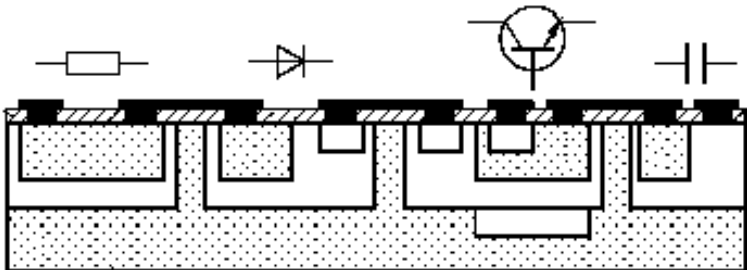


Т.В. Свистова

ПРИБОРЫ ТВЕРДОТЕЛЬНОЙ ЭЛЕКТРОНИКИ

Учебное пособие



Воронеж 2012

ФГБОУ ВПО «Воронежский государственный
технический университет»

Т.В. Свистова

ПРИБОРЫ ТВЕРДОТЕЛЬНОЙ ЭЛЕКТРОНИКИ

Утверждено Редакционно-издательским советом
университета в качестве учебного пособия

Воронеж 2012

УДК 621.382

Свистова Т.В. Приборы твердотельной электроники: учеб. пособие / Т.В. Свистова. Воронеж: ФГБОУ ВПО «Воронежский государственный технический университет», 2012. 294 с.

В учебном пособии рассматриваются основные физические процессы и явления, обеспечивающие работу приборов твердотельной электроники. Приводится анализ электронных процессов в объеме полупроводников, в электронно-дырочных переходах и в области пространственного заряда на поверхности полупроводников. Подробно рассмотрены характеристики и параметры различных видов приборов твердотельной электроники.

Издание соответствует требованиям Федерального государственного образовательного стандарта высшего профессионального образования по направлению подготовки бакалавров 210100 «Электроника и наноэлектроника» (профиль «Микроэлектроника и твердотельная электроника»), дисциплине «Физические основы электроники».

Учебное пособие подготовлено в электронном виде в текстовом редакторе Word 2003 и содержится в файле «Приборы ТЭ.doc».

Табл. 3. Ил. 138. Библиогр.: 20 назв.

Научный редактор д-р физ.-мат. наук, проф. С.И. Рембеза

Рецензенты: кафедра физики полупроводников и микроэлектроники Воронежского государственного университета (зав. кафедрой д-р физ.-мат. наук, проф. Е.Н. Бормонтов);
канд. физ.-мат. наук, доц. Е.В. Бордаков

© Свистова Т.В., 2012

© Оформление. ФГБОУ ВПО «Воронежский государственный технический университет», 2012

ВВЕДЕНИЕ

Твердотельная электроника решает задачи, связанные с изучением свойств полупроводников, диэлектриков, магнитных материалов, влиянием на эти свойства примесей и особенностей структуры, изучением свойств поверхностей и границ раздела между слоями различных материалов; созданием в кристалле областей с разными типами проводимости; созданием гетеропереходов и многослойных структур; созданием функциональных устройств микронных и субмикронных размеров, а также способами измерения их параметров. Основными направлениями твердотельной электроники являются: полупроводниковая электроника, связанная с разработкой различных видов полупроводниковых приборов, и микроэлектроника, связанная с разработкой интегральных схем.

Полупроводниковый прибор - это устройство, изготовленное частично или целиком из полупроводникового материала, которое может выполнять необходимые функции в электронной аппаратуре. Твердотельная электроника определяется физическими свойствами самих полупроводниковых материалов. Поэтому в первой главе учебного пособия представлены основные сведения из физики твердого тела и физики полупроводников. Во второй и третьей главах излагаются физические основы, связанные с контактными явлениями в электронно-дырочных переходах, барьерах Шоттки и гетеропереходах, а также основы физики поверхности полупроводников и МДП-структур. В четвертой главе рассмотрены полупроводниковые диоды, структура которых проще структур большинства других приборов. Поэтому материал этой главы необходим для понимания принципа работы других полупроводниковых приборов.

В учебном пособии рассмотрены физические основы работы твердотельных полупроводниковых приборов, использующих как явление инжекции носителей через p - n -переходы,

так и явления, связанные с эффектом поля. Глава 5 посвящена биполярному транзистору, работа которого основана на взаимодействии двух близко расположенных $p-n$ -переходов. Биполярный транзистор является одним из важнейших полупроводниковых приборов. Можно считать, что с его изобретения в 1948 году началась эра современной электроники. Работа тиристора, представляющего собой три близко расположенных и взаимодействующих $p-n$ -перехода ($p-n-p-n$ -структура), рассматривается в главе 6. Тиристоры обладают широким диапазоном уровня переключаемых мощностей. В зависимости от назначения они могут выдерживать токи от нескольких миллиампер до тысяч ампер и напряжение свыше 5000 В. Глава 7 посвящена униполярным приборам, работу которых определяют главным образом носители одного типа (основные). Рассмотрены родственные по принципу действия полевые транзисторы с $p-n$ -переходом в качестве затвора и полевые транзисторы металл-диэлектрик-полупроводник (МДП-транзисторы), представляющие собой важный элемент интегральных схем.

Область электроники вообще и полупроводниковых приборов в частности столь динамична и претерпевает столь быстрые изменения, что понятия сегодняшнего дня могут оказаться устаревшими уже завтра. Пособие направлено на то, чтобы оказать помощь в изучении основных физических процессов, лежащих в основе работы полупроводниковых приборов, и приобрести достаточные знания в области физики и математики, чтобы оценивать, понимать и отвечать на любые события в этих быстроразвивающихся областях.

Учебное пособие предназначено для студентов направления подготовки бакалавров 210100 «Электроника и наноэлектроника», профиля «Микроэлектроника и твердотельная электроника» очной формы обучения.

