а эффекта накопления неосновных носителей при малых прямых напряжениях практически нет. Поэтому обращенные диоды можно использовать на СВЧ. В-третьих, из-за относительно большой концентрации примесей в прилегающих к электронно-дырочному переходу областях обращенные диоды оказываются мало чувствительными к воздействиям проникающей радиации.

§ 3.30. ВАРИКАПЫ

Структура и принцип действия

Варикап — это полупроводниковый диод, в котором используется зависимость емкости от величины обратного напряжения и который предназначен для применения в качестве элемента с электрически управляемой емкостью.

Как было рассмотрено в § 2.8 и 3.6, диод обладает барьерной и диффузионной емкостями. В качестве варикапов используют только диоды при обратном постоянном смещении, когда проявляется только барьерная емкость. Диффузионная емкость проявляется при прямом смещении диода, когда проводимость его велика и велики потери мощности из-за относительно больших активных токов через диод.

Зависимость емкости от напряжения смещения различна для варикапов, изготовленных методом диффузии или методом сплавления примесей (см. § 2.8). В сплавных варикапах с резким электронно-дырочным переходом зависимость барьерной емкости от напряжения смещения получается более резкая. Связано это с тем, что глубина проникновения электрического поля или его приращения в материал зависит от удельного сопротивления этого материала. В сплавном варикапе слои базы, прилегающие к переходу, легированы равномерно, а в диффузионном — при удалении от перехода концентрация нескомпенсированных примесей увеличивается, т. е. уменьшается удельное сопротивление.

Поэтому для получения еще более резкой зависимости емкости варикапа

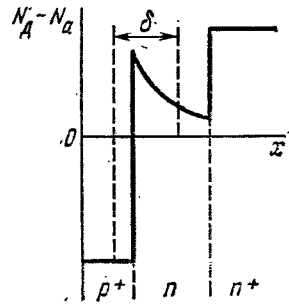


Рис. 3.60. Распределение концентрации нескомпенсированных примесей в структуре варикапа с резкой зависимостью емкости от напряжения

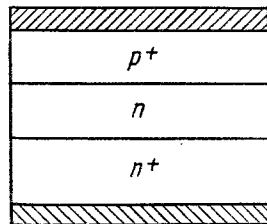


Рис. 3.61. Структура варикапа с малым сопротивлением базы

от напряжения смещения необходимо создавать в базе варикапа аномальное распределение нескомпенсированных примесей с градиентом концентрации другого знака по сравнению со знаком градиента концентрации в базе диффузионного диода (рис. 3.60).

Так же, как в других диодах, сопротивление базы варикапа должно быть по возможности меньшим. Одновременно для большего пробивного напряжения необходимо большее удельное сопротивление слоев базы, прилегающих к электронно-дырочному переходу. Таким образом, база варикапа должна состоять из двух слоев (рис. 3.60 и 3.61). Основная часть базы — подложка должна быть низкоомной. Тонкий слой базы, прилегающий к переходу, должен быть высокоомным.

Основные параметры

1. *Емкость варикапа* C_v — емкость, измеренная между выводами варикапа при заданном обратном напряжении. Для различных варикапов емкость может быть от нескольких единиц до нескольких сотен пикофарад.

2. *Коэффициент перекрытия по емкости* K_C — отношение емкостей варикапа при двух заданных значениях обратных напряжений. Значение этого параметра составляет обычно несколько единиц.

3. *Добротность варикапа* Q_v — отношение реактивного сопротивления варикапа на заданной частоте переменного сигнала к сопротивлению потерь при заданном значении емкости или обратного напряжения. Добротность — это величина, обратная тангенсу угла диэлектрических потерь. Добротность варикапов измеряют обычно при тех же напряжениях смещения, что и емкость. Значение добротности — от нескольких десятков до нескольких сотен.

Частотные свойства

Изменение барьерной емкости при изменении напряжения на электронно-дырочном переходе, а также заряд и разряд этой емкости под действием изменяющегося напряжения обусловлен смещением основных носителей заряда в прилегающих к переходу областях. Постоянная времени этого процесса — время релаксации $\tau = \epsilon \epsilon_0 \rho$.

Приняв $\rho = 10^{-1}$ Ом·см и $\epsilon = 16$ (для германия), получаем $\tau = 1,4 \cdot 10^{-13}$ с. Значит, инерционность процессов заряда и разряда барьерной емкости варикапа под действием изменяющегося напряжения может сказываться только на очень высоких частотах. Однако в варикапе (как и в любом другом полупроводниковом диоде) есть объемное сопротивление базы $r_б$ и активное сопротивление электронно-дырочного перехода $r_{пер}$ (рис. 3.62), которые могут привести к снижению добротности варикапа при высоких или низких частотах.

При низких частотах в эквивалентной схеме варикапа можно не учитывать малое сопротивление базы $r_б$ по сравнению с большим

емкостным сопротивлением барьерной емкости и большим активным сопротивлением перехода. Таким образом, при низких частотах упрощенная эквивалентная схема варикапа представляет собой параллельное соединение барьерной емкости и сопротивления перехода. Для параллельной схемы замещения тангенс угла диэлектрических потерь

$$\operatorname{tg} \delta = \frac{1}{\omega r_{\text{пер}} C_{\text{бар}}},$$

добротность

$$Q_{\text{пч}} = \omega r_{\text{пер}} C_{\text{бар}} = 2\pi f r_{\text{пер}} C_{\text{бар}}.$$

Следовательно, добротность варикапа при низких частотах растет с увеличением частоты, так как изменяется соотношение между реактивной и активной проводимостями варикапа (рис. 3.63).

Таким образом, варикапы, применяемые на очень низких частотах, должны иметь высокие значения произведения дифференциального сопротивления электронно-дырочного перехода на величину барьерной емкости $r_{\text{пер}} C_{\text{бар}}$. Величина этого произведения не зависит от площади перехода, так как барьерная емкость пропорциональна, а дифференциальное сопротивление обратно пропорционально площади электронно-дырочного перехода. Ясно, что величина дифференциального сопротивления электронно-дырочного перехода будет больше, если значения плотности постоянных обратных токов будут меньше. Поэтому для изготовления низкочастотных варикапов целесообразно использовать полупроводниковые материалы с большой шириной запрещенной зоны (кремний, арсенид галлия и т. д.).

При высоких частотах в эквивалентной схеме варикапа можно не учитывать большое активное сопротивление перехода по сравнению с малым (при высокой частоте) параллельно включенным емкостным сопротивлением барьерной емкости. Но при этом нельзя пренебрегать сопротивлением базы, которое может оказаться сравнимым с емкостным сопротивлением барьерной емкости. Таким образом, при высоких частотах упрощенная эквивалентная схема варикапа представляет собой последовательное соединение барьерной емкости и сопротивления базы. Для последовательной схемы замещения тангенс угла диэлектрических потерь

$$\operatorname{tg} \delta = \omega r_{\delta} C_{\text{бар}},$$

добротность

$$Q_{\text{пч}} = \frac{1}{\omega r_{\delta} C_{\text{бар}}} = \frac{1}{2\pi f r_{\delta} C_{\text{бар}}}.$$

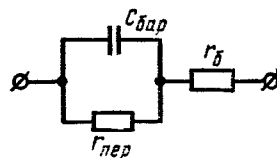


Рис. 3.62. Эквивалентная схема варикапа для широкого диапазона частот

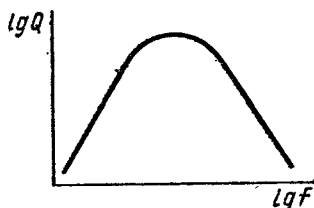


Рис. 3.63. Зависимость добротности варикапа от частоты

Следовательно, добротность варикапа при высоких частотах уменьшается с увеличением частоты, так как уменьшается отношение реактивного сопротивления варикапа к сопротивлению потерь (см. рис. 3.63).

Варикапы в основном применяют на высоких и сверхвысоких частотах. Поэтому определение одного из основных параметров варикапа — добротности соответствует именно последовательной схеме замещения.

Таким образом, варикапы, применяемые на высоких и сверхвысоких частотах, должны иметь малые значения дифференциального сопротивления базы r_6 . В рабочем диапазоне постоянных смещений варикапа, т. е. при обратных напряжениях сопротивление базы можно считать неизменным. Тогда дифференциальное сопротивление базы, равное статическому сопротивлению базы, для плоскостного варикапа можно определить по формуле

$$r_6 = \rho W/S, \quad (3.110)$$

где W — толщина базы варикапа; S — площадь электронно-дырочного перехода.

Из формулы (3.110) следует, что уменьшить дифференциальное или статическое сопротивление базы варикапа можно путем выбора исходного полупроводникового материала с большой подвижностью носителей заряда (арсенид галлия, германий с электропроводностью n -типа и т. д.).

Уменьшать сопротивление базы варикапа путем увеличения концентрации носителей заряда нецелесообразно, так как при увеличении концентрации носителей заряда будет уменьшаться подвижность носителей заряда (см. § 1.9) и пробивное напряжение варикапа (см. § 3.11).

Для уменьшения сопротивления базы без уменьшения пробивного напряжения базу варикапа делают двухслойной (см. рис. 3.61). При этом низкоомный слой базы представляет собой подложку относительно большей толщины, обеспечивающей необходимую механическую прочность структуре прибора. Высокоомный слой базы с тем же типом электропроводности имеет малую толщину (несколько микрометров). Он может быть создан методом эпитаксиального наращивания.

§ 3.31. НАДЕЖНОСТЬ ДИОДОВ

Общие понятия надежности

Надежность полупроводниковых диодов обычно значительно выше надежности других элементов радиоэлектронной аппаратуры. Однако в связи с усложнением схем требуется дальнейшее повышение надежности полупроводниковых диодов.

Количественно надежность связана с числом отказов, т. е. с числом нарушений работоспособности приборов в течение определенного промежутка времени. По характеру изменения параметров

отказы полупроводниковых диодов, как и других приборов, могут быть двух видов:

1) *катастрофические*, или *внезапные*, возникающие в результате скачкообразного изменения одного или нескольких параметров прибора;

2) *условные*, или *постепенные*, возникающие в результате плавного изменения основных параметров прибора. Условность отказа определяется выбором критериев годности. Условно отказавший прибор может быть фактически работоспособным в радиоэлектронных схемах или частично утратившим свою работоспособность.

Основные причины отказов

Катастрофические отказы полупроводниковых диодов могут быть обусловлены недостатками конструкции или технологии, а также неправильной эксплуатацией диодов.

Из-за различия коэффициентов теплового расширения соприкасающихся деталей может происходить механическое нарушение контакта или растрескивание кристалла полупроводника при изменении температуры диода. Эта причина катастрофического отказа относится к конструктивным недостаткам прибора.

В высокочастотных диодах размеры электродов составляют иногда от единиц до десятков микрометров. Внутренние выводы от таких электродов выполняются в виде очень тонкой проволоки, иногда в несколько микрометров. Такую проволоку трудно изготовить строго одного сечения по длине, что может обусловить перегорание внутреннего вывода даже при нормальных нагрузках по току. К перегоранию может привести и излишняя длина внутреннего вывода, так как по длине проволоки возникают значительные перепады температур.

Наиболее частой причиной катастрофических отказов является неправильная эксплуатация диодов. Так, например, даже очень кратковременные импульсы токов и напряжений, превышающие допустимые значения, могут привести к необратимому пробоем $p-n$ -перехода диода в связи с тем, что пробой часто происходит по неоднородностям в $p-n$ -переходе. При этом плотности тока в отдельных частях $p-n$ -перехода могут оказаться недопустимо большими, произойдет шнурование тока с резким увеличением удельной мощности, выделяющейся в шнуре. Подобные эффекты возможны в полупроводниковых диодах даже при прохождении коротких импульсов тока с амплитудой, не превышающей допустимого значения постоянного тока.

К неправильной эксплуатации полупроводниковых диодов следует отнести размещение их вблизи нагреваемых элементов схемы, что способствует перегреву диода, а также крепление диода способом, ухудшающим условия теплообмена с окружающей средой.

Условные отказы в полупроводниковых диодах, связанные с медленным изменением их параметров, чаще всего вызваны физи-

ческими и химическими процессами на поверхности и в объеме полупроводникового кристалла, сплавов и припоев.

В § 3.14 было рассмотрено влияние поверхностных состояний на основные параметры полупроводниковых диодов. При изменении поверхностных состояний могут существенно изменяться обратные токи и пробивные напряжения диодов. Для стабилизации поверхностных состояний кристалл полупроводника с *p-n*-переходом помещают внутрь герметичного корпуса.

Одним из основных факторов, определяющих нестабильность поверхности полупроводника, является влага. Какими бы тщательными ни были герметизация прибора и предварительный отжиг деталей, все же некоторое количество влаги попадает в прибор и влияет на стабильность его характеристик. Для поглощения остаточной влаги в корпус прибора часто помещают влагопоглощающее вещество — сорбент (силикагель, алюмогель или цеолит). Однако влагопоглотитель способен выполнить свои функции только при хорошей герметичности корпуса. При наличии же микротрещин в проходных изоляторах через сотни и более часов работы диода внутрь корпуса наберется такое количество влаги, которое вызовет нестабильность электрических характеристик диода.

Микротрещины в проходных изоляторах могут возникать из-за неудачного выбора материалов для спая стекла с коваром, неправильной конструкции изолятора, вследствие чего стекло работает не на сжатие, а на растяжение, из-за нарушения необходимых допусков либо из-за небрежного обращения с прибором (изгиб внешних выводов вблизи проходных стеклянных изоляторов).

Кроме того, причиной появления условных отказов являются физико-химические процессы в сплавах и припоях, используемых для создания *p-n*-переходов и невыпрямляющих контактов. Иногда кристаллизация этих сплавов при изготовлении прибора идет неравномерно и образуется не предусмотренная равновесной диаграммой состояния эвтектика с очень низкой температурой плавления, с ускоренными процессами диффузии различных элементов в сплаве, что приводит к старению сплавов и ухудшению качества *p-n*-переходов и невыпрямляющих контактов.

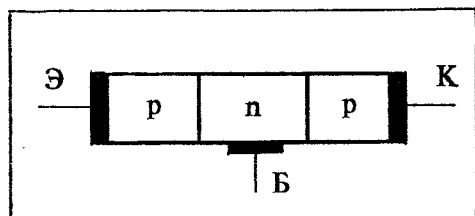
Контрольные вопросы

1. Почему по распределению неосновных носителей заряда в базе диода можно судить о величине плотности тока через диод?
2. Чем отличаются вольт-амперные характеристики диодов с толстой и тонкой базами?
3. Что такое диффузионная емкость диода?
4. Как связано время пролета неосновных носителей заряда через базу диода с толщиной базы и коэффициентом диффузии неосновных носителей?
5. В каких случаях следует учитывать процесс генерации носителей в *p-n*-переходе при исследовании вольт-амперных характеристик диодов?
6. Как влияет процесс рекомбинации носителей в *p-n*-переходе диода на его вольт-амперную характеристику?
7. Как связан коэффициент лавинного умножения с коэффициентом ударной ионизации?

8. Как зависит пробивное напряжение диодов при лавинном пробое от концентрации примесей в базе и от удельного сопротивления базы?
9. Как изменяется пробивное напряжение диода при лавинном и туннельном пробоях с увеличением температуры?
10. Каковы особенности теплового пробоя в реальных диодах?
11. Как могут влиять поверхностные состояния на вольт-амперные характеристики диодов?
12. Какие явления надо учитывать при работе диода на высоком уровне инжекции?
13. Как влияют процессы накопления и рассасывания неосновных носителей в базе диода, а также барьерная емкость на работу диода при быстром изменении напряжения или тока?
14. В чем состоят основные отличия свойств и параметров кремниевых и германиевых выпрямительных диодов? Как объяснить эти отличия?
15. Что такое сопротивление растекания и как его можно рассчитать?
16. Каков принцип действия стабилитронов и стабисторов?
17. Почему лавинно-пролетные диоды обладают отрицательным дифференциальным сопротивлением только на определенных частотах СВЧ-диапазона?
18. Каков принцип действия туннельных диодов?
19. Какие свойства обращенных диодов наиболее полезны?
20. Почему варикапы должны работать только при приложении к ним обратного постоянного смещения?
21. Перечислите основные причины катастрофических и условных отказов полупроводниковых диодов.

БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

4



§ 4.1. СТРУКТУРА И ОСНОВНЫЕ РЕЖИМЫ РАБОТЫ

Биполярный транзистор (обычно его называют просто транзистором) представляет собой полупроводниковый прибор, состоящий из трех областей с чередующимися типами электропроводности, пригодный для усиления мощности (рис. 4.1).

Эти области разделяются электронно-дырочными переходами. Особенность транзистора состоит в том, что между его электронно-дырочными переходами существует взаимодействие — ток одного из переходов может управлять током другого. Такое управление возможно, потому что носители заряда, инжектированные через один из электронно-дырочных переходов, могут дойти до другого перехода, находящегося под обратным напряжением, и изменить его ток.

Каждый из переходов транзистора можно включить либо в прямом, либо в обратном направлении. В зависимости от этого различают три режима работы транзистора:

1) *режим отсечки* — оба электронно-дырочных перехода закрыты, при этом через транзистор обычно идет сравнительно небольшой ток;

2) *режим насыщения* — оба электронно-дырочных перехода открыты;

3) *активный режим* — один из электронно-дырочных переходов открыт, а другой закрыт.

В режиме отсечки и режиме насыщения управление транзистором почти отсутствует. В активном режиме также управление осуществляется наиболее

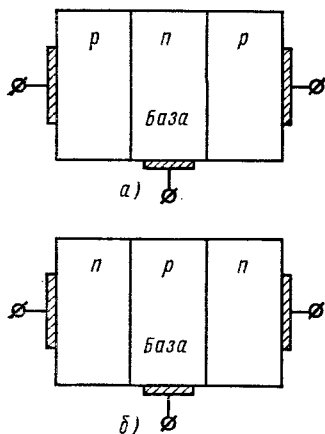


Рис. 4.1. Структуры биполярных транзисторов:

а — p-n-p; б — n-p-n

эффективно, причем транзистор может выполнять функции активного элемента электрической схемы (усиление, генерирование, переключение и т. п.).

Область транзистора, расположенную между электронно-дырочными переходами, называют базой. Примакивающие к базе области чаще всего делают неодинаковыми. Одну из областей изготавливают так, чтобы из нее наиболее эффективно происходила инжекция носителей в базу, а другую — так, чтобы соответствующий электронно-дырочный переход наилучшим образом осуществлял экстракцию инжектированных носителей из базы.

Область транзистора, основным назначением которой является инжекция носителей в базу, называют эмиттером, соответствующий электронно-дырочный переход — эмиттерным.

Область транзистора, основным назначением которой является экстракция носителей из базы, называют коллектором, соответствующий электронно-дырочный переход — коллекторным.

Если на эмиттерном переходе напряжение прямое, а на коллекторном обратное, то включение транзистора считают *нормальным*, при противоположной полярности напряжений — *инверсным*.

Часть базы, находящуюся между эмиттером и коллектором, через которую проходят носители заряда в активном режиме работы транзистора, называют активной частью, расположенную между эмиттером и выводом базы, называют пассивной, а ту часть, которая лежит за выводом базы, — периферической.

Основные характеристики транзистора определяются в первую очередь процессами, происходящими в базе. В зависимости от распределения примесей в базе может существовать или отсутствовать электрическое поле. Если при отсутствии токов в базе существует электрическое поле, которое способствует движению неосновных носителей заряда от эмиттера к коллектору, то транзистор назы-

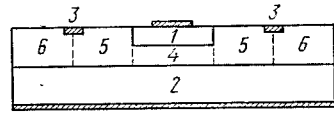


Рис. 4.2. Основные области транзистора:

1 — эмиттер; 2 — коллектор; 3 — невыпрямляющий контакт к базе; 4 — активная; 5 — пассивная; 6 — периферическая части базы

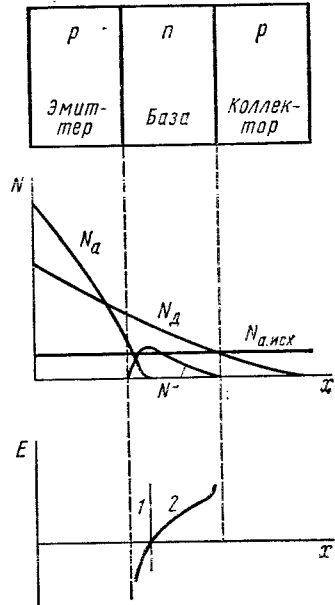


Рис. 4.3. Распределение примесей и электрического поля в базе транзистора, изготовленного методом диффузии:

1 — участок тормозящего поля; 2 — участок ускоряющего поля

(рис. 4.2). Часть базы, расположенную между эмиттером и выводом базы, называют пассивной, а ту часть, которая лежит за выводом базы, — периферической.

вают дрейфовым, если же поле в базе отсутствует — б е з дрейфовым.

Значение напряженности электрического поля в базе транзистора при отсутствии токов может быть определено из выражения для тока основных носителей. Например, для базы с электропроводностью n -типа

$$\mathbf{J}_n = qn_{n0}\mu_n\mathbf{E} + qD_n \text{grad } n_{n0} = 0; \quad (4.1)$$

отсюда

$$\mathbf{E} = -\frac{kT}{q} \cdot \frac{\text{grad } n_{n0}}{n_{n0}}. \quad (4.2)$$

Если учесть, что концентрация основных носителей заряда практически равна концентрации примесей, получим для полупроводника n -типа

$$\mathbf{E} = -\frac{kT}{q} \cdot \frac{\text{grad } N_d}{N_d}; \quad (4.3)$$

для полупроводника p -типа

$$\mathbf{E} = \frac{kT}{q} \cdot \frac{\text{grad } N_a}{N_a}. \quad (4.4)$$

Таким образом, возникновение электрического поля связано с наличием градиента концентрации примесей. Физически это объясняется тем, что градиент концентрации основных носителей заряда должен вызывать их перераспределение. При этом из области с большей концентрацией основные носители уходят, оставляя частично нескомпенсированные заряды ионов примеси, а в области с малой концентрацией примесей основные носители накапливаются. Получающееся электрическое поле всегда направлено так, что способствует движению неосновных носителей из области с большой концентрацией примесей в область с малой концентрацией. Следовательно, поле, способствующее движению неосновных носителей от эмиттера к коллектору, создается, если концентрация нескомпенсированных примесей в базе уменьшается по направлению от эмиттера к коллектору.

Такое распределение примесей в базе получается, например, при изготовлении транзистора методом диффузии (рис. 4.3). Однако в этом случае может возникнуть участок, где поле будет препятствовать движению носителей от эмиттера к коллектору (тормозящее поле), что обычно ухудшает свойства транзистора.

Различают три схемы включения транзистора — с общим эмиттером, общей базой и общим коллектором. Общим называют электрод, относительно которого измеряют и задают напряжения. Ток в цепи общего электрода не определяют. Усилительные свойства транзистора проявляются, если в схеме с общей базой в качестве входной цепи использовать эмиттерную, а в качестве выходной — коллекторную. В схеме с общим эмиттером входной является цепь базы, а выходной — цепь коллектора. Для схемы с общим коллектором входной является цепь базы, а выходной — цепь эмиттера.

Основные свойства транзистора определяются соотношениями токов и напряжений в различных его цепях и взаимным их влиянием друг на друга. Транзистор может работать на постоянном токе, малом переменном сигнале, большом переменном сигнале и импульсном сигнале.

Чтобы рассмотреть работу транзистора на постоянном токе, необходимо изучить стационарные потоки носителей в нем. Это дает возможность получить *статистические характеристики* и *статические параметры транзистора* — соотношения между его постоянными токами и напряжениями, выраженные графически или в виде численных значений.

На работу транзистора при малом переменном сигнале помимо стационарных потоков носителей заряда влияют также процессы накопления носителей, имеющиеся в транзисторе емкости, а также конечная скорость движения носителей. Свойства транзистора при работе на малом переменном сигнале описываются системами мало-сигнальных параметров.

При работе на большом сигнале и на импульсном сигнале помимо тех факторов, которые определяют работу транзистора на постоянном токе и малом переменном сигнале, существенно сказывается нелинейность статических характеристик транзистора.

§ 4.2. РАСПРЕДЕЛЕНИЕ СТАЦИОНАРНЫХ ПОТОКОВ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим распределение потоков носителей в транзисторе на примере структуры типа *p-n-p*.

Активный режим. Через эмиттерный переход транзистора, работающего в активном режиме, происходит инжекция носителей заряда в базу (рис. 4.4, *a*). Инжектированные носители (ток $I_{Эр}$) частично рекомбинируют в объеме базы и на ее поверхности, а некоторые из них могут доходить до контакта базы и рекомбинировать на нем (токи $I_{рек\ в}$, $I_{рек\ с}$, $I_{рек\ кон}$). Остальные инжектированные носители пересекают базу, доходят до коллекторного перехода и увеличивают его обратный ток.

К току носителей заряда, инжектированных эмиттером и дошедших до коллектора ($I_{Кр}$), добавляется ток носителей, образовавшихся в результате тепловой генерации в базе ($I_{Б, ген}$), в коллекторе ($I_{К, ген}$), а также в коллекторном переходе ($I_{ген}$). Кроме того, при достаточно больших напряжениях в коллекторном переходе происходит лавинное умножение носителей ($I_{л}$). Могут существовать и токи утечки по поверхности полупроводника. Все эти токи в сумме образуют ток коллектора.

Через эмиттер помимо тока носителей заряда, инжектируемых в базу ($I_{Эр}$), проходит ток носителей, инжектируемых из базы в эмиттер ($I_{Эн}$). В области эмиттера эти носители оказываются основными и рекомбинируют. Кроме того, через эмиттерный переход проходит ток, связанный с рекомбинацией носителей в области

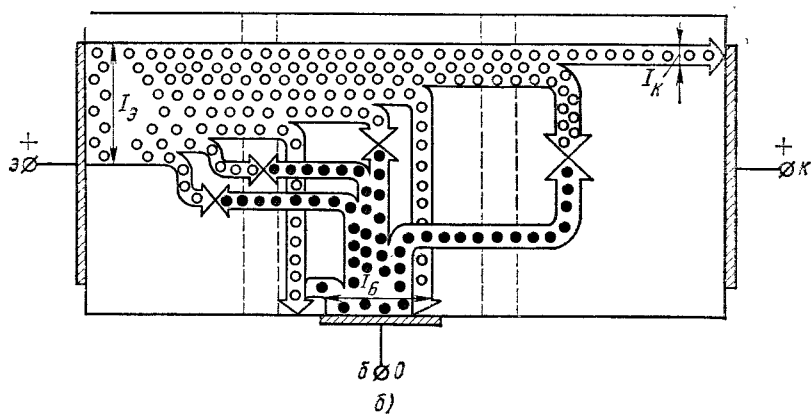
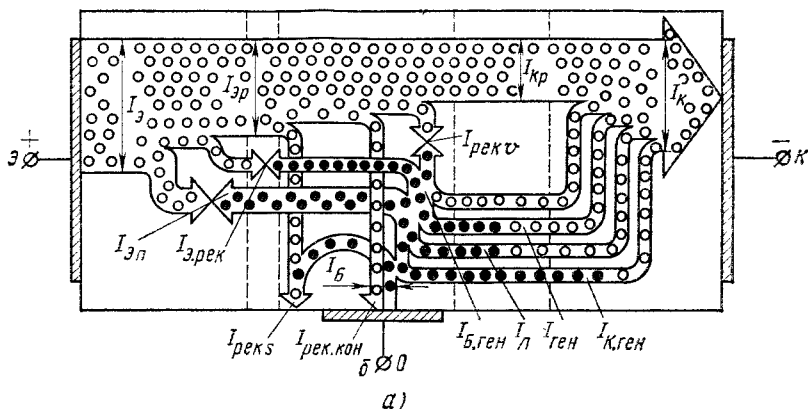


Рис. 4.4. Распределение стационарных потоков носителей в транзисторе:
 а — активный режим; б — режим насыщения

объемного заряда ($I_{Э, рек}$), а иногда (при малых общих токах) нужно учитывать и ток утечки.

Ток, проходящий к выводу базы, представляет собой алгебраическую сумму токов основных носителей, обуславливающих инжекцию носителей в эмиттер ($I_{Эп}$), рекомбинацию в эмиттерном переходе ($I_{Э, рек}$) и в базе ($I_{рек, в} + I_{рек, с}$), обратных токов коллекторного перехода ($I_{Б, ген} + I_{л} + I_{ген} + I_{К, ген}$), а также тока неосновных носителей заряда, дошедших до вывода базы ($I_{рек, конт}$). Величина и направление тока базы определяются соотношением этих составляющих.

Если рассматривать распределение токов в транзисторе с точки зрения взаимного влияния эмиттера и коллектора, то оказывается, что от тока эмиттера существенно зависит только составляющая $I_{Кр}$ тока коллектора, обусловленная инжектированными эмиттером в базу носителями, дошедшими до коллектора, а также составляющая тока, связанная с лавинным умножением в коллекторном

переходе (I_{Δ}). Все остальные составляющие тока коллектора либо совсем не зависят от тока эмиттера, либо эта зависимость слабая ($I_{K, \text{ген}}$).

Кроме того, на значение тока коллектора влияет не весь эмиттерный ток, а только его составляющая $I_{Эр}$, связанная с инъекцией неосновных носителей в базу. Составляющие тока эмиттера, связанные с инъекцией из базы в эмиттер ($I_{Эн}$) и рекомбинацией в области объемного заряда ($I_{Э, \text{рек}}$), хотя и зависят от напряжения на эмиттерном переходе, но на ток коллектора непосредственно не влияют.

Чтобы ток эмиттера эффективно управлял током коллектора, принимают меры для уменьшения всех составляющих тока, не принимающих участия в таком управлении. Прежде всего стараются сократить потери носителей заряда, инжектируемых в базу. Для этого:

1) толщину базы делают небольшой по сравнению с диффузионной длиной неосновных носителей заряда, что снижает потери на рекомбинацию в объеме материала ($I_{\text{рек } \nu}$);

2) поверхность полупроводника обрабатывают так, чтобы получить по возможности малое значение скорости поверхностной рекомбинации (уменьшить $I_{\text{рек } s}$);

3) вывод базы располагают настолько далеко от эмиттера, что до него дырки практически не доходят (уменьшение $I_{\text{рек, кол}}$);

4) размеры коллектора делают такими, чтобы он мог перехватить весь поток носителей, идущих от эмиттера, т. е. площадь коллекторного перехода должна быть значительно больше площади эмиттерного перехода.

В результате этих мер

$$I_{Эр} \approx I_{Kр}. \quad (4.5)$$

Для уменьшения составляющей тока носителей, инжектируемых из базы в эмиттер ($I_{Эн}$), концентрацию примесей в базе делают значительно меньшей, чем в эмиттере. Это приводит к тому, что ток через эмиттерный переход состоит практически из носителей заряда, инжектируемых в базу, т. е. коэффициент инжекции (отношение дырочного тока к полному току через переход) получается высоким.

Использование для увеличения тока коллектора явления ударной ионизации чаще всего нерационально, так как приводит к сильной зависимости тока I_{Δ} от напряжения на коллекторе, т. е. к нестабильной работе транзистора. Это явление используют только в лавинных транзисторах.

С учетом всех перечисленных мер можно считать, что ток коллектора транзистора в активном режиме примерно равен току эмиттера:

$$I_K \approx I_{Э}, \quad (4.6)$$

а ток базы значительно меньше как тока эмиттера, так и тока коллектора:

$$I_B = I_{Э} - I_K. \quad (4.7)$$

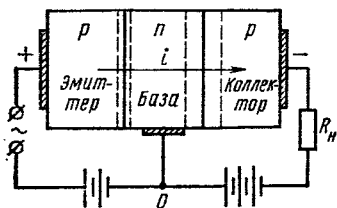


Рис. 4.5. Включение транзистора в схеме с общей базой

Рассмотрим, как происходит усиление в транзисторе при активном режиме его работы.

В схеме с общей базой (рис. 4.5) в выходной цепи (коллекторной) практически проходит тот же ток, что и во входной (эмиттерной), т. е. усиление по току в данном случае отсутствует. Однако эта схема дает возможность получить усиление по мощности.

Чтобы понять принцип усиления мощности в транзисторе, да и в других усилительных приборах, надо учесть взаимодействие носителей заряда с электрическим полем. Например, дырка, двигаясь по направлению электрического поля, разгоняется в этом поле и приобретает дополнительную энергию, забирая ее от электрического поля. Если же заставить дырку двигаться против электрического поля, то она будет тормозиться этим полем, отдавая ему часть своей энергии.

Электрическое поле в коллекторном переходе транзистора состоит из постоянной составляющей, созданной внешним источником питания в цепи коллектора, и переменной составляющей, возникающей при экстракции неосновных носителей из базы в коллекторный переход. Мгновенные значения переменной составляющей электрического поля в любой момент времени направлены в сторону, противоположную постоянной составляющей.

Поэтому дырка, проходя по коллекторному переходу, взаимодействует сразу с двумя составляющими электрического поля. От постоянной составляющей электрического поля дырка забирает энергию, двигаясь по направлению этой составляющей. Одновременно, двигаясь против мгновенных значений переменной составляющей электрического поля, дырка отдает часть своей энергии переменной составляющей.

Происходит своеобразное перекачивание энергии от постоянной составляющей электрического поля к переменной составляющей. Посредниками в этом перекачивании энергии являются носители заряда, инжектированные из эмиттера и дошедшие до коллекторного перехода. Для их инжекции требуется произвести относительно небольшую работу, так как высота потенциального барьера эмиттерного перехода мала.

В схеме с общим эмиттером входной цепью является цепь базы. Так как ток базы существенно меньше тока эмиттера, можно получить и усиление по току. Изменяя ток через вывод базы, меняем количество основных носителей в области базы, т. е. заряд базы, и, следовательно, потенциальный барьер между эмиттером и базой. Изменение высоты потенциального барьера вызывает соответствующую инжекцию неосновных носителей заряда. Большинство инжектированных носителей доходит до коллекторного перехода, изменяя его ток. Основным носителем заряда, введенный в базу из вывода базы,

либо может исчезнуть вследствие рекомбинации, либо может быть инжектирован в эмиттер. Как указывалось, в транзисторе приняты меры, чтобы вероятность этого была мала, и на один основной носитель заряда, вошедший в базу, приходится много неосновных носителей заряда, прошедших от эмиттера до коллектора. В этом и заключается усиление по току в схеме с общим эмиттером. Усиление по мощности в данном случае объясняется аналогично усилению в схеме с общей базой.

В схеме с общим коллектором выходной цепью является эмиттерная, входной — цепь базы. В связи с тем, что ток эмиттера почти равен току коллектора, здесь тоже имеет место усиление по току и по мощности.

Режим насыщения. При работе транзистора в режиме насыщения (см. рис. 4.4, б) в прямом направлении включен не только эмиттерный, но и коллекторный переход. Это приводит к тому, что не все носители, инжектированные эмиттером и дошедшие до коллекторного перехода, перехватываются им. Условно можно считать, что навстречу потоку неосновных носителей, идущих из базы в коллектор, идет поток таких же носителей из коллектора в базу, и суммарный их ток определяется разностью этих потоков.

В связи с тем, что в режиме насыщения коллекторный переход уже не осуществляет полной экстракции носителей из базы, там происходит их накопление и интенсивная рекомбинация.

В режиме насыщения соотношения (4.5) и (4.6) не выполняются, ток базы может оказаться сравнимым с током эмиттера.

Режим отсеки. Если на обоих переходах транзистора напряженное обратное, то через них проходят токи, обусловленные процессами тепловой генерации носителей заряда в объеме полупроводника, областях объемного заряда и на невыпрямляющих контактах, а также утечками. При достаточно больших напряжениях происходит лавинное умножение.

§ 4.3. РАСПРЕДЕЛЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА

Значения токов в транзисторе зависят от распределения в нем неосновных носителей заряда. На это распределение влияют многие факторы — геометрические размеры транзистора, параметры его материала, состояние поверхности, токи через переходы, напряжения на них и т. д.

Активный режим. Активный режим работы транзистора характеризуется тем, что из-за наличия обратного напряжения коллекторный переход перехватывает практически все носители заряда, подходящие к его границе. В результате носители в транзисторе движутся в небольшой области, которая по площади примерно равна эмиттерному переходу (рис. 4.6, а), т. е. только через активную часть базы. Лишь немногие носители отклоняются от направления, перпендикулярного эмиттерному переходу, и некоторые из них рекомбинируют на поверхности.

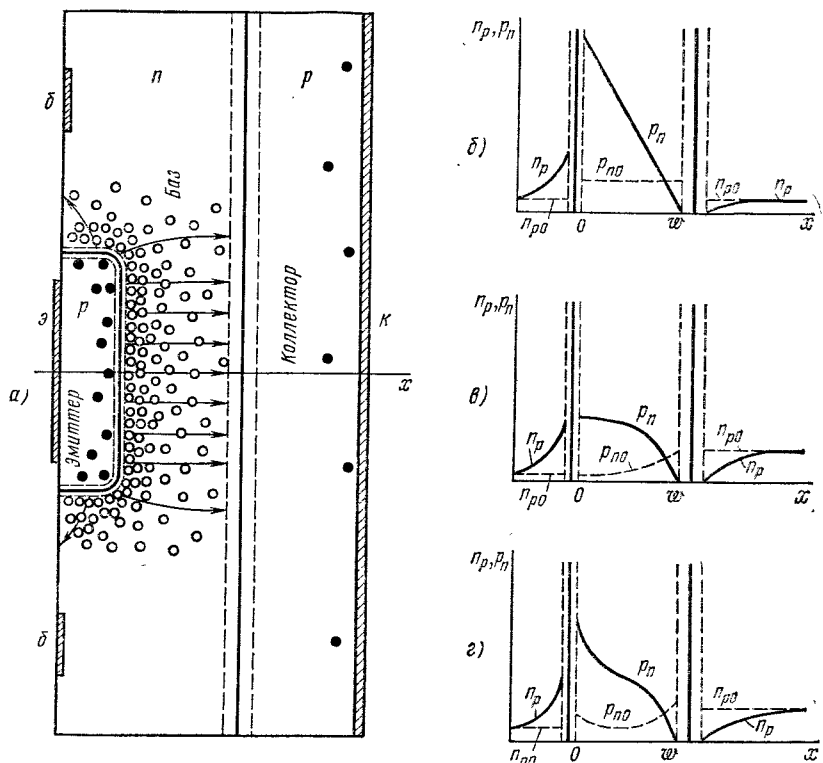


Рис. 4.6. Распределение неосновных носителей заряда в транзисторе при активном режиме:

а — в пространстве (схематично); б — в направлении оси x в бездрейфовом транзисторе; в — в направлении оси x в дрейфовом транзисторе; г — в направлении оси x в транзисторе с участком тормозящего поля в базе

В связи с тем, что большинство носителей в транзисторе, работающем в активном режиме, движется по параллельным путям, с достаточной для практических целей точностью их распределение можно рассматривать как одномерное, т. е. можно считать, что концентрация носителей изменяется только в направлении x , перпендикулярном плоскости эмиттера или коллектора.

Закон распределения носителей заряда в этом направлении определяется электрическим полем в области базы и ее толщиной w . Приближенное аналитическое выражение для такого закона можно получить, если учесть, что токи эмиттера и коллектора транзистора почти одинаковы. Так как носители заряда двигаются параллельными путями, что означает постоянство плотности тока неосновных носителей между эмиттером и коллектором, которая примерно равна плотности тока эмиттера $J_{Эр}$, уравнение для распределения, например, дырок в базе транзистора типа p - n - p примет вид

$$J_{Эр} = q p_n \mu_p E - q D_p \frac{dp_n}{dx} \quad (4.8)$$

или, с учетом (4.3),

$$J_{\Delta p} = -qD_p p_n \frac{1}{N} \cdot \frac{dN}{dx} - qD_p \frac{dp_n}{dx} = -qD_p \frac{1}{N} \cdot \frac{d(p_n N)}{dx}, \quad (4.9)$$

где N — концентрация примесей в базе.

Граничными условиями для случая, когда токи в транзисторе не очень велики, будут (см. рис. 4.6):

при $x = 0$

$$p_n = p_{n0}(0) \exp \frac{qU_{\Delta B}}{kT}; \quad N = N(0); \quad (4.10)$$

при $x = \omega$

$$p_n = \frac{J_{Kp}}{qv_p}; \quad N = N(\omega). \quad (4.11)$$

Координаты $x = 0$ и $x = \omega$ соответствуют границам областей объемного заряда эмиттерного и коллекторного переходов (см. рис. 4.6).

В связи с тем, что скорость движения носителей заряда в области коллекторного перехода v_p довольно велика, примем, что при $x = \omega$ $p_n = 0$.

Решив уравнение (4.9) относительно концентрации носителей в базе транзистора, получим

$$p_n = \frac{J_{\Delta n}}{qN} \int_x^{\omega} \frac{N}{D_p} dx. \quad (4.12)$$

Для бездрейфового транзистора $N = \text{const}$, следовательно, $D_p = \text{const}$ и

$$p_n = \frac{J_{\Delta p}}{qD_p} (\omega - x). \quad (4.13)$$

Таким образом, распределение носителей заряда, инжектированных эмиттером в базу бездрейфового транзистора, линейно (см. рис. 4.6, б). Этот вывод аналогичен полученному ранее для полупроводникового диода с тонкой базой (см. § 3.5) и так же, как для полупроводникового диода, является лишь предельным случаем. Реальное распределение носителей в базе бездрейфового транзистора отличается от линейного, хотя и незначительно.

Для транзисторов, как и для полупроводниковых диодов, можно найти распределение носителей в области эмиттера (как в диоде при прямом включении) и коллектора (как в диоде при обратном включении), что показано на рис. 4.6, б.

Для дрейфового транзистора при постоянном значении напряженности электрического поля в базе (это соответствует экспоненциальному распределению примесей) без учета зависимости D_p от N

$$\begin{aligned} p_n &= \frac{J_{\Delta p}}{q\mu_p E} \left\{ 1 - \exp \left[-\frac{qE}{kT} (\omega - x) \right] \right\} = \\ &= \frac{J_{\Delta p}}{qD_p} \left[1 - \left(\frac{N(\omega)}{N(0)} \right)^{1 - \frac{x}{\omega}} \right] \frac{\omega}{\ln \frac{N(0)}{N(\omega)}}. \end{aligned} \quad (4.14)$$

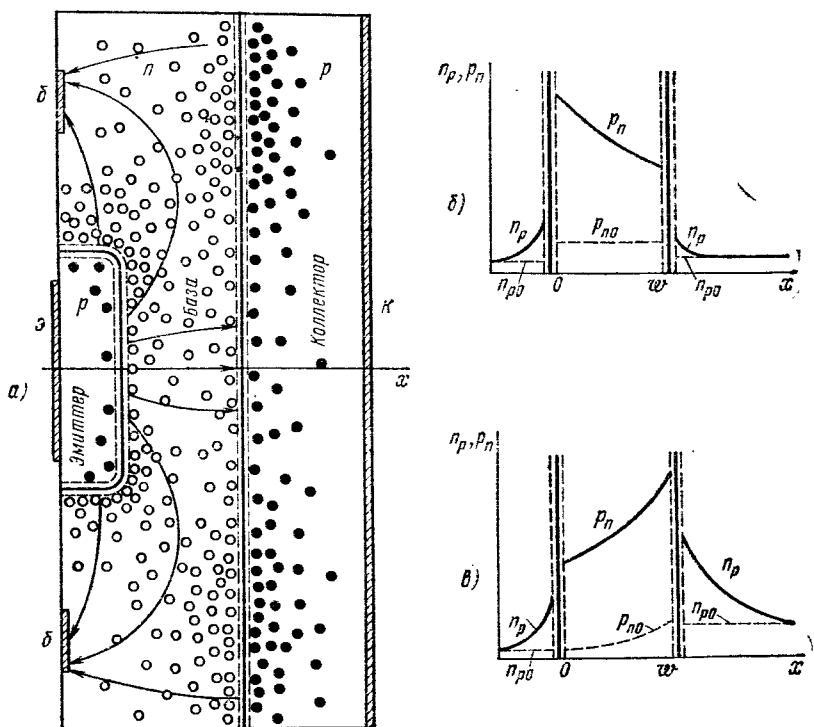


Рис. 4.7. Распределение неосновных носителей заряда в транзисторе при режиме насыщения:

а — в пространстве (схематично); б — в направлении оси x в бездрейфовом транзисторе; в — в направлении оси x в дрейфовом транзисторе

Соответствующее распределение концентрации носителей показано на рис. 4.6, в. В данном случае абсолютная величина градиента концентрации носителей растет по мере приближения к коллекторному переходу. При достаточно сильных электрических полях в базе градиент концентрации носителей вблизи эмиттера становится небольшим, т. е. ток здесь преимущественно дрейфовый. Вблизи коллектора концентрация инжектированных носителей падает, соответственно падает и дрейфовая составляющая тока, но зато растет диффузионная составляющая; поэтому полный ток остается постоянным.

При таком распределении примесей, когда создается участок тормозящего поля, распределение концентрации носителей в базе получается, как показано на рис. 4.6, г. Участку тормозящего электрического поля соответствует резкое возрастание концентрации носителей и ее градиента у эмиттера.

Режим насыщения. При режиме насыщения носители заряда в области базы транзистора двигаются не только от эмиттера к коллектору, но и в сторону вывода базы. Соответственно изменяется и их распределение. При этом значительная концентрация неоснов-

ных носителей заряда создается и в пассивной части базы. Кроме того, прямое напряжение на коллекторном переходе приводит к инжекции в область коллектора.

Все это не дает возможности распространить на режим насыщения полученные ранее расчетным путем распределения носителей. Поэтому ограничимся качественными характеристиками (рис. 4.7). Между эмиттером и коллектором распределение носителей определяется соотношениями напряжений на электронно-дырочных переходах и напряженностью электрического поля в базе. Для бездрейфового транзистора, если напряжение на коллекторном переходе меньше, чем на эмиттерном, концентрация неосновных носителей заряда у коллектора меньше (рис. 4.7, б). Из-за рекомбинации в базе закон распределения не вполне линейный.

Наличие электрического поля в базе дрейфового транзистора способствует движению носителей заряда в сторону коллектора. Кроме того, в этом случае концентрация примесей в базе у коллектора невелика, следовательно, и контактная разность потенциалов в коллекторном переходе меньше, чем в эмиттерном. Поэтому даже при меньших прямых напряжениях на коллекторе, чем на эмиттере, концентрация носителей в базе у коллекторного перехода может оказаться больше, чем у эмиттерного (рис. 4.7, в).

Распределение носителей в базе в направлении к ее выводу определяется инжекцией неосновных носителей через электронно-дырочные переходы и рекомбинацией носителей как в объеме, так и на поверхности базовой области.

Режим отсечки. При режиме отсечки части транзистора, примыкающие к его переходам, сильно обеднены неосновными носителями заряда.

§ 4.4. ЗНАЧЕНИЯ ПОСТОЯННЫХ ТОКОВ ПРИ АКТИВНОМ РЕЖИМЕ

Полученные в § 4.3 распределения концентрации носителей дают возможность записать аналитические выражения для постоянных токов в транзисторе, работающем в активном режиме.

Основную составляющую тока в транзисторе — ток носителей заряда, инжектированных из эмиттера в базу, находят согласно распределению носителей в базе (4.12) при подстановке $x = 0$ и соответствующего значения концентрации из (4.10). Кроме того, чтобы перейти от плотности тока к току, необходимо умножить полученное выражение на площадь эмиттера S_3 .

При расчете получим

$$I_{\text{ЭП}} = \frac{S_3 q n_i^2}{w} \exp \frac{qU_{\text{ЭБ}}}{kT} \int_0^N \frac{N}{D_p} dx \quad (4.15)$$

Для бездрейфового транзистора

$$I_{\text{ЭП}} \approx \frac{S_3 q p_{n0} D_p}{w} \exp \frac{qU_{\text{ЭБ}}}{kT}, \quad (4.16)$$

что почти аналогично выражению для прямого тока полупроводникового диода с тонкой базой (см. § 3.5).

Для дрейфового транзистора при $E = \text{const}$ и $D_p = \text{const}$

$$I_{\text{Эр}} \approx \frac{S_3 q p_{n0}(0) D_p}{\omega} \cdot \frac{\ln \frac{N(0)}{N(\omega)}}{1 - \frac{N(\omega)}{N(0)}} \exp \frac{qU_{\text{ЭБ}}}{kT}. \quad (4.17)$$

Сравнение полученных выражений показывает, что при одинаковых размерах и напряжениях ток эмиттера дрейфового транзистора больше, чем бездрейфового.

Для определения полного тока эмиттера необходимо в соответствии с изложенным в § 4.2 добавить к $I_{\text{Эр}}$ составляющие тока, связанные с инжекцией носителей из базы в эмиттер и рекомбинацией в p - n -переходе эмиттера. Обе составляющие рассчитывают так же, как для полупроводникового диода (см. § 3.5 и 3.10):

$$I_{\text{Эн}} = \frac{S_3 q n_{p0} D_p}{L_n} \left(\exp \frac{qU_{\text{ЭБ}}}{kT} - 1 \right), \quad (4.18)$$

$$I_{\text{Э, рек}} \approx S_3 \frac{q n_i \delta_3}{\tau} \cdot \frac{kT}{q(\varphi_{\text{кон}} - U_{\text{ЭБ}})} \exp \frac{qU_{\text{ЭБ}}}{2kT}. \quad (4.19)$$

Ток базы, обусловленный рекомбинацией носителей в ее объеме и на поверхности, можно также найти из распределения носителей. Область, где избыточная концентрация положительна, т. е. где идет процесс рекомбинации носителей, расположена между эмиттером и коллектором (рис. 4.6, а). Обычно концентрация инжектированных носителей заметно превосходит равновесную. При этом

$$\left(\frac{\partial p}{\partial t} \right)_v \approx \frac{p_n}{\tau_p}. \quad (4.20)$$

Принтегрировав по объему области, где создана избыточная концентрация, получим

$$I_{\text{рек } v} = q S_3 \int_0^{\omega} \frac{p_n}{\tau_p} dx = \frac{I_{\text{Эр}}}{\tau_p} \int_0^{\omega} \left(\frac{1}{N} \int_x^{\omega} \frac{N}{D_p} dx \right) dx. \quad (4.21)$$

На поверхности базы в узкой полосе вокруг эмиттера происходит поверхностная рекомбинация носителей заряда. В единицу времени на единице поверхности рекомбинирует число носителей, равное $s(p_n - p_{n0})_s$ (см. § 1.13). Приняв, что вблизи эмиттера избыточная подповерхностная концентрация

$$(p_n - p_{n0})_s \approx p_n(0) = p_{n0}(0) \exp \frac{qU_{\text{ЭБ}}}{kT}, \quad (4.22)$$

получим для тока базы, связанного с рекомбинацией на поверхности,

$$I_{\text{рек } s} \approx s S_{\text{рек}} \frac{\int_0^{\omega} \frac{N}{D_p} dx}{S_3 N(0)} I_{\text{Эр}}. \quad (4.23)$$

Здесь $S_{\text{рек}}$ — эффективная площадь полосы по периметру эмиттера, где происходит рекомбинация инжектированных носителей. Для бездрейфовых транзисторов можно считать, что ширина этой полосы примерно равна толщине пассивной области базы.

В активном режиме неосновные носители до вывода базы практически не доходят и можно считать $I_{\text{рек, экз}} = 0$.

Составляющая тока коллектора

$$I_{Kp} = I_{\text{Эр}} - (I_{\text{рек в}} + I_{\text{рек s}}). \quad (4.24)$$

В ток коллектора входят еще составляющие, связанные с генерацией носителей в объеме коллектора и в p - n -переходе коллектора. Их определяют примерно так же, как для полупроводникового диода (см. § 3.5 и 3.10).

Ток, связанный с генерацией в коллекторе,

$$I_{K, \text{ген}} = S_K \frac{qn_{p0}L_n}{\tau_n}. \quad (4.25)$$

Следует учесть, что в области коллектора действует электрическое поле, связанное с прохождением тока основных носителей заряда и направленное так, что оно способствует движению неосновных носителей к коллекторному переходу. Под влиянием этого поля носители заряда могут попадать в коллекторный переход не только из областей, равных по толщине L_n , но и из более далеких.

Расстояние, которое проходит носитель заряда путем диффузии за время жизни, равно диффузионной длине. Если к диффузии добавляется дрейф, то это расстояние увеличивается до

$$L_{n \text{эф}} = L_n + \mu_n E \tau_n = L_n \left(1 + \frac{qE}{kT} L_n \right). \quad (4.26)$$

При этом напряженность электрического поля

$$E = \frac{J_K}{q\rho_{p0}\mu_p}, \quad (4.27)$$

тогда

$$L_{n \text{эф}} = L_n \left(1 + L_n \frac{q}{kT} \cdot \frac{J_K}{q\rho_{p0}\mu_p} \right) \quad (4.28)$$

и обратный ток коллектора, связанный с генерацией носителей в его объеме,

$$I_{K, \text{ген}} = S_s \frac{qn_{p0}L_n}{\tau_n} \left(1 + L_n \frac{q}{kT} \cdot \frac{J_K}{q\rho_{p0}\mu_p} \right) + (S_K - S_s) \frac{qn_{p0}L_n}{\tau_n}, \quad (4.29)$$

а с учетом того, что $J_K \approx J_{Kp}$,

$$I_{K, \text{ген}} = S_K \frac{qn_{p0}L_n}{\tau_n} + I_{Kp} \frac{n_{p0}\mu_n}{\rho_{p0}\mu_p} \quad (4.30)$$

или

$$I_{K, \text{ген}} = I_{K, \text{ген}0} - (1 - \bar{\alpha}^*) I_{Kp}, \quad (4.31)$$

где

$$I_{K, \text{ген}0} = S_K \frac{qn_{p0}L_n}{\tau_n}; \quad \bar{\alpha}^* = 1 + \frac{n_{p0}\mu_n}{\rho_{p0}\mu_p}.$$

Значит, составляющая тока коллектора, связанная с тепловой генерацией носителей заряда, зависит от тока носителей, пришедших от эмиттера. Чтобы такое влияние было существенным, концентрации неосновных и основных носителей в коллекторе должны отличаться не очень сильно. Такие условия создаются при повышенной температуре в германиевых транзисторах, изготовленных методом диффузии.

Ток, связанный с генерацией носителей в области объемного заряда, находят обычным путем (см. § 3.10):

$$I_{\text{ген}} = \frac{qn_i \delta_k}{\tau} S_k. \quad (4.32)$$

К току коллектора добавляется еще ток носителей, генерируемых в объеме и на поверхности области базы, а также возникающих в результате лавинного умножения.

Ток генерации носителей в базе $I_{\text{Б, ген}}$ можно определить так же, как обратный ток полупроводникового диода. Лавинное умножение учитывают путем умножения всех проходящих через коллекторный переход токов на коэффициент лавинного умножения M .

Тогда общий ток коллектора

$$I_K = M (I_{Kp} \bar{\alpha}^* + I_{\text{Б, ген}} + I_{\text{К, ген 0}} + I_{\text{ген}}), \quad (4.33)$$

ток базы

$$I_B = I_{\text{Э}} - I_K = I_{\text{рек в}} + I_{\text{рек с}} + I_{\text{рек. кон}} + I_{\text{Эл}} + I_{\text{Э, рек}} - I_{Kp} (\bar{\alpha}^* M - 1) - M (I_{\text{Б, ген}} + I_{\text{К, ген 0}} + I_{\text{ген}}). \quad (4.34)$$

§ 4.5. ЯВЛЕНИЯ В ТРАНЗИСТОРАХ ПРИ БОЛЬШИХ ТОКАХ

При прохождении через транзистор больших токов соотношения, выведенные в § 4.3 и 4.4, нарушаются. Это связано с рядом явлений:

1) прохождение через транзистор больших токов (точнее токов с большой плотностью) часто сопровождается высоким уровнем инжекции. Практически в транзисторах, работающих в активном режиме, высокий уровень инжекции может наблюдаться только в базе. Действительно, эмиттер транзистора изготавливают из сильнолегированного материала, поэтому получить высокий уровень инжекции в нем довольно трудно. К тому же в общем токе эмиттера доля носителей заряда, инжектируемых из базы, невелика (см. § 4.2). В режиме насыщения высокий уровень инжекции возможен и в коллекторе;

2) прохождение больших токов через транзистор связано с перераспределением напряжений в нем из-за падений на объемных сопротивлениях;

3) при больших токах могут изменяться условия как на выпрямляющих, так и на невыпрямляющих контактах (последнее существенно в режиме насыщения);

4) прохождение больших токов приводит к перераспределению потенциала в областях объемного заряда электронно-дырочных переходов.

Высокий уровень инжекции в базе. Явления при высоком уровне инжекции в базе транзистора в принципе не отличаются от рассмотренных ранее для полупроводникового диода (см. § 3.15). В результате изменения концентрации основных носителей, связанного с условием электрической нейтральности, с одной стороны, растет инжекция носителей из базы в эмиттер (ток $I_{Эн}$), а с другой — уменьшается удельное сопротивление материала базы.

Возникающее при высоком уровне инжекции электрическое поле способствует переносу инжектированных носителей заряда через базу. При неравномерном легировании базы (когда транзистор дрейфовый) это электрическое поле накладывается на уже существующее в транзисторе. По мере повышения уровня инжекции распределение электрического поля все больше приближается к тому, которое определяется лишь инжектированными носителями, и перестает зависеть от распределения примесей.

При высоком уровне инжекции время жизни носителей в базе также изменяется. К изменению времени жизни в объеме добавляется изменение скорости поверхностной рекомбинации. В результате соотношения между током базы и током эмиттера (4.21) и (4.23) искажаются.

Перераспределение напряжений. При прохождении токов через объемные сопротивления на них создаются соответствующие падения напряжений. В области базы наиболее существенно падение напряжения из-за прохождения тока из пространства между эмиттером и коллектором к выводу базы (рис. 4.8, а). Так как толщина базы очень мала, ее сопротивление может оказаться довольно большим. В результате падение напряжения на эмиттерном переходе отличается от напряжения между выводами эмиттера и базы и не одинаково (что очень существенно) в разных точках эмиттера

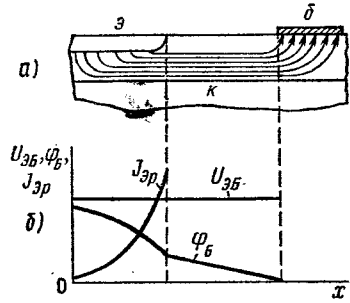


Рис. 4.8. Картина прохождения тока базы (а); распределение потенциала в базе ϕ_B и плотности инжекционного тока эмиттера $J_{Эр}$ при некотором напряжении $U_{ЭБ}$ между выводами эмиттера и базы (б)

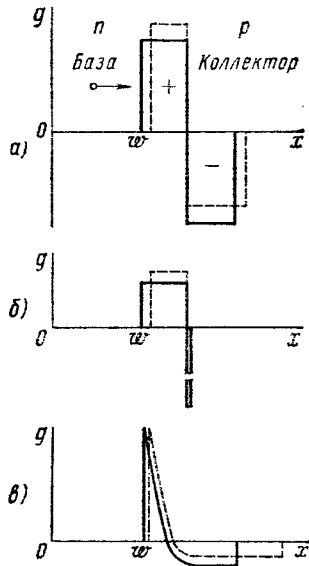


Рис. 4.9. Перераспределение плотности объемного заряда в коллекторном переходе при больших токах:

а — резкий переход; б — резкий переход с сильнолегированной областью коллектора; в — диффузионный переход

(рис. 4.8, б). Это приводит к тому, что и плотность тока эмиттера в разных точках различна; причем она возрастает к краю эмиттера, т. е. по мере роста тока средние части эмиттерного перехода как бы перестают работать — инжекция происходит преимущественно по краю эмиттера.

Изменение условий на контактах. В активном режиме работы транзистора при больших токах могут измениться условия на коллекторном переходе. Это связано (см. § 1.10) с насыщением дрейфовой скорости движения носителей. При этом концентрацию носителей в базе у коллектора нельзя считать равной нулю, как это было сделано в § 4.3, — ее значение стремится к $J_{Kp}/(qv_{p\max})$. Такое явление влечет за собой перераспределение носителей в базе. Если для бездрейфовых транзисторов из-за значительного градиента концентрации носителей во всей базе и обычно не очень больших плотностей токов влиянием такого перераспределения можно пренебречь, то для дрейфовых транзисторов оно может быть существенным.

Перераспределение потенциала в области объемного заряда при больших токах является также следствием насыщения скорости движения носителей. Действительно, из-за того что носители движутся с конечной скоростью, их концентрация в электронно-дырочном переходе отлична от нуля, а заряд добавляется к заряду ионизированных примесей или вычитается из него. Соответственно перераспределяется и поле в переходе.

Рассмотрим коллекторный переход транзистора (рис. 4.9, а). Нетрудно убедиться, что заряд, связанный с неосновными носителями, совпадает по знаку с зарядом ионизированных примесей в базе. Следовательно, из-за прохождения коллекторного тока плотность объемного заряда в части коллекторного перехода, приходящейся на базу, растет, а в части, приходящейся на коллектор, — падает. Такое изменение плотности объемного заряда эквивалентно соответствующему изменению концентрации примесей. Следовательно, при больших токах область объемного заряда в базе сужается, а в коллекторе расширяется.

Если коллекторный переход резкий, как обычно у бездрейфовых транзисторов, и область коллектора легирована очень сильно (сплавной переход), то при больших токах изменяется практически только толщина области базы (рис. 4.9, б). Когда коллекторный переход плавный (дрейфовый транзистор), изменение ширины базы сопровождается сдвигом обеих границ области объемного заряда в сторону коллектора (рис. 4.9, в). Если плотность тока удовлетворяет условию

$$J_{Kp} \geq qN_{исх}v_{p\max},$$

то носители заряда полностью компенсируют заряд примесей в коллекторной части перехода, и коллекторный переход как бы исчезает. При этом работа транзистора нарушается.

§ 4.6. СТАТИЧЕСКИЕ ПАРАМЕТРЫ

Статическими параметрами транзистора называют значения токов или напряжений, измеренных в определенных условиях, а также некоторые соотношения между этими величинами.

Параметры режима отсечки. В качестве статических параметров, характеризующих режим отсечки, обычно выбирают значения токов эмиттера и коллектора. В связи с тем, что в режиме отсечки наблюдается некоторое влияние одного перехода транзистора на другой, эти токи находят при определенных условиях включения транзистора.

Начальными токами переходов называют значения токов при обратном напряжении на каком-либо переходе транзистора при условии, что напряжение на другом переходе равно нулю. Практически начальные токи находят, применяя короткое замыкание выводов соответствующей области с базой. Так, например, *начальный ток коллектора* $I_{КБК}^*$ определяют при коротком замыкании выводов эмиттера и базы, а *начальный ток эмиттера* $I_{ЭБК}$ — при коротком замыкании выводов коллектора и базы. Так как начальные токи переходов транзистора обычно довольно малы, короткому замыканию выводов соответствует напряжение на переходе, практически равное нулю. Согласно определению начальных токов их находят на границе перехода транзистора из активного режима в режим отсечки.

Рассмотрим для примера происхождение тока $I_{КБК}$. При напряжении на эмиттере, равном нулю, обращаются в нуль только составляющие эмиттерного тока $I_{Эн}$ и $I_{Э, рек}$, в то время как в соответствии с (4.15) при $U_{ЭБ} = 0$

$$I_{Эр} = \frac{S_3 q n_i^2}{w} \int_0^w \frac{N}{D_p} dx \neq 0. \quad (4.35)$$

Это объясняется тем, что при подаче напряжения на коллекторный переход в базе перераспределяются неосновные носители, возникает градиент их концентрации и соответствующий ток. Начальный ток коллектора складывается из тока $I_{Кр}$, также не равного нулю, и обратных токов коллекторного перехода. При не очень больших напряжениях ($M = 1$)

$$I_{КБК} = \frac{S_3 q n_i^2}{w} \int_0^w \frac{N}{D_p} dx - I_{рек \text{ в}} - I_{рек \text{ с}} - I_{рек, \text{ кол}} + I_{Б, \text{ ген}} + I_{К, \text{ ген}} + I_{ген}. \quad (4.36)$$

* В индексах первая буква означает вывод, в цепи которого производится измерение (Э — эмиттер, К — коллектор, Б — база), вторая буква — общий вывод, третья буква — режим измерения (К — короткое замыкание, О — разомкнутая цепь).

Аналогично образуется и начальный ток эмиттера $I_{ЭБК}$.

Под обратными токами переходов транзистора понимают значения токов через какой-либо из переходов при обратном напряжении, в то время как ток в одном из двух оставшихся свободными выводов равен нулю. Так обратный ток коллектора транзистора в схеме с общей базой $I_{КБО}$ определяется при токе эмиттера, равном нулю, а обратный ток эмиттера для той же схемы $I_{ЭБО}$ — при токе коллектора, равном нулю. Аналогично можно определить и обратный ток коллектора для схемы с общим эмиттером $I_{КЭО}$ (при токе базы, равном нулю).

Значение тока $I_{КБО}$ можно найти из (4.33) с учетом того, что при этом ток эмиттера отсутствует, а значит, и $I_{Кр} = 0$. Тогда

$$I_{КБО} = (I_{Б.ген} + I_{К.ген} + I_{ген}) M. \quad (4.37)$$

Подставив это выражение в (4.33), получим

$$I_{К} = MI_{Кр} + I_{КБО}. \quad (4.38)$$

Следовательно, ток $I_{КБО}$ представляет собой суммарную неуправляемую со стороны эмиттера составляющую тока коллектора.

Ток базы

$$I_{Б} = I_{рек\ v} + I_{рек\ s} + I_{рек\ кон} + I_{Эн} + I_{Эрек} - I_{Кр}(M-1) - I_{КБО}. \quad (4.39)$$

Здесь тоже $I_{КБО}$ представляет собой составляющую, которая не зависит от тока эмиттера. Аналогично определяется и обратный ток эмиттера $I_{ЭБО}$.

Параметры активного режима. Статические параметры активного режима в основном характеризуют степень влияния входной цепи транзистора (эмиттерной, базовой) на выходную. К этим параметрам относится *статический коэффициент передачи тока базы*

$$h_{21Э} = \frac{I_{К} - I_{КБО}}{I_{Б} + I_{КБО}}. \quad (4.40)$$

Подставив полученные в § 4.4 выражения для токов коллектора и базы, получим

$$\begin{aligned} \frac{1}{h_{21Э}} \approx & \frac{1}{\bar{\alpha}^* M} \left[\frac{1}{\tau_p} \int_0^w \left(\frac{1}{N} \int_x^w \frac{N}{D_p} dx \right) dx + \right. \\ & + \frac{s S_{рек}}{S_0 N(0)} \int_0^w \frac{N}{D_p} dx + \frac{D_n}{N(0) L_n} \int_0^w \frac{N}{D_p} dx + \\ & \left. + \frac{\delta_3}{n_i \tau} \cdot \frac{kT}{q(\varphi_{кон} - U_{ЭБ})} \int_0^w \frac{N}{D_p} dx \exp\left(-\frac{qU_{ЭБ}}{2kT}\right) - \bar{\alpha}^* M + 1 \right] \quad (4.41) \end{aligned}$$

Значение статического коэффициента передачи тока базы определяется несколькими процессами, что в формуле (4.41) отражено соответствующими сомножителями и слагаемыми.

Множитель $\bar{\alpha}^*$, учитывающий изменение обратного тока коллектора при изменении тока эмиттера, обычно примерно равен единице. Коэффициент лавинного умножения M учитывает влияние этого процесса на общий ток коллектора.

Слагаемое

$$\frac{1}{\tau_p} \int_0^w \left(\frac{1}{N} \int_x^w \frac{N}{D_p} dx \right) dx = \frac{t_{\text{прол}}}{\tau_p} \quad (4.42)$$

учитывает потери носителей заряда в базе при объемной рекомбинации,
Слагаемое

$$\frac{sS_{\text{рек}}}{S_0 N(0)} \int_0^w \frac{N}{D_p} dx \quad (4.43)$$

учитывает потери носителей в базе из-за поверхностной рекомбинации,
Слагаемое

$$\frac{D_n}{N(0) L_n} \int_0^w \frac{N}{D_p} dx \quad (4.44)$$

учитывает влияние тока носителей заряда, инжектированных из базы в эмиттер,
Слагаемое

$$\frac{\delta_9}{n_i \tau} \cdot \frac{kT}{q(\Phi_{\text{кон}} - U_{\text{ЭБ}})} \int_0^w \frac{N}{D_p} dx \exp\left(-\frac{qU_{\text{ЭБ}}}{2kT}\right) \quad (4.45)$$

учитывает потери носителей из-за рекомбинации в эмиттерном переходе.

Из перечисленных процессов только лавинное умножение и изменение обратного тока коллектора под действием тока эмиттера приводят к росту значения $h_{21Э}$, остальные процессы обуславливают его снижение.

Для бездрейфового транзистора выражение (4.41) принимает более простой вид:

$$\frac{1}{h_{21Э}} = \frac{1}{\bar{\alpha}^* M} \left[\frac{\omega^2}{2L_p^2} + \frac{sS_{\text{рек}}\omega}{S_0 D_p} + \frac{D_n N \omega}{D_p N(0) L_n} + \frac{\delta_9 N \omega}{n_i \tau D_p} \cdot \frac{kT}{q(\Phi_{\text{кон}} - U_{\text{ЭБ}})} \exp\left(-\frac{qU_{\text{ЭБ}}}{2kT}\right) - M \left(1 + \frac{n_{p0} \mu_n}{p_{p0} \mu_p} \right) + 1 \right]. \quad (4.46)$$

Коэффициент передачи тока базы — один из важнейших параметров транзистора; поэтому целесообразно рассмотреть влияние на него значения постоянной составляющей тока.

При малых напряжениях, как известно (см. § 3.10), в образовании тока через электронно-дырочный переход большую роль играет рекомбинация в p - n -переходе. Так как рекомбинационный ток не связан с инжекцией в базу, он приводит к снижению $h_{21Э}$. В формуле (4.41) это отражено в том, что слагаемое, учитывающее рекомбинацию в p - n -переходе (4.45), растет с уменьшением $U_{\text{ЭБ}}$.

По мере роста напряжения на эмиттере и тока эмиттера роль рекомбинации в p - n -переходе уменьшается, соответственно растет $h_{21Э}$. К тому же с увеличением тока повышается уровень инжекции, а возникающее при этом поле способствует движению инжектированных носителей заряда через базу. Действие такого поля

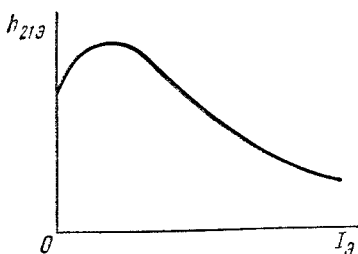


Рис. 4.10. Зависимость коэффициента передачи постоянного тока базы от постоянного тока эмиттера

эквивалентно росту D_p (см. § 3.15), что приводит к уменьшению первого и второго слагаемых в формуле (4.41).

Одновременно при повышении уровня инжекции начинает расти концентрация носителей в базе, что ведет к росту тока носителей, инжектированных из базы в эмиттер. В формуле (4.46) вместо N в третьем слагаемом следовало бы взять $N + p_n$. Соответственно уменьшается $h_{21Э}$.

К снижению $h_{21Э}$ при больших токах приводит и вытеснение носителей к краю эмиттера, рассмотренное

в § 4.5, в связи с чем возрастает роль поверхностной рекомбинации.

В некоторых транзисторах уменьшение $h_{21Э}$ при больших токах может быть связано с увеличением толщины базы. На зависимость $h_{21Э}$ от тока влияет также изменение времени жизни носителей в объеме и скорости поверхностной рекомбинации. Результирующая зависимость $h_{21Э}$ от $I_Э$ чаще всего имеет вид кривой с максимумом (рис. 4.10). Обычно значение $h_{21Э}$ составляет несколько десятков.

Иногда вводят *статический коэффициент передачи тока эмиттера*

$$h_{21Б} = -\frac{I_K - I_{КБ0}}{I_Э}, \quad (4.47)$$

при этом

$$h_{21Э} \approx \frac{|h_{21Б}|}{1 - |h_{21Б}|}. \quad (4.48)$$

Абсолютное значение $h_{21Б}$, как правило, лишь немного меньше единицы.

Помимо этого иногда используют значение крутизны передаточной характеристики

$$S = \frac{I_K - I_{КБ0}}{U_{ЭБ}}. \quad (4.49)$$

Формулы для расчета этих параметров можно вывести примерно так же, как для $h_{21Э}$.

Используя статические параметры и выражение для тока $I_{КБ0}$, перепишем выражение для тока $I_{КБК}$:

$$I_{КБК} = h_{21Б} \frac{qn_i^2 S_э}{w} + I_{КБ0} \int_0^w \frac{N}{D_p} dx. \quad (4.50)$$

Параметры режима насыщения. В качестве параметров режима насыщения обычно принимают измеренные при определенных токах значения падения напряжения между выводами транзистора. На-

пример, $U_{КБ,нас}$ — падение напряжения между выводами коллектора и базы в режиме насыщения (для схемы с общей базой) или $U_{КЭ,нас}$ — падение напряжения между выводами коллектора и эмиттера в режиме насыщения (для схемы с общим эмиттером).

Если эти падения напряжения отнести к току, проходящему через коллектор, то полученный параметр называют **сопротивлением насыщения**.

§ 4.7. ПРОБОЙ ТРАНЗИСТОРОВ

На процессы в транзисторе существенное влияние оказывает напряжение на коллекторе. Такое влияние обусловлено следующим. При изменении напряжения изменяется толщина области объемного заряда коллекторного перехода и соответственно толщина базы, а при достаточно больших коллекторных напряжениях начинает сказываться лавинное умножение.

С повышением напряжения на коллекторе толщина базы становится меньше, что приводит к увеличению градиента концентрации носителей заряда в базе, к уменьшению времени, в течение которого носители находятся в базе и, следовательно, к уменьшению роли рекомбинации в базе. Это ведет к росту коэффициентов передачи $h_{21Э}$ и $h_{21Б}$.

Смыкание переходов. При достаточно больших напряжениях на коллекторном переходе область объемного заряда коллекторного перехода может достигнуть эмиттерного перехода — произойдет так называемое смыкание переходов. При этом потенциальный барьер эмиттерного перехода понижается (рис. 4.11), возрастает ток эмиттера, а значит, и ток коллектора. По внешним признакам смыкание напоминает пробой или короткое замыкание эмиттера с коллектором. Таким образом, смыкание переходов является одной из причин, ограничивающих напряжение коллектора.

Лавинный пробой. Второй причиной, ограничивающей напряжение коллектора, является лавинное умножение. При этом существенную роль играет режим цепи базы. Если ток в цепи базы не ограничен, что имеет место, например, в схеме с общей базой, то пробой транзисторов не отличается от пробоя полупроводникового диода. В этом случае в коллекторном переходе произойдет лавинный пробой при напряжении $U_{проб}$.

Лавинный пробой коллекторного перехода представляет собой обратимый процесс, если ограничить возрастающий при пробое ток. С увеличением тока коллектора при лавинном умножении лавинный пробой может перейти в тепловой пробой с появлением отрицательного дифференциального сопротивления на выходе транзистора. Этот переход к тепловому пробую наиболее вероятен в транзисторах, изготовленных из германия (материала с малой шириной запрещенной зоны).

С увеличением напряжения на коллекторе коэффициент передачи постоянного тока эмиттера с учетом умножения в коллекторном переходе возрастает и при некотором напряжении на коллек-

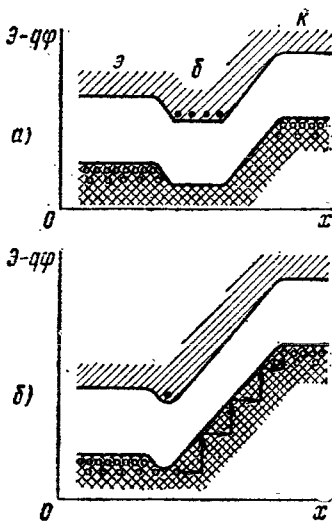


Рис. 4.11. Явление смыкания переходов в транзисторе:

а — энергетические зоны при небольших коллекторных напряжениях; б — смыкание областей объемного заряда

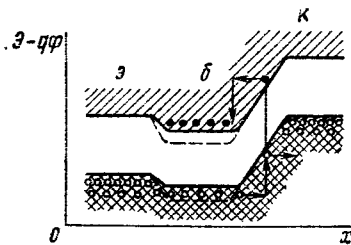


Рис. 4.12. Изменение потенциальных барьеров в транзисторе при лавинном умножении в коллекторном переходе и токе базы, равном нулю

торе $U(L)_{кЭ0}$ становится равным единице. Это напряжение иногда называют напряжением переворота фазы тока базы, так как при дальнейшем увеличении напряжения изменяется направление движения носителей заряда через вывод базы.

Ток коллектора в схеме с общим эмиттером

$$I_K = I_{КБ0} + |h_{21Б}| I_{Э} = \frac{I_{КБ0}}{1 - |h_{21Б}|} + \frac{h_{21Б}}{1 - |h_{21Б}|} I_{Б}. \quad (4.51)$$

При неизменном токе базы ($I_{Б} = \text{const}$) и при $U_{кЭ} \rightarrow U(L)_{кЭ0}$ ток коллектора будет стремиться к бесконечности, т. е. произойдет пробой транзистора, включенного по схеме с общим эмиттером.

Если ток базы фиксирован (например, при разомкнутой цепи базы или при включении в нее достаточно большого сопротивления), то в транзисторе начинает проявляться обратная связь. Образующиеся при лавинном умножении пары носителей заряда разделяются полем коллекторного перехода, при этом неосновные для базы носители уходят в коллектор, а основные — в базу (рис. 4.12). Таким образом, в базе создается избыточный заряд, в связи с чем изменяется ее потенциал. Получающееся при этом напряжение открывает эмиттерный переход.

Если вывод базы отсоединен, то основные носители заряда, накопившиеся в базе, могут исчезнуть только

двумя путями — либо уйти в эмиттер, либо рекомбинировать с носителями, инжектированными эмиттером. Однако транзистор делают так, что вероятность этих событий довольно мала — из эмиттера в базу проходит гораздо больше носителей, чем из базы в эмиттер, и носители, инжектированные в базу, почти не рекомбинируя, доходят до коллектора. Следовательно, на каждый основной носитель, попавший в базу, должно пройти из эмиттера в коллектор много неосновных носителей, прежде чем исчезнет избыточный заряд в базе. Значит, даже небольшое

возрастание коллекторного тока из-за лавинного умножения приводит к существенному росту тока эмиттера, что в свою очередь ведет к росту тока коллектора. Этот процесс может привести и к неустойчивой работе транзистора.

Таким образом, пробой транзистора, включенного по схеме с общим эмиттером, может происходить при меньших напряжениях, чем в схеме с общей базой. Соотношения между этими напряжениями можно найти, используя выражение для коэффициента умножения носителей в коллекторном переходе транзистора (3.70):

$$|h_{21Б}| = |h_{21Б}|_{M-1} M = \frac{|h_{21Б}|_{M-1}}{1 - \left(\frac{U(L)_{КЭО}}{U_{проб}}\right)^b} = 1. \quad (4.52)$$

Тогда

$$U(L)_{КЭО} = U_{проб} \sqrt[b]{1 - |h_{21Б}|_{M-1}}, \quad (4.53)$$

где $|h_{21Б}|_{M-1}$ — абсолютное значение $h_{21Б}$ без учета лавинного умножения.

Согласно выражению (4.53) разница между $U(L)_{КЭО}$ и $U_{проб}$ тем больше, чем больше величина коэффициента передачи постоянного тока эмиттера и чем меньше величина коэффициента b , который зависит от исходного материала (см. § 3.11).

Отсюда можно сделать следующие выводы:

1. Необходимо иметь в виду возможность пробоя транзистора, включенного по схеме с общим эмиттером, при значительно меньших напряжениях, чем $U_{проб}$, если в цепи базы включено относительно большое сопротивление. Эти процессы будут обратимыми, если ток коллектора ограничен параметрами внешней схемы в цепи коллектора (например, при большом сопротивлении нагрузки). В противном случае мощность, выделяющаяся в коллекторном переходе при резком увеличении тока коллектора, может превысить допустимую величину, тогда произойдет необратимый пробой транзистора, связанный с тепловым разрушением.

2. При включении транзистора в схему, находящуюся под напряжением (например, при измерении параметров транзисторов), вначале необходимо присоединить вывод базы, а затем выводы эмиттера и коллектора, чтобы не возникло условие нулевого базового тока.

Вторичный пробой. Под вторичным пробоем понимают явления, приводящие к внезапному переходу транзистора в состояние, которое характеризуется большим током и относительно малым напряжением. При вторичном пробое ток, проходящий через транзистор, локализуется в узкой области (шнуре).

Шнурование тока связано с наличием различного рода дефектов на поверхности и в объеме транзисторной структуры или каких-либо неоднородностей, вызывающих локальное увеличение плотности тока через коллекторный переход. Локальное увеличение плот-

ности тока приводит к локальному разогреву, что в свою очередь обуславливает:

1) увеличение тепловой генерации носителей в этом месте *p-n*-перехода коллектора и, следовательно, увеличение локальной плотности тока, еще больший локальный разогрев и т. д. При этом тепловая генерация носителей может возрасти настолько, что область объемного заряда вообще исчезнет на малой площади коллекторного перехода;

2) локальное уменьшение толщины *p-n*-перехода коллектора и повышение напряженности поля, а следовательно, и увеличение ударной ионизации. При этом увеличатся локальная плотность тока, локальный перегрев, тепловая генерация и т. д.;

3) увеличение локального коэффициента передачи тока эмиттера, так как с повышением температуры растет время жизни носителей; это вызовет еще большую концентрацию проходящего тока, еще больший локальный разогрев и т. д.

Все эти явления приводят к резкому увеличению тока и уменьшению напряжения. Инерционность данных явлений, связанных с тепловыми процессами, может быть очень малой из-за малости объема, где происходит шнурование.

Если ток через транзистор при вторичном пробое не ограничить, то локальный разогрев приведет к необратимым изменениям в транзисторе вплоть до локального расплавления. Чаще всего в результате вторичного пробоя базовая область транзистора проплавляется эмиттерным сплавом, т. е. сквозь базу протягивается тонкая область того же типа электропроводности, что и эмиттерная и коллекторная области, закорачивая эмиттер и коллектор. При этом характеристики как эмиттерного, так и коллекторного переходов могут остаться неизменными, а эмиттерная и коллекторная области окажутся закороченными.

Если же ток через транзистор при вторичном пробое ограничить, то локальный разогрев может не привести к расплавлению кристалла в тонкой области шнурования тока. В этом случае иногда может существовать относительно стабильное состояние с малым падением напряжения между коллектором и эмиттером. Однако сохранение подобного состояния в течение длительного времени или неоднократные повторения вторичного пробоя обычно приводят к необратимым изменениям параметров транзистора в связи с большой локализацией выделяющейся мощности.

Необходимо отметить, что отсутствие дефектов в структуре транзистора не гарантирует от возникновения вторичного пробоя. Так, базовая область под *p-n*-переходом может быть не эквипотенциальна из-за прохождения базовых токов. При различных направлениях тока базы наблюдается увеличение плотности тока эмиттера либо по периферии эмиттерного перехода, либо в центре перехода. Этот эффект может также создавать предпосылки для развития процессов шнурования тока.

Вероятность вторичного пробоя возрастает с увеличением рабочего тока транзистора и напряжения на коллекторе, так как при

этом неравномерность выделения мощности в транзисторе проявляется сильнее из-за большего перегрева и, следовательно, из-за большего влияния собственной электропроводности.

§ 4.8. СТАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Системы статических характеристик

Если обозначить напряжение и ток входного электрода транзистора через U_1 и I_1 , а напряжение и ток выходного — через U_2 и I_2 , то взаимозависимость этих четырех величин можно выразить двадцатью четырьмя семействами характеристик, относящихся к шести системам:

$$\begin{array}{ll}
 1) U_1 = f_1(I_1, I_2), & 4) I_1 = f_1(U_1, U_2), \\
 U_2 = f_2(I_1, I_2); & U_2 = f_2(U_1, I_2); \\
 2) I_1 = f_1(U_1, U_2), & 5) I_1 = f_1(I_2, U_2), \\
 I_2 = f_2(U_1, U_2); & U_1 = f_2(I_2, U_2); \\
 3) U_1 = f_1(I_1, U_2), & 6) I_2 = f_1(U_1, I_1), \\
 I_2 = f_2(I_1, U_2); & U_2 = f_2(U_1, I_1).
 \end{array}$$

Из четырех возможных семейств характеристик каждой системы, которые связывают между собой четыре величины I_1 , I_2 , U_1 и U_2 , два семейства являются основными, а два второстепенными. Их можно получить из основных путем перестроения. В качестве основных удобно выбирать семейства характеристик, связывающих ток и напряжение на входе — входные характеристики и ток и напряжение на выходе — выходные характеристики. Другие два семейства характеристик являются следствием входных и выходных. Семейства характеристик, которые связывают токи или напряжения на выходе с токами или напряжениями на входе, называют характеристиками передачи, а семейства, которые связывают напряжения или токи на входе с токами или напряжениями на выходе, — характеристиками обратной связи.

Из шести систем статических характеристик как наиболее удобные применяют первые три, но наибольшее распространение получила система

$$U_1, I_2 = f(I_1, U_2). \quad (4.54)$$

Важным преимуществом такой системы является то, что при исследовании статических характеристик в качестве независимых переменных берут ток на входе и напряжение на выходе транзистора. В схемах с общим эмиттером и с общей базой входной ток проходит через открытый эмиттерный переход, сопротивление которого мало, так что поддерживать заданное значение тока на входе удобно, включив последовательно относительно большое сопротивление. Необходимые напряжения на выходе будут поддерживаться постоянными, если выходную цепь питать от генератора напряжения.

На практике это осуществить легко, поскольку выходная цепь транзистора обладает высоким сопротивлением.