1. ФИЗИЧЕСКИЕ ОСНОВЫ РАБОТЫ ПРИБОРОВ ТВЕРДОТЕЛЬНОЙ ЭЛЕКТРОНИКИ

1.1. Зонная структура полупроводников

Твердое тело представляет собой ансамбль отдельных атомов, химическая связь между которыми объединяет их в кристаллическую решетку. Согласно постулатам Бора энергетические уровни для электронов в изолированном атоме имеют дискретные значения. Если твердое тело состоит из N атомов, то энергетические уровни оказываются N -кратно вырожденными. Электрическое поле ядер, или остовов атомов, выступает как возмущение, снимающее это вырождение. Дискретные моноэнергетические уровни атомов, составляющие твердое тело, расщепляются в энергетические зоны. Решение квантовых уравнений в приближении сильной или слабой связи дает качественно одну и ту же картину для структуры энергетических зон твердых тел. В обоих случаях разрешенные и запрещенные состояния для электронов чередуются и число состояний для электронов в разрешенных зонах равно числу атомов, что позволяет говорить о квазинепрерывном распределении энергетических уровней внутри разрешенных зон.

Наибольшее значение для электронных свойств твердых тел имеют верхняя и следующая за ней разрешенные зоны энергий. В том случае, если между ними нет энергетического зазора, то твердое тело с такой зонной структурой является *металлом*. Если величина энергетической щели между этими зонами (обычно называемой запрещенной зоной) больше 3 эВ, то твердое тело является *диэлектриком*. И, наконец, если ширина запрещенной зоны E_g лежит в диапазоне $(0,1 \div 3,0)$ эВ, то твердое тело принадлежит к классу *полупроводников*. В зависимости от сорта атомов, составляющих твердое тело, и конфигурации орбит валентных электронов реализуется тот или иной тип кристаллической решетки, а,

следовательно, и структура энергетических зон. На рис. 1.1 приведена структура энергетических уровней в изолированном атоме кремния, а также схематическая структура энергетических зон, возникающих при сближении этих атомов и образовании монокристаллического кремния с решеткой так называемого алмазоподобного типа.

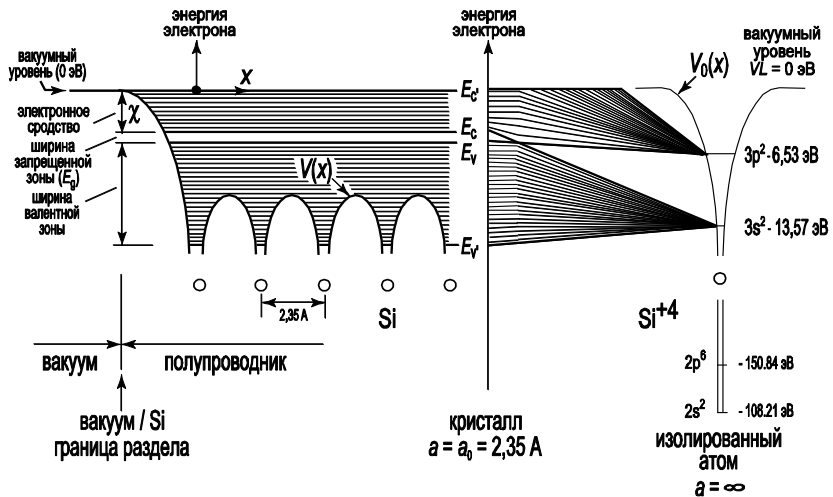


Рис. 1.1. Структура энергетических уровней в изолированном атоме кремния и схематическая структура энергетических зон, возникающих при сближении этих атомов и образовании монокристаллического кремния

Верхняя, не полностью заполненная, энергетическая зона в полупроводниках получила название *зоны проводимости*. Следующая за ней энергетическая зона получила название *валентной зоны*. Энергетическая щель запрещенных состояний между этими зонами называется *запрещенной зоной*. На зонных диаграммах положение дна зоны проводимости

обозначают значком E_c , положение вершины валентной зоны – E_v , а ширину запрещенной зоны – E_g .

Поскольку в полупроводниках ширина запрещенной зоны меняется в широком диапазоне, то вследствие этого в значительной мере меняется их удельная проводимость. По этой причине полупроводники классифицируют как вещества, имеющие при комнатной температуре удельную электрическую проводимость σ от 10^{-8} до 10^6 Ом $^{-1}$ ·см $^{-1}$, которая зависит в сильной степени от вида и количества примесей, структуры вещества и внешних условий: температуры, освещения (радиации), электрических и магнитных полей и т.д.

Для диэлектриков ширина запрещенной зоны $E_g > 3$ эВ, величина удельной проводимости $\sigma < 10^{-8}$ Ом $^{-1}$ ·см $^{-1}$, удельное сопротивление $\rho = \frac{1}{\sigma} > 10^8$ Ом·см. Для металлов величина удельной проводимости $\sigma > 10^6$ Ом $^{-1}$ ·см $^{-1}$.

1.2. Собственные и примесные полупроводники

Полупроводники, или полупроводниковые соединения, бывают собственными и примесными.

Собственные полупроводники – это полупроводники, в которых нет примесей (доноров и акцепторов). **Собственная концентрация** (n_i) – концентрация носителей заряда в собственном полупроводнике (электронов в зоне проводимости n и дырок в валентной зоне p , причем $n = p = n_i$). При $T = 0$ в собственном полупроводнике свободные носители отсутствуют ($n = p = 0$). При $T > 0$ часть электронов забрасывается из валентной зоны в зону проводимости. Эти электроны и дырки могут свободно перемещаться по энергетическим зонам. **Дырка** – это способ описания коллективного движения большого числа электронов (примерно 10^{23} см $^{-3}$) в не полностью заполненной валентной зоне. Электрон – это частица,

дырка – это квазичастица. Электрон можно инжектировать из полупроводника или металла наружу (например, с помощью фотоэффекта), дырка же может существовать только внутри полупроводника.