Кроме того, система уравнений (4.54) соответствует наиболее распространенной системе малосигнальных параметров (h -параметров), которая в свою очередь имеет ряд преимуществ перед другими системами малосигнальных параметров (см. § 4.10).

Характеристики транзистора, включенного по схеме с общей базой

Входные характеристики. Семейство входных статических характеристик транзистора, включенного по схеме с общей базой, представлено на рис. 4.13, а.

Общий характер этих зависимостей определяется p - n -переходом эмиттера, включенным в прямом направлении. Поэтому по внешнему виду входные характеристики похожи на прямые ветви вольт-амперной характеристики диода.

Смещение входных статических характеристик вниз в выбранной системе координат при увеличении абсолютной величины напряжения на коллекторе объясняется тем, что напряжение на коллекторе влияет на концентрацию носителей около него и изменяет толщину базы из-за изменения толщины коллекторного перехода. Эти причины приводят к увеличению градиента концентрации неосновных носителей (дырок) в базе с увеличением абсолютной величины напряжения на коллекторе при постоянном напряжении на эмиттере (рис. 4.13, б).

Следовательно, с увеличением абсолютной величины напряжения на коллекторе и при постоянном напряжении на эмиттере возрастает ток эмиттера, т. е. характеристики смещаются вниз.

Тот же результат можно получить, если рассмотреть уменьшение концентрации неосновных носителей в базе транзистора около

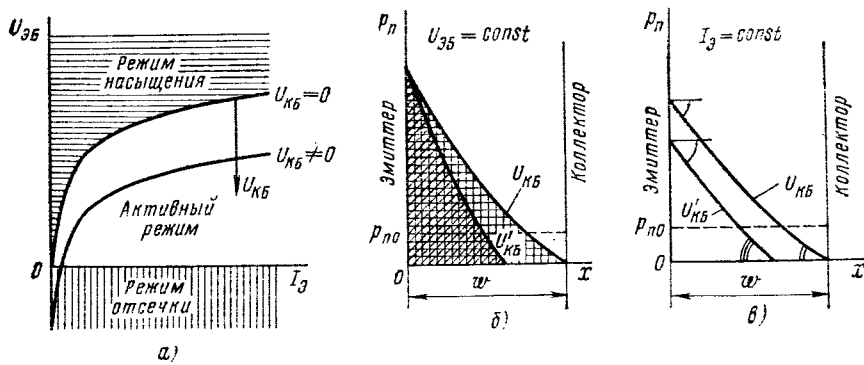


Рис. 4.13. Входные статические характеристики транзистора, включенного по схеме с общей базой (а), и распределение концентрации дырок в области базы при постоянном напряжении на эмиттерном переходе (б) и при постоянном токе эмиттера (в); $|U'_{кб}| > |U_{кб}|$

p - n -перехода эмиттера при увеличении абсолютной величины напряжения на коллекторе и при постоянном токе эмиттера (рис. 4.13, в).

Через эмиттерный p - n -переход протекает ток и при напряжении на эмиттере, равном нулю. Чтобы ток эмиттера стал равным нулю, на эмиттер должно быть подано обратное напряжение (работа в режиме отсечки).

Выходные характеристики. Семейство выходных статических характеристик транзистора, включенного по схеме с общей базой, представлено на рис. 4.14.

Общий характер этих зависимостей аналогичен обратной ветви вольт-амперной характеристики диода, так как коллекторный переход включен в обратном направлении.

Смещение выходных статических характеристик вверх в выбранной системе координат при увеличении тока эмиттера соответствует принципу действия транзистора.

Через коллекторный переход протекает ток и при напряжении на коллекторе, равном нулю, т. е. при коротком замыкании коллектора с базой. Это связано с наличием градиента концентрации неосновных носителей в базе транзистора при инжекции носителей из эмиттера, т. е. при существовании тока эмиттера. Чтобы ток коллектора стал равным нулю, на коллектор должно быть подано прямое напряжение (работа в режиме насыщения).

Характеристики передачи тока. Семейство статических характеристик передачи тока транзистора, включенного по схеме с общей базой, показано на рис. 4.15.

Общий характер этих зависимостей свидетельствует о том, что в транзисторе $I_K = I_{КБ0} + |h_{21Б}| I_E$. Характеристики передачи в первом приближении можно считать прямыми линиями. В действительности коэффициент передачи

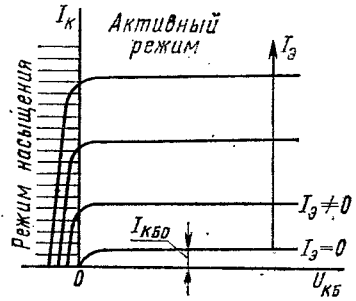


Рис. 4.14. Выходные статические характеристики транзистора, включенного по схеме с общей базой

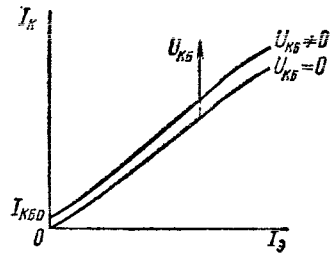


Рис. 4.15. Статические характеристики передачи тока транзистора, включенного по схеме с общей базой

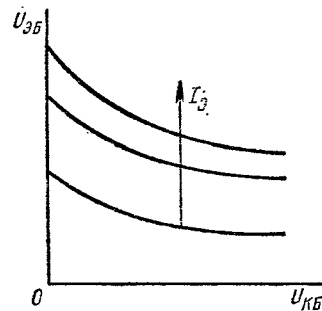


Рис. 4.16. Статические характеристики обратной связи транзистора, включенного по схеме с общей базой

постоянного тока эмиттера зависит от тока эмиттера (причины зависимости рассмотрены в § 4.5 и 4.6).

Смещение статических характеристик передачи вверх в выбранной системе координат при увеличении напряжения на коллекторе связано с уменьшением рекомбинации дырок при их прохождении через более тонкую базу (рис. 4.13, в).

При напряжении на коллекторе, отличном от нуля, характеристики передачи выходят не из начала координат, а из точек на оси ординат, соответствующих обратному току коллектора $I_{КБО}$. Однако величиной этого тока часто можно пренебречь.

Характеристики обратной связи. Семейство статических характеристик обратной связи транзистора, включенного по схеме с общей базой, приведено на рис. 4.16.

Общий характер этих зависимостей можно выяснить путем анализа рис. 4.13, в. Характеристики обратной связи могут быть легко получены из семейства входных статических характеристик путем графического перестроения.

Смещение статических характеристик обратной связи вверх при увеличении тока эмиттера очевидно и не требует специального пояснения.

Характеристики транзистора, включенного по схеме с общим эмиттером

Входные характеристики. Семейство входных статических характеристик транзистора, включенного по схеме с общим эмиттером, представлено на рис. 4.17.

Общий характер этих зависимостей аналогичен характеру подобных характеристик транзистора, включенного по схеме с общей базой, так как ток базы является суммой обратного тока коллектора $I_{КБО}$ и рекомбинационной составляющей, которая примерно пропорциональна току эмиттера и представляет собой малую часть тока эмиттера.

Смещение входных статических характеристик вверх в выбранной системе координат при увеличении абсолютного значения напряжения на коллекторе связано с уменьшением общего количества неосновных носителей в базе (заштрихованные площади на рис. 4.13, б) и, следовательно, с уменьшением количества рекомбинирующих носителей. Поэтому уменьшается составляющая тока базы, обусловленная рекомбинацией, при постоянном напряжении между базой и эмиттером.

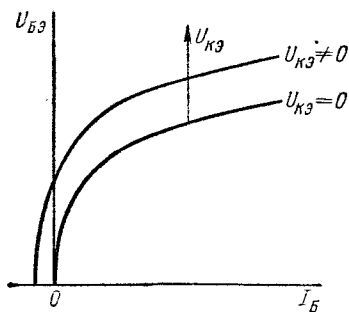


Рис. 4.17. Входные статические характеристики транзистора, включенного по схеме с общим эмиттером

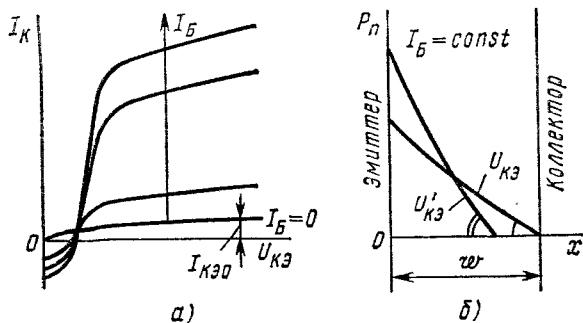


Рис. 4.18. Выходные статические характеристики транзистора, включенного по схеме с общим эмиттером (а), и распределение концентрации дырок в области базы при постоянном токе базы (б); $|U'_{КЭ}| > |U_{КЭ}|$

При наличии напряжения на коллекторе и токе базы, равном нулю, на выводе базы существует напряжение, которое объясняется падением напряжения на сопротивлении $p-n$ -перехода эмиттера. Поэтому входные статические характеристики для схемы с общим эмиттером заходят во второй квадрант при $U_{КЭ} \neq 0$.

Выходные характеристики. Семейство выходных статических характеристик транзистора, включенного по схеме с общим эмиттером, дано на рис. 4.18, а.

Общий характер этих зависимостей аналогичен характеру обратной ветви вольт-амперной характеристики диода, так как большая часть напряжения источника питания выходной цепи падает на $p-n$ -переходе коллектора, включенном в обратном направлении. Однако в отличие от выходных характеристик схемы с общей базой выходные характеристики схемы с общим эмиттером имеют значительно больший наклон, т. е. наблюдается большая зависимость тока коллектора от напряжения на коллекторе. Причины этого явления поясняет рис. 4.18, б, на котором видно, что с увеличением напряжения на коллекторе при постоянном токе базы увеличивается ток эмиттера и соответственно растет ток коллектора. Напомним, что основная составляющая тока базы (рекомбинационная) приблизительно пропорциональна общему числу дырок в базе и, следовательно, пропорциональна площади под кривыми распределения концентрации дырок в базе.

Смещение выходных статических характеристик вверх происходит в соответствии с формулой (4.40) и связано с увеличением тока эмиттера при условии постоянства напряжения на коллекторе и увеличения тока базы. Необходимо отметить различное расстояние между выходными характеристиками по оси токов при равных приращениях тока базы. При малых токах базы кривые располагаются часто, при больших токах базы — редко, а затем опять часто. Неравномерность расположения характеристик связана с измене-

нием коэффициента передачи постоянного тока базы ($h_{21Э}$) при изменении тока (см. рис. 4.10).

В транзисторе, включенном по схеме с общим эмиттером, ток $I_{КЭ0}$ превосходит ток $I_{КБ0}$ для схемы с общей базой ($I_{КЭ0} = \frac{I_{КБ0}}{1 - |h_{21Б}|}$). Объясняется это тем, что при токе базы, равном нулю, и при подаче напряжения на коллектор относительно эмиттера (в схеме с общим эмиттером) p - n -переход эмиттера оказывается включенным под небольшое прямое напряжение. Поэтому обратный ток коллектора из-за инжектированных из эмиттера в базу дырок увеличивается.

При напряжении на коллекторе, равном нулю, т. е. при коротком замыкании коллектора с эмиттером, и при наличии тока базы, или, другими словами, при подаче напряжения на базу относительно эмиттера, p - n -переход коллектора оказывается включенным в прямом направлении, так как он, по существу, включен параллельно p - n -переходу эмиттера. При токе коллектора, равном нулю, и токе базы, отличным от нуля, из эмиттера происходит инжекция дырок, что обеспечивает около коллектора в базе их концентрацию, превосходящую равновесное значение. Если же концентрация неосновных носителей в базе около p - n -перехода превышает равновесную, то это соответствует прямому включению перехода. Таким образом, транзистор работает в режиме насыщения при напряжении на коллекторе, равном нулю, и даже при небольшом запирающем напряжении на коллекторе относительно эмиттера.

Значение этого напряжения определяет сопротивление насыщения транзистора (см. § 4.6). На его величину влияют соотношение падений напряжения на эмиттерном и коллекторном переходах, объемное сопротивление коллектора и сопротивление базы.

На объемном сопротивлении коллектора при прохождении тока создается падение напряжения, направленное так, что оно открывает коллекторный переход. Из-за этого напряжение на внешнем выводе коллектора, соответствующее выходу транзистора из режима насыщения, возрастает.

Токи в базе транзистора направлены так, что создаваемое ими падение напряжения запирает наиболее удаленные от вывода базы части эмиттерного и коллекторного переходов (см. § 4.5). Вследствие этого даже тогда, когда средняя часть коллекторного перехода уже запирается, его периферические области еще остаются открытыми, и через них проходит ток к выводу базы (см. рис. 4.7, а). В результате ток эмиттера замыкается не через внешний вывод коллектора, а через область коллектора и вывод базы. Такое явление также приводит к росту напряжения, при котором транзистор выходит из режима насыщения. Диапазон этих значений напряжения коллектора тем больше, чем больше ток базы. Соответственно выходные статические характеристики транзистора в схеме с общим эмиттером при малых напряжениях на коллекторе и при $I_B \neq 0$ заходят в четвертый квадрант.

При изменении направления входного (базового) тока, т. е. при изменении полярности постоянного напряжения на базе относительно эмиттера уменьшается значение тока эмиттера, что приводит к уменьшению коэффициента передачи тока эмиттера. Следовательно, коэффициент передачи тока эмиттера достигнет величины, равной единице, при больших значениях напряжения на коллекторе, чем $U(L)_{кэ0}$, соответствующее $I_B = 0$. В этих условиях пробой транзистора будет происходить при напряжениях, больших $U(L)_{кэ0}$, но не превышающих $U_{проб}$.

Однако после того, как наступил пробой, ток через эмиттерный переход существенно возрастает, что является причиной роста коэффициента передачи тока эмиттера. Теперь для продолжающегося увеличения тока коллектора, т. е. для выполнения условия $|h_{21Б}| = 1$, можно уменьшить коэффициент умножения M , уменьшив напряжение на коллекторе. Коэффициент передачи тока эмиттера будет при этом сохраняться равным единице, так как растет ток эмиттера. Таким образом, на статических характеристиках появляются участки с отрицательным дифференциальным сопротивлением (рис. 4.19).

Характеристики передачи тока. Семейство статических характеристик передачи тока транзистора, включенного по схеме с общим эмиттером, показано на рис. 4.20.

Общий характер этих зависимостей свидетельствует о том, что в транзисторе $I_K = I_{кэ0} + h_{21Э} I_B$. В связи с большей зависимостью коэффициента передачи тока базы $h_{21Э}$ от режима работы транзистора по сравнению с аналогичной зависимостью коэффициента передачи постоянного тока эмиттера характеристики передачи тока в схеме с общим эмиттером отличаются большей нелинейностью.

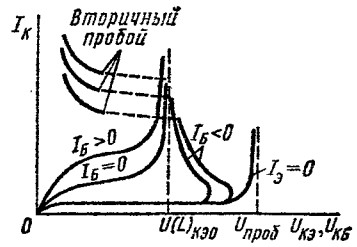


Рис. 4.19. Выходные статические характеристики транзистора, включенного по схеме с общим эмиттером, при больших коллекторных напряжениях

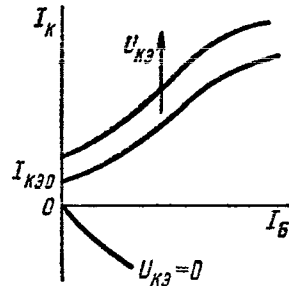


Рис. 4.20. Статические характеристики передачи тока транзистора, включенного по схеме с общим эмиттером

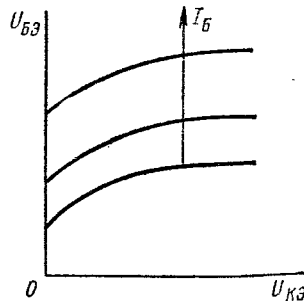


Рис. 4.21. Статические характеристики обратной связи транзистора, включенного по схеме с общим эмиттером

Смещение статических характеристик передачи тока вверх при увеличении напряжения на коллекторе также более значительно по сравнению со смещением подобных характеристик схемы с общей базой, так как связано в данном случае не с уменьшением рекомбинации дырок при уменьшении толщины базы, а с увеличением тока эмиттера при постоянной величине тока базы (см. рис. 4.18, б).

При напряжении на коллекторе, отличном от нуля, характеристики передачи тока выходят не из начала координат, а из точек на оси ординат, соответствующих токам $I_{КЭ0}$.

Характеристики обратной связи. Семейство статических характеристик обратной связи транзистора, включенного по схеме с общим эмиттером, приведено на рис. 4.21.

Общий характер этих зависимостей можно выяснить путем анализа рис. 4.18, б, либо простым графическим перестроением семейства входных статических характеристик схемы с общим эмиттером.

Смещение статических характеристик обратной связи при увеличении тока базы также не требует специального пояснения.

Практически больший интерес представляют статические характеристики схемы с общим эмиттером, так как в них указываются значения токов базы. Для схемы с общей базой ток базы не измеряется, а определить его по токам эмиттера и коллектора с достаточной точностью невозможно, так как токи эмиттера и коллектора мало отличаются.

§ 4.9. РАБОТА ТРАНЗИСТОРА НА МАЛОМ ПЕРЕМЕННОМ СИГНАЛЕ

Рассмотрим подробнее факторы, влияющие на работу транзистора при условии, что на его переходы помимо постоянных поданы малые переменные напряжения. При этом в транзисторе проходят постоянные и малые переменные токи. Под малыми переменными напряжениями и токами будем понимать такие, при которых связь между ними остается линейной.

Как отмечалось в § 4.1, при работе транзистора на малом переменном сигнале его свойства определяются процессами переноса, накопления и рассасывания носителей, а также перезарядом емкостей переходов. Чтобы рассмотреть это подробнее, представим себе, что через отдельные элементы транзистора проходит сигнал в виде небольшого прямоугольного импульса тока, либо небольшого тонкого пакета носителей заряда.

Барьерная емкость эмиттера. Пусть прямоугольный импульс тока проходит через цепь эмиттера транзистора, включенного по схеме с общей базой. Исходя из принципа работы транзистора, для того чтобы получить усиление, в базу должны быть инжектированы носители заряда, для чего требуется изменение напряжения на эмиттерном переходе. Однако в связи с тем, что эмиттерный переход обладает барьерной емкостью $C_{э, бар}$ и цепь, в которую он включен, — конечным сопротивлением, напряжение на эмиттерном переходе

изменится не мгновенно; значит, не мгновенно будет изменяться и инжекция носителей (рис. 4.22). Таким образом, наличие емкости эмиттера приводит к тому, что инжекция носителей в базу как бы замедляется и, если сигналы следуют один за другим через интервалы времени, сравнимые с этим замедлением, соответствующие пакеты инжектированных носителей заряда будут сливаться. Это первая причина искажения высокочастотных сигналов.

Пролет носителей через базу. Представим себе, что со стороны эмиттера в базу входит пакет инжектированных носителей заряда (рис. 4.23). Эти носители нарушают нейтральность базы и в нее из невыпрямляющего контакта должно войти соответствующее число основных носителей. Это происходит довольно быстро — за время диэлектрической релаксации. Следовательно, через вывод базы проходит импульс тока, равный импульсу тока эмиттера. Далее пакет носителей начинает двигаться в базе к коллектору. Если в базе носители не рекомбинируют и поле там неизменно, при таком движении отсутствует ток во всех электродах транзистора, так как инжектированный в базу заряд нейтрализован.

За время движения через базу пакет носителей заряда размывается из-за диффузии. Действительно, наличие градиентов концентрации приводит к тому, что носители, находящиеся на переднем крае такого пакета, обгоняют его, а на заднем крае отстают, и к коллектору подходит уже размытый пакет носителей.

Размытие пакета носителей заряда усиливается тем, что они проходят в транзисторе разные пути из-за непараллельности электронно-дырочных переходов эмиттера и коллектора. Особенно сильно это явление сказывается у краев эмиттера.

Когда пакет носителей доходит до коллекторного перехода, неосновные носители вытягиваются из базы, ее электрическая ней-

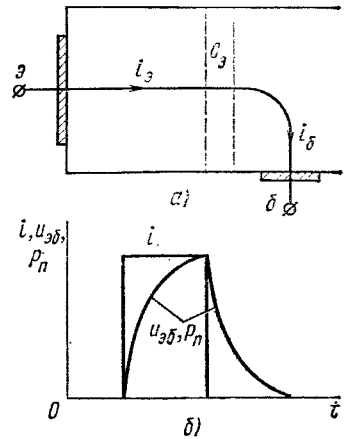


Рис. 4.22. Искажение сигнала из-за влияния барьерной емкости эмиттера:

а — путь тока; б — форма импульса тока и напряжения на эмиттерном переходе (граничной концентрации неосновных носителей в базе около эмиттера)

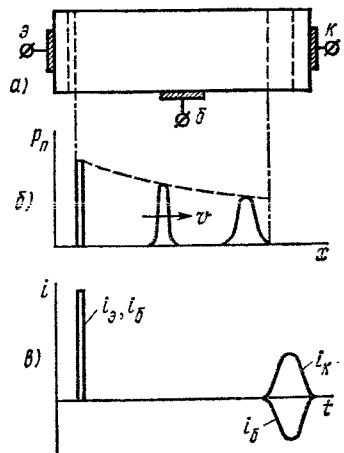


Рис. 4.23. Искажение сигнала при прохождении носителей заряда через базу транзистора:

а — структура транзистора; б — движение пакета носителей в базе; в — импульсы токов

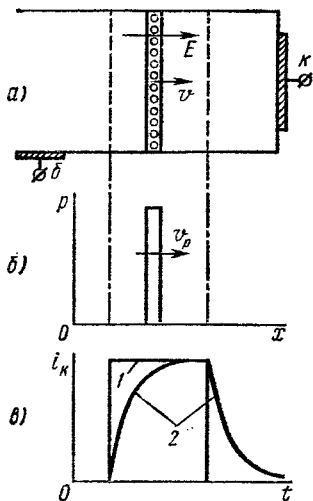


Рис. 4.24. Искажение сигнала при прохождении носителей через p - n -переход коллектора:

a — структура коллекторного перехода и прилегающих к нему областей; *б* — распределение носителей в коллекторном переходе; δ — импульсы тока коллектора (1 — без учета постоянной времени цепи коллектора; 2 — с учетом постоянной времени цепи коллектора)

тральность нарушается; следовательно, основные носители тоже выходят из базы. Если в базе не происходила рекомбинация, вышедший из базы заряд оказался бы равным вошедшему, т. е. среднее значение тока базы оказалось бы равным нулю. Однако через вывод базы при этом протекает переменный ток, связанный с изменением заряда носителей в базе. Так как в базе происходит рекомбинация, выходящий из нее заряд несколько меньше входящего, что и дает некоторое ненулевое значение среднего тока базы.

Когда время пролета носителей заряда через базу $t_{\text{прол}}$ оказывается большим, размытие импульсов тоже увеличивается. Тогда, если сигнал представляет собой последовательность импульсов, они перестают различаться в токе коллектора. Это вторая причина искажения высокочастотных сигналов транзистором.

Пролет носителей через область объемного заряда коллектора. Рассмотрим, как появляется ток в цепи коллектора, когда в область объемного заряда коллекторного перехода входит пакет носителей заряда (рис. 4.24). Природа

проходящего при этом тока состоит в том, что вошедшие в область объемного заряда носители искажают распределение электрического поля в нем. Изменение электрического поля приводит к появлению токов смещения, которые замыкаются через цепь коллектора. Такое изменение электрического поля, а значит, и прохождение тока в цепи коллектора имеет место до тех пор, пока носители находятся в области объемного заряда (см. § 3.27). Следовательно, длительность импульса тока в цепи коллектора определяется временем пролета носителей через p - n -переход коллектора (рис. 4.24, в). Если же сигналы следуют с интервалами времени, меньшими, чем время пролета через p - n -переход коллектора, они оказываются неразличимыми. Это третья причина искажения высокочастотных сигналов транзистором.

Барьерная емкость коллекторного перехода. Приведенные рассуждения основывались на предположении, что как только носители попадают в коллекторный переход, сразу в цепи коллектора появляется ток. Однако этот ток, как отмечалось, может появиться, если изменится электрическое поле в p - n -переходе коллектора. Но так как коллекторный переход обладает некоторой емкостью, а цепи, в которые он включен, — сопротивлением, появление тока происхо-

дит не мгновенно — сначала должно произойти изменение заряда барьерной емкости перехода. Это также приводит к искажению сигнала (рис. 4.24, в).

Таким образом, работа транзистора на переменном сигнале определяется как прохождением активных токов, аналогичных по физической природе постоянным токам, так и прохождением емкостных токов, связанных с наличием барьерных емкостей и накоплением зарядов в базе. На все эти явления оказывают влияние времена пролета носителей. Следовательно, для малого переменного сигнала транзистор представляет собой довольно сложный активный элемент электрической схемы, параметры которого оказываются комплексными и зависящими от частоты.

§ 4.10. МАЛОСИГНАЛЬНЫЕ ПАРАМЕТРЫ

Если переменные напряжения на переходах транзистора достаточно малы, токи в нем оказываются линейными функциями этих напряжений. Транзистор можно рассматривать как линейный четырехполюсник (рис. 4.25). При этом два внешних вывода четырехполюсника считают входными, соответствующие им ток и напряжение обозначают I_1 и U_1 . Два других вывода являются выходными, соответствующие им ток и напряжение обозначают I_2 и U_2 . За положительное принимают направление токов, входящих в четырехполюсник.

Чтобы показать связь между величинами U_1 и U_2 , I_1 и I_2 , можно составить шесть систем уравнений. Однако практически применяют только три системы. В первой из них напряжения рассматривают как линейные функции токов:

$$\left. \begin{aligned} U_1 &= z_{11}I_1 + z_{12}I_2; \\ U_2 &= z_{21}I_1 + z_{22}I_2. \end{aligned} \right\} \quad (4.55)$$

Коэффициенты z_{ik} , имеющие размерность сопротивления и являющиеся комплексными, можно выразить через токи и напряжения, измеренные в режиме холостого хода, следующим образом:

$$\left. \begin{aligned} z_{11} &= \frac{U_1}{I_1} \Big|_{I_2=0}; & z_{12} &= \frac{U_1}{I_2} \Big|_{I_1=0}; \\ z_{21} &= \frac{U_2}{I_1} \Big|_{I_2=0}; & z_{22} &= \frac{U_2}{I_2} \Big|_{I_1=0}. \end{aligned} \right\} \quad (4.56)$$

Здесь, как и в дальнейшем, индекс 11 (читается «один-один») означает входной параметр (характеризующий входную цепь), индекс 12 (читается «один-два») — параметр обратной связи, индекс 21 (читается «два-один») — параметр прямой передачи и индекс 22 (читается «два-два») — выходной параметр.

Для получения режима холостого хода в цепь включают сопротивление, значительно большее, чем соответ-

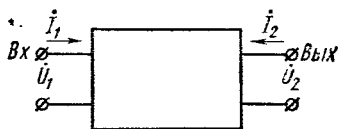


Рис. 4.25. Четырехполюсник, эквивалентный транзистору

ствующее сопротивление четырехполюсника (входное или выходное). При экспериментальном определении параметров транзистора необходимо обеспечить питание его электродов постоянным напряжением либо через очень большое активное сопротивление от достаточно высоковольтного источника питания, либо через индуктивные элементы.

Осуществление режима холостого хода в цепи эмиттера или базы (для схемы с общим эмиттером) не представляет особого труда, так как внутреннее сопротивление открытого перехода мало, а создание режима холостого хода в цепи коллектора (выходной) затруднено тем, что внутреннее сопротивление при этом очень велико (достигает нескольких мегом). Следовательно, определить экспериментально z -параметры транзистора трудно.

Если токи транзистора рассматривать как линейные функции напряжений, получаем систему уравнений:

$$\left. \begin{aligned} \dot{I}_1 &= y_{11}\dot{U}_1 + y_{12}\dot{U}_2; \\ \dot{I}_2 &= y_{21}\dot{U}_1 + y_{22}\dot{U}_2. \end{aligned} \right\} \quad (4.57)$$

Коэффициенты y_{ik} , имеющие размерность проводимости и тоже комплексные, определяют в режиме короткого замыкания:

$$\left. \begin{aligned} y_{11} &= \left. \frac{\dot{I}_1}{\dot{U}_1} \right|_{\dot{U}_2=0}; & y_{12} &= \left. \frac{\dot{I}_1}{\dot{U}_2} \right|_{\dot{U}_1=0}; \\ y_{21} &= \left. \frac{\dot{I}_2}{\dot{U}_1} \right|_{\dot{U}_2=0}; & y_{22} &= \left. \frac{\dot{I}_2}{\dot{U}_2} \right|_{\dot{U}_1=0}. \end{aligned} \right\} \quad (4.58)$$

Получение режима короткого замыкания состоит в том, что исследуемую цепь шунтируют сопротивлением, значительно меньшим внутреннего сопротивления соответствующей цепи. Учитывая необходимость обеспечения питания электродов транзистора постоянным напряжением, такое шунтирование можно производить емкостью.

Режим короткого замыкания можно легко осуществить в цепи коллектора, где внутреннее сопротивление транзистора велико. Выполнение режима короткого замыкания в цепи эмиттера затруднено, особенно в области низких частот. Это ограничивает применение y -параметров для описания свойств транзистора.

Во избежание перечисленных трудностей чаще всего для описания свойств транзистора используют так называемую смешанную систему:

$$\left. \begin{aligned} \dot{U}_1 &= h_{11}\dot{I}_1 + h_{12}\dot{U}_2; \\ \dot{I}_2 &= h_{21}\dot{I}_1 + h_{22}\dot{U}_2. \end{aligned} \right\} \quad (4.59)$$

Для определения h -параметров необходим режим короткого замыкания в выходной цепи и режим холостого хода во входной. Физический смысл h -параметров довольно прост:

$$\begin{aligned}
 h_{11} &= \left. \frac{\dot{U}_1}{I_1} \right|_{\dot{U}_2=0} && \text{— входное сопротивление при коротком замыкании} \\
 &&& \text{выходной цепи;} \\
 h_{12} &= \left. \frac{\dot{U}_1}{\dot{U}_2} \right|_{I_1=0} && \text{— коэффициент обратной связи по напряжению при} \\
 &&& \text{холостом ходе во входной цепи;} \\
 h_{21} &= \left. \frac{I_2}{I_1} \right|_{\dot{U}_2=0} && \text{— коэффициент передачи тока при коротком замы-} \\
 &&& \text{кании выходной цепи;} \\
 h_{22} &= \left. \frac{I_2}{\dot{U}_2} \right|_{I_1=0} && \text{— выходная проводимость при холостом ходе во} \\
 &&& \text{выходной цепи.}
 \end{aligned}$$

Преимущества h -параметров состоят в удобстве их экспериментального определения. Кроме того, h -параметры измеряют в режимах, близких к режимам работы транзисторов в практических схемах. Однако для расчета электрических схем часто удобнее применять другие параметры (например, y).

Переход от одной системы параметров к другой довольно прост. С этой целью уравнения системы, от которой осуществляется переход, следует решить относительно величин, являющихся функциями в системе, к которой осуществляется переход. Полученные коэффициенты при токах или напряжениях и дадут формулы перехода. Например, если необходимо перейти от системы y -параметров к системе z -параметров, уравнения (4.57) надо решить относительно \dot{U}_1 и \dot{U}_2 (функции в системе z). Полученные выражения имеют вид:

$$\left. \begin{aligned}
 \dot{U}_1 &= \frac{y_{22}}{\Delta y} I_1 - \frac{y_{12}}{\Delta y} I_2; \\
 \dot{U}_2 &= -\frac{y_{21}}{\Delta y} I_1 + \frac{y_{11}}{\Delta y} I_2,
 \end{aligned} \right\} \quad (4.60)$$

где $\Delta y = y_{11}y_{22} - y_{12}y_{21}$ — определитель системы y .

Сравнив систему (4.60) с (4.55), получим соотношения для перехода от одной системы параметров к другой. Найденные аналогичным путем формулы сведены в табл. 4.1.

Значения параметров транзистора, представленного в виде четырехполюсника, зависят от схемы его включения. Однако если эти параметры известны для какой-либо одной схемы, сравнительно легко произвести пересчет для другой. Для этого надо заменить напряжения и токи (имея в виду правило знаков), учитывая, что в транзисторе

$$I_9 + I_6 + I_k = 0; \quad (4.61)$$

$$\dot{U}_{69} + \dot{U}_{к6} + \dot{U}_{9к} = 0. \quad (4.62)$$

Сделав необходимые подстановки и преобразовав уравнения, получаем формулы перехода как коэффициенты в уравнениях. Например, если известны y -параметры транзистора для схемы с общей базой, а требуется найти их для схемы с общим эмиттером, в уравнения

$$I_9 = y_{116} \dot{U}_{96} + y_{126} \dot{U}_{к6}; \quad (4.63)$$

$$I_k = y_{216} \dot{U}_{96} + y_{226} \dot{U}_{к6} \quad (4.64)$$

Таблица 4.1

ФОРМУЛЫ ПЕРЕХОДА МЕЖДУ СИСТЕМАМИ
ПАРАМЕТРОВ

	$ z $	$ y $	$ h $
$ z $	$z_{11} \quad z_{12}$ $z_{21} \quad z_{22}$	$\frac{y_{22}}{\Delta_y} \quad -\frac{y_{12}}{\Delta_y}$ $-\frac{y_{21}}{\Delta_y} \quad \frac{y_{11}}{\Delta_y}$	$\frac{\Delta_h}{h_{22}} \quad \frac{h_{12}}{h_{22}}$ $-\frac{h_{21}}{h_{22}} \quad \frac{1}{h_{22}}$
$ y $	$\frac{z_{22}}{\Delta_z} \quad -\frac{z_{12}}{\Delta_z}$ $-\frac{z_{21}}{\Delta_z} \quad \frac{z_{11}}{\Delta_z}$	$y_{11} \quad y_{12}$ $y_{21} \quad y_{22}$	$\frac{1}{h_{11}} \quad -\frac{h_{12}}{h_{11}}$ $\frac{h_{21}}{h_{11}} \quad \frac{\Delta_h}{h_{11}}$
$ h $	$\frac{\Delta_z}{z_{22}} \quad \frac{z_{12}}{z_{22}}$ $-\frac{z_{21}}{z_{22}} \quad \frac{1}{z_{22}}$	$\frac{1}{y_{11}} \quad -\frac{y_{12}}{y_{11}}$ $\frac{y_{21}}{y_{11}} \quad \frac{\Delta_y}{y_{11}}$	$h_{11} \quad h_{12}$ $h_{21} \quad h_{22}$

подставляем токи и напряжения с помощью выражений (4.61) и (4.62). После преобразований получим:

$$\dot{I}_6 = (y_{116} + y_{126} + y_{216} + y_{226}) \dot{U}_{69} - (y_{126} + y_{226}) \dot{U}_{к9}, \quad (4.65)$$

$$\dot{I}_к = -(y_{216} + y_{226}) \dot{U}_{69} + y_{226} \dot{U}_{к9}, \quad (4.66)$$

где $\dot{U}_{69} = -\dot{U}_{96}$.

Следовательно,

$$\left. \begin{aligned} y_{119} &= y_{116} + y_{126} + y_{216} + y_{226}; \\ y_{129} &= -(y_{126} + y_{226}); \\ y_{219} &= -(y_{216} + y_{226}); \\ y_{229} &= y_{226}. \end{aligned} \right\} \quad (4.67)$$

Аналогично выводят соотношения и для других параметров.

§ 4.11. ЭКВИВАЛЕНТНЫЕ СХЕМЫ

Рассмотрение транзистора как активного линейного четырехполюсника (§ 4.10) удобно для расчета электрических схем. Однако оно имеет и ряд недостатков, которые связаны прежде всего с тем, что параметры четырехполюсника вводят в известной степени формально и каждый из них может отражать влияние сразу нескольких физических процессов. Поэтому получаются сложные зависимости параметров четырехполюсника от режима работы транзистора (постоянных напряжений и токов), от частоты и температуры. Чтобы упростить эти зависимости, свойства транзистора при малом переменном сигнале описывают с помощью эквивалентных схем. Под

эквивалентной понимают электрическую схему, составленную из линейных элементов электрических цепей (сопротивлений, емкостей, индуктивностей), генераторов тока или напряжения), которая по своим свойствам при данном сигнале (например, при малом переменном) не отличается от реального объекта (транзистора). Графическое изображение эквивалентных схем позволяет более экономно зафиксировать основные соотношения. При расчетах с помощью эквивалентных схем сначала определяют токи и напряжения в самой схеме и затем переходят к каким-то другим параметрам, например параметрам четырехполюсника.

Никакая эквивалентная схема из конечного числа элементов не может быть полностью эквивалентной реальному транзистору, т. е. все эквивалентные схемы оказываются приближенными. Чем проще эквивалентная схема, тем меньше она содержит элементов, тем легче ею пользоваться, но обычно тем менее точно она отражает свойства реального транзистора.

По способу построения различают формальные и физические эквивалентные схемы.

Формальные эквивалентные схемы строят на основе описания транзистора с помощью уравнений четырехполюсника (рис. 4.26). Как видно из рисунка, каждая схема содержит по четыре элемента — два сопротивления (комплексных) и два генератора тока или напряжения. Такие эквивалентные схемы не имеют никаких преимуществ по сравнению с описанием транзистора как четырехполюсника путем задания его параметров (или соответствующих уравнений).

Формальные эквивалентные схемы можно представить в разной форме: например, построить так, чтобы в схеме был только один активный элемент (рис. 4.27). Обычно эти схемы

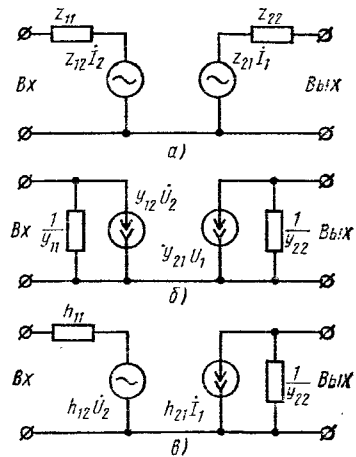


Рис. 4.26. Формальные эквивалентные схемы транзистора:

а — для системы z -параметров; б — для системы y -параметров; в — для системы h -параметров

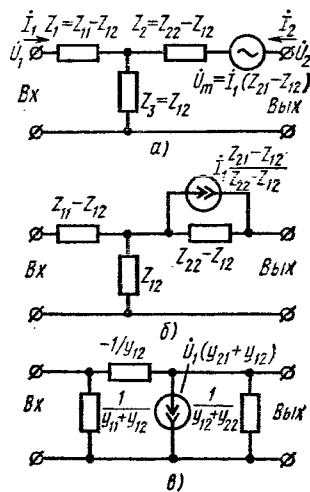


Рис. 4.27. Одногенераторные формальные эквивалентные схемы:

а — Т-образная эквивалентная схема с генератором в. д. с.; б — Т-образная эквивалентная схема с генератором тока; в — П-образная эквивалентная схема с генератором тока

по способу начертания подразделяют на Т- и П-образные. Чтобы установить связь элементов однопораторных формальных схем с параметрами четырехпоросника, можно записать выражения для токов и напряжений в них, а затем сопоставить со значениями токов и напряжений четырехпоросника. Например, для Т-образной эквивалентной схемы с генератором э. д. с. (рис. 4.27, а) в режиме холостого хода на выходе

$$\left. \begin{aligned} \dot{U}_1 &= \dot{I}_1(z_1 + z_3), \\ \dot{U}_2 &= \dot{U}_m + \dot{I}_1 z_3, \end{aligned} \right\} \quad (4.68)$$

в режиме холостого хода на входе

$$\left. \begin{aligned} \dot{U}_1 &= \dot{I}_2 z_3, \\ \dot{U}_2 &= \dot{I}_2(z_2 + z_3). \end{aligned} \right\} \quad (4.69)$$

Отсюда

$$\left. \begin{aligned} \left. \begin{aligned} \frac{\dot{U}_2}{\dot{I}_1} \Big|_{\dot{I}_2=0} &= z_3 + \frac{\dot{U}_m}{\dot{I}_1} = z_{21}; & \frac{\dot{U}_1}{\dot{I}_1} \Big|_{\dot{I}_2=0} &= z_1 + z_2 = z_{11}; \\ \frac{\dot{U}_1}{\dot{I}_2} \Big|_{\dot{I}_1=0} &= z_3 = z_{12}; & \frac{\dot{U}_2}{\dot{I}_2} \Big|_{\dot{I}_1=0} &= z_2 + z_3 = z_{22}. \end{aligned} \right\} \end{aligned} \right\} \quad (4.70)$$

После преобразования получим значения, указанные на рисунке. Аналогично получают параметры и других однопораторных схем.

Физические эквивалентные схемы составляют на основании физических соображений для определенных типов конструкций транзистора, для определенного частотного диапазона, ориентируясь на определенную схему включения транзистора (с общим эмиттером, общей базой, общим коллектором). Каждый вывод физической эквивалентной схемы соответствует электроду транзистора. Заметим, что в формальных эквивалентных схемах различают только входные и выходные зажимы, независимо от того, какими электродами транзистора они являются.

Физическую эквивалентную схему строят, выделяя мысленно некоторые части в транзисторе и рассматривая отдельно процессы в этих частях.

За основу построения обычно берут формальную эквивалентную схему идеализированного транзистора, называемого *одномерной теоретической моделью*. При изучении одномерной теоретической модели транзистора считают, что носители заряда в нем движутся по параллельным путям, а поверхностная рекомбинация только изменяет время жизни носителей. Кроме того, в одномерной теоретической модели не учитывают влияние объемных сопротивлений и токов, проходящих через барьерные емкости переходов. При таких предположениях получается, что параметры формальной эквивалентной схемы довольно просто выражаются через конструктивные параметры идеализированного транзистора (толщину базы), режим его работы, свойства материала.

К одномерной теоретической модели транзистора добавляют элементы, учитывающие другие процессы, например падения напряжений на объемных сопротивлениях, токи через емкости и т. д.

§ 4.12. ЭКВИВАЛЕНТНАЯ СХЕМА ОДНОМЕРНОЙ ТЕОРЕТИЧЕСКОЙ МОДЕЛИ

Как отмечалось, параметры формальной эквивалентной схемы легко можно выразить через параметры четырехполюсника, которые в свою очередь могут быть найдены по значениям токов и напряжений в транзисторе. Поэтому для построения формальной эквивалентной схемы одномерной теоретической модели транзистора нужно знать переменные составляющие токов и напряжений. Строгий расчет этих составляющих производится примерно так же, как и постоянных составляющих (см. § 4.4). Однако он сложнее, так как в данном случае менее обоснованно предположение о постоянстве плотности тока во всех сечениях базы, особенно сильно это проявляется в области высоких частот.

Для дальнейших расчетов воспользуемся упрощенным приемом. Учитывая, что все параметры формальной эквивалентной схемы комплексные, изобразим их как результат параллельного соединения сопротивлений и емкостей. Полученная эквивалентная схема показана на рис. 4.28. Эта эквивалентная схема должна быть справедлива в области низких частот. Тогда активные сопротивления эквивалентной схемы можно получить как отношения приращений напряжений в цепях транзистора к вызвавшим их приращениям токов (такие приращения играют на низких частотах роль переменных сигналов).

Согласно эквивалентной схеме получаем:

сопротивление эмиттера

$$r_e = \left. \frac{du_{эб}}{di_э} \right|_{i_k = \text{const}} - \left. \frac{du_{эб}}{di_k} \right|_{i_э = \text{const}}; \quad (4.71)$$

сопротивление коллектора

$$r_k \approx \left. \frac{du_{кб}}{di_k} \right|_{i_э = \text{const}}; \quad (4.72)$$

диффузионное сопротивление базы

$$r_b = \left. \frac{du_{эб}}{di_k} \right|_{i_э = \text{const}}; \quad (4.73)$$

коэффициент передачи тока эмиттера одномерной теоретической модели транзистора

$$\alpha = \left. \frac{di_k}{di_э} \right|_{u_{кб} = \text{const}}. \quad (4.74)$$

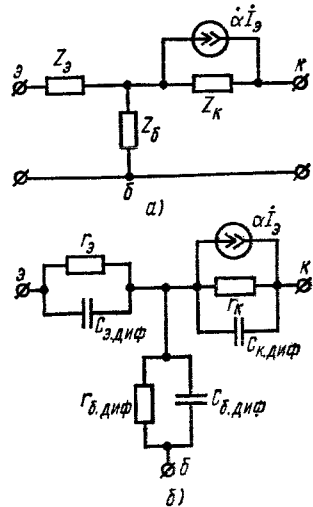


Рис. 4.28. Эквивалентная схема одномерной теоретической модели транзистора:

а — с комплексными сопротивлениями; б — с активными сопротивлениями и диффузионными емкостями

Подставляя значения $F_A(\xi)$ и $\operatorname{tg} \alpha$ в формулу (3.195), будем иметь для $\alpha = 14^\circ$

$$\tau_1 = \sqrt[3]{\left\{ 2 \cdot 0,495 + \sqrt{\left(\frac{0,73}{0,245}\right)^2} \right\}^2} = \sqrt[3]{(9,50 + 5,15)^2} \approx 6$$

и по формуле (3.196) значение интервала аппроксимации

$$\Delta t_1 = 6 - \frac{0,73}{0,249} \approx 3 \text{ сек.}$$

Для прямых, расположенных под углами 9° и 23° соответственно, получим $\tau_1 = 7,8$, $\Delta t_1 = 3,2 \text{ сек.}$, $\tau_1 = 4,5$, $\Delta t_1 = 2,8 \text{ сек.}$ Сравнивая полученные аппроксимации (рис. 3.12), выбираем интервал $\Delta t_1 = 3 \text{ сек.}$, что соответствует $\alpha = 14^\circ$. При этом

$$b = \operatorname{tg} \alpha = 0,245; \quad a = 0,73 - 2 \cdot 0,245 = 0,24.$$

Для 2-го участка имеем $F_A(t_1) = 1,47$. Строим прямые под углами α в 48° , 50° и 57° . Легко видеть, что для этого участка можно получить удовлетворительную аппроксимацию на интервале длительностью около $3 \div 3,5 \text{ сек.}$, а так как коэффициент $F(t)$ возрастает по сравнению с 1-м участком почти в 3 раза, то можно ожидать заметного уменьшения периода колебаний. По этой причине целесообразно найти сразу третий нуль процесса, положив в формуле (3.195) коэффициент равным 6π . Расчеты по формулам (3.195) и (3.196) в этом случае дадут для углов в 48° , 50° и 57° соответственно

$$\Delta t_2 = 3,25 \text{ сек.}, \quad \Delta t_2 = 3,20 \text{ сек.}, \quad \Delta t_2 = 3,16 \text{ сек.}$$

По рис. 3.12 выбираем участок аппроксимации $\Delta t_2 = 3,2 \text{ сек.}$

Рассмотрим предварительную оценку процесса. Прежде всего определяем λ_k для 1-го участка:

$$\lambda_1 = \frac{a_1}{b_1} + \xi = 3 \text{ сек.}$$

Так как $\Delta t_1 = 3 \text{ сек.}$, то

$$t_1 = \xi + \Delta t_1 = 5 \text{ сек.}, \quad \tau_1 = 6, \quad t_{m1} = \frac{t_1 - \xi}{2} + \xi = 3,5; \quad \tau_{m1} = 4,5.$$

Максимум первой полуволны процесса определяем по формуле (3.193), используя найденные значения t_{m1} и τ_{m1} :

$$g_{m1}(t, \xi) = \frac{e^{-\frac{1}{2} p \xi - \frac{1}{2} p t_{m1}}}{\sqrt[4]{b_1^2 \tau_{m1} \lambda_1}} = 0,91.$$

На 2-м участке

$$\lambda_2 = \frac{a_2}{b_2} + t_1 = 1 \text{ сек.}$$

По формуле (3.194) определяем значение τ_2

$$\tau_2^{\frac{3}{2}} = \frac{3\pi}{2 \sqrt{b_2}} + \lambda_2^{\frac{3}{2}} = 4,8,$$

откуда $\tau_2 = 2,84$ и $t_2 = 6,89 \text{ сек.}$

Далее находим примерное положение второго максимума

$$t_{m2} = \frac{t_2 - t_1}{2} + t_1 = 5,9 \text{ сек.}$$

По формуле (3.197) находим значение коэффициента β_0

$$\beta_0 = -\sqrt[4]{\frac{\tau_1}{\lambda_1}} = -1,19.$$

Максимум второй полуволны процесса определяем по формуле (3.193), используя найденные величины t_{m2} и τ_{m2}

$$g_{m2}(t, \xi) = \beta_0 \frac{e^{-\frac{1}{2} p \xi - \frac{1}{2} p t_{m2}}}{\sqrt[4]{b_2^2 \tau_{m2} \lambda_2}} = -0,6.$$

Положение третьего нуля процесса нами уже определено выше. Оно характеризуется моментом времени $t_3 = 8,2$ сек. Затем находим значения t_{m3}

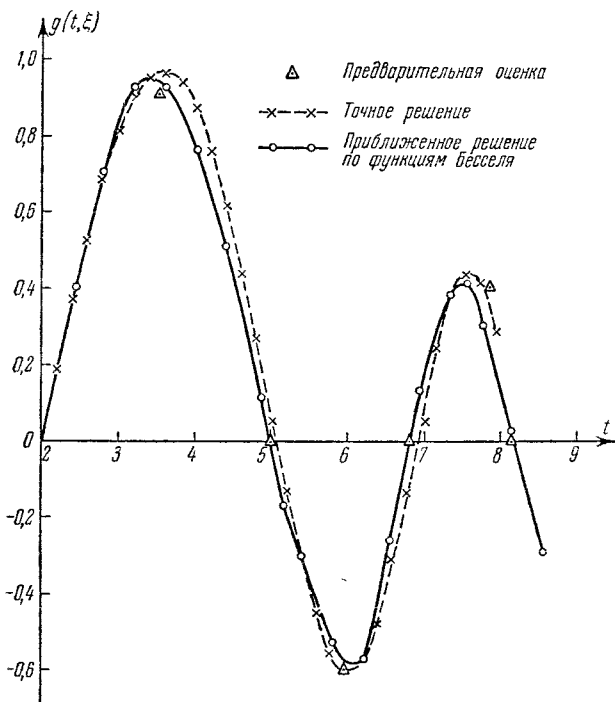


Рис. 3.13.

и τ_{m3} и по формуле (3.193) определяем максимум следующей полуволны процесса:

$$g_{m3}(t, \xi) = \beta_0 \frac{e^{-\frac{1}{2} p \xi - \frac{1}{2} p t_{m3}}}{\sqrt[4]{b_3^2 \tau_{m3} \lambda_2}} = 0,4.$$

Результаты расчетов по предварительной оценке процесса приведены в таблице 14, а соответствующие точки нанесены на рис. 3.13.

Таблица 14
Предварительная оценка процесса

Нули процесса	Максимумы процесса	Значения $g(t, \xi)$
2,0	—	0
—	3,5	0,91
5,0	—	0
—	5,9	-0,6
6,84	—	0
—	7,9	0,4
8,2	—	0

Рассмотрим приближенный метод построения процесса с помощью функций Бесселя. Искомый процесс определяется выражением

$$g(t, \xi) = e^{-\frac{1}{2} p(t-\xi)} r(t, \xi),$$

причем функция $r(t, \xi)$ находится по формуле (3.182).

Предварительные расчеты для 1-го участка:

$$\tau_1 = \frac{a_1}{b_1} + t = 0,98 + t, \quad e^{\frac{1}{2} p\xi} = 1,22,$$

$$\lambda_1 = \frac{a_1}{b_1} + \xi = 2,98, \quad \sqrt{\lambda_1} = 1,73, \quad J_{\frac{1}{3}}(y) = 0,57, \quad J_{-\frac{1}{3}}(y) = 0,10.$$

Результаты расчета процесса на 1-м участке сведены в таблицу 15, где обозначено $A = J_{\frac{1}{3}}(x)J_{-\frac{1}{3}}(y) - J_{-\frac{1}{3}}(x)J_{\frac{1}{3}}(y)$.

Таблица 15

Порядок расчета функции $g(t, \xi)$ для 1-го участка

t	τ	$\frac{2}{3} \frac{3}{\tau^2}$	x	$J_{\frac{1}{3}}(x)$	$J_{-\frac{1}{3}}(x)$	$\sqrt{\tau}$	A	$r(\tau, \lambda)$	$e^{-\frac{1}{2} p t}$	$g(t, \xi)$
2,0	2,98	3,46	1,71	0,57	-0,10	1,73	0	0	0,820	0
2,4	3,38	4,18	2,07	0,41	-0,12	1,84	1,1095	0,423	0,790	0,408
2,8	3,78	4,94	2,44	0,22	-0,29	1,95	0,1875	0,766	0,755	0,705
3,2	4,18	5,73	2,84	0,015	-0,40	2,05	0,2430	0,045	0,725	0,925
3,6	4,58	6,60	3,27	-0,16	-0,44	2,14	0,2345	0,050	0,700	0,900
4,0	4,98	7,45	3,69	-0,30	-0,40	2,24	0,1980	0,927	0,670	0,756
4,4	5,38	8,38	4,15	-0,38	-0,30	2,32	0,1335	0,648	0,645	0,510
4,8	5,78	9,30	4,60	-0,37	-0,14	2,41	0,0362	0,151	0,620	0,114
5,2	6,18	10,3	5,10	-0,28	+0,03	2,49	0,0451	0,235	0,596	-0,617

Аналогично рассчитывается и 2-й участок процесса. За начало участка примем время, равное $t_1 = 5$ сек.

Предварительные расчеты:

$$\tau_2 = \frac{a_2}{b_2} + t = -4,05 + t, \quad e^{\frac{1}{2}pt} = 1,65,$$

$$\lambda_2 = \frac{a_2}{b_2} + t_1 = 0,95, \quad \sqrt{\lambda_2} = 0,975.$$

Коэффициент β_0 находим непосредственно из графика предыдущего участка

$$\beta_0 = \frac{r(\tau_{k+1}, \lambda) - r(\tau_k, \lambda)}{\Delta t} = \frac{-0,235 - 0,151}{0,4} = -0,965.$$

Результаты расчета процесса на 2-м участке сведены в таблицу 16.

Таблица 16

Порядок расчета функции $g(t, \xi)$ для 2-го участка

t	τ	$\frac{3}{3-\tau^2}$	x	$\frac{J_1(x)}{3}$	$J - \frac{1}{3}(x)$	$\sqrt{\tau}$	A	$r(\tau, \lambda)$	$e^{-\frac{1}{2}pt}$	$g(t, \xi)$
5,00	0,95	0,616	0,763	0,725	0,770	0,975	0	0	0,607	0
5,4	1,35	1,045	1,295	0,690	0,378	1,162	0,257	0,353	0,583	-0,330
5,8	1,75	1,540	1,910	0,470	-0,020	1,323	0,377	0,589	0,560	-0,526
6,2	2,15	2,100	2,600	0,148	-0,333	1,465	0,356	0,667	0,538	-0,571
6,6	2,55	2,710	3,36	-0,200	-0,440	1,600	0,165	0,312	0,516	-0,257
7,0	2,95	3,37	4,17	-0,375	-0,290	1,720	-0,079	-0,16	0,498	0,128
7,4	3,35	4,10	5,08	-0,280	0,020	1,830	-0,230	-0,497	0,478	0,379
7,6	3,55	4,46	5,54	0,150	0,175	1,885	-0,243	-0,54	0,468	0,403
7,8	3,75	4,85	6,02	0,004	0,279	1,940	-0,171	-0,392	0,460	0,298
8,2	4,15	5,63	6,97	0,240	0,275	0,040	-0,015	-0,036	0,440	0,025
8,6	4,55	6,45	8,00	0,260	0,0350	2,130	0,175	0,440	0,424	-0,298
9,0	4,95	7,33	9,07	0,020	-0,220	2,222	0,175	0,460	0,406	-0,298

Таблица 17

Точное значение функции $g(t, \xi)$

t	$g(t, \xi)$	t	$g(t, \xi)$	t	$g(t, \xi)$	t	$g(t, \xi)$
2	0	3,6	0,960	5,2	-0,130	6,8	-0,142
2,2	0,194	3,8	0,935	5,4	-0,320	7,0	0,052
2,4	0,367	4,0	0,866	5,6	-0,453	7,2	0,238
2,6	0,525	4,2	0,753	5,8	-0,556	7,4	0,370
2,8	0,692	4,4	0,616	6,0	-0,600	7,6	0,431
3,0	0,810	4,6	0,442	6,2	-0,570	7,8	0,408
3,2	0,900	4,8	0,270	6,4	-0,480	8,0	0,283
3,4	0,950	5,0	0,052	6,6	-0,314		

Совершенно аналогично рассчитываются и последующие участки процесса.

Кривая функции $g(t, \xi)$, полученная расчетом, приведена на рис. 3.13. Для сравнения точности расчетов там же нанесена кривая, отвечающая

точному решению уравнения (3.198), определяемому формулой

$$g(t, \xi) = \frac{1}{\sqrt{q}} e^{-pt} \sin \left[\frac{\sqrt{q}}{p} (e^{pt} - e^{p\xi}) \right] \quad (3.198')$$

и таблицей 17.

§ 19. Система второго порядка. Применение аппроксимирующего решения

Если в уравнении

$$\frac{d^2 u}{dt^2} + F(t)u = 0 \quad (3.199)$$

коэффициент $F(t)$ изменяется незначительно относительно своего среднего большого значения (рис. 3.14), можно довольно просто найти приближенное значение импульсной переходной функции следующим путем.

Пусть частное решение некоторого однородного уравнения второго порядка имеет вид

$$u_1(t) = \frac{1}{\sqrt{N}} e^{-js(t)}, \quad (3.200)$$

причем

$$S(t) = \int_0^t N(t) dt. \quad (3.201)$$

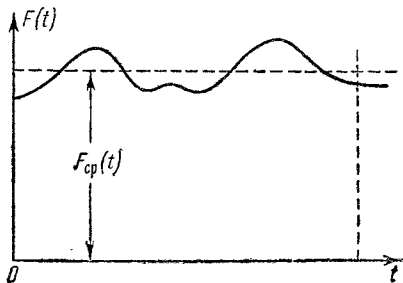


Рис. 3.14.

Найдем дифференциальное уравнение, которому удовлетворяет частное решение (3.200), для чего дважды продифференцируем соотношение (3.200) по времени и исключим промежуточные переменные. При этом будем иметь:

$$\frac{d^2 u}{dt^2} + \left[N^2 - \frac{3}{4} \left(\frac{N'}{N} \right)^2 + \frac{1}{2} \frac{N''}{N} \right] u = 0. \quad (3.202)$$

Сравнивая уравнения (3.199) и (3.202), видим, что если удовлетворяется тождество

$$F(t) = N^2 - \frac{3}{4} \left(\frac{N'}{N} \right)^2 + \frac{1}{2} \frac{N''}{N}, \quad (3.203)$$

то соотношение (3.200) будет являться частным решением уравнения (3.199). В общем случае можно было бы решить уравнение (3.203) относительно искомой функции $N(t)$, и тогда мы получили бы точное решение уравнения (3.199). Однако отыскание решения уравнения (3.203) является более трудной задачей, чем решение уравнения (3.199), так как уравнение (3.203) — нелинейное.

Если предположить, что отношения второго и третьего членов уравнения (3.203) к первому малы, т. е. удовлетворяются неравенства

$$\left| \frac{N'}{N^2} \right|^2 < \varepsilon, \quad \left| \frac{N''}{N^3} \right| < \varepsilon^2, \quad \varepsilon \ll 1, \quad (3.204)$$

решение уравнения (3.203) можно искать в виде быстро сходящегося ряда

$$N = N_0 + N_1 + N_2 + \dots \quad (3.205)$$

Подставляя ряд (3.205) в уравнение (3.203), получим следующие формулы для определения членов ряда (3.205):

$$\left. \begin{aligned} N_0(t) &= \sqrt{F(t)}, \\ 2N_0N_1 &= \frac{3}{4} \left(\frac{N'_0}{N_0} \right)^2 - \frac{1}{2} \cdot \frac{N''_0}{N_0}, \\ 2N_0N_2 + N_1^2 &= \frac{3}{2} \left(\frac{N'_0}{N_0} \right)^2 \left(\frac{N'_1}{N_0} - \frac{N_1}{N_0} \right) - \frac{1}{2} \cdot \frac{N''_0}{N_0} \left(\frac{N'_1}{N_0} - \frac{N_1}{N_0} \right). \end{aligned} \right\} (3.206)$$

Однако во многих практических случаях можно ограничиться лишь первым членом ряда (3.205), особенно когда функция $F(t)$ изменяется медленно, оставаясь в среднем большой. При этом непосредственно из формул (3.206) видно, что члены N_1 и N_2 и т. д. будут малы по сравнению с членом N_0 , и ими можно пренебречь.

Итак, будем считать, что функция $N(t)$ определяется соотношением

$$N(t) = \sqrt{F(t)}. \quad (3.207)$$

Тогда частное решение уравнения (3.199) непосредственно дается формулами (3.200), (3.201) и (3.207). Имея другое линейно независимое частное решение уравнения (3.199), можно найти импульсную переходную функцию $r(t, \xi)$ по формуле (3.175). Если считать, что коэффициент $F(t)$ на интересующем нас интервале времени положителен, естественно взять за второе частное решение комплексно-сопряженную относительно выражения (3.200) величину, т. е.

$$u_2(t) = \frac{1}{\sqrt{N}} e^{iS(t)}. \quad (3.208)$$

При этом импульсная переходная функция $r(t, \xi)$ определится, как и ранее, выражением

$$r(t, \xi) = -\frac{1}{\Delta(\xi)} \begin{vmatrix} u_1(t) & u_2(t) \\ u_1(\xi) & u_2(\xi) \end{vmatrix}. \quad (3.209)$$

Подставляя в выражение (3.209) значения u_1 и u_2 из формул (3.200) и (3.208) и раскрывая определитель Вронского, получим:

$$r(t, \xi) = -\frac{u_1(t)u_2(\xi) - u_1(\xi)u_2(t)}{u_1(\xi)u'_2(\xi) - u'_1(\xi)u_2(\xi)} = \frac{u_1(t)u_2(\xi) - u_1(\xi)u_2(t)}{2j}.$$

так как

$$\begin{aligned} u_1(\xi) u_2'(\xi) - u_1'(\xi) u_2(\xi) &= 2j; \\ u_1(t) u_2(\xi) &= \frac{1}{\sqrt{N(t)N(\xi)}} e^{-j[S(t)-S(\xi)]}; \\ u_1(\xi) u_2(t) &= \frac{1}{\sqrt{N(t)N(\xi)}} e^{j[S(t)-S(\xi)]}, \end{aligned}$$

или окончательно

$$r(t, \xi) = \frac{1}{\sqrt{F(t)F(\xi)}} \sin [S(t) - S(\xi)]. \quad (3.210)$$

Легко видеть, что формула (3.210) удовлетворяет и предельному случаю, когда функция $F(t)$ есть постоянная величина. Положив при этом $F(t) = \Omega^2$, получим:

$$S(t) = \int_0^t \Omega dt = \Omega t, \quad S(\xi) = \int_0^\xi \Omega d\xi = \Omega \xi,$$

и на основании формулы (3.210) будем иметь:

$$r(t, \xi) = r(\tau) = \frac{1}{\Omega} \sin \Omega \tau, \quad \tau = t - \xi,$$

как и должно быть.

Так как система второго порядка описывается, как это было принято выше, уравнением

$$\frac{d^2 x_{\text{ВЫХ}}}{dt^2} + p(t) \frac{dx_{\text{ВЫХ}}}{dt} + q(t) x_{\text{ВЫХ}} = x_{\text{ВХ}}(t),$$

то соответствующая импульсная переходная функция $g(t, \xi)$ будет определяться через функцию $r(t, \xi)$ на основании зависимости (3.159). Применяя эту зависимость и используя выражение (3.210), окончательно получим:

$$g(t, \xi) = \frac{1}{\sqrt{F(t)F(\xi)}} e^{-\frac{1}{2} \int_\xi^t p d} \sin [S(t) - S(\xi)]. \quad (3.211)$$

При практическом применении формулы (3.211) необходимо всегда вначале убедиться в том, что коэффициенты дифференциального уравнения системы удовлетворяют условиям (3.204), для чего удобно пользоваться следующей формулой для оценки «медленности» изменения коэффициента $F(t)$, полученной из соотношений (3.203) и (3.207):

$$\left| -\frac{1}{2} \frac{[F'(t)]^2}{F^3(t)} + \frac{1}{4} \frac{F''(t)}{F^2(t)} \right| \ll 1. \quad (3.212)$$

Рассмотренный способ получения приближенного решения уравнения второго порядка с переменными коэффициентами известен в задачах, приводящихся к уравнениям Хилла, как аппроксимация Бриллюина — Вентцеля — Крамера, или сокращенно «аппроксимация БВК» [3].

Пример 3.6. Найдем приближенное решение уравнения из примера предыдущего параграфа с помощью «аппроксимации БВК».

На основании формулы (3.211) будем иметь:

$$g(t, \xi) = \frac{e^{-\frac{1}{2}(t-\xi)}}{\sqrt[4]{F(t)F(\xi)}} \sin[S(t) - S(\xi)], \quad (3.213)$$

где

$$S(t) = \int_0^t \sqrt{F(t)} dt, \quad F(t) = 0,25e^{0,4t} - 0,01.$$

Оценка применимости метода по формуле (3.212) дает:

$$\left| -\frac{1}{2} \frac{[F'(t)]^2}{F^3(t)} + \frac{1}{4} \frac{F''(t)}{F^2(t)} \right| = | -0,16e^{-0,4t} |,$$

так как

$$\frac{[F'(t)]^2}{F^3(t)} \approx \frac{0,01e^{2 \cdot 0,4t}}{0,25^3 \cdot e^{3 \cdot 0,4t}} = 0,64e^{-0,4t}, \quad \frac{F''(t)}{F^2(t)} \approx \frac{0,04e^{0,4t}}{0,25^2 \cdot e^{2 \cdot 0,4t}} = 0,64e^{-0,4t}$$

и $| -0,16e^{-0,4t} | \ll 1$ для всех t , причем с ростом t точность решения растет. Полученная оценка показывает, что можно ожидать весьма близкого совпадения приближенного решения с точным.

Исходя из числовых значений $F(t)$, приближенно будем иметь $F(t) \approx 0,25e^{0,4t} = qe^{2pt}$, откуда $F(\xi) \approx qe^{2p\xi}$. Далее,

$$S(t) = \sqrt{q} \int_0^t e^{pt} dt = \frac{\sqrt{q}}{p} (e^{pt} - 1).$$

При этом получим:

$$S(t) - S(\xi) = \frac{\sqrt{q}}{p} (e^{pt} - e^{p\xi}).$$

Далее, будем иметь:

$$\frac{1}{\sqrt[4]{F(t)F(\xi)}} \approx \frac{1}{\sqrt{q^2 e^{2p(t+\xi)}}} = \frac{1}{\sqrt{q}} e^{-\frac{p}{2}(t+\xi)},$$

и решение примет следующий вид:

$$g(t, \xi) = \frac{1}{\sqrt{q}} e^{-\frac{1}{2} p(t+\xi)} e^{-\frac{1}{2} p(t-\xi)} \sin \left[\frac{\sqrt{q}}{p} (e^{pt} - e^{p\xi}) \right]$$

или окончательно

$$g(t, \xi) = \frac{1}{\sqrt{q}} e^{pt} \sin \left[\frac{\sqrt{q}}{p} (e^{pt} - e^{p\xi}) \right]. \quad (3.214)$$

Сравнивая соотношения (3.214) и (3.198'), видим, что аппроксимация в данном случае дала совпадение с точным решением.

ГЛАВА 4

СИСТЕМЫ С ПЕРЕМЕННЫМИ ПАРАМЕТРАМИ, НАХОДЯЩИЕСЯ ПОД ВОЗДЕЙСТВИЕМ СИГНАЛОВ В ВИДЕ ЗАДАННЫХ ФУНКЦИЙ

§ 20. Определение выходного сигнала системы с помощью импульсной переходной функции

В предыдущих главах были подробно рассмотрены вопросы, относящиеся к свойствам импульсной переходной функции и к методам ее отыскания, причем было показано, что с помощью импульсной переходной функции устанавливается связь между входным и выходным сигналами системы. Эта связь дается соотношением.

$$x_{\text{вых}}(t) = \int_0^t w(t, \xi) x_{\text{вх}}(\xi) d\xi. \quad (4.1)$$

Таким образом, принципиально задача определения выходного сигнала системы при заданном входном сигнале решается путем взятия интеграла (4.1) тем или иным способом. Хотя в сравнительно простых случаях непосредственное использование формулы (4.1) и ведет к желаемому результату, исчерпывая поставленную задачу, однако значительно чаще приходится изыскивать обходные пути, так как взятие интеграла (4.1) оказывается достаточно трудным делом.

Рассмотрим некоторые практические способы определения выходного сигнала системы с помощью импульсной переходной функции, если входной сигнал является заданной функцией времени. Предварительно заметим, что последнее определение является в некоторой степени условным и должно приниматься только в определенном смысле.

В самом деле, если бы входной сигнал системы был вполне определенной, заранее известной функцией времени, то отпала бы необходимость в передаче сигнала через систему именно как входного — он мог бы просто воспроизводиться программным устройством в нужной точке системы либо компенсироваться в случае системы стабилизации. В действительности входной сигнал является случайным.

и рассматривать его как заданную функцию можно либо в том случае, когда нас интересует реакция данной конкретной системы на данный конкретный сигнал, либо тогда, когда в среднем случайные входные сигналы близки к рассматриваемой заданной функции и их отклонения от этой функции не имеют в данной задаче существенного значения.

Одним из сравнительно простых и в большой степени наглядных методов, всегда приводящим к цели, является метод определения отдельных точек выходного процесса графически путем умножения входного сигнала на сопряженную импульсную реакцию системы и вычисления получающейся площади непосредственно по графику.

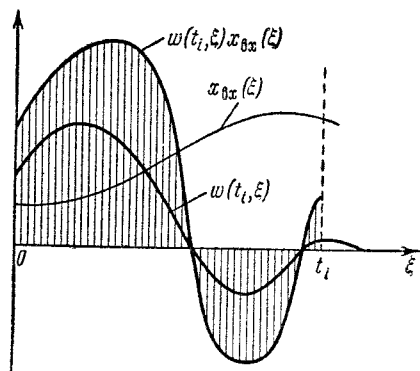


Рис. 4.1.

Если в выражении (4.1) вместо текущего времени подставить некоторое его фиксированное значение, то тогда ордината $x_{\text{вых}}(t_i)$ будет равна определенному интегралу

$$x_{\text{вых}}(t_i) = \int_0^{t_i} w(t_i, \xi) x_{\text{вх}}(\xi) d\xi, \quad (4.2)$$

который характеризует площадь, ограниченную кривой $w(t_i, \xi) x_{\text{вх}}(\xi)$, осью ξ и ординатой $\xi = t_i$ (рис. 4.1). Эта площадь легко вычисляется, если кривые построены на миллиметровой бумаге. Ясно, что для

получения n точек кривой выходного процесса необходимо иметь n кривых сопряженных импульсных реакций системы для n фиксированных значений t . Эти кривые могут быть найдены с помощью методов, изложенных в предыдущей главе. Возможно сочетание расчетных методов с методами моделирования, когда сопряженные импульсные реакции определяются на моделирующих устройствах (глава 6), а расчет выходных процессов производится вышеуказанным способом. Благодаря наглядности операций сравнительно легко усматривается влияние вида импульсной реакции на характер выходного процесса при данном входе, что позволяет производить выбор необходимого вида импульсной переходной функции.

Рассмотрим аналитический способ определения выходного процесса в виде степенного ряда для инерционных цепей

$$x_{\text{вых}}(t) = x_{\text{вых}}(0) + x'_{\text{вых}}(0)t + \frac{x''_{\text{вых}}(0)}{2!}t^2 + \dots \quad (4.3)$$

Очевидно, определение процесса сводится к определению коэффициентов ряда (4.3), причем в зависимости от порядка системы,

Выполняя в этом выражении операцию дифференцирования и группируя члены, получим:

$$\begin{aligned}
 x_{\text{вых}}^{(n+\nu)}(0) = & \{[\omega_t^{(n+\nu-1)}(t, \xi) x_{\text{вх}}(\xi)]_{\xi=t} + \\
 & + \left[\frac{d^\nu}{dt^\nu} \omega_t^{(n-1)} + \dots + \frac{d^2}{dt^2} \omega_t^{(n+\nu-3)} + \frac{d}{dt} \omega_t^{(n+\nu-2)} \right] x_{\text{вх}}(t) + \\
 & + \left[\nu \frac{d^{\nu-1}}{dt^{\nu-1}} \omega_t^{(n-1)} + \dots + 2 \frac{d}{dt} \omega_t^{(n+\nu-3)} + \omega_t^{(n+\nu-2)} \right] x'_{\text{вх}}(t) + \dots \\
 & \dots + \omega_t^{(n-1)} x_{\text{вх}}^{(\nu)}(t) \}_{t=0}. \quad (4.8)
 \end{aligned}$$

Если теперь ввести операторное выражение

$$Q(p, t) = p^\nu \omega^{(n-1)} + p^{\nu-1} \omega^{(n)} + \dots + p \omega^{(n+\nu-2)}, \quad (4.9)$$

где

$$p = \frac{d}{dt}, \quad \omega^{(n+k)} = [\omega_t^{(n+k)}(t, \xi)]_{\xi=t}, \quad (4.10)$$

и полагать, что $p^{1-\mu} = 0$, $\mu = 1, 2, \dots$, соотношение (4.8) можно записать в виде

$$\begin{aligned}
 x_{\text{вых}}^{(n+\nu)}(0) = & \{[\omega_t^{(n+\nu-1)}(t, \xi) x_{\text{вх}}(\xi)]_{\xi=t} + Q(p, t) x_{\text{вх}}(t) + \\
 & + \frac{dQ}{dp} x'_{\text{вх}}(t) + \frac{1}{2!} \frac{d^2 Q}{dp^2} x''_{\text{вх}}(t) + \dots + \frac{1}{\nu!} \frac{d^\nu Q}{dp^\nu} x_{\text{вх}}^{(\nu)}(t) \}_{t=0}, \quad (4.11) \\
 & \nu = 0, 1, 2, \dots
 \end{aligned}$$

По формулам (4.9) и (4.11) можно определить коэффициенты ряда (4.3), а следовательно, и искомый выходной процесс для случая, когда входной сигнал не содержит дельта-функций. Если же таковые имеются, то они выделяются из общего сигнала и реакция системы на них отыскивается отдельно методами, изложенными в предыдущей главе. Затем результаты обоих расчетов складываются.

Рассмотренный метод не ограничивается, конечно, только применением к отысканию выходных процессов системы. Он с успехом может быть использован для вычислений общих интегральных преобразований вида $F(s) = \int_0^b H(s, \lambda) F(\lambda) d\lambda$ в том случае, когда представление функции $F(s)$ в виде степенного ряда (вообще говоря, бесконечного) с N первыми членами является достаточным.

В качестве иллюстрации применения выведенных формул рассмотрим простой пример.

Пример 4.1. Пусть на вход системы первого порядка с импульсной переходной функцией

$$\omega(t, \xi) = e^{-\frac{1}{2}(t^2 - \xi^2)} \quad (4.12)$$

подается сигнал

$$x_{\text{вх}}(t) = 1 - \frac{1}{2}t^2 + \frac{1}{8}t^4. \quad (4.13)$$

Требуется рассчитать и построить процесс на выходе системы в виде полинома пятой степени

$$x_{\text{вых}}(t) = x_{\text{вых}}(0) + x'_{\text{вых}}(0)t + x''_{\text{вых}}(0)\frac{t^2}{2} + \dots \quad (4.13')$$

Задача, таким образом, состоит в определении коэффициентов этого ряда и решается с помощью формулы (4.11). Применительно к рассматриваемому случаю в формуле следует положить $n=1$ и задавать $\nu=0, 1, 2, 3, 4$.

Определим сначала составляющие искоемых коэффициентов, входящие в квадратную скобку выражения (4.11),

$$\{[\omega_t^{(n+\nu-1)}(t, \xi) x_{\text{вх}}(\xi)]_{\xi-t}\}_{t=\xi=0}.$$

Дифференцируя по t выражение (4.12) $k = n + \nu - 1 = \nu$ раз, где $\nu=0, 1, 2, 3, 4$, будем иметь:

$$\begin{aligned} [\omega(t, \xi)]_{\xi-t} &= [e^{-\frac{1}{2}(t^2-\xi^2)}]_{\xi-t} = 1, \\ [\omega'_t(t, \xi)]_{\xi-t} &= \left[-\frac{1}{2}2te^{-\frac{1}{2}(t^2-\xi^2)} \right]_{\xi-t} = -t \end{aligned}$$

и т. д.

Умножая выражения производных на $x_{\text{вх}}(\xi)$ и полагая в соответствии с формулой (4.11) $t = \xi = 0$, получим:

$$\begin{aligned} [\omega(t, \xi) x_{\text{вх}}(\xi)]_{\xi-t=0} &= 1, \\ [\omega'_t(t, \xi) x_{\text{вх}}(\xi)]_{\xi-t=0} &= 0 \end{aligned}$$

и т. д. Результаты расчетов приведены в таблице 18.

Таблица 18

Значения первого члена ряда (4.11)

$k = n + \nu - 1$	$[\omega_t^{(k)}(t, \xi)]_{\xi-t}$	$[\omega_t^{(k)}(t, \xi) x_{\text{вх}}(\xi)]_{\xi-t=0}$
0	1	1
1	$-t$	0
2	$t^2 - 1$	-1
3	$3t - t^3$	0
4	$3 - 6t^2 + t^4$	3

Для определения остальных составляющих $x_{\text{вых}}^{(n+\nu)}(0)$ нам потребуются значения оператора $Q(p, t)$ и его производных по p , а также значения производных $x_{\text{вх}}^k(t)$ в момент $\xi = t = 0$. Обращаясь к формуле (4.9) для $n=1$, будем иметь:

$$Q(p, t) = p^4 \omega^{(\nu-4)}(t, t) + p^3 \omega^{(\nu-3)}(t, t) + p^2 \omega^{(\nu-2)}(t, t) + p \omega^{(\nu-1)}(t, t).$$

Дифференцируя это выражение по p , запишем необходимые для расчета выражения производных оператора $Q(p, t)$:

$$\frac{dQ}{dp} = 4p^3 w^{(\nu-4)}(t, t) + 3p^2 w^{(\nu-3)}(t, t) + 2p w^{(\nu-2)}(t, t) + w^{(\nu-1)}(t, t),$$

$$\frac{d^2Q}{dp^2} = 12p^2 w^{(\nu-4)}(t, t) + 6p w^{(\nu-3)}(t, t) + 2w^{(\nu-2)}(t, t),$$

$$\frac{d^3Q}{dp^3} = 24p w^{(\nu-4)}(t, t) + 6w^{(\nu-3)}(t, t),$$

$$\frac{d^4Q}{dp^4} = 24w^{(\nu-4)}(t, t).$$

Задаваясь последовательно значениями $\nu = 0, 1, 2, 3, 4$ и учитывая, что производная по t от $[\omega(t, \xi)]_{\xi=t} = \omega(t, t)$ может быть только положительной, будем получать следующие выражения:

для $\nu = 0$:

$$Q = \frac{dQ}{dp} = \frac{d^2Q}{dp^2} = \frac{d^3Q}{dp^3} = \frac{d^4Q}{dp^4} = 0;$$

для $\nu = 1$:

$$Q = p\omega(t, t) = \frac{d}{dt} \omega(t, t) = 0,$$

$$\frac{dQ}{dp} = \omega(t, t) = 1; \quad \frac{d^2Q}{dp^2} = \frac{d^3Q}{dp^3} = \frac{d^4Q}{dp^4} = 0;$$

для $\nu = 2$:

$$Q = p^2\omega(t, t) + p\omega'(t, t) = \frac{d}{dt} [\omega'(t, t)] = -1,$$

$$\frac{dQ}{dp} = 2p\omega(t, t) + \omega'(t, t) = -t,$$

$$\frac{d^2Q}{dp^2} = 2\omega(t, t) = 2; \quad \frac{d^3Q}{dp^3} = \frac{d^4Q}{dp^4} = 0;$$

для $\nu = 3$:

$$Q = p^3\omega(t, t) + p^2\omega'(t, t) + p\omega''(t, t) = \frac{d}{dt} [\omega''(t, t)] = 2t,$$

$$\frac{dQ}{dp} = 3p^2\omega(t, t) + 2p\omega'(t, t) + \omega''(t, t) = -3 + t^2,$$

$$\frac{d^2Q}{dp^2} = 6p\omega(t, t) + 2\omega'(t, t) = -2t,$$

$$\frac{d^3Q}{dp^3} = 6\omega(t, t) = 6; \quad \frac{d^4Q}{dp^4} = 0;$$

для $\nu = 4$:

$$Q = p^4\omega(t, t) + p^3\omega'(t, t) + p^2\omega''(t, t) + p\omega^{(3)}(t, t) = 5 - 3t^2,$$

$$\frac{dQ}{dp} = 4p^3\omega(t, t) + 3p^2\omega'(t, t) + 2p\omega''(t, t) + \omega^{(3)}(t, t) = 7t - t^3,$$

$$\frac{d^2Q}{dp^2} = 12p^2\omega(t, t) + 6p\omega'(t, t) + 2\omega''(t, t) = -8 + 2t^2,$$

$$\frac{d^3Q}{dp^3} = 24p\omega(t, t) + 6\omega'(t, t) = -6t,$$

$$\frac{d^4Q}{dp^4} = 24\omega(t, t) = 24.$$

Результаты расчетов коэффициента Q и его производных для текущего момента времени t приведены в таблице 19.

Таблица 19

Значения коэффициента Q и его производных

ν	Q	$\frac{dQ}{dp}$	$\frac{d^2Q}{dp^2}$	$\frac{d^3Q}{dp^3}$	$\frac{d^4Q}{dp^4}$
0	0	0	0	0	0
1	0	1	0	0	0
2	-1	- t	2	0	0
3	$2t$	$-3 + t^2$	$-2t$	6	0
4	$5 - 3t^2$	$7t - t^3$	$-8 + 2t^2$	$6t$	24

Производные по времени от входного сигнала определяются последовательным дифференцированием выражения (4.13):

$$x'_{\text{вх}}(t) = -t + \frac{1}{2}t^3,$$

$$x''_{\text{вх}}(t) = -1 + \frac{3}{2}t^2,$$

$$x'''_{\text{вх}}(t) = 3t,$$

$$x^{(4)}_{\text{вх}}(t) = 3.$$

Необходимые для расчетов произведения $Q(p, t) x_{\text{вх}}(t)$, $\frac{dQ}{dp} x'_{\text{вх}}(t)$ и т. д. приведены в таблице 20.

Таблица 20

Значения членов ряда (4.11) при произвольном t

ν	$Q x_{\text{вх}}$	$\frac{dQ}{dp} x'_{\text{вх}}$	$\frac{d^2Q}{dp^2} x''_{\text{вх}}$	$\frac{d^3Q}{dp^3} x^{(3)}_{\text{вх}}$	$\frac{d^4Q}{dp^4} x^{(4)}_{\text{вх}}$
0	0	0	0	0	0
1	0	$-t + \frac{1}{2}t^3$	0	0	0
2	$-1 + \frac{1}{2}t^2 - \frac{1}{8}t^4$	$t^2 - \frac{1}{2}t^4$	$-2 + 3t^2$	0	0
3	$2t^2 - t^3 + \frac{1}{4}t^5$	$3t - \frac{5}{2}t^3 + \frac{1}{2}t^5$	$2t^2 - 3t^3$	$18t$	0
4	$(5 - 3t^2) \times$ $\times \left(1 - \frac{1}{2}t^2 + \frac{1}{8}t^4\right)$	$(5t - t^3) \times$ $\times \left(-t + \frac{1}{2}t^3\right)$	$(-8 + 2t^2) \times$ $\times \left(-1 + \frac{3}{2}t^2\right)$	$-18t^2$	72

Теперь нетрудно вычислить и значения коэффициентов искомого ряда (4.13), если в соответствии с формулой (4.11) умножить коэффициенты $\frac{d^v Q}{d\rho^v}$

Таблица 21
Значения коэффициентов ряда (4.3)

ν	$n + \nu = i$	$x_{\text{вых}}^{(i)}(0)$
0	1	1
1	2	0
2	3	-3
3	4	0
4	5	15

на $\frac{1}{\nu!}$ и во всех членах положить $t = \xi = 0$. Полученные в результате расчета значения коэффициентов $x_{\text{вых}}^{(i)}(0)$ приведены в таблице 21. В соответствии с этой таблицей искомым выходной процесс можно записать в окончательном виде:

$$x_{\text{вых}}(t) = t - \frac{3}{3!} t^3 + \frac{15}{5!} t^5 = \\ = t \left(1 - \frac{1}{2} t^2 + \frac{1}{8} t^4 \right). \quad (4.14)$$

Результаты расчета в нашем случае могут быть проверены и непосредственным применением формулы (4.1) к заданным функциям (4.12) и (4.13)

$$x_{\text{вых}}(t) = \int_0^t e^{-\frac{1}{2}(t^2 - \xi^2)} \left(1 - \frac{1}{2} \xi^2 + \frac{1}{8} \xi^4 \right) d\xi = \\ = e^{-\frac{1}{2} t^2} \int_0^t e^{-\frac{1}{2} \xi^2} \left(1 - \frac{1}{2} \xi^2 + \frac{1}{8} \xi^4 \right) d\xi.$$

Легко видеть, что подынтегральное выражение с точностью до трех членов разложения функции $e^{-\frac{1}{2} \xi^2}$ в ряд равно единице. Следовательно,

$$x_{\text{вых}}(t) \approx e^{-\frac{1}{2} t^2} \cdot t \approx t \left(1 - \frac{1}{2} t^2 + \frac{1}{8} t^4 \right). \quad (4.15)$$

Выражения (4.14) и (4.15) совпадают вплоть до пятой степени t , что и требовалось.