Легирование – введение примеси в полупроводник, в этом случае полупроводник называется **примесным**. Если в полупроводник, состоящий из элементов 4 группы (например, кремний или германий), ввести в качестве примеси элемент 5 группы, то получим **донорный** полупроводник (у него будет **электронный тип проводимости**), или полупроводник *n*-типа. Если же ввести в качестве примеси элемент 3 группы, то получится **акцепторный** полупроводник, обладающий **дырочной проводимостью** (*p*-тип) (рис. 1.2).

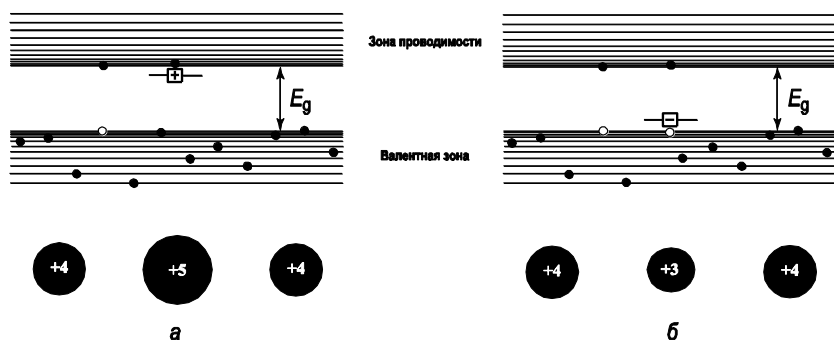


Рис. 1.2. Энергетические схемы полупроводников *n*-типа (а) и *p*-типа (б)

1.3. Статистика электронов и дырок в полупроводниках. Концентрация носителей заряда и положение уровня Ферми

Равновесные процессы – процессы, происходящие в телах, которые не подвергаются внешним воздействиям. В состоянии термодинамического равновесия для кристалла полу-

проводника при заданной температуре существует определенное распределение электронов и дырок по энергиям, а также значения их концентраций. Вычисление концентраций основных и неосновных носителей заряда составляет главную задачу статистики электронов и дырок в кристаллах.

Электроны как частицы, обладающие полуцелым спином, подчиняются статистике Ферми – Дирака. Вероятность того, что электрон будет находиться в квантовом состоянии с энергией E , выражается функцией Ферми – Дирака:

$$f(E, T) = \frac{1}{1 + \exp\left(\frac{E - F}{kT}\right)}. \quad (1.1)$$

Здесь F – электрохимический потенциал, или уровень Ферми. Из (1.1) видно, что уровень Ферми можно определить как энергию такого квантового состояния, вероятность заполнения которого равна $1/2$.

Вид функции Ферми – Дирака схематически показан на рис. 1.3. При $T = 0$ она имеет вид разрывной функции. Для $E < F$ она равна 1, а значит, все квантовые состояния при $E < F$ заполнены электронами. Для $E > F$ функция $f = 0$ и соответствующие квантовые состояния совершенно не заполнены. При $T > 0$ функция Ферми изображается непрерывной кривой и в узкой области энергий, порядка нескольких kT , в окрестности точки $E = F$ быстро изменяется от 1 до 0. Размытие функции Ферми тем больше, чем выше температура.

Вычисление различных статистических величин значительно упрощается, если уровень Ферми F лежит в запрещенной зоне энергий и удален от края зоны E_C хотя бы на $2kT$ (или $E_C - E > kT$). Тогда в распределении (1.1) единицей в знаменателе можно пренебречь, и оно переходит в распределение Максвелла – Больцмана классической статистики. Это случай невырожденного полупроводника:

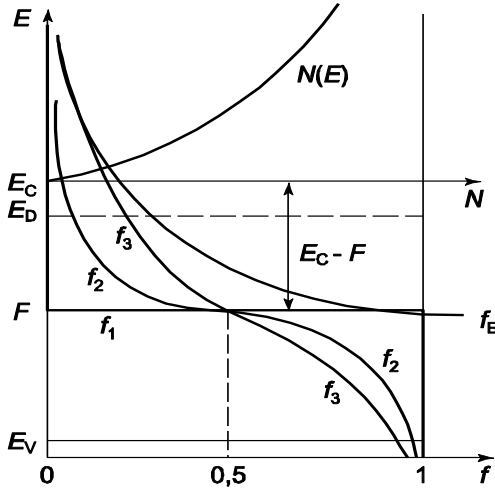


Рис. 1.3. Функция распределения плотности состояний в зоне проводимости $N(E)$, функции Ферми – Дирака f и Больцмана f_B

$$f(E, T) = \exp\left(-\frac{E - F}{kT}\right). \quad (1.2)$$

Функция Ферми – Дирака для электронов f_n имеет вид

$$f_n = \exp\left(-\frac{E - F}{kT}\right) \quad (1.3)$$

Концентрация электронов в зоне проводимости равна

$$n = N_C \exp\left(-\frac{E_C - F}{kT}\right), \quad (1.4)$$

где

$$N_C = 2 \left(\frac{2\pi m_n kT}{h^2} \right)^{3/2}. \quad (1.5)$$

Величина N_C получила название эффективной плотности состояний в зоне проводимости.

В случае невырожденного полупроводника, когда уровень Ферми лежит выше потолка валентной зоны хотя бы на $2kT$, то есть $F - E_C > 2kT$ (или $F - E_C > kT$), функция Ферми – Дирака для дырок f_p имеет вид

$$f_p = \exp\left(-\frac{F - E}{kT}\right), \quad (1.6)$$

а концентрация дырок в валентной зоне

$$p = N_V \exp\left(-\frac{F - E_V}{kT}\right), \quad (1.7)$$

где E_V – энергия, соответствующая потолку валентной зоны, а N_V – эффективная плотность состояний в валентной зоне, которая рассчитывается по уравнению (1.5), если вместо m_n взять эффективную массу дырки m_p .

Для расчета n и p по уравнениям (1.4) и (1.7) необходимо знать положение уровня Ферми F . Однако произведение концентраций электронов и дырок для невырожденного полупроводника не зависит от уровня Ферми, хотя зависит от температуры:

$$n \cdot p = (n_i)^2 = N_C \cdot N_V \cdot \exp\left(-\frac{E_{\text{гг}}}{kT}\right). \quad (1.8)$$

Это уравнение используется для расчета p при известном n или, наоборот, для расчета n при известном p . Величина n_i при некоторых температурах для конкретных полупроводников приводится в справочниках.

1.4. Концентрация электронов и дырок в собственном полупроводнике

Полупроводник называется собственным, если в нем отсутствуют донорные и акцепторные примеси. В этом случае электроны появляются в зоне проводимости только за счет теплового заброса из валентной зоны, тогда $n = p$ (рис. 1.4).

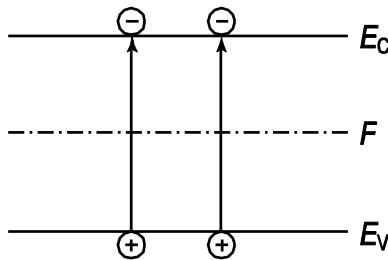


Рис. 1.4. Заброс электронов из валентной зоны в собственном полупроводнике

При отсутствии внешних воздействий (освещение, электрическое поле и т.д.) будем обозначать концентрации свободных электронов и дырок с индексом нуль, то есть n_0 и p_0 соответственно. При $n_0 = p_0$ из (1.8) получаем

$$n_0 = p_0 = n_i = \sqrt{N_C \cdot N_V} \exp\left(-\frac{E_g}{2kT}\right), \quad (1.9)$$

где n_i - концентрация собственных носителей заряда в зоне проводимости и в валентной зоне. Для расчета N_C и N_V используется формула (1.5). Как следует из соотношения (1.9), концентрация собственных носителей определяется в основном температурой и шириной запрещенной зоны полупроводника.

1.5. Концентрация электронов и дырок в примесном полупроводнике

Уравнение (1.8) справедливо только для равновесных носителей заряда, то есть в отсутствие внешних воздействий. В наших обозначениях

$$n_0 \cdot p_0 = (n_i)^2. \quad (1.10)$$

Пусть полупроводник легирован донорами с концентрацией N_D . При комнатной температуре в большинстве полупроводников все доноры ионизованы, так как энергии активации доноров составляют всего несколько сотых электронвольта. Тогда для донорного полупроводника

$$n_0 = N_D. \quad (1.11)$$

Концентрацию дырок в донорном полупроводнике найдем из (1.10):

$$p_0 = \frac{n_i^2}{N_D}. \quad (1.12)$$

На рис. 1.5 приведена зонная диаграмма полупроводника n -типа, показывающая положение энергетических уровней донорной примеси E_D и схематическое соотношение концентраций основных n_0 и неосновных p_0 носителей заряда.

Соответственно, если полупроводник легирован акцепторами с концентрацией N_A , то концентрации основных p_0 и неосновных n_0 носителей заряда будут равны

$$p_0 = N_A \text{ и } n_0 = \frac{n_i^2}{N_A}. \quad (1.13)$$

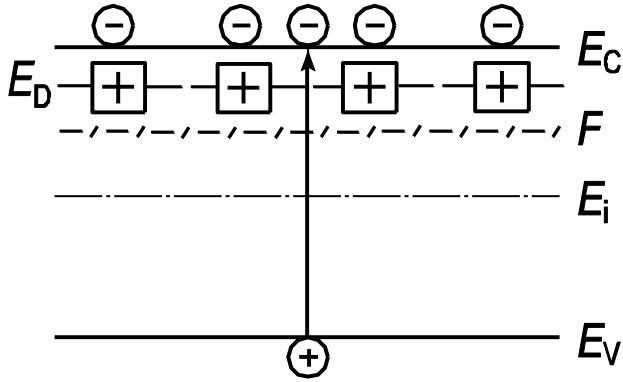


Рис. 1.5. Зонная диаграмма полупроводника *n*-типа

На рис. 1.6 приведена зонная диаграмма полупроводника *p*-типа, показывающая положение энергетических уровней акцепторной примеси E_A и схематическое соотношение концентраций основных p_0 и неосновных n_0 носителей заряда.

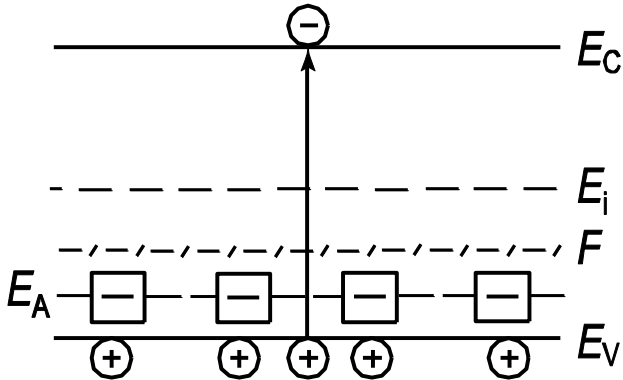


Рис. 1.6. Зонная диаграмма полупроводника *p*-типа

1.6. Определение положения уровня Ферми

Для собственного полупроводника уравнение электронейтральности приобретает вид $p - n = 0$ или $p = n$. Если ширина запрещенной зоны полупроводника достаточно велика (E_g много больше kT) и если эффективные массы электронов m_n и дырок m_p одного порядка, то уровень Ферми будет достаточно удален от краев зон ($E_C - F > 2kT$ и $F - E_V > 2kT$) и полупроводник будет невырожденным.