§ 21. Применение преобразований Лапласа и Фурье. Параметрическая передаточная функция системы и ее определение

При анализе систем с переменными параметрами интегральная связь между входным и выходным сигналами непосредственно почти никогда не используется из-за возникающих при этом вычислительных трудностей, а применяется ее изображение по Лапласу или Фурье. Операции определения выходного сигнала системы с помощью упомянутых изображений в значительной степени упрощаются. Кроме того (что является немаловажным для инженерной практики), в этом случае представляется возможность использовать таблицы преобразований Лапласа и применять приближенные методы, связанные со спектральными представлениями. Поэтому распространение указанных мето-

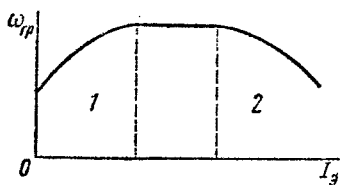


Рис. 4.39. Зависимость граничной частоты коэффициента передачи тока базы от постоянного тока эмиттера:

1 — диапазон частот с преимущественным влиянием цепи эмиттера; 2 — диапазон частот с преимущественным влиянием изменения толщины коллекторного перехода

малых токов определяющую роль играет $r_э C_э$, бар. С увеличением тока сопротивление эмиттера падает и определяющим процессом становится перенос носителей через базу. При этом возникновение электрического поля в базе приводит к дальнейшему росту $\omega_{гр}$. При очень больших токах начинает сказываться сдвиг границы области объемного заряда коллекторного перехода и увеличение его толщины (см. § 4.5). Это ведет к снижению $\omega_{гр}$.

Уточненная частотная характеристика. В некоторых случаях для

аппроксимации частотной характеристики транзистора используют выражение, аналогичное (4.127):

$$\dot{h}_{21б} = \frac{h_{21б} \exp\left(-jm \frac{\omega}{\omega_{h21б}}\right)}{1 + j \frac{\omega}{\omega_{h21б}}}. \quad (4.153)$$

Как и ранее, зависимость модуля $\dot{h}_{21б}$ от частоты имеет вид

$$|\dot{h}_{21б}| = \frac{h_{21б}}{\sqrt{1 + \left(\frac{\omega}{\omega_{h21б}}\right)^2}}, \quad (4.154)$$

а аргумента

$$\arg \dot{h}_{21б} = -\left(m \frac{\omega}{\omega_{h21б}} + \arctg \frac{\omega}{\omega_{h21б}}\right). \quad (4.155)$$

Физический смысл *предельной частоты коэффициента передачи тока эмиттера* $\omega_{h21б}$ — частота, на которой модуль коэффициента передачи тока эмиттера падает в $\sqrt{2}$ раз по сравнению с его низкочастотным значением.

Из равенства аргументов выражений (4.155) и (4.148) получим

$$m \frac{\omega}{\omega_{h21б}} + \arctg \frac{\omega}{\omega_{h21б}} = \arctg \frac{\omega}{\omega_{гр}} \quad (4.156)$$

или, взяв приближенно

$$\arctg \frac{\omega}{\omega_{h21б}} \approx \frac{\omega}{\omega_{h21б}} \quad \text{и} \quad \arctg \frac{\omega}{\omega_{гр}} \approx \frac{\omega}{\omega_{гр}}, \quad (4.157)$$

$$\omega_{h21б} = (1 + m) \omega_{гр}. \quad (4.158)$$

Частотная зависимость $\dot{h}_{21э}$. Воспользовавшись выражением (4.153), найдем частотную зависимость коэффициента передачи тока для схемы с общим эмиттером $\dot{h}_{21э}$. Из условия баланса токов в транзисторе

$$\dot{h}_{21э} = \frac{-\dot{h}_{21б}}{1 + \dot{h}_{21б}}. \quad (4.159)$$

Подставив значение h_{216} из (4.148), получим

$$\dot{h}_{215} = \frac{-h_{216}}{\left(1 + j \frac{\omega}{\omega_{rp}}\right) \left(1 + \frac{h_{216}}{1 + j \frac{\omega}{\omega_{rp}}}\right)} = \frac{-h_{216}}{1 + h_{216} + j \frac{\omega}{\omega_{rp}}}, \quad (4.160)$$

или

$$\dot{h}_{215} = \frac{h_{215}}{1 + j \frac{\omega}{\omega_{h215}}}, \quad (4.161)$$

где

$$\omega_{h215} = \omega_{rp} (1 - |h_{216}|) \approx \frac{\omega_{rp}}{h_{215}}. \quad (4.162)$$

Физический смысл *предельной частоты коэффициента передачи тока базы* ω_{h215} — частота, на которой модуль коэффициента передачи тока базы уменьшается в $\sqrt{2}$ раз по сравнению с низкочастотным значением.

Следует отметить, что если частотные свойства транзистора определяются только процессами переноса через базу (например, для бездрейфовых транзисторов), то

$$\omega_{rp} = 1/t_{прол}; \quad h_{215} = \tau_p/t_{прол} \quad (4.163)$$

и

$$\omega_{h215} = 1/\tau_p. \quad (4.164)$$

Практически частота ω_{h215} сравнительно невелика — транзистор может работать и на существенно более высоких частотах. В этом случае можно считать, что

$$|\dot{h}_{215}| \approx \left| \frac{h_{215} \omega_{h215}}{\omega} \right| = \frac{\omega_{rp}}{\omega}. \quad (4.165)$$

Таким образом, частота ω_{rp} — *граничная частота коэффициента передачи тока базы* — это частота, при которой модуль коэффициента передачи тока базы становится равным единице.

Из соотношения (4.165) следует, что на частотах, превышающих предельную частоту коэффициента передачи тока базы ω_{h215} , модуль коэффициента передачи тока базы обратно пропорционален частоте. Поэтому для экспериментального определения граничной частоты коэффициента передачи тока базы ω_{rp} достаточно измерить модуль коэффициента передачи тока базы $|\dot{h}_{215}|$ на какой-либо частоте $\omega > \omega_{h215}$ и далее использовать соотношение

$$\omega_{rp} = |\dot{h}_{215}| \omega. \quad (4.166)$$

Графически зависимость коэффициентов передачи от частоты представлена на рис. 4.40. Из графика видно, что $|\dot{h}_{215}|$ падает на значительно более низких частотах, чем $|\dot{h}_{216}|$. Физически это связано с влиянием фазового сдвига между токами эмиттера и

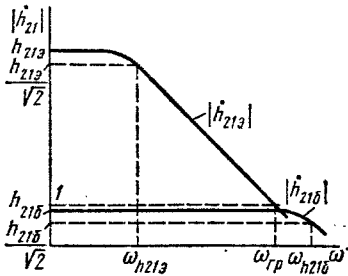


Рис. 4.40. Зависимости модулей коэффициентов передачи от частоты (масштаб логарифмический)

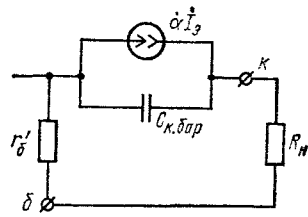


Рис. 4.42. Включение нагрузки в цепь коллектора

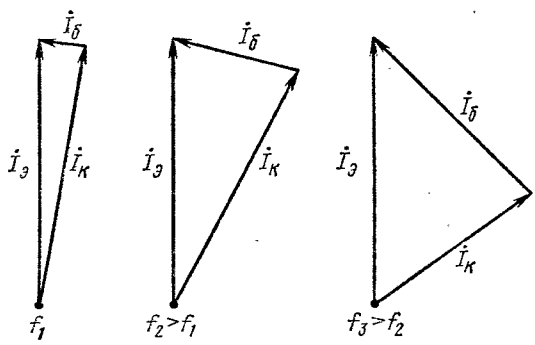


Рис. 4.41. Векторные диаграммы токов транзистора на разных частотах

коллектора (рис. 4.41). С увеличением частоты фазовый сдвиг растет и это приводит к росту тока базы, даже если ток коллектора и ток эмиттера неизменны по абсолютной величине. Следовательно, и

$$|h_{21\beta}| = I_k / I_b \quad (4.167)$$

уменьшается.

Усиление по мощности на высоких частотах.

С ростом частоты сигнала коэффициент усиления транзистора по мощности падает. Это связано как с уменьшением коэффициента усиления по току, рассмотренным ранее, так и с влиянием цепи $r'_b C_{к. бap}$. С ростом частоты шунтирующее действие этой цепи на нагрузку усиливается, в нагрузку ответвляется меньший ток, что и приводит к снижению коэффициента усиления по мощности (рис. 4.42).

Для борьбы с этим явлением следует выбирать определенное сопротивление нагрузки, т. е. согласовывать нагрузку. Однако на очень высоких частотах и это не дает возможности получить достаточный коэффициент усиления по мощности. На некоторой частоте коэффициент усиления транзистора по мощности даже при согласованной нагрузке становится равным единице. Это означает, что транзистор на такой частоте уже нельзя рассматривать как активный элемент электрической схемы.

При коэффициенте усиления по мощности, равном (или меньшем) единице, нельзя осуществить режим самовозбуждения в генераторе на транзисторе. Поэтому ту частоту, на которой коэффициент усиления по мощности становится равным единице, называют **максимальной частотой генерации**.

Расчет показывает, что максимальную частоту генерации можно найти по формуле

$$f_{\max} = \sqrt{\frac{h_{21} \beta f_{\text{гп}}}{8\pi r_0' C_{\text{к. баз}}}}. \quad (4.168)$$

Максимальная частота генерации наиболее полно характеризует свойства транзистора, так как определяет область частот, в которой транзистор остается активным элементом электрической схемы.

§ 4.15. РАБОТА ТРАНЗИСТОРА НА ИМПУЛЬСАХ

Особенности работы

Транзисторы часто используют в переключающих и импульсных схемах. При этом от транзистора требуется, как правило, неискаженное воспроизведение усиленного импульса на выходе. Работа транзистора в качестве усилителя малых импульсных сигналов в принципе ничем не отличается от работы транзистора как усилителя малых синусоидальных сигналов. Импульс можно представить в виде суммы ряда гармонических составляющих и, зная частотные свойства транзистора, определить искажения формы импульса, которые могут иметь место при усилении.

В данном параграфе рассмотрим работу транзистора как усилителя больших импульсных сигналов. В таких случаях нельзя использовать малосигнальные параметры транзистора, а также его обычные эквивалентные схемы. Кроме того, транзистор в переключающих и импульсных схемах часто работает не только в активном режиме, но и в режимах отсечки и насыщения.

Схема с общей базой

Рассмотрим процессы, происходящие в транзисторе, включенном по схеме с общей базой, при подаче через эмиттер импульса тока длительностью $t_{\text{имп}}$ в прямом направлении с последующим изменением полярности (рис. 4.43). В исходном состоянии транзистор находится в режиме отсечки, т. е. эмиттерный и коллекторный переходы закрыты.

После подачи на эмиттер импульса тока в прямом направлении ток коллектора появляется не сразу из-за конечного времени передвижения инжектированных носителей заряда до коллекторного перехода и наличия барьерных емкостей (рис. 4.43, б). Время, на которое появление коллекторного тока отстает от эмиттерного, называют в р е м е н е м з а д е р ж к и t_3 . Процесс установления тока коллектора характеризуется длительностью переднего фронта импульса $t_{\text{фд}}$ (рис. 4.43, б). За время переднего фронта импульса тока коллектора в базе транзистора происходит накопление неосновных носителей заряда, что можно пояснить с помощью рис. 4.43, в, где показаны кривые распределения неосновных носителей в базе транзистора в различные моменты времени. Напряжение на эмит-

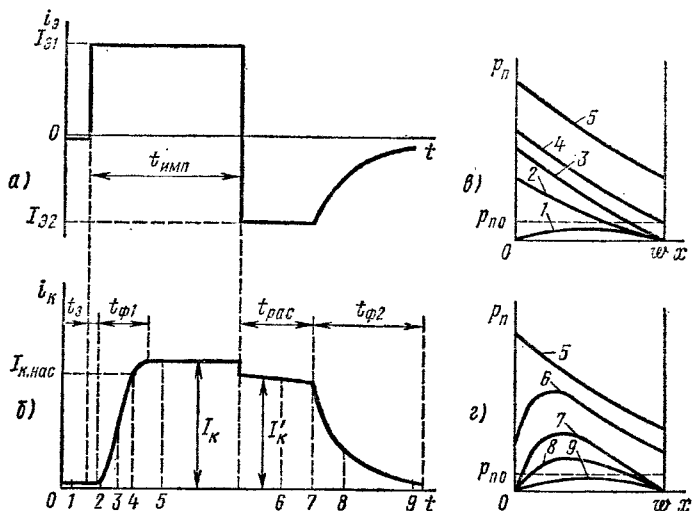


Рис. 4.43. Временные зависимости тока эмиттера (а) и тока коллектора (б) при включении транзистора по схеме с общей базой и распределение неосновных носителей в базе транзистора в различные моменты времени (в и г)

терном переходе растет замедленно из-за заряда емкости эмиттерного перехода, что соответствует увеличению со временем ординаты, отсекаемой кривыми распределения неосновных носителей со стороны эмиттерного перехода. Градиент концентрации неосновных носителей около эмиттера, соответствующий величине инжекционной составляющей тока, растет со временем в связи с уменьшением емкостной составляющей тока эмиттера.

В процессе накопления неосновных носителей в базе транзистора происходит увеличение тока коллектора: Однако ток коллектора не может возрастать неограниченно, так как в практически осуществляемых схемах в цепь коллектора включается сопротивление нагрузки. Действительно, на сопротивление нагрузки приходится какая-то часть напряжения источника питания выходной цепи транзистора, в результате чего доля напряжения, приходящаяся на коллекторный переход, уменьшается по мере увеличения тока коллектора. При определенных токе эмиттера и напряжении на эмиттере концентрация неосновных носителей в базе около коллектора может превзойти равновесное значение, что будет соответствовать изменению знака напряжения на коллекторном переходе, т. е. переходу транзистора в режим насыщения. В этот момент (кривая 4 на рис. 4.43, в) ток коллектора определяется сопротивлением нагрузки и э. д. с. источника питания в цепи коллектора

$$I_{к. нас} \approx \mathcal{E}_{КБ}/R_{н.} \quad (4.169)$$

Значение установившегося тока коллектора транзистора, находящегося в режиме насыщения, несколько превышает значение тока

насыщения, вычисленного по (4.169). Значение тока коллектора определяется не только э. д. с. источника питания в цепи коллектора, но и падением напряжения на объемном сопротивлении базы при прохождении по нему тока. При прохождении через эмиттер тока в прямом направлении падение напряжения на объемном сопротивлении базы, как видно из рис. 4.44, должно складываться с э. д. с. источника питания в коллекторной цепи:

$$I_k \approx (\mathcal{E}_{КБ} + \Delta U) / R_n. \quad (4.170)$$

После входа транзистора в режим насыщения процесс накопления неосновных носителей заряда в базе и коллекторе транзистора еще некоторое время продолжается. Длительность переднего фронта импульса тока коллектора зависит от амплитуды импульса прямого тока эмиттера и от частотных свойств транзистора.

В момент изменения направления тока эмиттера происходит изменение полярности падения напряжения на объемном сопротивлении базы. При этом скачкообразно уменьшается величина тока коллектора, так как

$$I'_k \approx (\mathcal{E}_{КБ} - \Delta U) / R_n. \quad (4.171)$$

Одновременно начинается процесс рассасывания неосновных носителей заряда, накопленных в области базы транзистора. В первый момент после изменения направления тока эмиттера концентрация неосновных носителей в базе около p - n -переходов эмиттера и коллектора велика. Поэтому сопротивления этих p - n -переходов для обратных токов малы. Значит, величина обратного тока эмиттера и тока коллектора после переключения определяется сопротивлениями во внешних цепях. Концентрация неосновных носителей в базе около p - n -переходов не может мгновенно уменьшиться до нуля. Это соответствовало бы бесконечно большому величинам градиентов концентрации неосновных носителей заряда в базе около p - n -переходов и бесконечно большим токам, чего практически быть не может из-за конечных величин сопротивлений во внешних цепях транзистора. До тех пор, пока в процессе рассасывания концентрации неосновных носителей около p - n -переходов не достигнут нуля, обратные токи через соответствующие p - n -переходы будут оставаться постоянными, т. е. токи эмиттера и коллектора будут неизменными, пока транзистор находится в режиме насыщения. Время, в течение которого транзистор находится в режиме насыщения после окончания импульса прямого тока эмиттера, называют в р е м е н е м р а с с а с ы в а н и я ($t_{рас}$ на рис. 4.43, б).

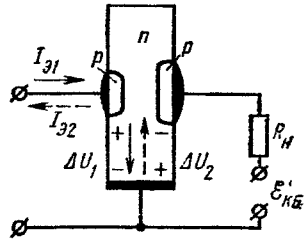


Рис. 4.44. Пояснение скачкообразного изменения тока коллектора из-за изменения полярности падения напряжения на объемном сопротивлении базы при изменении направлении тока эмиттера

В некоторый момент времени концентрации неосновных носителей в базе около p - n -переходов коллектора и эмиттера достигают нуля. С этого момента токи коллектора и эмиттера будут уменьшаться со временем, так как процесс рассасывания неосновных носителей продолжается и уменьшается абсолютное значение градиентов концентрации неосновных носителей около соответствующих p - n -переходов. Изменения в распределении неосновных носителей заряда в базе транзистора в различные моменты времени процесса рассасывания показаны на рис. 4.43, г. Скорость уменьшения тока коллектора в процессе рассасывания характеризуется длительностью заднего фронта t_{q2} .

Значения времени рассасывания $t_{\text{рас}}$ и длительности заднего фронта импульса тока коллектора зависят от величины обратного тока эмиттера $I_{\text{Э2}}$ и от частотных свойств транзистора, т. е. определяются его геометрией и временем жизни неосновных носителей в базе. Время рассасывания зависит также от величины прямого тока эмиттера $I_{\text{Э1}}$, т. е. от полного количества неосновных носителей заряда, накопленных в базовой области до переключения.

Увеличить быстродействие транзистора, работающего в качестве переключателя, т. е. уменьшить время рассасывания $t_{\text{рас}}$, можно путем введения в исходный монокристалл полупроводника примесей рекомбинационных ловушек (золото для кремния). При этом будет уменьшено время жизни неосновных носителей заряда. Однако наряду с высоким быстродействием такие транзисторы обладают рядом недостатков. Во-первых, коэффициенты передачи тока у них оказываются меньше из-за более интенсивной рекомбинации неосновных носителей в базе транзистора. Во-вторых, обратный ток коллектора у них оказывается больше из-за более интенсивной тепловой генерации носителей заряда в электронно-дырочном переходе коллектора и в прилегающих к нему областях базы и коллектора. В-третьих, у них оказывается большая зависимость времени рассасывания от температуры из-за сильной зависимости времени жизни от температуры.

Более удачным методом повышения быстродействия транзистора, работающего в качестве переключателя, является шунтирование коллекторного перехода диодом Шоттки, в котором отсутствует инжекция неосновных носителей и их накопление при прямом смещении. Структура и принцип действия такого транзистора будут рассмотрены в § 7.4, так как наибольшее применение bipolarные транзисторы с диодом Шоттки получили в интегральных микросхемах.

Схема с общим эмиттером

В транзисторе, включенном по схеме с общим эмиттером, при работе на импульсах с большой амплитудой происходят те же процессы накопления неосновных носителей заряда в базе и их рассасывания. На рис. 4.45 показаны временные зависимости тока базы и тока коллектора при включении транзистора по схеме с общим эмиттером. Особенностью временной зависимости в этом случае

по сравнению с аналогичной зависимостью для схемы с общей базой является скачкообразное увеличение тока коллектора при перемене направления входного тока базы с $I_{Б1}$ до $I_{Б2}$. Входному току $I_{Б1}$ соответствует отрицательный потенциал базового вывода $U_{Б1}$ по отношению к общему эмиттерному выводу (рис. 4.46).

Поэтому

$$I_K \approx (\mathcal{E}_{КЭ} - U_{Б1})/R_H.$$

Выключающему току базы $I_{Б2}$ соответствует положительный потенциал базового вывода $U_{Б2}$, поэтому

$$I'_K \approx \frac{\mathcal{E}_{КЭ} + U_{Б2}}{R_H}.$$

Изменение тока коллектора в момент переключения входа транзистора в схеме с общим эмиттером обычно небольшое по сравнению с изменением того же тока в схеме с общей базой.

§ 4.16. ШУМЫ В ТРАНЗИСТОРАХ

Транзистор, как и другие полупроводниковые приборы, обладает собственными шумами, т. е. при отсутствии сигналов на входе транзистора на выходе у него имеется некоторое хаотически изменяющееся напряжение.

Физические явления, обуславливающие появление шумов в транзисторе, по существу, те же, что и в других полупроводниковых приборах. В транзисторе наблюдаются *тепловой*, *дробовой* и *избыточный* шум. Мощность теплового шума, как и в полупроводниковых диодах, определяется формулой

$$P = kT \Delta f. \quad (4.172)$$

Среднеквадратичное значение тока дробового шума

$$\bar{i}^2 = 2kTI_{\ominus} \Delta f. \quad (4.173)$$

Избыточные шумы в транзисторе связаны с процессами генерации и рекомбинации носителей заряда в основном на поверхности полупроводника. Как и в других полупроводниковых приборах, в транзисторах наблюдается уменьшение избыточных шумов с ростом частоты (примерно обратно пропорционально частоте).

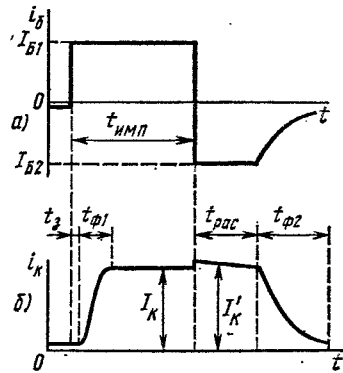


Рис. 4.45. Временные зависимости тока базы (а) и тока коллектора (б) при включении транзистора по схеме с общим эмиттером

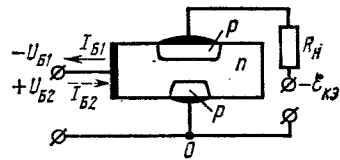


Рис. 4.46. Пояснение скачкообразного изменения тока коллектора при перемене полярности входного напряжения

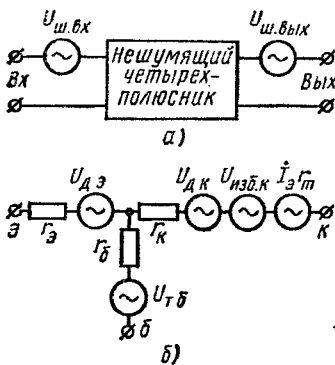


Рис. 4.47. Шумовые эквивалентные схемы транзисторов:
 а — формальная; б — физическая

При описании шумовых свойств транзистора можно использовать его шумовую эквивалентную схему, принцип построения которой примерно тот же, что и обычных эквивалентных схем. При построении формальной эквивалентной схемы транзистор можно рассматривать как идеальный нешумящий четырехполюсник, во входную и выходную цепи которого включены генераторы шумов (рис. 4.47, а). Этот четырехполюсник можно заменить эквивалентной схемой, например, как в § 4.11. Значения э. д. с. (или тока) генераторов шумов можно определить экспериментально.

Физические шумовые эквивалентные схемы дают возможность связать источники шумов в транзисторе с коэффициентом шума. При этом генераторы шумов, включаемые в эквивалентную схему, связывают с определенными процессами (рис. 4.47, б). Шумы эмиттерного перехода транзистора являются дробовыми и отражаются соответствующим генератором, шумы активного сопротивления базы — тепловыми. На коллекторном переходе имеются и дробовые, и избыточные шумы. Э. д. с. генераторов шумов можно определить следующим образом.

Через эмиттерный переход идет ток; значит, там должны быть дробовые шумы. В соответствии с формулой (4.173) э. д. с. от этих шумов, выделяемая на сопротивлении эмиттера (для единичной полосы частот),

$$\mathcal{E}_{дэ} = \sqrt{2qI_{э}r_{э}^2}. \quad (4.176)$$

Аналогично э. д. с. дробовых шумов в цепи коллектора

$$\mathcal{E}_{дк} = \sqrt{2qI_{к}r_{к}^2}. \quad (4.177)$$

База представляет собой активное сопротивление; следовательно, здесь должны наблюдаться тепловые шумы с э. д. с.

$$\mathcal{E}_{тб} = \sqrt{4kTr_{б}}. \quad (4.178)$$

Основным параметром, характеризующим шумы в транзисторах, считают коэффициент шума

$$F = \frac{(P_{\text{сигн}}/P_{\text{ш}})_{\text{вх}}}{(P_{\text{сигн}}/P_{\text{ш}})_{\text{вых}}}, \quad (4.174)$$

где $P_{\text{сигн}}$ — мощность сигнала; $P_{\text{ш}}$ — мощность шумов.

Коэффициент шума показывает, во сколько раз ухудшается отношение сигнал/шум из-за наличия собственных шумов усилителя (в данном случае транзистора). Для идеального усилителя $F = 1$. Иногда коэффициент шума выражают в децибелах. Тогда

$$F = 10 \lg \frac{(P_{\text{сигн}}/P_{\text{ш}})_{\text{вх}}}{(P_{\text{сигн}}/P_{\text{ш}})_{\text{вых}}}. \quad (4.175)$$

Э. д. с. избыточных шумов, отнесенных полностью к цепи коллектора, можно определить по эмпирической формуле

$$\mathcal{E}_{\text{изб. к}} = KU_{\text{КБ}}^a \left(\frac{U_{\text{КБ}}}{I_{\text{К}}} \right)^b \frac{1}{f}, \quad (4.179)$$

где K — коэффициент, зависящий от материала и обработки поверхности полупроводника; a и b — показатели степени, имеющие значения в пределах 1,2—1,8.

Пользуясь эквивалентной схемой (рис. 4.47, б), можно определить шумы, создаваемые транзистором в схеме.

Коэффициент шума транзистора зависит от режима его работы и частоты. Зависимость коэффициента шума от тока транзистора представлена на рис. 4.48, а. На рисунке видно, что коэффициент шума при малых $I_{\text{Э}}$, когда преобладают тепловые и избыточные шумы, сначала сравнительно слабо зависит от тока, а при дальнейшем увеличении тока в связи с увеличением роли дробовых шумов растет примерно пропорционально $I_{\text{Э}}$.

С изменением напряжения коллектора коэффициент шума также сначала изменяется мало (рис. 4.48, б), а затем растет довольно быстро. Это объясняется тем, что при низких напряжениях коэффициент шума определяется тепловыми и дробовыми шумами, а при больших напряжениях начинают преобладать избыточные шумы.

При росте частоты (рис. 4.48, в) коэффициент шума сначала падает из-за снижения роли избыточных шумов, затем в некотором диапазоне частот остается постоянным. Здесь он определяется тепловыми и дробовыми шумами. На частотах, где уменьшается коэффициент усиления транзистора, коэффициент шума снова растет. Это следует из определения коэффициента шума (4.174) — при очень высоких частотах мощность сигнала на выходе падает, в то время как мощность шумов, создаваемых в транзисторе, не изменяется.

Состояние поверхности полупроводника существенно влияет на избыточные шумы (рис. 4.48, в). Поэтому создание малошумящих транзисторов основано на совершенствовании обработки поверхности.

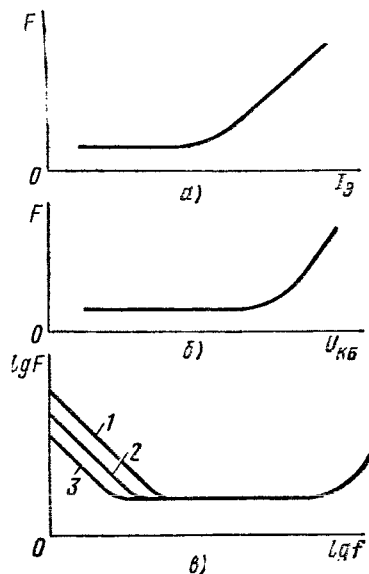


Рис. 4.48. Зависимости коэффициента шума транзистора:

а — от тока эмиттера; б — от напряжения коллектора; в — от частоты (1 — загрязненная поверхность; 2 — нормально обработанная поверхность; 3 — поверхность с улучшенной обработкой)

§ 4.17. НИЗКОЧАСТОТНЫЕ МАЛОМОЩНЫЕ ТРАНЗИСТОРЫ

Низкочастотными считают транзисторы с рабочими частотами до 3 МГц, а маломощными — с допустимой мощностью рассеяния на коллекторе до 0,3 Вт. Большинство относящихся к этой группе транзисторов изготовляют методом сплавления, поэтому их называют с п л а в н ы м и. Так как при изготовлении низкочастотных сплавных транзисторов обычно используют равномерно легированный исходный материал, то при малых токах электрическое поле в области базы таких транзисторов отсутствует и по механизму движения носителей заряда они относятся к бездрейфовым.

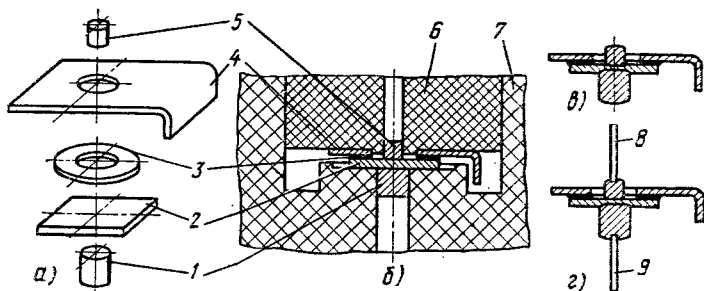


Рис. 4.49. Схема технологического процесса изготовления сплавного транзистора:

a — заготовки; *б* — размещение заготовок в cassette; *в* — транзистор после сплавления; *г* — транзистор с выводами эмиттера и коллектора; 1 — коллекторная навеска; 2 — пластинка полупроводника; 3 — кольцо припоя для получения невыпрямляющего контакта с выводом базы; 4 — кристаллодержатель (вывод базы); 5 — эмиттерная навеска; 6 — пробка cassette; 7 — корпус cassette; 8 — вывод эмиттера; 9 — вывод коллектора

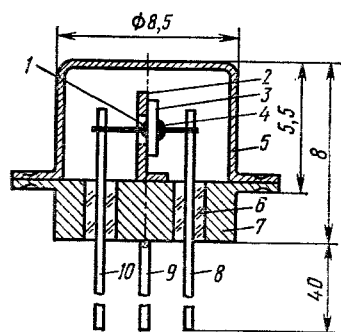


Рис. 4.50. Конструкция сплавных маломощных транзисторов в корпусе с герметизацией холодной сваркой:

1 — эмиттерная навеска; 2 — кристаллодержатель; 3 — пластинка полупроводника; 4 — коллекторная навеска; 5 — баллон; 6 — проходные изоляторы; 7 — ножка корпуса; 8 — вывод коллектора; 9 — вывод базы; 10 — вывод эмиттера

Общая технологическая схема изготовления сплавного транзистора напоминает схему изготовления диода, за исключением того, что в полупроводниковую пластинку производят сплавление двух навесок примесей с двух сторон (рис. 4.49). Вырезанные из монокристалла германия или кремния пластинки шлифуют и травят до необходимой толщины. Эти операции очень ответственны, так как получающаяся после сплавления эмиттера и коллектора толщина базы зависит от толщины исходной пластинки, а свойства транзистора в значительной степени определяются толщиной базы. Поэтому перед сплавлением переходов толщину пластинок тщательно контролируют.

Вплавление эмиттерной и коллекторной навесок производят в защитной атмосфере. Материал навесок выбирают в зависимости от того, какого типа переход следует получить (см. § 2.4). В соответствии с материалом выбирают и температуру сплавления. Детали транзистора (полупроводниковую пластинку, эмиттерную и коллекторную навески, кристаллодержатель — вывод базы и кольцо припоя для получения невыпрямляющего контакта с выводом базы) помещают в графитовую кассету (рис. 4.49, б). Кассета обеспечивает центровку деталей.

После сплавления транзистор практически готов (рис. 4.49, в). Остается припаять выводы к эмиттеру и коллектору (рис. 4.49, г), иногда эту операцию выполняют одновременно с сплавлением навесок.

Так как параметры транзистора зависят от состояния поверхности полупроводника, его помещают в корпус. Сплавные транзисторы выпускают в металлических корпусах, детали которых соединяют холодной сваркой (рис. 4.50). При холодной сварке соединение деталей производят путем сильного их сжатия, сопровождающегося пластической деформацией. Для этого требуется увеличение толщины ножки во избежание растрескивания изоляторов выводов.

§ 4.18. ВЫСОКОЧАСТОТНЫЕ МАЛОМОЩНЫЕ ТРАНЗИСТОРЫ

К высокочастотным относят транзисторы с рабочими частотами свыше 30 МГц. Для обеспечения работы транзистора на таких частотах требуется, как это следует из § 4.9, уменьшить время пролета носителей заряда через базу и область объемного заряда коллектора, уменьшить барьерные емкости и объемные сопротивления базы и коллектора. Выполнить все это на основе технологии сплавления невозможно. Так, уменьшение толщины базы сплавного транзистора не только сложно технологически, но и обуславливает снижение напряжения смыкания переходов. Если для того чтобы избежать смыкания и снизить сопротивление базы, взять исходный материал с большей концентрацией примесей, возрастет барьерная емкость коллектора и снизится его пробивное напряжение. Технология сплавления примесей не позволяет также изготавливать электроннодырочные переходы малой площади, что требуется для уменьшения емкости.

Диффузионные транзисторы

Основным методом изготовления высокочастотных транзисторов является метод, в основе которого лежит диффузия примесей. Такие транзисторы называют диффузионными. При диффузии примеси в базе распределяются неравномерно, там создается электрическое поле; следовательно, по механизму движения носителей заряда диффузионные транзисторы могут относиться к дрейфовым.

Если произвести диффузию доноров и акцепторов, например, в электронный кремний, они распределятся в материале, как показано на рис. 4.3. Акцепторные примеси проникнут на большую глубину, так как коэффициент диффузии для них больше, а донорные примеси создадут большую концентрацию у поверхности. Разность концентрации доноров и акцепторов, определяющая концентрацию носителей, при этом будет изменяться, как показано на рис. 4.3. Таким образом, получается структура *n-p-n*. После присоединения выводов к соответствующим областям получают транзистор.

Диффузионный транзистор имеет ряд преимуществ перед сплавным. Так, диффузионная технология позволяет довольно точно контролировать глубину залегания переходов, т. е. получать базу транзистора малой толщины. Концентрация примесей в базе, если пренебречь небольшой областью вблизи эмиттерного перехода, уменьшается в сторону коллектора, что создает в базе электрическое поле, способствующее движению носителей от эмиттера к коллектору и соответственно уменьшающее время пролета носителей через базу.

Благоприятным для создания высокочастотного транзистора является то обстоятельство, что при диффузионной технологии исходный материал образует область коллектора, концентрация примесей в которой может быть небольшой. В результате область объемного заряда коллекторного перехода при повышении напряжения распространяется в основном не в базу, а в коллектор, что снижает возможность смыкания даже при очень малой толщине базы. Это же способствует тому, что концентрацию примесей в базе диффузионного транзистора можно повысить для снижения ее сопротивления без существенного повышения барьерной емкости коллекторного перехода и падения его пробивного напряжения.

Однако использование исходного высокоомного материала приводит к увеличению последовательного сопротивления коллектора, что нежелательно. Во избежание этого в качестве исходного материала используют пластинки из сильнолегированного полупроводника, на которых с помощью эпитаксии (см. § 2.4) наращивают сравнительно тонкий слой полупроводника с малой концентрацией примесей. Затем двухслойную пластинку используют для изготовления транзистора по диффузионной технологии. Коллекторный электронно-дырочный переход такого транзистора создают в эпитаксиальном слое, что обеспечивает малую емкость и высокое пробивное напряжение коллекторного перехода. В то же время из-за малой толщины

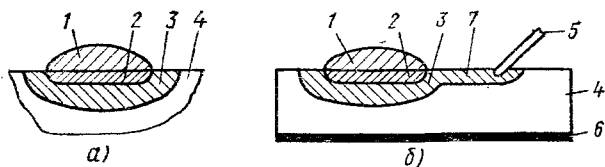


Рис. 4.51. Сплавно-диффузионный транзистор:

a — схема структуры; *б* — структура с выводом базы через соединительный слой; 1 — эмиттерная навеска; 2 — рекристаллизованная *p*-область (эмиттер); 3 — диффузионная область базы; 4 — исходный материал (коллектор); 5 — вывод базы; 6 — вывод коллектора; 7 — соединительный слой

эпитаксиального слоя последовательное сопротивление коллектора оказывается небольшим. Транзисторы, изготовленные по такой технологии, называют **э п и т а к с и а л ь н ы м и**.

Сплавно-диффузионные транзисторы

Способ изготовления диффузионных транзисторов требует совершенных методов обработки поверхности полупроводника. Для получения хороших частотных характеристик толщина базы должна быть порядка долей микрометра. Так как базовый слой располагается довольно близко от поверхности, на его однородность могут повлиять различные поверхностные дефекты, размеры которых могут быть также порядка микрометра. Поэтому для изготовления германиевых транзисторов с рабочими частотами в десятки-сотни мегагерц используют технологию, при которой область базы и коллекторный переход получают путем диффузии примесей, а эмиттерный переход — вплавлением (рис. 4.51).

Диффузия примесей происходит из эмиттерной навески, находящейся в расплавленном состоянии. Транзисторы, полученные таким методом, называют **с п л а в н о - д и ф ф у з и о н н ы м и**.

Для изготовления сплавно-диффузионного транзистора на основе германия *p*-типа в эмиттерную навеску вводится сурьма, имеющая большой коэффициент диффузии, и индий, у которого коэффициент диффузии мал. При вплавлении такой навески сурьма диффундирует из нее на довольно большую глубину, а индий — на малую. Из-за разных растворимостей концентрация индия в рекристаллизованной области оказывается много большей, чем сурьмы. Таким образом образуется распределение примесей, близкое к идеальному для дрейфового транзистора. При такой технологии требования к состоянию поверхности исходной пластинки значительно ниже, чем при чисто диффузионной, так как эмиттерная навеска проплавляет пластинку полупроводника на некоторую глубину и диффузия происходит в места, не нарушенные поверхностной обработкой.

При изготовлении сплавно-диффузионного и чисто диффузионного транзисторов значительные трудности может вызвать необходимость присоединения вывода к базе, имеющей малую толщину.

Трудности получения надежного вывода базы в значительной степени устраняются, если на поверхности пластинки полупровод-

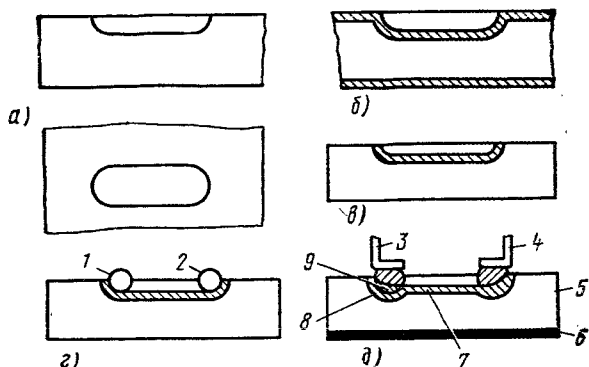


Рис. 4.52. Схема технологического процесса изготовления сплавно-диффузионного транзистора:

a — исходная пластина полупроводника с лункой; *б* — получение соединительного слоя путем диффузии сурьмы; *в* — пластина после шлифовки; *г* — пластина с навесками перед вплавлением; *д* — структура после вплавления; 1 — эмиттерная навеска ($\text{In} + \text{Ga} + \text{Sb}$); 2 — базовая навеска ($\text{Sb} + \text{Pb}$); 3 — вывод эмиттера; 4 — вывод коллектора; 5 — коллектор (исходный материал); 6 — вывод коллектора; 7 — диффузионный соединительный слой; 8 — диффузионная базовая область; 9 — рекристаллизованная *p*-область (эмиттер)

ника создан слой с электропроводностью того же типа, что и база. Такой слой может быть образован, например, путем диффузии примесей из газообразной фазы (в рассмотренном случае такой примесью является сурьма). Вывод базы сравнительно легко можно присоединить к этому соединительному слою с помощью впаивания (рис. 4.51, б).

Схема технологического процесса изготовления сплавно-диффузионного транзистора приведена на рис. 4.52. Заготовку из германия *p*-типа обрабатывают так, что на ней образуются лунки продолговатой формы (рис. 4.52, *a*), затем пластинку подвергают обработке в парах сурьмы, при диффузии которой образуется *n*-область (рис. 4.52, *б*). Далее производят шлифовку пластинки, при которой слой *n*-типа снимается везде, кроме лунок (рис. 4.52, *в*). В лунку вкладывают две навеси: эмиттерную и базовую (рис. 4.52, *г*). Эмиттерная состоит из сплава $\text{In} + \text{Ga} + \text{Sb}$, а базовая — из сплава $\text{Pb} + \text{Sb}$. Во время впаивания из эмиттерной навески происходит диффузия сурьмы, а индий и галлий при рекристаллизации дают *p*-область эмиттера. Базовая навеска при впаивании образует обычный невыпрямляющий контакт с *n*-областью (рис. 4.52, *д*).

Корпуса сплавно-диффузионных транзисторов могут иметь конструкцию, аналогичную рассмотренной в § 4.17 конструкции сплавных транзисторов. Существенным их отличием является то, что здесь кристалл удобнее крепить к кристаллодержателю коллектором, и поэтому у диффузионных транзисторов корпус часто является выводом коллектора, а не базы, как у большинства маломощных сплавных транзисторов.

Транзисторы «меза»

Транзисторы «меза» в результате соответствующей технологии имеют столбовую структуру (рис. 4.53). В исходную пластинку полупроводника осуществляется диффузия вещества, изменяющего тип электропроводности исходного материала. Затем производят шлифовку одной стороны так, что получается структура p - n ; на пластинку испарением в вакууме наносят полоски для получения эмиттерного и базового контактов. Для кремниевых транзисторов, например, эмиттерный контакт делают из алюминия, а базовый — из золота. После этого производят сплавление. Транзистор, изготовленный таким способом, еще не обладает хорошими частотными свойствами, так как механическим путем трудно получить малую площадь коллектора, а следовательно, и малую его емкость. Поэтому производится травление части кристалла, при этом область его около контактов эмиттера и базы защищается. После травления получается столбовая структура (рис. 4.53, *д*). Затем присоединяют выводы и герметизируют транзистор в корпусе.

Одним из преимуществ транзисторов «меза» является хороший теплоотвод от коллектора, так как контакт с его выводом осуществляется на большой площади. Столбовая структура имеет также большую механическую прочность, а пластинка полупроводника — большой размер, что облегчает ее монтаж. Транзисторы «меза» позволяют получить граничные частоты усиления по току

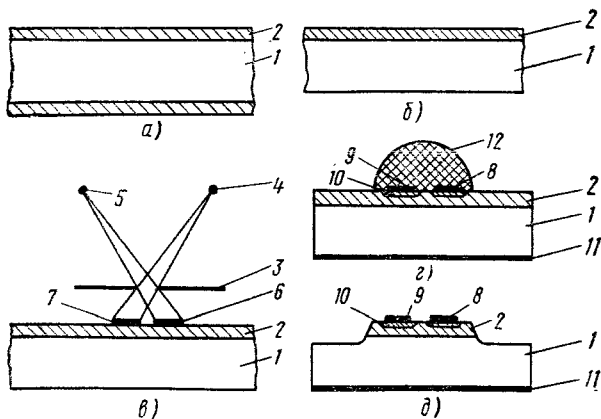


Рис. 4.53. Технологический процесс изготовления транзисторов с «меза»-структурой:

а — заготовка после диффузии; *б* — заготовка после шлифовки одной стороны; *в* — нанесение эмиттерной и базовой полосок; *г* — транзистор «меза» после травления; 1 — исходный материал; 2 — диффузионный базовый слой; 3 — маска; 4 — испаритель материала эмиттерной навески; 5 — испаритель материала базовой навески; 6 — базовая навеска; 7 — эмиттерная навеска; 8 — базовый контакт; 9 — эмиттерный контакт; 10 — рекристаллизованная область эмиттера; 11 — вывод коллектора; 12 — защита от травления.

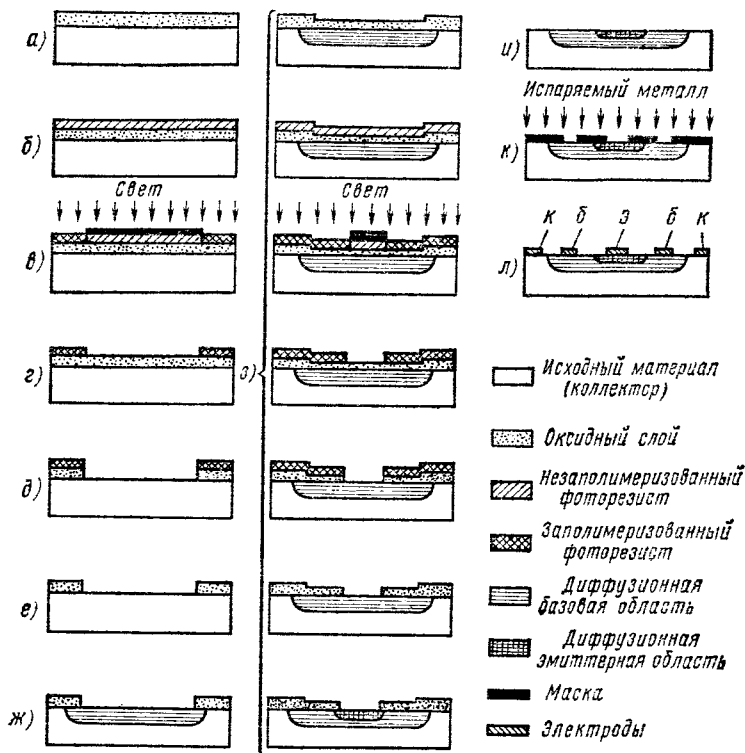


Рис. 4.54. Схема технологического процесса изготовления планарного транзистора:

а — окисления пластинка полупроводника; б — окисленная пластинка с нанесенным фоторезистом; в — засветка фоторезиста; г — пластинка после растворения незаполимеризованного фоторезиста; д — травление пленки оксида; е — пластинка для проведения диффузии; ж — пластинка после проведения диффузии; з — получение эмиттера; и — удаление окисной пленки; к — нанесение выводов; л — структура готового транзистора

500 МГц, а емкости коллектора менее 1 пФ. Однако при производстве таких транзисторов неизбежны трудности в получении базы, если поверхность кристалла обработана плохо.

Планарные транзисторы

Благодаря совершенствованию обработки полупроводниковых материалов появилась возможность изготавливать транзисторы путем чисто диффузионной технологии, не применяя операций впавления. Наиболее широко распространена планарная технология (см. § 2.4). Выполненный по данной технологии транзистор называют п л а н а р н ы м; основные этапы этой технологии показаны на рис. 4.54.

Отличие изготовления планарного транзистора состоит в том, что покрытие пластинки полупроводника слоем окисла, нанесение фоторезиста, его засветка, травление и диффузия повторяется с тем, чтобы получить базовую и эмиттерную область. Чаще планарную технологию используют при производстве кремниевых транзисторов, так как в этом случае слои SiO_2 получают наиболее простым методом — окислением исходной пластинки полупроводника. Однако можно делать и германиевые планарные транзисторы — в этом случае слой окиси кремния наносят на германий каким-либо другим методом. Исходные пластинки полупроводника для изготовления планарных транзисторов могут иметь высокоомный эпитаксиальный слой (планарно-эпитаксиальные транзисторы).

При планарной технологии существенно улучшаются параметры транзисторов, особенно их частотные характеристики. Это обусловлено тем, что в данном случае можно проводить селективную диффузию, т. е. вводить примеси только в небольшие ограниченные области. В то же время оптические методы, применяемые при фотолитографии, позволяют с большой точностью совмещать эти области. В результате оказалось возможным изготовление транзисторов с толщиной базы в доли микрометра и размерами переходов в единицы микрометров. Рабочие частоты планарных транзисторов достигают примерно 10 ГГц.

При таких высоких рабочих частотах устанавливаются очень жесткие ограничения на значения сопротивлений и емкостей отдельных областей транзистора. Чтобы уменьшить сопротивление базы в планарных транзисторах, часто используют многоэмиттерную конструкцию, где вместо одного имеется несколько эмиттеров в виде узких полос, расположенных между выводами базы.

Создание СВЧ-транзисторов потребовало изменений конструкции их корпусов. В области сверхвысоких частот большую роль играют емкости между выводами и индуктивность выводов. Поэтому для СВЧ-транзисторов применяют корпуса с ленточными выводами и пониженной индуктивностью.

Однако преимущества планарного транзистора не ограничиваются быстротой. Например, границы его электронно-дырочных переходов оказываются под слоем окисла. Так как окись кремния является диэлектриком с очень высокими характеристиками, она может служить защитой поверхности полупроводника от внешних воздействий. Такая защита усиливается при покрытии окисла тонким слоем легкоплавкого стекла. Эта особенность планарных транзисторов дала возможность в некоторых случаях отказаться от помещения их в корпус, что резко снизило габариты прибора (рис. 4.55). Бескорпусные транзисторы широко применяют в герметизированных гибридных микросхемах. Изготавливают планарные транзисторы и в пластмассовых корпусах.

Дальнейшее совершенствование технологии СВЧ-транзисторов направлено на еще более точное изготовление базы со сравнительно большой концентрацией примесей. С этой целью можно применять метод легирования полупроводника путем внедрения ионов примеси

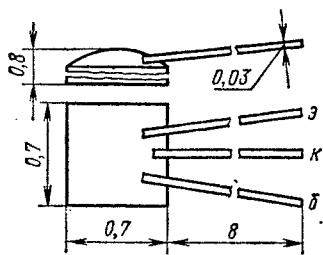


Рис. 4.55. Бескорпусный планарный транзистор

(имплантация). Ионы легирующих материалов разгоняют электрическим полем в вакууме до определенной энергии, а затем пучок таких ионов направляют на поверхность полупроводниковой пластинки. В полупроводнике ионы проходят определенный путь, а затем тормозятся. Таким образом, в глубине полупроводника образуется легированный слой (например, базовый). Возникающие при этом дефекты кристаллической решетки устраняют отжигом.

Ионное легирование — один из перспективных методов создания СВЧ-транзисторов.

§ 4.19. МОЩНЫЕ ТРАНЗИСТОРЫ

К мощным относят транзисторы, способные рассеивать не менее 1,5 Вт. При разработке мощных транзисторов помимо тех задач, которые встречаются при создании маломощных транзисторов, приходится решать следующие проблемы:

1) мощные транзисторы работают при довольно больших токах, поэтому в них особенно сильно проявляются связанные с этим эффекты (см. § 4.5);

2) для обеспечения достаточной мощности в нагрузке обычно используют источники питания с большим напряжением; поэтому чаще всего мощные транзисторы должны быть рассчитаны на большие рабочие напряжения, чем маломощные;

3) при использовании мощных транзисторов необходимо учитывать обеспечиваемые ими значения коэффициента полезного действия. В частности, требуется по возможности уменьшать падение напряжения на открытом транзисторе, т. е. стремиться снизить его сопротивление насыщения;

4) конструкция мощного транзистора должна обеспечивать эффективный отвод рассеиваемого тепла;

5) значительный перегрев активной области мощного транзистора и большие размеры применяемых в нем пластин полупроводника обуславливают необходимость борьбы с возможными разрушениями из-за механических напряжений.

Кроме того, мощные транзисторы должны быть и достаточно быстродействующими.

Сплавные транзисторы

Среди низкочастотных мощных транзисторов значительное место занимают сплавные германиевые транзисторы. Для обеспечения больших токов в транзисторах необходимы большие площади эмиттера. Однако увеличение размеров эмиттера привело бы к значительному вытеснению тока на его края из-за падения напряжения на сопро-

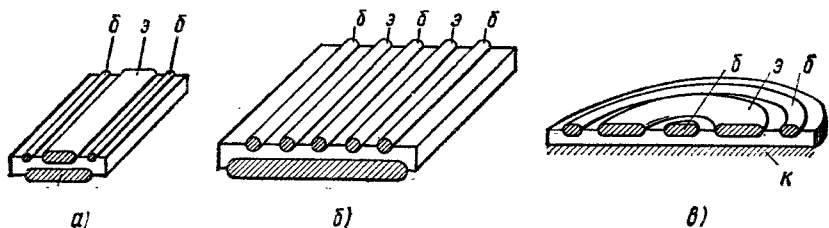


Рис. 4.56. Конфигурации переходов мощных сплавных транзисторов:
а, б — эмиттерный переход и вывод базы в виде полос; *в* — эмиттерный переход и вывод
 базы в виде колец

тивлению базы. До некоторой степени это явление можно устранить, применяя материал базы с большой диффузионной длиной и повышая эффективность эмиттера. Наиболее же существенным является изменение конфигурации эмиттера — изготовление его в виде ряда полос (рис. 4.56, *а, б*) или в виде колец (рис. 4.56, *в*). Но изготовление таких транзисторов ограничено тем, что при технологии сплавления затруднено выполнение достаточно большого числа переходов в виде узких полос.

Для обеспечения лучшего теплоотвода пластинку полупроводника припаивают к корпусу коллекторной стороной. Основание корпуса выполняют из меди. Так как теплоотвод от корпуса должен обычно осуществляться на шасси всего устройства или на радиатор, нижнюю поверхность корпуса для уменьшения теплового сопротивления не окрашивают.

Конверсионные транзисторы

Технологически конверсионные транзисторы напоминают сплавно-диффузионные. Для их изготовления применяют германий, в который вводят сразу два типа примесей — донорную и акцепторную (медь). Медь отличается высоким коэффициентом диффузии в германии, благодаря этому при сплавлении эмиттера происходит диффузия меди из германия в эмиттерную навеску. В результате из слоя, примыкающего к эмиттеру, акцепторная примесь удаляется, изменяется тип электропроводности, т. е. происходит так называемая *к о н в е р с и я*. Такой метод позволяет получать тонкие базовые слои большой площади, а следовательно, изготавливать мощные высокочастотные транзисторы.

Планарные транзисторы

Создание мощных транзисторов с помощью планарной технологии позволяет изготавливать эмиттеры сложной формы. Планарная технология при использовании пластин с эпитаксиальными слоями дала возможность разработать мощные СВЧ-транзисторы.

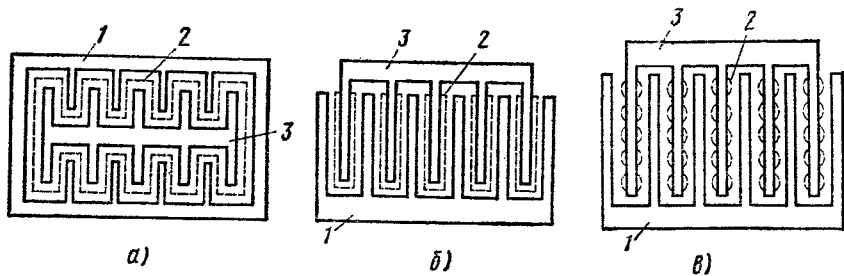


Рис. 4.57. Геометрия электродов мощного СВЧ-транзистора:

а — конструкция с гребенчатым эмиттером; *б* — многоэмиттерная конструкция с плосковыми эмиттерами; *в* — многоэмиттерная конструкция с круглыми эмиттерами; 1 — вывод базы; 2 — эмиттер; 3 — вывод эмиттера

Одним из вариантов конструкции, применяемой в планарных мощных транзисторах, является гребенчатый эмиттер (рис. 4.57, *а*); число зубьев «гребенки» может быть большим — до нескольких десятков. Так как ширина каждого зубца небольшая, эффекты вытеснения тока к краю эмиттера сильно подавляются. Одновременно снижается и сопротивление базы, что улучшает усиление по мощности на высокой частоте и снижает сопротивление насыщения. Однако при слишком узких и длинных эмиттерных полосках может сказываться падение напряжения вдоль них при больших токах.

Хорошее распределение токов получают при создании многоэмиттерных транзисторов, в которых используют большое число не связанных непосредственно эмиттеров в виде полосок (рис. 4.57, *б*), либо кругов, квадратов (рис. 4.57, *в*). Эти эмиттеры соединяют путем металлизации поверх слоя окисла. Так как отдельные эмиттеры включены параллельно в прямом направлении, распределение токов между ними может быть неравномерным. Для их выравнивания отдельные эмиттеры подключают через сопротивления, представляющие собой также слои металла, нанесенного на окисел.

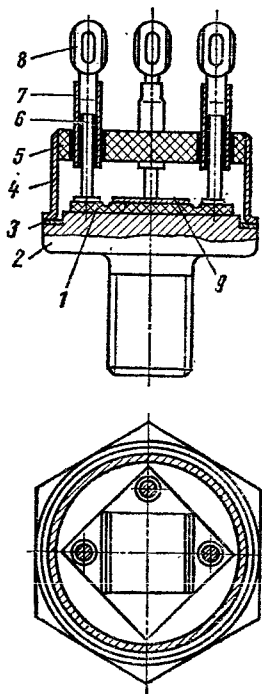


Рис. 4.58. Конструкция мощного СВЧ-транзистора:

1 — кристаллодержатель из бериллиевой керамики; 2 — фланец; 3 — кольцо; 4 — баллон; 5 — керамический изолятор; 6 — внутренний вывод; 7 — трубка; 8 — наружный вывод; 9 — кристалл полупроводника

Наиболее совершенными являются многоструктурные мощные транзисторы, изготовленные в виде нескольких мощных многоэмиттерных транзисторов на одной пластинке полупроводника. Такая конструкция позволяет улучшить теплоотвод, так как тепловые потоки от отдельных разнесенных достаточно далеко структур не перекрываются.

Лучшие из современных мощных СВЧ-транзисторов рассчитаны на мощность в несколько ватт на частоте в несколько гигагерц.

Чтобы радиатор транзистора не был включен в цепь нагрузки, в СВЧ планарных транзисторах желательно изолировать коллектор от корпуса. При этом должен быть обеспечен хороший теплоотвод. Достигают этого путем крепления кристалла полупроводника на пластинке из бериллиевой керамики, имеющей очень высокий коэффициент теплопроводности (рис. 4.58).

Для повышения допустимого напряжения коллектора планарного транзистора иногда применяют травление края коллекторного перехода; при этом создается структура «меза», а транзистор называют мезопланарным. При травлении удаляются части базы и коллектора, имеющие наибольшую концентрацию примесей и дефектов; на краю перехода получается скос, который приводит к повышению напряжения поверхностного пробоя.

§ 4.20. ОДНОПЕРЕХОДНЫЙ ТРАНЗИСТОР

Однопереходный транзистор — это трехэлектродный полупроводниковый прибор с одним p - n -переходом и двумя невыпрямляющими контактами к базовой области, предназначенный для усиления и генерации электромагнитных колебаний при модуляции сопротивления базы в результате инжекции через p - n -переход неосновных носителей заряда.

Структура однопереходного транзистора и его эквивалентная схема показаны на рис. 4.59. Область эмиттера (область с электропроводностью p -типа) должна быть легирована сильнее, чем область базы (область с электропроводностью n -типа), для того, чтобы при прямом включении эмиттерного перехода прямой ток через него имел в основном лишь дырочную составляющую. В этом случае из-за инжекции неосновных носителей заряда в базу транзистора и из-за накопления основных носителей, которые входят в базу через один из невыпрямляющих контактов к базе для компенсации инжектированного заряда неосновных носителей, будет происходить уменьшение сопротивления базы (модуляция) и увеличение тока между невыпрямляющими контактами к базе или тока в цепи нагрузки.

Если на базовые выводы прибора подано межбазовое напряжение $U_{Б2Б1}$, то из-за прохождения тока $I_{Б2}$ вдоль базы будет существовать продольное падение напряжения. Обозначим падение напряжения на части базы протяженностью l_1 (рис. 4.59, a) как $U_{вн}$. Это падение напряжения смещает p - n -переход в обратном направлении.

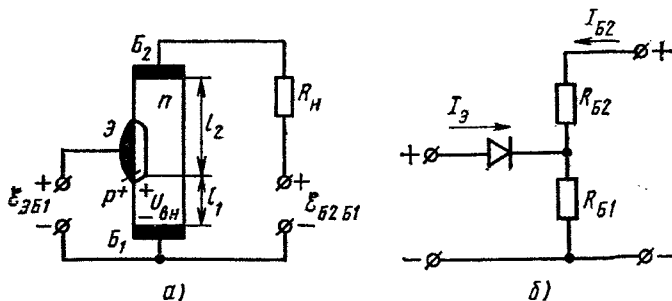


Рис. 4.59. Структура (а) и эквивалентная схема (б) однопереходного транзистора

Поэтому при напряжении на эмиттере $U_{ЭБ1} < U_{вн}$ p - n -переход смещен в обратном направлении и во входной цепи проходит небольшой обратный ток $I_{ЭБ0}$ (рис. 4.60, а).

Если же входное напряжение, поданное на эмиттер относительно базы $Б_1$, превысит $U_{вн}$, то p - n -переход открывается и начинается инжекция неосновных носителей заряда (дырок) в базу. Вначале инжекция происходит только через часть p - n -перехода, расположенную ближе к первому контакту базы $Б_1$. В результате сопротивление части базы протяженностью l_1 уменьшается, что приводит к еще большему смещению p - n -перехода эмиттера в прямом направлении и к появлению на входной статической характеристике участка с отрицательным дифференциальным сопротивлением (рис. 4.60, а).

Таким образом, однопереходный транзистор может находиться в двух устойчивых состояниях — в закрытом, которое характеризуется относительно большими сопротивлениями между различными выводами однопереходного транзистора, и в открытом (или в состоянии насыщения), характеризующемся относительно малыми сопротивлениями. В открытом состоянии однопереходный транзистор будет находиться до тех пор, пока инжекция носителей заряда через эмиттерный переход будет поддерживать в базе избыточную концентрацию неосновных и основных носителей заряда, т. е. до тех пор, пока ток эмиттера будет превышать значение тока выключения $I_{выкл.}$

Выходные, или межбазовые, статические характеристики однопереходного транзистора представляют собой зависимости тока через второй базовый вывод от межбазового напряжения при определенных значениях тока эмиттера (рис. 4.60, б). При $I_Э = 0$ выходная характеристика — прямая линия. При прямых токах эмиттера, отличных от нуля, выходные характеристики оказываются нелинейными, так как суммарное напряжение на эмиттерном переходе изменяется с изменением выходного тока $I_{Б2}$.

Однопереходный транзистор, как всякий прибор с отрицательным дифференциальным сопротивлением, может быть использован в качестве переключателей, генераторов и усилителей. При этом он

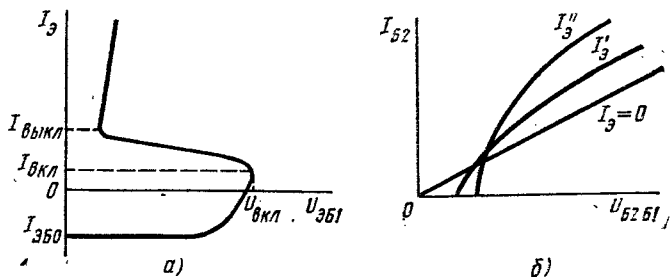


Рис. 4.60. Входная (а) и выходные (б) статические характеристики однопереходного транзистора

может обеспечить усиление как по мощности и напряжению, так и по току.

Определим коэффициент передачи тока однопереходного транзистора. При большой напряженности электрического поля в базе можно пренебречь диффузией носителей заряда и учитывать только дрейфовые токи:

$$I_{Б2} = I_p + I_n = qS(p_{нб}\mu_p + n_{нб}\mu_n)E,$$

где S — площадь поперечного сечения базы (для тока $I_{Б2}$).

Допустим, что ток эмиттера увеличился на $\Delta I_{Э}$, тогда в базе увеличится концентрация дырок на $\Delta p_{нб}$ и одновременно для сохранения электрической нейтральности в базу из невыпрямляющих контактов войдет такое же количество основных носителей заряда ($\Delta n_{нб} = \Delta p_{нб}$).

Так как

$$\Delta I_{Э} = qS \Delta p_{нб} \mu_p E$$

и

$$\Delta I_{Б2} = qS (\Delta p_{нб} \mu_p + \Delta n_{нб} \mu_n) E,$$

то коэффициент передачи тока

$$k_i = \Delta I_{Б2} / \Delta I_{Э} = 1 + (\mu_n / \mu_p).$$

При этом предполагалось, что эффективность эмиттера близка к единице и коэффициент переноса носителей заряда по базе тоже близок к единице в связи с малым временем дрейфа носителей в базе по сравнению со временем их жизни.

Так как подвижность электронов в большинстве полупроводников превышает подвижность дырок в том же материале, то для однопереходных транзисторов выгоднее применять исходный материал с электропроводностью n -типа.

Как видно, принцип действия однопереходного транзистора хотя и связан с инжекцией неосновных носителей в базу, но существенно отличается от принципа действия обычного биполярного транзистора. Кроме того, в связи с относительно большим объемом базы однопереходные транзисторы значительно уступают обычным биполярным транзисторам по частотным свойствам.

Катастрофические отказы

Катастрофические отказы транзисторов являются чаще всего следствием либо недостаточной обработки их конструкции, либо нарушений в технологии. Как и в полупроводниковых диодах (см. § 3.31), катастрофические отказы транзисторов вызываются обрывами выводов, растрескиванием кристалла, пробоем переходов или окисного слоя, короткими замыканиями.

Обрывы или перегорание внутренних выводов чаще всего случаются при использовании очень тонкой проволоки, которая бывает неоднородная по толщине. Такую проволоку применяют, например, для создания выводов планарных транзисторов. Вероятность такого отказа растет, если прибор подвергается воздействию вибрации, ударов и т. п., а также при изменениях температуры. Иногда наблюдаются обрывы вывода сплавного транзистора в месте припайки к навеске. Причиной такого обрыва может быть нарушение технологии пайки, либо отсутствие термокомпенсирующего изгиба (тогда обрыв происходит при низкой температуре).

В планарных транзисторах короткие замыкания иногда возникают между металлическими слоями на поверхности окисла и полупроводником. Причиной таких замыканий могут быть отверстия в окисном слое, появившиеся в результате нарушений слоя в процессе фотолитографии. Например, пылинки, осевшая на поверхность фоторезиста, приводит к тому, что под ней фоторезист не засвечивается, следовательно, не полимеризуется. После травления окисла на месте, где находилась пылинка, получается отверстие. Не допустить дефекты такого рода можно, создав соответствующую чистоту в производственных помещениях.

Растрескивание кристалла чаще всего является следствием разницы в температурных коэффициентах линейного расширения полупроводника и электродного материала. Такой вид отказов характерен для мощных транзисторов, и, чтобы не допустить эти дефекты, применяют, например, термокомпенсирующие прокладки.

К особому виду катастрофических отказов транзисторов относятся те, которые связаны с эффектами шнурования тока (см. § 3.13). Так как в месте шнурования развивается очень высокая температура, она может приводить к плавлению электродного материала. Расплавленная область прорастает через базу и обуславливает короткое замыкание эмиттера с коллектором.

Условные отказы

Катастрофические отказы не являются основными для транзисторов — на их долю приходится примерно 20 % всех наблюдаемых отказов. В основном же отказы транзисторов обусловлены постепенным изменением параметров сверх допустимых отклонений.

В транзисторах практически не наблюдается каких-либо объемных процессов, приводящих к изменению параметров. Токи транзи-

стора являются чисто электронными, в них нет каких-либо ионных составляющих, а используемые примеси имеют при рабочей температуре транзистора ничтожные коэффициенты диффузии. Поэтому постепенные изменения параметров транзисторов связаны с явлениями на поверхности полупроводника.

В принципе влияние состояния поверхности на параметры транзистора подобно аналогичному влиянию у полупроводниковых диодов (см. § 3.14). Однако следует заметить, что помимо изменения токов переходов $I_{КБ0}$, $I_{КЭ0}$, $I_{КБК}$ для транзисторов характерно и изменение коэффициента передачи, связанное с изменением скорости поверхностной рекомбинации.

Одной из причин изменения состояния поверхности, вызывающей изменения параметров транзисторов, является наличие следов влаги в баллоне. Особенно сильно это сказывается на параметрах сплавных транзисторов. Стабильность параметров существенно улучшается, когда в корпусе помещается влагопоглотитель (цеолит), но даже такая мера иногда может быть недостаточно эффективной, если прибор подвергается циклическим изменениям температуры. При этом происходит перераспределение влаги между влагопоглотителем и поверхностью полупроводника, что достаточно для изменения параметров транзистора.

Другой причиной изменения состояния поверхности, характерной для планарных транзисторов, является движение ионов щелочных металлов в слое окисла, особенно заметное при подаче напряжения и высокой температуре.

К снижению надежности транзисторов могут приводить неблагоприятные условия работы — большая рассеиваемая мощность, повышенная температура, вибрация и т. д. Особенно сильным оказывается такое влияние, когда режим работы транзистора приближается к предельно допустимому. На рис. 4.61 показана обобщенная зависимость интенсивности отказов от отношения рабочей мощности рассеяния, отнесенной к максимально допустимой мощности на коллекторе, и от нормализованной температуры:

$$T_n = \frac{T_{р\text{аб}} - T_{сн}}{T_{\text{max}} - T_{сн}},$$

где $T_{р\text{аб}}$ — рабочая температура; T_{max} — максимально допустимая температура; $T_{сн}$ — температура, начиная с которой по техническим условиям требуется снижение электрического режима.

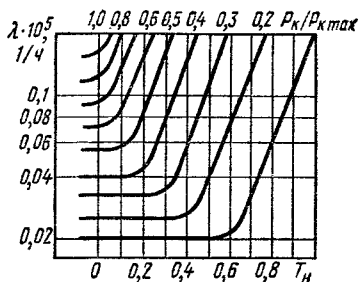


Рис. 4.61. Обобщенная зависимость интенсивности отказов транзисторов от нормализованной температуры при различных отношениях рабочей мощности рассеяния к максимально допустимой мощности на коллекторе

ВЛИЯНИЕ ВНЕШНИХ ФАКТОРОВ НА ТРАНЗИСТОРЫ

Вид внешнего воздействия	Вызываемые или ускоряемые процессы	Типичные дефекты
Повышенная температура	Высыхание защитных покрытий и деформация их. Выделение газов. Расплавление. Растрескивание кристаллов. Миграция ионов примесей и влаги. Изменение электрических характеристик полупроводника	Тепловой пробой. Увеличение обратных токов. Потеря герметичности. Обрывы и короткие замыкания
Пониженная температура	Конденсация влаги. Растрескивание кристаллов. Изменение электрических характеристик полупроводника	Лавинный пробой. Снижение коэффициента передачи. Потеря герметичности. Обрывы и короткие замыкания
Повышенная влажность	Адсорбция и абсорбция влаги. Химические реакции с влагой Электролиз, Коррозия	Изменение электрических параметров. Появление нестабильности. Коррозия выводов и корпуса. Повреждения лакокрасочных покрытий
Резкие изменения температуры	Механические напряжения в местах спаев. Растрескивание кристаллов. Растрескивание и деформация покрытий	Потеря герметичности. Обрывы и короткие замыкания. Изменение электрических параметров
Пониженное давление	Ухудшение теплоотдачи	Перегрев
Механические усилия	Механические напряжения. Усталость	Обрывы и короткие замыкания. Потеря герметичности

Причиной отказов транзистора бывают механические перегрузки при постоянных ускорениях свыше 10 000 g и ударных нагрузках свыше 1000 g.

В табл. 4.2 обобщены факторы внешней среды, влияющие на параметры транзисторов.

Отказы полупроводниковых приборов могут происходить не только при их эксплуатации в неблагоприятных условиях, но и при хранении.

При изучении влияния радиации на надежность полупроводниковых приборов следует различать временные и необратимые изменения параметров.

Временные изменения наблюдаются во время воздействия γ -радиации и связаны в основном с появлением избыточной концентрации носителей заряда в объеме полупроводника, что обуславливает, например, рост обратного тока. Обратимым может быть также влияние ионизации газа в объеме баллона транзистора. Однако при этом возможно оседание ионов на поверхность полупроводника, что приводит к длительным изменениям параметров транзисторов.

Облучение тяжелыми частицами (нейтронами) приводит к образованию различных дефектов в объеме полупроводника (см. § 10.5). При этом могут происходить необратимые изменения параметров транзисторов; например, из-за появления новых рекомбинационных ловушек может снижаться коэффициент передачи (рис. 4.62). Более высокая радиационная стойкость высокочастотных транзисторов связана с меньшей толщиной базы и, следовательно, с меньшей вероятностью появления в ней дефектов.

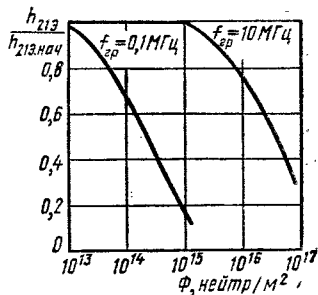


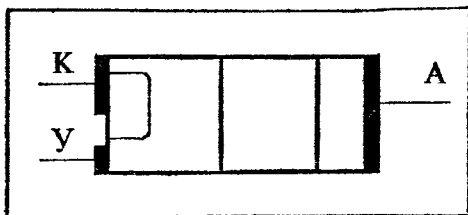
Рис. 4.62. Относительное изменение коэффициента передачи тока базы в зависимости от дозы нейтронов для низкочастотных и высокочастотных транзисторов

Контрольные вопросы

1. Каким образом в транзисторе происходит усиление электромагнитных колебаний по мощности?
2. Почему транзистор, включенный по схеме с общим эмиттером, может обеспечить усиление по току?
3. По каким причинам в базе транзистора возникает электрическое поле?
4. Какие физические процессы происходят в транзисторе при его пробое, если он включен по схемам с общей базой и с общим эмиттером?
5. Как объяснить вид входных и выходных статических характеристик транзистора, включенного по схемам с общей базой и общим эмиттером?
6. Какие факторы определяют инерционность транзистора при его работе на высоких частотах?
7. Какие существуют системы малосигнальных параметров транзистора и в чем преимущества системы h -параметров?
8. Какие существуют эквивалентные схемы транзистора?
9. Что такое одномерная теоретическая модель транзистора?
10. В чем отличие коэффициента передачи тока эмиттера одномерной теоретической модели транзистора и коэффициента передачи тока эмиттера транзистора?
11. Как и почему коэффициент передачи тока эмиттера зависит от величины постоянного тока эмиттера?

ТИРИСТОРЫ

5



§ 5.1. СТРУКТУРА И ПРИНЦИП ДЕЙСТВИЯ

Тиристор — это полупроводниковый прибор с тремя или более $p-n$ -переходами, вольт-амперная характеристика которого имеет участок с отрицательным дифференциальным сопротивлением и который используется для переключения.

Двухэлектродные тиристоры (динисторы)

Двухэлектродный тиристор (динистор) — это тиристор, имеющий два внешних вывода.

Структура двухэлектродного тиристора состоит из четырех слоев кристалла полупроводника с чередующимся типом электропроводности (рис. 5.1). Внешнее напряжение (прямое) на такой прибор подается минусом на крайнюю область с электропроводностью n -типа (на катодный электрод) и плюсом на крайнюю область с электропроводностью p -типа (на анодный электрод). В этом случае крайние $p-n$ -переходы включены в прямом направлении, поэтому их называют эмиттерными; средний $p-n$ -переход включен в обратном направлении, поэтому его называют коллекторным. Соответственно в таком приборе существуют две эмиттерные области (n - и p -эмиттеры) и две базовые (p - и n -базы).

Большая часть внешнего напряжения падает на коллекторном переходе, поэтому первый участок вольт-амперной характеристики тиристора аналогичен обратной ветви вольт-амперной характеристики диода (см. рис. 5.1). С увеличением анодного напряжения, приложенного к тиристорному между анодом и катодом, увеличивается прямое напряжение и на эмиттерных переходах. Электроны, инжектированные из n -эмиттера в p -базу, диффундируют к $p-n$ -переходу коллектора, проходят коллекторный переход и попадают в n -базу. Дальнейшему прохождению электронов по тиристорной структуре препятствует небольшой потенциальный барьер правого эмиттерного перехода (рис. 5.1). Поэтому часть электронов, оказавшись

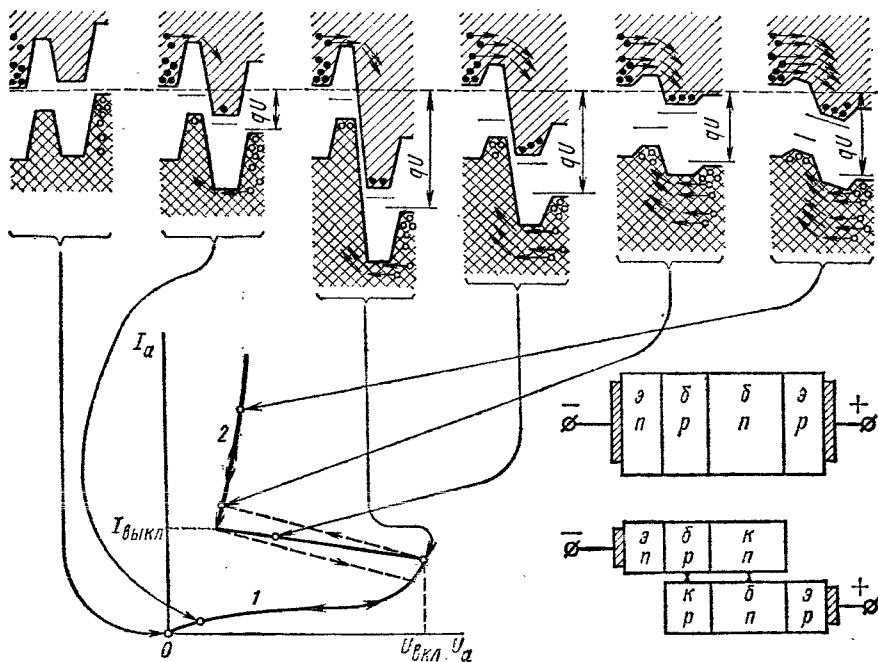


Рис. 5.1. Структура, вольт-амперная характеристика и энергетические диаграммы двухэлектродного тиристора

в потенциальной яме n -базы, образует избыточный отрицательный заряд, который, понижая высоту потенциального барьера правого эмиттерного перехода, вызывает увеличение инжекции дырок из p -эмиттера в n -базу. Инжектированные дырки диффундируют к p - n -переходу коллектора, проходят через этот переход и попадают в p -базу. Дальнейшему их продвижению по тиристорной структуре препятствует небольшой потенциальный барьер левого эмиттерного перехода. Следовательно, в p -базе происходит накопление избыточного положительного заряда, что обуславливает увеличение инжекции электронов из n -эмиттера. В результате накопления избыточного положительного заряда в p -базе и отрицательного заряда в n -базе при напряжении на тиристоре $U_{вкл}$ (напряжении включения) происходит резкое увеличение тока, проходящего через тиристор, и одновременное уменьшение падения напряжения на тиристоре. Второй участок вольт-амперной характеристики тиристора аналогичен прямой ветви вольт-амперной характеристики p - n -перехода (рис. 5.1). В режиме, соответствующем второму участку, напряжение на коллекторном переходе оказывается прямым из-за большого заряда, накопленного в базах.

Таким образом, тиристор может находиться в двух состояниях: *выключенном*, или *закрытом*, которое характеризуется большим падением напряжения на тиристоре и прохождением малых токов

через него, т. е. большим сопротивлением, и *включенном*, или *открытом*, которое характеризуется малым падением напряжения на тиристоре и прохождением больших токов через него, т. е. малым сопротивлением.

В открытом состоянии тиристор находится до тех пор, пока за счет проходящего тока поддерживается количество избыточных зарядов в базах, необходимое для понижения высоты потенциального барьера коллекторного перехода до величины, соответствующей прямому его включению. Если же ток через тиристор уменьшить до некоторой величины ($I_{\text{выкл}}$), то в результате рекомбинации и рассасывания количество избыточных зарядов в базах уменьшится, p - n -переход коллектора окажется включенным в обратном направлении, произойдет перераспределение падений напряжения на p - n -переходах тиристорной структуры, уменьшится инжекция из эмиттера и тиристор перейдет в закрытое состояние (рис. 5.1).

Структуру тиристора можно рассматривать как два эквивалентных транзистора, соединенных между собой (рис. 5.1). Постоянный ток коллектора этих транзисторов можно выразить через эмиттерные токи по аналогии с (4.47):

$$I_{n2} = \bar{\alpha}_1 I_{n1} + \bar{\alpha}_2 I_{n3} + I_{\text{КБО}},$$

где I_{n1} , I_{n2} и I_{n3} — соответственно токи первого, второго и третьего p - n -переходов; $\bar{\alpha}_1$ и $\bar{\alpha}_2$ — коэффициенты передачи постоянных токов эмиттера первого и второго эквивалентных транзисторов; $I_{\text{КБО}}$ — обратный ток коллектора первого и второго транзисторов при токах эмиттера, равных нулю.

Очевидно, что для двухэлектродной структуры

$$I_{n1} = I_{n2} = I_{n3} = I_a.$$

Тогда анодный ток

$$I_a = I_{\text{КБО}} / (1 - \bar{\alpha}), \quad (5.1)$$

где $\bar{\alpha} = \bar{\alpha}_1 + \bar{\alpha}_2$ — суммарный коэффициент передачи тока.

Таким образом, переключение тиристора в открытое состояние с резким увеличением тока будет происходить при условии $\bar{\alpha} = 1$.

Напомним, что коэффициент передачи тока эмиттера транзистора возрастает с увеличением тока эмиттера в результате уменьшения влияния рекомбинации в p - n -переходе эмиттера и появления электрического поля в базе транзистора. Коэффициент передачи тока эмиттера также растет при увеличении напряжения на коллекторе из-за уменьшения толщины базы и увеличения коэффициента умножения в коллекторном переходе. Все эти процессы происходят и в тиристорной структуре при увеличении прямого напряжения.

Для тиристора, находящегося в состоянии, соответствующем переходному участку вольт-амперной характеристики, суммарный коэффициент передачи тока стремится возрасти из-за увеличения проходящего тока. Но суммарный коэффициент передачи тока для переходного участка вольт-амперной характеристики должен быть равен единице. Дальнейшее возрастание суммарного коэффициента

передачи тока предотвращается уменьшением напряжения на коллекторном переходе и, следовательно, на всем тиристоре, т. е. ослаблением ударной ионизации или уменьшением коэффициента умножения носителей заряда M .

В кремниевых транзисторах зависимость коэффициента передачи тока эмиттера выражена сильнее, чем в германиевых. Объясняется это большей шириной запрещенной зоны кремния по сравнению с шириной запрещенной зоны германия. Следовательно, в кремниевом тиристоре также более сильная зависимость суммарного коэффициента передачи тока, чем в германиевом.

Преимуществом кремния как исходного материала для изготовления тиристоров по сравнению с германием является также меньшее значение токов $I_{КБ0}$, что обуславливает меньшую мощность, рассеиваемую в тиристоре в закрытом состоянии и, кроме того, повышенные значения напряжения включения $U_{вкл}$. Поэтому тиристоры делают только из кремния.

В соответствии с (5.1) переключение тиристора в открытое состояние происходит при возрастании суммарного коэффициента передачи тока до единицы. В то же время в каждой из транзисторных структур, составляющих тиристор, коэффициенты передачи тока эмиттера могут быть близки к единице уже при малых напряжениях и токах. Для уменьшения начальных величин коэффициента передачи одну из баз тиристора делают довольно толстой по сравнению с диффузионной длиной соответствующих неосновных носителей. Кроме того, один из эмиттерных переходов шунтируется объемным сопротивлением базы (рис. 5.2). Такое шунтирование обеспечивает малые значения коэффициента передачи тока при малых напряжениях на тиристоре, так как почти весь ток при этих условиях идет по шунтирующему сопротивлению базы, минуя левый $p-n$ -переход эмиттера. При больших напряжениях в связи с уменьшением сопротивления $p-n$ -перехода левого эмиттера почти весь ток будет проходить через $p-n$ -переход эмиттера, минуя шунтирующее сопротивление базы, что связано с инжекцией и накоплением неравновесных носителей заряда в соответствующих базах, или, другими словами, приводит к резкому увеличению суммарного коэффициента передачи тока.

Таким образом, шунтирование эмиттерного перехода тиристора дает возможность получить так называемую жесткую характеристику переключения тиристора, т. е. в связи с более резкой зависимостью коэффициента передачи тока от напряжения и от тока тиристор будет переходить в открытое состояние каждый раз при одном и том же напряжении включения. Наоборот, при слабой зависимости коэффициента передачи тока от напряжения и от тока тиристор может переходить в открытое состояние при несколько отличных напряжениях на тиристоре, т. е. будет иметь так называемую мягкую характеристику переключения.

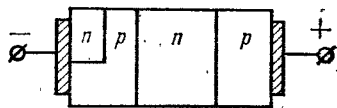


Рис. 5.2. Структура диодного тиристора с шунтированным эмиттерным переходом

Трехэлектродные тиристоры (тринисторы)

Трехэлектродный тиристор (тринистор) — это тиристор, имеющий три внешних вывода.

Чтобы перевести тиристор в открытое состояние, необходимо накопление избыточного отрицательного заряда электронов в n -базе и избыточного положительного заряда дырок в p -базе. Осуществляется это путем увеличения уровня инжекции через эмиттерные переходы при увеличении анодного напряжения на тиристоре до величины напряжения включения. Однако тот же эффект можно получить, если у одной из баз есть невыпрямляющий контакт с электродом (рис. 5.3, а), который называют управляющим электродом. При подаче на управляющий электрод напряжения такой полярности, чтобы прилегающий к этой базе эмиттерный переход был включен в прямом направлении, можно увеличить инжекцию из n -эмиттера, что приведет к накоплению избыточных зарядов в базовых областях тиристора и к переключению его в открытое состояние при анодном напряжении на тиристоре, меньшем напряжения включения $U_{вкл}$ (рис. 5.3, в). Другими словами, при подаче на управляющий электрод напряжения можно изменять ток эмиттера и, следовательно, регулировать значение коэффициента передачи тока эмиттера одной из транзисторных структур, что обусловит увеличение суммарного коэффициента передачи всей тиристорной структуры и переключение ее в открытое состояние.