Подставляя (1.4) и (1.7) в уравнение $p + p_D - n - n_A = 0$, имеем:

$$N_C \exp\left(\frac{F - E_C}{kT}\right) = N_V \exp\left(\frac{E_V - F}{kT}\right). \quad (1.14)$$

Отсюда вычисляем F . Уравнение (1.14) – это уравнение первого порядка относительно $e^{\frac{F}{kT}}$.

Это дает

$$F = E_i - \frac{1}{2}kT \ln\left(\frac{N_C}{N_V}\right) = E_i - \frac{3}{4}kT \ln\left(\frac{m_n^*}{m_p^*}\right), \quad (1.15)$$

где через $E_i = \frac{1}{2}(E_V + E_C)$ обозначена энергия середины запрещенной зоны. При выводе правого выражения для F величина (N_C/N_V) была заменена на (m_n/m_p) с помощью уравнения (1.5).

Для случая $m_n^* = m_p^*$ энергия Ферми в собственном полупроводнике находится посреди запрещенной зоны $F = (E_C + E_V)/2$.

Если известны концентрации носителей заряда в зонах n и p , то значение F можно определить из формул (1.4) и (1.7). Так, для невырожденного полупроводника n -типа имеем

$$F = E_C - kT \cdot \ln\left(\frac{N_C}{n}\right). \quad (1.16)$$

Аналогично для невырожденного полупроводника p -типа

$$F = E_V + kT \cdot \ln\left(\frac{N_V}{p}\right). \quad (1.17)$$

Из выражений (1.16) и (1.17) видно, что чем больше концентрация основных носителей, тем ближе уровень Ферми к краю соответствующей зоны. Для донорного полупроводника $n = N_D$, тогда

$$F = E_C - kT \cdot \ln\left(\frac{N_C}{N_D}\right). \quad (1.18)$$

Для акцепторного полупроводника $p = N_A$, тогда

$$F = E_V + kT \cdot \ln\left(\frac{N_V}{N_A}\right). \quad (1.19)$$

1.7. Проводимость полупроводников

При приложении электрического поля к однородному полупроводнику в нем протекает электрический ток. При наличии двух типов свободных носителей – электронов и дырок – проводимость σ полупроводника будет определяться суммой электронной σ_n и дырочной σ_p компонент проводимости: $\sigma = \sigma_n + \sigma_p$. Величина электронной и дырочной компонент в полной проводимости определяется классическим соотношением:

$$\sigma_n = \mu_n n_0 q; \quad \sigma_p = \mu_p p_0 q, \quad (1.20)$$

где μ_n и μ_p – подвижности электронов и дырок соответственно.

Для легированных полупроводников концентрация основных носителей всегда существенно больше, чем концентрация неосновных носителей, поэтому проводимость таких полупроводников будет определяться только компонентой проводимости основных носителей. Так, для полупроводника n -типа

$$\sigma = \sigma_n + \sigma_p = \sigma_n. \quad (1.21)$$

Величина, обратная удельной проводимости, называется удельным сопротивлением

$$\rho = \frac{1}{\sigma} = \frac{1}{\sigma_n + \sigma_p} = \frac{1}{\sigma_n}. \quad (1.22)$$

Здесь ρ – удельное сопротивление, обычно измеряемое в единицах [Ом·см]. Для типичных полупроводников, используемых в производстве интегральных схем, величина удельного сопротивления находится в диапазоне $\rho = (1 - 10)$ Ом·см.

Подставляя соотношение (1.20) в (1.22), получаем $\rho = \frac{1}{\sigma_n} = \frac{1}{\mu_n n_0 q} = \frac{1}{\mu_n N_D q}$, где N_D – концентрация доноров в полупроводнике n -типа в условиях полной ионизации доноров, равная концентрации свободных электронов n_0 .

1.8. Токи в полупроводниках

Как уже отмечалось выше, проводимость, а, следовательно, и ток в полупроводниках обусловлены двумя типами свободных носителей. Кроме этого, также есть две причины, обуславливающие появление электрического тока, – наличие

электрического поля и наличие градиента концентрации свободных носителей. С учетом сказанного плотность тока в полупроводниках в общем случае будет суммой четырех компонент:

$$\bar{J} = \bar{J}_p + \bar{J}_n = \bar{j}_{pE} + \bar{j}_{pD} + \bar{j}_{nE} + \bar{j}_{nD}, \quad (1.23)$$

где \bar{J} – плотность тока; \bar{j}_{nE} – дрейфовая компонента электронного тока; \bar{j}_{nD} – диффузионная компонента электронного тока; \bar{j}_{pE} – дрейфовая компонента дырочного тока; \bar{j}_{pD} – диффузионная компонента дырочного тока.

Выражение для каждой из компонент тока дается следующими соотношениями:

$$\bar{j}_{nE} = q\mu_n nE = \sigma_n E; \quad (1.24, a)$$

$$\bar{j}_{pE} = q\mu_p pE = \sigma_p E;$$

$$\bar{j}_{nD} = -qD_n \frac{dn}{dx}; \quad (1.24, б)$$

$$\bar{j}_{pD} = -qD_p \frac{dp}{dx},$$

где D_n – коэффициент диффузии электронов, связанный с подвижностью электронов μ_n соотношением $D_n = \frac{kT}{q} \mu_n$. Аналогичные соотношения существуют для коэффициентов диффузии дырок D_p и подвижности дырок μ_p .

Таким образом, плотность тока в полупроводнике в одномерном случае

$$\bar{j} = q\mu_n nE + q\mu_p pE - qD_n \frac{dn}{dx} - qD_p \frac{dp}{dx}. \quad (1.25)$$

Это выражение называется уравнением плотности тока.