Очевидно, что с ростом тока управления анодное напряжение, необходимое для переключения тиристора в открытое состояние, понижается (рис. 5.3, в). Для удобства управления тиристором управляющий электрод делают у более тонкой базы, так как коэффициентом передачи тока эмиттера такой транзисторной структуры легче управлять, чем коэффициентом передачи транзистора с толстой базой.

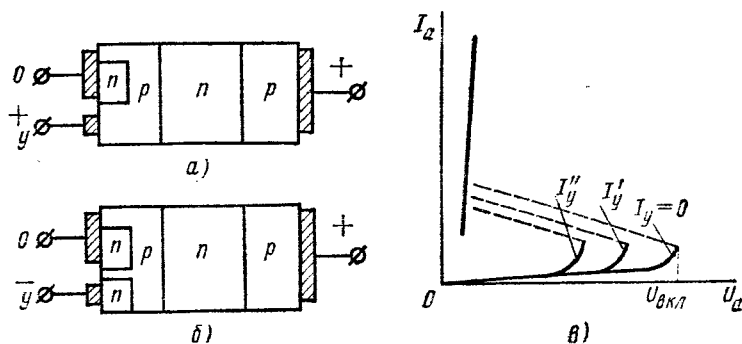


Рис. 5.3. Структуры трехэлектродных тиристоров с невыпрямляющим контактом между управляющим электродом и базой (а), с дополнительным p - n -переходом под управляющим электродом (б) и вольт-амперные характеристики трехэлектродного тиристора при различных токах ($I_y'' > I_y' > 0$) через управляющий электрод (в)

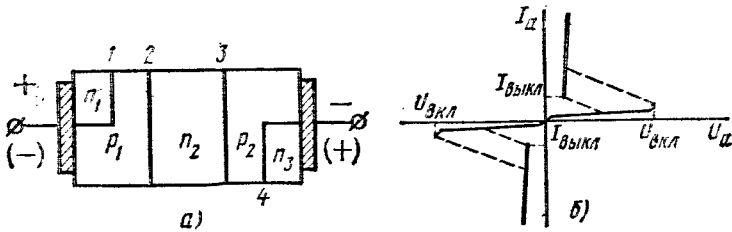


Рис. 5.4. Структура симметричного тиристора (а) и его вольт-амперная характеристика (б)

Управляющий контакт можно сделать не только невыпрямляющим, но и с дополнительным p - n -переходом (рис. 5.3, б). При определенной полярности напряжения на управляющем электроде через дополнительный переход будет происходить инжекция неосновных носителей заряда и их накопление в соответствующей базе, что также приведет к переключению тиристора в открытое состояние.

Симметричные тиристоры

Симметричный тиристор — это тиристор, имеющий практически одинаковые вольт-амперные характеристики при различных полярностях приложенного напряжения.

Основой симметричного тиристора является кремниевая монокристаллическая структура, состоящая из пяти областей с чередующимся типом электропроводности, которые образуют четыре p - n -перехода (рис. 5.4, а). Крайние переходы зашунтированы объемными сопротивлениями прилегающих к этим p - n -переходам областей с электропроводностью p -типа.

Если на такой тиристор подать напряжение плюсом на область n_1 и минусом на область n_3 , то p - n -переход 1 окажется включенным в обратном направлении, и ток, проходящий через него, будет пренебрежимо мал. Весь ток через тиристор при такой полярности приложенного напряжения будет проходить по шунтирующему сопротивлению области p_1 . Четвертый p - n -переход будет включен в прямом направлении и при достаточном значении напряжения через него будет происходить инжекция электронов. Значит, при выбранной полярности внешнего напряжения рабочая часть тиристора представляет собой структуру p - n - p - n , в которой могут происходить те же процессы, что и в обычном двухэлектродном тиристоре, приводящие к переключению его из закрытого состояния в открытое и обратно.

При перемене полярности внешнего напряжения четвертый p - n -переход окажется включенным в обратном направлении и, обладая при этом большим сопротивлением, будет целиком зашунтирован относительно малым сопротивлением области p_2 . Следовательно, при изменении полярности внешнего напряжения рабочая часть тиристора представляет собой структуру n - p - n - p , способную переключаться из закрытого состояния в открытое и обратно.

Таким образом, симметричный тиристор можно представить в виде двух тиристоров, включенных встречно и шунтирующих друг друга.

При изготовлении такой тиристорной структуры с геометрической симметрией относительно центра средней области n_2 и при одинаковых электрических параметрах p - n -переходов 1 и 4, 2 и 3 вольт-амперная характеристика симметричного тиристора также будет одинаковой при разных полярностях приложенного напряжения (рис. 5.4, б).

Симметричный тиристор можно сделать управляемым, если у одной из областей с электропроводностью p -типа осуществить невыпрямляющий контакт с соответствующим управляющим выводом.

§ 5.2. СПОСОБЫ ПЕРЕКЛЮЧЕНИЯ

Включение тиристорov

Включение тиристора путем увеличения анодного напряжения до напряжения включения $U_{вкл}$. Процесс включения тиристора или переключения его из закрытого в открытое состояние при медленном увеличении анодного напряжения на нем был рассмотрен в § 5.1.

Включение тиристора с помощью тока управления. Как отмечалось в § 5.1, увеличение тока, проходящего через один из эмиттерных переходов тиристора, путем подачи соответствующего напряжения на управляющий электрод приводит к накоплению неравновесных носителей заряда в базовых областях тиристора и к переключению его в открытое состояние при анодном напряжении на тиристоре значительно меньшем, чем $U_{вкл}$. При этом процесс накопления неравновесных носителей происходит не мгновенно; поэтому для включения тиристора необходимо, чтобы импульс управляющего тока имел определенную длительность и амплитуду.

Время включения тиристора можно представить в виде двух интервалов времени, соответствующих различным законам изменения тока через тиристор (рис. 5.5).

Время задержки t_z определяется временем диффузии инжектированных из n -эмиттера электронов через p -базу до p - n -перехода коллектора. Ток через коллекторный переход и, следовательно, через тиристор будет возрастать только тогда, когда инжектированные из n -эмиттера электроны дойдут до коллекторного перехода. Для удобства измерения обычно время задержки определяют как время от момента подачи импульса тока управления до момента, когда анодный ток через тиристор увеличится до 0,1 установившегося значения или напряжение на тиристоре снизится до 0,9 начального значения анодного напряжения (см. рис. 5.5).

Время нарастания $t_{нар}$ связано с инерционностью процесса накопления неравновесных носителей заряда в базах тиристора. Во время этого интервала времени происходит резкое увеличение анод-

ного тока через тиристор до условно принятой величины $0,9 I_a$ или резкое снижение анодного напряжения на тиристоре от $0,9$ до $0,1 U_a$. В интервал времени, определяемый временем нарастания, изменения тока, проходящего через тиристор, и падения напряжения на тиристоре соответствуют переходному участку вольт-амперной характеристики, который теоретически должен соответствовать условию равенства единице суммарного коэффициента передачи тиристорной структуры. Практически тиристор обычно включен в цепь, имеющую величину сопротивления меньшую, чем абсолютная величина отрицательного дифференциального сопротивления переходного участка его вольт-амперной характеристики.

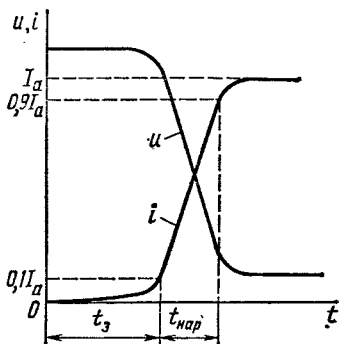


Рис. 5.5. Временные зависимости напряжения на тиристоре и тока через тиристор, характеризующие процесс его включения

Поэтому в период времени нарастания суммарный коэффициент передачи тиристора превышает единицу и все время возрастает из-за быстрого роста тока через тиристор. Таким образом, в этот период времени тиристор находится в неустойчивом режиме, что соответствует активному процессу переключения из закрытого в открытое состояние. Если в это время ток управления прекратится, то тиристор «самостоятельно» перейдет в открытое состояние. Значит, длительность импульса тока управления для перевода тиристора в открытое состояние должна быть больше времени задержки.

Завершением процесса включения тиристора считают момент изменения знака напряжения на коллекторном переходе.

Необходимо отметить, что во время переходных процессов через тиристор проходят большие токи при больших падениях напряжения на нем, что ведет к большим значениям выделяющейся в тиристоре так называемой мощности коммутационных потерь. Средняя мощность коммутационных потерь может привести к недопустимому перегреву тиристора при большой частоте переключения и при большом времени нарастания.

Включение тиристора путем быстрого увеличения анодного напряжения. При быстром нарастании анодного напряжения на тиристоре через него будет проходить емкостный ток, обусловленный наличием барьерных емкостей $p-n$ -переходов.

Рассмотрим влияние барьерной емкости коллекторного перехода. Емкостный ток через коллекторный переход

$$i_c = C_k \frac{du}{dt}, \quad (5.2)$$

где C_k — барьерная емкость коллекторного перехода, значение которой зависит от напряжения на переходе, а следовательно, и от времени, так как напряжение изменяется со временем.

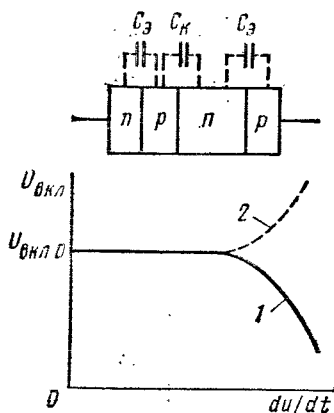


Рис. 5.6. Зависимость напряжения включения тиристора от скорости увеличения подаваемого на него напряжения с учетом только барьерной емкости коллекторного перехода (кривая 1) и только барьерных емкостей эмиттерных переходов (кривая 2)

Чем больше скорость изменения анодного напряжения на тиристоре, тем больше величина емкостного тока через коллекторный переход. Этот ток, проходя через эмиттерные переходы, вызывает увеличение коэффициентов передачи токов эмиттера, что приводит к переключению тиристора при анодном напряжении, меньшем напряжения включения на постоянном токе $U_{вкл0}$ (рис. 5.6).

Барьерные емкости эмиттерных переходов при быстром изменении приложенного напряжения вызывают появление емкостных токов через эти переходы. Емкостные токи не связаны с инжекцией носителей заряда, поэтому с увеличением скорости изменения приложенного к тиристору анодного напряжения переключение тиристора в открытое состояние должно происходить при напряжениях, больших, чем $U_{вкл0}$ (рис. 5.6).

Практически барьерная емкость коллекторного перехода сказывается сильнее, так как она шунтирует большое сопротивление коллекторного перехода, включенного в обратном направлении. Барьерные емкости эмиттерных переходов хотя и имеют большие значения по сравнению с барьерной емкостью коллекторного перехода, но на процесс включения тиристора влияют значительно меньше, так как они шунтируют малые сопротивления эмиттерных переходов, включенных в прямом направлении. Поэтому напряжение включения тиристора в открытое состояние с увеличением скорости изменения анодного напряжения уменьшается.

Выключение тиристорov

Выключение тиристора путем разрыва цепи анодного тока. Тиристор оказывается в выключенном состоянии только после рассасывания неравновесных носителей заряда в базах. Если до окончания процесса выключения вновь подать анодное напряжение на тиристор, он окажется во включенном состоянии. Таким образом, для выключения тиристора необходимо некоторое время, так как процесс рассасывания происходит не мгновенно.

При выключении тиристора путем разрыва цепи рассасывание неравновесных носителей заряда происходит только в результате рекомбинации. Поэтому время выключения тиристора в этом случае может быть довольно большим и зависит от времени жизни носителей заряда.

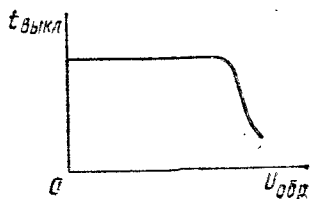


Рис. 5.7. Зависимость времени выключения тиристора от величины обратного напряжения

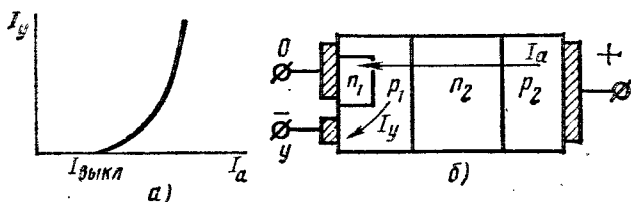


Рис. 5.8. Зависимость тока управления, необходимого для выключения тиристора, от величины анодного тока (а) и структура трехэлектродного тиристора (б)

Выключение тиристора путем изменения полярности анодного напряжения. Стремление ускорить процесс рассасывания неравновесных носителей заряда в базах путем изменения полярности приложенного к тиристорному анодному напряжению не всегда приводит к желаемому результату.

Для ускорения процесса рассасывания неравновесных носителей заряда в базах необходимо понизить потенциальный барьер коллекторного перехода. Однако коллекторный переход при прохождении тока через открытый тиристор до переключения уже был включен в прямом направлении из-за накопленных неравновесных носителей заряда в базах и, следовательно, имел низкое сопротивление. Поэтому на долю коллекторного перехода при переключении тиристора на обратное напряжение приходится малая часть всего внешнего напряжения и, значит, практически не происходит понижения высоты потенциального барьера этого перехода. Кроме того, наличие барьерной емкости также препятствует быстрому изменению напряжения на p - n -переходе.

Чтобы увеличить долю внешнего напряжения, приходящуюся на коллекторный переход, эмиттерные переходы при переключении необходимо довести до пробоя. Таким образом, время выключения тиристора уменьшается только при больших значениях обратных напряжений на аноде (рис. 5.7).

Выключение тиристора путем подачи напряжения на управляющий электрод или с помощью тока управления. Для выключения тиристора необходимо отвести неравновесные носители заряда из базы, у которой имеется управляющий электрод. В то же время анодный ток, проходящий через еще открытый тиристор, непрерывно восполняет количество неравновесных носителей заряда в базах тиристора. Таким образом, значение тока управления, необходимого

для выключения тиристора, зависит от значения анодного тока через тиристор (рис. 5.8, а). Кроме того, некоторые тиристоры (с большой площадью p - n -переходов) невозможно выключить с помощью тока управления при больших анодных токах. Объясняется это тем, что при движении носителей к управляющему электроду (дырок в тиристоре, структура которого показана на рис. 5.8, б) база тиристора под эмиттерным переходом становится неэквипотенциальной, и дальние от управляющего электрода части эмиттерного перехода остаются включенными в прямом направлении. Инжекция электронов из этих частей эмиттерного перехода поддерживает соответствующую часть тиристорной структуры в открытом состоянии. Стремление же еще больше увеличить ток управления путем увеличения напряжения на управляющем электроде относительно эмиттера может привести к пробое эмиттерного перехода, так как он имеет обычно малое пробивное напряжение из-за большой концентрации примесей с обеих сторон этого перехода.

§ 5.3. КОНСТРУКЦИЯ И ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ

Многослойные тиристорные структуры чаще всего создают в монокристаллических пластинках кремния путем последовательной диффузии различных примесей. Применяют также метод сплавления примесей. Одним из перспективных методов изготовления тиристорных структур является метод эпитаксиального наращивания.

Рассмотрим кратко метод изготовления и конструкцию мощного тиристора. В отшлифованную и протравленную пластинку кремния с электропроводностью n -типа толщиной около 0,25 мм проводится с двух сторон диффузия бора, галлия или алюминия при температуре 1300 °С, в течение 8—10 ч. В результате происходит диффузия акцепторной примеси в пластинку кремния на глубину 110—120 мкм и образуется p - n - p -структура с двумя p - n -переходами.

Следующую операцию — создание n -области эмиттера в одном из диффузионных слоев p -типа можно осуществить либо методом диффузии при защите противоположной стороны пластинки крем-

ния слоем двуокиси кремния, либо сплавлением донорной примеси (например, золота с сурьмой, где сурьма и является донором).

Для получения большей зависимости суммарного коэффициента передачи от тока эмиттера производят шунтирование эмиттерного перехода. При создании n -области эмиттера методом диффузии примеси с использованием планарной технологии можно довольно просто осуществить распределенное шунтирование, если n -область сделать не сплошной (рис. 5.9),

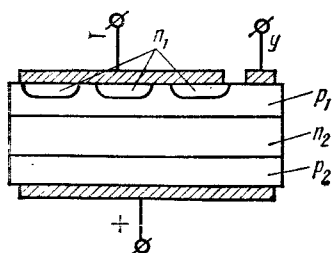


Рис. 5.9. Структура трехэлектродного тиристора с распределенным шунтированием эмиттерного перехода

т. е. провести диффузию доноров не по всей площади кремниевого диска или пластинки.

Остальные технологические операции (создание невыпрямляющих контактов, в том числе и управляющего, присоединение электродов и монтаж в корпусе) принципиально ничем не отличаются от аналогичных операций при изготовлении диодов и транзисторов. Управляющий контакт создают в месте выхода базовой области на верхнюю поверхность пластинки (рис. 5.9) либо в центре пластинки, либо по периферии, либо распределенным по площади.

Кремниевую пластинку с тиристорной структурой припаивают к массивному медному основанию корпуса прибора с использованием термокомпенсирующей вольфрамовой прокладки для уменьшения механических напряжений, которые могут возникать при работе тиристора в связи с различными температурными коэффициентами линейного расширения меди и кремния (рис. 5.10). Баллон корпуса со стеклянным проходным изолятором с помощью завальцованного шва герметически изолирует кремниевую пластинку с тиристорной структурой от воздействия окружающей среды. Через стеклянный изолятор проходят выводы управляющего электрода и одного из силовых электродов. Вторым силовым выводом является основание корпуса прибора.

Для повышения эффективности охлаждения кремниевой пластинки с тиристорной структурой применяют конструкции тиристорных таблеточного типа. Такие конструкции изготавливают с металлическими корпусами и стеклянными проходными изоляторами или с керамическими корпусами и изоляторами из заполимеризованных эпоксидных компаундов, или с металлокерамическими корпусами. Улучшенный теплоотвод в тиристорах

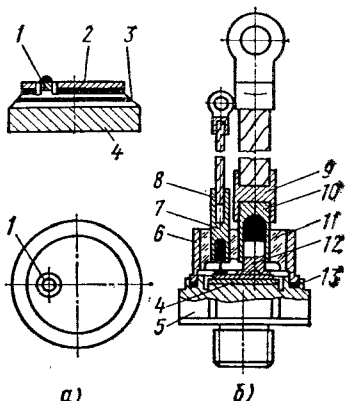


Рис. 5.10. Конструкция мощного тиристора:

a — кремниевый диск с электродами; *b* — разрез тиристора; 1 — управляющий электрод; 2 — вольфрамовая прокладка силового вывода; 3 — кремниевый диск с тиристорной структурой; 4 — термокомпенсирующая вольфрамовая прокладка; 5 — основание корпуса; 6 — крышка корпуса; 7 — переходный стакан вывода управляющего электрода; 8 — вывод управляющего электрода; 9 — переходная втулка силового вывода; 10 — переходный стакан силового вывода; 11 — стеклянный проходной изолятор; 12 — силовой электрод; 13 — фторопластовая прокладка

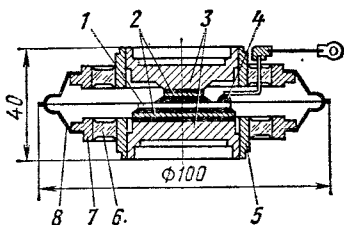


Рис. 5.11. Конструкция тиристора таблеточного типа, рассчитанного на допустимый анодный ток 300 А:

1 — кремниевый диск с тиристорной структурой; 2 — термокомпенсирующая вольфрамовая прокладка; 3 — контактные втулки; 4 — управляющий электрод; 5 — металлические электродные кольца; 6 — проходные стеклянные изоляторы; 7 — внешние металлические кольца; 8 — металлические пружинящие гофрированные шайбы, сваренные между собой после сборки тиристора

таблеточной конструкции достигается благодаря использованию двух радиаторов, присоединенных с обеих сторон корпуса тиристора (рис. 5.11).

§ 5.4. ПАРАМЕТРЫ И ХАРАКТЕРИСТИКИ

Сущность основных параметров тиристорov рассмотрена в предыдущих параграфах, поэтому приведем численные значения и температурные зависимости этих параметров.

Напряжение включения $U_{\text{вкл}}$ — это прямое анодное напряжение, при котором тиристор переходит из закрытого в открытое состояние при разомкнутом управляющем выводе. Для разных типов тиристорov напряжение включения различно — от 10 до 2500 В.

Основным фактором, влияющим на параметры тиристорov, в том числе и на напряжение включения, является изменение времени жизни носителей заряда при изменении температуры. С повышением температуры в рабочем диапазоне температур время жизни носителей возрастает, что приводит, как и в обычных транзисторах, к росту коэффициентов передачи токов. Поэтому напряжение включения с повышением температуры окружающей среды уменьшается.

У некоторых тиристорov первоначальным процессом, приводящим к увеличению суммарного коэффициента передачи тиристорной структуры и к переключению тиристора в открытое состояние, является лавинное умножение носителей заряда в сильном электрическом поле коллекторного перехода. Напряжение включения таких тиристорov увеличивается при увеличении температуры, так как при этом уменьшается длина свободного пробега носителей заряда.

Ток включения $I_{\text{вкл}}$ — это такое значение прямого анодного тока через тиристор, выше которого тиристор переключается в открытое состояние при разомкнутой цепи управляющего вывода.

Ток включения уменьшается с повышением температуры в связи с увеличением времени жизни носителей заряда и с увеличением коэффициентов передачи токов.

Отпирающий ток управления $I_{\text{у.вкл}}$ — наименьший ток в цепи управляющего вывода, который обеспечивает переключение тиристора в открытое состояние при данном напряжении на тиристоре. Отпирающий ток управления протекает в цепи управляющего вывода при некотором напряжении управления $U_{\text{у.вкл}}$, которое необходимо приложить между управляющим выводом и выводом от эмиттерной области, прилегающей к соответствующей базе.

С повышением температуры тиристора из-за увеличения времени жизни носителей и из-за соответствующего роста коэффициентов передачи токов отпирающий ток управления, а значит, и напряжение управления уменьшаются.

Время задержки t_z — время, в течение которого анодный ток через тиристор возрастает до величины 0,1 установленного значения с момента подачи на тиристор управляющего импульса, или время, в течение которого анодное напряжение на тиристоре умень-

шается до 0,9 начального значения с момента подачи на тиристор управляющего импульса. Оба определения равноценны при чисто активной нагрузке во внешней цепи тиристора.

Время задержки и время включения зависят от управляющего тока, уменьшаясь с увеличением амплитуды управляющего импульса тока. Связано это с возникновением больших градиентов концентрации носителей в тиристорной структуре и, следовательно, с возникновением больших диффузионных токов, которые обеспечивают более быстрое накопление неравновесных носителей в базовых областях тиристора, т. е. обеспечивают включение тиристора.

Время включения $t_{вкл}$ — время, в течение которого ток через тиристор возрастает до 0,9 установившегося значения с момента подачи на тиристор управляющего импульса. Если внешняя цепь представляет собой чисто активную нагрузку для тиристора, то во время включения тиристора с увеличением тока соответственно уменьшается напряжение на тиристоре. Поэтому время включения в этом случае принято определять как время, в течение которого напряжение на тиристоре уменьшается до 0,1 начального значения с момента подачи на тиристор управляющего импульса.

Максимально допустимая скорость нарастания прямого напряжения $(du/dt)_{max}$ — скорость нарастания прямого анодного напряжения, при которой еще не происходит переключения тиристора в открытое состояние при отключенном управляющем выводе. Кроме максимально допустимой скорости нарастания прямого напряжения используется иногда другой параметр тиристора — критическая скорость нарастания прямого анодного напряжения, т. е. такая минимальная скорость нарастания прямого анодного напряжения, при которой происходит переключение тиристора в открытое состояние. Критическое напряжение, при котором происходит самовключение тиристора, при больших величинах du/dt называют *н а п р я ж е н и е м с а м о в к л ю ч е н и я*. Оно меньше напряжения включения. Поэтому во всех случаях, чтобы исключить ложное переключение тиристора, амплитуда подаваемого напряжения и скорость его нарастания должны быть меньше допустимых значений напряжения самовключения и $(du/dt)_{max}$.

Максимально допустимая скорость нарастания прямого тока $(di/dt)_{max}$ — скорость нарастания прямого анодного тока через тиристор, не вызывающая необратимых процессов в тиристорной структуре и связанного с ними ухудшения электрических параметров тиристора. При любом способе включения тиристора процесс переключения начинает развиваться в ограниченном объеме, а затем распространяется по всей структуре. Таким образом, происходит неравномерное распределение тока по площади *p-n*-переходов тиристора. В тиристорах с большой площадью *p-n*-переходов, т. е. в мощных тиристорах, высокую проводимость весь объем тиристорной структуры может приобрести иногда только через несколько десятков микросекунд с момента подачи управляющего импульса.

При высоких скоростях нарастания прямого тока вблизи управляющего электрода (при включении тиристора с помощью тока уп-

равления) плотность прямого тока достигает больших значений и может произойти локальный разогрев структуры до температуры плавления кремния с последующим разрушением кремниевой пластинки или с последующим выходом тиристора из строя.

Бороться с этим явлением можно двумя способами: 1) управляющий электрод делают распределенным по всей площади тиристорной структуры; 2) ограничивают скорость нарастания прямого тока путем соответствующих схемных решений, например включая последовательно с тиристором небольшую индуктивность.

Максимально допустимый ток в открытом состоянии $I_{пр\ max}$ — максимальное значение тока в открытом состоянии, при котором обеспечивается заданная надежность тиристора. Для тиристоров разных типов максимально допустимый ток в открытом состоянии от 40 мА до 1000 А. Максимально допустимый ток определяется из условия, что температура любой части тиристорной структуры при работе тиристора не должна превышать допустимого значения (обычно 140 °С). Нагрев тиристора во время работы происходит из-за электрических потерь при прохождении прямого тока, токов утечки в закрытом тиристоре, тока в цепи управляющего вывода и от коммутационных потерь.

Остаточное напряжение $U_{пр}$ — значение напряжения на тиристоре, находящемся в открытом состоянии, при прохождении через него максимально допустимого тока. Остаточное напряжение обычно не превышает 2 В.

Ток выключения $I_{выкл}$ — значение прямого тока через тиристор при разомкнутой цепи управления, ниже которого тиристор выключается. В зависимости от типа тиристора ток выключения может быть от нескольких миллиампер до нескольких десятых долей ампера.

С повышением температуры величина тока выключения уменьшается в связи с ростом времени жизни носителей и увеличением коэффициентов передачи токов при тех же величинах проходящего через тиристор тока.

Время выключения $t_{выкл}$ — время от момента перемены тока, проходящего через тиристор, с прямого на обратный до момента, когда тиристор полностью восстановит запирающую способность в прямом направлении. Процесс выключения тиристора связан с исчезновением избыточных зарядов неравновесных носителей в базовых областях вследствие рекомбинации и ухода неравновесных носителей через *p-n*-переходы. Для ускорения процесса выключения в большинстве случаев к тиристору прикладывают обратное анодное напряжение. На рис. 5.12 изображена кривая тока через тиристор для случая, когда в момент времени t_0 к тиристору, находящемуся в открытом состоянии, приложено обратное напряжение.

В первый момент после переключения анодного напряжения на обратное ток через тиристор определяется сопротивлением нагрузки и равен $I'_{обр} \approx U_{обр}/R_H$ (при чисто активном сопротивлении нагрузки). Сопротивление тиристора в это время мало, так как велики концентрации неосновных носителей заряда около *p-n*-переходов. К моменту t_1 сильнолегированная *p*-база теряет избыточный

заряд неравновесных носителей, и напряжение u_{a1} на p - n -переходе 1 падает до нуля, а затем изменяет знак. К моменту времени t_2 ток через тиристор уменьшается до величины $I''_{обр} = \frac{U_{обр} - U_{проб\ э1}}{R_H}$. С двух сторон эмиттерного p - n -перехода 1 находятся низкоомные (из-за большой концентрации примесей) области кристалла кремния. Поэтому пробивное напряжение этого p - n -перехода мало. Следовательно, обратное напряжение на этом p - n -переходе достигает пробивного и затем стабилизируется (в момент времени t_2).

При дальнейшем рассасывании напряжения на эмиттерном p - n -переходе 3 достигает нуля. Ток через тиристор с этого момента опять быстро уменьшается до величины, соответствующей установившемуся обратному току $I_{обр}$.

Значит, в момент времени t_4 тиристор восстановил запирающую способность в обратном направлении. Однако в момент времени t_4 в базовых областях около p - n -перехода 2 сохраняется еще значительная избыточная концентрация носителей заряда, что обеспечит переключение тиристора в открытое состояние при подаче в этот момент на тиристор даже небольшого прямого напряжения, т. е. тиристор еще не восстановил запирающей способности для прямого напряжения. Таким образом, время выключения превышает время восстановления запирающей способности тиристора для обратного напряжения. Время выключения тиристоров может достигать до нескольких десятков микросекунд. Оно зависит от геометрии тиристора и от времени жизни неосновных носителей заряда в базовых областях (в основном от времени жизни неосновных носителей в толстой высокоомной области базы).

Время выключения тиристоров увеличивается с повышением температуры, что также связано с ростом времени жизни носителей заряда, с замедлением процесса рекомбинации неравновесных носителей в базовых областях тиристора.

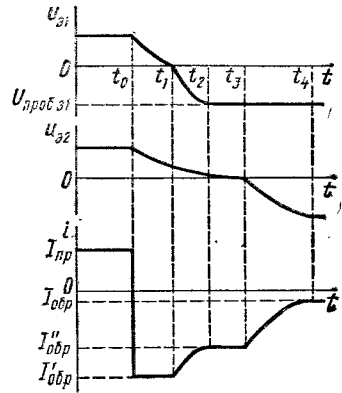


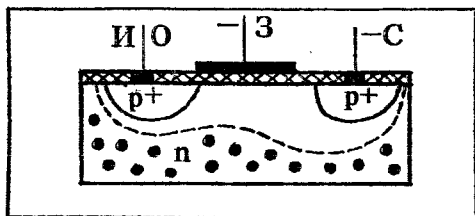
Рис. 5.12. Временные зависимости напряжений на первом и втором эмиттерных переходах и тока через тиристор во время переключения тиристора из открытого в закрытое состояние

Контрольные вопросы

1. Почему коллекторный переход тиристора оказывается смещенным в прямом направлении при переходе тиристора в открытое состояние?
2. Какие методы используют для уменьшения коэффициентов передачи тока транзисторных структур, составляющих тиристор?
3. Какими способами можно перевести тиристор из закрытого в открытое состояние?
4. Почему не всегда и не все тиристоры можно перевести из открытого в закрытое состояние с помощью тока управляющего электрода?

ПОЛЕВЫЕ ТРАНЗИСТОРЫ И ПРИБОРЫ С ЗАРЯДОВОЙ СВЯЗЬЮ

6



Полевой транзистор — это электропреобразовательный прибор, в котором ток через канал управляется электрическим полем, возникающим при приложении напряжения между затвором и истоком, и который предназначен для усиления мощности электромагнитных колебаний.

§ 6.1. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С УПРАВЛЯЮЩИМ *p-n*-ПЕРЕХОДОМ

Структура и принцип действия

Полевой транзистор с управляющим электронно-дырочным переходом имеет два невыпрямляющих контакта к области, по которой проходит управляемый ток основных носителей заряда, и один или два управляющих электронно-дырочных перехода, смещенных в обратном направлении (рис. 6.1). При изменении обратного напряжения на *p-n*-переходе изменяется его толщина и, следовательно, толщина области, по которой проходит управляемый ток основных носителей заряда. Область, толщина и поперечное сечение которой управляется внешним напряжением на управляющем *p-n*-переходе и по которой проходит управляемый ток основных носителей, называют каналом. Электрод, из которого в канал входят основные носители заряда, называют истоком. Электрод, через который из канала уходят основные носители заряда, называют стоком. Электрод, служащий для регулирования поперечного сечения канала, называют затвором.

Полевой транзистор с управляющим *p-n*-переходом — это полевой транзистор, затвор которого отделен в электрическом отношении от канала *p-n*-переходом, смещенным в обратном направлении.

Электропроводность канала может быть как *n*-, так и *p*-типа. Поэтому по электропроводности канала различают полевые транзисторы с *n*-каналом и *p*-каналом. Все полярности напряжений смещения, подаваемых на электроды транзисторов с *n*- и *p*-каналом, противоположны.

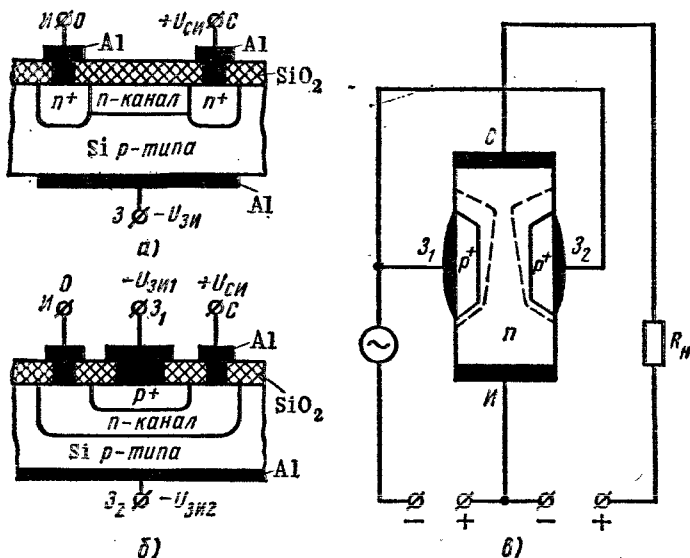


Рис. 6.1. Структуры полевых транзисторов с одним управляющим p - n -переходом (а), с двумя управляющими p - n -переходами (б и в) и принципиальная схема включения с общим истоком (с)

Управление током стока, т. е. током от внешнего относительно мощного источника питания в цепи нагрузки, происходит при изменении обратного напряжения на p - n -переходе затвора (или на двух p - n -переходах одновременно). В связи с малостью обратных токов мощность, необходимая для управления током стока и потребляемая от источника сигнала в цепи затвора, оказывается ничтожно малой. Поэтому полевой транзистор может обеспечить усиление электромагнитных колебаний как по мощности, так и по току и напряжению.

Таким образом, полевой транзистор по принципу действия аналогичен вакуумному триоду. Исток в полевом транзисторе подобен катоду вакуумного триода, затвор — сетке, сток — аноду. Но при этом полевой транзистор существенно отличается от вакуумного триода. Во-первых, для работы полевого транзистора не требуется подогрева катода. Во-вторых, любую из функций истока и стока может выполнять каждый из этих электродов. В-третьих, полевые транзисторы могут быть сделаны как с n -каналом, так и с p -каналом, что позволяет удачно сочетать эти два типа полевых транзисторов в схемах.

От биполярного транзистора и в том числе от однопереходного транзистора полевой транзистор отличается, во-первых, принципом действия: в биполярном транзисторе управление выходным сигналом производится входным током, а в полевом транзисторе — входным напряжением или электрическим полем. Во-вторых, полевые транзисторы имеют значительно большие входные сопротивления, что

связано с обратным смещением p - n -перехода затвора в рассматриваемом типе полевых транзисторов. В-третьих, полевые транзисторы могут обладать низким уровнем шума (особенно на низких частотах), так как в полевых транзисторах не используется явление инжекции неосновных носителей заряда и канал полевого транзистора может быть отделен от поверхности полупроводникового кристалла (рис. 6.1, б и в). Процессы рекомбинации носителей в p - n -переходе и в базе биполярного транзистора, а также генерационно-рекомбинационные процессы на поверхности кристалла полупроводника сопровождаются возникновением низкочастотных шумов.

Статические характеристики

Полевые транзисторы имеют большие входные и выходные сопротивления. Поэтому в отличие от биполярных транзисторов статические характеристики полевых транзисторов удобнее и правильнее исследовать с помощью генераторов напряжения в качестве источников питания. Следовательно, статические характеристики полевых транзисторов должны соответствовать системе уравнений

$$\left. \begin{aligned} I_1 &= f_1(U_1; U_2); \\ I_2 &= f_2(U_1; U_2). \end{aligned} \right\} \quad (6.1)$$

Важнейшими семействами статических характеристик для полевого транзистора являются семейство выходных статических характеристик и семейство характеристик передачи.

Выходные статические характеристики полевого транзистора представляют собой зависимости тока стока от напряжения на стоке относительно истока при различных постоянных напряжениях на затворе (рис. 6.2, а). Рассмотрим вначале характер одной зависимости $I_c = f(U_{си})$ при $U_{зи} = 0$. Напряжение на затворе относительно истока будет равно нулю только в том случае, если затвор закорочен с истоком. Характеристика выходит из начала координат под углом, соответствующим начальному статическому сопротивлению канала и сопротивлениям $R_{и}$ и R_c , прилегающих к каналу областей полупроводникового кристалла с тем же типом электропроводности. Статическое сопротивление канала определяется его длиной и поперечным сечением, зависящим от толщины p - n -перехода (или p - n -переходов).

Первая часть характеристики, которую называют *крутой частью*, сублинейна, т. е. ток стока растет замедленно с напряжением на стоке. Объясняется эта нелинейность характеристики увеличением толщины p - n -перехода затвора около стока, так как с увеличением напряжения на стоке растет по абсолютному значению обратное напряжение на p - n -переходе затвора. Ток стока, проходя по каналу, создает его неэквипотенциальность. Таким образом, наибольшая толщина p - n -перехода и соответственно наименьшее поперечное сечение канала получается со стороны стока (см. рис. 6.1, в и рис. 6.3).

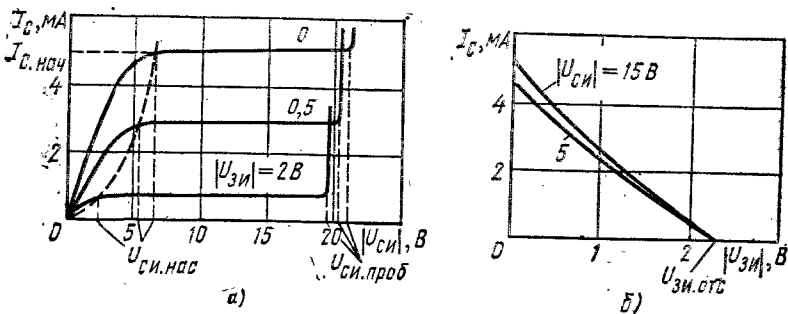
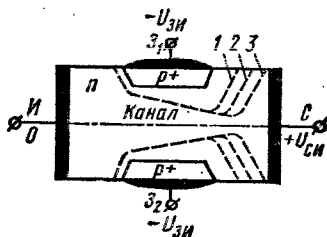


Рис. 6.2. Выходные статические характеристики (а) и статические характеристики передачи (б) полевого транзистора с управляющим p - n -переходом

Рис. 6.3. Увеличение толщины управляющего p - n -перехода со стороны стока и соответствующее увеличение длины перекрытой части канала при увеличении напряжения на стоке и неизменном напряжении на затворе относительно истока:

1 — напряженье на стоке равно напряжению насыщения $U_{СИ,нас}$; 2 — $U_{СИ} > U_{СИ,нас}$; 3 — $U_{СИ} > U_{СИ}$



Другой физической причиной, приводящей к сублинейности выходной характеристики, является уменьшение подвижности носителей заряда в канале при увеличении в нем напряженности электрического поля (см. § 1.10).

При некотором напряжении на стоке $U_{СИ,нас}$ — напряжении насыщения — происходит перекрытие канала из-за увеличения толщины p - n -перехода затвора. Ток стока при дальнейшем увеличении напряжения на стоке почти не растет.

При напряжении между затвором и истоком, равном нулю, и при напряжении на стоке, равном или превышающем напряжение насыщения, ток стока называют начальным током стока $I_{С,нач}$. Часть характеристики, соответствующую насыщению тока стока, называют пологой частью. Следует учитывать условность понятия «перекрытие» канала при увеличении напряжения на стоке и неизменном напряжении на затворе относительно истока, так как перекрытие канала при указанных условиях является следствием увеличения тока стока. Таким образом, можно считать, что в результате увеличения тока стока или напряжения на стоке автоматически устанавливается некоторое малое сечение канала со стороны стокового электрода.

При дальнейшем увеличении напряжения на стоке увеличивается длина перекрытой части канала и растет статическое сопротивление канала. Если бы длина перекрытой части канала увеличивалась пропорционально напряжению на стоке, то ток стока не

изменялся бы при напряжениях на стоке, превышающих напряжение насыщения. Однако длина перекрытой части канала увеличивается из-за увеличения толщины p - n -перехода с ростом напряжения на стоке (см. рис. 6.3), а толщина p - n -перехода пропорциональна либо корню квадратному, либо корню кубическому из напряжения (см. § 2.6 и 2.7). Поэтому в пологой части характеристики наблюдается некоторое увеличение тока стока при увеличении напряжения на стоке.

Теперь рассмотрим смещение и изменение статических характеристик с изменением напряжения на затворе. При подаче на затвор напряжения такой полярности относительно истока, которая соответствует обратному смещению p - n -перехода затвора, и при увеличении этого напряжения по абсолютному значению уменьшается начальное поперечное сечение канала. Поэтому начальные участки выходных статических характеристик при напряжениях на затворе, отличных от нуля, имеют другой наклон, соответствующий большему начальным статическим сопротивлениям канала.

При меньших начальных поперечных сечениях перекрытие канала из-за увеличения напряжения на стоке происходит при меньших напряжениях насыщения (см. рис. 6.2, а).

При больших напряжениях на стоке может возникнуть пробой p - n -перехода затвора. Обратное напряжение на p - n -переходе затвора изменяется вдоль длины канала, достигая максимального значения у стокового конца канала. Напряжение, приложенное к p - n -переходу затвора в этом месте, является суммой напряжений на стоке и на затворе. Таким образом, пробой полевого транзистора будет происходить при разных напряжениях на стоке в зависимости от напряжения на затворе. Чем больше напряжение на затворе, тем меньше напряжение на стоке, при котором произойдет пробой p - n -перехода затвора (см. рис. 6.2, а). Полевые транзисторы делают обычно на основе кремния. Поэтому пробой таких транзисторов имеет лавинный характер.

Характеристики передачи полевого транзистора в соответствии с (б.1) представляют собой зависимости тока стока от напряжения на затворе при различных постоянных напряжениях на стоке. Так как основным рабочим режимом полевых транзисторов является режим насыщения тока стока, что соответствует пологим частям выходных статических характеристик, то наибольший интерес представляет зависимость тока насыщения от напряжения на затворе при постоянном напряжении на стоке. Характер этой зависимости ясен из принципа действия полевого транзистора с управляющим p - n -переходом. При изменении напряжения на стоке смещением характеристик передачи практически можно пренебречь в связи с малым изменением тока стока в пологой части выходных статических характеристик (рис. 6.2, б).

Напряжение между затвором и истоком полевого транзистора с управляющим p - n -переходом, при котором ток стока достигает заданного низкого значения, называют напряжением отсечки полевого транзистора $U_{зи.отс}$.

При рассмотрении статических характеристик полевого транзистора были отмечены его основные статические параметры. По статической характеристике передачи можно определить еще один основной параметр полевого транзистора, характеризующий его усилительные свойства — *крутизну характеристики* полевого транзистора S , которая представляет собой отношение изменения тока стока к изменению напряжения на затворе при коротком замыкании по переменному току на выходе транзистора в схеме с общим истоком:

$$S = \left. \frac{dI_C}{dU_{3И}} \right|_{U_{СИ} = \text{const}}$$

Крутизна характеристики полевого транзистора составляет обычно несколько миллиампер на вольт.

§ 6.2. РАСЧЕТ ВЫХОДНЫХ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЕВОГО ТРАНЗИСТОРА С УПРАВЛЯЮЩИМ p - n -ПЕРЕХОДОМ

Пренебрегая объемными сопротивлениями кристалла полупроводника на участках между концами канала и контактами истока и стока, рабочую часть полевого транзистора можно представить в упрощенном виде (рис. 6.4).

Плотность тока в канале

$$J = \gamma E = -\gamma \frac{d\varphi}{dx}, \quad (6.2)$$

где γ — удельная проводимость канала.

В первом приближении будем считать удельную проводимость материала канала не зависящей от напряженности электрического поля, т. е. не будем учитывать изменение подвижности. Плотность тока в канале изменится по его длине в связи с изменением сечения канала и соответствующим изменением напряженности.

Ток в канале транзистора, неизменный по всему каналу,

$$I_C = -\gamma b w \frac{d\varphi}{dx}, \quad (6.3)$$

где b — ширина канала.

Толщина канала w зависит от толщины p - n -переходов:

$$w = a - 2\delta. \quad (6.4)$$

Толщина p - n -перехода зависит от напряжения на нем (см. § 2.6). В выражении (2.27) для толщины резкого p - n -перехода можно пренебречь контактной разностью потенциала на p - n -переходе по сравнению с относительно большим обратным напряжением, поданным на затвор полевого транзистора. Однако необходимо учесть неэквипотенциальность канала, возникающую из-за прохождения тока по каналу от истока к стоку. Тогда

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0}{qN}(\varphi - U_{3И})}. \quad (6.5)$$

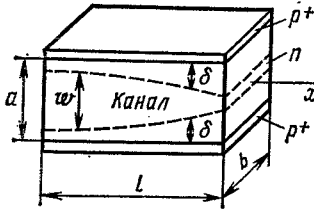


Рис. 6.4. Структура рабочей части полевого транзистора с управляющими p - n -переходами

Для более компактной записи найдем напряжение отсечки $U_{\text{зи.отс}}$. Из определения напряжения отсечки ясно, что толщина канала около стока при этом напряжении равна нулю, а толщина p - n -перехода

$$\delta = \frac{a}{2} = \sqrt{\frac{2\epsilon\epsilon_0}{qN}} U_{\text{зи.отс}}^{1/2}. \quad (6.6)$$

Используя (6.4), (6.5) и (6.6), получим

$$w = a \left(1 - \sqrt{\frac{\varphi + U_{\text{зи.отс}}}{U_{\text{зи.отс}}}} \right). \quad (6.7)$$

После подстановки (6.7) в (6.3) абсолютное значение тока стока

$$|I_C| = \gamma b a \left(1 - \sqrt{\frac{\varphi + U_{\text{зи.отс}}}{U_{\text{зи.отс}}}} \right) \frac{d\varphi}{dx}. \quad (6.8)$$

Решая это дифференциальное уравнение с разделяющимися переменными при граничных условиях:

$$1) \ x = 0, \ \varphi = 0 \quad \text{и} \quad 2) \ x = l, \ \varphi = U_{\text{си}},$$

получим искомую функцию

$$I_C = \frac{1}{R_{\text{си.отк}}} \left[U_{\text{си}} + \frac{2}{3} \cdot \frac{U_{\text{зи}}^{3/2} - (U_{\text{си}} + U_{\text{зи}})^{3/2}}{U_{\text{зи.отс}}^{1/2}} \right], \quad (6.9)$$

где $R_{\text{си.отк}} = l/(\gamma ab)$ — сопротивление сток — исток в открытом состоянии, т. е. при $U_{\text{зи}} = 0$ и при малом напряжении на стоке (меньшем напряжения насыщения).

Выражение (6.9) дает возможность найти ток насыщения полевого транзистора. Как отмечалось ранее, перекрытие канала происходит при напряжении отсечки $U_{\text{зи.отс}}$. Режим насыщения наступит при условии

$$U_{\text{зи}} + U_{\text{си}} = U_{\text{зи.отс}},$$

т. е. при напряжении на стоке

$$U_{\text{си.нас}} = U_{\text{зи.отс}} - U_{\text{зи}}. \quad (6.10)$$

Если в выражении (6.9) заменить напряжение на затворе $U_{\text{зи}}$ напряжением на стоке, при котором происходит насыщение, и напряжением отсечки $U_{\text{зи.отс}}$, пользуясь соотношением (6.10), то получим связь между током и напряжением насыщения:

$$I_{\text{с.нас}} = \frac{1}{R_{\text{си.отк}}} \left[U_{\text{си.нас}} + \frac{2}{3} (U_{\text{зи.отс}} - U_{\text{си.нас}}) \times \right. \\ \left. \times \sqrt{\frac{U_{\text{зи.отс}} - U_{\text{си.нас}}}{U_{\text{зи.отс}}}} - \frac{2}{3} U_{\text{зи.отс}} \right]. \quad (6.11)$$

На рис. 6.2 штриховая кривая показывает зависимость $I_{\text{с.нас}} = f(U_{\text{си.нас}})$. Эта зависимость является геометрическим местом

точек, соответствующих точкам и напряжениям, при которых наступает режим насыщения полевого транзистора.

Важно знать также зависимость тока насыщения от напряжения на затворе полевого транзистора, т. е. характеристику передачи. Эту зависимость можно найти, если в выражение (6.9) подставить из условия насыщения (6.10) напряжение на стоке $U_{СИ.нас}$:

$$I_{С. нас} = \frac{1}{R_{СИ. отк}} \left[\frac{1}{3} U_{ЗИ. отс} - U_{ЗИ} \left(1 - \frac{2}{3} \sqrt{\frac{U_{ЗИ}}{U_{ЗИ. отс}}} \right) \right]. \quad (6.12)$$

Поскольку пологий участок выходных статических характеристик полевых транзисторов является, как и в вакуумных пентодах, основным рабочим участком, определим крутизну характеристики S именно в этой области. Дифференцируя (6.12) по $U_{ЗИ}$, получим

$$S = \left| \frac{dI_{С. нас}}{dU_{ЗИ}} \right| = \frac{1}{R_{СИ. отк}} \left(1 - \sqrt{\frac{U_{ЗИ}}{U_{ЗИ. отс}}} \right). \quad (6.13)$$

Из выражения (6.13) следует, что для получения больших значений этого параметра необходимо иметь меньшую величину сопротивления сток—исток в открытом состоянии транзистора $R_{СИ.отк}$ или большую удельную проводимость исходного материала. В то же время концентрация примесей и соответственно носителей заряда в канале должна быть небольшой, чтобы при увеличении напряжения на p - n -переходе он расширялся в сторону канала. Таким образом, для получения больших значений крутизны характеристики желательно при изготовлении полевого транзистора выбирать материал с большей подвижностью носителей заряда.

Значение крутизны характеристики полевого транзистора в значительной мере определяется отношением ширины канала b к его длине l . Увеличение отношения b/l позволяет повысить крутизну характеристики S и ток насыщения полевого транзистора [см. (6.11) и (6.13)].

Согласно выражению (6.13) для повышения крутизны характеристики S необходимо увеличивать толщину канала a . Однако с увеличением толщины канала недопустимо растет напряжение отсечки и напряжение насыщения, соответствующее входу полевого транзистора в режим насыщения. Так как режим насыщения является основным рабочим режимом полевого транзистора, напряжение отсечки должно быть малым. Поэтому толщину канала стараются сделать небольшой, несмотря на некоторое уменьшение крутизны характеристики.

§ 6.3. ЭКВИВАЛЕНТНЫЕ СХЕМЫ ПОЛЕВОГО ТРАНЗИСТОРА С УПРАВЛЯЮЩИМ p - n -ПЕРЕХОДОМ

Исходя из принципа действия и структуры полевого транзистора, можно составить его эквивалентную схему для низких частот (рис. 6.5).

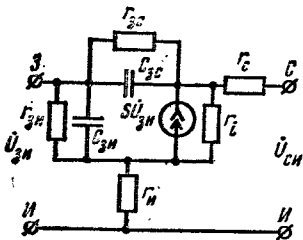


Рис. 6.5. Физическая эквивалентная схема полевого транзистора

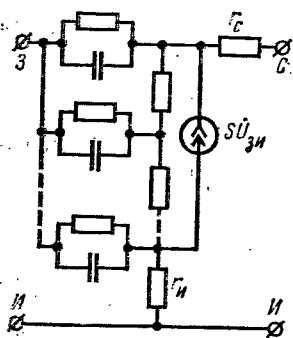


Рис. 6.6. Физическая эквивалентная схема с распределенными параметрами полевого транзистора

ному напряжению U_{zn} ; коэффициентом пропорциональности является крутизна характеристики S .

Нужно учитывать, что емкость и сопротивление затвора распределены по всей его площади и что сопротивление канала также является распределенным. В этом случае эквивалентная схема полевого транзистора должна быть представлена в виде схемы с распределенными параметрами (рис. 6.6). Однако такая схема значительно сложнее для выяснения свойств и характеристик полевых транзисторов.

Кроме физических эквивалентных схем полевого транзистора можно представить и формальные эквивалентные схемы с y -, z - или h -параметрами. Так как входные и выходные сопротивления полевых транзисторов велики, то практически удобнее измерять и задавать комплексные параметры проводимостей его формальной эквивалентной схемы (рис. 6.7). Токи и напряжения на выводах полевого транзистора в режиме малого сигнала для схемы с общим

Сопротивления r_c и r_n представляют собой объемные сопротивления кристалла полупроводника на участках между концами канала и контактами стока и истока соответственно. Эти сопротивления зависят от конструкции транзистора и технологии его изготовления. На низких частотах влиянием сопротивления r_c можно пренебречь по сравнению с обычно большим сопротивлением нагрузки в цепи стока и большим дифференциальным сопротивлением канала r_i . Общее для входной и выходной цепей сопротивление r_n является сопротивлением внутренней обратной связи в полевом транзисторе, включенном по схеме с общим истоком. Падение напряжения на этом сопротивлении при прохождении тока стока оказывается обратным для p - n -перехода. В свою очередь, увеличение обратного напряжения на p - n -переходе затвора транзистора приводит к уменьшению тока стока.

Емкости C_{zn} и C_{zc} , сопротивления r_{zn} и r_{zc} замещают в этой эквивалентной схеме p - n -переход с его барьерной емкостью и большим активным дифференциальным сопротивлением при обратном смещении.

Генератор тока, включенный параллельно сопротивлению канала, отражает усилительные свойства транзистора. Ток этого генератора пропорционален вход-

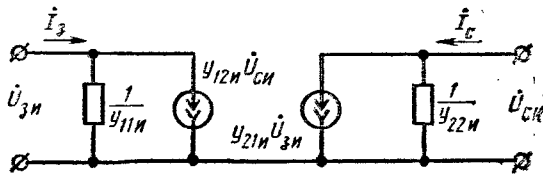


Рис. 6.7. Формальная эквивалентная схема полевого транзистора, соответствующая системе y -параметров

истоком соответствуют следующим характеристическим уравнениям четырехполюсника:

$$\left. \begin{aligned} I_3 &= y_{11n} \dot{U}_{3n} + y_{12n} \dot{U}_{сн}; \\ I_с &= y_{21n} \dot{U}_{3n} + y_{22n} \dot{U}_{сн}. \end{aligned} \right\} \quad (6.14)$$

Определяются y -параметры при режимах короткого замыкания по переменному току на выходе и входе транзистора:

$$\left. \begin{aligned} y_{11n} &= \left. \frac{I_3}{\dot{U}_{3n}} \right|_{\dot{U}_{сн}=0}; & y_{12n} &= \left. \frac{I_3}{\dot{U}_{сн}} \right|_{\dot{U}_{3n}=0}; \\ y_{21n} &= \left. \frac{I_с}{\dot{U}_{3n}} \right|_{\dot{U}_{сн}=0}; & y_{22n} &= \left. \frac{I_с}{\dot{U}_{сн}} \right|_{\dot{U}_{3n}=0} \end{aligned} \right\} \quad (6.15)$$

Если эти режимы воспроизвести на эквивалентной схеме рис. 6.5, то можно найти формулы перехода от параметров сосредоточенных элементов физической эквивалентной схемы к y -параметрам. Пренебрегая малыми проводимостями p - n -перехода затвора и объемными сопротивлениями полупроводника около истока и стока, получим:

$$\left. \begin{aligned} y_{11n} &= j\omega(C_{зс} + C_{зп}); & y_{12n} &= -j\omega C_{зс}; \\ y_{21n} &= S - j\omega C_{зп}; & y_{22n} &= \frac{1}{r_i} + j\omega C_{зс}. \end{aligned} \right\} \quad (6.16)$$

Все эти параметры зависят от значений постоянных смещений на электродах полевого транзистора.

§ 6.4. ЧАСТОТНЫЕ СВОЙСТВА ПОЛЕВЫХ ТРАНЗИСТОРОВ С УПРАВЛЯЮЩИМ p - n -ПЕРЕХОДОМ

Принцип действия полевого транзистора не связан с инжекцией неосновных носителей заряда в базу и их относительно медленным движением до коллекторного перехода. Полевой транзистор — это прибор без инжекции. Поэтому инерционность и частотные свойства полевого транзистора с управляющим p - n -переходом обусловлены инерционностью процесса заряда и разряда барьерной емкости p - n -перехода затвора. Напряжение на затворе измениться мгновенно не может, так как барьерная емкость p - n -перехода затвора перезаряжается токами, проходящими через распределенное сопротивление канала и через объемные сопротивления кристалла полу-

проводника на участках между концами канала и контактами истока и стока. Поэтому не может мгновенно измениться и сечение канала.

На низких частотах полное входное сопротивление полевого транзистора с управляющим *p-n*-переходом определяется большой величиной $r_{\text{зи}}$. С ростом частоты входного сигнала полное входное сопротивление транзистора уменьшается в связи с наличием емкости $C_{\text{зи}}$. Следовательно, для управления полевым транзистором при высоких частотах необходима большая мощность входного сигнала.

Кроме того, наличие в полевом транзисторе проходной емкости $C_{\text{зс}}$, аналогичной емкости анод—сетка в вакуумной лампе, приводит к возникновению в полевом транзисторе частотно-зависимой обратной связи. С ростом частоты увеличивается обратная связь через цепь $r_c C_{\text{зс}}$ (см. рис. 6.5), что эквивалентно уменьшению полного входного сопротивления полевого транзистора и уменьшению его усиления.

Для получения оптимального усиления в реальных схемах на полевых транзисторах необходимо согласование внешних сопротивлений с входным и выходным сопротивлениями транзистора. Поэтому во внешней цепи входа и выхода полевого транзистора обычно есть большие сопротивления, которые значительно увеличивают постоянные времени перезаряда емкостей полевого транзистора.

В связи с перечисленными причинами максимальные рабочие частоты реальных схем на полевых транзисторах с управляющим *p-n*-переходом не превышают нескольких сотен мегагерц.

§ 6.5. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С ИЗОЛИРОВАННЫМ ЗАТВОРОМ (МДП-ТРАНЗИСТОРЫ)

Полевой транзистор с изолированным затвором (МДП-транзистор) — это полевой транзистор, затвор которого отделен в электрическом отношении от канала слоем диэлектрика.

Структуры полевых транзисторов с изолированным затвором показаны на рис. 6.8.

В кристалле полупроводника с относительно высоким удельным сопротивлением, который называют *п о д л о ж к о й*, созданы две сильнолегированные области с противоположным типом электропроводности. На эти области нанесены металлические электроды — исток и сток. Расстояние между сильнолегированными областями истока и стока может составлять всего несколько микрометров. Поверхность кристалла полупроводника между истоком и стоком покрыта тонким слоем (порядка 0,1 мкм) диэлектрика. Так как исходным полупроводником для полевых транзисторов обычно является кремний, то в качестве диэлектрика используется слой двуокиси кремния SiO_2 , выращенный на поверхности кристалла кремния путем высокотемпературного окисления. На слой диэлектрика нанесен металлический электрод — затвор. Получается структура, состоящая из металла, диэлектрика и полупроводника. Поэтому

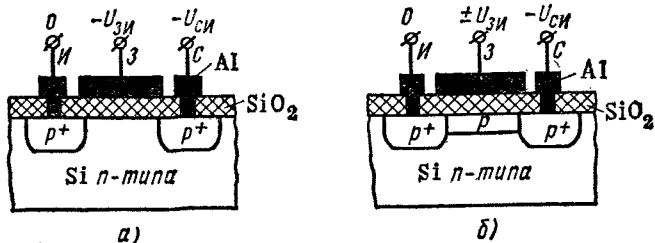


Рис. 6.8. Структуры полевых транзисторов с изолированным затвором (с p -каналом):

a — с индуцированным каналом; b — со встроенным каналом

полевые транзисторы с изолированным затвором часто называют МДП-транзисторами.

Существуют две разновидности МДП-транзисторов: с индуцированным каналом и со встроенным каналом.

В МДП-транзисторах с индуцированным каналом (рис. 6.8, a) проводящий канал между сильнолегированными областями истока и стока и, следовательно, заметный ток стока появляется только при определенной полярности и при определенном значении напряжения на затворе относительно истока, которое называют пороговым напряжением ($U_{зи.пор}$).

В МДП-транзисторах со встроенным каналом (рис. 6.8, b) у поверхности полупроводника под затвором при нулевом напряжении на затворе относительно истока существует инверсный слой — канал, который соединяет исток со стоком.

Изображенные на рис. 6.8 структуры полевых транзисторов с изолированным затвором имеют подложку с электропроводностью n -типа. Поэтому сильнолегированные области под истоком и стоком, а также индуцированный и встроенный канал имеют электропроводность p -типа. Если же аналогичные транзисторы созданы на подложке с электропроводностью p -типа, то канал у них будет иметь электропроводность n -типа.

МДП-транзисторы с индуцированным каналом

Принцип действия. При напряжении на затворе относительно истока, равном нулю, и при наличии напряжения на стоке ток стока оказывается ничтожно малым. Он представляет собой обратный ток p - n -перехода между подложкой и сильнолегированной областью стока. При отрицательном потенциале на затворе (для структуры, показанной на рис. 6.8, a) в результате проникновения электрического поля через диэлектрический слой в полупроводник при малых напряжениях на затворе (меньших $U_{зи.пор}$) у поверхности полупроводника под затвором возникает обедненный основными носителями слой и область объемного заряда, состоящая из ионизированных нескомпенсированных примесных атомов. При напряже-

ниях на затворе, больших $U_{зи,пор}$, у поверхности полупроводника под затвором возникает инверсный слой (см. § 1.12), который и является каналом, соединяющим исток со стоком. Толщина и поперечное сечение канала будут изменяться с изменением напряжения на затворе, соответственно будет изменяться и ток стока, т. е. ток в цепи нагрузки и относительно мощного источника питания (схема включения полевого транзистора с изолированным затвором аналогична схеме включения полевого транзистора с управляющим $p-n$ -переходом, показанной на рис. 6.1, *в*, но полярности внешних источников питания различны для транзисторов с p - и n -каналом). Так происходит управление током стока в поле-вом транзисторе с изолированным затвором и с индуцированным каналом.

В связи с тем, что затвор отделен от подложки диэлектрическим слоем, ток в цепи затвора ничтожно мал, мала и мощность, потребляемая от источника сигнала в цепи затвора и необходимая для управления относительно большим током стока. Таким образом, МДП-транзистор с индуцированным каналом может производить усиление электромагнитных колебаний по напряжению и по мощности.

Принцип усиления мощности в МДП-транзисторах можно рассматривать с точки зрения передачи носителями заряда энергии постоянного электрического поля (энергии источника питания в выходной цепи) переменному электрическому полю. Этот общий для различных приборов принцип усиления мощности был рассмотрен в § 4.2 для объяснения усиления по мощности биполярных транзисторов. В МДП-транзисторе до возникновения канала почти все напряжение источника питания в цепи стока падало на полупроводнике между истоком и стоком, создавая относительно большую постоянную составляющую напряженности электрического поля. Под действием напряжения на затворе в полупроводнике под затвором возникает канал, по которому от истока к стоку движутся носители заряда — дырки. Дырки, двигаясь по направлению постоянной составляющей электрического поля, разгоняются этим полем и их энергия увеличивается за счет энергии источника питания в цепи стока. Одновременно с возникновением канала и появлением в нем подвижных носителей заряда уменьшается напряжение на стоке, т. е. мгновенное значение переменной составляющей электрического поля в канале направлено противоположно постоянной составляющей. Поэтому дырки тормозятся переменным электрическим полем, отдавая ему часть своей энергии.

Выходные статические характеристики. Характер зависимостей $I_c = f(U_{си})$ при $U_{зи} = \text{const}$ для МДП-транзистора с индуцированным каналом аналогичен характеру таких же зависимостей для полевого транзистора с управляющим $p-n$ -переходом (см. § 6.1). Сублинейность крутых частей характеристик (рис. 6.9, *а*) объясняется уменьшением толщины канала около стока при увеличении напряжения на стоке и неизменном напряжении на затворе; так как на сток и на затвор подаются потенциалы одного знака отно-

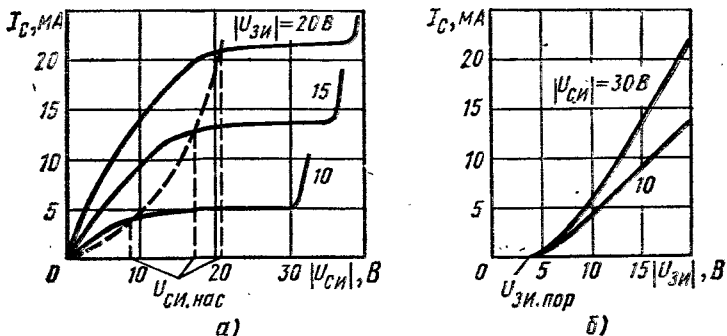


Рис. 6.9. Выходные статические характеристики (а) и статические характеристики передачи (б) МДП-транзистора с индуцированным каналом

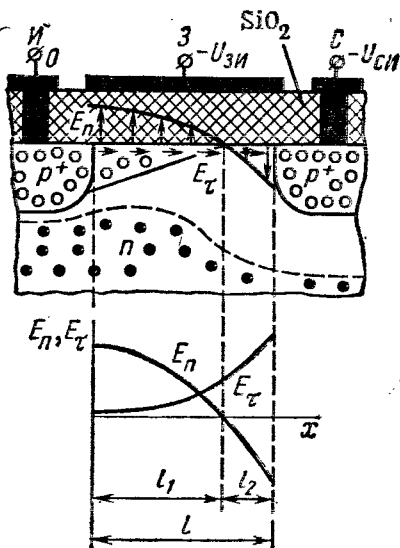


Рис. 6.10. Распределение напряженности электрического поля у поверхности полупроводника под затвором при напряжении на стоке, превышающем напряжение насыщения

сительно истока. Следовательно, разность потенциалов между стоком и затвором или между затвором и прилегающей к стоку частью канала уменьшается. Другими словами, из-за прохождения по каналу тока стока получается неэквипотенциальность канала по его длине. Поэтому при увеличении тока стока происходит уменьшение поперечного сечения канала около стока.

При напряжении насыщения $U_{СИ,нас}$ происходит перекрытие канала около стока, и дальнейшее увеличение напряжения на стоке вызывает очень малое увеличение тока стока. Распределение напряженности электрического поля у поверхности полупроводника при напряжении на стоке, превышающем напряжение насыщения, т. е. для пологой части выходных статических характеристик, показано на рис. 6.10. На расстоянии l_1 от сильнолегированной области истока преобладает нормальная составляющая напряженности электрического поля, созданная напряжением на затворе. На этом

участке существует инверсный слой у поверхности полупроводника. На расстоянии l_2 от сильнолегированной области стока преобладает касательная составляющая электрического поля, созданная напряжением на стоке относительно истока. Несмотря на то что на участке канала протяженностью l_2 нормальная составляющая напряженности имеет другое направление и отталкивает дырки от поверхности полупроводника, через этот перекрытый участок канала идет ток, связанный с движением дырок под действием сильного тянущего поля (касательной составляющей).

Сублинейный характер зависимостей $I_C = f(U_{СИ})$ вызван также эффектом насыщения дрейфовой скорости носителей заряда или уменьшением их подвижности в сильных полях, как и в полевых транзисторах с управляющим p - n -переходом.

При увеличении напряжения на затворе (по абсолютному значению) выходные статические характеристики смещаются в область больших токов стока (см. рис. 6.9, а), что легко понять на основе принципа действия МДП-транзистора с индуцированным каналом.

При больших напряжениях на стоке может произойти пробой МДП-транзистора, при этом может быть два вида пробоя — пробой p - n -перехода под стоком и пробой диэлектрика под затвором.

Пробой p - n -перехода обычно имеет лавинный характер, так как МДП-транзисторы изготавливают обычно на кремнии. При этом на пробивное напряжение $U_{СИ,проб}$ может влиять напряжение на затворе: так как на сток и на затвор МДП-транзистора с индуцированным каналом подаются потенциалы одной полярности, то с увеличением напряжения на затворе будет увеличиваться $U_{СИ,проб}$ (см. рис. 6.9, а).

Пробой диэлектрика под затвором может происходить при напряжении на затворе всего в несколько десятков вольт, так как толщина слоя двуокиси кремния около 0,1 мкм. Пробой обычно имеет тепловой характер, происходит при шунтировании тока (см. § 3.13), и поэтому даже при небольших энергиях импульсов напряжения могут произойти необратимые изменения в диэлектрике. Этот вид пробоя может возникать в результате накопления статических зарядов, так как входное сопротивление МДП-транзисторов велико. Для исключения возможности такого вида пробоя вход МДП-транзистора часто защищают стабилитроном, ограничивающим напряжение на затворе.

Статические характеристики передачи. Характер зависимостей $I_C = f(U_{ЗИ})$ при $U_{СИ} = \text{const}$ ясен из принципа действия МДП-транзистора с индуцированным каналом. Характеристики для разных напряжений на стоке выходят из точки на оси абсцисс, соответствующей пороговому напряжению $U_{ЗИ,пор}$ (см. рис. 6.9, б). С увеличением напряжения на стоке при неизменном напряжении на затворе ток стока возрастает даже в пологой части статических выходных характеристик (см. рис. 6.9, а), что приводит к смещению характеристик передачи вверх в выбранной системе координат.

Интересным и важным с точки зрения применения МДП-транзисторов является температурное изменение статических характери-

стик передачи. Эти изменения вызваны в основном двумя физическими процессами. Во-первых, с увеличением температуры в рабочем диапазоне температур уменьшается подвижность носителей заряда, что приводит к уменьшению тока стока. Во-вторых, происходит перераспределение носителей по энергиям и смещение уровня Ферми к середине запрещенной зоны (см. § 1.8). В связи с таким смещением уровня Ферми инверсный слой образуется у поверхности полупроводника при меньших напряженностях электрического поля (см. рис. 1.21). Поэтому с увеличением температуры пороговое напряжение $U_{зи.пор}$ уменьшается. В результате статические характеристики передачи для неизменного напряжения на стоке, но для разных температур пересекаются (рис. 6.11).

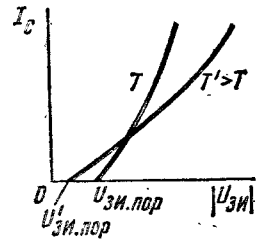


Рис. 6.11. Изменение статических характеристик передачи и порогового напряжения на затворе $U_{зи.пор}$ МДП-транзистора с изменением температуры

Таким образом, температурные изменения тока стока при неизменных напряжениях на МДП-транзисторе могут быть как отрицательными, так и положительными, а также нулевыми в определенной рабочей точке статических характеристик. Обычно эффект температурной компенсации получается при напряжениях на затворе, незначительно превышающих пороговое напряжение $U_{зи.пор}$. Кроме того, еще надо учитывать, что крутизна характеристики S , определяющая усилительные свойства МДП-транзистора, изменяется с температурой даже при неизменном постоянном токе стока.

МДП-транзисторы со встроенным каналом

В связи с наличием встроенного канала в таком МДП-транзисторе при нулевом напряжении на затворе (см. рис. 6.8, б) поперечное сечение и проводимость канала будут изменяться при изменении напряжения на затворе как отрицательной, так и положительной полярности. Таким образом, МДП-транзистор со встроенным кана-

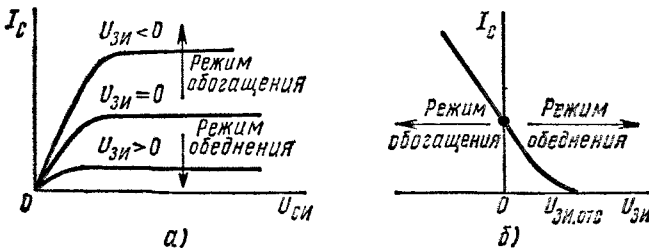


Рис. 6.12. Выходные статические характеристики (а) и статические характеристики передачи (б) МДП-транзистора со встроенным каналом

лом может работать в двух режимах: в режиме обогащения и в режиме обеднения канала носителями заряда. Эта особенность МДП-транзисторов со встроенным каналом отражается и на смещении выходных статических характеристик при изменении напряжения на затворе и его полярности (рис. 6.12, а).

Статические характеристики передачи (рис. 6.12, б) выходят из точки на оси абсцисс, соответствующей напряжению отсечки $U_{зи.отс}$, т. е. напряжению между затвором и истоком МДП-транзистора со встроенным каналом, работающего в режиме обеднения, при котором ток стока достигает заданного низкого значения.

§ 6.6. РАСЧЕТ ВЫХОДНЫХ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЕВОГО ТРАНЗИСТОРА С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Рассмотрим вначале удельную проводимость канала γ_s , т. е. проводимость квадрата канала вдоль оси x единичной длины ($x = 1$) и единичной ширины ($y = 1$) (рис. 6.13). Удельная проводимость канала является величиной, обратной удельному сопротивлению слоя (см. § 1.14):

$$\gamma_s = \frac{1}{\rho_s} = q \int_0^w p \mu_p dz = q p_s \mu_{ps}, \quad (6.17)$$

где $p_s = \int_0^w p dz$ — количество дырок в единичном квадрате канала, которое можно назвать также поверхностной плотностью дырок в канале, так как p_s имеет размерность см^{-2} ; μ_{ps} — эффективная подвижность дырок в канале, которая из-за рассеяния на поверхности полупроводника обычно в несколько раз меньше подвижности дырок в объеме; w — толщина канала.

Удельную проводимость канала можно представить еще в таком виде:

$$\gamma_s = q p_s \mu_{ps} = Q_{ps} \mu_{ps}, \quad (6.18)$$

где Q_{ps} — плотность подвижного заряда (дырок) в канале.

Кроме дырок в канале есть неподвижные заряды — ионизированные примеси в канале, есть неподвижные заряды в слое двуокиси кремния вблизи границы раздела двуокись кремния — кремний. Таким образом, плотность полного заряда, изменяющаяся по длине канала из-за изменения потенциала канала при прохождении тока стока, может быть определена следующим образом:

$$Q_{\text{полн}} = Q_{\text{неподв}} + Q_{ps} = C_{зк} (U_{зи} - U), \quad (6.19)$$

где U — изменяющийся по длине канала потенциал или разность потенциалов между некоторой точкой канала с координатой x и истоком; $C_{зк}$ — удельная емкость между затвором и каналом.

При напряжениях на затворе, не превышающих порогового $U_{\text{зи.пор}}$, канала под затвором еще нет, т. е. под затвором находится обедненный основными носителями слой, содержащий только неподвижные заряды. Поэтому

$$Q_{\text{исподв}} = C_{\text{зк}} U_{\text{зи.пор}} \quad (6.20)$$

Из соотношений (6.19) и (6.20) плотность подвижного заряда

$$Q_{\text{ps}} = C_{\text{зк}} (U_{\text{зи}} - U_{\text{зи.пор}} - U). \quad (6.21)$$

Ток, проходящий по каналу от истока к стоку и одинаковый во всех сечениях по длине канала, или ток стока,

$$I_{\text{с}} = \gamma_{\text{s}} b E_x = \gamma_{\text{s}} b \left| \frac{dU}{dx} \right|, \quad (6.22)$$

где b — ширина канала; E_x — изменяющаяся по длине канала продольная составляющая напряженности электрического поля.

Отсюда, используя (6.18) и (6.21), получим

$$I_{\text{с}} \int_0^l dx = \mu_{\text{ps}} C_{\text{зк}} b \int_0^{U_{\text{си}}} (U_{\text{зи}} - U_{\text{зи.пор}} - U) dU.$$

Тогда для крутой части выходных статических характеристик

$$I_{\text{с}} = \frac{\mu_{\text{ps}} C_{\text{зк}} b}{l} \left[(U_{\text{зи}} - U_{\text{зи.пор}}) U_{\text{си}} - \frac{U_{\text{си}}^2}{2} \right]. \quad (6.23)$$

Для получения выражений, соответствующих пологим частям выходных статических характеристик, надо учесть, что при потенциале стокового конца канала $U(l) \geq U_{\text{си.нас}}$ происходит перекрытие канала около стока. При этом дырок (подвижного заряда) в перекрытой части канала почти нет. Они проходят перекрытую часть канала под действием большой продольной составляющей напряженности электрического поля со скоростью, равной скорости насыщения. Поэтому приблизительно можно считать, что в канале около стока $Q_{\text{ps}} \approx 0$. Тогда, используя (6.21), получим:

$$Q_{\text{ps}} = C_{\text{зк}} (U_{\text{зи}} - U_{\text{зи.пор}} - U_{\text{си.нас}}) \approx 0,$$

отсюда

$$U_{\text{си.нас}} = U_{\text{зи}} - U_{\text{зи.пор}} \quad (6.24)$$

Для пологой части выходных статических характеристик, подставляя (6.24) в (6.23), получим

$$I_{\text{с}} = \frac{\mu_{\text{ps}} C_{\text{зк}} b}{2l} (U_{\text{зи}} - U_{\text{зи.пор}})^2. \quad (6.25)$$

Таким образом, ток стока в первом приближении не зависит от напряжения на стоке при $U_{\text{си}} > U_{\text{си.нас}}$.

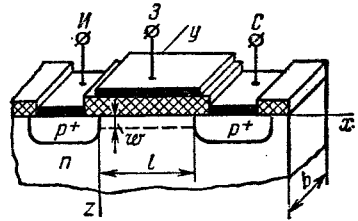


Рис. 6.13. Структура части МДП-транзистора, принятая для расчета выходных статических характеристик

§ 6.7. ПАРАМЕТРЫ И СВОЙСТВА ПОЛЕВЫХ ТРАНЗИСТОРОВ С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Основным параметром полевого транзистора с изолированным затвором, отражающим его усилительные свойства, является крутизна характеристики (см. § 6.1). Крутизна характеристики передачи при низкой частоте, соответствующая крутой части выходных статических характеристик, может быть определена путем дифференцирования (6.23) по напряжению на затворе при неизменном напряжении на стоке:

$$S = \frac{dI_C}{dU_{ЗИ}} \Big|_{U_{СН}=\text{const}} = \frac{\mu_{ps} C_{зк} b}{l} U_{СИ}. \quad (6.26)$$

Для пологой части выходных статических характеристик крутизна характеристики передачи может быть получена путем дифференцирования (6.25) по напряжению на затворе:

$$S = \frac{\mu_{ps} C_{зк} b}{l} (U_{ЗИ} - U_{ЗИ. \text{пор}}). \quad (6.27)$$

Как видно, для увеличения крутизны характеристики исходный полупроводник должен обладать большей подвижностью носителей заряда. Транзистор с *n*-каналом будет иметь большее значение крутизны характеристики по сравнению с транзистором с *p*-каналом, так как подвижность электронов превышает обычно подвижность дырок.

Крутизна характеристики будет больше в полевых транзисторах с меньшей длиной канала. Нижний предел длины канала ограничен технологией изготовления. Обычно для изготовления полевых транзисторов с изолированным затвором применяют планарную технологию и метод фотолитографии, разрешающая способность которого не позволяет получать длину канала меньше 3—4 мкм.

Крутизну характеристики можно увеличить путем увеличения удельной емкости между затвором и каналом. Эта емкость определяется относительной диэлектрической проницаемостью и толщиной слоя диэлектрика под затвором. Использование диэлектрика с большей относительной диэлектрической проницаемостью приведет к увеличению крутизны характеристики, но одновременно увеличатся и паразитные емкости между затвором и стоком, между затвором и истоком, что отрицательно повлияет на частотные свойства полевого транзистора. Уменьшение толщины слоя диэлектрика под затвором может также привести к недопустимому уменьшению пробивного напряжения этого слоя между затвором и стоком.

Увеличение ширины канала приводит к увеличению крутизны характеристики, но одновременно и к ухудшению частотных свойств полевого транзистора в связи с увеличением паразитных емкостей.

Физическая эквивалентная схема полевого транзистора с изолированным затвором аналогична физической эквивалентной схеме полевого транзистора с управляющим *p-n*-переходом (см. рис. 6.5).

Однако в связи с тем, что затвор изолирован от полупроводника слоем диэлектрика, активные сопротивления между затвором и истоком, между затвором и стоком оказываются очень большими. Поэтому ими можно пренебречь даже на относительно малых частотах по сравнению с параллельно включенными емкостными сопротивлениями. Можно пренебречь также очень малыми сопротивлениями r_n и r_c , которые представляют собой дифференциальные сопротивления сильнолегированных областей полупроводника под истоком и стоком.

Быстродействие полевых транзисторов с изолированным затвором определяется временем перезаряда распределенной емкости между затвором и каналом. Постоянные времени процесса перезаряда этой емкости при малом внешнем сопротивлении в цепи затвора ограничивают рабочий диапазон частот полевого транзистора с изолированным затвором частотами около 10 ГГц, т. е. принципиально такие транзисторы могут работать приблизительно до тех же частот, что и биполярные транзисторы. Однако основной особенностью полевых транзисторов является очень большое входное сопротивление. Активная составляющая этого сопротивления может иметь величину до 10^{15} Ом. Поэтому полевые транзисторы применяют в схемах, имеющих также большие сопротивления. Из-за наличия больших внешних сопротивлений в схеме постоянные времени процесса перезаряда емкости между затвором и каналом оказываются увеличенными. В результате полевые транзисторы работают обычно до частот в несколько мегагерц, т. е. рабочий диапазон частот у полевых транзисторов значительно меньше, чем у биполярных транзисторов.

§ 6.8. ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ С ЗАРЯДОВОЙ СВЯЗЬЮ

Прибор с зарядовой связью — это полупроводниковый прибор, в котором может происходить накопление неосновных носителей заряда под электродами МДП-структур (под электродами затворов) и перемещение этих носителей от одного электрода к другому.

Структура и принцип действия приборов с зарядовой связью

Полупроводниковые приборы с зарядовой связью (ПЗС, *CCD* — *charge coupled devices*) были предложены в 1970 г.

Принцип действия ПЗС основан на хранении заряда неосновных носителей заряда в потенциальных ямах, образуемых у поверхности полупроводника под действием внешнего поля, и на перемещении этого заряда вдоль поверхности при сдвиге потенциальных ям.

Рассмотрим принцип действия ПЗС на примере трехтактной схемы сдвигового регистра, которую можно представить как схему МДП-транзистора со многими затворами (рис. 6.14, а). Этот прибор состоит из трех секций.

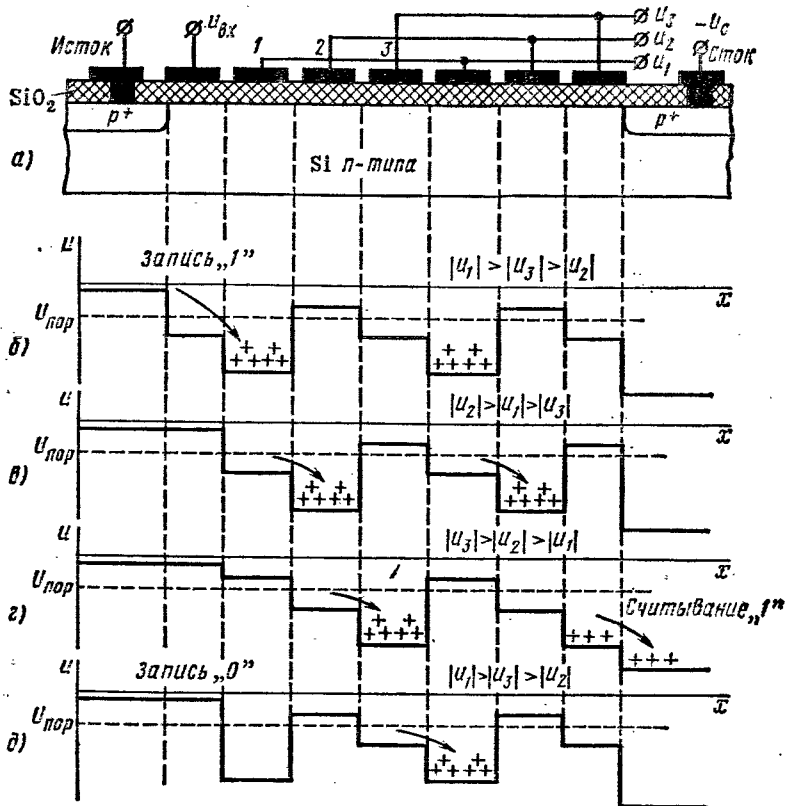


Рис. 6.14. Структура ПЗС с трехзатвортным питанием затворов секции переноса (а) и пояснение принципа его действия (б, в, г, д):

б — запись логической единицы путем нижесекции пакета дырок в потенциальную яму под первым затвором секции переноса; в — перенос информационного заряда — пакета дырок в последующие потенциальные ямы при изменении потенциалов на электродах затвора; г — считывание логической единицы на выходе прибора при экстракции дырок из потенциальной ямы в p^+ -область стока; д — запись логического нуля при отсутствии отрицательного потенциала на входном затворе

Входная секция включает в себя исток с p^+ -областью под ним и входной затвор, выполняющий роль ключа для управления движением дырок из диффузионной p^+ -области истока в первую потенциальную яму.

Вторая секция — секция переноса состоит из ряда затворов, управляющих потенциалом на границе кремний — двуокись кремния. Эти затворы соединены между собой через два. Напряжения на затворах секции переноса имеют вид импульсов различной амплитуды, которые сменяют друг друга циклической перестановкой (рис. 6.14, б — д). При таком изменении напряжения на затворах потенциальные ямы перемещаются к выходу прибора, увлекая за собой пакеты носителей заряда — дырок.

Третья секция прибора — *выходная секция* представляет собой *p-n*-переход стока, смещенный в обратном направлении, напряжение на котором будет изменяться при поступлении пакета носителей заряда — дырок (рис. 6.14, *з*).