1.9. Неравновесные носители. Генерация и рекомбинация носителей

Образование свободных носителей заряда в полупроводниках связано с переходом электронов из валентной зоны в зону проводимости. Для осуществления такого перехода электрон должен получить энергию, достаточную для преодоления запрещенной зоны. Эту энергию электрон получает от ионов решетки, совершающих тепловые колебания. Таким образом, преодоление запрещенной зоны электроном происходит обычно за счет тепловой энергии решетки. Концентрация носителей заряда, вызванная термическим возбуждением в состоянии теплового равновесия, называется *равновесной*.

Однако помимо теплового возбуждения появление свободных носителей заряда может быть связано с другими причинами, например, в результате облучения фотонами или частицами большой энергии, ударной ионизации, введения носителей заряда в полупроводник из другого тела (инжекция). *Инжекцией* неосновных носителей называется введение неосновных носителей заряда. Возникшие таким образом избыточные носители заряда называются *неравновесными*. Таким образом, полная концентрация носителей заряда равна

$$n = n_0 + \Delta n ; \quad (1.26)$$

$$p = p_0 + \Delta p , \quad (1.27)$$

где n_0 и p_0 – равновесная концентрация, n и p – неравновесная концентрация электронов и дырок, а Δn и Δp - концентрация избыточных электронов и избыточных дырок. Если возбуждение избыточных электронов производилось из валентной зоны, а полупроводник однородный и не содержит объемного заряда, то концентрация избыточных электронов равна концентрации избыточных дырок:

$$\Delta n = \Delta p . \quad (1.28)$$

После прекращения действия механизма, вызвавшего появление неравновесной концентрации носителей, происходит постепенное возвращение к равновесному состоянию. Процесс установления равновесия заключается в том, что каждый избыточный электрон при встрече с вакантным местом (дыркой) занимает его, в результате чего пара неравновесных носителей исчезает. Явление исчезновения пары носителей получило название *рекомбинации*. В свою очередь, возбуждение электрона из валентной зоны или примесного уровня, сопровождающееся появлением дырки, называется *генерацией* носителей заряда. На рис. 1.7 G – это темп генерации, а R – темп рекомбинации свободных носителей заряда в собственном полупроводнике.

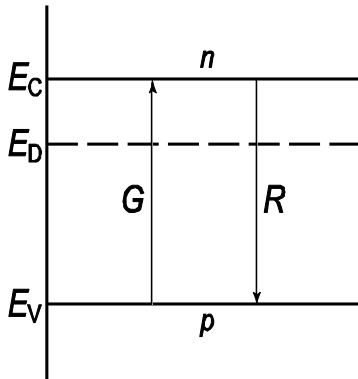


Рис. 1.7. Генерация и рекомбинация свободных электронов и дырок в полупроводниках

Для количественного описания процессов рекомбинации носителей заряда вводятся понятия времени жизни τ и диффузионной длины L .

Время жизни неосновных носителей заряда – это время, в течение которого избыточная концентрация в месте введения носителей уменьшается в e раз вследствие рекомбинации.

Диффузионная длина равна среднему расстоянию, на котором избыточная концентрация носителей уменьшается в e раз вследствие рекомбинации.

Диффузионная длина связана с временем жизни носителей соотношениями:

$$L_n = \sqrt{D_n \tau_n}; \quad L_p = \sqrt{D_p \tau_p}, \quad (1.29)$$

где D_n и D_p – коэффициенты диффузии электронов и дырок, связанные с подвижностью соотношениями Эйнштейна

$$D_n = \frac{kT}{q} \mu_n; \quad D_p = \frac{kT}{q} \mu_p. \quad (1.30)$$

1.10. Уравнение непрерывности

Динамика изменения концентрации неравновесных носителей по времени при наличии генерации и рекомбинации в полупроводнике, а также при протекании электрического тока определяется уравнением непрерывности. Для полупроводника n -типа уравнение непрерывности будет описывать динамику изменения концентрации дырок p_n :

$$\frac{\partial p_n}{\partial t} = G_p - R_p - \frac{1}{q} \operatorname{div}(J_p), \quad (1.31)$$

где J_p – дырочный ток, включающий дрейфовую и диффузионную компоненту; G_p – темп генерации неравновесных носителей; R_p – темп рекомбинации.

Уравнение непрерывности – это уравнение сохранения числа частиц в единице объема. Это уравнение показывает, как и по каким причинам изменяется концентрация неравновесных дырок со временем. Во-первых, концентрация дырок может изменяться из-за генерации (ударная ионизация, ионизация под действием света). Во-вторых, концентрация дырок может изменяться из-за их рекомбинации, что учитывает второе слагаемое. В-третьих, концентрация дырок может изменяться из-за дивергенции потока дырок, что учитывает третье слагаемое.

Уравнение непрерывности для электронов имеет вид

$$\frac{\partial n_p}{\partial t} = G_n - R_n - \frac{1}{q} \operatorname{div}(J_n), \quad (1.32)$$

где J_n – электронный ток, включающий дрейфовую и диффузионную компоненту; G_n – темп генерации неравновесных носителей; R_n – темп рекомбинации.

В условиях термодинамического равновесия рекомбинация электронов полностью уравнивает их тепловую генерацию, поэтому $G_n = R_n$. Если нет ударной и оптической генерации, то генерация возможна только за счет тепловой энергии. В этом случае можно говорить о результирующем эффекте генерации - рекомбинации, введя обозначение $RG = R_n - G_n$.

Из многих известных механизмов рекомбинации в кремнии и германии доминирующим является механизм рекомбинации носителей заряда через центры рекомбинации (ловушки).

Центры рекомбинации - это атомы золота, платины, серебра, дефекты кристаллической решетки. Ловушки создают уровни в запрещенной зоне. Наиболее эффективны с точки зрения рекомбинации те уровни, которые близки к середине запрещенной зоны.