Пусть на начальном такте работы на входной затвор подано напряжение $U_{вх}$, достаточное для образования проводящего канала под входным затвором ($U_{вх} > U_{пор}$). Если при этом на первом затворе секции переноса существует достаточно большое отрицательное напряжение, т. е. если под первым затвором секции переноса существует глубокая потенциальная яма для дырок, то дырки будут выходить из истока, проходить по каналу под входным затвором и накапливаться в потенциальной яме под первым затвором секции переноса (рис. 6.14, *б*).

К началу следующего такта напряжение на входном затворе $U_{вх}$ снимается. Поэтому проводящий канал под входным затвором исчезает. Таким образом, произошла запись информации, например логической единицы; так как под первым затвором секции переноса остался пакет дырок. (Для записи логического нуля на начальном такте работы прибора на входной затвор не должно быть подано отрицательное напряжение.)

После смены напряжений на затворах секции переноса самое отрицательное напряжение будет на втором затворе, поэтому пакет дырок передвинется в потенциальную яму под вторым затвором секции переноса (рис. 6.14, *в*). При следующих тактах изменения напряжения на затворах секции переноса будет происходить дальнейшее продвижение пакета дырок к выходной секции (рис. 6.14, *з, д*).

Если в потенциальных ямах, подходящих к *p-n*-переходу стока, нет носителей заряда — дырок, то не будет и изменения тока в цепи стока. И только в том случае, когда потенциальная яма, содержащая дырки, подойдет к *p-n*-переходу стока, дырки втянутся в него и в цепи стока пройдет импульс тока или изменится напряжение на стоке (рис. 6.14, *з*).

Параметры приборов с зарядовой связью

Прежде всего следует отметить, что ПЗС является типично динамическим устройством и имеет нижний и верхний пределы тактовых частот.

Нижний предел тактовой частоты определяется тем, что между потенциальной ямой у поверхности и остальным объемом полупроводника проходят токи, связанные в основном с тепловой генерацией носителей заряда и ничем в принципе не отличающиеся от обратных токов электронно-дырочного перехода. Эти токи прежде всего влияют на уровень логического нуля, повышая заряд дырок в пустых потенциальных ямах. В зависимости от температуры и свойств материала заметное накопление дырок в пустых потенциальных ямах может произойти за время от сотых долей до единиц секунд. Поэтому нижний предел рабочих частот для ПЗС лежит где-то в диапазоне единиц — десятков килогерц.

Верхний предел тактовой частоты определяется временем перетекания заряда из одной потенциальной ямы в другую (порядка единиц наносекунд). За более короткое время весь заряд не успевает перейти из одной потенциальной ямы в соседнюю. Поэтому верхний предел тактовых частот для ПЗС определяется обычно десятками мегагерц.

В диапазоне рабочих частот, далеких от предельных, в ПЗС тоже не происходит полной передачи информационного заряда из одной потенциальной ямы в другую. Связано это с явлениями захвата носителей поверхностными энергетическими уровнями ловушек захвата. В результате, во-первых, уменьшается общий заряд передаваемого пакета дырок, т. е. уменьшается уровень логической единицы. Во-вторых, захваченные ловушками захвата и освободившиеся через некоторое время носители заряда могут попасть в пустые потенциальные ямы, искажив тем самым уровень логического нуля. Для уменьшения влияния этого эффекта необходимо обеспечить плотность поверхностных состояний примерно на два порядка меньшую, чем допускается при производстве МДП-транзисторов.

Другим, более простым, методом борьбы с рассмотренным эффектом является кодирование логического нуля отсутствием заряда в потенциальной яме, а некоторым малым зарядом. Этот заряд во время продвижения по секции переноса ПЗС приводит к заполнению большей части поверхностных состояний, которые уже не могут больше захватывать носители заряда при прохождении пакета дырок. При этом, однако, уменьшается амплитуда сигнала на выходе ПЗС.

Третьим методом борьбы с явлением захвата является использование так называемого углубленного канала. При этом путем введения соответствующих примесей в поверхностный слой полупроводника создают такое распределение электрического поля вблизи поверхности, при котором носители заряда движутся не вблизи поверхности, а на некоторой глубине (до нескольких микрометров). Это резко снижает потери на захват носителей. К тому же в ПЗС с углубленным каналом удастся сдвинуть верхний предел рабочих частот, по крайней мере, за 100 МГц. Однако приборы с углубленным каналом имеют значительно меньшую эффективность управления со стороны затворов и соответственно меньший информационный заряд по сравнению с обычными ПЗС.

Для оценки рассмотренного эффекта захвата носителей в ПЗС используют параметр *эффективности передачи заряда* $\eta = \frac{(Q_{\langle 1 \rangle} - Q_{\langle 0 \rangle})_{i+1}}{(Q_{\langle 1 \rangle} - Q_{\langle 0 \rangle})_i}$, показывающий, какая доля заряда переносится из одной потенциальной ямы i в другую $i + 1$. Значение эффективности передачи η обычно близко к 1, поэтому удобнее пользоваться *коэффициентом потерь* (неэффективностью передачи) $K_n = 1 - \eta$.

Для хороших ПЗС коэффициент потерь составляет менее 10^{-4} . Частотная зависимость коэффициента потерь показана на рис. 6.15.

Спад коэффициента потерь K_n в диапазоне низких частот определяется уменьшением влияния обратных токов между потенциальной ямой и остальным объемом полупроводника (подложкой). Подъем коэффициента потерь при высоких частотах связан с неполным переносом заряда из одной потенциальной ямы в другую из-за быстрого изменения тактовых импульсов на затворах. Конечное значение коэффициента потерь в диапазоне средних частот определяется эффектом захвата носителей заряда ловушками захвата.

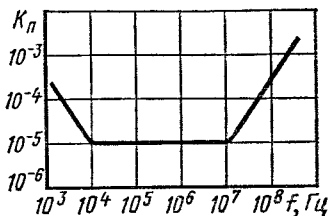


Рис. 6.15. Частотная зависимость коэффициента потерь ПЗС

Как видно, в ПЗС неизбежно происходит потеря информационного заряда, так что достаточная величина выходного сигнала может быть получена при передаче заряда на небольшое число тактов (порядка сотни) и в течение небольшого времени. Чтобы устранить этот недостаток, используют схемы регенерации, которые представляют собой, по существу, усилители. Считываемый с ПЗС сигнал усиливается с соответствующим формированием его уровней, а затем производится запись информации в цепочку ПЗС. Для длительного хранения информации цепочки ПЗС замыкают в кольцо. Регенерация информационного заряда может быть совмещена с выводом информации — тогда получается прибор с неразрушающим считыванием информации.

Через управляющие электроды затворов ПЗС практически не проходит ток электропроводности, так как затворы изолированы от полупроводника (подложки) слоем хорошего диэлектрика — двуокиси кремния. Но в цепи электродов секции переноса все-таки расходуется некоторая мощность на перенос информационного заряда. Эта мощность пропорциональна тактовой частоте. С точки зрения конструкции и технологии ПЗС отличаются от других приборов тем, что имеют малое число диффузионных областей и контактов металла с полупроводником, т. е. малое число потенциально ненадежных элементов структуры. Это приводит к высокому проценту выхода годных приборов при изготовлении ПЗС, к низкой их себестоимости и большой надежности.

Применение приборов с зарядовой связью

К настоящему времени выявились три основных направления использования ПЗС: 1) запоминающие устройства электронных вычислительных машин; 2) устройства преобразования изображения в электрические сигналы; 3) устройства обработки аналоговой информации.

Запоминающие устройства ЭВМ. По принципу действия ПЗС представляют собой запоминающие устройства типа линий задержки. Запоминающие устройства на базе ПЗС доведены до промышленного

выпуска, так как они лучше всего соответствуют природе ПЗС — сдвигающие регистры с последовательным вводом и выводом информации.

Информация непрерывно циркулирует в таком запоминающем устройстве на ПЗС с регенерацией. При обращении к запоминающему устройству производится выборка записанной информации с регенерацией или без нее, т. е. с неразрушающим считыванием или с разрушением записанной информации. Часто в режиме хранения информация циркулирует по прибору относительно медленно, при тактовой частоте порядка десятка килогерц, чтобы только обеспечить регенерацию и не тратить большой мощности на перенос информационного заряда. При обращении к запоминающему устройству тактовая частота повышается до предельной — около нескольких мегагерц, что обеспечивает быструю выборку записанной информации.

Устройства преобразования изображения в электрические сигналы. Принцип действия таких устройств основан на том, что при освещении ПЗС в полупроводнике около его поверхности образуются пары носителей заряда электрон — дырка, которые разделяются электрическим полем потенциальной ямы под затвором секции переноса. Образующиеся при поглощении квантов света носители заполняют потенциальные ямы пропорционально освещенности данной области ПЗС. Если затем произвести обычным путем сдвиг записанной световой информации, то сигнал на выходе ПЗС будет повторять распределение освещенности, т. е. будет выведена строка изображения. Так же может быть выведена следующая строка и т. д. В настоящее время созданы передающие камеры с ПЗС, достигающие обычного телевизионного стандарта по разрешающей способности, в том числе и для цветного телевидения.

Устройства обработки аналоговой информации. С помощью ПЗС могут запоминаться и аналоговые сигналы, но в этом случае становится невозможной регенерация записанной информации. Однако и простое запоминание открывает большие возможности использования ПЗС, так как эти приборы позволяют регулировать задержку переноса информации. Простейшим вариантом использования ПЗС для обработки аналоговой информации оказались линии фиксированной задержки для телевизионных приемников цветного изображения.

§ 6.9. РАЗНОВИДНОСТИ ПРИБОРОВ С ЗАРЯДОВОЙ СВЯЗЬЮ

Принцип действия приборов с зарядовой связью был рассмотрен в § 6.8 на примере ПЗС с трехтактным питанием затворов секции переноса. Конструкция такого прибора имеет ряд недостатков:

1) металлические электроды — затворы необходимо располагать на очень малом расстоянии друг от друга (около 2—3 мкм), что затрудняет их изготовление;

2) при малом расстоянии между затворами возможно загрязнение тонкого слоя двуокиси кремния, расположенного между электродами, примесями из атмосферы;

3) неустранимые пересечения линий металлизации для соединения затворов между собой усложняют технологию.

Эти недостатки частично можно устранить в других конструкциях подобных приборов.

Двухтактные приборы с зарядовой связью

Структура двухтактного ПЗС или ПЗС со ступенчатым диэлектриком отличается чередованием областей полупроводника (кремния) с толстым и тонким слоями двуокиси кремния (рис. 6.16, а). Электрод каждого затвора секции переноса расположен на поверхности толстого и тонкого слоя двуокиси кремния. При приложении напряжения на затвор под ним автоматически получается асимметричная потенциальная яма, конфигурация которой обеспечивает направленное перемещение пакета носителей заряда к стоку (рис. 6.16, б, в). Понятно, что схема управления секцией переноса такого ПЗС значительно проще, чем у трехтактного ПЗС. Соединение затворов между собой в двухтактных ПЗС может быть выполнено без пересечений линий металлизации, что упрощает технологию изготовления таких приборов.

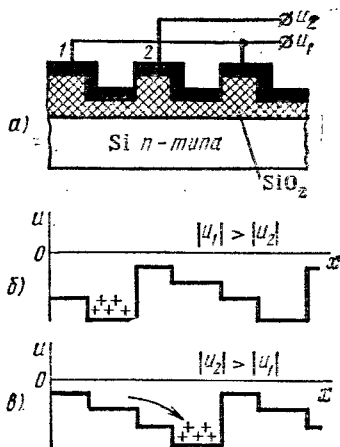


Рис. 6.16. Структура секции переноса ПЗС со ступенчатым диэлектриком (а) и пояснение принципа переноса информационного заряда при двухтактном питании затворов секции переноса (б, в)

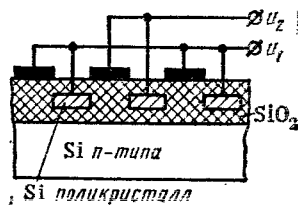
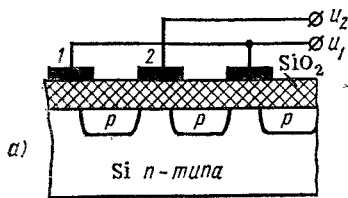


Рис. 6.17. Структура секции переноса ПЗС с рядом скрытых затворов

ПЗС с рядом скрытых затворов (с двухспойной металлизацией)

В структуре секции переноса такого прибора часть затворов сделана из поликристаллического кремния (или из тугоплавкого металла, например молибдена). В зазорах между кремниевыми затворами на поверхности слоя двуокиси кремния расположены металлические затворы (рис. 6.17), которые отделены от затворов из поликристаллического кремния слоем двуокиси кремния, нанесенным после создания затворов из поликристаллического кремния. В такой структуре удается сократить расстояние между



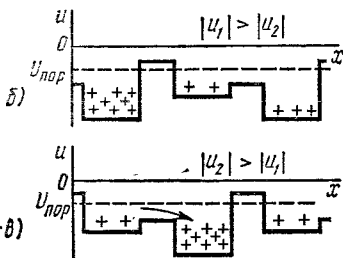
электродами затворов до величины, равной толщине слоя двуокиси кремния.

ПЗС на цепочках МДП-транзисторов

Структура секции переноса таких ПЗС представляет собой цепочку МДП-транзисторов с двухтактными цепями управления (рис. 6.18, а). Диффузионные области с электропроводностью p -типа, являющиеся стоком для одного и истоком для другого МДП-транзистора, не подключаются к источнику питания в отличие от обычных МДП-транзисторов.

Информационный заряд в виде пакета дырок может храниться в диффузионных областях с электропроводностью p -типа, так как потенциальный барьер электронно-дырочной областью и подложкой препятствует растеканию дырок в режиме хранения,

Рис. 6.18. Структура секции переноса ПЗС на цепочках МДП-транзисторов (а) и распределение поверхностного потенциала в режиме переноса информационного заряда (б, в)



несмотря на некоторое прямое смещение этого перехода из-за положительного заряда пакета дырок в диффузионной p -области (рис. 6.18, б).

В режиме передачи информации (пакета дырок) на каждый второй затвор поступает импульс отрицательной полярности, амплитуда которого выше значения $U_{пор}$. По этой причине под каждым вторым затвором в полупроводнике возникает канал. Из-за несимметрии расположения металлического электрода относительно диффузионных областей каждая вторая диффузионная область оказывается под большим отрицательным потенциалом, что обеспечивает направленное движение дырок из одной диффузионной области в следующую, т. е. в более глубокую потенциальную яму (рис. 6.18, в).

Таким образом, принцип действия ПЗС на цепочках МДП-транзисторов аналогичен принципу действия ПЗС других конструкций и структур, которые были рассмотрены ранее. Отличие ПЗС на цепочках МДП-транзисторов от других ПЗС состоит в том, что структура их секций переноса содержит достаточно большое число электронно-дырочных переходов, что делает ее менее совершенной с технологической точки зрения.

Поверхностно-зарядовый транзистор

Простейшая структура поверхностно-зарядового транзистора состоит из кремниевой подложки и трех изолированных друг от друга и от подложки электродов (рис. 6.19). Первый и третий электроды

могут быть скрытыми. Их делают обычно из поликристаллического кремния или тугоплавкого металла (например, молибдена). При подаче на первый или третий электрод отрицательного потенциала в полупроводнике под соответствующим электродом образуется потенциальная яма для неосновных носителей — дырок. Информация в виде пакета дырок может быть введена в эту потенциальную яму путем освещения. Таким образом, первый и третий электроды вместе с областями полупроводника под этими электродами могут выполнять функции приема и хранения информационного заряда.

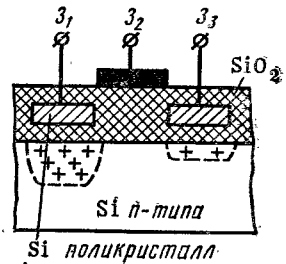


Рис. 6.19. Структура поверхностно-зарядового транзистора при хранении информационного заряда под первым электродом

Второй электрод выполняет функцию затвора — при подаче на него отрицательного потенциала образуется канал, соединяющий потенциальные ямы под первым и третьим электродами. При этом может происходить перетекание информационного заряда из потенциальной ямы, например, под первым электродом в потенциальную яму под третьим электродом, если потенциал третьего электрода более отрицателен.

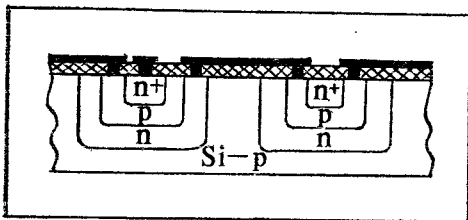
Так же как в ПЗС других конструкций, при статическом режиме работы поверхностно-зарядового транзистора будет происходить накопление паразитного заряда в пустой потенциальной яме из-за тепловой генерации носителей, т. е. в статическом режиме такой прибор работать не может.

Контрольные вопросы

1. Как с помощью полевого транзистора с управляющим *p-n*-переходом производится преобразование энергии относительно мощности источника питания выходной цепи в энергию электромагнитных колебаний?
2. Почему свойства и характеристики полевого транзистора следует описывать системой уравнений, в которых токи являются функциями напряжений?
3. Какие физические факторы влияют на характер зависимости тока стока от напряжения на стоке полевого транзистора с управляющим *p-n*-переходом?
4. Какие параметры характеризуют основные свойства полевых транзисторов?
5. Как можно объяснить усиление по мощности в схеме с полевым транзистором?
6. Какие отличия существуют в структуре МДП-транзисторов с индуцированным и встроенным каналами? Как эти отличия отражаются на статических характеристиках передачи и каковы специфические параметры тех и других транзисторов?
7. Каков принцип действия приборов с зарядовой связью?
8. Какие существуют разновидности структур секции переноса приборов с зарядовой связью?
9. Каков смысл основных параметров приборов с зарядовой связью?

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ КАК ЭЛЕМЕНТЫ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

7



Микроэлектроника — это новое научно-техническое направление электроники, охватывающее проблемы исследования, конструирования и изготовления высоконадежных и экономичных микроминиатюрных электронных схем и устройств с помощью сложного комплекса физических, химических, схемотехнических, технологических и других методов.

§ 7.1. ЗАДАЧИ И ПРИНЦИПЫ МИКРОЭЛЕКТРОНИКИ

Первой задачей микроэлектроники является создание максимально надежных электронных схем и устройств. Эта задача решается в основном путем использования качественно новых принципов изготовления электронной аппаратуры, т. е. путем отказа от использования дискретных компонентов электронной аппаратуры и создания интегральных микросхем, в которых формирование активных (транзисторы, диоды), пассивных (резисторы, конденсаторы) и соединительных элементов электронной схемы происходит на поверхности или в объеме полупроводникового монокристалла или на поверхности диэлектрической подложки в едином технологическом цикле. Минимальное количество внутрисхемных соединений дает возможность резко повысить надежность микроэлектронной аппаратуры. Именно этим преодолеваются сложные противоречия между возросшими требованиями к надежности электронной аппаратуры и ее усложнением. Так, первые радиоприемники имели всего 10—15 элементов (2—3 лампы). Современные радиоприемники имеют на порядок больше элементов. Число элементов в электронной вычислительной машине порядка 10^7 . При таком большом числе дискретных элементов и соответственно большом числе внутрисхемных соединений надежность всего электронного устройства на дискретных элементах оказывается малой. Каждая же интегральная микросхема выполняет функции многих элементов, таких, как резисторы, конденсаторы, диоды, транзисторы и т. д.

Второй задачей микроэлектроники является снижение стоимости электронных схем и устройств. Эта задача решается путем исключения нерациональных технологических операций, сокращения числа внутрисхемных соединений, исключения раздельной герметизации отдельных элементов. Снижение стоимости интегральных микросхем по сравнению со стоимостью аналогичных схем на дискретных элементах происходит из-за сокращения количества сборочных операций и из-за комплексного изготовления различных элементов электронной схемы в едином технологическом цикле. Эти преимущества интегральных микросхем становятся более значительными по мере усложнения интегральных микросхем и возрастания в них количества элементов.

Наряду с решением этих двух важнейших задач микроэлектроники создание и использование интегральных микросхем приводит к резкому уменьшению массы и объема электронной аппаратуры по сравнению с аппаратурой на дискретных элементах, а также к уменьшению потребляемой мощности. Дальнейшее уменьшение массы и объема интегральных микросхем рассматривается как задача второстепенной важности.

Интегральная микросхема — это микроэлектронное изделие, выполняющее определенную функцию преобразования и обработки сигнала и имеющее не менее пяти элементов (транзисторов, диодов, резисторов, конденсаторов), которые нераздельно связаны и электрически соединены между собой так, что устройство рассматривается как единое целое.

С точки зрения интеграции, т. е. объединения многих элементов электронной схемы в одном монокристалле полупроводника или в одном блоке, основными параметрами интегральных микросхем являются плотность упаковки и степень интеграции. *Плотность упаковки* — это число элементов электронной схемы в одном кубическом сантиметре объема интегральной микросхемы. *Степень интеграции* определяется количеством элементов, входящих в состав интегральной микросхемы: если интегральная микросхема содержит до 10 элементов, то ее называют интегральной микросхемой первой степени интеграции; от 10 до 100 элементов — второй степени интеграции; от 100 до 1000 элементов — третьей степени интеграции и т. д.

§ 7.2. КЛАССИФИКАЦИЯ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

Основные направления в микроэлектронике в настоящее время определяются технологией, используемой для создания интегральных микросхем. Поэтому основная классификация интегральных микросхем проводится обычно по конструктивно-технологическому признаку. Различают полупроводниковые и гибридные интегральные микросхемы.

Полупроводниковая интегральная микросхема — это интегральная микросхема, все элементы и межэлементные соединения которой выполнены в объеме и на поверхности полупроводника (рис. 7.1).

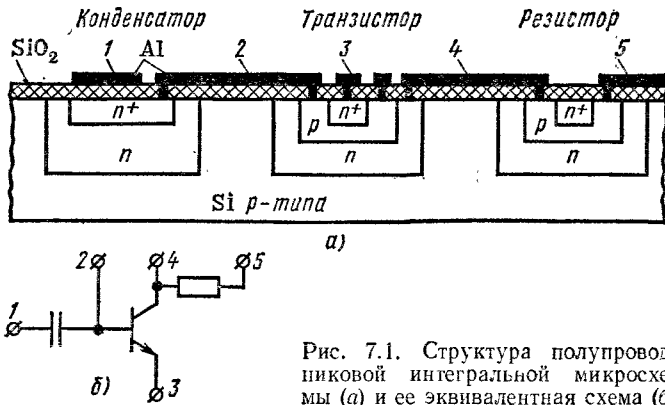


Рис. 7.1. Структура полупроводниковой интегральной микросхемы (а) и ее эквивалентная схема (б)

При изготовлении полупроводниковых интегральных микросхем обычно используют планарную технологию (см. § 2.4). Активные и пассивные элементы полупроводниковой интегральной микросхемы избирательно формируют в одном монокристалле. Изоляцию отдельных элементов в полупроводниковой интегральной микросхеме осуществляют одним из способов, рассмотренных в § 7.3. Соединение элементов между собой в полупроводниковой интегральной микросхеме может быть выполнено как в объеме, так и на поверхности монокристалла полупроводника путем создания инверсных слоев определенной конфигурации с использованием планарной технологии или путем создания на окисленной поверхности полупроводника токоведущих дорожек, например, методом вакуумного напыления металла (алюминия, золота).

Активные элементы (транзисторы и диоды) полупроводниковых интегральных микросхем имеют хорошие электрические характеристики, так как они изготовлены в монокристалле полупроводника. Пассивные элементы (резисторы и конденсаторы) обладают рядом недостатков: диапазон их номиналов ограничен, мала точность изготовления необходимого номинала и получается большой разброс параметров. Кроме того, полупроводниковые резисторы обладают существенной температурной зависимостью, что осложняет конструирование микросхем. Еще одним недостатком полупроводниковых интегральных микросхем является наличие большого числа паразитных связей между элементами и наличие емкостей, отделяющих резистор от полупроводниковой подложки.

Однако несмотря на отмеченные недостатки, полупроводниковые интегральные микросхемы в настоящее время являются одним из наиболее перспективных направлений микроэлектроники, так как они позволяют создавать надежные и достаточно сложные в функциональном отношении схемы малых размеров при малой их стоимости.

Плотность упаковки полупроводниковых интегральных микросхем до 10^4 эл/см², степень интеграции — 3 и выше. Линейные раз-

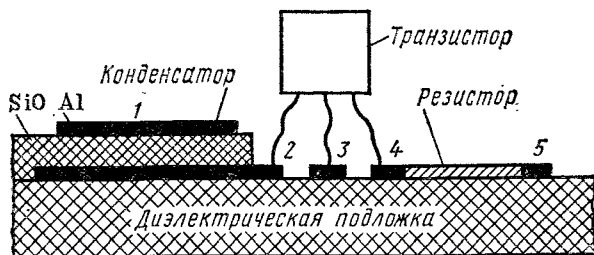


Рис. 7.2. Структура гибридной интегральной микросхемы, эквивалентная схема которой показана на рис. 7.1, б

меры отдельных элементов полупроводниковых интегральных микросхем могут быть очень малы (около 2 мкм), а расстояние между отдельными элементами около 5 мкм.

Гибридная интегральная микросхема — это интегральная микросхема, пассивные элементы которой выполнены посредством нанесения различных пленок на поверхности диэлектрической подложки [из стекла, керамики или ситалла, а активные элементы — навесные полупроводниковые приборы без своих корпусов (рис. 7.2).

Гибридные интегральные микросхемы позволяют использовать преимущества пленочной технологии в сочетании с полупроводниковой технологией. Резисторы и конденсаторы, созданные методами пленочной технологии, могут при малой занимаемой площади иметь большие номиналы и малые температурные изменения параметров, зависящие от выбранного резистивного или диэлектрического материала пленок. Контроль скорости осаждения или нанесения пленок в процессе создания резисторов и конденсаторов позволяет изготовлять их с большой точностью и малым разбросом параметров.

Плотность упаковки гибридных интегральных микросхем несколько меньшая — до 150 эл/см², степень интеграции также меньше аналогичного параметра полупроводниковых интегральных микросхем.

Изготовление гибридных интегральных микросхем перспективно для устройств с относительно небольшим числом активных элементов.

Полупроводниковая интегральная микросхема может быть изготовлена *по совмещенной технологии*: активные элементы выполнены в объеме полупроводникового монокристалла, а пассивные элементы — на защищенной поверхности монокристалла (например, окислом) в тонкопленочном исполнении. На этой же поверхности сделаны и токопроводящие дорожки и площадки (рис. 7.3).

Как для гибридных интегральных микросхем, так и для полупроводниковых интегральных микросхем, изготовленных по совмещенной технологии, процессы создания активных и пассивных элементов разнесены во времени. Поэтому характеристики и параметры пассивных элементов в значительной мере независимы от

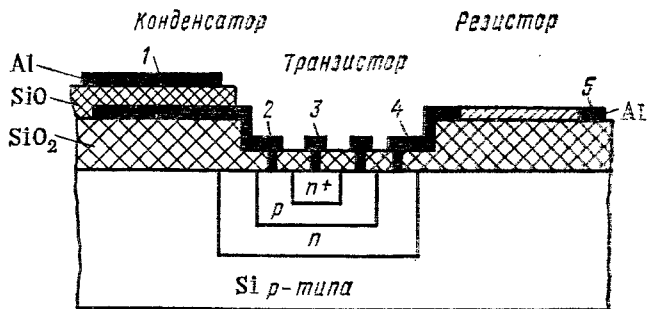


Рис. 7.3. Структура полупроводниковой интегральной микросхемы, изготовленной по совмещенной технологии; эквивалентная схема ее показана на рис. 7.1, б

характеристик и параметров активных элементов. Они определяются выбором материала, толщиной пленок и их конфигурацией.

Поскольку транзисторы и диоды полупроводниковой интегральной микросхемы, изготовленной по совмещенной технологии, находятся внутри монокристалла — подложки, размеры такой интегральной микросхемы могут быть значительно уменьшены по сравнению с размерами гибридной интегральной микросхемы, в которой используются дискретные активные элементы, занимающие сравнительно много места на подложке.

По функциональному назначению интегральные микросхемы делятся на две большие группы: логические (или цифровые) и аналоговые (или линейно-импульсные).

Логические интегральные микросхемы используют в электронных вычислительных машинах, устройствах дискретной обработки информации, системах автоматики.

Аналоговые интегральные микросхемы используют для усиления сигналов низкой и высокой частоты, видеоусилителей, генераторов, смесителей, детекторов и других устройств, где активные элементы работают в линейном режиме или осуществляют нелинейные преобразования входных сигналов.

§ 7.3. МЕТОДЫ ИЗОЛЯЦИИ ЭЛЕМЕНТОВ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

При изготовлении полупроводниковых интегральных микросхем все элементы формируют в одной, чаще всего кремниевой пластине. Для нормальной работы интегральных микросхем необходимо обеспечить хорошую изоляцию между отдельными элементами, с тем чтобы исключить паразитное взаимодействие между ними. Для этой цели используют изоляцию элементов электронно-дырочным переходом, смещенным в обратном направлении, и диэлектрическую изоляцию.

Изоляция элементов электронно-дырочным переходом

Принцип такой изоляции заключается в том, что для каждого элемента в кристалле формируется свой так называемый островок, окруженный электронно-дырочным переходом, который при работе микросхемы смещается в обратном направлении. Ток утечки такой изоляции обычно не превышает 10^{-7} А.

Электронно-дырочный переход, изолирующий отдельные элементы полупроводниковой интегральной микросхемы друг от друга, может быть создан различными способами, которых в настоящее время разработано более десятка. Рассмотрим самый распространенный из них — *планарно-эпитаксиальный с разделительной диффузией* (рис. 7.4). Вначале на пластине кремния выращивают эпитаксиальный слой с электропроводностью n -типа, противоположного типу электропроводности объема полупроводника, т. е. создают эпитаксиальный электронно-дырочный переход по всей площади пластины кремния. Затем проходят обычные этапы планарной технологии: окисление поверхности эпитаксиальной пленки, нанесение фоторезиста, засветка его через маску — фотошаблон, вскрытие окон в двуокиси кремния, локальная диффузия акцепторов через весь эпитаксиальный слой для подложки с электропроводностью p -типа (рис. 7.4, а).

В полученных островках с электропроводностью n -типа (рис. 7.4, б) в дальнейшем формируют структуры транзисторов или других полупроводниковых приборов. Для улучшения некоторых параметров и характеристик транзисторов перед выращиванием эпитаксиального слоя может быть проведена дополнительная диффузия донорной примеси методом планарной технологии с целью создания скрытых легированных слоев (n^+ -слоев). В результате под коллектором будущего транзистора появляется хорошо проводящая область, что и приводит к улучшению характеристик транзистора.

Диэлектрическая изоляция

Метод диэлектрической изоляции также имеет много вариантов. Рассмотрим *метод изоляции элементов двуокисью кремния*. На рис. 7.5 показана последовательность операций при использовании этого метода: окисление монокристаллической пластины кремния (рис. 7.5, а); фотолитография; вытравливание канавок в кремнии

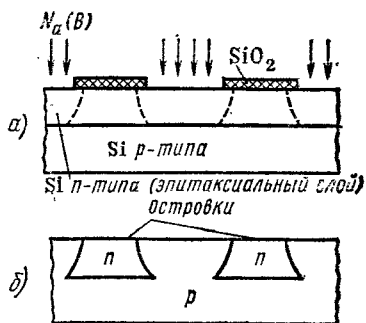


Рис. 7.4. Формирование островков с электропроводностью n -типа в монокристаллической подложке кремния с электропроводностью p -типа планарно-эпитаксиальным методом с разделительной диффузией

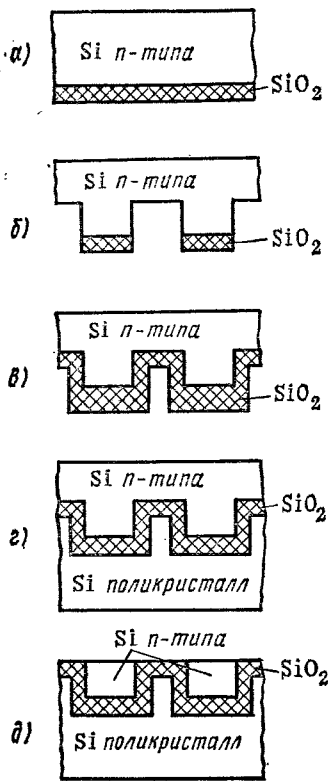


Рис. 7.5. Последовательность основных технологических этапов формирования островков монокристаллического кремния с электропроводностью *n*-типа на поликристаллической подложке кремния методом диэлектрической изоляции:

a — окисление монокристаллического кремния; *б* — вытравливание канавок в кремнии; *в* — повторное окисление кремния; *г* — наращивание поликристаллического кремния; *д* — шлифовка монокристаллического кремния до разделения островков

через вскрытые в двуокиси кремния окна — глубина канавок около 20 мкм (рис. 7.5, б); повторное окисление кремния при высокой температуре или нанесение двуокиси кремния другим способом — толщина слоя окисла около 1 мкм (рис. 7.5, в); выращивание на слое двуокиси кремния поликристаллического кремния толщиной 100—200 мкм путем, например, пиролизического разложения силана (рис. 7.5, г); шлифовка или стравливание с противоположной стороны монокристаллического кремния до разделения островков (рис. 7.5, д).

Метод диэлектрической изоляции позволяет получить хорошую изоляцию как по постоянному, так и по переменному току, поскольку емкость, связанная с окисным слоем, может быть очень малой (300 пФ/мм² при слое окисла толщиной 1 мкм). Пробивное напряжение для диэлектрической изоляции получается значительно большим по сравнению с пробивным напряжением для изоляции электронно-дырочным переходом (выше 800 В).

Основным недостатком этого метода изоляции является его большая стоимость по сравнению с планарно-эпитаксиальным.

§ 7.4. АКТИВНЫЕ ЭЛЕМЕНТЫ

В качестве активных элементов интегральных микросхем используют обычно различные транзисторные структуры, сформированные в монокристаллических пластинках кремния методами планарной технологии. Транзисторы интегральных микросхем могут отличаться не только по структуре, но и по принципу действия.

Биполярный транзистор

Биполярный транзистор является наиболее распространенным активным элементом в современных интегральных микросхемах. Структура биполярного транзистора в интегральных микросхемах (интегрального транзистора) отличается от структуры дискретного

транзистора изоляцией от подложки. Другая особенность связана с тем, что вывод от коллекторной области интегрального транзистора осуществляется на верхней поверхности кристалла. Поэтому для уменьшения объемного сопротивления области коллектора перед эпитаксиальным наращиванием производится обычно подлегирование подложки в тех местах, где будут сформированы транзисторные структуры, т. е. создается скрытый n^+ -слой (рис. 7.6).

Однако даже при наличии скрытого n^+ -слоя сопротивление коллекторной области интегрального транзистора оказывается больше аналогичного сопротивления дискретного транзистора, так как скрытый n^+ -слой отделен от коллекторного электрода высокоомным слоем коллекторной области. Это приводит к некоторому ухудшению частотных свойств интегрального транзистора в связи с увеличением постоянной времени цепи коллектора (времени перезаряда барьерной емкости коллектора). Поэтому граничные частоты ($f_{гр}$) биполярных транзисторов в интегральных микросхемах обычно не превышают 500 МГц.

Кроме того, в связи с увеличенным сопротивлением коллекторной области интегральный транзистор имеет повышенное значение напряжения между коллектором и эмиттером в режиме насыщения $U_{КЭ,нас}$.

Интегральный транзистор отличается от аналогичного дискретного транзистора, изготовленного по той же планарной технологии, значительно меньшими размерами. Объясняется это тем, что в дискретном планарном транзисторе должны быть контактные площадки для подсоединения проволочных выводов эмиттера, базы и коллектора. Размер контактных площадок должен быть не менее 100×100 мкм². В интегральной микросхеме контактные площадки нужны только для подсоединения выводов от всей схемы. Отдельные же элементы интегральной микросхемы соединены между собой межэлементными соединениями в виде тонких и узких (несколько микрометров) металлических полос.

Биполярный транзистор с диодом Шоттки

Биполярный транзистор, особенно в логических интегральных микросхемах, часто выполняет функцию переключающего элемента. При этом он работает не только в активном режиме, но и в режимах насыщения и отсечки. В режиме насыщения происходит накопление

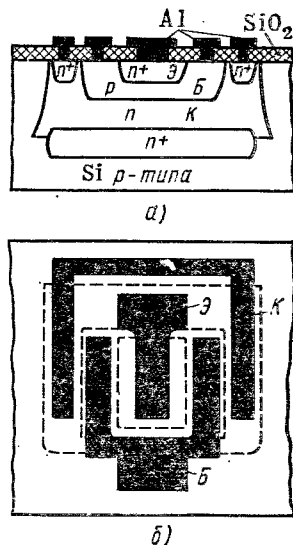


Рис. 7.6. Структура биполярного транзистора со скрытым n^+ -слоем (а) и конфигурацией электродов этого транзистора (б)

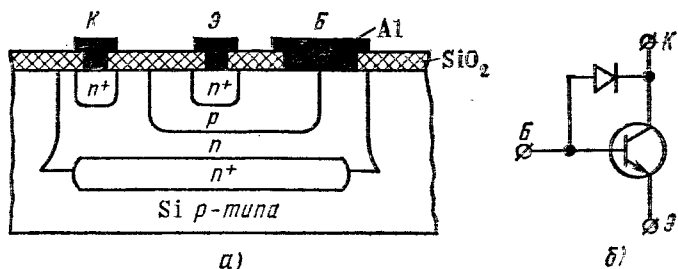


Рис. 7.7. Структура $n-p-n$ -транзистора с диодом Шоттки (а) и его эквивалентная схема (б)

неосновных носителей заряда в базе транзистора, а также в коллекторной области (см. § 4.15). Процессы накопления неосновных носителей и их последующего рассасывания при переводе транзистора в режим отсечки или в выключенное состояние связаны с относительно медленным процессом диффузии неосновных носителей заряда. Инерционность этих процессов определяет скорость перевода транзистора из включенного в выключенное состояние и обратно, т. е. скорость срабатывания схемы.

Для ускорения процесса рассасывания неосновных носителей заряда целесообразно ограничить их накопление. Достичь этого можно путем шунтирования коллекторного перехода транзистора диодом Шоттки, т. е. диодом с выпрямляющим электрическим переходом в виде контакта металла с полупроводником (см. § 3.23). Структура этого транзистора и его эквивалентная схема показаны на рис. 7.7.

Контакт алюминиевого электрода с p -областью базы оказывается невыпрямляющим, а контакт алюминиевого электрода с относительно высокоомной n -областью коллектора получается выпрямляющим. Из-за неравенства работ выхода электронов из алюминия и из кремния с электропроводностью n -типа и в результате химической обработки поверхности кремния на контакте возникает потенциальный барьер для электронов высотой около 0,6 эВ (рис. 7.8, а), что несколько меньше высоты потенциального барьера $p-n$ -перехода коллектора. Поэтому при прямом смещении коллекторного перехода и соответственно при прямом смещении выпрямляющего контакта алюминий — кремний с электропроводностью n -типа (диод Шоттки) основная часть прямого тока коллектора будет проходить через диод Шоттки. Этот ток связан с движением электронов из n -области коллектора в металл и не сопровождается инжекцией дырок в n -область коллектора. Поэтому в высокоомной области коллектора не происходит накопления неосновных носителей заряда (рис. 7.9).

Кроме того, из-за меньшей высоты потенциального барьера на контакте металл — полупроводник (на диоде Шоттки) по сравнению с высотой потенциального барьера на $p-n$ -переходе коллектора при тех же прямых токах коллектора на коллекторном пере-

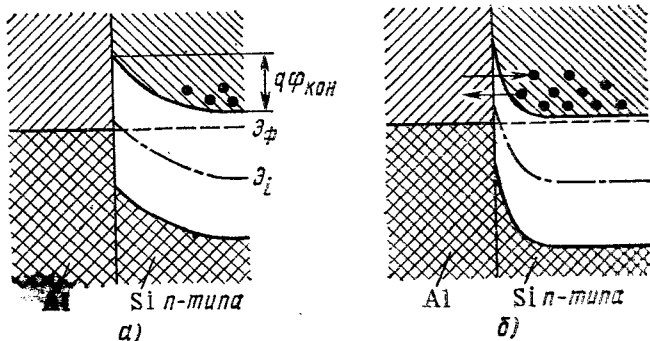


Рис. 7.8. Энергетические диаграммы выпрямляющего контакта алюминиевого электрода с высокоомной n -областью коллектора (а) и невыпрямляющего контакта алюминиевого электрода с сильнолегированной n^+ -областью эмиттера (б) кремниевого транзистора

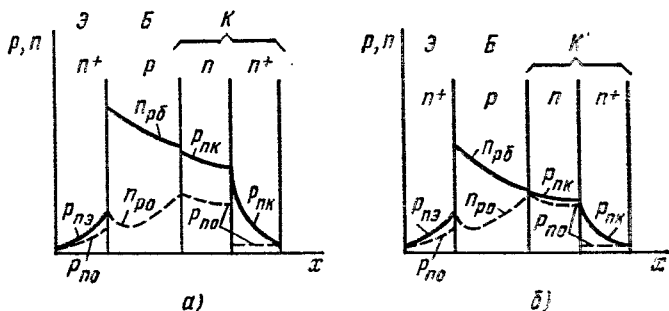


Рис. 7.9. Распределение неосновных носителей заряда в различных областях транзистора при его работе в режиме насыщения:

а — в структуре обычного планарного транзистора; б — в структуре аналогичного транзистора с диодом Шоттки, включенным параллельно коллекторному переходу

ходе будет меньше прямое напряжение, что соответствует меньшему количеству накопленных неосновных носителей заряда в базе транзистора при режиме насыщения (рис. 7.9). В результате время рассасывания в транзисторе с диодом Шоттки оказывается значительно меньше (несколько наносекунд), чем время рассасывания в транзисторе аналогичной структуры, но без шунтирующего диода Шоттки.

Отметим, что на контакте алюминиевых электродов с сильнолегированной n^+ -областью эмиттера и с сильнолегированной частью n^+ -области коллектора также возникают потенциальные барьеры, но их толщина оказывается настолько малой, что сквозь такие узкие потенциальные барьеры электроны могут проходить практически беспрепятственно путем туннелирования (см. рис. 7.8, б). Таким образом, контакты алюминиевых электродов с эмит-

терной областью и с сильнолегированной частью коллекторной области получаются невыпрямляющими, а их формирование и формирование выпрямляющего контакта диода Шоттки осуществляется во время одного процесса металлизации.

Изготовление интегрального транзистора с диодом Шоттки не требует введения дополнительных технологических операций. Необходимо лишь изменить соответствующим образом фотошаблон, применяемый при проведении фотолитографии для снятия двуокиси кремния под контакты, и расширить слой напыляемого алюминия за металлургическую границу коллекторного перехода. Однако при снятии двуокиси кремния в месте выхода коллекторного перехода на поверхность монокристалла кремния и при обработке этой поверхности перед нанесением алюминиевой металлизации следует предотвратить возможность загрязнения *p-n*-перехода коллектора неконтролируемыми примесями.

Полевой транзистор с изолированным затвором

В связи с особенностями структуры МДП-транзисторов их можно формировать без специальных островков в монокристалле интегральной микросхемы, что упрощает технологию — уменьшает число технологических операций, удешевляет интегральные микросхемы и дает возможность увеличить плотность упаковки. Другая особенность и преимущество МДП-транзисторов в качестве активных элементов интегральных микросхем состоит в том, что при нулевом напряжении на затворе МДП-транзистора с индуцированным каналом ток стока практически отсутствует, т. е. транзистором потребляется мощность только во время подачи напряжения на затвор. Это уменьшение потребляемой мощности интегральных микросхем на МДП-транзисторах с индуцированным каналом особенно существенно для создания логических интегральных микросхем. Важным также является то обстоятельство, что логические интегральные микросхемы могут быть построены целиком на гальванически соединенных между собой МДП-транзисторах без использования других элементов.

Диод

Этот активный элемент широко используется в интегральных микросхемах, особенно в логических интегральных микросхемах. Для создания диода вообще достаточно сформировать только один электронно-дырочный переход. Однако диодам в интегральных микросхемах придают транзисторную структуру и в зависимости от конкретного назначения используют тот или иной электронно-дырочный переход путем применения одного из пяти возможных вариантов включения (рис. 7.10).

В первом варианте (*I*) используется эмиттерный переход, а коллекторный короткозамкнут. Такое включение используют в логи-

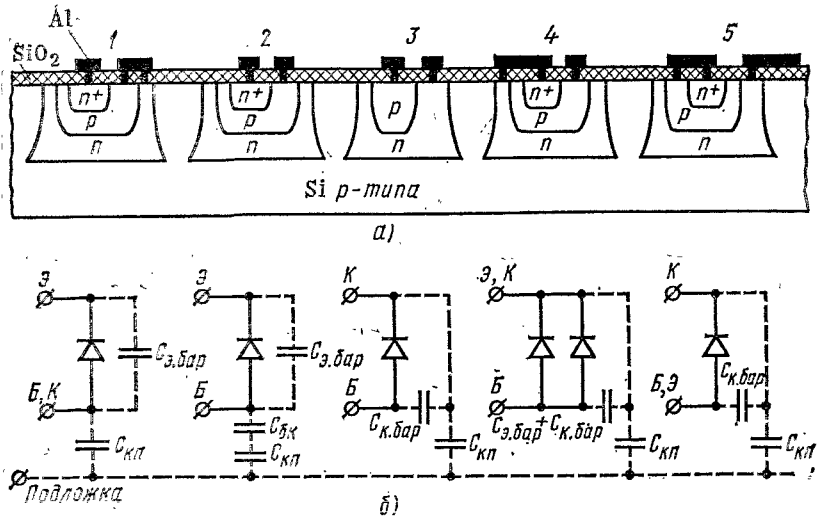


Рис. 7.10. Возможные варианты использования электронно-дырочных переходов транзисторной структуры в качестве диода (а) и их эквивалентные схемы (б)

ческих микросхемах, так как в этом случае достигается наибольшее быстродействие: накопление носителей заряда может происходить только в базовой области, а она очень тонкая. Возможность накопления носителей заряда в коллекторной области исключена шунтированием коллекторного перехода. Время переключения может быть около 1 нс.

Во втором варианте (2) используется эмиттерный переход, а коллекторная цепь разомкнута.

В третьем варианте (3) используется коллекторный переход, а эмиттерной области при этом может и не быть, т. е. этап диффузии примесей для формирования эмиттерной области может быть исключен из технологического процесса. Если же эмиттерная область сформирована, то цепь эмиттера остается разомкнутой. Коллекторная область обычно является относительно высокоомной, поэтому такой диод имеет достаточно высокое пробивное напряжение (~ 50 В). Площадь коллекторного перехода значительно больше площади эмиттерного перехода, поэтому использование коллекторного перехода в качестве диодной структуры дает возможность пропускать большие прямые токи.

В четвертом варианте (4) эмиттерную и коллекторную области соединяют между собой, т. е. эмиттерный и коллекторный переходы включают параллельно. Допустимый прямой ток при этом оказывается еще больше, но увеличивается также и суммарная барьерная емкость.

В пятом варианте (5) используется коллекторный переход, а эмиттерный короткозамкнут.

Диффузионные резисторы

В полупроводниковых интегральных микросхемах биполярный транзистор является элементом с самой сложной структурой. Для его формирования необходимо провести последовательно несколько этапов диффузии примесей. Чтобы не усложнять технологию изготовления интегральной микросхемы, целесообразно для создания резисторов использовать одну из областей транзисторной структуры: эмиттер, базу или коллектор (рис. 7.11).

Эмиттерная область содержит наибольшую концентрацию примесей и обладает наименьшим удельным сопротивлением слоя. Поэтому эмиттерная область пригодна для формирования диффузионных резисторов только с малым сопротивлением (около 10 Ом). Из-за большой концентрации примесей температурные коэффициенты сопротивления таких резисторов будут малы (см. рис. 1.11).

Коллекторная область транзисторной структуры содержит наименьшую концентрацию примесей. Поэтому коллекторная область вообще пригодна для формирования диффузионных резисторов с большим сопротивлением, но из-за малой концентрации примесей температурные коэффициенты сопротивления таких резисторов велики.

Таким образом, для формирования диффузионных резисторов обычно используют базовую область транзисторной структуры. Без существенного увеличения площади, занимаемой диффузионным резистором, в базовой области могут быть созданы резисторы с сопротивлением до 50 кОм. В то же время такие диффузионные резисторы имеют приемлемые температурные зависимости сопротивления, во всяком случае эти зависимости более слабые, чем для диффузионных резисторов, сформированных в коллекторной области.

Диффузионные резисторы, как и другие резистивные элементы, характеризуют следующие параметры: диапазон номинальных значений сопротивления, допуски по сопротивлению, температурный коэффициент сопротивления, допустимая мощность рассеяния и максимальное напряжение.

Для диффузионных резисторов *диапазон номинальных значений сопротивления*, как отмечалось, ограничен сверху.

В массовом производстве интегральных микросхем диффузионные резисторы не могут быть изготовлены с достаточно малыми пределами допусков по ряду причин. Например, трудно выдерживать необходимую поверхностную концентрацию примесей и глубину диффузии с высокой точностью. Различные этапы фотолитографического процесса, несмотря на совершенство применяемого оборудования, вносят в абсолютную величину сопротивления диффузионного резистора существенную ошибку, особенно для резисторов с узкой шириной резистивной диффузионной полоски.

Температурный коэффициент сопротивления диффузионного резистора зависит от концентрации примесей в использованном диф-

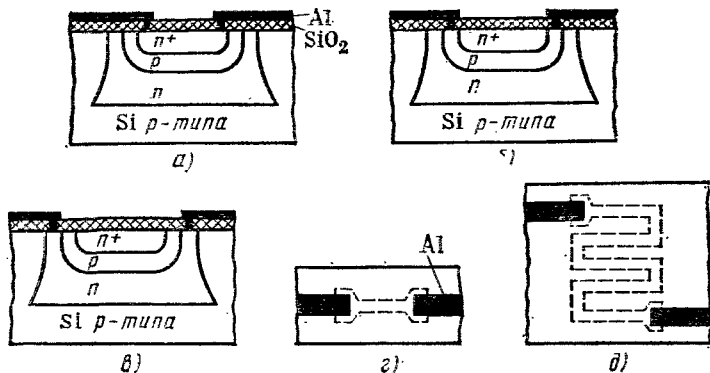


Рис. 7.11. Возможные варианты использования различных областей транзисторной структуры в качестве резистивных слоев для формирования диффузионных резисторов интегральной микросхемы (а, б, в) и некоторые варианты конфигурации резистивного слоя и электродов резистора (г, д)

фузионном слое. Так как диффузионные резисторы формируют в базовых областях транзисторных структур, то концентрация примесей в этих областях определяется необходимыми параметрами и свойствами транзисторов.

Допустимая мощность рассеяния диффузионных резисторов ограничена малыми размерами и топологией резистивной диффузионной полоски резистора (рис. 7.11, г, д), а также связана с температурным коэффициентом сопротивления резистора, так как нагрев резистора проходящим током вызывает изменение сопротивления и приводит к нелинейности вольт-амперной характеристики.

Максимальное напряжение — это специфический параметр диффузионных резисторов. Диффузионный резистор, как отмечалось, представляет собой диффузионную полоску с определенным типом электропроводности, отделенную от подложки интегральной микросхемы $p-n$ -переходом. Электронно-дырочный переход должен быть смещен в обратном направлении для изоляции диффузионной полоски от подложки. Поэтому максимальное падение напряжения на резисторе не может быть больше напряжения смещения, которое в свою очередь не может превышать пробивного напряжения изолирующего $p-n$ -перехода.

Кроме перечисленных недостатков диффузионных резисторов и трудностей пресектирования и создания интегральных микросхем с диффузионными резисторами необходимо отметить, что при формировании интегральных микросхем вообще и микросхем с диффузионными резисторами в частности в структуре интегральной микросхемы образуются паразитные элементы, которые могут нарушить нормальную работу интегральной микросхемы. Диффузионный резистор отделен от подложки электронно-дырочным переходом, который имеет собственную барьерную емкость. Эта

емкость может оказывать существенное влияние на работу схемы при высоких частотах.

Другим паразитным элементом, возникающим возле диффузионного резистора, может быть паразитный биполярный транзистор, эмиттером которого является диффузионный резистивный слой, базой — коллекторная область исходной транзисторной структуры, коллектором — подложка интегральной микросхемы (рис. 7.11, б). Если изолирующий $p-n$ -переход между резистивной полоской и подложкой (эмиттерный переход паразитного транзистора) окажется смещенным в прямом направлении, то паразитный транзистор может нарушить работу интегральной микросхемы. Поэтому и необходимо, чтобы падение напряжения на диффузионном резисторе не превышало максимального напряжения.

Несмотря на отмеченные недостатки, диффузионные резисторы широко применяют в интегральных микросхемах, так как их формирование не требует дополнительных технологических операций и не удорожает схему.

Пленочные резисторы

Основой пленочного резистора является резистивная пленка из металла (хром, тантал, палладий), металлического сплава (нихром) или металлокерамики (кермет). Резистивную пленку определенной конфигурации тем или иным способом наносят на диэлектрическую подложку гибридной интегральной микросхемы или на окисленный кристалл полупроводниковой интегральной микросхемы, изготовленной по совмещенной технологии (рис. 7.12).

Материал, используемый для резистивных пленок, должен обеспечивать возможность получения широкого диапазона номинальных значений сопротивления, обладать низким температурным коэффициентом сопротивления, высокой коррозионной стойкостью и стабильностью параметров во времени. Наибольшее распространение получили нихромовые резистивные пленки из-за их стабильности и возможности их создания с малым температурным коэффициентом сопротивления.

После осаждения пленки нихрома на диэлектрическую подложку производят термообработку путем отжига пленки на воздухе. При этом нихромовая пленка покрывается слоем окисла, который значительно улучшает стабильность пленочного резистора.

При создании пленочных резисторов на диэлектрической подложке не образуются паразитных элементов, которые получаются в полупроводниковых интегральных микросхемах с диффузионными резисторами.

Для нанесения резистивной пленки могут быть использованы различные способы: термическое испарение резистивного материала в вакууме (из резистивного испарителя или с помощью электронного луча), катодное распыление, анодирование осажденных на подложку металлических пленок и химическое осаждение при

химической реакции в газовой фазе. Термическое испарение в вакууме — наиболее простой и производительный метод получения резистивных пленок, а также других металлических и диэлектрических пленок, необходимых для изготовления интегральных микросхем. Унификация технологических операций получения различных элементов интегральной микросхемы, возможность изготовления этих различных элементов в едином технологическом цикле — это существенное преимущество термического испарения в вакууме среди других способов нанесения резистивных пленок.

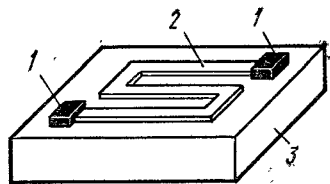


Рис. 7.12. Возможная конфигурация пленочного резистора:

1 — металлические электроды резистора; 2 — резистивная пленка; 3 — диэлектрическая подложка

Диффузионные конденсаторы

В качестве конденсаторов интегральных микросхем часто используют барьерную емкость электронно-дырочного перехода, смещенного в обратном направлении. Такой пассивный элемент интегральной микросхемы удобно формировать одновременно с формированием транзисторных структур или использовать непосредственно электронно-дырочные переходы транзисторных структур (рис. 7.13). Барьерная емкость p - n -перехода может быть использована как для создания конденсатора постоянной емкости, так и для конденсатора переменной емкости, которой можно управлять путем изменения постоянного смещения на переходе (см. § 3.30).

Качество диффузионных конденсаторов, как и других конденсаторов, и их пригодность для применения в интегральных микросхемах необходимо оценивать по их технологической совместимости с технологией других (в первую очередь активных элементов) интегральных микросхем и по таким основным параметрам, как диапазон номинальных значений емкости (или удельная емкость), пробивное напряжение, величина добротности, допуски по емкости.

Диапазон номинальных значений емкости диффузионных конденсаторов, которые могут быть сформированы на отведенных для них площадях монокристалла полупроводника, определяется концентрацией примесей в прилегающих к переходу областях. Диффузионные конденсаторы, использующие эмиттерную емкость транзисторной структуры, имеют большую удельную емкость по сравнению с конденсаторами на коллекторном переходе.

Однако при большой концентрации примесей в прилегающих к переходу областях и, следовательно, при малой толщине перехода будет мало пробивное напряжение такого перехода, а значит, и диффузионного конденсатора. Таким образом, удельную емкость и пробивное напряжение диффузионных конденсаторов надо рассматривать совместно. Взаимосвязь между этими параметрами оказывается неблагоприятной для диффузионных конденсаторов.

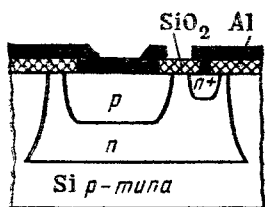


Рис. 7.13. Сечение диффузионного конденсатора с использованием барьерной емкости коллекторного перехода транзисторной структуры (диффузия примесей для формирования эмиттерной области исключена из технологического процесса)

Добротность — величина, обратная тангенсу угла диэлектрических потерь диффузионных конденсаторов, обычно значительно ниже добротности дискретных конденсаторов с диэлектрической изоляцией. Однако стоимость формирования диффузионных конденсаторов мала, так как они создаются в едином технологическом процессе с другими элементами интегральной микросхемы. Поэтому диффузионные конденсаторы широко используют в интегральных микросхемах, когда можно мириться с низкой добротностью.

Диффузионный конденсатор, как и другие элементы интегральной микросхемы, должен быть изолирован от стальных элементов и от подложки интегральной микросхемы. Часто эта изоляция осуществ-

ляется электронно-дырочным переходом. Поэтому при формировании диффузионного конденсатора одновременно образуется и структура паразитного транзистора, эмиттером которого является одна из областей (обкладок) диффузионного конденсатора, базой — другая область (обкладка), коллектором — подложка (рис. 7.13).

В связи с зависимостью барьерной емкости от напряжения смещения на электронно-дырочном переходе диффузионные конденсаторы могут быть использованы для усиления электромагнитных колебаний, т. е. могут быть активными элементами интегральных микросхем.

МДП-конденсаторы

В качестве диэлектрика такого конденсатора используют слой двуокиси кремния, которым покрыт кристалл полупроводника (рис. 7.14). Одной обкладкой конденсатора является слой металла (обычно алюминия), нанесенный на поверхность слоя двуокиси кремния одновременно с созданием межэлементных соединений и контактных площадок. Другой обкладкой — сильнолегированная область полупроводника, которая формируется одновременно с формированием эмиттерных областей транзисторных структур интегральных микросхем. Таким образом, процесс изготовления МДП-конденсаторов также не требует проведения дополнительных операций для их формирования.

В островке, предназначенном для МДП-конденсатора, не формируют базовую область транзисторной структуры, т. е. не проводят диффузию примесей для создания базовой области. Поэтому под МДП-конденсатором есть только один электронно-дырочный переход между коллекторной областью транзисторной структуры и подложкой, который необходим для изоляции МДП-конденсатора от других элементов, расположенных на одной с ним полупроводниковой пластине.

Следует отметить, что емкость МДП-конденсатора может иметь сложную зависимость от напряжения постоянного смещения и от частоты переменного напряжения. Связано это с возможностью образования у поверхности полупроводника (в данном примере у поверхности n^+ -области) обедненных и инверсных слоев под действием проникающего в полупроводник электрического поля. Частотная зависимость емкости МДП-конденсатора может появиться, если в полупроводнике образуется инверсный слой, в котором накопление и рассасывание неосновных для исходного полупроводника носителей заряда происходит в результате процессов тепловой генерации и рекомбинации, а инерционность этих процессов может быть велика.

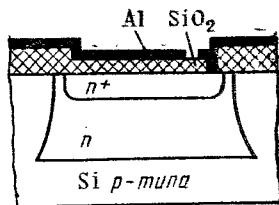


Рис. 7.14. Сечение МДП-конденсатора



Рис. 7.15. Сечение пленочного конденсатора:

1 — диэлектрическая подложка; 2 — нижняя обкладка; 3 — диэлектрическая пленка; 4 — верхняя обкладка

Пленочные конденсаторы

Пленочные конденсаторы формируют на диэлектрической подложке гибридных интегральных микросхем. При этом необходимо провести, по крайней мере, три операции вакуумного напыления: нижней проводящей обкладки конденсатора, диэлектрической пленки и верхней проводящей обкладки (рис. 7.15). Такой пленочный конденсатор называют однослойным. Для получения большей емкости или для уменьшения площади, занимаемой конденсатором на подложке, можно делать многослойные пленочные конденсаторы, секции которых располагают «этажами» — одна над другой. Однако создание «многоэтажных» конденсаторов затрудняет процесс их изготовления, так как надо вводить дополнительные операции нанесения различных слоев, повышает стоимость, уменьшает надежность, увеличивает процент брака из-за увеличения краевого эффекта, уменьшения плотности и электрической прочности верхних диэлектрических слоев.

В качестве диэлектрика пленочных конденсаторов могут быть использованы различные материалы, но наибольшее применение находит моноокись кремния.

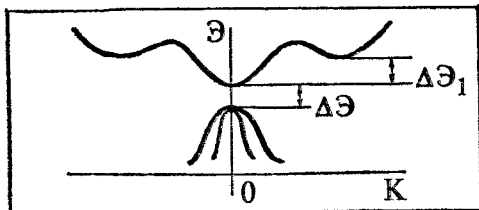
Контрольные вопросы

1. Что такое микроэлектроника?
2. Какие задачи стоят перед микроэлектроникой?
3. Что такое полупроводниковая интегральная микросхема?
4. Что такое гибридная интегральная микросхема?
5. Что такое полупроводниковая интегральная микросхема, изготовленная по совмещенной технологии?

6. Какими методами производят изоляцию элементов интегральных микросхем?
7. Какие элементы интегральных микросхем относят к активным элементам?
8. Какие элементы интегральных микросхем считают пассивными?
9. Почему в качестве диодов интегральных микросхем используют транзисторные структуры?
10. Почему и как емкость диффузионного конденсатора зависит от напряжения постоянного смещения?
11. Какие недостатки и преимущества имеют диффузионные резисторы?
12. Почему не используют прямого смещения электронно-дырочного перехода диффузионного конденсатора?
13. Какие разновидности конденсаторов применяют в интегральных микросхемах?

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ С ИСПОЛЬЗОВА- НИЕМ МЕЖДОЛИННОГО ПЕРЕХОДА НОСИТЕЛЕЙ ЗАРЯДА

8



§ 8.1. ПРИНЦИП ДЕЙСТВИЯ ГЕНЕРАТОРА ГАННА

Генератор Ганна — это полупроводниковый прибор, предназначенный для преобразования энергии источника питания постоянного напряжения в энергию сверхвысокочастотных электромагнитных колебаний в результате возникновения в полупроводнике домена под действием сильного электрического поля.

Физические основы отрицательного дифференциального сопротивления при междолинном переходе носителей заряда

Энергетическая диаграмма некоторых полупроводников (например, арсенида галлия), построенная в пространстве квазиимпульсов (в k -пространстве), может иметь несколько минимумов (см. рис. 1.15). В таком полупроводнике могут существовать электроны с разными подвижностями — «легкие» и «тяжелые» (см. § 1.10). Соотношение между концентрациями «легких» n_1 и «тяжелых» n_2 электронов изменяется при изменении напряженности электрического поля, так как в сильном электрическом поле (при напряженности, большей порогового значения $E > E_{\text{пор}}$) электроны, приобретая дополнительную энергию, превышающую $\Delta\mathcal{E}_1$ (см. рис. 1.15), переходят в боковые долины и становятся «тяжелыми». Если при этом еще не происходит заметной ударной ионизации, то общая концентрация электронов остается неизменной и равной равновесной концентрации:

$$n_1 + n_2 = n_0.$$

Обозначив подвижность «легких» электронов μ_1 , подвижность «тяжелых» электронов μ_2 , выражение для плотности тока через кристалл полупроводника можно записать так:

$$J = q(n_1\mu_1 + n_2\mu_2)E.$$

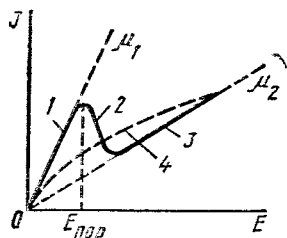


Рис. 8.1. Зависимость плотности тока через полупроводник с многодолинной структурой зоны проводимости от напряженности электрического поля:

1 — при наличии электронов только с большой подвижностью μ_1 (в центральной долине); 2 — переходный участок; 3 — при наличии электронов только с малой подвижностью $\mu_2 \ll \mu_1$ (в боковой долине); 4 — при наличии в кристалле локальных напряженностей, отличных от средней напряженности электрического поля

При слабых электрических полях ($E < E_{\text{пор}}$) практически все электроны находятся в центральной долине, $n_1 \approx n_0$ и плотность тока при этом $J = qn_0\mu_1E$, что соответствует участку 1 вольт-амперной характеристики кристалла полупроводника (рис. 8.1). При сильных электрических полях ($E \gg E_{\text{пор}}$) можно предположить, что практически все электроны приобретут добавочную энергию, большую, чем $\Delta\varepsilon_1$, и окажутся в боковой долине. В этом случае $n_2 \approx n_0$ и $J = qn_0\mu_2E$, что соответствует участку 3 вольт-амперной характеристики рис. 8.1. При средних напряженностях электрического поля, лишь немного превышающих пороговую напряженность, плотность тока определяется соотношением концентрации «легких» и «тяжелых» электронов (участок 2 вольт-амперной характеристики рис. 8.1).

Для возникновения отрицательного дифференциального сопротивления необходим одновременный переход большинства электронов из центральной долины в боковую

при пороговой напряженности электрического поля. Но практически получить статическую вольт-амперную характеристику, соответствующую сплошной кривой рис. 8.1, не удастся, так как в кристалле или около невыпрямляющих контактов всегда есть неоднородности, в результате чего возникают локальные напряженности электрического поля, превышающие среднюю напряженность. Превращение в этих местах «легких» электронов в «тяжелые» еще больше увеличивает неоднородность электрического поля. Поэтому практически не получается одновременного перехода большинства электронов в кристалле из центральной долины в боковую, и статическая вольт-амперная характеристика получается без участка с отрицательным дифференциальным сопротивлением (штриховая кривая 4 на рис. 8.1).

Эффект Ганна

В 1963 г. американский ученый Дж. Б. Ганн обнаружил возникновение электромагнитных колебаний в однородном кристалле полупроводника при приложении к нему постоянного напряжения. Приборы на основе этого эффекта называют либо генераторами Ганна, либо диодами Ганна (хотя такие приборы и не имеют p - n -перехода). Пусть на однородно легированный кристалл (рис. 8.2), имеющий два невыпрямляющих контакта (катод и анод), подано постоянное напряжение, создающее в кристалле напряженность электрического поля, несколько меньшую пороговой

напряженности. При этом все свободные электроны в кристалле являются «легкими», и плотность тока через кристалл имеет максимальное значение:

$$J_{\max} = qn_0\mu_1 E_0 = qn_0 v_0.$$

Локальная напряженность электрического поля около невыпрямляющих контактов из-за наличия различных дефектов может превышать пороговую напряженность электрического поля. Это обеспечит образование «тяжелых» электронов около катода, которые, двигаясь относительно медленно к аноду, создают отрицательный заряд. «Легкие» электроны в остальной части кристалла движутся к аноду быстрее «тяжелых». Поэтому около пакета «тяжелых» электронов со стороны анода получается недостаток электронов, что равносильно образованию некоторого положительного заряда, состоящего из нескомпенсированных ионизированных доноров (рис. 8.3). Таким образом образуется домен, состоящий из двух слоев: слой со стороны катода из-за избытка «тяжелых» электронов имеет отрицательный заряд, слой со стороны анода из-за недостатка электронов имеет положительный заряд. Домен обладает своим электрическим полем $E_{\text{дом}}$, направленным в ту же сторону, что и поле, созданное внешним напряжением. В результате по мере образования домена поле в нем растет, а за пределами домена уменьшается, т. е. скорость движения «тяжелых» электронов внутри домена увеличивается, а скорость движения «легких» электронов за пределами домена уменьшается. В некоторый момент времени скорость движения «тяжелых» электронов (скорость домена) оказывается равной скорости движения «легких» электронов:

$$v_1 = v_2$$

или

$$\mu_1 E_1 = \mu_2 E_2,$$

где v_1 — скорость движения электронов за пределами домена; v_2 — скорость движения электронов внутри домена, что соответ-

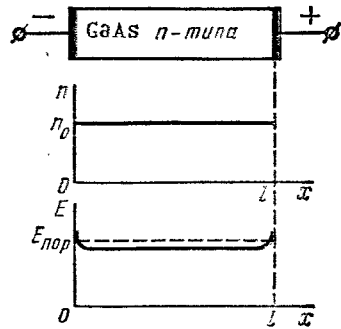


Рис. 8.2. Распределение электронов и напряженности электрического поля в кристалле генератора Ганна в первый момент после подачи постоянного напряжения

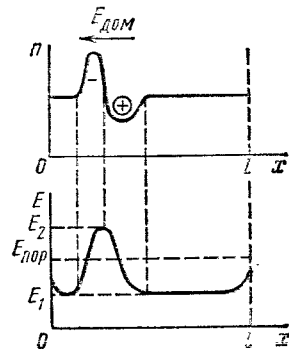


Рис. 8.3. Распределение электронов и напряженности электрического поля в кристалле после формирования домена

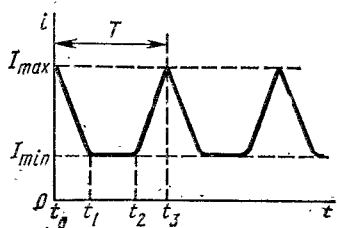


Рис. 8.4. Зависимость тока, проходящего через генератор Ганна, от времени:

t_0 — момент времени начала образования домена; t_1 — момент времени, соответствующий окончанию образования домена; t_2 — момент времени начала исчезновения домена на аноде; t_3 — момент времени, соответствующий полному исчезновению первого домена на аноде и зарождению второго домена на катоде

отсутствию домена. Сразу после этого у катода формируется новый домен, и процесс повторяется. На рис. 8.4 показана зависимость тока, проходящего через кристалл, от времени.

§ 8.2. ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ ГЕНЕРАТОРОВ ГАННА

Несмотря на принципиальную простоту конструкции генератора Ганна, представляющего собой пластину однородного кристалла арсенида галлия с электропроводностью n -типа с невыпрямляющими контактами, нанесенными на ее противоположные стороны, в производстве таких приборов встречаются трудности, связанные прежде всего с воспроизводимостью однородного легирования арсенида галлия.

Удельное сопротивление исходного арсенида галлия выбирают от 1 до 10 Ом·см. Время диэлектрической релаксации, ограничивающее скорость образования доменов, в этом случае 10^{-12} — 10^{-11} с.

В генераторах Ганна, имеющих толщину кристалла между электродами более 200 мкм, колебания, как правило, некогерентны. Это связано с наличием в каждом из кристаллов нескольких дефектов, на которых могут зарождаться домены. Путь, пробегаемый доменом от места его зарождения до анода, определяет период колебаний. Поэтому если домены зарождаются на различных неоднородностях кристалла, т. е. на различных расстояниях от анода, то колебания имеют шумовой характер. Применение генераторов Ганна практически оправдано в диапазоне частот более 1 ГГц, что соответствует толщине кристаллов полупроводника между электродами $l \leq 100$ мкм.

В свою очередь наименьшая толщина кристаллов полупроводника для генераторов Ганна ограничивается технологическими трудностями. Так, при использовании традиционных методов

ствует скорости движения домена от катода к аноду.

Очевидно, что $v_1 < v_0$, так как $E_1 < E_0$. Поэтому после образования домена плотность тока через кристалл уменьшится до

$$J_{\min} = qn_0v_1.$$

Минимальное значение плотности тока через кристалл будет сохраняться в течение всего времени движения домена через кристалл или в течение времени пролета

$$t_{\text{прол}} = l/v_2,$$

где l — длина кристалла.

При достижении анода домен исчезает и плотность тока возрастает до величины J_{\max} , соответствующей

изготовления тонких пластинок из арсенида галлия (шлифовки, полировки и травления) удается получить генераторы Ганна с $l \approx 2$ мкм, величиной порогового напряжения 1 В и частотой генерации около 30 ГГц. Однако такой метод производства генераторов Ганна малой длины чрезвычайно труден и слишком трудоемок.

Более приемлем метод фотолитографии, который позволяет уменьшить расстояние между контактами до 1 мкм. Перспективным в изготовлении тонких пленок арсенида галлия является метод эпитаксиального наращивания.

При создании низкоомных омических контактов, необходимых для работы генераторов Ганна, существуют два подхода.

Первый из них заключается в поисках приемлемой технологии нанесения таких контактов непосредственно на высокоомный арсенид галлия.

Второй подход заключается в изготовлении многослойной конструкции генератора (типа «сэндвич»). В генераторах такой структуры на слой сравнительно высокоомного арсенида галлия, служащего рабочей частью генератора, наращивают с двух сторон эпитаксиальные слои относительно низкоомного арсенида галлия с электропроводностью n -типа. Эти высоколегированные слои служат переходными прослойками от рабочей части прибора к металлическим электродам.

Такие многослойные генераторы имеют два преимущества по сравнению с обычной конструкцией генератора Ганна. Во-первых, нанесение металлических контактов на сильнолегированный материал является сравнительно простой и хорошо освоенной в серийном производстве операцией. Во-вторых, при весьма тонкой рабочей части генератор может иметь сравнительно большую общую толщину, что упрощает обращение с ним при напайке внешних выводов, монтаже в СВЧ-патрон и т. д.

Проводились эксперименты по изготовлению генераторов Ганна на основе таких материалов, как теллурид кадмия, арсенид и фосфид индия и др. Однако техника получения этих материалов отработана хуже, чем техника получения арсенида галлия и эксперименты имели чисто научное значение. Арсенид галлия является пока единственным материалом, практически применяемым для изготовления генераторов Ганна.

§ 8.3. СВОЙСТВА И ПАРАМЕТРЫ ГЕНЕРАТОРОВ ГАННА

Как любой генератор СВЧ-диапазона, генератор Ганна характеризуется генерируемой мощностью (в импульсном и непрерывном режимах), частотным диапазоном, коэффициентом полезного действия, уровнем частотных и амплитудных шумов и возможностью частотной перестройки.

Мощности СВЧ-колебаний, генерируемых на генераторах Ганна в непрерывном режиме, достигают нескольких сотен милливольт, в импульсном режиме — нескольких сотен ватт.

Частота генерации генераторов Ганна зависит только от расстояния между электродами. Поэтому ограничения частотного диапазона генераторов Ганна определяются тем, в каких пределах можно практически изменить длину кристалла. Как отмечалось в § 8.2, существуют определенные трудности в создании генераторов Ганна с длиной более 100 и менее 2 мкм. В связи с этим диапазон частот таких генераторов ограничен частотами 1 и 30 ГГц.

Коэффициент полезного действия генераторов Ганна лабораторного изготовления достигает 24%.

В связи с возможным наличием в кристалле генератора Ганна нескольких неоднородностей зарождение домена, как отмечалось в § 8.2, может происходить в различные моменты времени на разном расстоянии от анода. Поэтому частота колебаний будет изменяться, т. е. могут возникать частотные шумы. Уровень частотных и амплитудных шумов генераторов Ганна довольно высок.

Важным для практического применения генераторов Ганна является вопрос о возможности их *частотной перестройки* в достаточно широком диапазоне. Из принципа действия генератора Ганна ясно, что частота его должна слабо зависеть от приложенного напряжения. С увеличением приложенного напряжения несколько возрастает толщина домена, а скорость его движения изменяется незначительно. В результате при изменении напряжения от порогового до пробивного частота колебаний увеличивается всего на десятые доли процента.

Срок службы генераторов Ганна относительно мал, что связано с одновременным воздействием на кристалл полупроводника таких факторов, как сильное электрическое поле и перегрев кристалла из-за выделяющейся в нем мощности.

§ 8.4. ГЕНЕРАТОРЫ С ОГРАНИЧЕНИЕМ НАКОПЛЕНИЯ ОБЪЕМНОГО ЗАРЯДА

Генератор с ограничением накопления объемного заряда — это генератор Ганна, работающий в режиме неполного формирования домена.

Предположим, что кроме постоянного смещения, большего $E_{пор}l$, где $E_{пор}$ — пороговая напряженность поля, к прибору приложено также синусоидальное напряжение достаточно большой амплитуды (рис. 8.5). В этом случае часть периода напряжение на приборе будет меньше порогового значения $E_{пор}l$.

Пока мгновенные значения напряжения в приборе больше, чем пороговое напряжение $E_{пор}l$, у катода формируется домен. Однако если прибор поместить в резонатор, настроенный на достаточно большую частоту, то домен, не успев сформироваться, начнет рассасываться, так как суммарное напряжение смещения и колебаний резонатора в следующую часть периода будет меньше, чем пороговое напряжение $E_{пор}l$. Во время формирования домена ток, проходящий через прибор, уменьшается, а во время рассасывания возрастает. Таким образом, возникают периодические колебания тока, но период этих колебаний определяется уже не пролетным време-

нем домена, а частотой резонатора. Такие генераторы называют генераторами с ограничением накопления объемного заряда (ОНОЗ). Этот режим работы генератора Ганна (в английской транскрипции — *LSA (Limited Spacecharge Accumulation)*) был обнаружен в 1966 г.

Для установления режима ОНОЗ необходимо выполнение ряда условий. Прежде всего нужно, чтобы домен не успел сформироваться за время, пока напряжение на генераторе превышает пороговую величину $E_{пор}$, период колебаний T должен быть меньше $(2-3) \tau'$, где $\tau' = \rho' \epsilon_0$ — максвелловское время релаксации для материала в состоянии с отрицательным дифференциальным сопротивлением.

За то время, пока мгновенные значения напряжения на генераторе меньше величины порогового напряжения $E_{пор}$, возникший у катода домен должен успеть полностью исчезнуть. Для выполнения этого условия необходимо, чтобы период колебаний был много больше максвелловского времени релаксации τ для материала, находящегося в слабом электрическом поле. Таким образом, для существования режима ОНОЗ необходимо выполнение условий:

$$T = \frac{1}{f} < (2 \div 3) \tau' = \frac{(2 \div 3) \epsilon_0}{q n_0 \mu'} \quad \text{и} \quad T \gg \tau = \rho \epsilon_0 = \frac{\epsilon_0}{q n_0 \mu_1}$$

или

$$\frac{\epsilon_0}{q \mu_1} \ll \frac{n_0}{f} < \frac{(2 \div 3) \epsilon_0}{q \mu'}$$

где μ_1 и μ' — подвижности электронов в слабом поле и при наличии отрицательного дифференциального сопротивления.

Одним из преимуществ генераторов с ОНОЗ перед генераторами Ганна является независимость частоты генерируемых колебаний от толщины кристалла. Другое, более важное их преимущество — возможность получения больших мощностей генерации на более высоких частотах. Генераторы с ОНОЗ могут быть изготовлены из относительно толстых пластинок высокоомного арсенида галлия, что позволяет подавать на такие приборы значительные напряжения.

С помощью генераторов с ОНОЗ можно достичь мощности непрерывных колебаний порядка десятых долей ватта в диапазоне 30—300 ГГц с к. п. д., равном 20%, и сотен ватт в импульсном режиме.

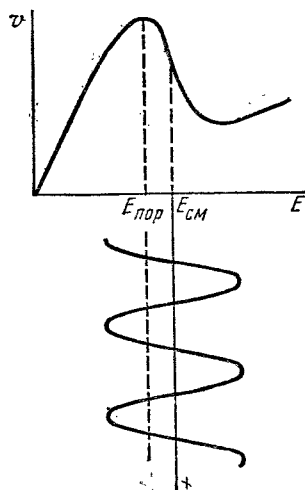
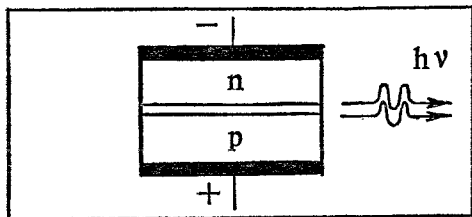


Рис. 8.5. Пояснение принципа действия генератора с ограничением накопления объемного заряда

ИЗЛУЧАЮЩИЕ ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ

9



§ 9.1. СВЕТОДИОДЫ

Принцип действия

Излучающий полупроводниковый прибор — это полупроводниковый прибор, предназначенный для непосредственного преобразования электрической (или световой) энергии в энергию светового излучения.

Полупроводниковый светодиод — это излучающий полупроводниковый прибор с одним или несколькими электрическими переходами, предназначенный для непосредственного преобразования электрической энергии в энергию некогерентного светового излучения.

Обычно излучение светодиода вызвано самопроизвольной рекомбинацией носителей заряда при их инжекции под действием прямого напряжения на электронно-дырочном переходе. Рекомбинация носителей может происходить как в прилегающих к p - n -переходу областях, так и в самом p - n -переходе. Чтобы кванты энергии — фотоны, освободившиеся при рекомбинации, соответствовали квантам видимого света, ширина запрещенной зоны исходного полупроводника должна быть относительно большой ($\Delta E > 1,7$ эВ). По этой причине рекомбинационная составляющая прямого тоска через такие диоды оказывается существенной по сравнению с инжекционной составляющей (см. § 3.10), т. е., по крайней мере, при малых прямых напряжениях на светодиоде процесс рекомбинации идет в основном в p - n -переходе. Иногда возбужденное состояние носителей заряда в светодиоде создают подачей обратного напряжения, при котором происходит ударная ионизация атомов полупроводника в p - n -переходе с последующей рекомбинацией неравновесных носителей заряда.

Исходными полупроводниковыми материалами для светодиодов являются арсенид галлия GaAs, карбид кремния SiC, фосфид галлия GaP или твердые растворы фосфида и арсенида галлия, которые образуют непрерывный ряд твердых растворов $\text{GaAs}_x\text{P}_{1-x}$, где x принимает значение от 0 до 1.

Перспективным материалом для изготовления светодиодов является нитрид галлия GaN, который имеет большую ширину запрещенной зоны ($\Delta E = 3,4$ эВ), и поэтому энергии квантов света, возникающих в этом материале при рекомбинации носителей заряда, могут перекрывать всю видимую область спектра.

Если бы рекомбинация электронов и дырок после создания тем или иным способом неравновесного состояния происходила только с излучением квантов света, то внутренний квантовый выход (отношение излученных фотонов к числу рекомбинировавших пар носителей) был бы близок к 100 %. Однако значительная часть процессов рекомбинации заканчивается выделением энергии в виде элементарных квантов тепловых колебаний кристаллической решетки — фононов. Такие переходы электронов между энергетическими уровнями называют безызлучательными. Соотношение между излучательными и безызлучательными переходами зависит от ряда причин, в частности от структуры энергетических зон полупроводника, наличия в кристалле примесей с большой эффективностью безызлучательных захватов электронов и дырок. Наилучшими с точки зрения величины внутреннего квантового выхода являются светодиоды из арсенида галлия, где получен внутренний квантовый выход, близкий к 100%. В светодиодах на основе других полупроводниковых материалов внутренний квантовый выход значительно ниже (иногда составляет единицы процентов), но и при таких значениях внутреннего квантового выхода излучение светодиода оказывается достаточным для практического использования.

Конструкция и технология изготовления

Даже при высоком внутреннем квантовом выходе внешний квантовый выход светодиодов оказывается значительно ниже, так как из-за высокого показателя преломления полупроводника большая часть квантов света испытывает полное внутреннее отражение на границе раздела полупроводника с окружающим воздухом. После отражения в полупроводнике может происходить поглощение квантов света. Через границу раздела проходит лишь та небольшая доля света, которая падает на границу раздела под углом, меньшим критического угла полного внутреннего отражения $\varphi_{кр} = \arcsin(1/n_r)$, где n_r — коэффициент преломления света в полупроводнике, равный, например, 3,3 и 3,6 для фосфида и арсенида галлия соответственно. Как правило, $\varphi_{кр} \leq 10-16^\circ$.

Таким образом, из светодиода, имеющего простейшую плоскую структуру (рис. 9.1, а), в окружающее пространство выходят только сотые доли всех квантов света, возникших в $p-n$ -переходе и вблизи него. Внешний квантовый выход удастся увеличить при использовании более сложных конструкций светодиодов со структурой в виде полусферического монокристалла полупроводника (рис. 9.1, б) или плоской структуры с прозрачным полусферическим покрытием (рис. 9.1, в).

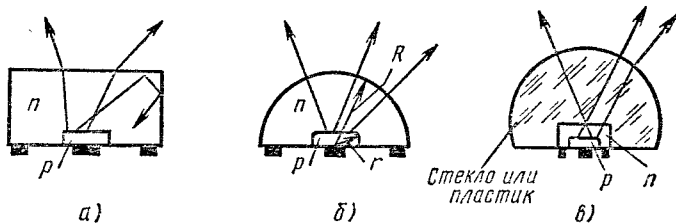


Рис. 9.1. Структура светодиодов:

а — плоская планарная; б — полусферическая; в — плоская с прозрачным полусферическим покрытием

В светодиодах с полусферической структурой для всей поверхности угол падения квантов света оказывается меньше критического угла полного внутреннего отражения при большом отношении радиусов R и r (рис. 9.1, б), т. е. при

$$R/r \geq n_r/n_{\text{ср}},$$

где n_r и $n_{\text{ср}}$ — коэффициенты преломления света в полупроводнике и среде, окружающей светодиод ($n_{\text{ср}} = 1$ для воздуха).

Однако в светодиодах с полусферической конструкцией несколько возрастают потери квантов света в результате поглощения в самом полупроводнике из-за увеличения длины пути квантов света от места возникновения до поверхности. Все же светодиоды с полусферической структурой имеют внешний квантовый выход на порядок выше по сравнению со светодиодами с плоской структурой.