В модели рекомбинации носителей заряда через ловушки, имеющие один единственный уровень $E_{л}$ в запрещенной зоне (модель Шокли - Рида - Холла), темп генерации-рекомбинации носителей заряда определяется выражением

$$RG = \frac{p \cdot n - n_i^2}{\tau_{no}(p + p_{л}) + \tau_{po}(n + n_{л})}, \quad (1.33)$$

где $n = n_0 + \Delta n$, $p = p_0 + \Delta p$, $n_i^2 = n_0 p_0$; n_0, p_0 - равновесные концентрации электронов в зоне проводимости и дырок в валентной зоне; $\Delta n, \Delta p$ - концентрация неравновесных (избыточных) носителей; n, p - полная концентрация электронов и дырок; n_i - собственная концентрация носителей; $n_{л}, p_{л}$ - равновесная концентрация электронов в зоне проводимости и дырок в валентной зоне, когда уровень Ферми совпадает с уровнем ловушек $E_{л}$; τ_{po}, τ_{no} - время жизни неосновных неравновесных носителей в полупроводниках n и p - типа.

Допустим, что избыточные дырки и электроны введены в равном количестве $\Delta n = \Delta p$. При таких условиях выражение (1.33) можно упростить. Допустим, что $\Delta p \ll n_0$, то есть избыточная концентрация неосновных носителей дырок - много меньше равновесной концентрации основных носителей. В этом случае говорят о низком уровне инжекции (НУИ). Выражение (1.33) принимает вид:

в полупроводнике n - типа

$$RG = \frac{p - p_0}{\tau_p} = \frac{\Delta p}{\tau_p}; \quad (1.34)$$

в полупроводнике p - типа

$$RG = \frac{n - n_0}{\tau_n} = \frac{\Delta n}{\tau_n}, \quad (1.35)$$

где τ_p, τ_n - время жизни дырок и электронов при низком уровне инжекции.

При высоком уровне инжекции (ВУИ) дырок в полупроводнике n - типа, когда $\Delta p \gg n_0$, преобразование (1.33) дает: в полупроводнике n - типа

$$RG = \frac{p - p_0}{\tau_{pв}} = \frac{\Delta p}{\tau_{pв}}; \quad (1.36)$$

в полупроводнике p - типа

$$RG = \frac{n - n_0}{\tau_{nв}} = \frac{\Delta n}{\tau_{nв}}, \quad (1.37)$$

где $\tau_{pв}, \tau_{nв}$ - время жизни дырок и электронов при высоком уровне инжекции.

Таким образом, после упрощения уравнения непрерывности (1.31) и (1.32) для дырок и электронов имеют следующий вид:

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

$$\frac{\partial n_p}{\partial t} = -\frac{n - n_0}{\tau_n} - \frac{1}{q} \operatorname{div}(J_n). \quad (1.39)$$

Подставим в эти выражения значения плотностей дрейфового и диффузионного тока электронов и дырок, получим уравнения непрерывности в следующем виде:

$$\frac{\partial p_n}{\partial t} = -\frac{p - p_0}{\tau_p} + D_p \frac{d^2 p}{dx^2} + \mu_p E \frac{dp}{dx} + p \mu_p \frac{dE}{dx}; \quad (1.40)$$

$$\frac{\partial n_p}{\partial t} = -\frac{n-n_0}{\tau_n} + D_n \frac{d^2 n}{dx^2} + \mu_n E \frac{n}{dx} + n\mu_n \frac{dE}{dx}. \quad (1.41)$$

Решение этих уравнений не всегда возможно и обычно допускаются упрощения. В случае, когда электрическое поле отсутствует или его влиянием можно пренебречь, уравнения (1.40) и (1.41) упрощаются:

$$\frac{\partial p_n}{\partial t} = -\frac{p-p_0}{\tau_p} + D_p \frac{d^2 p}{dx^2}; \quad (1.42)$$

$$\frac{\partial n_p}{\partial t} = -\frac{n-n_0}{\tau_n} + D_n \frac{d^2 n}{dx^2}. \quad (1.43)$$

Эти уравнения называют уравнениями диффузии и используют для анализа физических явлений в некоторых полупроводниковых приборах.

1.11. Электрические поля в кристаллах

На характеристики и свойства полупроводниковых приборов заметно влияет электрическое поле, которое возникает в результате создания в кристалле разности потенциалов от внешнего источника напряжения («внешнее» поле) или неравномерного распределения зарядов в кристалле за счет неравномерного распределения примеси («встроенное» поле).

Электрическое поле и его связь с зарядом описывается уравнением Пуассона

$$\operatorname{div} \bar{E} = -\frac{\rho_{ид}}{\varepsilon \varepsilon_0}, \quad (1.44)$$

где E - напряженность электрического поля, $E = \operatorname{grad} \varphi$, φ - электростатический потенциал; ε - относительная диэлектри-

ческая проницаемость среды; ε_0 - диэлектрическая проницаемость вакуума; $\rho_{об}$ - плотность объемного заряда, определяемая концентрациями электронов, дырок, доноров, акцепторов:

$$\rho_{об} = q(p + N_d - n - N_a). \quad (1.45)$$

В одномерном случае, когда напряженность электрического поля меняется вдоль оси x (приближение справедливо для большинства типов приборов), уравнение Пуассона имеет вид

$$\operatorname{div} \bar{E} = -\frac{q}{\varepsilon \varepsilon_0} (p + N_d - n - N_a). \quad (1.46)$$