Значительно проще технология изготовления светодиодов с прозрачным полусферическим покрытием из стекла или из различных пластических материалов с высоким показателем преломления для увеличения критического угла полного внутреннего отражения в полупроводнике.

Для повышения внешнего квантового выхода применяют различные просветляющие покрытия внешней поверхности светодиодов, т. е. прозрачные слои вещества толщиной в четверть длины волны с показателем преломления, равным $n_1^{1/2}$. Просветляющие покрытия увеличивают внешний квантовый выход примерно в 1,5 раза вне зависимости от структуры светодиода. Применение многослойных просветляющих покрытий позволяет повысить внешний квантовый выход, но усложняет технологию изготовления светодиодов.

В связи с большим коэффициентом поглощения арсенида галлия с электропроводностью p -типа по сравнению с коэффициентом поглощения арсенида галлия с электропроводностью n -типа вывод излучения из светодиодов на основе арсенида галлия делают обычно через область диода с электропроводностью n -типа.

Площадь излучающего p - n -перехода светодиода обычно не превышает 1 мм^2 . Она ограничивается параметрами источника питания, т. е. током, который может обеспечить источник питания (обычно $10 - 50 \text{ мА}$).

Светодиоды могут иметь один p - n -переход или несколько p - n -переходов (*матричные светодиоды*), расположенных на одном монокристалле (рис. 9.2). При соответствующем включении отдельных p - n -переходов матричного светодиода можно получить светящуюся цифру, букву или знак.

Еще одной разновидностью светодиодов является светодиод с управляемой геометрией светящегося поля (*управляемый светодиод*) (рис. 9.3, а). Область с электропроводностью n -типа такого светодиода низкоомна и поэтому является практически эквипотенциальной. Область с электропроводностью p -типа сравнительно высокоомна и поэтому при приложении внешних напряжений к электродам не будет эквипотенциальной. Распределение потенциала в p -области будет зависеть от величины напряжения, поданного на управляющий электрод (рис. 9.3, б). Соответственно от величины поданного на управляющий электрод напряжения будет зависеть и размер светящегося поля управляемого светодиода. Такие светодиоды могут быть использованы в качестве индикаторов настройки транзисторных приемников, для записи аналоговой информации на фотопленку, в качестве шкалы различных измерительных приборов и для других целей.

Основным методом формирования p - n -переходов при создании светодиодов из фосфида и арсенида галлия является метод эпитаксиального наращивания. Обычно это жидкофазная эпитаксия, иногда — эпитаксия из газовой фазы. Для формирования p - n -переходов в карбиде кремния используют метод диффузии примесей, а иногда методы планарной технологии, позволяющие получать p - n -переходы различной конфигурации.

Основные характеристики и параметры

Яркость излучения. Единицей яркости в системе СИ является кандела на квадратный метр ($\text{кд}/\text{м}^2$) — яркость источника излучения, каждый квадратный метр излучающей поверхности которого имеет в данном направлении силу света, равную одной канделе. Следует отметить, что световые измерения, строго говоря, не вполне объективны. Основным «прибором», с помощью которого можно

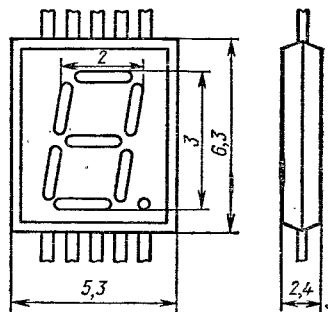


Рис. 9.2. Конструкция матричного светодиода — цифрового индикатора АЛ1304

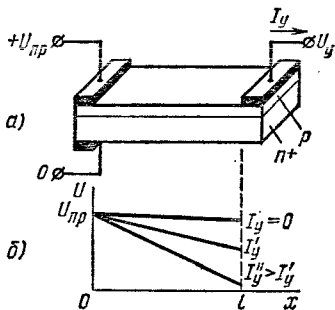


Рис. 9.3. Структура управляемого светодиода (а) и распределение потенциала в p -области (б)

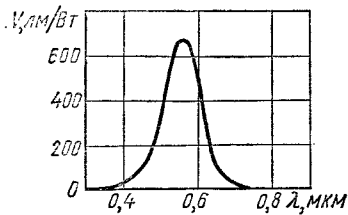


Рис. 9.4. Спектральная характеристика видности глаза человека

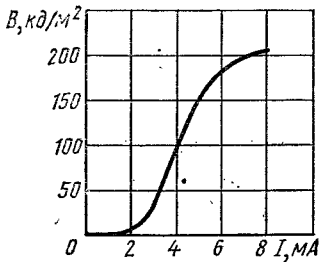


Рис. 9.5. Яркостная характеристика светодиода

измерять светотехнические величины, в конечном счете является глаз человека. Чувствительность глаза характеризуют специальной величиной, которая получила название видности. Видность V — это отношение светового потока Φ (т. е. оцениваемой нашим глазом мощности) к соответствующей истинной, полной мощности лучистой энергии Φ_3 :

$$V = \Phi / \Phi_3.$$

Единицей видности является люмен на ватт (лм/Вт), т. е. видность представляет собой величину, позволяющую переходить от световых величин к энергетическим.

Глаз человека по-разному реагирует на излучения разных длин волн. Спектральная чувствительность глаза или спектральная характеристика видности показана на рис. 9.4. Максимальная видность (при длине волны $\lambda = 554$ нм, лежащей в зеленой области спектра) составляет 683 лм/Вт.

Таким образом, излучающий прибор, который всю свою энергию отдает только в виде излучения с длиной волны 554 нм, обладает наибольшей яркостью и экономичностью с точки зрения глаза человека при использовании светодиодов для визуального отображения информации. Однако часто светодиоды используют для записи информации на фоточувствительные материалы, для передачи информации в виде светового сигнала к какому-нибудь приемнику излучения и для других целей. В этих случаях яркость излучения не является основным параметром светодиода, потому что фоточувствительный материал или приемник излучения могут иметь свои спектральные характеристики чувствительности, существенно отличающиеся от спектральной характеристики видности глаза человека.

Яркостная характеристика. Светодиод, p - n -переход которого включен в прямом направлении, обладает относительно малым сопротивлением. Поэтому светодиоды следует считать токовыми приборами, питаемыми от генераторов тока. Соответственно яркостной характеристикой светодиода является зависимость яркости от проходящего через светодиод тока (рис. 9.5).

Вид яркостной характеристики определяется структурой p - n -перехода и зависит от слоя или места, в котором происходит преимущественная рекомбинация носителей заряда (p - n -переход, сильнолегированная или слаболегированная области, прилегающие к p - n -переходу). Все эти слои содержат различные примеси и в разных

концентрация и не могут быть равноценны с точки зрения излучательной рекомбинации. Трудно создать структуру p - n -перехода, в которой рекомбинация носителей заряда происходила бы в слое, оптимальном с точки зрения излучательной рекомбинации или с точки зрения квантового выхода электролюминесценции в широком диапазоне изменения плотности тока.

При малых токах и соответственно при малых напряжениях не светодиоде излучения принципиально быть не должно вплоть до напряжения на светодиоде, численно равного энергии излучаемого кванта света. Поэтому начальный участок яркостной характеристики нелинеен.

При больших токах яркостная характеристика сублинейна, так как при больших плотностях тока в слое, где происходит рекомбинация носителей заряда, будет существенное отклонение от равновесного состояния. Поэтому положение квазиуровней Ферми для электронов и для дырок будет отличаться от положения уровня Ферми в условиях термодинамического равновесия. С увеличением отклонения от равновесного термодинамического состояния аналогично смещаются и демаркационные уровни $\mathcal{E}_{\text{дем}n}$ и $\mathcal{E}_{\text{дем}p}$. Таким образом, число примесных уровней ловушек захвата уменьшается, а число рекомбинационных уровней увеличивается. Теперь часть носителей может рекомбинировать без излучения квантов света или с излучением фотонов с энергией, не соответствующей энергии квантов видимого света.

Оптимальный режим работы светодиода определяется видом его яркостной характеристики, и, наоборот, при конструировании светодиода определенного назначения необходимо создать определенную структуру p - n -перехода для обеспечения наибольшего квантового выхода излучения в соответствующем диапазоне плотностей тока.

Нижний и верхний пределы диапазона рабочих напряжений. Нижний предел рабочего напряжения светодиода или пороговое напряжение светодиода определяется энергией излучаемых квантов света. Численно значение этого параметра в вольтах должно быть несколько выше энергии излучаемых квантов света в электронвольтах. Верхний предел рабочего напряжения определяется допустимой мощностью рассеяния светодиода. Его значение зависит от величины контактной разности потенциала p - n -перехода светодиода и от сопротивления базы светодиода. В связи с тем, что светодиод является прибором, управляемым током, правильнее было бы характеризовать этот прибор диапазоном рабочих токов. Однако диапазон рабочих напряжений или его верхний предел является важным параметром светодиода с точки зрения выбора источника питания для этого прибора.

Спектральная характеристика светодиода — зависимость интенсивности излучения от длины волны излучаемого света или от энергии излучаемых квантов. В первом приближении спектральный состав излучения можно характеризовать цветом свечения (рис. 9.6). Но более подробные сведения дает спектральная характеристика светодиода.

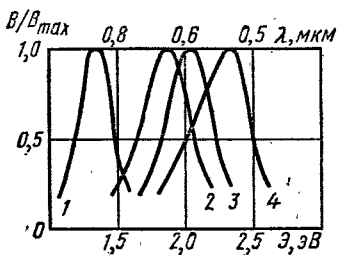
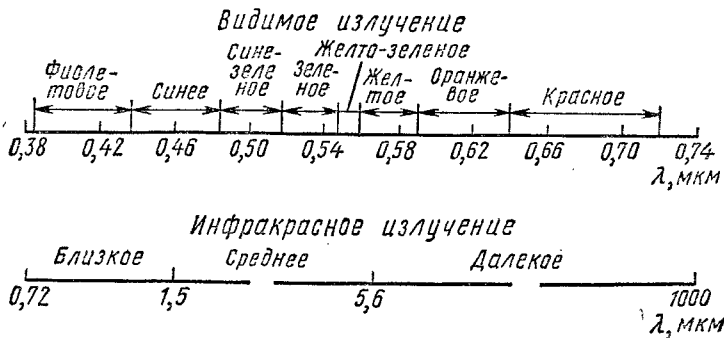


Рис. 9.6. Диапазоны длин волн видимого и инфракрасного излучения

Рис. 9.7. Спектральные характеристики светодиодов на основе фосфида галлия (1, 3) и карбида кремния (2, 4) с различными примесями

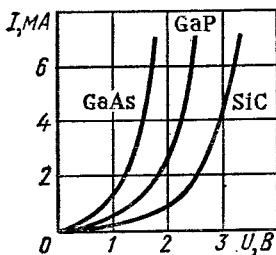


Рис. 9.8. Вольт-амперные характеристики светодиодов, изготовленных на основе различных полупроводниковых материалов

Длина волны излучаемого света определяется разностью энергий двух энергетических уровней, между которыми происходит переход электронов на излучательном этапе процесса рекомбинации. В связи с разной шириной запрещенной зоны у различных полупроводниковых материалов и разной глубиной залегания в запрещенной зоне энергетических уровней различных примесей, обеспечивающих излучательную рекомбинацию носителей заряда, длина волны излучаемого света различна в разных светодиодах. Так, максимум спектральной характеристики светодиодов, изготовленных на основе монокристаллов твердых растворов фосфида и арсенида галлия $\text{GaAs}_x\text{P}_{1-x}$, лежит в пределах от 565 нм (желтое излучение) для светодиодов из фосфида галлия до 920 нм (ближкое инфракрасное излучение) для светодиодов из арсенида галлия.

Таким образом, светодиоды из арсенида галлия не пригодны для визуальной индикации, но могут быть использованы для других

целей. Различные модификации карбида кремния имеют ширину запрещенной зоны от 2,4 до 3,3 эВ. Эффективную излучательную рекомбинацию носителей заряда в карбиде кремния обеспечивают примеси с разной глубиной залегания в запрещенной зоне (азот, алюминий, бор, бериллий, скандий, галлий). Поэтому максимум спектральной характеристики излучения светодиодов из карбида кремния может лежать в пределах, соответствующих практически всему диапазону видимого излучения.

Примеры спектральных характеристик светодиодов на основе различных материалов приведены на рис. 9.7. Так как переход электронов при рекомбинации носителей заряда обычно происходит не между двумя определенными энергетическими уровнями, а между двумя группами близко расположенных энергетических уровней, то спектр излучения оказывается размытым. Размытие спектра излучения определяют полушириной спектральной характеристики, измеряемой в нанометрах или микрометрах на высоте 0,5 от максимума.

Параметры светодиода как элемента электрической схемы определяются его *вольт-амперной характеристикой*. Различия прямых ветвей вольт-амперных характеристик светодиодов из разных полупроводниковых материалов связаны с различной шириной запрещенной зоны (рис. 9.8). Обратные ветви вольт-амперных характеристик соответствуют относительно малым пробивным напряжениям, что объясняется малой толщиной *p-n*-переходов в светодиодах, в основном не предназначенных для включения в обратном направлении.

Инерционность светодиодов характеризуется постоянными времени разгорания и затухания электролюминесценции при импульсном возбуждении светодиодов, которые принято измерять между яркостями излучения 0,1 и 0,9 от максимальной. Как и частотные свойства выпрямительных диодов, инерционность светодиодов может определяться временем перезаряда барьерной емкости *p-n*-перехода и процессами накопления и рассасывания неосновных носителей заряда. В связи с малыми временами жизни неосновных носителей, характерных для исходных материалов светодиодов, инерционность их оказывается достаточно малой: постоянные времени затухания электролюминесценции не превышают долей микросекунды. Таким образом, постоянные времени, характеризующие инерционность светодиодов, оказываются несущественными параметрами для светодиодов, предназначенных для визуальной индикации, так как инерционность глаза человека достаточно велика (около 50 мс). Для систем записи и считывания информации без визуализации постоянные времени могут быть, наоборот, одними из наиболее существенных параметров светодиода.

Коэффициент полезного действия, или эффективность, светодиода зависит от внутреннего квантового выхода и конструкции светодиода. Обычно к. п. д. светодиода не превышает сотых долей, т. е. оказывается либо меньше, либо сравним с к. п. д. ламп накаливания. Основные потери световой энергии происходят в резуль-

тате полного внутреннего отражения и из-за поглощения квантов света в полупроводниковом материале светодиода, на контактах и других частях конструкции светодиода.

Срок службы и надежность. Явления электролюминесценции, обуславливающие излучение, связаны с электронными процессами, происходящими в объеме монокристалла полупроводника. На срок службы, определяемый изменением параметров светодиода со временем, могли бы влиять эффекты переноса ионов внутри кристалла. Однако в таком материале, как карбид кремния, существуют очень прочные связи между атомами, и при обычных температурах он слабо взаимодействует с окружающей средой. Поэтому срок службы светодиодов из карбида кремния практически не имеет принципиальных ограничений.

Яркость излучения светодиодов из фосфида галлия несколько убывает с течением времени, однако срок службы их оценивается десятками тысяч часов.

Таким образом, отличительными свойствами светодиодов являются их высокая надежность, большой срок службы, малые инерционность, габариты, масса, потребляемая мощность, возможность изготовления светодиодных матриц и светодиодов с различным цветом свечения, совместимость с интегральными микросхемами в связи с низкими рабочими напряжениями.

§ 9.2. ЛАЗЕРЫ

Принцип действия

Полупроводниковый лазер — это излучающий полупроводниковый прибор, предназначенный для непосредственного преобразования электрической энергии или энергии некогерентного излучения в энергию когерентного излучения.

В полупроводниковых лазерах или в оптических квантовых генераторах (полупроводниковых ОКГ) излучение, так же как и в светодиодах, порождается рекомбинацией электронов и дырок. Однако эта рекомбинация в лазерах оказывается в основном не самопроизвольной, как в светодиодах, а вынужденной (индуцированной или стимулированной). Именно поэтому источники вынужденного излучения называли лазерами*. Излучение при вынужденной рекомбинации получается когерентным (см. § 1.11), что является принципиальным отличием полупроводниковых лазеров от светодиодов. Явление вынужденной рекомбинации дает возможность управлять излучением возбужденных атомов полупроводника с помощью электромагнитных волн и таким образом усиливать и генерировать когерентный свет.

Для работы лазера необходимо преобладание вынужденной излучательной рекомбинации над поглощением квантов света. Преобла-

* *Laser — lightwave amplification by stimulated emission of radiation* (усиление световых волн с помощью стимулированного излучения).

дание излучения над поглощением или поглощения над излучением зависит от соотношения в кристалле полупроводника возбужденных и невозбужденных атомов, т. е. от населенности энергетических уровней полупроводника. В равновесных условиях на более высоких энергетических уровнях при любой температуре полупроводника число электронов меньше, чем на более низких энергетических уровнях. При этом нельзя получить усиления света в результате вынужденной рекомбинации. Для преобладания вынужденной рекомбинации над поглощением квантов света необходимо, чтобы верхние энергетические уровни были более заполнены электронами по сравнению с нижними. Состояние полупроводника, в котором число электронов на одном из энергетических уровней с большей энергией больше числа электронов на уровне с меньшей энергией, называют состоянием с инверсной населенностью. Поглощение квантов света в полупроводнике с инверсной населенностью энергетических уровней будет мало, так как около потолка валентной зоны почти нет электронов, которым может быть передана энергия кванта света. С другой стороны, в полупроводнике с инверсной населенностью может происходить вынужденная рекомбинация.

Инверсную населенность в полупроводнике можно создать различными способами: 1) с помощью инжекции носителей заряда при прямом включении *p-n*-перехода, что используют в так называемых инжекционных лазерах; 2) путем электронного возбуждения, т. е. путем бомбардировки полупроводника пучком быстрых электронов; 3) с помощью оптической накачки, т. е. путем возбуждения атомов полупроводника квантами света от мощного излучателя некогерентного или когерентного света; 4) путем использования эффектов сильного электрического поля, т. е. лавинного умножения носителей заряда или туннелирования электронов при их переходе с энергетических уровней, расположенных вблизи потолка валентной зоны, на энергетические уровни, расположенные вблизи дна зоны проводимости.

Наибольший практический интерес представляет первый из перечисленных способов создания инверсной населенности. Поэтому рассмотрим инжекционные лазеры.

Конструкция и технология изготовления инжекционных лазеров

Инверсную населенность в инжекционном лазере с *p-n*-переходом проще получить, если одна из областей диодной структуры является вырожденной, т. е. содержит большую концентрацию соответствующих примесей. При прямом включении *p-n*-перехода прямой ток состоит из двух составляющих — электронной и дырочной. Чем больший ток проходит через *p-n*-переход, тем с большим запасом выполняется условие инверсной населенности. Минимальный ток, при котором происходит преимущественно вынужденная рекомбинация, называют пороговым током.

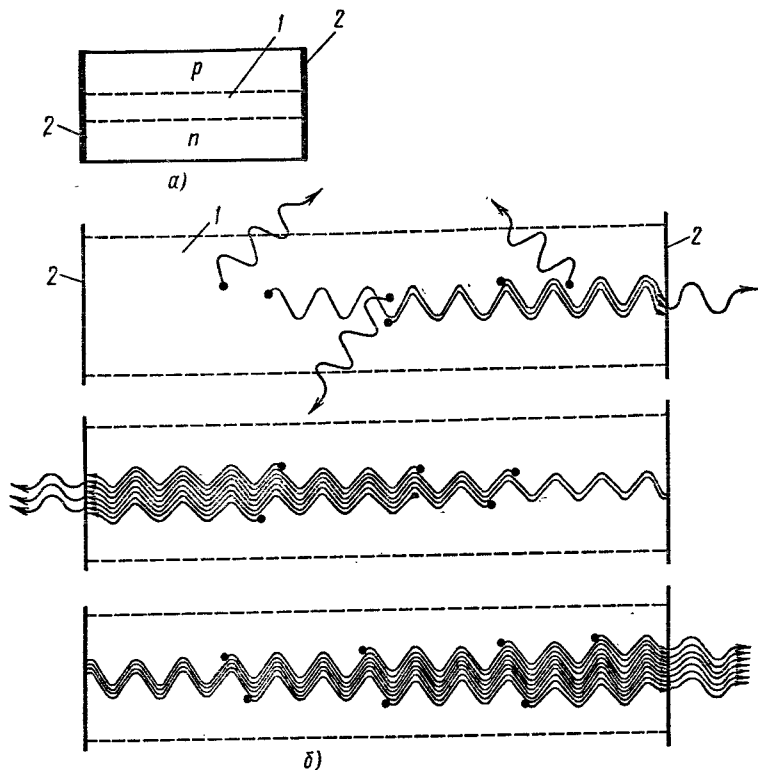


Рис. 9.9. Структура полупроводникового инжекционного лазера (а) и схема процесса образования фотонной лавины в оптическом резонаторе лазера (б):

1 — активная область с инверсной населенностью; 2 — отражающие поверхности кристалла полупроводника

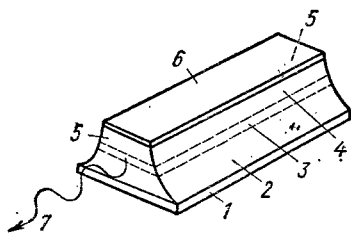


Рис. 9.10. Конструкция полупроводникового инжекционного лазера:

1 — молибденовая пластина — нижний электрод; 2 — область электропроводности n -типа; 3 — активная область с инверсной населенностью; 4 — область электропроводности p -типа; 5 — полированные торцевые поверхности кристалла полупроводника; 6 — верхний электрод; 7 — излучение

Если ток, проходящий через p - n -переход, больше порогового, то p - n -переход является усиливающей средой для света, распространяющегося в плоскости p - n -перехода. Число актов вынужденной рекомбинации можно увеличить, если обеспечить прохождение каждого кванта света несколько раз в плоскости p - n -перехода. Для этого две противоположные грани монокристалла полупроводника делают строго параллельными и тщательно отполированными.

Чтобы обеспечить необходимый коэффициент отражения от торцов, можно их не металлизировать, так как большой коэффициент преломления полупроводникового материала обеспечивает отражение от этих торцов до 35 % квантов света. После многократного отражения от полированных торцов и соответствующего многократного прохождения вдоль p - n -перехода свет выходит из полупроводника (рис. 9.9). Кванты света, двигающиеся строго перпендикулярно торцам кристалла, могут много раз пройти через активную область с инверсной населенностью и тем самым создать большую лавину квантов света. Две другие боковые грани должны быть скошены под некоторым углом, чтобы воспрепятствовать возникновению генерации света между ними (рис. 9.10). Те кванты света, которые начали двигаться не вдоль p - n -перехода и не перпендикулярно торцам кристалла, уходят из активной области с инверсной населенностью и не вызывают вынужденной рекомбинации.

Для изготовления инжекционных лазеров используют арсенид галлия, твердые растворы арсенид-фосфид галлия $\text{GaAs}_x\text{P}_{1-x}$, арсенид индия, фосфид индия и другие полупроводниковые материалы. Наибольшее распространение получили инжекционные лазеры из арсенида галлия. В качестве исходной заготовки для таких лазеров служит монокристалл арсенида галлия, по форме приближающийся к кубу или параллелепипеду, со сторонами длиной в несколько десятых долей миллиметра.

В арсениде галлия, легированном донорами (Te, Se и др.), p - n -переход обычно создают путем диффузии акцепторов (Zn, Cd и др.). Области с электропроводностью p - и n - типа должны иметь концентрации соответствующих примесей, при которых энергетические состояния электронов и дырок близки к вырождению. Для создания невыпрямляющего контакта с n -областью монокристалл с диодной структурой припаивают к пластинке молибдена, покрытой слоем золота (см. рис. 9.10). На поверхность p -области наносят слой сплава золота с серебром.

Инверсную населенность значительно проще можно создать в полупроводниковом инжекционном лазере с гетеропереходами (рис. 9.11). Базовую область такой структуры делают из полупроводника с меньшей шириной запрещенной зоны и большей диэлектрической проницаемостью, чем у эмиттерных областей. Инжектированные в базу носители заряда оказываются в потенциальных ямах. Различие показателей преломления (диэлектрических проницаемостей) базовой и эмиттерных областей приводит к полному внутрен-

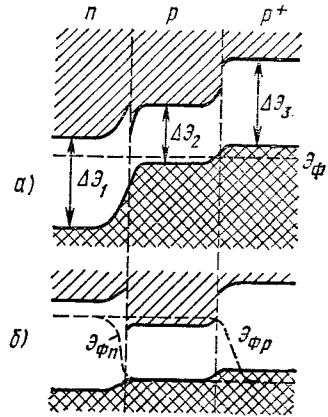


Рис. 9.11. Энергетическая диаграмма полупроводникового инжекционного лазера с гетеропереходами без приложенного напряжения (а) и при прямом напряжении (б)

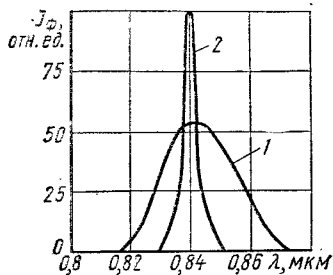


Рис. 9.12. Спектральные характеристики излучающего диода на основе арсенида галлия при температуре 77 К:

1 — при токе ниже порогового значения (режим работы светодиода); 2 — при токе выше порогового значения (режим работы инжекционного лазера)

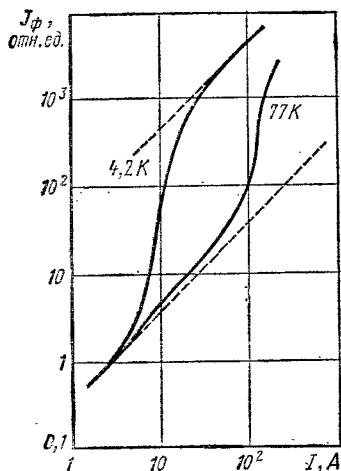


Рис. 9.13. Ярковая характеристика полупроводникового излучающего диода на основе арсенида галлия при различных температурах

чения значительно больше, так как излучение получается когерентным и строго направленным.

Диаграмма направленности излучения лазера характеризует пространственное распределение интенсивности излучения. Излучение полупроводниковых лазеров обладает достаточно малым (как правило, не превышающим нескольких градусов) углом расходимости светового пучка. Но по этому параметру полупроводниковые лазеры значительно уступают газовым и твердотельным диэлектрическим

нему отражению квантов света на гетеропереходах, т. е. область базы является по существу световодом. Все это обеспечивает значительно меньшие пороговые плотности токов и большие эффективности или к. п. д. инжекционных лазеров с гетеропереходами.

Основные характеристики и параметры

Пороговая плотность тока существенно зависит от температуры инжекционного лазера: для лазеров на основе арсенида галлия пороговая плотность тока порядка 10^2 А/см² при температуре 4,2 К и порядка 10^4 А/см² при 77 К. Таким образом, для уменьшения пороговой плотности тока необходимо глубокое охлаждение инжекционного лазера. Инжекционные лазеры с использованием гетеропереходов, имеющие значительно меньшие пороговые плотности токов, могут работать при комнатной температуре в непрерывном режиме.

Спектральная характеристика лазера, как и любого другого источника света, представляет собой зависимость интенсивности излучения (чаще в относительных единицах) от длины волны (рис. 9.12). При малых токах (меньше порогового) излучение, возникающее в основном из-за самопроизвольной рекомбинации, является некогерентным. Поэтому спектральная характеристика получается широкой, т. е. лазер работает как светодиод. При больших токах (больше порогового) интенсивность излу-

лазерам, что связано с малыми размерами кристалла полупроводника и особенно с малыми размерами активной области, где происходит вынужденная рекомбинация.

Коэффициент полезного действия полупроводникового инжекционного лазера на основе арсенида галлия достигает 70 %, в то время как величина внутреннего квантового выхода, возможно, достигает 100 %, т. е. каждый инжектированный электрон при рекомбинации с дыркой создает фотон. По к. п. д. полупроводниковые инжекционные лазеры превосходят газовые и твердотельные диэлектрические лазеры, у которых он равен соответственно тысячным и сотым долям процента.

Яркостная характеристика лазера, т. е. зависимость интенсивности излучения от проходящего через лазер тока (рис. 9.13) представляет собой почти линейные зависимости в диапазоне токов, соответствующих преобладанию самопроизвольной рекомбинации (режим работы светодиода) и преобладанию вынужденной рекомбинации (режим работы лазера).

§ 9.3. ЭЛЕКТРОЛЮМИНЕСЦЕНТНЫЕ ПОРОШКОВЫЕ ИЗЛУЧАТЕЛИ

Принцип действия

Электролюминесцентные порошковые излучатели — это одна из разновидностей излучающих полупроводниковых приборов, в которых используется электролюминесценция электролюминофора (см. § 1.11). Электролюминесцентный порошковый излучатель представляет собой многослойную систему, состоящую из стеклянной подложки, на которую последовательно наносят проводящий прозрачный электрод из окислов различных металлов (SnO_2 , In_2O_3 , CdO и др.), слой электролюминофора, защитный диэлектрический слой в виде лакового покрытия или тонкого слоя двуокиси, либо окиси кремния (SiO_2 , SiO) и второго электрода (рис. 9.14).

Одним из наиболее распространенных электролюминофоров является сульфид цинка (ZnS), активированный для получения яркого свечения примесями меди, марганца и других элементов. Зерна или поликристаллы сульфида цинка скрепляются между собой диэлектрической связкой. Так как отдельные зерна порошка сульфида цинка разделены диэлектрической связующей прослойкой, электролюминесцентные конденсаторы могут работать только при переменном напряжении, т. е. возбуждение электролюминофора происходит под действием электрического поля.

Под действием приложенного напряжения происходит ионизация примесных атомов электролюминофора либо в результате туннелирования электронов с примесных уровней в зону проводимости (электростатическая ионизация), либо в результате ударной ионизации в сильном электрическом поле в обедненных поверхностных слоях зерен сульфида цинка.

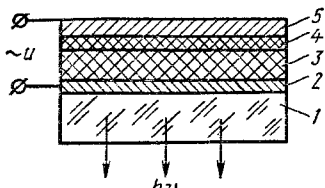


Рис. 9.14. Структура электролюминесцентного порошкового излучателя (конденсатора):

1 — стеклянная подложка; 2 — прозрачный электрод; 3 — слой электролюминофора; 4 — защитный слой; 5 — металлический электрод

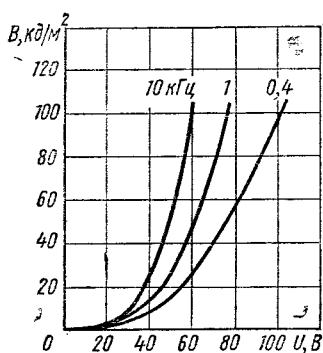


Рис. 9.15. Яркостная характеристика электролюминесцентного порошкового излучателя

После возбуждения поверхностных слоев зерен электролюминофора происходит процесс высвечивания электролюминофора — рекомбинация носителей заряда с выделением избыточной энергии в виде квантов света. Наряду с излучательной рекомбинацией происходит и безызлучательная рекомбинация, при которой избыточная энергия выделяется в виде квантов тепловой энергии.

Цвет излучения определяется шириной запрещенной зоны электролюминофора и глубиной залегающих энергетических уровней рекомбинационных ловушек в запрещенной зоне. Длительность процесса высвечивания (послесвечение) зависит от времени жизни неосновных носителей заряда и от наличия в электролюминофоре ловушек захвата, которые могут существенно увеличить эффективное время жизни носителей.

Основные характеристики и параметры

Одним из важнейших параметров электролюминесцентного порошкового излучателя является *эффективная*

яркость при определенной частоте переменного напряжения и при определенном значении этого напряжения или плотности тока.

Эффективная яркость электролюминесцентных порошковых излучателей зависит от приложенного напряжения (рис. 9.15). *Яркостная характеристика* нелинейна, так как процесс увеличения избыточной концентрации носителей заряда при ударной ионизации и при туннельном эффекте характеризуется степенными зависимостями от напряжения (или от напряженности электрического поля).

Большая нелинейность яркостных характеристик оказывается полезной при создании электролюминесцентных матричных экранов и преобразователей изображения, так как дает возможность получить большую контрастность изображения и большую разрешающую способность.

Крутизну яркостной характеристики иногда оценивают *кратностью изменения яркости* при уменьшении напряжения на электролюминесцентном излучателе в два раза от номинального значения. Кратность изменения яркости электролюминесцентных порошковых излучателей не превышает 25.

Зависимость эффективной яркости от частоты переменного напряжения (см. рис. 9.15) объясняется увеличением числа волн яркости в единицу времени при увеличении частоты.

Спектр излучаемого электролюминесцентным порошковым излучателем света характеризуют длиной волны, соответствующей максимуму спектральной характеристики излучения. Эта длина волны зависит от разности энергий уровней, между которыми происходит переход электронов при излучательной рекомбинации.

§ 9.4. ЭЛЕКТРОЛЮМИНЕСЦЕНТНЫЕ ПЛЕНОЧНЫЕ ИЗЛУЧАТЕЛИ

Принцип действия

Электролюминесцентные пленочные излучатели отличаются от электролюминесцентных порошковых излучателей тем, что между двумя электродами в них находится однородная поликристаллическая пленка электролюминофора, созданная термическим испарением с последующим осаждением в вакууме. Так как в электролюминесцентных пленочных излучателях отсутствует диэлектрическая связка в электролюминофоре, они могут работать и на постоянном токе.

Возбуждение электролюминофора, т. е. создание неравновесного состояния поверхностных слоев отдельных кристаллов электролюминофора, происходит из-за инжекции носителей заряда через потенциальные барьеры на поверхности отдельных, контактирующих между собой кристаллов электролюминофора и из-за инжекции из электродов. При рекомбинации инжектированных носителей избыточная энергия может выделяться в виде квантов света. Возбуждение электролюминофора может происходить также из-за эффектов сильного поля (туннелирование и ударная ионизация) в обедненных поверхностных слоях кристаллов электролюминофора.

Основные характеристики и параметры

Толщина пленки электролюминофора в электролюминесцентных пленочных излучателях мала. Поэтому рабочие напряжения таких излучателей (25—30 В) значительно меньше *рабочих напряжений* электролюминесцентных порошковых излучателей, сделанных из порошкообразного электролюминофора с диэлектрической связкой.

Яркостная характеристика электролюминесцентных пленочных излучателей, т. е. зависимость яркости свечения от приложенного напряжения нелинейна, так как нелинейны вольт-амперные характеристики этих приборов. В связи с большей нелинейностью яркостных характеристик электролюминесцентных пленочных излучателей по сравнению с нелинейностью аналогичных характеристик электролюминесцентных порошковых излучателей пленочные излучатели обладают большей разрешающей способностью, которая ограничи-

вается практически размерами отдельных кристаллов электролюминофора (10^{-5} — 10^{-4} мм). Разрешающая способность пленочных излучателей велика еще и потому, что тонкие пленки практически не рассеивают свет. *Кратность изменения яркости* электролюминесцентных пленочных излучателей достигает 1000.

Электролюминесцентные пленочные излучатели уступают электролюминесцентным порошковым излучателям по *экономичности* и *сроку службы*. Низкий срок службы, характерный для большинства полупроводниковых поликристаллических приборов, связан с одновременным действием больших напряженностей электрического поля и повышенной температурой в местах точечных контактов между отдельными кристаллами полупроводника.

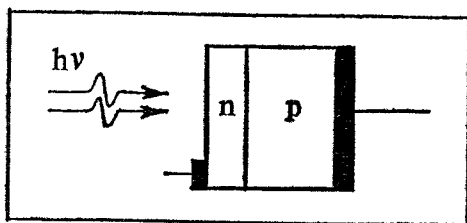
Другим недостатком электролюминесцентных пленочных излучателей, как и электролюминесцентных порошковых излучателей, является большой *разброс параметров*. Этот недостаток также характерен для всех поликристаллических полупроводниковых приборов.

Контрольные вопросы

1. Каким образом в светодиоде происходит непосредственное преобразование электрической энергии в энергию световую?
2. Как должна быть выполнена структура и конструкция светодиода для получения наибольшего внешнего квантового выхода?
3. Какими параметрами можно характеризовать различные свойства светодиодов?
4. Каков принцип действия полупроводникового лазера?
5. Каковы отличия в принципе действия и в свойствах полупроводникового лазера и светодиода?
6. Перечислите сходства и отличия в принципе действия и в свойствах электролюминесцентных порошковых излучателей и электролюминесцентных пленочных излучателей.

ПОЛУПРОВОДНИКОВЫЕ ПРИЕМНИКИ ИЗЛУЧЕНИЯ

10



§ 10.1. ФОТОРЕЗИСТОРЫ

Фоторезистор — это полупроводниковый резистор, сопротивление которого зависит от освещенности.

Таким образом, принцип действия фоторезистора основан на фоторезистивном эффекте (см. § 1.11).

Технология изготовления и конструкция

Основной частью конструкции фоторезистора является полупроводниковый фоточувствительный слой полупроводника, который может быть выполнен в виде монокристаллической или поликристаллической пластинки полупроводника или в виде поликристаллической пленки полупроводника, нанесенной на диэлектрическую подложку. В качестве полупроводникового материала для фоторезисторов обычно используют сульфид кадмия, селенид кадмия или сульфид свинца. На поверхность фоточувствительного слоя наносят металлические электроды. Иногда электроды наносят непосредственно на диэлектрическую подложку перед осаждением полупроводникового слоя.

Поверхность полупроводникового фоточувствительного слоя, расположенную между электродами, называют рабочей площадью. Фоторезисторы делают с рабочими площадками прямоугольной формы, в виде меандра или в виде кольца. Площадь рабочих площадок различных фоторезисторов составляет обычно от десятых долей до десятков квадратных миллиметров. Исходя из площади рабочей площадки можно правильно выбрать размер светового пучка, оценить световой поток, при котором должен работать фоторезистор, и т. д. При эксплуатации фоторезистора рекомендуется его рабочую площадку засвечивать полностью, так как при этом эффект изменения сопротивления фоторезистора будет максимален.

Подложку с нанесенным на нее полупроводниковым фоточувствительным слоем или пластинку полупроводника помещают в пластмассовый или металлический корпус.

Основные характеристики и параметры

Вольт-амперные характеристики фоторезистора представляют собой зависимости светового тока $I_{\text{св}}$ при неизменной величине светового потока, а также темнового тока $I_{\text{тем}}$ от приложенного к фоторезистору напряжения (рис. 10.1). В рабочем диапазоне напряжений вольт-амперные характеристики фоторезисторов при различных значениях светового потока практически линейны. Однако у большинства пленочных фоторезисторов и у фоторезисторов с фоточувствительным слоем из поликристаллического полупроводникового материала линейность вольт-амперной характеристики нарушается при малых напряжениях (характеристика является сверхлинейной). Эта нелинейность связана с явлениями на контактах между отдельными зернами или кристаллами полупроводника. При малых напряжениях сопротивление фоторезистора определяется в основном сопротивлением этих контактов. Напряжение, приложенное к фоторезистору, падает в основном на контактах между зернами полупроводника. Поэтому напряженность электрического поля на контактах получается большой даже при малых напряжениях на фоторезисторе. В связи с этим при увеличении приложенного напряжения сопротивление контактов уменьшается либо из-за эффектов сильного поля (например, туннелирование сквозь тонкие потенциальные барьеры на контактах), либо из-за разогрева приконтактных областей отдельных зерен полупроводника. При дальнейшем увеличении напряжения сопротивление фоторезистора будет определяться уже объемным сопротивлением зерен полупроводника и поэтому будет оставаться постоянным, что соответствует линейному участку вольт-амперной характеристики.

При больших напряжениях на фоторезисторе вольт-амперная характеристика опять может отклоняться от линейной, становясь или сверхлинейной или сублинейной. Сверхлинейность связана с повышением температуры всего фоточувствительного слоя из-за большой выделяющейся мощности. Сублинейность может наблюдаться у фоторезисторов, изготовленных из монокристаллического полупроводника и имеющих малое расстояние между электродами по сравнению с дрейфовым сдвигом носителей заряда. При этом носители, возникшие в фоточувствительном слое, за время жизни успевают перейти в металлический электрод и перестают участвовать, таким образом, в фоторезистивном эффекте.

Световая или локс-амперная характеристика фоторезистора представляет собой зависимость фототока $I_{\text{ф}} = I_{\text{св}} - I_{\text{тем}}$ от освещенности, или от падающего на фоторезистор светового потока. Фоторезисторы имеют обычно сублинейную световую характеристику (рис. 10.2). Сублинейность световой характеристики объясняется смещением демаркационных уровней, или квазиуровней

Ферми, для электронов и для дырок с увеличением отклонения от равновесного состояния при увеличении освещенности: электронный демаркационный уровень (квазиуровень Ферми для электронов) смещается к зоне проводимости в результате увеличения концентрации свободных электронов, дырочный демаркационный уровень (квазиуровень Ферми для дырок) одновременно смещается к валентной зоне из-за увеличения концентрации дырок (см. рис. 1.8). Вследствие смещения демаркационных уровней часть уровней ловушек захвата становится уровнями рекомбинационных ловушек. С ростом концентрации рекомбинационных ловушек уменьшается время жизни носителей заряда, что и является первой причиной сублинейности световой характеристики. Закономерности возрастания фототока от освещенности различны у различных фоторезисторов и определяются концентрацией тех или иных примесей в полупроводнике и распределением примесных уровней по запрещенной зоне энергетической диаграммы полупроводника.

Второй причиной, приводящей к сублинейности световой характеристики фоторезистора, является уменьшение подвижности носителей заряда при увеличении освещенности из-за увеличения концентрации ионизированных атомов в полупроводнике и, следовательно, из-за увеличения рассеяния носителей заряда ионизированными атомами.

В узком диапазоне освещенностей для аппроксимации световой характеристики часто используют зависимость

$$I_{\phi} = AE^x,$$

где A и x — коэффициенты, являющиеся постоянными для данного фоторезистора в выбранном диапазоне освещенностей; E — освещенность.

Спектральная характеристика фоторезистора — это зависимость фототока от длины волны падающего на фоторезистор света (рис. 10.3). При больших длинах волн, т. е. при малых энергиях квантов света по сравнению с шириной запрещенной зоны полупроводника энергия кванта оказывается недостаточной для переброса электрона из валентной зоны в зону проводимости. Поэтому

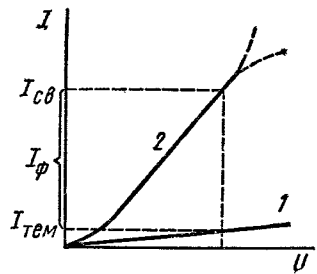


Рис. 10.1. Вольт-амперная характеристика фоторезистора:

1 — в темноте; 2 — при освещении

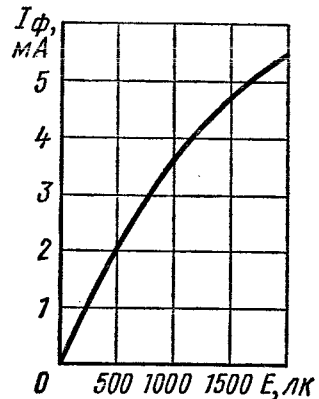


Рис. 10.2. Световая характеристика фоторезистора ФСК-1

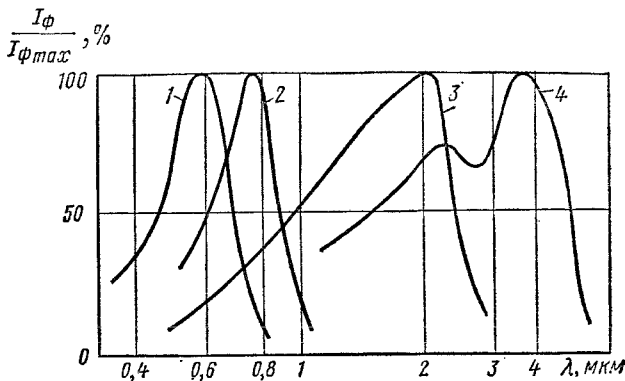


Рис. 10.3. Усредненные спектральные характеристики различных фоторезисторов:

1 — ФСК; 2 — ФСД; 3 — ФСА; 4 — СФ4

для каждого полупроводника и соответственно для каждого фоторезистора существует пороговая длина волны (наибольшая), которую обычно определяют как длину волны, соответствующую спаду фототока на 50 % со стороны больших длин волн.

При малых длинах волн с уменьшением длины волны падающего на фоторезистор света растет показатель поглощения. Поэтому глубина проникновения квантов света в полупроводник уменьшается, т. е. основная часть неравновесных носителей заряда возникает вблизи освещаемой поверхности фоточувствительного слоя. При этом увеличивается роль поверхностной рекомбинации и уменьшается среднее время жизни неравновесных носителей. Таким образом, спектральная характеристика имеет спад и при малых длинах волн.

Различные полупроводники имеют ширину запрещенной зоны от десятых долей до трех электронвольт. Поэтому максимум спектральной характеристики различных фоторезисторов может находиться в инфракрасной, видимой или ультрафиолетовой частях электромагнитного спектра.

Постоянная времени — это время, в течение которого фототок фоторезистора изменяется после освещения или после затемнения фоторезистора на 63 % (в e раз) по отношению к установившемуся значению. Таким образом, постоянные времени характеризуют скорость реакции фоторезистора на изменение светового потока, т. е. характеризуют инерционность фоторезистора.

В связи с тем, что скорость нарастания фототока при освещении несколько отличается от скорости его спада после затемнения фоторезистора, различают постоянные времени нарастания τ_n и спада τ_{cn} . Численные значения постоянных времени различных фоторезисторов от десятков микросекунд до десятков миллисекунд. Постоянные времена измеряют при освещенности 200 лк, темпера-

туре окружающей среды 20 °С и сопротивлении нагрузки, включенном в схему измерения, меньше 1 кОм. Освещение при определении постоянных времени производят обычно от источника излучения с цветовой температурой 2840 К.

Все эти условия необходимы при измерении постоянных времени для однозначности получаемых результатов, так как постоянные времени зависят от концентрации ловушек захвата и от скорости их заполнения и опустошения, что, в свою очередь, изменяется при изменении освещенности, температуры и других условий, в которых работает фоторезистор. Так, с увеличением освещенности уменьшается число ловушек захвата и растет число рекомбинационных ловушек вследствие расщепления уровня Ферми на квазиуровни или смещения демаркационных уровней (см. рис. 1.8). Оба эти фактора приводят к уменьшению времени жизни носителей заряда и соответственно к уменьшению постоянных времени фоторезистора.

С повышением температуры возрастает вероятность ионизации ловушек захвата, что означает более быстрое их опустошение и уменьшение постоянных времени.

Наличие существенной инерционности у фоторезисторов приводит к тому, что с увеличением частоты модуляции светового потока эффективное значение возникающего переменного фототока уменьшается. Максимальная частота модуляции светового потока для фоторезисторов не превосходит десятков килогерц.

Темновое сопротивление — это сопротивление фоторезистора при отсутствии освещения. Темновое сопротивление принято определять через 30 с после затемнения фоторезистора, предварительно находящегося под освещенностью 200 лк. Обусловлено это инерционностью опустошения ловушек захвата после прекращения освещения. Например, у фоторезисторов ФСК-1 отношение темновых сопротивлений, измеренных после затемнения через 30 с и через 16 ч, может достигать трех порядков.

Удельная интегральная чувствительность — это отношение фототока к световому потоку и к приложенному напряжению:

$$K_0 = I_{\phi} / (\Phi U).$$

Чувствительность называют интегральной, потому что измеряют ее при освещении фоторезистора светом сложного спектрального состава: от источника света с цветовой температурой 2840 К при освещенности 200 лк. Удельные интегральные чувствительности различных фоторезисторов от 1 до 600 мА/(В·лм).

§ 10.2. ФОТОДИОДЫ

Полупроводниковый фотодиод — это полупроводниковый диод, обратный ток которого зависит от освещенности.

Обычно в качестве фотодиодов используют полупроводниковые диоды с *p-n*-переходом, смещенным в обратном направлении внешним источником питания. При поглощении квантов света в *p-n*-переходе или в прилегающих к нему областях кристалла полупровод-

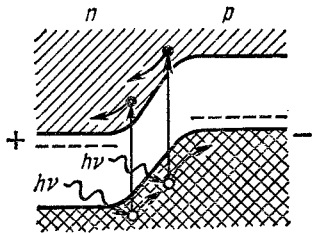


Рис. 10.4. Движение неравновесных носителей заряда, образованных квантами света в p - n -переходе или вблизи него, при обратном смещении p - n -перехода

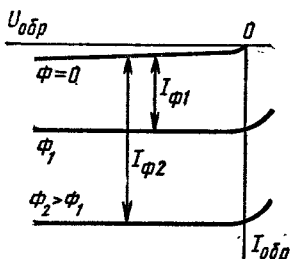


Рис. 10.5. Обратные ветви вольт-амперных характеристик фотодиода при различных световых потоках

ника образуются новые носители заряда (пары электрон—дырка). Неосновные носители, возникшие в прилегающих к p - n -переходу областях на расстоянии, не превышающем диффузионной длины, диффундируют к p - n -переходу и проходят через него под действием электрического поля или с точки зрения энергетической диаграммы скатываются с потенциального барьера (рис. 10.4). Поэтому обратный ток через фотодиод возрастает при освещении. К аналогичному результату приводит поглощение квантов света непосредственно в p - n -переходе. В результате при освещении фотодиода обратный ток через него возрастает на величину, называемую фототоком (рис. 10.5).

В рабочем диапазоне обратных напряжений при освещении фотодиода обратные токи практически не зависят от приложенного напряжения, хотя обратная ветвь вольт-амперной характеристики фотодиода в затемненном состоянии может не иметь участка насыщения тока.

В конструкции фотодиода, естественно, должна быть предусмотрена необходимость освещения кристалла полупроводника с одновременной защитой этого кристалла от других

внешних воздействий (рис. 10.6).

Свойства фотодиодов можно характеризовать параметрами и зависимостями, аналогичными параметрам и зависимостям фоторезисторов. Однако у фотодиодов есть существенные отличительные особенности. Так, световая характеристика фотодиода, т. е. зависимость фототока от освещенности, соответствует прямой пропорциональности фототока от освещенности. Связано это с тем, что толщина базы фотодиода значительно меньше диффузионной длины неосновных носителей заряда. Поэтому практически все неосновные носители, возникшие в базе в результате световой генерации, доходят до p - n -перехода и принимают участие в образовании фототока. Во всяком случае потери неосновных носителей на рекомбинацию в базе и на поверхности базы практически не зависят от освещенности, так как исходный полупроводник содержит малое количество неконтролируемых примесей, которые могли бы выполнять роль рекомбинационных ловушек и ловушек захвата.

Следствием линейности световой характеристики фотодиода является независимость интегральной чувствительности фотодиода от приложенного обратного напряжения. Поэтому одним из основ-

ных параметров фотодиода является не удельная интегральная чувствительность, а просто интегральная чувствительность:

$$K = I_{\Phi} / \Phi.$$

Другой особенностью фотодиодов и важным преимуществом их по сравнению с фоторезисторами является малая инерционность. Вообще на инерционность фотодиодов могут влиять три физических фактора: время диффузии или дрейфа неравновесных носителей заряда через базу τ_d ; время их пролета через p - n -переход τ_i ; время перезаряда барьерной емкости p - n -перехода, характеризующееся постоянной времени $rC_{бар}$.

Время диффузии носителей заряда через базу фотодиода можно определить аналогично времени пролета носителей через базу транзистора по формулам (4.134) или (4.135). В сплавных германиевых фотодиодах толщина базы составляет 20—30 мкм и $\tau_d \approx 50$ нс.

Время пролета носителей через p - n -переход

$$\tau_i \approx \delta / v_{\max},$$

где δ — толщина p - n -перехода; v_{\max} — максимальная скорость дрейфа носителей заряда (см. § 1.10).

В германии и кремнии $v_{\max} \approx 5 \cdot 10^6$ см/с, толщина p - n -перехода, зависящая от обратного напряжения и концентрации примесей в базе, обычно меньше 5 мкм. Следовательно, время пролета носителей через p - n -переход $\tau_i \approx 0,1$ нс.

Постоянная времени фотодиода $rC_{бар}$ определяется барьерной емкостью p - n -перехода, зависящей от напряжения, и сопротивлением базы фотодиода при малом сопротивлении нагрузки во внешней цепи. Сопротивление базы у фотодиодов значительно больше, чем у других диодов, так как невыпрямляющий контакт к базе фотодиода расположен по краям базы, а не по всей поверхности (см. рис. 10.6). Поэтому уменьшение толщины базы может привести не к уменьшению, а к увеличению сопротивления базы. Постоянная времени фотодиодов $rC_{бар}$ получается порядка наносекунд.

Таким образом, инерционность сплавных фотодиодов определяется временем диффузии носителей заряда через базу. В диффузионных фотодиодах, создав ускоряющее электрическое поле в базе из-за неравномерного распределения примесей, можно понизить время пролета носителей через базу до нескольких наносекунд. В таких фотодиодах все три фактора будут оказывать приблизительно одинаковое влияние на инерционность.

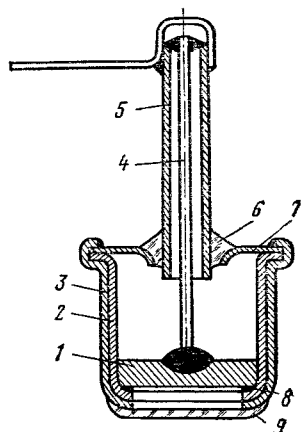


Рис. 10.6. Конструкция фотодиода ФД-1 в металлическом корпусе:

1 — круглая пластинка германия с p - n -переходом; 2 — кристаллодержатель; 3 — коваровый корпус; 4 — вывод; 5 — коваровая трубка; 6 — стеклянный проходной изолятор; 7 — ножка корпуса; 8 — кольцо припоя; 9 — стеклянная линза

Спектральная характеристика фотодиодов также определяется со стороны больших длин волн шириной запрещенной зоны исходного полупроводникового материала, при малых длинах волн — большим показателем поглощения и увеличением влияния поверхностной рекомбинации носителей заряда с уменьшением длины волны квантов света. Таким образом, коротковолновая граница фоточувствительности фотодиода зависит от толщины базы и от скорости поверхностной рекомбинации. Уменьшая эти величины, можно существенно сдвигать коротковолновую границу фоточувствительности фотодиодов в сторону меньших длин волн.

Положение максимума на спектральной характеристике фотодиода сильно зависит от степени роста коэффициента поглощения в данном полупроводнике. При резком увеличении коэффициента поглощения с уменьшением длины волны падающего света, например в германии, положение максимума определяется шириной запрещенной зоны ($\lambda_{\max} = 1,55$ мкм) и практически не зависит от толщины базы.

Если зависимость коэффициента поглощения от длины волны слабая, как, например, в кремнии, то эффект уменьшения проникновения квантов света в глубь полупроводника и увеличения роли поверхностной рекомбинации будет сказываться слабее с уменьшением длины волны. Поэтому максимум спектральной характеристики может смещаться при изменении толщины базы и скорости поверхностной рекомбинации. Так, максимум спектральной характеристики кремниевых фотодиодов можно смещать в диапазоне от 0,6 до 1 мкм путем изменения технологии их изготовления и конструкции.

Фотодиоды на основе контакта металл — полупроводник

Структура фотодиода на основе выпрямляющего контакта металл — полупроводник и его энергетическая диаграмма при обратном напряжении приведены на рис. 10.7. Чтобы основная часть квантов света проникла через металлический верхний электрод 1, его толщина должна быть малой (около 10 нм для Au). Уменьшения потерь на отражение можно достичь, используя просветляющие покрытия.

Принцип действия фотодиода на основе выпрямляющего контакта металл — полупроводник аналогичен принципу действия фотодиода с *p-n*-переходом. Однако есть некоторые различия, которые сказываются на характеристиках и параметрах. Первым отличием является возможность поглощения квантов света с энергией, меньшей ширины запрещенной зоны, для которых полупроводник оказывается прозрачным, в металле верхнего электрода 1. При этом, если энергия кванта света превышает высоту потенциального барьера, то возбужденные электроны из металла могут перейти в полупроводник через потенциальный барьер (рис. 10.7), обеспечивая тем самым возникновение фототока. Поэтому длинноволновая граница спектральной характеристики фотодиодов на основе контакта металл — полупроводник определяется высотой потенциального барьера на

этом контакте и расположена при более длинных волнах электромагнитного спектра.

Другим отличием рассматриваемых фотодиодов является то, что с уменьшением длины волны квантов света (с ростом энергии квантов) и с увеличением показателя поглощения в полупроводнике кванты света продолжают поглощаться в слое объемного заряда, где существует электрическое поле. Поэтому коротковолновая граница спектральной характеристики фотодиодов на основе контакта металл — полупроводник расположена при более коротких волнах электромагнитного спектра.

Таким образом, спектральная характеристика фотодиода на основе контакта металл — полупроводник значительно шире, чем спектральная характеристика фотодиода с *p-n*-переходом из того же полупроводника.

Кроме того, сопротивление базы фотодиода на основе контакта металл — полупроводник намного меньше. Поэтому постоянная времени τ_{rc} оказывается малой и инерционность определяется в основном только временем пролета носителей через область объемного заряда на выпрямляющем контакте металл — полупроводник. Это время пролета может быть порядка 0,1—0,01 нс, что позволяет использовать фотодиоды на основе контакта металл — полупроводник при СВЧ-модуляции светового потока.

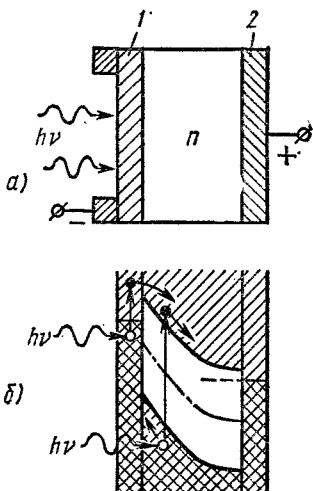


Рис. 10.7. Структура фотодиода на основе выпрямляющего контакта металл — полупроводник (*a*) и его энергетическая диаграмма при обратном напряжении (*б*):

1 — тонкий металлический электрод; 2 — нижний электрод — подложка

Фотодиоды на основе гетероперехода

Энергетическая диаграмма гетероперехода, смещенного в обратном направлении, показана на рис. 10.8. При освещении фотодиода с таким гетеропереходом со стороны широкозонного полупроводника квантами света с энергией $h\nu$ ($\Delta\mathcal{E}_1 > h\nu > \Delta\mathcal{E}_2$) свет поглощается в узкозонном полупроводнике. Широкозонный полупроводник оказывается прозрачным для таких квантов света. Возникшие при этом неосновные носители заряда, проходя через гетеропереход, создают фототок.

С уменьшением длины волны падающего на фотодиод света растет показатель поглощения узкозонного полупроводника. Глубина проникновения квантов в этот полупроводник уменьшается. Генерация неравновесных носителей происходит только вблизи гетероперехода.

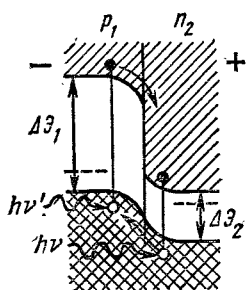


Рис. 10.8. Энергетическая диаграмма гетероперехода при обратном напряжении и при освещении его квантами света с различной энергией ($h\nu' > h\nu$)

При малых длинах волн падающего света ($h\nu' > \Delta\varepsilon_1$) кванты света поглощаются в широкозонном полупроводнике. Таким образом, спектральная характеристика фотодиода на основе гетероперехода получается более широкой по сравнению со спектральными характеристиками фотодиодов на основе обычных $p-n$ -переходов.

§ 10.3. ПОЛУПРОВОДНИКОВЫЕ ФОТОЭЛЕМЕНТЫ

Принцип действия

Полупроводниковый фотоэлемент — это полупроводниковый прибор с выпрямляющим электрическим переходом, предназначенный для непосредственного преобразования световой энергии в электрическую.

Фотоэлемент работает без внешних источников питания, а сам является источником электрической энергии.

Рассмотрим принцип действия фотоэлемента с $p-n$ -переходом в качестве выпрямляющего контакта. При освещении фотоэлемента из-за поглощения квантов света в $p-n$ -переходе и областях полупроводника, прилегающих к $p-n$ -переходу, происходит генерация новых носителей заряда. Диффузионное электрическое поле, существующее в $p-n$ -переходе, производит разделение неравновесных носителей. Другими словами, с точки зрения энергетической диаграммы $p-n$ -перехода (рис. 10.9) неравновесные электроны скатываются с потенциального барьера и попадают в n -область, а дырки, наоборот, в p -область. В результате накопления электронов в n -области и дырок в p -области между этими областями возникает дополнительная разность потенциалов — фото-э. д. с.

Накопление неравновесных носителей в соответствующих областях не может продолжаться беспрестанно, так как одновременно с накоплением дырок в p -области и электронов в n -области происходит понижение высоты потенциального барьера на величину возникшей фото-э. д. с. Уменьшение высоты потенциального барьера или уменьшение суммарной напряженности электрического поля в $p-n$ -переходе ухудшает «разделительные свойства» перехода.

Кроме составляющей фото-э. д. с., которая возникает из-за разделения носителей заряда электрическим полем $p-n$ -перехода или другого потенциального барьера и которая является основной

в фотоэлементах, могут быть и другие составляющие. При освещении полупроводника появляется градиент концентрации электронов и дырок, которые диффундируют от освещаемой поверхности в глубь полупроводника. Но коэффициенты диффузии электронов и дырок различны. Поэтому возникает вторая составляющая фото-э. д. с. (см. § 1.11). Кроме того, при наличии на освещаемой поверхности полупроводника ловушек захвата носителей одного знака возникает третья составляющая фото-э. д. с. в результате диффузии в глубь полупроводника носителей заряда другого знака.

Технология изготовления и конструкция

В настоящее время фотоэлементы широко применяются в виде солнечных батарей (совокупность электрически соединенных фотоэлементов) для преобразования энергии солнечного света непосредственно в электрическую энергию, питающую установки космических аппаратов. Обычно для этих целей используют кремниевые фотоэлементы. Электронно-дырочный переход в монокристаллической пластинке кремния с электропроводностью p -типа создают диффузией фосфора или сурьмы (рис. 10.10). При большой концентрации доноров (фосфор или сурьма) в поверхностном слое кремния проводимость n -области получается высокой. Поэтому невыпрямляющий контакт к этой области можно сделать в виде кольца или рамки, оставив всю поверхность кристалла доступной для освещения.

Основные характеристики и параметры

Вольт-амперные характеристики. Режиму работы фотоэлемента (режиму генерации фото-э. д. с.) при разных освещенностях или световых потоках соответствуют части вольт-амперных характеристик, расположенные в четвертом квадранте (рис. 10.11). Точки пересечения вольт-амперных характеристик с осью напряжения соответствуют значениям фото-э. д. с. или напряжениям холостого хода при разных освещенностях. У кремниевых фотоэлементов фото-э. д. с. 0,5—0,55 В.

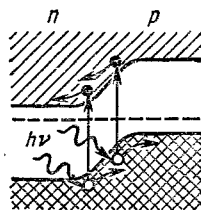


Рис. 10.9. Разделение неравновесных носителей заряда потенциальным барьером p - n -перехода при поглощении квантов света в p - n -переходе или вблизи него

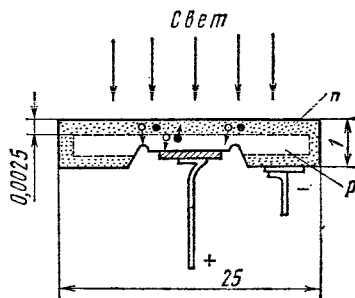


Рис. 10.10. Структура кремниевого фотоэлемента, изготовленного методом диффузии примесей

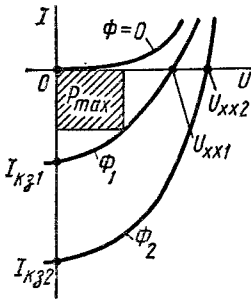


Рис. 10.11. Вольт-амперные характеристики фотоэлемента при различных световых потоках, падающих на фотоэлемент

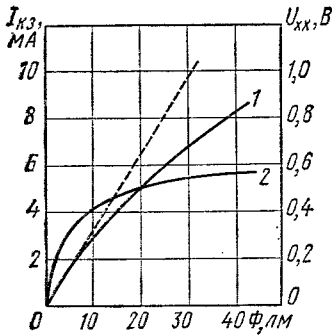


Рис. 10.12. Световые характеристики фотоэлемента:

1 — тока короткого замыкания;
2 — напряжения холостого хода

Сублинейность световых характеристик связана с уменьшением высоты потенциального барьера при накоплении избыточного заряда электронов в n -области и дырок в p -области.

Спектральная характеристика фотоэлемента — это зависимость тока короткого замыкания от длины волны падающего света. Спектральные характеристики фотоэлементов аналогичны спектральным характеристикам фотодиодов, изготовленных на основе того же полупроводника. Максимум спектральной характеристики кремниевых фотоэлементов почти соответствует максимуму спектрального распределения энергии солнечного света. Именно поэтому кремниевые фотоэлементы широко используют для создания солнечных батарей.

Коэффициент полезного действия фотоэлемента — это отношение максимальной мощности, которую можно получить от фотоэлемента, к полной мощности лучистого потока, падающего на рабочую поверхность фотоэлемента:

$$\eta = P_{\max}/P.$$

Точки пересечения вольт-амперных характеристик с осью токов соответствуют значениям токов короткого замыкания, которые зависят от площади выпрямляющего электрического перехода фотоэлемента. Поэтому сравнивают и оценивают фотоэлементы по плотностям тока короткого замыкания. У кремниевых фотоэлементов плотность тока короткого замыкания при средней освещенности солнечным светом 20—25 $\text{мА}/\text{см}^2$.

По вольт-амперным характеристикам при разных освещенностях фотоэлемента можно выбрать оптимальный режим работы фотоэлемента, т. е. оптимальное сопротивление нагрузки, при котором в нагрузке будет выделяться наибольшая мощность. Оптимальному режиму работы фотоэлемента соответствует наибольшая площадь прямоугольника с вершиной на вольт-амперной характеристике при данной освещенности (см. рис. 10.11).

Для кремниевых фотоэлементов при оптимальной нагрузке напряжение на нагрузке 0,35—0,4 В, плотность тока через фотоэлемент 15—20 $\text{мА}/\text{см}^2$.

Световые характеристики фотоэлемента — это зависимости фото-э. д. с. и тока короткого замыкания от светового потока или от освещенности фотоэлемента (рис. 10.12).

К основным процессам, приводящим к уменьшению к. п. д. фотоэлементов, относят: отражение части излучения от поверхности полупроводника, фотоэлектрически неактивное поглощение квантов света в полупроводнике (без образования пар носителей электрон—дырка), рекомбинацию неравновесных носителей еще до их разделения электрическим полем выпрямляющего электрического перехода (особенно на поверхности кристалла полупроводника), потери мощности при прохождении тока через объемное сопротивление базы фотоэлемента.

В результате к. п. д. кремниевых фотоэлементов при преобразовании солнечного света в электрическую энергию не превышает в настоящее время 12%. Однако его можно существенно повысить, используя в качестве исходного полупроводника теллурид кадмия, арсенид галлия и другие материалы с несколько большей шириной запрещенной зоны по сравнению с кремнием, а также используя фотоэлементы на основе гетеропереходов.

§ 10.4. ФОТОТРАНЗИСТОРЫ И ФОТОТИРИСТОРЫ

Транзистор, реагирующий на облучение световым потоком и способный одновременно усиливать фототок, называют фототранзистором.

Биполярные фототранзисторы

Биполярный фототранзистор может быть включен в схему по-разному. Если подать напряжение между базой и коллектором, сместив коллекторный переход в обратном направлении и оставив эмиттерный вывод неподключенным к схеме, то такое включение биполярного фототранзистора ничем не будет отличаться от схемы включения фотодиода. При поглощении квантов света в базовой и в коллекторной областях образуются неравновесные пары носителей заряда (электроны и дырки). Неосновные носители (дырки в n -базе и электроны в p -коллекторе для транзистора p - n - p -типа) диффундируют к коллекторному переходу, втягиваются существующим там электрическим полем в коллекторный переход и проходят через него, создавая тем самым фототок $I_{\text{ф}}$.

Однако биполярный фототранзистор обычно применяют при включении его по схеме с общим эмиттером. Поэтому рассмотрим принцип действия биполярного фототранзистора, включенного по схеме с общим эмиттером. Предположим вначале, что базовый вывод не подключен к схеме, т. е. ток базы равен нулю ($I_{\text{Б}} = 0$). В этом случае неосновные носители заряда, проходя через p - n -переход коллектора, создают тот же фототок $I_{\text{ф}}$. Неравновесные основные носители — электроны в n -базе, возникшие из-за поглощения там квантов света, и электроны, пришедшие в базу из коллектора, оказываются в своеобразной потенциальной яме (рис. 10.13). Накопление в базе неравновесных основных носителей заряда понижает высоту потенциальных барьеров эмиттерного и коллекторного переходов. Из-за уменьшения высоты потенциального барьера эмит-

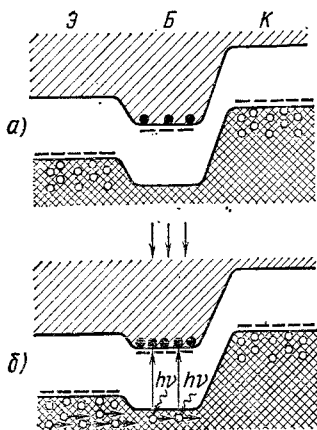


Рис. 10.13. Энергетическая диаграмма фототранзистора без освещения (а) и при освещении (б)

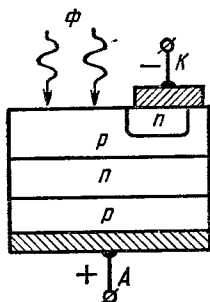


Рис. 10.14. Структура фототиристора

действий (например, изменение параметров, вызванное изменением температуры в процессе работы).

Фототиристоры

Фототиристором называют тиристор, напряжение включения которого уменьшается с увеличением освещенности.

Соответственно отличительной особенностью структуры и конструкции фототиристора является возможность освещения одной из базовых областей (рис. 10.14).

Переключение фототиристора из закрытого в открытое состояние происходит, как и у обычного тиристора, при увеличении суммарного коэффициента передачи по току тиристорной структуры до единицы (см. § 5.1). В фототиристере увеличение этого параметра может происходить в результате увеличения тока через тиристор-

терного перехода увеличивается инжекция дырок из эмиттера в базу (рис. 10.13, б). Соответственно возрастает и ток коллектора. Таким образом, накопленный в базе биполярного фототранзистора дополнительный заряд неравновесных основных носителей обеспечивает усиление фототока, т. е. при освещении результирующий ток коллектора $I'_ф \approx \approx h_{21Э} I_ф$.

При подключении вывода базы к внешней схеме ток базы может изменяться при освещении фототранзистора. Степень изменения этого тока зависит от сопротивлений в цепи базы. Изменение тока базы происходит в результате выхода неравновесных электронов из нее во внешнюю базовую цепь. В результате накопленный в базе заряд основных носителей уменьшается, что уменьшает усиление фототока.

Таким образом, биполярный фототранзистор обладает наибольшей чувствительностью к облучению светом базовой области при включении по схеме с общим эмиттером и отключенной базой. Поэтому у первых конструкций биполярных фототранзисторов вывод базы вообще отсутствовал. Однако наличие вывода базы у биполярных фототранзисторов позволяет использовать не только оптическое, но и электрическое управление фототранзистором, осуществить компенсацию посторонних внешних воз-

ную структуру при поглощении квантов света в базовых областях, т. е. из-за генерации носителей заряда в базовых областях при их освещении. Таким образом, фототиристор является аналогом управляемого тиристора, включение которого в открытое состояние может быть произведено импульсом света.

§ 10.5. ПРИЕМНИКИ ПРОНИКАЮЩЕЙ РАДИАЦИИ И КОРПУСКУЛЯРНО-ПРЕОБРАЗОВАТЕЛЬНЫЕ ПРИБОРЫ

Взаимодействие частиц высокой энергии и проникающей радиации с полупроводником

Частицы с высокой энергией и проникающая радиация, проходя через полупроводник, могут производить ионизацию атомов полупроводника. Каждая частица с высокой энергией или каждый квант проникающей радиации могут образовать в полупроводнике большое число свободных электронно-дырочных пар — носителей заряда. Этот процесс является основой принципа действия различных полупроводниковых приемников проникающей радиации и корпускулярно-преобразовательных приборов (датчиков проникающей радиации без электронно-дырочного перехода и с электронно-дырочным переходом, атомных электроэлементов).

Однако при воздействии проникающей радиации или частиц высокой энергии, кроме образования свободных носителей заряда, в полупроводнике могут происходить и нежелательные процессы. Так, атомы кристаллической решетки могут быть смещены из положения равновесия с образованием различных дефектов кристаллической решетки. Например, облучение германия электронами высокой энергии аналогично введению в кристаллическую решетку германия атомов акцепторной примеси — в запрещенной зоне полупроводника образуются акцепторные примесные уровни.

После облучения германия *n*-типа электронами с энергией около 1 МэВ проводимость уменьшается. Это связано с уменьшением концентрации носителей заряда в связи с компенсацией. Когда концентрация образованных облучением акцепторных дефектов равняется концентрации основных носителей заряда (электронов), наблюдается минимум проводимости. Дальнейшее облучение обусловливает возрастание концентрации дырок и увеличение проводимости (рис. 10.15). Необходимо отметить, что возникновение дополнительных акцепторных примесей в кристалле приводит к уменьшению подвижности носителей. Однако относительное уменьшение проводимости из-за уменьшения подвижности значительно меньше увеличения проводимости из-за изменения концентрации дырок. После облучения германия *p*-типа электронами высокой энергии его проводимость возрастает.

При облучении полупроводника медленными нейтронами в нем происходят радиоактивные изменения, в результате которых в исходном кристалле появляются примеси посторонних элементов.

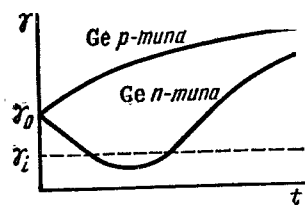


Рис. 10.15. Зависимость удельной проводимости германия от времени облучения электронами больших энергий

В германии, например, при нейтронном облучении появляются примеси галлия и мышьяка. Так как образование атомов галлия при этом происходит в три раза быстрее, чем атомов мышьяка, то общая электропроводность определяется избытком галлия, т. е. бомбардировка германия нейтронами приводит к появлению дырочной электропроводности.

При облучении полупроводника быстрыми нейтронами наряду с радиоактивными превращениями в кристалле происходит образование дефектов, подобных тем, какие образуются под действием быстрых электронов, и приводящих к появлению акцепторных уровней в кристалле.

Приемники проникающей радиации на основе поликристаллических полупроводников

Как принцип действия, так и метод изготовления приемников проникающей радиации на основе поликристаллических полупроводников аналогичен методу изготовления поликристаллических фоторезисторов. Основой приемника проникающей радиации является поликристаллический слой селенистого или сернистого кадмия, нанесенный на проводящую подложку путем, например, возгонки в вакууме. Эта подложка служит одновременно одним из электродов приемника проникающей радиации. Второй электрод наносят путем термического испарения в вакууме почти на всю поверхность поликристаллического полупроводникового слоя. Этот второй электрод в связи с малой его толщиной является прозрачным для рентгеновского излучения и тем более для гамма-излучения. Таким образом, в отличие от фоторезисторов в приемниках проникающей радиации электроды расположены на противоположных поверхностях поликристаллического полупроводникового слоя.

Отечественная промышленность выпускает приемники проникающей радиации различных марок: РГД-0, РГД-1, РГД-2 и ГД-Г1 (рентгено-гамма-датчик и гамма-датчик). Отличаются эти приемники проникающей радиации друг от друга в основном внешним конструктивным оформлением. Чувствительностью к рентгено- и гамма-излучению помимо перечисленных приемников проникающей радиации обладают сернисто-кадмиевые фоторезисторы ФСК-М и ФСК, а также селенисто-кадмиевые ФСД. Однако удельная чувствительность фоторезисторов к проникающей радиации примерно на два порядка меньше, чем приемников проникающей радиации. Свойства приемников проникающей радиации характеризуются зависимостями и параметрами, аналогичными зависимостям и параметрам фоторезисторов. Основным недостатком приемников проникающей радиации является их большая инерционность, особенно

при малых интенсивностях облучения. Постоянные времена приемников проникающей радиации составляют несколько десятков секунд при мощности облучения 5 Р/ч.

Приемники проникающей радиации и частиц высокой энергии на основе монокристаллов

Принцип действия такого приемника проникающей радиации и частиц высокой энергии аналогичен принципу действия фотодиода. Приемник проникающей радиации или частиц высокой энергии на основе монокристаллов полупроводника с *p-n*-переходом используют при подаче на него обратного напряжения. Поглощение в полупроводнике частиц высокой энергии или проникающей радиации приводит к возникновению новых неосновных носителей заряда и, следовательно, к увеличению обратного тока.

Рабочим объемом такого приемника в основном является объем *p-n*-перехода, т. е. область, обедненная подвижными носителями заряда. При проникновении в кристалл полупроводника частицы высокой энергии вдоль ее трека образуется колонка ионизации, геометрические размеры которой могут быть относительно большими, так как пробег частицы высокой энергии в полупроводнике может составлять десятки сантиметров. Поэтому толщина *p-n*-перехода приемника проникающей радиации или частиц высокой энергии должна существенно превышать толщину *p-n*-перехода фотодиода и других диодов.

Наиболее существенные результаты в создании диодов с большой толщиной *p-n*-перехода достигнуты при использовании диффузии атомов лития в кремний или германий с электропроводностью *p*-типа с последующим дрейфом ионов лития. Литий, представляющий собой донорную примесь для кремния и германия, имеет чрезвычайно высокий коэффициент диффузии, а ионы лития — большую подвижность при наличии электрического поля в монокристаллах кремния или германия. При изготовлении, например, кремниевых приемников проникающей радиации используют толстые (более 10 мм) монокристаллические пластинки высокоомного кремния с примесью бора. Диффузия лития с одной стороны пластинки кремния проводится при относительно низкой температуре (300—600 °С) в течение нескольких минут. В дальнейшем к полученному *p-n*-переходу прикладывают обратное напряжение при одновременном нагреве кристалла до температуры около 200 °С. Из-за большой подвижности ионы лития дрейфуют в электрическом поле, компенсируя исходную проводимость или заряды ионов бора в *p-n*-переходе. В результате получается распределение примесей по толщине кристалла, показанное на рис. 10.16. Таким методом создают диоды с толщиной *p-n*-перехода до 10 мм.

Приемники проникающей радиации на основе монокристаллов полупроводников с *p-n*-переходом, имеющим большую толщину, позволяют регистрировать и спектрометрировать длиннопробежные частицы высокой энергии. Основным недостатком таких приемников,

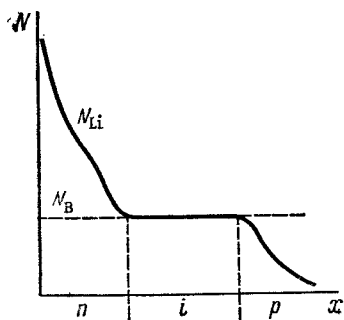


Рис. 10.16. Распределение концентрации примесей бора и лития по толщине кристалла кремния после диффузии и дрейфа лития

созданных с использованием диффузии и дрейфа лития, является необходимость их глубокого охлаждения при хранении и транспортировке во избежание самопроизвольной диффузии и самопроизвольного дрейфа лития, т. е. во избежание перераспределения примесей.

Атомные электроэлементы

С генерацией э. д. с. при облучении электронно-дырочного перехода частицами высокой энергии связан один из возможных путей создания атомного электроэлемента, преобразующего энергию радиоактивных излучений в электрическую энергию.

Первые атомные электроэлементы работали с радиоактивным стронцием (^{90}Sr), представляющим собой «золу» атомных реакторов. Стронций при распаде испускает электроны больших энергий, которые проникают в толщу полупроводниковой пластинки с p - n -переходом и производят там ионизацию атомов полупроводника. Возникшие пары носителей разделяются диффузионным полем перехода, что обеспечивает появление нескомпенсированных зарядов на электродах атомного электроэлемента. В результате в первых атомных электроэлементах создавалась разность потенциалов до 0,2 В. Ток короткого замыкания атомного электроэлемента площадью 0,32 см² составлял 5 мкА, коэффициент полезного действия — 1 %. Мощность, отдаваемая одним атомным электроэлементом, невелика, поэтому их соединяли последовательно и параллельно, создавая таким образом атомные батареи. Поскольку период полураспада стронция равен 28 годам, предполагалось, что срок службы атомных батарей должен исчисляться, по крайней мере, 28 годами. Однако оказалось, что эффективность атомных батарей значительно снижалась в течение двух-трех недель из-за того, что быстрые электроны, излучаемые стронцием, создавали дефекты в кристаллической решетке полупроводника, портили структуру электронно-дырочного перехода.

Дальнейшая разработка атомных батарей в основном ведется по трем направлениям: 1) подбирают радиоактивные вещества, спектр получения которых не содержит очень быстрых частиц; 2) предполагают использовать замедлители частиц с тем, чтобы p - n -переход подвергался облучению, но не выходил из строя; 3) энергию испускаемых радиоактивным веществом частиц пытаются предварительно преобразовать в световую энергию. Для этого на пути движения частиц высокой энергии помещают слой люминофора. Поглощаемые люминофором частицы вызывают его свечение, которое воздействует на p - n -переход фотоэлемента.

Полупроводниковый оптрон — это прибор, состоящий из оптически связанных между собой элементов оптроной пары (управляемого полупроводникового излучателя света и полупроводникового приемника излучения) и предназначенный для выполнения различных функциональных преобразований электрических или оптических сигналов.

В качестве одного элемента оптроной пары — излучателя света может быть использован светодиод, электролюминесцентный порошковый или пленочный излучатель, а также полупроводниковый лазер. В качестве второго элемента оптроной пары — приемника излучения может быть использован фоторезистор, фотодиод, фототранзистор или фототиристор. Наибольшее распространение получили следующие комбинации оптронных пар: светодиод — фоторезистор и светодиод — фототиристор (рис. 10.17).

Между элементами оптроной пары может быть осуществлена как оптическая, так и электрическая связь. На рис. 10.17 показана схема оптрона с внутренней прямой оптической связью. Усиление или просто преобразование входного электрического сигнала в таком оптроне происходит следующим образом. Изменение входного тока — тока через светодиод сопровождается изменением яркости его свечения и изменением освещенности приемника излучения (например, фоторезистора или фототиристора). Это в свою очередь приводит к изменению сопротивления приемника излучения и соответственно к изменению тока через выход оптрона или к изменению напряжения на нагрузочном сопротивлении во внешней цепи оптрона. Естественно, что для существования хорошей оптической связи между элементами оптроной пары кроме соответствующего расположения элементов оптроной пары необходимо по возможности более близкое совпадение спектральных характеристик этих элементов.

Интересным и полезным свойством такого оптрона кроме возможного усиления электрического сигнала по мощности, по напряжению и по току является полная электрическая развязка входа и выхода прибора, что позволяет полностью исключить обратную электрическую связь с его выхода на вход, т. е. изолировать друг от друга две части электрического тракта какой-либо системы.

Внутренняя оптическая связь в оптроне может быть не только прямой, но и обратной положительной или отрицательной (рис. 10.18 и 10.19). При наличии той или иной обратной оптической связи в оптроне между элементами оптроной пары должна существовать и электрическая связь. Поэтому такие оптроны иногда называют оптронами с электрооптической связью.

В *оптроне с положительной обратной оптической связью* (рис. 10.18, а) изменение яркости свечения излучателя света может происходить как из-за изменения сопротивления приемника излучения под действием входного оптического сигнала при неизменном входном напряжении на оптроне, так и из-за изменения входного напряжения при неизменной освещенности приемника излучения

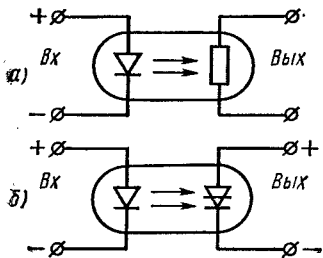


Рис. 10.17. Схемы оптронов АОР104А и АОР104Б, состоящих из светодиода и фоторезистора (а), а также оптронов АОУ103А и АОУ103Б, состоящих из светодиода и фототиристора

входным оптическим сигналом. Таким образом, этот оптрон имеет два входа: оптический $B_{вх}$ и электрический $U_{вх}$. Конструктивно такой оптрон должен быть выполнен так, чтобы, по крайней мере, часть светового потока от излучателя света поступала на приемник излучения, вызывая уменьшение его сопротивления. С уменьшением сопротивления приемника излучения возрастает яркость свечения излучателя света, что приводит к дальнейшему уменьшению сопротивления приемника излучения и т. д. Этот процесс при определенных условиях ($B_{вх} > B_{вх.пор}$ или $U_{вх} > U_{вх.пор}$) может привести к резкому скачкообразному возрастанию тока, проходящего через оптрон, и его выходного излучения (см. рис. 10.18, б).

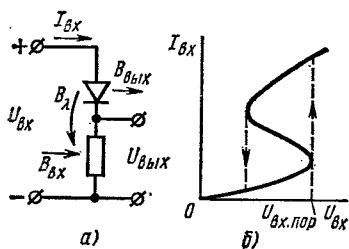


Рис. 10.18. Оптрон с положительной обратной связью между его элементами:

а — принципиальная схема; б — вольт-амперная характеристика

Кроме двух входов (оптического и электрического) оптрон имеет два аналогичных выхода, что позволяет разнообразить его области применения. Он может выполнять функции своеобразного ключа или триггера.

Пример схемы *оптрона с отрицательной обратной оптической связью* приведен на рис. 10.19, а. В таком оптроне излучатель света и приемник излучения включены параллельно. Поэтому увеличение входного напряжения $U_{вх}$ на оптроне и соответствующее увеличение тока через излучатель света вызывает уменьшение сопротивления приемника излучения, что приводит к еще большему увеличению входного тока через оптрон (рис. 10.19, б). Оптрон с отрицательной обратной оптической связью обладает нелинейным сопротивлением, характеристики которого могут изменяться в широких пределах, так как оптрон имеет электрический и оптический вход и выход.

На рис. 10.20 приведена схема *оптрона с электрической связью*. Входным управляющим сигналом оптрона является световой поток $B_{вх}$, подаваемый от внешнего излучателя на приемник излучения оптрона, сопротивление которого изменяется с изменением освещенности. При этом происходит перераспределение напряжения на элементах оптрона и изменение тока в последовательной цепи, что приводит к изменению яркости свечения излучателя света.

Такой оптрон может быть использован для усиления света без изменения его спектрального состава при соответствующих спект-

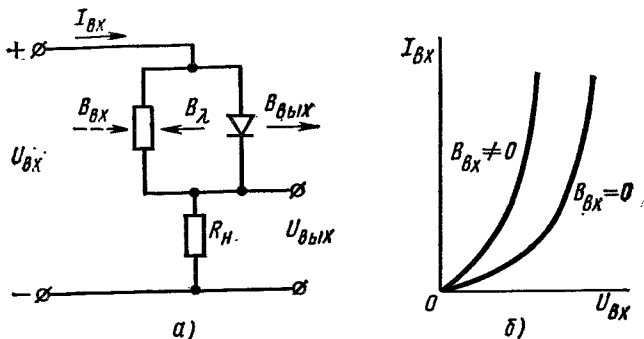


Рис. 10.19. Оптрон с отрицательной оптической обратной связью между его элементами:

а — принципиальная схема; б — вольт-амперная характеристика

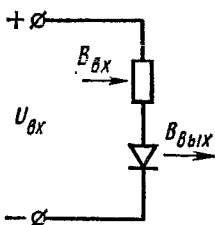


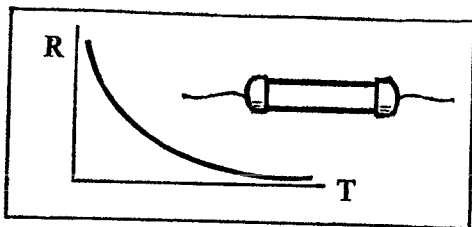
Рис. 10.20. Схема оптрона с электрической связью между его элементами

ральных характеристиках приемника входного излучения и излучателя оптрона. Кроме того, подбирая соответствующим образом спектральные характеристики приемника излучения и излучателя света, в таком оптроне с электрической связью можно осуществить преобразование видимого излучения одной длины волны в видимое излучение другой длины волны, невидимое инфракрасное излучение — в видимое, рентгеновское — в видимое и т. д.

Контрольные вопросы

1. Как объяснить спектральную характеристику фоторезистора?
2. Какими параметрами характеризуется фоторезистор?
3. Каковы отличия в свойствах фотодиода и фоторезистора?
4. Какие структуры могут иметь фотодиоды и каковы основные отличия в свойствах фотодиодов на основе различных выпрямляющих электрических переходов?
5. Как в фотоэлементе происходит непосредственное преобразование световой энергии в электрическую?
6. Каковы отличия в принципах действия и свойствах фотодиода и биполярного фототранзистора?
7. Почему фототиристор может управлять относительно большими мощностями по сравнению с допустимой мощностью рассеяния самого фототиристора?
8. В чем сходства и различия полупроводниковых приемников проникающей радиации и фоторезисторов или фотодиодов?
9. Как могут быть связаны между собой элементы оптронной пары в оптроне?

ТЕРМИСТОРЫ



Термистор — это полупроводниковый резистор, в котором используется зависимость электрического сопротивления полупроводника от температуры.

§ 11.1. ТЕРМИСТОРЫ ПРЯМОГО ПОДОГРЕВА

Принцип действия термисторов с отрицательным температурным коэффициентом сопротивления

В термисторах прямого подогрева сопротивление изменяется или под влиянием тепла, выделяющегося в них при прохождении электрического тока, или в результате изменения температуры термистора из-за изменения теплового облучения термистора (например, при изменении температуры окружающей среды). Наиболее широко применяются термисторы, основной особенностью которых является значительное уменьшение сопротивления с увеличением температуры, т. е. термисторы с отрицательным температурным коэффициентом сопротивления.

Уменьшение сопротивления полупроводника с увеличением температуры может быть обусловлено разными причинами — увеличением концентрации носителей заряда, увеличением их подвижности или фазовыми превращениями полупроводникового материала.

1. Первое явление характерно для термисторов, изготовленных из монокристаллов ковалентных полупроводников (кремний, германий, карбид кремния, соединения типа $A^{III}B^V$ и др.). Такие полупроводники обладают отрицательным температурным коэффициентом сопротивления в диапазоне температур, соответствующих примесной электропроводности, когда не все примеси ионизированы, а также в диапазоне температур собственной электропроводности, когда концентрация носителей изменяется из-за ионизации собственных атомов полупроводника (см. рис. 1.9). И в том и в другом случае зависимость сопротивления полупроводника определяется в основном изменением концентрации носителей заряда, так как

температурные изменения подвижности при этом пренебрежимо малы.

В этих диапазонах температур зависимость сопротивления полупроводника от температуры соответствует уравнению

$$R = R_{\infty} \exp \frac{B}{T}, \quad (11.1)$$

где B — коэффициент температурной чувствительности; R_{∞} — постоянная, зависящая от материала и размеров термистора.

При неполной ионизации примесей и отсутствии компенсации

$$B \approx \Delta \mathcal{E}_n / (2k),$$

где $\Delta \mathcal{E}_n$ — энергия ионизации примесей (доноров или акцепторов).

Для скомпенсированного полупроводника при неполной ионизации примесей

$$B \approx \Delta \mathcal{E}_n / k. \quad (11.2)$$

При собственной электропроводности

$$B \approx \Delta \mathcal{E} / (2k), \quad (11.3)$$

где $\Delta \mathcal{E}$ — ширина запрещенной зоны полупроводника.

2. Основная часть термисторов, выпускаемых промышленностью, изготовлена из поликристаллических окисных полупроводников — из окислов так называемых металлов переходной группы таблицы Менделеева (от титана до цинка). Термисторы в форме стержней, трубок, дисков или пластинок получают методами керамической технологии, т. е. путем обжига заготовок при высокой температуре.

Электропроводность окисных полупроводников с преобладающей ионной связью между атомами отличается от электропроводности ковалентных полупроводников. Для металлов переходной группы характерны наличие незаполненных электронных оболочек и переменная валентность. В результате при образовании окисла в определенных условиях (наличие примесей, отклонение от стехиометрии) в одинаковых кристаллографических положениях оказываются ионы с разными зарядами. Электропроводность таких материалов связана с обменом электронами между соседними ионами. Энергия, необходимая для такого обмена, мала. Поэтому все электроны (или дырки), которые могут переходить от одного иона к другому, можно считать свободными носителями заряда, а их концентрацию — постоянной в рабочем диапазоне температур термистора.

Из-за сильного взаимодействия носителей заряда с ионами подвижность носителей заряда в окисном полупроводнике оказывается малой и экспоненциально возрастающей с ростом температуры. В результате температурная зависимость сопротивления термистора из окисного полупроводника оказывается такой же, как и у термисторов из ковалентных полупроводников (рис. 11.1), но коэффициент температурной чувствительности характеризует в этом случае изменение подвижности носителей заряда, а не изменение их концентрации.

3. В окислах ванадия V_2O_4 и V_2O_3 при температуре фазовых превращений (68 и -110°C) наблюдается уменьшение удельного сопротивления на несколько порядков. Это явление также может быть использовано для создания термисторов с большим отрицательным температурным коэффициентом сопротивления в диапазоне температур, соответствующих фазовому превращению.

Характеристики и параметры термисторов прямого подогрева

Температурная характеристика термистора — это зависимость его сопротивления от температуры. Пример температурной характеристики одного из термисторов приведен на рис. 11.1.

Номинальное сопротивление термистора — это его сопротивление при определенной температуре (обычно 20°C). Термисторы изготавливают с допустимым отклонением от номинального сопротивления $\pm 20, 10$ и 5% . Номинальные сопротивления различных типов термисторов имеют значения от нескольких ом до нескольких сотен килоом.

Коэффициент температурной чувствительности B — это коэффициент в показателе экспоненты температурной характеристики термистора (11.1). Значение этого коэффициента, зависящее от свойств материала термистора, практически постоянно для данного термистора в рабочем диапазоне температур и для различных типов термисторов находится в пределах от 700 до 15 000 К. Коэффициент температурной чувствительности может быть найден экспериментально путем измерения сопротивлений термистора при двух температурах T_0 и T по формуле

$$B = \frac{\ln \frac{R}{R_0}}{1/T_0 - 1/T}.$$

Температурный коэффициент сопротивления термистора показывает относительное изменение сопротивления термистора при изменении температуры на один градус:

$$\text{TK } R|_T = \frac{1}{R} \cdot \frac{dR}{dT}. \quad (11.4)$$

Температурный коэффициент сопротивления зависит от температуры, поэтому его необходимо записывать с индексом, указывающим температуру, при которой имеет место данное значение. Зависимость температурного коэффициента сопротивления от температуры можно получить, используя уравнения (11.4) и (11.1):

$$\text{TK } R = -B/T^2.$$

Значения температурного коэффициента сопротивления при комнатной температуре различных термисторов находятся в пределах $-(0,8-6,0) \cdot 10^{-2} \text{ K}^{-1}$.

Коэффициент рассеяния термистора H численно равен мощности, рассеиваемой термистором при разности температур термистора

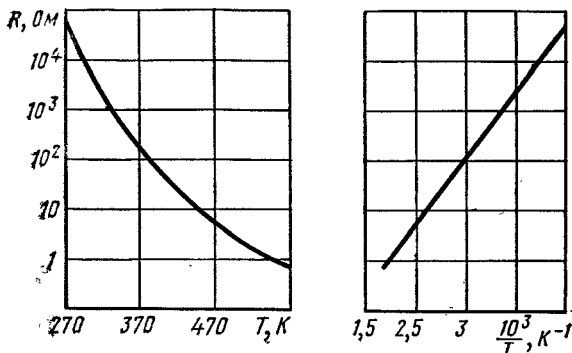
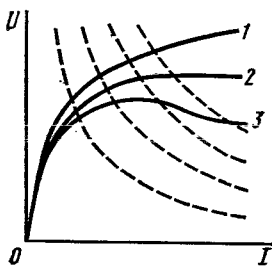


Рис. 11.1. Температурная характеристика одного из термисторов

Рис. 11.2. Три типа статических вольт-амперных характеристик термисторов прямого подогрева (сплошные линии) и гиперболы равной мощности (штриховые линии)



и окружающей среды в один градус, или, другими словами, численно равен мощности, которую надо выделить в термисторе, чтобы нагреть его на один градус.

Статическая вольт-амперная характеристика термистора — это зависимость падения напряжения на термисторе от проходящего через него тока в условиях теплового равновесия между термистором и окружающей средой.

На рис. 11.2 показаны статические вольт-амперные характеристики термисторов. Линейность характеристик при малых токах и напряжениях объясняется тем, что выделяемая в термисторе мощность недостаточна для существенного изменения его температуры. При увеличении тока, проходящего через термистор, выделяемая в нем мощность повышает его температуру. Сопротивление термистора определяется, таким образом, суммарной температурой — температурой окружающей среды и температурой перегрева термистора. При этих токах сопротивление термистора уменьшается с увеличением тока и температуры в соответствии с (11.1), линейность статической вольт-амперной характеристики нарушается. При дальнейшем увеличении тока и большой температурной чувствительности термистора может наблюдаться падающий участок статической вольт-амперной характеристики, т. е. уменьшение напряжения на термисторе с увеличением проходящего через него тока.

Мощность, выделяющаяся в термисторе, непрерывно возрастает с увеличением проходящего через термистор тока, несмотря на уменьшение падающего на термисторе напряжения. В результате

гиперболы равной мощности пересекают статическую вольт-амперную характеристику термистора только в одной точке (рис. 11.2).

Для каждой точки статической вольт-амперной характеристики термистора справедливо уравнение теплового баланса между мощностью, выделяющейся в термисторе из-за проходящего тока, и мощностью, которую он рассеивает в окружающей среде:

$$U^2/R = I^2 R = H(T - T_{\text{окр}}), \quad (11.5)$$

где T — температура термистора; $T_{\text{окр}}$ — температура окружающей среды.

Из уравнения (11.5) с учетом (11.1) можно получить уравнения статической вольт-амперной характеристики термистора в параметрическом виде:

$$U = \sqrt{HR_{\infty}(T - T_{\text{окр}}) \exp \frac{B}{T}}; \quad (11.6)$$

$$I = \sqrt{\frac{H}{R_{\infty}}(T - T_{\text{окр}}) \exp \left(-\frac{B}{T}\right)}. \quad (11.7)$$

Вид статической вольт-амперной характеристики термистора определяется коэффициентом рассеяния H , коэффициентом температурной чувствительности B , номинальным сопротивлением термистора и температурой окружающей среды. Так, при уменьшении коэффициента рассеяния H (например, при уменьшении давления воздуха, окружающего термистор) происходит более интенсивный разогрев термистора, и, следовательно, те же температуры достигаются при меньших мощностях, выделяемых в термисторе при прохождении тока, т. е. статическая вольт-амперная характеристика смещается вниз (в область меньших напряжений).

При увеличении температуры окружающей среды уменьшается сопротивление термистора, снижается максимум статической вольт-амперной характеристики и уменьшается ее крутизна. Такую зависимость используют в системах автоматического контроля и регулирования температуры.

Увеличение коэффициента температурной чувствительности B приводит к смещению максимума статической вольт-амперной характеристики в сторону меньших мощностей, а крутизна падающего участка возрастает.

Исследуем уравнение (11.6) на максимум функции, полагая коэффициент температурной чувствительности и коэффициент рассеяния постоянными. Для этого первую производную от напряжения по току приравняем нулю. В связи с параметрическим видом вольт-амперной характеристики в данном случае производную от напряжения по вспомогательной переменной, т. е. по параметру T , умножим на производную температуры по току и приравняем полученное произведение нулю:

$$\frac{dU}{dI} = \frac{dU}{dT} \cdot \frac{dT}{dI} = 0.$$

В результате получим

$$T_{U_{\text{экстр}}}^2 - BT_{U_{\text{экстр}}} + BT_{\text{окр}} = 0; \quad (11.8)$$

отсюда

$$T_{U_{\text{экстр}}} = \frac{B \pm \sqrt{B(B - 4T_{\text{окр}})}}{2}. \quad (11.9)$$

Из уравнения (11.9) следует:

1. Статическая вольт-амперная характеристика термистора будет иметь экстремальные значения напряжения (кривая 3 на рис. 11.2) только при условии

$$B > 4T_{\text{окр}}. \quad (11.10)$$

2. Существует два решения, соответствующие максимуму и минимуму напряжения на термисторе. Однако при минимуме напряжения на термисторе температура его оказывается выше допустимой, т. е. практически минимум напряжения находится за пределами рабочего диапазона температур термистора.

3. Температура, а значит, и сопротивление термистора при экстремальных значениях напряжения определяются только величинами B и $T_{\text{окр}}$. Температура термистора при экстремальных значениях напряжения не зависит, например, от коэффициента рассеяния. Поэтому максимумы (и минимумы) статических вольт-амперных характеристик термистора, помещенного в различные среды, должны наблюдаться при одних и тех же сопротивлениях термистора. Коэффициент рассеяния сказывается только на значениях напряжения и тока в точке максимума (и минимума) статической вольт-амперной характеристики, что следует из уравнений (11.6) и (11.9).

Максимально допустимая температура термистора — это температура, при которой еще не происходит необратимых изменений параметров и характеристик термистора. Максимально допустимая температура определяется не только свойствами исходных материалов термистора, но и его конструктивными особенностями.

Максимально допустимая мощность рассеяния термистора — это мощность, при которой термистор, находящийся в спокойном воздухе при температуре 20 °С, разогревается при прохождении тока до максимально допустимой температуры. При уменьшении температуры окружающего воздуха, а также при работе термистора в средах, обеспечивающих лучший теплоотвод, мощность рассеяния может превышать максимально допустимое значение.

Коэффициент энергетической чувствительности термистора G численно равен мощности, которую необходимо подвести к термистору для уменьшения его сопротивления на 1 %. Коэффициент энергетической чувствительности связан с коэффициентом рассеяния и температурным коэффициентом сопротивления соотношением

$$G = H/TKR.$$

Величина коэффициента энергетической чувствительности зависит от режима работы термистора, т. е. будет различна в каждой точке статической вольт-амперной характеристики.

Постоянная времени термистора — это время, в течение которого температура термистора уменьшится на 63 % (в e раз) по отношению к разности температур термистора и окружающей среды (например, при переносе термистора из воздушной среды с температурой 120 °С в воздушную среду с температурой 20 °С). Тепловая инерционность термистора, характеризуемая его постоянной времени, определяется конструкцией и размерами термистора и зависит от теплопроводности среды, в которой находится термистор. Для разных типов термисторов постоянная времени лежит в пределах от 0,5 до 140 с.

§ 11.2. БОЛОМЕТРЫ

Полупроводниковый болометр — это прибор, предназначенный для индикации и измерения теплового излучения (оптического или инфракрасного диапазона частот электромагнитного излучения).

Обычно болометр состоит из двух пленочных термисторов (толщиной до 10 мкм). Один из термисторов болометра является активным, т. е. непосредственно подвергается воздействию измеряемого излучения. Сопротивление этого термистора изменяется в результате нагрева при облучении электромагнитным излучением оптического или инфракрасного диапазона частот. Второй термистор — компенсационный, служит для компенсации возможных изменений температуры окружающей среды. Компенсационный термистор должен быть экранирован от измеряемого излучения. Активный и компенсационный термисторы помещают в один герметичный корпус.

Болометры обычно имеют три внешних вывода — от активного и компенсационного термисторов и от средней точки.

Для характеристики болометров используют следующие параметры: 1) сопротивление активного термистора болометра при комнатной температуре; 2) рабочее напряжение; 3) чувствительность при определенной частоте модуляции лучистого потока, равная отношению полезного сигнала, снимаемого с болометра на вход усилителя, к мощности излучения, падающего на болометр; 4) порог чувствительности, численно равный мощности излучения, которая вызывает сигнал, эквивалентный уровню собственных шумов болометра, т. е. порог чувствительности определяется минимальной мощностью излучения, которую при данных условиях способен зарегистрировать болометр; 5) постоянная времени, характеризующая тепловую инерционность активного термистора; 6) уровень собственных шумов.

Полупроводниковые болометры применяют в различных системах ориентации; для бесконтактного и дистанционного измерения температур и т. д.

§ 11.3. ТЕРМИСТОРЫ КОСВЕННОГО ПОДОГРЕВА

Термистор косвенного подогрева — это термистор, имеющий дополнительный источник тепла — подогреватель.

Конструктивное выполнение термисторов косвенного подогрева может быть различным. Часто подогреватель делают в виде обмотки на изоляционной трубке, внутри которой расположен термистор. В других случаях сам термистор сделан в виде трубки, внутри которой проходит нить подогрева. Общим для всех конструкций термисторов косвенного подогрева является наличие у них двух электрически изолированных друг от друга цепей — управляющей и управляемой.

Кроме таких параметров, как номинальное сопротивление и коэффициент температурной чувствительности, термисторы косвенного подогрева имеют свои специфические характеристики и параметры.

Статические вольт-амперные характеристики термистора косвенного подогрева снимают при различных токах через подогреватель (рис. 11.3).

Подогревная характеристика — это зависимость сопротивления термистора косвенного подогрева от мощности, выделяемой в спирали подогревной обмотки (рис. 11.4).

Для получения наибольшей чувствительности термистора косвенного подогрева, т. е. наибольшего изменения сопротивления, его следует использовать в режимах, при которых мощностью, выделяемой в самом термочувствительном элементе проходящим через него током, можно было бы пренебречь.

Коэффициент тепловой связи — это отношение мощности P_T , необходимой для разогрева термочувствительного элемента до некоторой температуры при прямом нагреве, к мощности $P_{под}$, необходимой для разогрева до той же температуры при косвенном подогреве, т. е. путем пропускания тока через подогреватель:

$$K = P_T / P_{под}$$

Обычно для определения коэффициента тепловой связи термистор косвенного подогрева разогревают до так называемого горячего

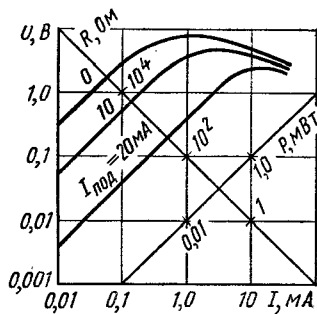


Рис. 11.3. Статические вольт-амперные характеристики термистора косвенного подогрева при различных токах через подогреватель

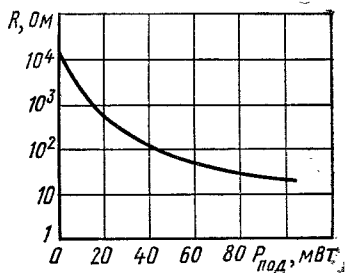


Рис. 11.4. Подогревная характеристика термистора косвенного подогрева

сопротивления термистора при максимальной мощности, выделяемой в подогревателе. Коэффициент тепловой связи обычно находится в пределах от 0,5 до 0,97, т. е. меньше единицы, так как часть тепла, выделяемая подогревателем, неизбежно теряется.

Постоянные времени. Тепловая инерционность термисторов косвенного подогрева характеризуется двумя постоянными времени. За величину первой постоянной времени принимают время, в течение которого температура термочувствительного элемента изменяется в e раз по отношению к установившемуся значению при мгновенном изменении мощности в цепи подогревателя. Вторая постоянная времени характеризует задержку в изменении температуры термочувствительного элемента по отношению к изменению температуры подогревателя. Таким образом, первая постоянная времени характеризует тепловую инерционность всей конструкции термистора косвенного подогрева; вторая постоянная времени — тепловую инерционность термочувствительного элемента.

§ 11.4. ПОЗИСТОРЫ

Позисторы — это полупроводниковые резисторы с положительным температурным коэффициентом сопротивления.

В массовом производстве позисторы делают на основе титанатобариевой керамики, удельное сопротивление которой значительно уменьшено в результате добавления примесей. Титанат бария ($BaTiO_3$) — диэлектрик с удельным сопротивлением при комнатной температуре 10^{10} — 10^{12} Ом·см. Введение в титанат бария малых количеств (0,1—0,3 ат. %) примесей редкоземельных элементов (лантана, церия или др.) приводит к уменьшению его удельного сопротивления до 10 — 10^2 Ом·см (рис. 11.5).

При дальнейшем увеличении концентрации примесей степень тетрагональности материала уменьшается, происходит разукрупнение кристаллов, и поэтому удельное сопротивление увеличивается. Такой материал обладает аномальной температурной зависимостью: в узком диапазоне температур его удельное сопротивление увеличивается на несколько порядков с увеличением температуры.

Резкое увеличение удельного сопротивления титаната бария происходит из-за тетрагонально-кубического фазового превращения, т. е. в диапазоне температур выше точки Кюри, при которых титанат бария переходит из сегнетоэлектрического в параэлектрическое состояние.

При изменении катионного состава сегнетоэлектриков точка Кюри смещается. Поэтому, увеличивая относительное содержание стронция и олова в твердых растворах типа $BaTiO_3$ — $SrTiO_3$ и $BaTiO_3$ — $BaSnO_3$, можно смещать точку Кюри в сторону более низких температур.

Конструктивно позисторы оформляют в виде дисков с металлическими электродами на торцах и с проволочными выводами от электродов.

Иногда для создания позисторов используют монокристаллические кремний, германий и другие полупроводниковые материалы. Принцип действия таких позисторов основан на уменьшении подвижности носителей заряда с увеличением температуры в результате увеличения их рассеяния на тепловых колебаниях атомов кристаллической решетки. Так, позисторы, изготовленные из монокристаллического кремния с малой концентрацией примесей (10^{15} — 10^{17} см⁻³), могут иметь температурный коэффициент сопротивления $(0,7-1) \cdot 10^{-2}$ К⁻¹ с положительным знаком в диапазоне температур от 20 до ~ 100 °С.

Одним из преимуществ монокристаллических позисторов по сравнению с поликристаллическими является возможность изготовления их с относительно малым разбросом параметров и характеристик. При использовании для изготовления монокристаллических позисторов широко распространенной планарной технологии можно создавать монокристаллические позисторы с разбросом номинальных сопротивлений 1—2%. Однако из-за меньшей стоимости и из-за больших температурных коэффициентов сопротивления поликристаллические позисторы нашли более широкое применение.

Свойства позисторов оцениваются характеристиками и параметрами, аналогичными характеристикам и параметрам термисторов с отрицательным температурным коэффициентом сопротивления.

Температурная характеристика. Зависимость сопротивления различных позисторов от температуры показана на рис. 11.6. При относительно малых и больших температурах температурные характеристики позисторов соответствуют температурным зависимостям удельного сопротивления обычных ионных полупроводников,

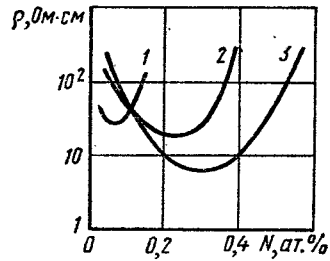


Рис. 11.5. Зависимости удельного сопротивления титаната бария от концентрации различных примесей:

1 — неодим; 2 — церий, лантан, ниобий; 3 — иттрий

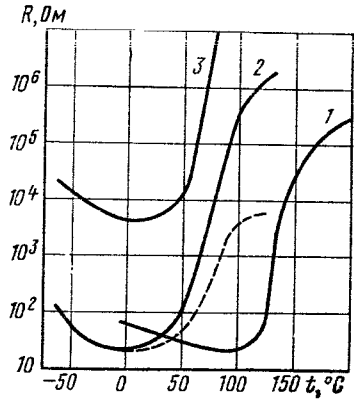


Рис. 11.6. Температурные характеристики некоторых позисторов:

1 — СТ5-1; 2 — СТ6-2Б; 3 — СТ6-3Б; штриховая кривая для СТ6-2Б получена при нагреве его проходящим током

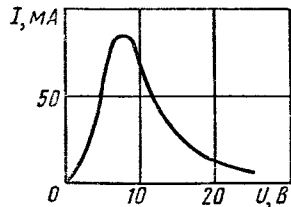


Рис. 11.7. Статическая вольт-амперная характеристика позистора СТ5-1

т. е. позисторы имеют отрицательный температурный коэффициент сопротивления.

Существенным недостатком поликристаллических позисторов является неоднозначность температурных характеристик при различных напряжениях, приложенных к позистору: увеличение напряжения снижает величину сопротивления и уменьшает изменение сопротивления при разогреве позистора (рис. 11.6).

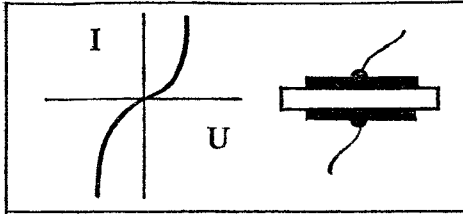
Температурный коэффициент сопротивления для позисторов оказывается не очень удобным параметром, так как его величина очень сильно зависит от температуры.

Статическая вольт-амперная характеристика позистора (рис. 11.7), так же как и обычного термистора, представляет собой зависимость напряжения на позисторе от проходящего через него тока при условии теплового равновесия между теплом, выделяемым в позисторе, и теплом, отводимым от позистора.

Контрольные вопросы

1. Какие физические явления могут быть положены в основу создания термисторов с отрицательным температурным коэффициентом сопротивления?
2. При каких условиях термистор будет обладать отрицательным дифференциальным сопротивлением?
3. Что характеризует постоянная времени термистора?
4. Почему полупроводниковые болометры состоят из двух термисторов?
5. На каких физических явлениях может быть основан принцип действия позисторов?

ВАРИСТОРЫ



§ 12.1. ПРИНЦИП ДЕЙСТВИЯ

Варистор — это полупроводниковый резистор, сопротивление которого зависит от приложенного напряжения, и поэтому его динамическая вольт-амперная характеристика нелинейна (рис. 12.1.).

Варисторы изготавливают методами керамической технологии, т. е. путем высокотемпературного обжига заготовок из порошкообразного карбида кремния со связкой, в качестве которой используют глину. Конструктивное оформление варисторов может быть различным в зависимости от назначения и необходимых параметров.

Нелинейность вольт-амперных характеристик варисторов обусловлена явлениями на точечных контактах между кристаллами карбида кремния: увеличение в сильных электрических полях проводимости поверхностных потенциальных барьеров или окисных пленок на кристаллах карбида кремния и увеличение проводимости точечных контактов между кристаллами из-за разогрева в связи с выделяющейся на контактах мощностью.

Поскольку толщина поверхностных потенциальных барьеров и окисных пленок на кристаллах карбида кремния мала, там могут возникать сильные электрические поля даже при малых напряжениях на варисторе, что приводит к туннелированию носителей заряда сквозь потенциальные барьеры или сквозь тонкие окисные пленки. Таким образом, при малых напряжениях на варисторе нелинейность вольт-амперной характеристики связана с зависимостью проводимости поверхностных потенциальных барьеров и окисных пленок от величины напряжения.

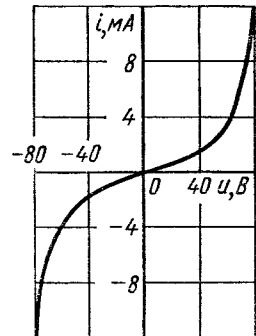


Рис. 12.1. Вольт-амперная характеристика одного из варисторов

При больших напряжениях на варисторе и соответственно, при больших токах, проходящих через варистор, плотность тока в точечных контактах оказывается очень большой. Все напряжение, приложенное к варистору, падает на точечных контактах. Поэтому удельная мощность (мощность в единице объема), выделяющаяся в точечных контактах, достигает таких значений, которые нельзя не учитывать. Разогрев точечных контактов приводит к уменьшению их сопротивления и к нелинейности вольт-амперной характеристики.

Сопротивление точечных контактов определяется сопротивлением растекания (см. § 3.21), т. е. сопротивлением малых активных областей полупроводника под точечными контактами. Из-за малости активных областей их разогрев практически не приводит к повышению температуры всего варистора. Кроме того, малые объемы активных областей обеспечивают малую инерционность тепловых процессов — разогрева и охлаждения этих областей. Теоретические расчеты показывают, что тепловая постоянная времени активных областей может составлять 10^{-6} — 10^{-7} с.

Считая разогрев активных областей под точечными контактами одним из основных процессов, приводящих к нелинейности вольт-амперной характеристики в рабочем диапазоне напряжений и токов для варистора, можно получить ряд важных зависимостей и характеристик варисторов.

Температурная зависимость удельной проводимости полупроводников соответствует уравнению

$$\gamma = \gamma_{\infty} \exp(-B/T). \quad (12.1)$$

Сопротивление растекания двух контактирующих кристаллов [см. (3.103)]

$$R_{\text{кон}} \approx \frac{2}{\pi d \gamma} = \frac{2}{\pi d \gamma_{\infty}} \exp \frac{B}{T}, \quad (12.2)$$

где d — диаметр точечного контакта; B — коэффициент температурной чувствительности поверхностных слоев кристаллов карбида кремния.

Тогда статическое сопротивление варистора, состоящего из a параллельно включенных цепочек, имеющих в свою очередь b последовательно включенных контактирующих кристаллов,

$$R = \frac{b}{a} R_{\text{кон}} = \frac{2b}{a\pi d \gamma_{\infty}} \exp \frac{B}{T}. \quad (12.3)$$

Уравнение теплового баланса для активных областей варистора

$$U^2/R = I^2 R = H(T - T_0), \quad (12.4)$$

где H — коэффициент рассеяния активных областей; T — температура активных областей; T_0 — температура среды, окружающей активные области.

Из уравнения (12.4), учитывая (12.3), можно по аналогии со сказанным в § 11.1 получить уравнения вольт-амперной характери-

стики варистора в параметрической форме:

$$U = \sqrt{\frac{2Hb(T-T_0)}{a\pi d\gamma_\infty}} \exp \frac{B}{T}; \quad (12.5)$$

$$I = \sqrt{\frac{a\pi dH\gamma_\infty(T-T_0)}{2b}} \exp \left(-\frac{B}{T} \right). \quad (12.6)$$

Уравнения (12.5) и (12.6) неудобны для расчета цепей с варисторами, так как содержат ряд величин (a , b , d , H), значения которых практически невозможно определить непосредственно. Однако с помощью этих уравнений можно выяснить некоторые требования, предъявляемые к исходному материалу для изготовления варисторов с заданными свойствами. Кроме того, эти уравнения дают возможность установить связь между различными параметрами варисторов.

§ 12.2. ХАРАКТЕРИСТИКИ

Коэффициент нелинейности варистора — это отношение статического R и дифференциального r сопротивлений при заданном постоянном напряжении на варисторе:

$$\beta = \frac{R}{r} = \frac{U}{I} \cdot \frac{dI}{dU}. \quad (12.7)$$

Из уравнений (12.5) и (12.6) можно найти дифференциальное сопротивление варистора:

$$r = \frac{dU}{dI} = \frac{dU}{dT} / \frac{dI}{dT} = \frac{2b}{a d\pi\gamma_\infty} \left(\frac{T^2 - BT + BT_0}{T^2 + BT - BT_0} \right) \exp \frac{B}{T}. \quad (12.8)$$

Тогда коэффициент нелинейности варистора с учетом соотношений (12.7) и (12.8)

$$\beta = \frac{R}{r} = \frac{T^2 + BT - BT_0}{T^2 - BT + BT_0}. \quad (12.9)$$

На рис. 12.2 показаны расчетные зависимости коэффициента нелинейности от напряжения и температуры активных областей варисторов из материалов с различной величиной коэффициента температурной чувствительности B поверхностных слоев кристаллов.

Для определения положения максимумов этих зависимостей продифференцируем (12.9) по температуре и приравняем производную нулю. Тогда получим условие $T = 2T_0$, при котором коэффициент нелинейности будет иметь максимальное значение:

$$\beta_{\max} = \frac{4T_0 + B}{4T_0 - B}. \quad (12.10)$$

Из соотношения (12.10) следует, что при $B < 4T_0$ должны наблюдаться нелинейные вольт-амперные характеристики с положительным коэффициентом нелинейности. Максимальный коэффициент нелинейности должен быть равен бесконечности при $B = 4T_0$. Если $B > 4T_0$, то у варистора должно наблюдаться отрицательное

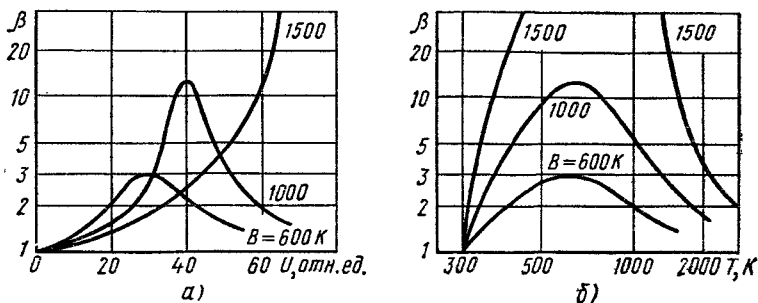


Рис. 12.2. Расчетные зависимости коэффициента нелинейности варисторов из материалов с различной величиной коэффициента температурной чувствительности B :

a — от напряжения; b — от температуры активной области

дифференциальное сопротивление, коэффициент нелинейности при этом тоже будет отрицательным.

На основе этих расчетов (рис. 12.2, б) можно заключить, что температура активных областей варистора может превышать температуру окружающей среды на несколько сотен градусов. Следовательно, для изготовления варисторов со стабильными параметрами необходим термостойкий материал. Именно поэтому при массовом производстве варисторов используют карбид кремния — один из самых термостойких материалов. Одновременно поликристаллический карбид кремния является очень дешевым материалом.

Основными примесями в техническом карбиде кремния являются азот и алюминий. Энергия ионизации этих примесей в карбиде кремния невелика (особенно при большой концентрации основных и компенсирующих примесей, что имеет место в техническом карбиде кремния), соответственно мала и величина коэффициента температурной чувствительности B . Поэтому коэффициент нелинейности варисторов не превышает 6, что ограничивает возможные применения варисторов.

Увеличение температуры окружающей среды должно привести к уменьшению коэффициента нелинейности (12.9) и незначительному сдвигу максимума кривой $\beta = f(U)$ по напряжению (рис. 12.3).

Вольт-амперная характеристика варистора, как было отмечено, должна соответствовать уравнениям (12.5) и (12.6). Если же варистор работает в узком диапазоне изменений напряжений и токов, то коэффициент нелинейности в этом диапазоне можно считать постоянным:

$$\beta = \frac{U}{I} \frac{dI}{dU} = \text{const.}$$

Тогда

$$\int \frac{dI}{I} = \beta \int \frac{dU}{U}; \quad \ln I = \beta \ln U + \ln A$$

и вольт-амперная характеристика варистора будет соответствовать уравнению

$$I = AU^\beta, \quad (12.11)$$

где A — коэффициент, величина которого зависит от типа варистора и от температуры.

Иногда вольт-амперные характеристики варисторов аппроксимируют уравнением

$$U = A_1 I^\alpha, \quad (12.12)$$

где $\alpha = 1/\beta$ и $A_1 = A^{-1/\beta} = A^{-\alpha}$.

Используя уравнения (12.11) и (12.12), статическое сопротивление варистора можно выразить как функцию тока или напряжения:

$$R(I) = A_1 I^{\frac{1}{\beta} - 1}; \quad (12.13)$$

$$R(U) = A_1^\beta U^{1-\beta}. \quad (12.14)$$

Температурные коэффициенты статического сопротивления, напряжения и тока. В связи с нелинейностью вольт-амперной характеристики следует различать температурные коэффициенты статического сопротивления варистора, измеренные при постоянном напряжении и при постоянном токе, а также температурные коэффициенты напряжения и тока. Из уравнений (12.11) — (12.14) с учетом температурного изменения коэффициентов A и A_1 можно получить:

$$\begin{aligned} \text{TKR} \Big|_{U=\text{const}} &= \frac{1}{R} \cdot \frac{\partial R}{\partial T} \Big|_{U=\text{const}} = \\ &= \text{TK} A_1 + \left(\frac{1}{\beta} - 1 \right) \text{TKI} \Big|_{U=\text{const}}; \end{aligned} \quad (12.15)$$

$$\begin{aligned} \text{TKR} \Big|_{I=\text{const}} &= \frac{1}{R} \cdot \frac{\partial R}{\partial T} \Big|_{I=\text{const}} = \\ &= \beta \text{TK} A_1 + (1 - \beta) \text{TKU} \Big|_{I=\text{const}}; \end{aligned} \quad (12.16)$$

$$\text{TKI} \Big|_{U=\text{const}} = \frac{1}{I} \cdot \frac{\partial I}{\partial T} \Big|_{U=\text{const}} = -\beta \text{TK} A_1; \quad (12.17)$$

$$\text{TKU} \Big|_{I=\text{const}} = \frac{1}{U} \cdot \frac{\partial U}{\partial T} \Big|_{I=\text{const}} = \text{TK} A_1. \quad (12.18)$$

При малых напряжениях на варисторе, когда коэффициент нелинейности $\beta = 1$, т. е. на линейном участке вольт-амперной характеристики

$$\text{TK} A_1 = \frac{1}{A_1} \cdot \frac{dA_1}{dT} = \text{TKR} \Big|_{U \approx 0} = -\frac{B}{T^2}. \quad (12.19)$$

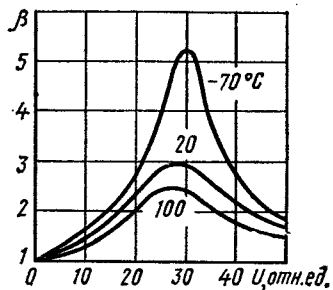


Рис. 12.3. Расчетная зависимость коэффициента нелинейности варистора от напряжения при различных температурах ($B = 600 \text{ К}$)

Используя уравнения (12.15) — (12.18), можно определить соотношения между различными температурными коэффициентами варистора:

$$\begin{aligned}TKR|_{U=\text{const}} &= \beta TKR|_{I=\text{const}}; \\TKI|_{U=\text{const}} &= -\beta TKU|_{I=\text{const}}.\end{aligned}$$

У варисторов, выпускаемых отечественной промышленностью, в диапазоне температур от -40 до $+100$ °C

$$\begin{aligned}-TKR|_{U=\text{const}} &= TKI|_{U=\text{const}} < 7 \cdot 10^{-3} \text{K}^{-1}; \\-TKR|_{I=\text{const}} &= -TKU|_{I=\text{const}} < 1,4 \cdot 10^{-3} \text{K}^{-1}.\end{aligned}$$

Из уравнений (12.15) — (12.19) и (12.10) следует, что у варисторов с большим значением максимального коэффициента нелинейности должны быть большие по абсолютной величине значения температурных коэффициентов сопротивления, тока и напряжения. Между максимальным коэффициентом нелинейности и температурными коэффициентами сопротивления, тока и напряжения должна быть взаимосвязь, так как они зависят от коэффициента температурной чувствительности B :

$$TKI|_{U=\text{const}} = \frac{4\beta_{\text{max}}(\beta_{\text{max}} - 1)}{(\beta_{\text{max}} + 1)T_0}. \quad (12.20)$$

Экспериментально такая взаимосвязь наблюдается, но расчет температурного коэффициента тока варистора по уравнению (12.20) дает несколько завышенные результаты. Численные расхождения можно объяснить, во-первых, наличием добавочных сопротивлений, включенных последовательно и параллельно сопротивлению активных областей варистора, и, во-вторых, тем, что температура среды, окружающей активные области варистора, несколько выше температуры среды, окружающей весь варистор.

Частотные свойства варисторов могут определяться либо инерционностью процессов, приводящих к нелинейности вольт-амперной характеристики, либо собственной емкостью варистора. Доминирующее значение с точки зрения частотных свойств варисторов имеет их собственная емкость.

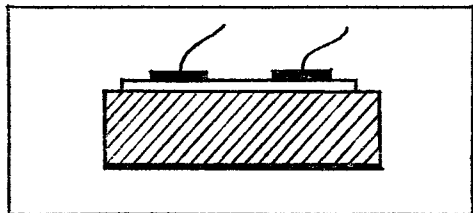
В качестве исходного материала для изготовления варисторов иногда используют окислы цинка и других металлов. Принцип действия таких оксидных варисторов имеет свои особенности. Поэтому свойства оксидных варисторов несколько отличаются от свойств варисторов из карбида кремния.

Контрольные вопросы

1. Какой физический процесс можно считать основным при работе варисторов в диапазоне средних и больших токов?
2. Какими уравнениями можно аппроксимировать вольт-амперную характеристику варистора? Когда удобнее использовать ту или иную форму аппроксимации?
3. Почему у варисторов с большим коэффициентом нелинейности получаются большие температурные коэффициенты сопротивления?

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ

13



§ 13.1. ПЕРЕКЛЮЧАТЕЛИ

Структура переключателей

Аморфные, или стеклообразные, полупроводники можно разделить на *оксидные, элементные и халькогенидные* (или бескислородные). Оксидные аморфные полупроводники или стекла состоят из окислов различных элементов. Элементные аморфные полупроводники или элементные стекла могут быть изготовлены из таких элементов, как селен, сера, теллур, германий и фосфор. Из элементных стекол наиболее широко применяется аморфный селен. Халькогенидные стекла состоят из сульфидов, селенидов и теллуридов различных элементов, т. е. из соединений, аналогичных окислам, в которых роль кислорода выполняет его аналоги — сера, селен и теллур.

Переключатели делают обычно на основе халькогенидных стекол, например, такого состава: $\text{Te}_{48}\text{As}_{30}\text{Ge}_{10}\text{Si}_{12}$. Халькогенидное стекло наносят на полированную подложку из графита методом вакуумного осаждения при испарении. Толщина пленки халькогенидного стекла от единиц до сотен микрометров. Одним электродом переключателя является графитовая подложка. Вторым электродом служит тонкий слой металла, нанесенного на пленку халькогенидного стекла также методом испарения в вакууме с последующей конденсацией.

Механизм переключения

Переключатель может находиться в двух состояниях: в закрытом состоянии с большим сопротивлением и в открытом состоянии с малым сопротивлением (рис. 13.1). Переключение в открытое состояние происходит при некотором напряжении переключения $U_{\text{н}}$. При переключении прибор обладает отрицательным дифференциальным сопротивлением (ОДС). В открытом состоянии переключатель

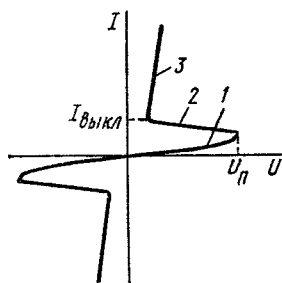


Рис. 13.1. Вольт-амперная характеристика переключателя на аморфном полупроводнике:

1 — участок, соответствующий закрытому состоянию переключателя; 2 — переходный участок; 3 — участок, соответствующий открытому состоянию переключателя

находится до тех пор, пока через него идет ток, превышающий ток выключения $I_{\text{выкл}}$.

Механизм переключения связан с разогревом пленки аморфного полупроводника под действием выделяющейся мощности. Явления, происходящие при переключении, аналогичны явлениям теплового пробоа (см. § 3.13) и явлениям в термисторах с отрицательным температурным коэффициентом сопротивления (см. § 11.1). Следовательно, отрицательное дифференциальное сопротивление и эффект переключения в пленках аморфных полупроводников будут существовать, если удельная проводимость аморфного полупроводника растет с увеличением температуры и если коэффициент температурной чувствительности $B > 4T_0$ (см. § 11.1).

В связи с аналогией в принципах действия переключателей на аморфных полупроводниках и термисторов вольт-амперная характеристика переключателя может быть аппроксимирована той же системой уравнений [(11.6) и (11.7)]. Разница в механизмах действия переключателей и термисторов связана со значительно меньшим объемом нагреваемой области в аморфной пленке переключателя — проводящего канала или шнура, образующегося при шнуровании тока (см. § 3.13). Проводящий канал возникает перед переключением прибора в открытое состояние и из-за разогрева имеет большую удельную проводимость, чем остальная пассивная часть пленки аморфного полупроводника.

Поэтому величины, входящие в уравнения [(11.6) и (11.7)], являются параметрами проводящего канала в пленке аморфного полупроводника, т. е. H — это коэффициент рассеяния проводящего канала; T — температура проводящего канала; R_∞ — коэффициент, зависящий от площади сечения проводящего канала, от толщины пленки аморфного полупроводника и от его свойств; B — коэффициент температурной чувствительности аморфного полупроводника.

В открытом состоянии переключателя, т. е. при малом сопротивлении проводящего канала, необходимо учитывать еще и сопротивление растекания в графитовой подложке под проводящим каналом. Так, при удельном сопротивлении графита $\rho = 0,2 \text{ Ом} \cdot \text{см}$ и радиусе сечения проводящего канала $r = 1 \text{ мкм}$ сопротивление растекания R_s в графитовой подложке, вычисленное по формуле (3.103), составляет 320 Ом. Такое сопротивление может существенно повлиять на вольт-амперную характеристику переключателя в открытом состоянии. С учетом падения напряжения на сопротивлении растекания вольт-амперная характеристика переключателя должна соответст-

вовать уравнению

$$U = \sqrt{HR_{\infty}(T - T_0) \exp \frac{B}{T}} + R_s I.$$

Результаты численного расчета по этому уравнению показывают, что значения тока и температуры проводящего канала при минимальном напряжении на переключателе в открытом состоянии в сильной степени зависят от величины сопротивления растекания в графитовой подложке. Например, без учета сопротивления растекания расчетные значения тока и температуры при минимальном напряжении оказываются равными соответственно 1165 мА и 4400 К, что является нереальным и не соответствует экспериментальным значениям. С учетом сопротивления растекания ($R_s = 250$ Ом) те же величины равны 2,2 мА и 430 К, что вполне допустимо. При расчетах в этом примере принято $B = 5250$ К, $HR_{\infty} = 3,7 \cdot 10^{-6}$ Ом·Вт·К⁻¹.

Удельная проводимость разогретого проводящего канала превышает удельную проводимость остальной пассивной части пленки аморфного полупроводника. Однако площадь поперечного сечения проводящего канала на несколько порядков меньше площади пассивной части пленки аморфного полупроводника между электродами. Поэтому полная проводимость пассивной части пленки может быть больше абсолютного значения дифференциальной проводимости, характерной для переходного участка вольт-амперной характеристики проводящего канала. В таком случае переключатель будет иметь так называемую у-образную вольт-амперную характеристику (см. § 3.13).

Таким образом, переключатели на аморфных полупроводниках могут иметь у-образные вольт-амперные характеристики, которые, на первый взгляд противоречат тепловому механизму переключения.

Переходный участок у-образной вольт-амперной характеристики нельзя экспериментально снять по точкам даже при использовании идеального генератора тока в качестве источника питания, т. е. источника питания с бесконечно большим собственным сопротивлением.

Характеристики и свойства

Напряжение переключения — это минимальное напряжение, при котором происходит переключение переключателя из закрытого в открытое состояние. Для различных переключателей на аморфных полупроводниках напряжение переключения составляет от единиц до десятков вольт.

Напряжение переключения переключателей на аморфных полупроводниках уменьшается с увеличением температуры окружающей среды так же, как и пробивное напряжение при тепловом пробое (см. § 3.13). Однако в переключателях с малой толщиной пленки аморфного полупроводника (несколько микрометров) из-за большой напряженности электрического поля тепловому пробоею может

предшествовать лавинный пробой. Пробивное напряжение при лавинном пробое растет с увеличением температуры окружающей среды (см. § 3.11). Поэтому у переключателей с малой толщиной пленки аморфного полупроводника может быть сложная зависимость напряжения переключения от температуры. Тем не менее механизм переключения из закрытого состояния в открытое связан только с тепловым пробоем.

Ток выключения — это минимальный ток, при котором переключатель еще находится в открытом состоянии.

При работе переключателя на переменном напряжении или в импульсном режиме необходимо учитывать инерционность процесса разогрева и охлаждения проводящего канала. Ток через переключатель, зависящий от температуры проводящего канала, будет запаздывать относительно приложенного напряжения, т. е. будет сдвинут по фазе. Однако тепловые постоянные времени в связи с малостью объема проводящего канала оказываются также очень малыми (10^{-6} — 10^{-8} с). Следовательно, переключатели на аморфных полупроводниках могут работать при частотах до десятков, а иногда и сотен мегагерц.

На время переключения влияет собственная емкость переключателя и емкость внешних элементов схемы. При переключении прибора из закрытого в открытое состояние разряд собственной емкости и емкостей, подключенных параллельно переключателю, происходит через проводящий канал, ускоряя его разогрев и уменьшая тем самым время переключения.

§ 13.2. ЗАПОМИНАЮЩИЕ УСТРОЙСТВА

Запоминающие устройства на аморфных полупроводниках имеют ту же структуру, что и переключатели, но в качестве аморфного полупроводника обычно используют халькогенидные стекла другого состава (например, $\text{Te}_{81}\text{Ge}_{15}\text{As}_4$).

Запоминающие устройства отличаются от переключателей тем, что открытое состояние запоминающего устройства может сохраняться даже в обесточенном положении, т. е. после отключения прибора от источника питания. Такой прибор обладает двумя вольт-амперными характеристиками, соответствующими закрытому и открытому состояниям (рис. 13.2).

Механизм переключения запоминающего устройства из закрытого в открытое состояние так же, как и в обычных переключателях на аморфных полупроводниках, связан с тепловым пробоем при шунировании тока и образовании проводящего канала. Особенность принципа действия запоминающих устройств на аморфных полупроводниках состоит в переходе всего или части материала проводящего канала из аморфного в кристаллический полупроводник при относительно медленном охлаждении проводящего канала после отключения прибора от источника питания.

Таким образом, эффект памяти или состояние памяти может наблюдаться у приборов, изготовленных на основе материала,

способного к фазовому переходу из аморфного в кристаллическое состояние с резким изменением при этом удельного сопротивления.

Чтобы процесс кристаллизации материала проводящего канала успел произойти за время его охлаждения, необходимо разогреть проводящий канал и прилегающие к нему области до необходимой для кристаллизации температуры, т. е. для переключения запоминающего устройства в открытое состояние с эффектом памяти необходим импульс тока большой длительности (обычно более 10^{-3} с) и достаточно большой амплитуды. Если это условие выполнено, то из-за относительно большого разогретого объема полупроводника его охлаждение будет происходить за время, достаточное для прохождения процесса кристаллизации материала в проводящем канале.

Для перевода запоминающего устройства из открытого в закрытое состояние необходим импульс тока относительно малой длительности (обычно 10^{-7} — 10^{-4} с), который за время своего действия успеет разогреть только проводящий канал малого объема. Из-за малости разогретого объема его охлаждение будет происходить быстро, за время, в течение которого материал проводящего канала не успевает закристаллизоваться. Тогда проводящий канал будет представлять собой аморфный полупроводник, что соответствует закрытому состоянию запоминающего устройства.

§ 13.3. НАДЕЖНОСТЬ, СТАБИЛЬНОСТЬ И СРОК СЛУЖБЫ ПРИБОРОВ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ

Надежность, стабильность и срок службы приборов на аморфных полупроводниках довольно низки по сравнению с другими полупроводниковыми приборами. Особенно низки эти параметры при прохождении через приборы на аморфных полупроводниках тока только в одном направлении.

Нестабильность обычно проявляется в уменьшении со временем напряжения переключения и с переходом прибора на аморфных полупроводниках в состояние памяти, т. е. в открытое состояние с эффектом памяти. Связано это явление с постепенной кристаллизацией аморфного полупроводника в проводящем канале, которая происходит тем быстрее, чем выше температура и напряженность поля в проводящем канале. При этих условиях облегчается дрейф ионов аморфного полупроводника в проводящем канале и их перегруппировка, способствующая кристаллизации.

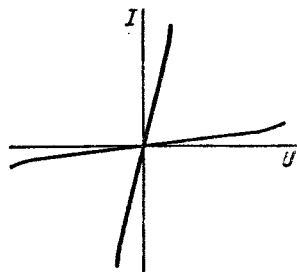


Рис. 13.2. Вольт-амперная характеристика запоминающего устройства (элемента памяти) на аморфном полупроводнике

Число переключений, которое может выдержать переключатель, зависит от величины тока, проходящего через открытый переключатель, и от длительности прохождения этого тока. Чем больше амплитуда и длительность импульсов тока, тем за меньшее число переключений происходит перегруппировка ионов в проводящем канале, приводящая к кристаллизации аморфного полупроводника в проводящем канале.

Максимальное число переключений, которое выдерживает переключатель до выхода его из строя, обычно определяется величиной 10^9 — 10^{12} . Для правильной оценки такого, казалось бы большого, числа срабатываний переключателя необходимо учесть, что время рассасывания современных транзисторов может составлять всего несколько наносекунд. Это означает, что транзистор, работающий в схеме переключателя, может, по крайней мере, за несколько часов выполнить более 10^{12} переключений.

Таким образом, малый срок службы приборов на аморфных полупроводниках, связанный с относительно большой ионной составляющей проводимости этих материалов при повышенных температурах, ставит под сомнение их перспективность и значительно ограничивает области применения таких приборов.

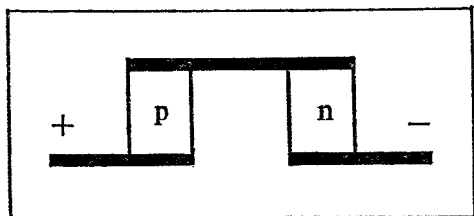
Для повышения стабильности параметров переключателей, их надежности и числа срабатываний до выхода из строя необходимо использовать полупроводниковые материалы с большим коэффициентом температурной чувствительности ($B > 4 T_0$) и одновременно с более высокой температурой плавления или размягчения, с меньшей ионной составляющей проводимости.

Контрольные вопросы

1. Почему и при каких условиях переключатели на аморфных полупроводниках могут находиться в двух состояниях — в закрытом и в открытом?
2. Почему у некоторых переключателей не удается исследовать переходный участок вольт-амперной характеристики, соответствующий отрицательному дифференциальному сопротивлению?
3. Как может изменяться при изменении температуры напряжение переключения переключателей на аморфных полупроводниках?
4. Почему в запоминающих устройствах на аморфных полупроводниках открытое состояние может сохраняться даже после отключения прибора от источника питания?

ПОЛУПРОВОДНИКОВЫЕ ТЕРМОЭЛЕКТРИЧЕСКИЕ ПРИБОРЫ

14



§ 14.1. ПРИНЦИП ДЕЙСТВИЯ

В полупроводниковых термоэлектрических приборах используют *термоэлементы*, состоящие из двух последовательно соединенных ветвей, изготовленных из различных материалов. Обычно одна ветвь, называемая *p*-ветвью или положительной ветвью, имеет электропроводность *p*-типа; другая, называемая *n*-ветвью или отрицательной ветвью, — электропроводность *n*-типа (рис. 14.1). Ветви термоэлемента электрически соединены между собой металлическими пластинами. Контакт ветви термоэлемента с металлической пластиной называют спаем.

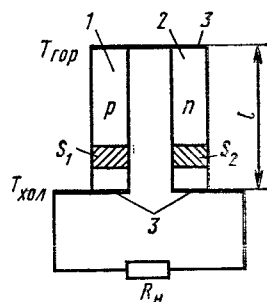


Рис. 14.1. Схема полупроводникового термоэлемента с сопротивлением нагрузки:

- 1 — положительная ветвь;
- 2 — отрицательная ветвь;
- 3 — соединительные металлические пластины горячего и холодного спаев

Возникновение термо-э.д.с. [эффект Зеебека]

При различной температуре спаев в цепи с термоэлементом появляется термоэлектродвижущая сила (термо-э. д. с.), которая состоит из трех составляющих.

Первая составляющая термо-э. д. с. обусловлена диффузией носителей заряда от горячих спаев к холодным. Диффузия носителей заряда в ветвях термоэлемента может возникать по двум причинам. Во-первых, в ветвях термоэлемента у горячего спаива оказывается большее число ионизированных примесей по сравнению с холодными концами ветвей. При дополнительной ионизации примесей на горячих концах ветвей увеличивается концентрация соответствующих носителей. Таким образом, первая причина диф-

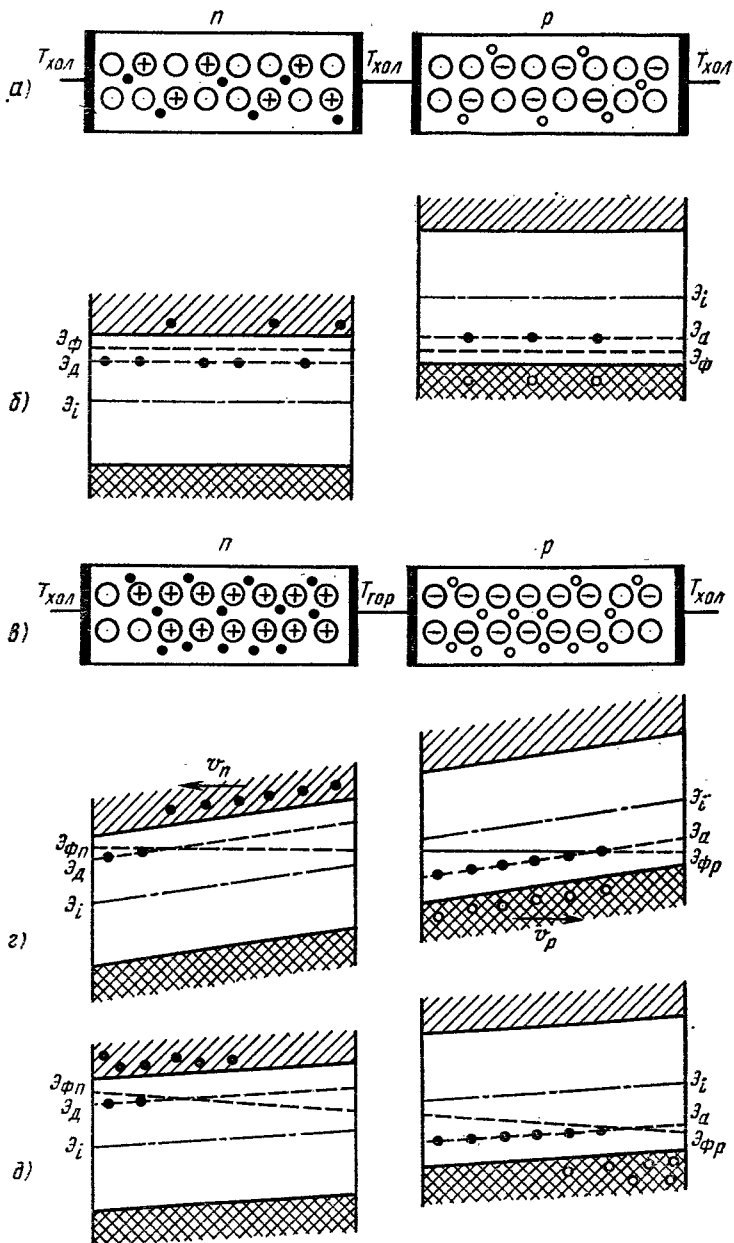


Рис. 14.2. Появление термо-э. д. с. в результате диффузии носителей заряда по ветвям термоэлемента;

a — спаи термоэлемента находятся при одной температуре $T_{хол}$ (не все доноры и акцепторы ионизированы); *б* — энергетические диаграммы ветвей термоэлемента в состоянии термодинамического равновесия (при $T_{хол}$); *в* — спаи термоэлемента находятся при разных температурах (около горячих спаев все примеси ионизированы); *г* — энергетические диаграммы ветвей термоэлемента при сопротивлении нагрузки $R_H = 0$ (идет ток, равный току короткого замыкания термоэлемента); *д* — энергетические диаграммы при сопротивлении нагрузки $R_H = \infty$ (возникла термо-э. д. с.)

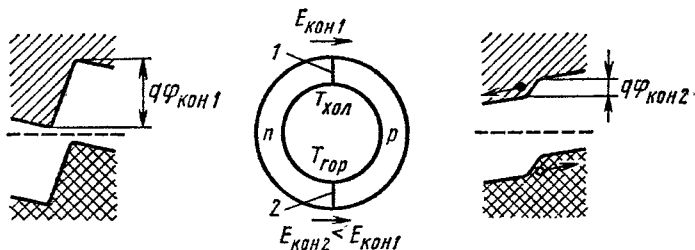


Рис. 14.3. Появление термо-э. д. с. из-за температурной зависимости контактной разности потенциалов при коротком замыкании ветвей термоэлемента и энергетические диаграммы спаев, находящихся при разных температурах

фузии носителей — возникновение градиента концентрации носителей (рис. 14.2).

Во-вторых, если в ветвях термоэлемента при температуре, соответствующей температуре холодных спаев, все примеси уже ионизированы (температура выше температуры истощения примесей), то исходная концентрация носителей заряда на горячем и холодном конце каждой ветви одинакова. Но на горячих концах ветвей носители приобретают большие энергии. Поэтому опять будет происходить диффузия носителей заряда от горячего к холодному концу в каждой ветви, связанная с выравниванием средней энергии, приходящейся на один носитель определенного знака.

Диффузия, например, электронов в n -ветви может происходить только от горячего конца именно n -ветви к холодному и не может происходить в p -ветвь, так как переходу электронов в p -ветвь препятствует потенциальный барьер горячего спая термоэлемента. Аналогично дырки могут диффундировать только по p -ветви к холодному ее концу.

Перемещение носителей заряда, связанное с диффузией этих носителей, нарушает электрическую нейтральность горячих и холодных концов каждой ветви — на горячих концах ветвей остаются нескомпенсированные ионизированные примесные атомы, а на холодных концах ветвей образуется избыток основных носителей заряда. В результате возникает первая составляющая термо-э. д. с., которую можно назвать диффузионной, так как она возникает из-за процессов диффузии.

Вторая составляющая термо-э. д. с. — это следствие температурной зависимости контактной разности потенциалов. Если оба спаи термоэлемента имеют одну и ту же температуру, то контактные разности потенциалов на этих спаих равны, направлены в противоположные стороны при обходе контура с термоэлементом и не дают результирующей термо-э. д. с. Если же температура спаев различна, то величина контактной разности потенциалов будет также различна (рис. 14.3). Поэтому в цепи термоэлемента появляется вторая составляющая термо-э. д. с. с той же полярностью, что и первая составляющая.

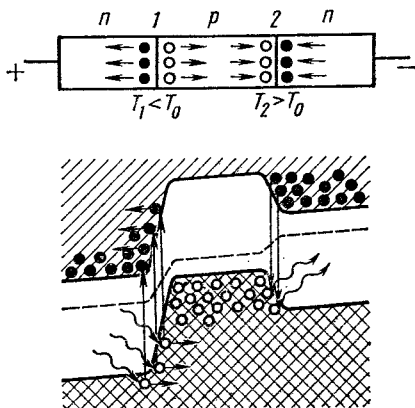


Рис. 14.4. Охлаждение и нагрев спаев полупроводниковых термоэлементов при прохождении тока и энергетическая диаграмма, поясняющая эти явления

лодного $T_{\text{хол}}$ спаев и от электрофизических свойств материалов, образующих ветви термоэлемента. В небольшом диапазоне температур термо-э. д. с. \mathcal{E} можно считать (с достаточной для практических целей точностью) пропорциональной разности температур и некоторому коэффициенту $\alpha_{1,2}$, называемому коэффициентом термо-э. д. с.:

$$\mathcal{E} = \alpha_{1,2} (T_{\text{гор}} - T_{\text{хол}}). \quad (14.1)$$

Поглощение и выделение тепла в спаих термоэлемента (эффект Пельтье)

При прохождении постоянного тока через термоэлемент в спаих его ветвей поглощается или выделяется (в зависимости от направления тока) некоторое количество тепла, пропорциональное времени, току и коэффициенту Пельтье:

$$Q_{\text{П}} = \pm \Pi_{1,2} I t. \quad (14.2)$$

Величина коэффициента Пельтье зависит от электрофизических свойств материалов ветвей термоэлемента.

Особенно сильно явление Пельтье, как и явление Зеебека, проявляется в полупроводниковых термоэлементах, состоящих из ветвей с электропроводностью *n*- и *p*-типа. Если направление тока в цепи термоэлементов такое, как показано на рис. 14.4, то на спае *I* свободные электроны и дырки, возникшие в результате тепловой генерации, двигаются в разные стороны под действием суммарного электрического поля (диффузионного поля спаия и внешнего поля). При тепловой генерации носителей в области спаия *I* на переброс электрона из валентной зоны в зону проводимости затрачивается некоторая тепловая энергия кристаллической решетки полупро-

Третья составляющая термо-э. д. с. возникает в термоэлементе вследствие увлечения электронов (дырок) фононами, т. е. квантами тепловой энергии. Если в ветвях термоэлемента есть градиент температуры, то будет существовать направленное движение фононов от горячего спаия к холодному. В результате столкновений фононов с носителями заряда фононы увлекают за собой в *n*-ветви электроны, а в *p*-ветви — дырки. Этот эффект может оказаться преобладающим при низких температурах.

Результирующая термо-э. д. с., состоящая из трех рассмотренных составляющих, зависит от температуры горячего $T_{\text{гор}}$ и холодного $T_{\text{хол}}$ спаев.

водника. Поэтому при прохождении тока спай 1 будет охлаждаться, а спай 2 будет нагреваться, так как к нему с разных сторон подходят электроны и дырки, которые отдают при рекомбинации некоторую энергию в виде тепла кристаллической решетке полупроводника.

В результате при прохождении тока термоэлемент работает как своеобразный тепловой насос, забирая тепловую энергию на спае 1 и выделяя ее на спае 2.

Рассмотренные явления, приводящие к охлаждению спае 1 и к нагреву спае 2 при указанной полярности приложенного внешнего напряжения, будут наблюдаться только в том случае, если генерация и рекомбинация носителей происходит непосредственно на спае или вблизи спаев. Это условие справедливо для полупроводниковых термоэлементов и может не выполняться в других полупроводниковых приборах (например, в коллекторном электронно-дырочном переходе биполярного транзистора).

Эффект Пельтье является обратным эффекту Зеебека. Поэтому для одного и того же термоэлемента существует связь между коэффициентом Пельтье и коэффициентом термо-э. д. с.:

$$\Pi_{1,2} = \alpha_{1,2} T. \quad (14.3)$$

Это соотношение может быть получено при применении к термоэлектрическим явлениям первого и второго закона термодинамики.

§ 14.2. ТЕРМОЭЛЕКТРИЧЕСКИЕ ГЕНЕРАТОРЫ

Термоэлектрический генератор (ТЭГ) — это агрегат энергопитания, состоящий из источника тепловой энергии (горелка, атомный реактор и т. п.), термоэлектрической батареи, токопроводящих цепей и радиатора, охлаждающего холодные спаи термоэлементов. Термоэлектрическая батарея — это совокупность электрически соединенных термоэлементов. Термоэлементы могут быть соединены между собой последовательно, параллельно и смешанно.

Коэффициент полезного действия термоэлемента

Термоэлектрические генераторы предназначены для непосредственного преобразования тепловой энергии в электрическую. Поэтому основным параметром термоэлектрического генератора является коэффициент полезного действия, т. е. отношение полезной мощности, выделяемой в нагрузку генератора, к количеству тепла, поступающего в единицу времени на горячие спаи термоэлементов.

Рассмотрим работу отдельного термоэлемента в качестве преобразователя тепловой энергии в электрическую (см. рис. 14.1).

Если удельные сопротивления ветвей термоэлемента ρ_1 и ρ_2 , то полное сопротивление термоэлемента

$$R = \frac{\rho_1 l}{S_1} + \frac{\rho_2 l}{S_2}. \quad (14.4)$$

Переходными сопротивлениями в местах спаев ветвей термоэлемента с металлическими соединительными пластинами пренебрегаем по сравнению с объемным сопротивлением ветвей термоэлемента.

Полную теплопроводность термоэлемента можно выразить через удельные теплопроводности ветвей K_1 и K_2 :

$$K = \frac{K_1 S_1}{l} + \frac{K_2 S_2}{l}. \quad (14.5)$$

В результате теплопроводности от горячего спая или от источника тепловой энергии к холодному спаю в единицу времени передается тепло

$$Q_K = K (T_{\text{гор}} - T_{\text{хол}}). \quad (14.6)$$

Используя выражение (14.1), для тока, протекающего в цепи термоэлемента, можно записать:

$$I = \frac{\alpha_{1,2} (T_{\text{гор}} - T_{\text{хол}})}{R + R_n}. \quad (14.7)$$

При этом в нагрузке выделяется полезная мощность

$$P_n = I^2 R_n = \frac{\alpha_{1,2}^2 (T_{\text{гор}} - T_{\text{хол}})^2 R_n}{(R + R_n)^2}. \quad (14.8)$$

Ток, проходящий в цепи термоэлемента и обусловленный эффектом Зеебека, вызывает, во-первых, выделение и поглощение тепла из-за эффекта Пельтье на спаях термоэлемента и, во-вторых, выделение джоулева тепла в ветвях термоэлемента.

Направление тока в термоэлементе соответствует поглощению на горячем спае в единицу времени теплоты

$$Q_{\text{п.гор}} = \Pi_{1,2} I = \alpha_{1,2} I T_{\text{гор}}. \quad (14.9)$$

Эффект Пельтье, заключающийся в перекачивании некоторого количества тепла от горячего спая к холодному, как бы увеличивает теплопроводность термоэлемента. Теплота, отдаваемая холодному спаю,

$$Q_{\text{п.хол}} = \alpha_{1,2} I T_{\text{хол}}. \quad (14.10)$$

Из-за разных значений $T_{\text{гор}}$ и $T_{\text{хол}}$ величины $Q_{\text{п.гор}}$ и $Q_{\text{п.хол}}$ также различны. Вычитая (14.10) из (14.9) и учитывая (14.1), получим

$$Q_{\text{п.гор}} - Q_{\text{п.хол}} = \alpha_{1,2} I (T_{\text{гор}} - T_{\text{хол}}) = \xi I = I^2 (R + R_n). \quad (14.11)$$

Таким образом, вся электрическая мощность, вырабатываемая термоэлементом, — это разница теплоты, поглощенной на горячем спае, и теплоты, перенесенной на холодный спай. Одна часть этой электрической мощности ($I^2 R_n$) выделяется в виде полезной мощности в нагрузке, другая часть ($I^2 R$) — в виде джоулева тепла в ветвях термоэлемента. В первом приближении можно считать, что джоулево тепло распределяется пополам между холодным и

горячим спаями, т. е. к горячему спаю или источнику тепловой энергии возвращается мощность

$$\frac{1}{2} Q_{\text{Дж}} = \frac{1}{2} I^2 R = \frac{\alpha_{1,2}^2 (T_{\text{гор}} - T_{\text{хол}})^2 R}{2(R - R_{\text{н}})^2} \quad (14.12)$$

Определив все количество тепла, поступающее в единицу времени на горячий спай термоэлемента от источника тепловой энергии, и учтя часть джоулева тепла, возвращенную горячему спаю при прохождении тока по цепи термоэлемента, коэффициент полезного действия термоэлемента можно записать следующим образом:

$$\eta = \frac{P_{\text{н}}}{Q_{\text{К}} + Q_{\text{П. гор}} - \frac{1}{2} Q_{\text{Дж}}} \quad (14.13)$$

Если учесть (14.6), (14.8), (14.9) и (14.12), то

$$\eta = \frac{\frac{\alpha_{1,2}^2 (T_{\text{гор}} - T_{\text{хол}})^2 R_{\text{н}}}{(R + R_{\text{н}})^2}}{K(T_{\text{гор}} - T_{\text{хол}}) + \frac{\alpha_{1,2}^2 T_{\text{гор}} (T_{\text{гор}} - T_{\text{хол}})}{R + R_{\text{н}}} - \frac{\alpha_{1,2}^2 (T_{\text{гор}} - T_{\text{хол}})^2 R}{2(R + R_{\text{н}})^2}} \quad (14.14)$$

Обозначим отношение сопротивлений

$$R_{\text{н}}/R = m.$$

Тогда

$$R_{\text{н}} = mR = ml(\rho_1/S_1 + \rho_2/S_2) \quad (14.15)$$

$$\text{и} \quad R + R_{\text{н}} = (m + 1)R = (m + 1)l(\rho_1/S_1 + \rho_2/S_2). \quad (14.16)$$

Используя (14.5), (14.15) и (14.16), преобразуем (14.14):

$$\eta = \frac{T_{\text{гор}} - T_{\text{хол}}}{T_{\text{гор}}} \cdot \frac{m/(m + 1)}{1 + \frac{KR(m + 1)}{\alpha_{1,2}^2 T_{\text{гор}}} - \frac{T_{\text{гор}} - T_{\text{хол}}}{2(m + 1)T_{\text{гор}}}} \quad (14.17)$$

Первый множитель в (14.17) представляет собой коэффициент полезного действия обратимой тепловой машины, второй множитель характеризует уменьшение к. п. д. вследствие необратимых потерь на теплопроводность и джоулево тепло в термоэлементе.

В величину KR в знаменателе (14.17) входят кроме параметров материала ветвей термоэлемента ρ_1 и ρ_2 , K_1 и K_2 , площади поперечного сечения ветвей термоэлемента S_1 и S_2 [(14.4) и (14.5)]. Для обеспечения наибольшего к. п. д. (при данных $T_{\text{гор}}$, $T_{\text{хол}}$, $\alpha_{1,2}$, m , ρ и K ветвей) необходимо выбрать площади поперечных сечений ветвей так, чтобы величина KR была минимальной. Для этого продифференцируем

$$KR = (K_1 S_1 + K_2 S_2) \left(\frac{\rho_1}{S_1} + \frac{\rho_2}{S_2} \right) = \left(\frac{K_1 S_1}{S_2} + K_2 \right) \left(\frac{\rho_1 S_2}{S_1} + \rho_2 \right) \quad (14.18)$$

по (S_1/S_2) и приравняем нулю производную. Функция (14.18) имеет минимальное значение при оптимальном отношении площадей поперечных сечений

$$(S_1/S_2)_{\text{опт}} = \sqrt{\rho_1 K_2 / (\rho_2 K_1)}. \quad (14.19)$$

При этом

$$(KR)_{\min} = (\sqrt{K_1\rho_1} + \sqrt{K_2\rho_2})^2. \quad (14.20)$$

Величину $\alpha_{1,2}^2/KR$, обратную величине, входящей в знаменатель выражения (14.17), с учетом оптимального отношения площадей поперечного сечения ветвей (14.19) обычно обозначают Z и называют эффективностью или добротностью термoeлементa:

$$Z = \frac{\alpha_{1,2}^2}{(\sqrt{K_1\rho_1} + \sqrt{K_2\rho_2})^2}. \quad (14.21)$$

Таким образом, к. п. д. термoeлементa оптимальных размеров зависит от следующих факторов: 1) эффективности термoeлементa Z , определяемой только электрофизическими параметрами полупроводниковых материалов ветвей термoeлементa; 2) разности температур горячего и холодного спаев термoeлементa; 3) отношения сопротивления нагрузки к сопротивлению термoeлементa.

Для нахождения максимального к. п. д. необходимо выбрать оптимальное отношение сопротивления нагрузки к сопротивлению термoeлементa. Дифференцируя выражение для к. п. д. (14.17) по m и приравнявая нулю производную, получим:

$$m_{\text{opt}} = (R_H/R)_{\text{opt}} = \sqrt{1 + Z\bar{T}}, \quad (14.22)$$

где $\bar{T} = (T_{\text{гор}} - T_{\text{хол}})/2$ — средняя температура горячего и холодного спаев термoeлементa.

Подставляя (14.21) и (14.22) в общую формулу (14.17), находим максимальное значение к. п. д. термoeлементa, зависящее лишь от температуры спаев в эффективности термoeлементa:

$$\eta_{\max} = \frac{T_{\text{гор}} - T_{\text{хол}}}{T_{\text{гор}}} \cdot \frac{m_{\text{opt}} - 1}{m_{\text{opt}} + (T_{\text{хол}}/T_{\text{гор}})}. \quad (14.23)$$

Из соотношений (14.23) и (14.22) видно, что при определенных температурах $T_{\text{гор}}$ и $T_{\text{хол}}$ и при $Z \rightarrow \infty$ $\eta_{\max} \rightarrow \frac{T_{\text{гор}} - T_{\text{хол}}}{T_{\text{гор}}}$, т. е. максимальный к. п. д. стремится к к. п. д. идеальной тепловой машины (цикл Карно). Как видно, приближение максимального к. п. д. термoeлементa к к. п. д. идеальной тепловой машины требует больших значений эффективности термoeлементa Z .

Часто от термoeлементa желательно получить максимальную полезную мощность даже в ущерб к. п. д. При таком режиме работы необходимо, чтобы сопротивление нагрузки было равно сопротивлению термoeлементa, т. е. $R_H/R = m = 1$. При этом

$$\eta \Big|_{m=1} = \frac{1}{2} \cdot \frac{T_{\text{гор}} - T_{\text{хол}}}{T_{\text{гор}} + \frac{2}{Z} - \frac{1}{4}(T_{\text{гор}} - T_{\text{хол}})}. \quad (14.24)$$

В большинстве случаев максимальный к. п. д. (14.23) и к. п. д. в режиме максимальной отдаваемой мощности (14.24) отличаются друг от друга на 2—3% (относительных).

Основные требования к материалам для термоэлементов

В соответствии с (14.23) и (14.24) для термоэлементов целесообразно использовать материалы, обеспечивающие большое значение эффективности Z и выдерживающие наибольшую температуру $T_{гор}$, точнее — с наибольшим значением ZT (Z — среднее значение эффективности материала в диапазоне температур от $T_{хол}$ до $T_{гор}$).

Как видно из (14.21), величина эффективности термоэлемента Z является функцией коэффициента термо-э. д. с. $\alpha_{1,2}$, удельных сопротивлений ρ_1 и ρ_2 и удельных теплопроводностей κ_1 и κ_2 материалов ветвей термоэлемента. Все эти параметры зависят от концентрации носителей заряда (электронов или дырок). Удельное сопротивление обратно пропорционально концентрации носителей (рис. 14.5). Удельная теплопроводность складывается из теплопроводности кристаллической решетки κ_r и электронной теплопроводности $\kappa_{эл}$: $\kappa = \kappa_r + \kappa_{эл}$. Удельная теплопроводность решетки в первом приближении не зависит от концентрации носителей, а удельная электронная теплопроводность пропорциональна концентрации носителей заряда.

У термоэлементов из металлов или сплавов (термопары) величина эффективности Z мала из-за низкого коэффициента термо-э. д. с. и больших значений удельной теплопроводности. У термоэлементов из диэлектрических материалов эффективность Z стремится к нулю из-за больших значений удельного сопротивления. У термоэлементов из материалов со средней концентрацией носителей заряда, что соответствует полупроводниковым материалам, эффективность Z достигает максимального значения. Поэтому полупроводниковые термоэлементы экономичнее металлических. Максимум эффективности Z , как показывают расчеты, лежит в диапазоне концентраций носителей заряда $(2-3) \cdot 10^{19} \text{ см}^{-3}$. Такие концентрации на три порядка меньше концентрации носителей в металлах.

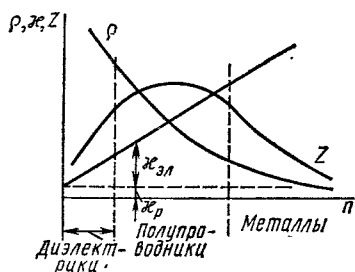


Рис. 14.5. Зависимости удельного сопротивления ρ , теплопроводности κ и эффективности термоэлемента Z от концентрации носителей заряда

Конструкция, свойства и параметры термоэлектрических батарей

Коэффициент термо-э. д. с. большинства термоэлементов ($\alpha_{1,2} = \alpha_1 - \alpha_2$) составляет 300—500 мкВ/К. Учитывая, что почти все термоэлектрические генераторы работают при разности температур спаев, не превышающей 300—600 К, от одного термоэлемента можно получить термо-э. д. с. 0,15—0,30 В. При этом согласно (14.24) оптимальные условия получения электрической мощности на наг-

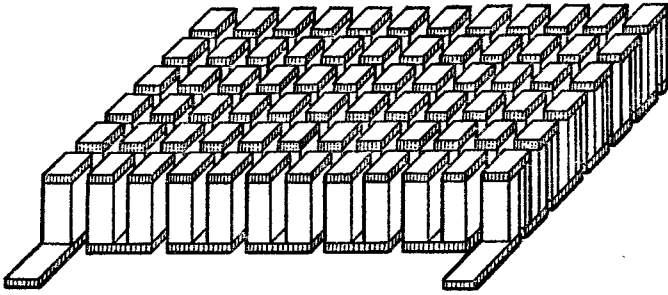


Рис. 14.6. Внешний вид одного из вариантов термобатарей

рузке достигаются при сопротивлении нагрузки, сравнимом с сопротивлением термоэлемента, т. е. около нескольких миллиом. Для получения значительных напряжений и для согласования с реальными нагрузочными сопротивлениями необходимо последовательное соединение многих термоэлементов.

В большинстве ТЭГ термоэлектрическая батарея представляет собой плоскую конструкцию, состоящую из нескольких десятков термоэлементов (рис. 14.6). По форме термоэлементы могут быть разнообразными в зависимости от конструктивных особенностей источника тепловой энергии. Наиболее распространены термоэлементы из ветвей в виде параллелепипедов с различным соотношением площади поперечного сечения и высоты.

Соединение отдельных ветвей между собой металлическими перемычками должно быть сделано так, чтобы контактные сопротивления были намного меньше сопротивления ветвей термоэлемента. Иначе контактные сопротивления могут уменьшать эффективность термоэлемента из-за возрастания общего сопротивления термоэлемента при неизменной величине термо-э. д. с.

В качестве источников тепловой энергии в ТЭГ применяют атомные реакторы, источники тепла с радиоактивными изотопами, используют химические реакции горения или тепловую энергию, излучаемую Солнцем.

В настоящее время существуют три основные области применения термоэлектрических генераторов: для катодной защиты магистральных газо- и нефтепроводов от коррозии, для питания космических аппаратов электрической энергией и для питания электрической энергией навигационных и метеорологических установок.

§ 14.3. ХОЛОДИЛЬНИКИ И ТЕПЛОВЫЕ НАСОСЫ

Максимальное снижение температуры холодных спаев термоэлемента

При прохождении тока через термоэлементы наряду с поглощением на спае I (см. рис. 14.4) в единицу времени некоторого количества тепла

$$Q_{II} = - \Pi_{1,2} I \quad (14.25)$$

происходит еще выделение джоулева тепла, пропорционального квадрату силы тока и сопротивлению спая I . В первом приближении можно считать, что сопротивления спаев 1 и 2 равны, т. е. каждое из них составляет половину полного сопротивления термоэлемента R . Тогда джоулево тепло, выделяющееся на первом спаяе,

$$Q_{Дж} = \frac{1}{2} I^2 R. \quad (14.26)$$

Алгебраическая сумма соотношений (14.25) и (14.26) для первого спая

$$Q = Q_{П} + Q_{Дж} = -\Pi_{1,2} I + \frac{1}{2} I^2 R. \quad (14.27)$$

Графическое пояснение уравнения (14.27) дано на рис. 14.7.

При токе, равном нулю, оба эффекта отсутствуют и охлаждения не происходит, а при достаточно большом токе $Q_{Дж}$ превысит $Q_{П}$, и охлаждение первых спаев перейдет в нагрев. Поэтому существует оптимальный ток, при котором эффект охлаждения максимален (от первого спая отводится максимальное количество тепла). Дифференцированием (14.27) находим, что Q достигает максимума при оптимальном токе:

$$I_{\text{опт}} = \Pi_{1,2} / R. \quad (14.28)$$

При этом токе

$$Q_{I \text{ опт}} = -\Pi_{1,2}^2 / (2R). \quad (14.29)$$

Из (14.29) следует, что чем меньше сопротивление термоэлемента, тем больше количество тепла $Q_{I \text{ опт}}$, отводимое от холодного спая. Однако из этого не следует, что, увеличивая площадь поперечного сечения или уменьшая длину ветвей термоэлемента, можно достичь более низких температур на холодном спаяе, так как одновременно в такой же мере возрастают теплопроводность и поток тепла по ветвям термоэлемента от горячего спая к холодному. Температура холодного спая будет понижаться до тех пор, пока количество тепла $Q_{К}$, переносимое от горячего спая к холодному в результате теплопроводности ветвей термоэлемента, не сбалансирует отводимое от спая I количество тепла Q_I при токе I . (В данном случае не учитываем поток тепла из окружающей среды на холодный спай термоэлемента, т. е. предполагаем, что теплоизоляция холодного спая идеальна).

Условием теплового баланса, следовательно, будет

$$-Q_I = Q_K = K (T_{\text{гор}} - T_{\text{хол}})$$

или

$$T_{\text{гор}} - T_{\text{хол}} = -Q_I / K.$$

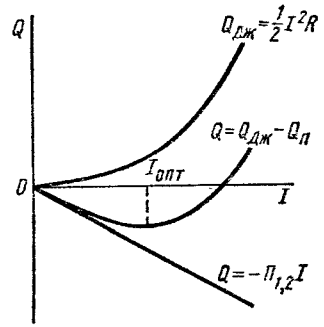


Рис. 14.7. Зависимость количества тепла Q , поглощаемого на холодном спаяе термоэлемента, от величины тока, проходящего по термоэлементу

В оптимальном режиме работы термоэлемента, т. е. с учетом (14.29), получим

$$(T_{\text{гор}} - T_{\text{хол}})_{\text{max}} = -Q_{I \text{ опт}}/K = \Pi_{1,2}^2/(2KR). \quad (14.30)$$

Используя (14.3) и (14.21), находим

$$(T_{\text{гор}} - T_{\text{хол}})_{\text{max}} = \frac{\alpha_{1,2}^2}{KR} \cdot \frac{1}{2} T_{\text{хол}}^2 = \frac{1}{2} Z T_{\text{хол}}^2. \quad (14.31)$$

Таким образом, требования, которым должны удовлетворять параметры полупроводниковых материалов ветвей термоэлемента, работающего в холодильнике, сводятся к достижению возможно большего значения эффективности Z (как и для термоэлемента, работающего в термоэлектрическом генераторе).

Применение полупроводниковых холодильников

Полупроводниковые холодильники с достаточной эффективностью применяют в радиоэлектронике, медицине, сельском хозяйстве, метеорологии и для бытовых целей.

Однако экономичность полупроводниковых холодильников при охлаждении объемов более десятков литров ниже экономичности существующих компрессионных холодильных машин. Поэтому полупроводниковые холодильники пока могут конкурировать с компрессионными только при охлаждении малых объемов.

Использование принципа полупроводникового теплового насоса

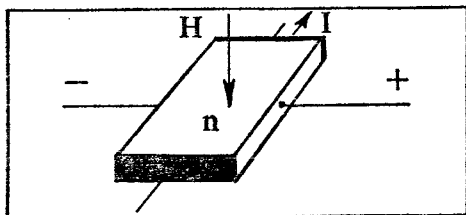
В термоэлектрической батарее при прохождении через нее постоянного тока происходит охлаждение одних и нагрев других спаев из-за поглощения и выделения теплоты на соответствующих спаях в связи с эффектом Пельтье. Во-первых, это явление оказывается удобным для использования в термостатах: простое изменение направления тока позволяет либо охлаждать, либо нагревать спаи термоэлементов, находящиеся внутри термостата.

Во-вторых, в отличие от прямого подогрева джоулевым теплом электрическая энергия в термоэлементе служит средством переноса тепловой энергии от холодных к горячим спаям. Таким образом, на горячих спаях может выделяться большее количество тепла, чем было затрачено электрической энергии. Так, при разности температур $T_{\text{гор}} - T_{\text{хол}} = 10 \text{ К}$ на каждый ватт выделяемого на горячих спаях тепла необходимо затратить не более 0,2 Вт электрической энергии, при разности температур 20 К — 0,3 Вт, а при

разности температур $30\text{ К} \text{---} 0,5\text{ Вт}$. Таким образом, принцип теплового насоса может оказаться экономически целесообразным для обогрева и охлаждения жилых и производственных помещений. Но это возможно только при небольшой стоимости термоэлементов, когда выигрыш в расходе энергии компенсирует повышенную стоимость термоэлектрического подогревателя по сравнению с обычным электрическим нагревательным устройством.

Контрольные вопросы

1. По каким причинам в полупроводниковых термоэлементах может возникать термо-э. д. с.?
2. Что необходимо для получения максимального к. п. д. термоэлектрического генератора?
3. Каков принцип действия полупроводниковых холодильников?
4. Как можно использовать эффект Пельтье для подогрева и почему этот эффект не используют для обогрева жилых и производственных помещений?



§ 15.1. ПРИНЦИП ДЕЙСТВИЯ

Принцип действия различных полупроводниковых гальваномагнитных приборов основан на двух эффектах: на *эффекте Холла*, т. е. на эффекте возникновения поперечной разности потенциалов в полупроводнике, по которому проходит электрический ток, в том случае, когда есть магнитное поле, перпендикулярное направлению тока, и на *магниторезистивном эффекте*, т. е. на эффекте изменения электрического сопротивления полупроводника под действием магнитного поля (эффект Гаусса). Оба эффекта вызваны тем, что на движущийся со скоростью v носитель заряда в магнитном поле с индукцией B действует сила Лоренца

$$F_{\text{л}} = q [vB]. \quad (15.1)$$

Рассмотрим подробнее эти эффекты.

Эффект Холла

Предположим, что в пластинке полупроводника, находящейся в магнитном поле, идет ток, обусловленный движением только электронов (рис. 15.1, а). Пренебрежем пока статистическим разбросом электронов по скоростям. Тогда сила Лоренца будет смещать движущиеся электроны к левой грани пластинки полупроводника. Направление смещения определяется направлением силы Лоренца, т. е. векторным произведением (15.1) с учетом знака носителей, или правилом левой руки, относящимся к техническому направлению тока. В результате смещения движущихся электронов между боковыми гранями пластинки полупроводника возникает э. д. с. Холла.

В полупроводнике с электропроводностью p -типа при том же техническом направлении тока вектор скорости дырок направлен противоположно вектору скорости электронов, знак носителей

заряда также другой. Поэтому сила Лоренца (15.1) действует на дырки в ту же сторону, смещая их также к левой грани пластинки полупроводника (рис. 15.1, б). Полярность э. д. с. Холла при этом получается другой.

Накопление носителей заряда у боковой грани пластинки полупроводника прекратится, когда сила Лоренца уравновесится силой холловского электрического поля. При перпендикулярном направлении напряженности магнитного поля к поверхности пластинки полупроводника условием такого динамического равновесия будет равенство:

$$qvB = qE_x. \quad (15.2)$$

Считая холловское электрическое поле однородным и учитывая геометрические размеры пластинки полупроводника, можно записать для э. д. с. Холла, т. е. для поперечной разности потенциалов между боковыми гранями пластинки полупроводника с электропроводностью p -типа:

$$\mathcal{E}_x = E_x a = vBa.$$

Значение скорости дырок можно определить из формулы для тока:

$$I = JS = qr\mu_p ES = qrvad.$$

Тогда

$$\mathcal{E}_x = \frac{1}{qp} \cdot \frac{IB}{\delta} = X \frac{IB}{\delta}, \quad (15.3)$$

где $X = 1/(qp)$ — коэффициент Холла для полупроводника с электропроводностью p -типа.

В действительности носители заряда в полупроводнике распределены по скоростям. Это распределение зависит от преобладающего механизма рассеяния носителей в конкретном полупроводнике. Поэтому более точное значение коэффициента Холла отличается от имеющегося в выражении (15.3) множителем A :

$$X = A/(qp). \quad (15.4)$$

Величина множителя A находится в диапазоне от 1 до 2 и зависит от механизма рассеяния носителей заряда. Так, для вырожден-

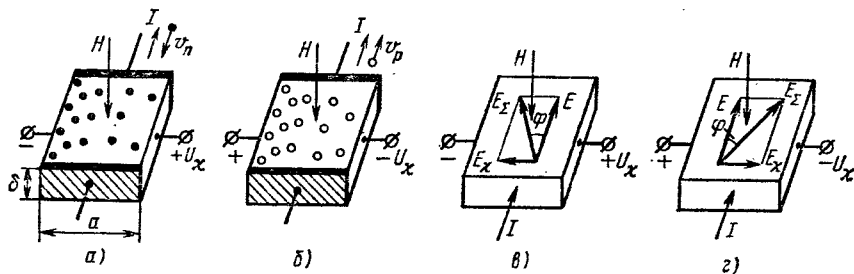


Рис. 15.1. Возникновение э. д. с. Холла в полупроводнике с электропроводностью n -типа (а, в) и p -типа (б, г)

ного полупроводника $A = 1$, для полупроводника с преобладающим рассеянием носителей на тепловых колебаниях кристаллической решетки $A = 1,18$, для полупроводника с преобладающим рассеянием на ионизированных примесях $A = 1,93$.

Для полупроводника с электропроводностью n -типа полярность э. д. с. Холла противоположна. Поэтому коэффициент Холла для такого полупроводника имеет другой знак:

$$X = -A/(qn).$$

В полупроводниках с приблизительно равными концентрациями электронов и дырок (например, в собственных полупроводниках) расчет коэффициента Холла получается более сложным:

$$X = -\frac{A}{q} \cdot \frac{n\mu_n^2 - p\mu_p^2}{(n\mu_n + p\mu_p)^2}.$$

После возникновения холловской напряженности электрического поля и установления динамического равновесия между силой Лоренца и силой холловского электрического поля все носители заряда, имеющие скорость v , будут двигаться по прямолинейным траекториям в соответствии с направлением внешнего электрического поля E (рис. 15.1, θ и γ). При этом направление вектора суммарного электрического поля ($E_{\Sigma} = E + E_X$) отличается от технического направления вектора тока на некоторый угол φ (рис. 15.1, θ и γ), который называют углом Холла. Величину угла Холла определяют по формуле

$$\operatorname{tg} \varphi = E_X/E.$$

Холловская напряженность электрического поля в полупроводнике с электропроводностью p -типа, с учетом (15.3),

$$E_X = \frac{1}{qp} \cdot \frac{IB}{\delta a} = \frac{1}{qp} JB.$$

Напряженность в пластинке полупроводника от внешнего источника питания

$$E = J/\gamma = J/(q\mu_p).$$

Поэтому

$$\operatorname{tg} \varphi = E_X/E = \mu_p B. \quad (15.5)$$

Очевидно, что для полупроводниковой пластинки с электропроводностью n -типа получится аналогичное соотношение между углом Холла, подвижностью электронов и величиной магнитной индукции. При малых магнитных полях и, следовательно, при малых углах Холла

$$\operatorname{tg} \varphi \approx \varphi$$

или

$$\varphi \approx \mu B. \quad (15.6)$$

Магниторезистивный эффект

Под действием силы Лоренца траектория движения носителя заряда искривляется, что равносильно уменьшению длины свободного пробега в направлении внешнего поля между токовыми контактами или увеличению удельного сопротивления полупроводника в магнитном поле.

Эффект Холла мы рассматривали на примере пластинки полупроводника, помещенной в магнитное поле. В этом случае при установлении динамического равновесия возникшая холловская напряженность электрического поля компенсирует действие силы Лоренца и, следовательно, не происходит искривления траекторий носителей заряда, двигающихся со скоростью v . Казалось бы, что в таком случае сопротивление полупроводника не должно изменяться под действием магнитного поля. Однако эти рассуждения справедливы только для носителей, двигающихся со скоростью v , соответствующей средней скорости. В действительности же, как было отмечено, носители в полупроводнике распределены по скоростям. Поэтому носители со скоростью, превышающей среднюю скорость, будут смещаться к одной грани пластинки полупроводника, так как на них будет действовать большая сила Лоренца (15.1). Носители, обладающие скоростью, меньшей средней скорости, будут смещаться к другой грани пластинки полупроводника, так как на них будет действовать большая сила холловской напряженности электрического поля. Таким образом, в рассмотренном примере удельное сопротивление полупроводника изменяется в магнитном поле из-за искривления траекторий носителей заряда, движущихся со скоростью, отличной от средней скорости.

Холловская напряженность электрического поля частично или полностью компенсирует действие силы Лоренца в зависимости от скорости носителей заряда. Поэтому наибольший магниторезистивный эффект можно получить в полупроводнике такой формы и конструкции, при которых возникновение холловской напряженности электрического поля будет затруднено или вообще невозможно. Теоретически такую конструкцию можно представить, если предположить неограниченность полупроводниковой пластинки в направлении, параллельном токовым контактам. Практическая возможность реализации таких условий рассмотрена в § 15.3.

В неограниченном полупроводнике не происходит накопления носителей заряда на боковых гранях пластинки полупроводника, не образуется э. д. с. Холла, а траектория движения носителей заряда отклоняется от направления внешнего электрического поля в направлении силы Лоренца (рис. 15.2). В этом случае вектор плотности тока совпадает по направлению со скоростью движения дырок и поэтому он оказывается сдвинутым от направления внешнего электрического поля на тот же угол φ — угол Холла.

Рассмотрим теперь магниторезистивный эффект с количественной точки зрения при оптимальных условиях, т. е. для неограниченного полупроводника.

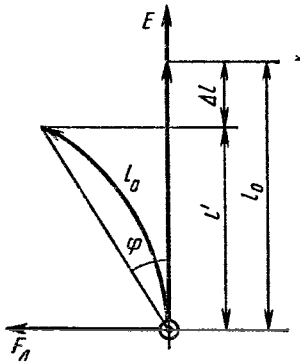


Рис. 15.2. Изменение длины свободного пробега носителя заряда (дырки) вдоль вектора электрического поля в полупроводнике, находящемся в магнитном поле

угла Холла φ можно

Отклонение траектории движения носителей заряда от направления внешнего электрического поля в неограниченном полупроводнике равносильно уменьшению длины свободного пробега носителей заряда в направлении электрического поля на величину

$$\Delta l = l_0 - l' = l_0 - l_0 \cos \varphi,$$

где l_0 — длина свободного пробега носителей заряда при отсутствии магнитного поля; l' — проекция пути, пройденного носителем заряда между двумя последовательными столкновениями (длина свободного пробега) на направление внешнего электрического поля (рис. 15.2).

При малых магнитных полях и, следовательно, при малых значениях

$$\cos \varphi \approx 1 - \frac{\varphi^2}{2!} + \dots$$

Тогда

$$\Delta l \approx l_0 - l_0 + \frac{l_0 \varphi^2}{2}.$$

Используя (15.6), получим

$$\Delta l \approx (l_0/2) \mu^2 B^2.$$

Так как за время свободного пробега носитель заряда проходит при наличии магнитного поля меньший путь вдоль электрического поля E , то это эквивалентно уменьшению дрейфовой скорости или подвижности, а следовательно, и удельной проводимости полупроводника. Относительное изменение удельного сопротивления при этом

$$\frac{\rho - \rho_0}{\rho_0} = \frac{\gamma_0 - \gamma}{\gamma_0} = \frac{\Delta l}{l_0} = \frac{\mu^2 B^2}{2}. \quad (15.7)$$

Для ограниченного по размерам кристалла полупроводника обычно пользуются соотношением

$$\Delta \rho / \rho_0 = C \mu^2 B^2,$$

где C — коэффициент, зависящий от геометрических размеров пластинки полупроводника.

§ 15.2. ПРЕОБРАЗОВАТЕЛИ ХОЛЛА

Преобразователь Холла — это гальваномагнитный полупроводниковый прибор, основанный на использовании эффекта Холла.

Технология изготовления и конструкция

Выражение (15.3) справедливо для пластинки полупроводника, длина которой значительно больше ширины. Э. д. с. преобразователей Холла конечной длины зависит от отношения l/a :

$$\mathcal{E}_x = X \frac{IB}{\delta} f\left(\frac{l}{a}\right). \quad (15.8)$$

Результаты расчета поправочной функции $f(l/a)$ приведены на рис. 15.3. Как видно из рисунка, поправка к э. д. с. Холла, связанная с конечной длиной преобразователя, составляет всего несколько процентов при $l/a = 2 \div 3$. Поэтому нет смысла сильно увеличивать длину преобразователя Холла и такое соотношение длины и ширины имеют практически все преобразователи Холла.

Преобразователи Холла могут быть изготовлены в виде пластинок прямоугольной формы, вырезанных из монокристалла полупроводника (например, германия) или поликристаллического полупроводника, или в виде тонких (1 — 10 мкм) пленок, нанесенных на диэлектрическую подложку методом испарения в вакууме. В качестве материала тонких пленок используют такие соединения, как селенид и теллурид ртути (HgSe, HgTe) и их твердые растворы. Эти соединения обладают большим значением подвижности электронов [около 10^4 см²/(В·с)] даже при малых толщинах пленок. Для изготовления преобразователей Холла применяют также сурьмянистый индий InSb — соединение типа A^{III}B^V, имеющее рекордно высокую подвижность электронов, и другие материалы.

К пластинке полупроводника припаивают четыре электрода: два входных (токовых) и два выходных (холловских). Выходные контакты преобразователей Холла обычно не удается расположить строго на эквипотенциальной линии (при $B = 0$) из-за возможной несимметрии выходных контактов или из-за неоднородности полупроводникового материала. К искажению эквипотенциалов и, в результате, к неэквипотенциальности выходных контактов приводят также механические напряжения, которые могут возникнуть в различных частях пластинки полупроводника при изготовлении преобразователя Холла. В этом случае при прохождении тока через датчик и при отсутствии магнитного поля на выходных электродах будет некоторая разность потенциалов, возникновение

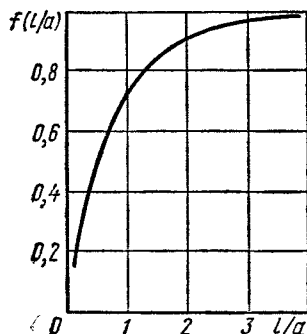


Рис. 15.3. Поправочная функция к э. д. с. Холла, связанная с величиной отношения длины и ширины преобразователя Холла

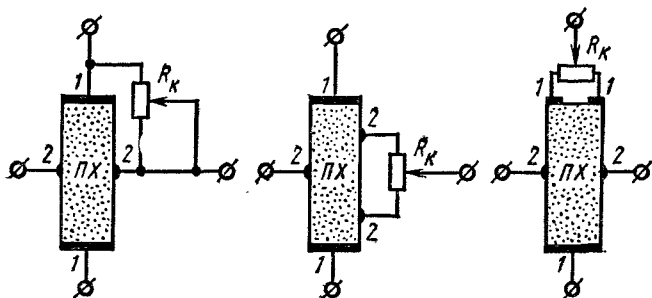


Рис. 15.4. Схемы для компенсации неэквипотенциальности выходных электродов:

ПХ — преобразователь Холла; 1 — входные электроды; 2 — выходные электроды; R_K — компенсационный резистор

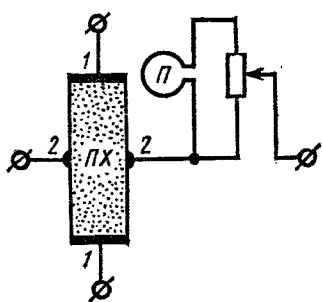


Рис. 15.5. Схема компенсации э. д. с., индуцируемой в цепи выходных электродов:

ПХ — преобразователь Холла; 1 — входные электроды; 2 — выходные электроды; П — петля для компенсации э. д. с. индукции в цепи выходных электродов

которой не связано с эффектом Холла. Эта разность потенциалов устраняется с помощью компенсирующих схем (рис. 15.4).

Искажения показаний преобразователя Холла при работе в переменном магнитном поле могут быть также вызваны э. д. с. индукции в контуре выходных электродов. Величину этого паразитного сигнала можно уменьшить, если скрутить вместе провода, идущие от выходных электродов, т. е. уменьшить площадь контура, пронизываемую магнитным потоком. Э. д. с., наводимую в контуре выходных электродов, можно скомпенсировать с помощью дополнительной петли, помещенной в то же самое магнитное поле (рис. 15.5).

Основные параметры и свойства

Входным сопротивлением преобразователя Холла называют сопротивление между входными электродами. При отсутствии магнитного поля его можно определить по удельному сопротивлению исходного материала и по геометрическим размерам:

$$R_{\text{вх}} = \rho l / (ad). \quad (15.9)$$

Сопротивление полупроводников и в том числе входное сопротивление преобразователя Холла изменяется в магнитном поле, возрастая при увеличении напряженности магнитного поля.

Первой причиной изменения входного сопротивления преобразователя Холла, помещенного в магнитное поле, является изменение подвижности носителей (см. § 15.1).

Второй причиной изменения входного сопротивления преобразователя Холла в магнитном поле является изменение концентрации носителей заряда из-за влияния поверхностной рекомбинации носителей, которые отклоняются силой Лоренца к боковой поверхности преобразователя.

Выходным сопротивлением преобразователя Холла называют сопротивление между выходными электродами. Если выходные контакты точечные, то выходное сопротивление определяется удвоенным сопротивлением растекания точечного контакта (см. § 3.21):

$$R_{\text{вых}} \approx \rho / (\pi a_k), \quad (15.10)$$

где a_k — радиус точечного выходного контакта.

При выходных контактах, выполненных в виде полуцилиндра, вплавленных в боковые грани полупроводниковой пластинки, выходное сопротивление преобразователя Холла

$$R_{\text{вых}} = \frac{2\rho}{\pi\delta} \ln \frac{a}{2a_k},$$

где a_k — радиус полуцилиндрического выходного контакта.

При наличии тока в цепи холловских электродов выходное сопротивление преобразователя Холла так же, как и входное сопротивление, возрастает с увеличением напряженности магнитного поля.

Коэффициент использования или *к. п. д.* преобразователя Холла определяется отношением отдаваемой и подводимой мощностей:

$$\eta = P_{\text{н}} / P_{\text{вх}}.$$

Мощность, подводимая к преобразователю Холла, с учетом (15.9),

$$P_{\text{вх}} = I^2 \frac{\rho l}{a\delta}. \quad (15.11)$$

Мощность, выделяемая на сопротивлении нагрузки,

$$P_{\text{н}} = I_{\text{х}}^2 R_{\text{н}} = \frac{U_{\text{х}}^2 R_{\text{н}}}{(R_{\text{вых}} + R_{\text{н}})^2}.$$

При согласовании выходного сопротивления преобразователя Холла и сопротивления нагрузки, т. е. при равенстве этих сопротивлений ($R_{\text{вых}} = R_{\text{н}}$), преобразователь отдает в нагрузку максимальную мощность при фиксированной подводимой мощности. Тогда для преобразователя с точечными выходными контактами

$$P_{\text{н}} = U_{\text{х}}^2 \pi a_k / (4\rho).$$

Коэффициент использования такого преобразователя Холла с учетом соотношений (15.3) и (15.4)

$$\eta = \frac{\pi a a_k}{4\rho^2 \delta l} X^2 B^2 = \left(\frac{A}{2} \right)^2 \mu^2 B^2 \frac{\pi a a_k}{\delta l}.$$

При условии согласования сопротивления нагрузки и выходного сопротивления для преобразователя Холла с полужилиндрическими выходными контактами и с учетом выражения (15.11)

$$P_n = \frac{\pi \delta U_X^2}{8\rho \ln \frac{a}{2a_k}},$$

коэффициент использования

$$\eta = \frac{\pi a X^2 B^2}{8\rho^2 l \ln \frac{a}{2a_k}} = \left(\frac{A}{2}\right)^2 \mu^2 B^2 \frac{\pi a}{2l \ln \frac{a}{2a_k}}.$$

Таким образом, при различных выходных контактах коэффициент использования преобразователей Холла пропорционален квадрату подвижности носителей заряда.

Максимально допустимый ток. Для увеличения э. д. с. Холла и выходной мощности необходимо увеличивать входную мощность. При подводимой к преобразователю Холла мощности, равной допустимой мощности рассеяния, температура преобразователя повышается до максимально допустимой. Поэтому

$$I_{\max}^2 \frac{\rho l}{a\delta} = \beta S \Delta T, \quad (15.12)$$

где S — площадь поверхности преобразователя Холла; β — коэффициент теплоотдачи; ΔT — разность температур между максимальной допустимой и температурой окружающей среды.

Если пренебречь площадью боковых граней и считать, что $S = 2al$, то на основании уравнения теплового баланса (15.12) максимально допустимый ток через преобразователь Холла

$$I_{\max} = a \sqrt{\frac{2\beta \Delta T \delta}{\rho}}. \quad (15.13)$$

Согласно выражению (15.13) увеличить максимально допустимый ток, а следовательно, и входную мощность, не изменяя температуры преобразователя, можно, увеличивая интенсивность теплообмена между преобразователем и окружающей средой.

Следует отметить, что максимальная мощность рассеяния преобразователя определяется теплоотдачей с его поверхности только у очень тонких пленочных преобразователей, а также у кристаллических преобразователей, имеющих поверхность, превышающую 70—80 мм². У кристаллических преобразователей малых размеров даже тонкие электроды отводят тепла больше, чем рассеивается с поверхности самого преобразователя.

Максимальная э. д. с. Холла при заданной индукции магнитного поля возникает в преобразователе при прохождении через него максимально допустимого тока. Как следует из (15.3) и (15.13), максимальная э. д. с. Холла

$$\mathcal{E}_{X \max} = XBa \sqrt{\frac{2\beta \Delta T}{\rho \delta}}.$$

Вольтовая чувствительность представляет собой отношение максимальной э. д. с. Холла и индукции магнитного поля:

$$\gamma = \frac{\mathcal{E}_{X \max}}{B} = aX \sqrt{\frac{2\beta\Delta T}{\rho\delta}}.$$

При оценке качества преобразователя Холла вольтовая чувствительность является более удобным и объективным параметром, чем максимальная э. д. с. Холла, которая зависит от величины магнитной индукции.

Вольтовую чувствительность преобразователя Холла можно значительно увеличить при использовании импульсного режима питания. Максимально допустимый ток в импульсном режиме можно увеличить в десятки и сотни раз, соответственно увеличится и максимальная э. д. с. Холла при заданной индукции магнитного поля.

Кэффициент передачи равен отношению напряженности э. д. с. Холла к напряженности электрического поля между входными электродами при определенной магнитной индукции. С учетом соотношений (15.3), (15.4) и (15.9) коэффициент передачи преобразователя Холла

$$K = E_X/E = |\operatorname{tg} \varphi| = \mu B.$$

При рассмотрении основных параметров преобразователей Холла можно подумать, что качество преобразователя определяется только подвижностью носителей заряда в полупроводниковой пластинке. Однако это не совсем так. Действительно, при высокой подвижности э. д. с. Холла составляет существенную часть входного напряжения, т. е. коэффициент передачи велик. Но при этом само входное напряжение должно быть достаточно большим, чтобы полезный выходной сигнал имел необходимую величину. Например, для преобразователя Холла из полупроводника с электропроводностью p -типа э. д. с. Холла при допустимой мощности рассеяния, т. е. с учетом выражения (15.11),

$$\mathcal{E}_X = \frac{A}{qr} \cdot \frac{IB}{a} = AB \sqrt{\frac{P_{\text{вх. доп}}}{q\delta l}} \sqrt{\frac{\mu}{p}}. \quad (15.14)$$

Поэтому э. д. с. на выходе преобразователя Холла или его вольтовая чувствительность в соответствии с (15.14) зависит не только от подвижности носителей заряда, но и от их концентрации. В антимониде индия с рекордно большой подвижностью носителей заряда концентрация носителей не может быть малой из-за малой ширины запрещенной зоны. В результате преобразователи Холла из германия обладают лучшими параметрами.

Частотные свойства. Время релаксации процессов, определяющих эффект Холла, порядка $10^{-12} - 10^{-13}$ с, что и определяет теоретический частотный предел использования преобразователей Холла. Однако при их практическом использовании на очень высоких частотах встречаются определенные трудности.

В соответствии с тем, что на преобразователь Холла воздействуют две управляющие величины — магнитная индукция и входной ток, при рассмотрении частотных свойств необходимо различать два случая.

В первом случае преобразователь Холла находится в постоянном магнитном поле, а входная цепь питается высокочастотным управляющим током. В поведении преобразователя на постоянном и переменном токах до частот в несколько десятков мегагерц не наблюдается какой-либо разницы в пределах точности измерений. Чем больше частота входного тока, тем сильнее сказывается нелинейность контактов полупроводника с металлическими электродами.

Во втором случае преобразователь Холла питается постоянным управляющим током и находится в высокочастотном магнитном поле. При таком режиме работы в пластинке полупроводника индуцируются вихревые токи, которые вызывают ее нагрев. Вследствие этого максимально допустимый ток необходимо уменьшать при использовании преобразователя Холла в высокочастотном магнитном поле.

Применение преобразователей Холла

Применение преобразователей Холла основано на том, что их выходной сигнал (э. д. с. Холла) пропорционален произведению тока на магнитную индукцию. В связи с этим преобразователи Холла можно применять для измерения магнитных полей, токов, а также в качестве множительных устройств, фазочувствительных детекторов, анализаторов спектра и т. п.

§ 15.3. МАГНИТОРЕЗИСТОРЫ

Магниторезистор — это полупроводниковый резистор, в котором используется зависимость электрического сопротивления от магнитного поля.

Принцип действия магниторезисторов основан на магниторезистивном эффекте (см. § 15.1). Поскольку холловская напряженность электрического поля, возникающая в полупроводнике с током при наличии магнитного поля, снижает магниторезистивный эффект, то конструкция магниторезистора должна быть такой, чтобы уменьшить или полностью устранить э. д. с. Холла. Наилучшей формой магниторезистора является *диск Корбино* (рис. 15.6, а).

При отсутствии магнитного поля ток в таком магниторезисторе проходит в радиальном направлении от центра диска ко второму электроду, расположенному по периметру диска, или наоборот. Под действием магнитного поля носители заряда отклоняются в направлении, перпендикулярном радиусу. Так как не существует граней, на которых может происходить накопление зарядов, то э. д. с. Холла в таком магниторезисторе не возникает.

Другой конструкцией магниторезистора является пластинка полупроводника, ширина которой много больше ее длины (рис. 15.6, б). Однако существенным недостатком магниторезистора такой конструкции является его малое сопротивление, для увеличения которого применяют последовательное соединение нескольких магниторезисторов или нанесение на поверхности пластины полупроводника металлических полос (рис. 15.6, в). Каждая часть пластины полупроводника между двумя металлическими полосами представляет собой отдельный магниторезистор. Можно также считать, что металлические полосы выполняют роль шунтов, уменьшающих э. д. с. Холла, возникающую на боковых гранях пластины полупроводника.

Основным полупроводниковым материалом для магниторезисторов является антимонид индия (InSb) и арсенид индия (InAs) — материалы с большой подвижностью носителей заряда.

§ 15.4. МАГНИТОДИОДЫ И МАГНИТОТРАНЗИСТОРЫ

Магнитодиоды

Магнитодиод — это полупроводниковый диод, в котором используется изменение вольт-амперной характеристики под действием магнитного поля.

Для диодов с тонкой базой ($W < L$) и с несимметричным p - n -переходом (например, p^+ - n) формулу тока насыщения можно записать в соответствии с (3.31) и с учетом (1.29):

$$I_{\text{нас}} = S \frac{p_n o_n^M p k T}{W_n}. \quad (15.15)$$

Следовательно, сопротивление диода с тонкой базой под действием магнитного поля будет изменяться только в результате изменения подвижности неосновных носителей заряда в базе диода. Из сравнения выражений (15.15) и (15.7) видно, что изменение тока, проходящего через диод с тонкой базой, или изменение его сопротивления в магнитном поле значительно меньше, чем в магниторезисторе. Таким образом, диоды с тонкой базой нецелесообразно использовать в качестве магниточувствительных полупроводниковых приборов — магнитодиодов.

В диодах с толстой базой ($W > L$) прямое напряжение, приложенное к диоду, распределяется между p - n -переходом и сопротивлением базы диода:

$$U_{\text{пр}} = U_{p-n} + I_{\text{пр}} R_{\delta}.$$

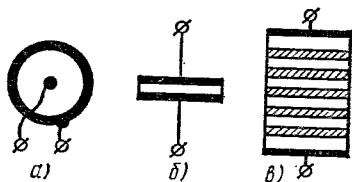


Рис. 15.6. Различные конструкции магниторезисторов:

а — магниторезистор в виде диска Корбино; б — магниторезистор, изготовленный из пластинки полупроводника с малой длиной; в — магниторезистор из пластинки или пленки полупроводника с поперечными металлическими полосками на поверхности

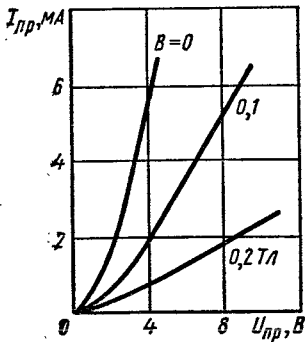


Рис. 15.7. Прямые ветви вольт-амперных характеристик германиевого магнитодиода, находящегося в магнитных полях с различной магнитной индукцией

Сопротивление базы диода увеличивается в поперечном магнитном поле в результате уменьшения подвижности основных и неосновных носителей заряда, как и в обычном магниторезисторе. Увеличение сопротивления базы диода с толстой базой может быть связано также с уменьшением времени жизни неосновных носителей, если из-за искривления траектории движения неосновные носители будут достигать поверхности базовой области, где велика скорость их рекомбинации.

В результате увеличения сопротивления базы прямое напряжение, приложенное к диоду с толстой базой, перераспределяется: уменьшается доля напряжения, приходящаяся на $p-n$ -переход. Этот процесс приводит к резкому уменьшению тока, проходящего через диод, так как этот ток связан экспоненциальной зависимостью с напряжением на $p-n$ -переходе. Кроме того, ток, проходящий через $p-n$ -переход, уменьшается из-за уменьшения тока насыщения, как и в диоде с тонкой базой.

Таким образом, диод с толстой базой может быть использован в качестве магнитодиода при соответствующем выборе геометрических размеров базы диода и электрофизических параметров исходного материала. Обычно магнитодиоды изготавливают с толщиной базы, соответствующей нескольким диффузионным длинам неосновных носителей, т. е. толщиной в несколько миллиметров. Полупроводниковый материал базы, так же как и для магниторезисторов, должен обладать большой подвижностью носителей заряда [см. соотношение (15.7)]. Этим требованиям удовлетворяют германий и кремний.

Прямые ветви вольт-амперных характеристик германиевого магнитодиода в магнитных полях с различной магнитной индукцией показаны на рис. 15.7. Для оценки чувствительности магнитодиода к магнитному полю по аналогии с преобразователями Холла используют вольтовую чувствительность

$$\gamma = \Delta U / (BI),$$

где ΔU — изменение напряжения на магнитодиоде при внесении его в магнитное поле.

Вольтовая чувствительность магнитодиодов может быть значительно выше вольтовой чувствительности преобразователей Холла из того же материала.

Биполярные магнитотранзисторы

Магнитотранзистор — это транзистор, в котором используется зависимость его характеристик и параметров от магнитного поля.

Обычно биполярные транзисторы мало чувствительны к магнитному полю, так как поперечное магнитное поле приводит только к искривлению траекторий движения неосновных носителей заряда, идущих через базу от эмиттера к коллектору, что эквивалентно уменьшению эффективной подвижности неосновных носителей в базе транзистора. В связи с малой толщиной базы в обычных биполярных транзисторах практически все инжектированные эмиттером носители достигают коллектора, несмотря на искривление траекторий их движения магнитным полем. Другой физической причиной изменения параметров биполярных транзисторов в магнитном поле является изменение сопротивления базы транзистора.

Для увеличения чувствительности к магнитному полю биполярные магнитотранзисторы делают с двумя коллекторными переходами (рис. 15.8). Без магнитного поля половина инжектированных носителей заряда попадает на один коллектор, половина — на другой. Магнитное поле отклоняет носители от одного коллектора к другому. По изменению токов первого и второго коллекторов можно оценить или измерить магнитную индукцию поперечного магнитного поля, используя для этого, например, своеобразную мостовую схему (см. рис. 15.8).

Биполярные магнитотранзисторы в диапазоне слабых магнитных полей могут иметь магнитную чувствительность на несколько порядков большую по сравнению с магнитной чувствительностью преобразователей Холла.

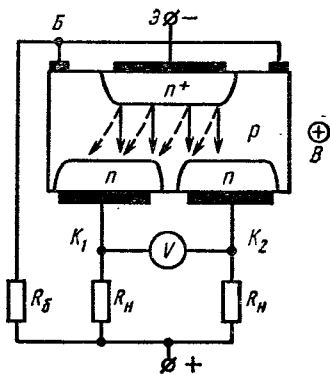
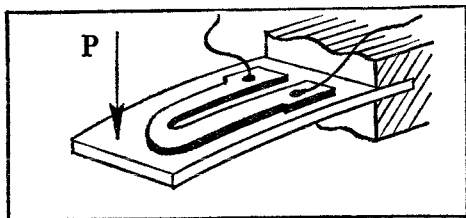


Рис. 15.8. Структура биполярного магнитотранзистора с двумя коллекторными переходами и схема его включения

Контрольные вопросы

1. В чем заключается эффект Холла?
2. Что такое угол Холла? От чего зависит угол Холла?
3. Что такое магниторезистивный эффект?
4. Почему вольтовая чувствительность преобразователей Холла, изготовленных из материала с большой подвижностью носителей заряда, оказывается иногда меньше, чем вольтовая чувствительность таких же преобразователей из материала с меньшей подвижностью носителей?
5. Какую конструкцию должны иметь магниторезисторы?
6. Какие диоды можно использовать в качестве магнитодиодов?
7. Что такое биполярный магнитотранзистор?



§ 16.1. ТЕНЗОРЕЗИСТОРЫ

Принцип действия

Полупроводниковый тензорезистор — это полупроводниковый резистор, в котором используется зависимость электрического сопротивления от механических деформаций.

Принцип действия полупроводникового тензорезистора основан на *тензорезистивном эффекте* — на изменении электрического сопротивления полупроводника под действием механических деформаций.

У полупроводниковых кристаллов со структурой алмаза, таких, как германий и кремний, обладающих высокой степенью симметрии, удельная проводимость при обычных условиях изотропна. Однако если по одной из кристаллографических осей германия или кремния приложено давление, то происходит анизотропное изменение его удельной проводимости. Тензорезистивный эффект можно объяснить только на основе многодолинной модели энергетических зон кристаллов. Рассмотрим кратко физические основы этой модели.

Энергия электрона в кристалле связана с абсолютным значением квазинимпульса P или волнового вектора k соотношением

$$\mathcal{E} = \frac{m^* v^2}{2} = \frac{P^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*}, \quad (16.1)$$

где

$$P = \sqrt{P_x^2 + P_y^2 + P_z^2}; \quad k = \sqrt{k_x^2 + k_y^2 + k_z^2}.$$

Формула (16.1) свидетельствует о том, что зависимость энергии электрона в кристалле от волнового вектора имеет параболический характер, показанный на рис. 16.1, а. Это наиболее простая форма энергетических зон в кристаллах, когда минимум

зоны проводимости и максимум валентной зоны расположены при $k = 0$. Если в k -пространстве с координатами k_x, k_y, k_z (или в пространстве квазиимпульсов с координатами $\hbar k_x, \hbar k_y, \hbar k_z$) построить изоэнергетические поверхности (поверхности равной энергии), то в соответствии с (16.1) они будут представлять собой сферы с центром в точке $k = 0$ ($k_x = k_y = k_z = 0$). Сечение такой сферической поверхности плоскостью $k_x k_y$ показано на рис. 16.1, б.

Однако обычно для полупроводниковых кристаллов зависимость энергии от волнового вектора имеет более сложный вид: в зоне проводимости имеется не один, а несколько минимумов энергии (рис. 16.2, в). Соответствующие минимумам изоэнергетические поверхности в k -пространстве являются эллипсоидами вращения, вытянутыми вдоль осей симметрии кристалла. В кремнии — это направления [100], т. е. имеется шесть эквивалентных эллипсоидов, в германии — направления [111], т. е. четыре эллипсоида. Сечения эллипсоидов равной энергии плоскостью $k_x k_y$ для кремния показаны на рис. 16.1, г.

Эллипсы с центрами на оси k_x соответствуют уравнению

$$\mathcal{E} = \frac{\hbar^2}{2} \left(\frac{k_x^2}{m_{\parallel}^*} + \frac{k_y^2}{m_{\perp}^*} \right); \quad (16.2)$$

эллипсы с центрами на оси k_y — уравнению

$$\mathcal{E} = \frac{\hbar^2}{2} \left(\frac{k_y^2}{m_{\parallel}^*} + \frac{k_x^2}{m_{\perp}^*} \right). \quad (16.3)$$

В уравнения (16.2) и (16.3) введены различные эффективные массы. Эффективную массу электрона m_{\parallel}^* называют продольной, а m_{\perp}^* — поперечной в соответствии с осями эллипсоидов, к которым они относятся.

Так как эффективная масса больше у электронов, движущихся, например, вдоль осей [100], и меньше у электронов, движущихся в перпендикулярном направлении, то отношение подвижностей будет обратным:

$$(\mu_{\perp} = q\bar{\tau}/m_{\perp}^*) > (\mu_{\parallel} = q\bar{\tau}/m_{\parallel}^*),$$

где $\bar{\tau}$ — среднее время свободного пробега электрона.

Различие в подвижностях изображено на рис. 16.2, а в виде векторов разной длины. Аналогичная картина имеет место в гер-

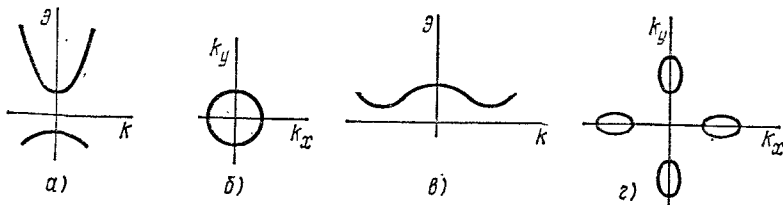


Рис. 16.1. Структура энергетических зон кристаллов полупроводника

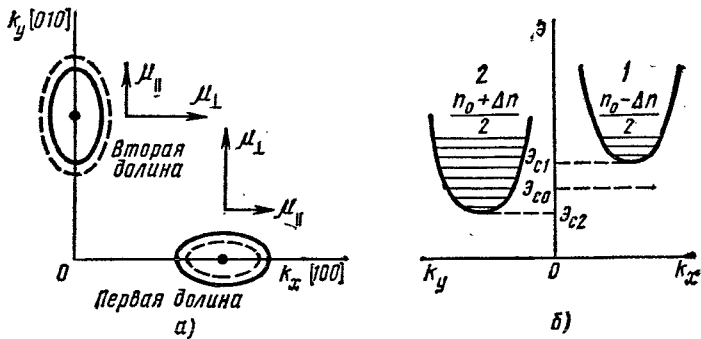


Рис. 16.2. Изоэнергетические поверхности кремния n -типа в плоскости $k_x k_y$ (а) и смещение долин энергетической диаграммы при одноосной деформации кристалла кремния (б)

мании n -типа с той лишь разницей, что продольные оси эллипсоидов совпадают с кристаллографическими направлениями [111].

В недеформированных кристаллах долины (эллипсоиды) симметричны и концентрация электронов в них одинакова. Суммарная проводимость, обусловленная наличием электронов в этих долинах, при токе в направлении оси [100] определяется по формуле

$$\gamma = q \left(\frac{n_0}{2} \mu_{\parallel} + \frac{n_0}{2} \mu_{\perp} \right) = \frac{qn_0}{2} (\mu_{\parallel} + \mu_{\perp}) = qn_0 \mu_{\text{ср}}. \quad (16.4)$$

При токе в направлении [010] результат будет таким же. Можно показать, что соотношение (16.4) справедливо при любом направлении тока, т. е. удельная проводимость недеформированного кристалла изотропна.

Если кристалл подвергается растяжению вдоль оси [100] и ток течет в том же направлении, то минимумы, расположенные по осям k_x и k_y , находятся в разных условиях по отношению к деформации, так как в направлении [100] кристалл растягивается, а в направлении [010] сжимается. Вследствие этого минимум 1 поднимается, а минимум 2 опускается (рис. 16.2, б). При этом часть электронов из первой долины перейдет во вторую долину. Количество электронов с большой подвижностью, движущихся в направлении электрического поля, возрастает, а значит, увеличится и удельная проводимость. Тогда

$$\frac{\gamma_{\text{деф}} - \gamma_0}{\gamma_0} = \frac{\Delta\gamma}{\gamma_0} = \frac{\Delta n}{2n_0} \cdot \frac{\mu_{\perp} - \mu_{\parallel}}{\mu_{\text{ср}}} = \frac{\Delta n}{n_0} \cdot \frac{\mu_{\perp} - \mu_{\parallel}}{\mu_{\perp} + \mu_{\parallel}}.$$

Следовательно,

$$\gamma_{\text{деф}} = q \left(\frac{n_0 + \Delta n}{2} \mu_{\perp} + \frac{n_0 - \Delta n}{2} \mu_{\parallel} \right) = \gamma_0 + \frac{q\Delta n}{2} (\mu_{\perp} - \mu_{\parallel}).$$

Таким образом, эффект изменения удельной проводимости при деформации кристалла определяется количеством электронов, перешедших из долины в долину, и анизотропией эффективных масс.

Если растяжение кристалла производится вдоль направления $[111]$, то деформация вдоль осей $[100]$ одинакова и минимумы смещаются в одном и том же направлении на одну и ту же величину. При этом нет перехода электронов из долины в долину и нет изменения удельной проводимости.

Переход электронов из долины в долину не дает изменения удельной проводимости, если нет анизотропии эффективных масс, т. е. если $m_{\parallel}^* = m_{\perp}^*$, что соответствует сферическим изоэнергетическим поверхностям кристалла. В кремнии n -типа анизотропия эффективных масс составляет $m_{\parallel}^*/m_{\perp}^* \approx 5$, а в германии n -типа $m_{\parallel}^*/m_{\perp}^* \approx 20$.

Величиной, численно характеризующей изменение удельной проводимости или удельного сопротивления при деформации, является *тензочувствительность*, которая представляет собой отношение относительного изменения удельного сопротивления полупроводника к относительной деформации в данном направлении:

$$m = \frac{\Delta\rho/\rho}{\Delta l/l}. \quad (16.5)$$

Тензочувствительность в полупроводниках зависит от кристаллографического направления, типа электропроводности и удельного сопротивления (рис. 16.3). В германии с электропроводностью n - и p -типа наибольшая тензочувствительность имеет место в направлении $[111]$, в кремнии n -типа — в направлении $[100]$ и в кремнии p -типа — в направлении $[111]$. Знак тензочувствительности в полупроводниках n -типа отрицательный, а в полупроводниках p -типа положительный.

Технология изготовления и конструкция. Наиболее сильно тензорезистивный эффект выражен в таких полупроводниковых материалах, как кремний, германий, антимонид индия, фосфид индия, арсенид галлия, антимонид галлия. Для тензорезисторов широко используют кремний. Он обладает высокой тензочувствительностью, химически достаточно инертен, имеет относительно высокое (по сравнению с другими полупроводниками) значение разрушающего напряжения ($9 \cdot 10^8$ Па), позволяет изготавливать тензорезисторы с большой допустимой рабочей температурой.

Простым и широко распространенным способом изготовления тензорезисторов является выпиливание тензочувствительных эле-

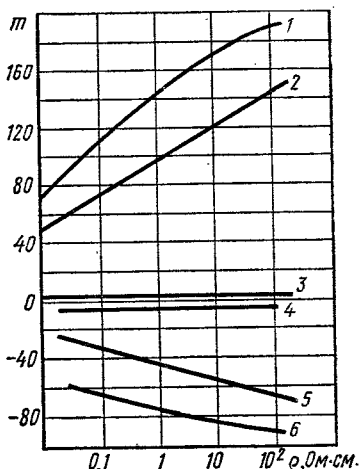


Рис. 16.3. Зависимость тензочувствительности кремния от удельного сопротивления при различных типах электропроводности и в различных кристаллографических направлениях:

1 — p -тип $[111]$; 2 — p -тип $[110]$;
3 — p -тип $[100]$; 4 — n -тип $[111]$;
5 — n -тип $[110]$; 6 — n -тип $[100]$;

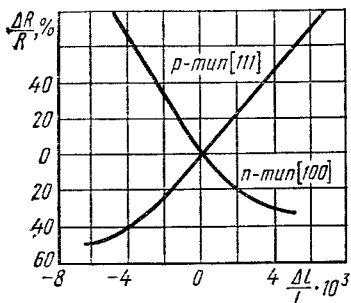


Рис. 16.4. Деформационные характеристики тензорезисторов из кремния с электропроводностью p - и n -типов

высокоомного кремния с p - n -переходами и удобные для использования в мостовых схемах.

Основные параметры, характеристики и свойства

Номинальным сопротивлением тензорезистора называют его сопротивление без деформации при температуре 20°C ; обычно оно имеет величину от нескольких десятков до нескольких тысяч ом.

Деформационная характеристика — это зависимость относительного изменения сопротивления тензорезистора от относительной деформации. Как видно из рис. 16.4, тензорезисторы из кремния p -типа имеют линейную деформационную характеристику в диапазоне $[(-2) \div (+6)] \cdot 10^{-3}$ единиц относительной деформации, а тензорезисторы из кремния n -типа — только при сжатии от $-4 \cdot 10^{-3}$ до 0.

Коэффициент тензочувствительности представляет собой отношение относительного изменения сопротивления к относительному изменению длины тензорезистора:

$$K = \frac{\Delta R/R}{\Delta l/l}. \quad (16.6)$$

Рассмотрим тензорезистор как однородное по удельному сопротивлению тело с постоянным сечением S , длиной l .

Величину коэффициента тензочувствительности тензорезистора можно выразить через удельное сопротивление ρ и коэффициент Пуассона $\mu_{\text{П}}$. Для этого в выражение (16.6) вместо $\Delta R/R$ подставим дифференциал прологарифмированного выражения $R = \rho(l/S)$.

Тогда

$$K = \frac{dR/R}{dl/l} = 1 + \frac{d\rho/\rho}{dl/l} - \frac{dS/S}{dl/l}.$$

Но так как

$$\frac{dS}{S} = -2\mu_{\text{П}} \frac{dl}{l},$$

то

$$K = \frac{dR/R}{dl/l} = 1 + 2\mu_{\Pi} + \frac{d\rho/\rho}{dl/l}.$$

Учитывая выражение для тензочувствительности материала (16.5), получим

$$K = 1 + 2\mu_{\Pi} + m. \quad (16.7)$$

Величина $1 + 2\mu_{\Pi}$ характеризует изменение сопротивления, связанное с изменением длины и сечения тензорезистора. Однако коэффициент тензочувствительности в значительной степени определяется большей для полупроводниковых материалов величиной тензочувствительности m .

Коэффициент тензочувствительности полупроводниковых тензорезисторов зависит от деформации, так как деформационные характеристики тензорезисторов не являются прямыми линиями (рис. 16.4). Изменение коэффициента тензочувствительности, связанное с деформацией, тем больше, чем больше удельное сопротивление полупроводникового материала тензорезистора.

Предельная деформация тензорезисторов зависит от площади поперечного сечения полупроводникового стержня, а также от обработки поверхности этого стержня. При малой площади поперечного сечения ($0,3 \text{ мм}^2$) предельная относительная деформация тензорезистора может достигать порядка 10^{-2} (при тщательной обработке поверхностей полупроводниковых стержней). Превышение предельной деформации приводит к выходу тензорезистора из строя.

Основными свойствами полупроводниковых тензорезисторов в отличие от металлических проволочных и фольговых тензорезисторов являются малые размеры, очень высокие коэффициенты тензочувствительности (на два порядка выше, чем у проволочных тензорезисторов), высокий уровень выходного сигнала измерительных схем. Важнейшая особенность полупроводниковых тензорезисторов — это возможность изготовления их с различными механическими и электрическими свойствами, что неосуществимо в проволочных тензорезисторах. Например, при одних и тех же размерах сопротивление полупроводникового тензорезистора может быть $100 \text{ Ом} - 50 \text{ кОм}$, а коэффициент тензочувствительности — от -100 до $+200$.

§ 16.2. ТЕНЗОДИОДЫ И ПОЛИКРИСТАЛЛИЧЕСКИЕ ТЕНЗОРЕЗИСТОРЫ

Тензодиоды

Тензодиод — это полупроводниковый диод, в котором используется изменение вольт-амперной характеристики под действием механических деформаций.

В тензодиодах используется изменение величины потенциального барьера p - n -перехода, обусловленное изменением ширины запрещенной зоны при механической деформации. Коэффициент тензочувствительности тензодиода к всестороннему давлению достигает нескольких сотен, а при одноосной деформации он значительно выше.

В качестве тензoeлектрических полупроводниковых приборов можно применять туннельные диоды, крутизна вольт-амперной характеристики которых на отдельных участках зависит от деформации. Используя такие участки кривой, можно получить стабильный эффект с хорошей линейностью и $K \approx 200$.

Существенным преимуществом туннельных диодов является слабая зависимость коэффициента тензочувствительности и сопротивления от температуры. Чувствительность к деформации можно повысить, шунтируя диод положительным сопротивлением, значение которого близко к величине отрицательного дифференциального сопротивления диода. Это позволяет получить $K \approx 30\,000$, но за счет ухудшения линейности и температурной стабильности. Величины измеряемых давлений и чувствительность к давлению в рабочем интервале температур легко регулируются изменением величины шунтирующего сопротивления и тока через диод. Тензодиоды можно использовать для измерения малых давлений, в качестве гидрофонов, сейсмографов и т. п.

Поликристаллические тензорезисторы

Более сложными неоднородными системами, чувствительными к деформации, являются поликристаллические полупроводниковые тензорезисторы. Сопротивление таких тензорезисторов определяется не только удельным сопротивлением отдельных кристаллов, но и сопротивлением контактов между ними. При деформации этих тензорезисторов их сопротивление может изменяться вследствие изменения площадей контактов между кристаллами. Однако в большинстве случаев на поверхности кристаллов существуют потенциальные барьеры или *p-n*-переходы. При деформации их сопротивление резко изменяется в связи с большими местными механическими напряжениями, поэтому коэффициент тензочувствительности поликристаллических тензорезисторов может достигать больших величин. Так, у тензорезисторов, выполненных из поликристаллического сернистого свинца, коэффициент тензочувствительности достигает нескольких тысяч, т. е. на два порядка больше, чем в монокристаллических тензорезисторах из того же материала.

Поликристаллические тензорезисторы пока не получили широкого распространения из-за плохой воспроизводимости параметров, нестабильности, явлений старения, зависимости от состояния окружающей среды и т. д. Но не следует считать, что разработка поликристаллических тензорезисторов бесперспективна, а их недостатки неустранимы.

Контрольные вопросы

1. Какие физические процессы используются в различных тензoeлектрических полупроводниковых приборах?
2. Что такое тензочувствительность и коэффициент тензочувствительности? Чем эти параметры отличаются друг от друга?
3. Почему при использовании туннельных диодов в качестве тензодиодов можно получить очень высокий коэффициент тензочувствительности?

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

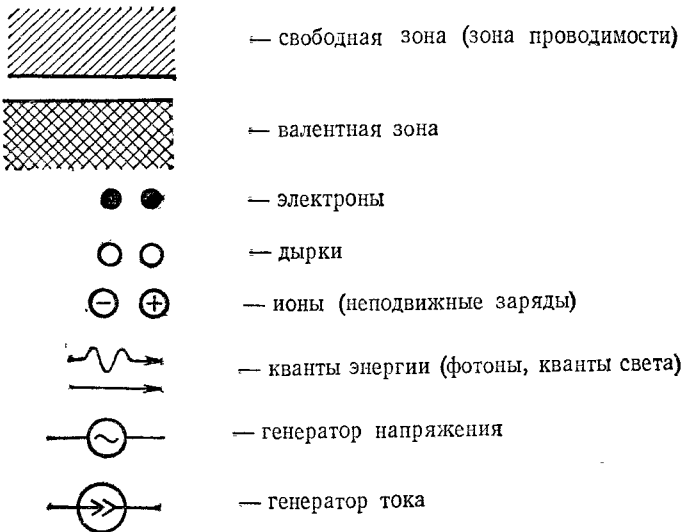
1. Степаненко И. П. Основы теории транзисторов и транзисторных схем. — М.: Энергия, 1977.
2. Федотов Я. А. Основы физики полупроводниковых приборов. — М.: Советское радио, 1969.
3. Спиридонов Н. С. Основы теории транзисторов. — Киев: Техніка, 1975.
4. Трутко А. Ф. Методы расчѐта транзисторов. — М.: Энергия, 1971.
5. Колесников В. Г. и др. Кремниевые планарные транзисторы / Под ред. Я. А. Федотова. — М.: Советское радио, 1973.
6. Челноков В. Е., Евсеев Ю. А. Физические основы работы силовых полупроводниковых приборов. — М.: Энергия, 1973.
7. Агаханян Т. М. Основы транзисторной электроники. — М.: Энергия, 1974.
8. Степаненко И. П. Основы микроэлектроники. — М.: Советское радио, 1980.
9. Справочник по полупроводниковым диодам, транзисторам и интегральным схемам / Под ред. Н. Н. Горюнова. — М.: Энергия, 1976.
10. Чебовский О. Г., Моисеев Л. Г., Сахаров Ю. В. Силовые полупроводниковые приборы (спраночник). — М.: Энергия, 1975.
11. Носов Ю. Р. Оптоэлектроника. — М.: Советское радио, 1977.
12. Викулин И. М., Стафеев В. И. Полупроводниковые датчики, — М.: Советское радио, 1975.
13. Котенко Г. И. Магниторезисторы. — Л.: Энергия, 1972.

ОБОЗНАЧЕНИЯ ОСНОВНЫХ ВЕЛИЧИН, ПРИНЯТЫЕ В КНИГЕ

- B* — магнитная индукция, коэффициент температурной чувствительности, яркость
C — емкость
D — коэффициент диффузии
E — напряженность электрического поля, освещенность
E — э. д. с.
F — коэффициент шума
f — частота
g — плотность объемного заряда
h — постоянная Планка ($\hbar = h/2\pi$)
I — постоянное, эффективное значение тока, интенсивность света
i — мгновенное значение тока
J — плотность тока
 $i = \sqrt{-I}$

k — постоянная Больцмана, волновое число
 L — диффузионная длина носителей заряда, индуктивность
 l — длина, расстояние, длина свободного пробега носителя заряда
 M — коэффициент лавинного умножения
 m — масса, тензочувствительность
 m^* — эффективная масса носителя заряда
 N — концентрация примесей, эффективная плотность энергетических уровней
 n — концентрация электронов
 n_0 — равновесная концентрация электронов
 P — постоянное, эффективное значение мощности, импульс
 p — концентрация дырок
 p_0 — равновесная концентрация дырок
 Q — заряд, добротность, количество тепла
 q — элементарный заряд
 Φ — световой поток
 R — статическое сопротивление, радиус
 r — дифференциальное сопротивление, радиус
 S — площадь, крутизна характеристики
 s — скорость поверхностной рекомбинации
 T — абсолютная температура, период колебаний
 TKI — температурный коэффициент тока
 TKR — температурный коэффициент сопротивления
 TKU — температурный коэффициент напряжения
 t — температура, время
 U — постоянное, эффективное значение напряжения
 u — мгновенное значение напряжения
 V — объем
 v — скорость
 W — толщина базы
 \mathcal{E} — энергия
 α — коэффициент передачи тока эмиттера одномерной теоретической модели транзистора, коэффициент ударной ионизации, коэффициент термо- ε д. с., показатель поглощения
 α^* — эффективность коллектора
 α_n — коэффициент переноса
 β — коэффициент передачи тока базы одномерной теоретической модели транзистора, коэффициент нелинейности
 γ — удельная проводимость, эффективность эмиттера
 δ — толщина p - n -перехода
 ε — относительная диэлектрическая проницаемость
 ε_0 — электрическая постоянная
 η — коэффициент полезного действия
 K — удельная теплопроводность
 λ — длина волны
 μ — подвижность
 μ_n — коэффициент Пуассона
 ν — частота в оптическом диапазоне
 Π — коэффициент Пельтье
 ρ — удельное сопротивление
 σ — проводимость
 τ — время жизни носителей заряда, постоянная времени
 φ — потенциал, угол, фазовый сдвиг
 X — коэффициент Холла
 ω — угловая частота

На рисунках приняты следующие обозначения:



УНИВЕРСАЛЬНЫЕ ФИЗИЧЕСКИЕ ПОСТОЯННЫЕ

Постоянная Планка
 $h = 6,62 \cdot 10^{-34}$ Дж·с = $4,14 \cdot 10^{-15}$ эВ·с,
 Постоянная Больцмана
 $k = 1,38 \cdot 10^{-23}$ Дж/К = $0,86 \cdot 10^{-4}$ эВ/К,
 Масса покоя электрона
 $m \approx 9 \cdot 10^{-28}$ г,
 Заряд электрона (элементарный заряд)
 $q \approx 1,6 \cdot 10^{-19}$ Кл.
 Электрическая постоянная
 $\epsilon_0 = 8,86 \cdot 10^{-12}$ Ф/см.
 Магнитная постоянная
 $\mu_0 = 1,257 \cdot 10^{-8}$ Гн/см.
 Скорость света в вакууме
 $c \approx 3 \cdot 10^{10}$ см/с.

АЛФАВИТНЫЙ УКАЗАТЕЛЬ

В алфавитный указатель включены основные термины и понятия, встречающиеся в учебнике. Рядом с термином указаны только те страницы, на которых можно найти определение или расшифровку данного термина, а также сведения, поясняющие его смысл.

Термины, состоящие из прилагательного и существительного, помещены в алфавитном указателе в большинстве случаев с инверсией, т. е. за основное слово принято существительное. Однако некоторые термины, используемые в технической литературе обычно без инверсии (например, «электронно-дырочный переход»), приведены в алфавитном указателе дважды.

- Активная часть базы транзистора 167
Активный режим работы транзистора 166, 169
Акцептор 14
Атомная батарея 356
Атомный элемент 356
- База полупроводникового прибора 41, 167
Барьерная емкость *p-n*-перехода 54
Барьер Шоттки 139
Бездрейфовый транзистор 168
Безызлучательная рекомбинация 9
Биполярный транзистор 166, 302
— — с диодом Шоттки 303
Блок выпрямительный 133
Болометр полупроводниковый 366
Быстрое поверхностное состояние 34
- Валентная зона 6
Варикап 159
Варистор 371
- Видеодетектор 143
Видность 326
Внезапный отказ 163
Внутренний квантовый выход 30, 323
Время жизни неравновесных носителей заряда 15
— пролета неосновных носителей заряда через базу диода 90
— — неосновных носителей через базу транзистора 220
Вторичный пробой в транзисторе 189
Вынужденная рекомбинация 32
Выпрямительный блок 133
— полупроводниковый диод 127
— столб 133
Вырожденный полупроводник 11
Высокий уровень инжекции в базе диода 112
— — — в базе транзистора 181
Высота потенциального барьера *p-n*-перехода 40
- Генератор Ганна 315
— оптический квантовый 330

- с ограничением накопления объемного заряда 320
- термоэлектрический 387
- Генерационный ток 92
- Генерация носителей заряда 7
 - — — световая 8
 - — — тепловая 7
- Гетеропереход 62
- Гибридная интегральная микросхема 299
- Граничная частота коэффициента передачи тока базы 225

Демаркационный энергетический уровень 17

Динистор 252

Диод выпрямительный 127

- Ганна 316

- измерительный 144

- импульсный 137

- лавинно-пролетный 149

- лавинный 132

- модуляторный 146

- обращенный 158

- переключательный 144

- — перерезонансный 145

- — резонансный 144

- плоскостной 70

- полупроводниковый 69

- с барьером Шоттки 139

- с контролируемым лавинообразованием 132

- смесительный 145

- с толстой базой 79

- с тонкой базой 80

- точечный 70

- туннельный 152

- умножительный 146

- частотный 132

- Шоттки 304

Диск Корбино 406

Диффузионная длина 20

Диффузионная емкость базы транзистора 208, 210

- — диода 82, 87, 88

- — коллектора транзистора 208, 212

- — эмиттера транзистора 208, 210

Диффузионное сопротивление базы

транзистора 207, 210

- электрическое поле 38

Диффузионный конденсатор 311

- *p-n*-переход 44

- резистор 308

Диффузия носителей заряда 20

Длина диффузионная 20

- свободного пробега носителя заряда 21

Дно разрешенной энергетической зоны 6

Донор 14

Дрейф носителей заряда 19

Дрейфовый транзистор 168

Запоминающее устройство на аморфном полупроводнике 380

Запрещенная зона 6

Затвор полевого транзистора 268

Зона валентная 6

- запрещенная 6

- примесная 14

- проводимости 6

- разрешенная 6

- свободная 6

- энергетическая 6

Избыточная концентрация носителей заряда 8

Излучательная рекомбинация 9

Излучатель электролюминесцентный пленочный 337

- — порошковый 335

Излучающий полупроводниковый прибор 322

Измерительный диод 144

Импульсный диод 137

Инверсная населенность энергетических уровней 331

Инверсное включение транзистора 167

Инверсный слой 34

Инжекция неосновных носителей заряда 41

Интегральная микросхема 297

- — аналоговая 300

- — гибридная 299

— — логическая 300
— — полупроводниковая 297
Ионизация ударная 25
Исток полевого транзистора 268

Канал в полевом транзисторе 268
— поверхностной электропроводности 111
Катастрофический отказ 163
Катодолюминесценция 32
Квазиуровень Ферми для электронов (дырок) 18
Квантовый выход внутренний 30, 323
— — внешний 323, 324
Коллектор 167
Коллекторный переход 167
Конверсионный *p-n*-переход 45
Конденсатор диффузионный 311
— МДП 312
— пленочный 313
Контакт металл — полупроводник 60
— металлургический 38
— полупроводников с одним типом электропроводности 58
Контактная разность потенциалов 38
Концентрация носителей заряда избыточная 8
— — — неравновесная 8
— — — равновесная 8
— — — собственная 12
Корпускулярно-преобразовательный прибор 353
Коэффициент выпрямления 67
— диффузии 20
— лавинного умножения 95
— нелинейности 67, 373
— передачи тока базы одномерной теоретической модели транзистора 209
— — — базы транзистора 184
— — — транзистора 203
— — — эмиттера одномерной теоретической модели транзистора 207
— — — эмиттера транзистора 186
— переноса 208
— ударной ионизации 25
— Холла 397
— шума 143, 232

Крутизна характеристики полевого транзистора 273, 275, 286

Лавинно-пролетный диод 149
Лавинный диод 132
— пробой *p-n*-перехода 95
Лазер полупроводниковый 330
Ловушка захвата 17
— рекомбинационная 9, 18
Люминесценция 31

Магнитодиод 407
Магниторезистивный эффект 396, 399
Магниторезистор 406
Магнитотранзистор 409
Максимальная частота генерации транзистора 226
Масса эффективная носителя заряда 27
Матричный светодиод 325
МДП-транзистор 278
— — с индуцированным каналом 279
— — со встроенным каналом 279, 283
МДП-конденсатор 312
Медленное поверхностное состояние 35
Межзонная рекомбинация 8
Металлургический контакт 38
Микросхема интегральная 297
— — аналоговая 300
— — гибридная 299
— — логическая 300
— — полупроводниковая 297
Микроэлектроника 296
Модель транзистора одномерная теоретическая 206
Модуляторный диод 146

Напряжение обратное 41
— переворота фазы тока базы транзистора 188
— прямое 41
Наращивание эпитаксиальное 46
Начальный ток коллектора 183
— — эмиттера 183
Невырожденный полупроводник 11
Непосредственная рекомбинация 8
Непрямой переход электронов 29
Неравновесная концентрация носителей заряда 8

- Нерезонансный переключаемый диод 145
- Несимметричный *p-n*-переход 46
- Низкий уровень инжекции 112
- Нормальное включение транзистора 167
- Обедненный слой 34**
- Обогащенный слой 34
- Обратное напряжение 41
- Обратный ток коллектора 184
- — эмиттера 184
- Обращенный диод 158
- Одномерная теоретическая модель транзистора 206
- Однопереходный транзистор 245
- Оптический квантовый генератор 330
- Оптрон полупроводниковый 357
- Отказ внезапный 163
- катастрофический 163
- постепенный 163
- условный 163
- Пассивная часть базы транзистора 167**
- Переключатель на аморфном полупроводнике 377
- Переключаемый диод 144
- Переход электронов прямой 29
- — не прямой 29
- электрический 64
- электронно-дырочный 38
- — — диффузионный 44
- — — коллекторный 167
- — — конверсионный 45
- — — несимметричный 46
- — — плавный 46
- — — планарный 44
- — — резкий 46
- — — симметричный 46
- — — сплавной 44
- — — эмиттерный 167
- — — эпитаксиальный 46
- Периферическая часть базы транзистора 167
- ПЗС (прибор с зарядовой связью) 287
- Плотность упаковки 297
- Поверхностная рекомбинация 9
- Поверхностное состояние быстрое 34
- — медленное 35
- Поверхностно-зарядовый транзистор 294
- Поверхностный пробой *p-n*-перехода 111
- Поглощение фотонов 28
- — носителями заряда 28
- — примесное 28
- — собственное 28
- Подвижность носителей заряда 19
- Позистор 368
- Показатель поглощения 29
- Полевой транзистор 268
- — с изолированным затвором 278
- — с управляющим *p-n*-переходом 268
- Поле электрическое диффузионное 38
- Полупроводник вырожденный 11
- невырожденный 11
- примесный 13
- простой 13
- скомпенсированный 15
- сложный 13
- собственный 12
- *n*-типа 15
- *p*-типа 15
- Полупроводниковая интегральная микросхема 297
- Полупроводниковый болометр 366
- диод 69
- лазер 330
- оптрон 357
- прибор излучающий 322
- прибор с зарядовой связью 287
- приемник проникающей радиации 353
- светодиод 322
- стабилитрон 146
- тензодиод 415
- тензорезистор 410
- тепловой насос 394
- термоэлектрический прибор 383
- фотодиод 343
- Полупроводниковый фотоэлемент 348
- холодильник 392
- Постепенный отказ 163
- Потолок разрешенной зоны 6

- Предельная частота коэффициента передачи тока базы 225
 — — — передачи тока эмиттера 224
 Преобразовательный диод 145
 Преобразователь Холла 401
 Примесная зона 14
 Примесное поглощение 28
 Примесный полупроводник 13
 Пробивное напряжение *p-n*-перехода 95
 Пробой диода 95
 — — лавинный 95
 — — поверхностный 111
 — — по дефектам 100
 — — тепловой 103
 — — туннельный 101
 — транзистора 187
 — — вторичный 189, 197
 — — лавинный 187
 Проводимость удельная 19
 Прямой переход электронов 29
- Работа выхода электронов 60
 Равновесная концентрация электронов (дырок) 8
 Разность потенциалов контактная 38
 «Разогрев» носителей заряда 26
 Разрешенная энергетическая зона 6
 Режим работы транзистора активный 166, 169
 — — — насыщения 166, 173
 — — — отсечки 166, 173
 Резистор диффузионный 308
 — пленочный 310
 Резкий *p-n*-переход 46
 Рекомбинационный ток 94
 Рекомбинация безызлучательная 9
 — вынужденная 32
 — излучательная 9
 — индуцированная 32
 — линейная 16
 — межзонная 8
 — непосредственная 8
 — поверхностная 9
 — самопроизвольная 32
 — спонтанная 32
 — стимулированная 32
- с участием рекомбинационных ловушек 9
 Световая генерация носителей заряда 8
 Светодиод полупроводниковый 322
 — матричный 325
 — управляемый 325
 Свободная зона (энергетическая) 6
 Симметричный стабилитрон 148
 — тиристор 257
 Скомпенсированный полупроводник 15
 Скорость поверхностной рекомбинации носителей заряда 35
 — рекомбинации на контакте 65
 Слой обедненный 34
 — обогащенный 34
 — инверсный 34
 Смесительный диод 145
 Смыкание переходов транзистора 187
 Собственная концентрация носителей заряда 12
 Собственный полупроводник 12
 Соотношение Эйнштейна 21
 Сопротивление базы транзистора диффузионное 207, 210
 — — — объемное 213
 — коллектора 207, 211
 — контакта 66
 — контактное удельное 66
 — насыщения транзистора 187
 — растекания 136
 — слоя удельное 36
 — эмиттера 207, 209
 Сплавной *p-n*-переход 44
 Средний уровень инжекции 112
 Стабилитрон 146
 — симметричный 148
 — термокомпенсированный 148
 Стабистор 149
 Статический коэффициент передачи тока базы транзистора 184
 — — — тока эмиттера транзистора 186
 Степень интеграции 297
 Сток полевого транзистора 268
- Тензодиод 415
 Тензорезистор полупроводниковый 410

— — поликристаллический 416
Тепловая генерация носителей заряда
7
Тепловой пробой диода 103
Термистор 360
— косвенного подогрева 367
— прямого подогрева 360
Термокомпенсированный стабилитрон
148
Термоэлектрическая батарея 387
Термоэлектрический генератор 387
Термоэлемент 383
Тиристор 252
— двухэлектродный 252
— симметричный 257
— трехэлектродный 256
Ток генерационный 92
— коллектора начальный 183
— насыщения 72
— рекомбинационный 94
— эмиттера начальный 183
Точечный диод 70
Транзистор биполярный 166, 302
— — бездрейфовый 168
— — с диодом Шоттки 303
— — диффузионный 236
— — дрейфовый 168
— — интегральный 302
Транзистор биполярный конверсион-
ный 243
— — планарный 240, 243
— — с мезаструктурой 239
— — сплавно-диффузионный 237
— — сплавной 234, 242
— — эпитаксиальный 237
— полевой 268
— — поверхностно-зарядовый 294
— — с изолированным затвором 278
— — с индуцированным каналом
279
— — со встроенным каналом 279, 283
— — с управляющим *p-n*-переходом
268
Трехэлектродный тиристор 256
Тринистор 256
Туннелирование носителей заряда
25
Туннельный диод 152
— пробой *p-n*-перехода 101

Угол Холла 398
Ударная ионизация 25
Удельная проводимость полупровод-
ника 19
Удельное контактное сопротивление 66
— сопротивление слоя 36
Умножительный диод 146
Управляемый светодиод 325
Уравнение непрерывности 74
— Пуассона 46
Уровень инжекции высокий 112
— — низкий 112
— — средний 112
Уровень энергетический демаркацион-
ный 17
— — ловушки захвата 17, 18
— — — рекомбинационной 9, 18
— — поверхностный 33
— — Ферми 10
Условие лавинного пробоя *p-n*-пере-
хода 97
Условный отказ 163

Физическая эквивалентная схема тран-
зистора 206
Формальная эквивалентная схема тран-
зистора 205
Фотодиод 343
Фотолюминесценция 32
Фоторезистивный эффект 30
Фоторезистор 339
Фототиристор 352
Фототранзистор 351
Фотоэлемент 348

Холодильник полупроводниковый 392

Частотный диод 132
Часть базы транзистора активная 167
— — — пассивная 167
— — — периферическая 167

Ширина запрещенной зоны 6

Э. д. с. Дембера 31
Экстракция носителей заряда 41

- Электрический переход 64
- Электролюминесцентный пленочный излучатель 337
- порошоквый излучатель 335
- Электролюминесценция 32
- Электронно-дырочный переход 38
- Электроэлемент атомный 356
- Элемент памяти на аморфном полупроводнике 380
- Эмиттер 167
- Эмиттерный *p-n*-переход 167
- Энергетическая зона 6
- Энергетический уровень ловушки захвата 17, 18
- — рекомбинационной ловушки 9, 18
- Энергия ионизации донора (акцептора) 14
- сродства к электрону 63
- Эпитаксиальное наращивание 46
- Эпитаксиальный *p-n*-переход 46
- Эффект Ганна 316
- Гаусса 396
- Зеебека 383
- магниторезистивный 396, 399
- Пельтье 386
- тензорезистивный 410
- фоторезистивный 30
- Холла 396
- Эффективная масса носителя заряда 27
- Эффективность коллектора 208
- эмиттера 208

ОГЛАВЛЕНИЕ

	Предисловие	3
	Введение	4
1		
ОСНОВНЫЕ СВЕДЕНИЯ ПО ФИЗИКЕ ПОЛУПРОВОДНИКОВ	§ 1.1. Энергетические зоны полупроводников	6
	§ 1.2. Генерация и рекомбинация носителей заряда	7
	§ 1.3. Концентрация носителей заряда в полупроводнике при термодинамическом равновесии	9
	§ 1.4. Собственные полупроводники	12
	§ 1.5. Примесные полупроводники	13
	§ 1.6. Время жизни неравновесных носителей заряда	15
	§ 1.7. Процессы электропроводности в полупроводниках	18
	§ 1.8. Температурные зависимости концентрации носителей заряда и положения уровня Ферми	22
	§ 1.9. Температурные зависимости подвижности носителей заряда и удельной проводимости	23
	§ 1.10. Полупроводники в сильных электрических полях	24
	§ 1.11. Оптические свойства полупроводников	28
	§ 1.12. Обогащенные, инверсные и обедненные поверхностные слои	32
	§ 1.13. Поверхностная рекомбинация	35
	§ 1.14. Проводимость канала поверхностной электропроводности	36
2		
КОНТАКТНЫЕ ЯВЛЕНИЯ	§ 2.1. Электронно-дырочный переход	38
	§ 2.2. Токи через электронно-дырочный переход	40
	§ 2.3. Концентрация неосновных носителей заряда у границ электронно-дырочного перехода	41
	§ 2.4. Методы формирования и классификация электронно-дырочных переходов	44
	§ 2.5. Распределение напряженности электрического поля и потенциала в электронно-дырочном переходе	46
	§ 2.6. Аналитический расчет резкого электронно-дырочного перехода	50
	§ 2.7. Аналитический расчет плавного электронно-дырочного перехода с линейным распределением концентрации примесей	53

§ 2.8. Барьерная емкость электронно-дырочного перехода	54
§ 2.9. Контакт полупроводников с одним типом электропроводности	58
§ 2.10. Контакт металл—полупроводник	60
§ 2.11. Гетеропереходы	62
§ 2.12. Свойства и параметры невыпрямляющих контактов	64

3 ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

§ 3.1. Структура и основные элементы	69
§ 3.2. Вольт-амперная характеристика диода при инжекции и экстракции носителей заряда	70
§ 3.3. Расчет распределения неосновных носителей заряда в базе диода	73
§ 3.4. Расчет постоянных токов, проходящих через диод и связанных с инжекцией и экстракцией носителей заряда	77
§ 3.5. Частные случаи расчета распределения неосновных носителей заряда и тока насыщения	79
§ 3.6. Расчет переменных токов и полной проводимости диода	82
§ 3.7. Графики частотных зависимостей параметров диода	86
§ 3.8. Физический смысл параметров диода	87
§ 3.9. Пределы применимости частных случаев расчета параметров диода	91
§ 3.10. Генерация и рекомбинация носителей заряда в электронно-дырочном переходе	92
§ 3.11. Лавинный пробой	95
§ 3.12. Туннельный пробой	101
§ 3.13. Тепловой пробой	103
§ 3.14. Влияние поверхностных состояний на вольт-амперную характеристику диода	109
§ 3.15. Процессы в диодах при больших прямых токах	111
§ 3.16. Расчет вольт-амперной характеристики диода при больших прямых токах	115
§ 3.17. Вольт-амперная характеристика диода в полулогарифмических координатах	122
§ 3.18. Переходные процессы в диодах	123
§ 3.19. Выпрямительные плоскостные диоды	127
§ 3.20. Селеновые выпрямители	133
§ 3.21. Выпрямительные точечные высокочастотные диоды	135
§ 3.22. Импульсные диоды	137

§ 3.23. Плоскостные диоды с выпрямлением на контакте металл—полупроводник	139
§ 3.24. СВЧ-диоды	141
§ 3.25. Стабилитроны	146
§ 3.26. Стабисторы	149
§ 3.27. Лавинно-пролетные диоды	149
§ 3.28. Туннельные диоды	152
§ 3.29. Обращенные диоды	158
§ 3.30. Варикапы	159
§ 3.31. Надежность диодов	162

4 БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

§ 4.1. Структура и основные режимы работы	166
§ 4.2. Распределение стационарных потоков носителей заряда	169
§ 4.3. Распределение носителей заряда	173
§ 4.4. Значения постоянных токов при активном режиме	177
§ 4.5. Явления в транзисторах при больших токах	180
§ 4.6. Статические параметры	183
§ 4.7. Пробой транзисторов	187
§ 4.8. Статические характеристики	191
§ 4.9. Работа транзистора на малом переменном сигнале	198
§ 4.10. Малосигнальные параметры	201
§ 4.11. Эквивалентные схемы	204
§ 4.12. Эквивалентная схема одномерной теоретической модели	207
§ 4.13. Барьерные емкости переходов и сопротивление базы	212
§ 4.14. Частотные характеристики	218
§ 4.15. Работа транзистора на импульсах	227
§ 4.16. Шумы в транзисторах	231
§ 4.17. Низкочастотные маломощные транзисторы	234
§ 4.18. Высокочастотные маломощные транзисторы	235
§ 4.19. Мощные транзисторы	242
§ 4.20. Однопереходный транзистор	245
§ 4.21. Надежность транзисторов	248

5 ТИРИСТОРЫ

§ 5.1. Структура и принцип действия	252
§ 5.2. Способы переключения	258
§ 5.3. Конструкция и технология изготовления	262
§ 5.4. Параметры и характеристики	264
	429

6			
ПОЛЕВЫЕ ТРАНЗИСТОРЫ			
И ПРИБОРЫ			
С ЗАРЯДОВОЙ СВЯЗЬЮ			
	§ 6.1.	Полевые транзисторы с управляющим <i>p-n</i> -переходом	268
	§ 6.2.	Расчет выходных статических характеристик полевого транзистора с управляющим <i>p-n</i> -переходом	273
	§ 6.3.	Эквивалентные схемы полевого транзистора с управляющим <i>p-n</i> -переходом	275
	§ 6.4.	Частотные свойства полевых транзисторов с управляющим <i>p-n</i> -переходом	277
	§ 6.5.	Полевые транзисторы с изолированным затвором (МДП-транзисторы)	278
	§ 6.6.	Расчет выходных статических характеристик полевого транзистора с изолированным затвором	284
	§ 6.7.	Параметры и свойства полевых транзисторов с изолированным затвором	286
	§ 6.8.	Полупроводниковые приборы с зарядовой связью	287
	§ 6.9.	Разновидности приборов с зарядовой связью	292
7			
ПОЛУПРОВОДНИКОВЫЕ			
ПРИБОРЫ			
КАК ЭЛЕМЕНТЫ			
ИНТЕГРАЛЬНЫХ			
МИКРОСХЕМ			
	§ 7.1.	Задачи и принципы микроэлектроники	296
	§ 7.2.	Классификация интегральных микросхем	297
	§ 7.3.	Методы изоляции элементов интегральных микросхем	300
	§ 7.4.	Активные элементы	302
	§ 7.5.	Пассивные элементы	308
8			
ПОЛУПРОВОДНИКОВЫЕ			
ПРИБОРЫ			
С ИСПОЛЬЗОВАНИЕМ			
МЕЖДОЛИННОГО			
ПЕРЕХОДА			
НОСИТЕЛЕЙ ЗАРЯДА			
	§ 8.1.	Принцип действия генератора Ганна	315
	§ 8.2.	Технология изготовления генераторов Ганна	318
	§ 8.3.	Свойства и параметры генераторов Ганна	319
	§ 8.4.	Генераторы с ограничением накопления объемного заряда	320
9			
ИЗЛУЧАЮЩИЕ			
ПОЛУПРОВОДНИКОВЫЕ			
ПРИБОРЫ			
	§ 9.1.	Светодиоды	322
	§ 9.2.	Лазеры	330
	§ 9.3.	Электролюминесцентные порошковые излучатели	335
	§ 9.4.	Электролюминесцентные пленочные излучатели	337
10			
ПОЛУПРОВОДНИКОВЫЕ			
ПРИЕМНИКИ ИЗЛУЧЕНИЯ			
	§ 10.1.	Фоторезисторы	339
	§ 10.2.	Фотодиоды	343
	§ 10.3.	Полупроводниковые фотоэлементы	348
	§ 10.4.	Фототранзисторы и фототристоры	351

	§ 10.5. Приемники проникающей радиации и корпускулярно-преобразовательные приборы	353
	§ 10.6. Оптроны	357
11		
ТЕРМИСТОРЫ	§ 11.1. Термисторы прямого подогрева	360
	§ 11.2. Болометры	366
	§ 11.3. Термисторы косвенного подогрева	367
	§ 11.4. Позисторы	368
12		
ВАРИСТОРЫ	§ 12.1. Принцип действия	371
	§ 12.2. Характеристики	373
13		
ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ	§ 13.1. Переключатели	377
	§ 13.2. Запоминающие устройства	380
	§ 13.3. Надежность, стабильность и срок службы приборов на аморфных полупроводниках	381
14		
ПОЛУПРОВОДНИКОВЫЕ ТЕРМОЭЛЕКТРИЧЕСКИЕ ПРИБОРЫ	§ 14.1. Принцип действия	383
	§ 14.2. Термоэлектрические генераторы	387
	§ 14.3. Холодильники и тепловые насосы	392
15		
ПОЛУПРОВОДНИКОВЫЕ ГАЛЬВАНОМАГНИТНЫЕ ПРИБОРЫ	§ 15.1. Принцип действия	396
	§ 15.2. Преобразователи Холла	401
	§ 15.3. Магниторезисторы	406
	§ 15.4. Магнитодиоды и магнитотранзисторы	407
16		
ТЕНЗОЭЛЕКТРИЧЕСКИЕ ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ	§ 16.1. Тензорезисторы	410
	§ 16.2. Тензодиоды и поликристаллические тензорезисторы	415
	Список рекомендуемой литературы	417
	Обозначения основных величин, принятые в книге	417
	Универсальные физические постоянные	419
	Алфавитный указатель	420