Основная система уравнений для анализа работы полупроводниковых приборов определяет статику и динамику поведения носителей тока в полупроводниках, которые находятся под действием внешних полей. Основная система уравнений состоит из пяти уравнений: два уравнения плотности токов, два уравнения непрерывности и уравнение Пуассона. Эту систему уравнений часто называют **фундаментальной системой уравнений** (ФСУ) физики полупроводниковых приборов. В одномерном приближении ФСУ принимает вид:

$$\begin{aligned} \bar{j}_n &= q\mu_n nE - qD_n \frac{dn}{dx}; \\ \bar{j}_p &= q\mu_p pE - qD_p \frac{dp}{dx}; \\ \frac{\partial p_n}{\partial t} &= -\frac{p - p_0}{\tau_p} + D_p \frac{d^2 p}{dx^2}; \\ \frac{\partial n_p}{\partial t} &= -\frac{n - n_0}{\tau_n} + D_n \frac{d^2 n}{dx^2}; \\ \varepsilon \varepsilon_0 \frac{dE}{dx} &= q(p + N_d - n - N_a). \end{aligned}$$

2. КОНТАКТНЫЕ ЯВЛЕНИЯ

2.1. Разновидности электрических переходов и контактов

Основным элементом структуры большинства полупроводниковых приборов является электрический переход. *Электрический переход* – это переходный слой между областями твердого тела с различными типами проводимости или различными значениями удельной электрической проводимости.

Переход между двумя областями полупроводника, одна из которых имеет электропроводность p -типа, а другая n -типа называют *электронно-дырочным* или *p - n -переходом*.

Переходы между двумя областями с одним типом электропроводности (n - или p -типом), отличающиеся концентрацией примесей и соответственно значением удельной проводимости, называют *изотипными* переходами: *электронно-электронными* (n^+ - n -переход) или *дырочно-дырочными* (p^+ - p -переход). Термином n^+ и p^+ обозначают сильнолегированные области, следовательно, слои n^+ и p^+ имеют меньшее удельное сопротивление, поэтому большую удельную проводимость.

Переходы, образованные в одном полупроводниковом материале, например германии, кремнии, арсениде галлия, называют *гомпереходами*.

Переходы между двумя полупроводниковыми материалами, имеющими различную ширину запрещенной зоны, называют *гетеропереходами*.

Если одна из областей, образующих переход, является металлом, а другая полупроводником, то такие переходы называют *переходом металл-полупроводник* или *переходом Шоттки*.

К электрическим переходам относят также структуру *металл – диэлектрик – полупроводник* (МДП), в которой

между металлом и полупроводником расположен тонкий слой диэлектрика.

По функциональному назначению электрические переходы делят на *выпрямляющие*, электрическое сопротивление которых при одном (прямом) направлении тока меньше, чем при другом – обратном и *омические*, электрическое сопротивление которых мало и практически не зависит от направления и значения тока в заданном диапазоне токов.

2.2. Электронно-дырочный переход

Электронно-дырочный переход (*p-n*-переход) находится на границе между двумя областями полупроводника, одна из которых имеет электронную (*n*), а другая – дырочную (*p*) электрические проводимости, то есть соответственно *n*- или *p*-области. Однако его нельзя создать простым соприкосновением полупроводниковых пластин *n*- или *p*-типов, так как при этом неизбежен промежуточный слой воздуха, оксидов или поверхностных загрязнений. Переход создается в кристалле полупроводника с помощью технологических процессов (например, сплавления, диффузии), в результате которых граница раздела между областями *p*- и *n*-типов находится внутри полупроводникового монокристалла. Классическим примером *p-n* перехода являются: *nSi – pSi*, *nGe – pGe*.

Различают симметричные и несимметричные *p-n*-переходы. В *симметричных* переходах концентрация электронов в полупроводнике *n*-типа n_n и концентрация дырок в полупроводнике *p*-типа p_p равны, то есть $n_n = p_p$. Другими словами, концентрации основных носителей зарядов по обе стороны симметричного *p-n*-перехода, равны. На практике используются, как правило, *несимметричные* переходы, в которых концентрация, например, электронов в полупроводнике *n*-типа больше концентрации дырок в полупроводнике *p*-типа, то есть $n_n > p_p$, при этом различие в концентрациях может со-

ставлять 100 - 1000 раз. Низкоомная область, сильно легированная примесями (например, n -область в случае $n_n > p_p$), называется **эмиттером**; высокоомная, слаболегированная (p -область в случае перехода $n_n > p_p$), - **базой**. Для случая, когда концентрации электронов в полупроводнике n -типа больше концентрации электронов в полупроводнике p -типа, то есть $p_p > n_n$, эмиттером будет p -область, а базой - n -область.

В зависимости от характера примесей, обеспечивающих требуемый тип электропроводности в областях, различают два типа переходов: резкий (ступенчатый) и плавный (линейный). В **резком** переходе концентрация примесей на границе раздела областей изменится на расстоянии, соизмеримом с диффузионной длиной, в **плавном** – на расстоянии, значительно большем диффузионной длины. Лучшим выпрямительными (вентильными) свойствами обладают резкие p - n -переходы. Резкий p - n -переход образуется при сплавлении, плавный – получается методом диффузии или методом выращивания из расплава.

В зависимости от площади p - n -переходы разделяются на **точечные** и **плоскостные**. Точечные p - n -переходы образуются точечно-контактным способом. Плоскостные переходы в зависимости от метода их изготовления бывают сплавными, диффузионными, эпитаксиальными и т.п.

2.2.1. p - n -переход в состоянии термодинамического равновесия

В полупроводнике с областями p - и n -типов, образующих переход можно выделить следующие пространственные области (рис. 2.1): 1 – металлургический переход (контакт) – воображаемая плоскость, разделяющая p - и n - области; 2 – область перехода или область пространственного заряда (ОПЗ) или обедненная область, располагается по обе стороны металлургического перехода и имеет толщину от 10^{-6} до

10^{-4} см (в зависимости от технологии производства); 3 – нейтральные p - и n -области, лежащие между областью пространственного и границами полупроводников p - и n -типов; 4 – омические контакты, которыми оканчиваются нейтральные области.

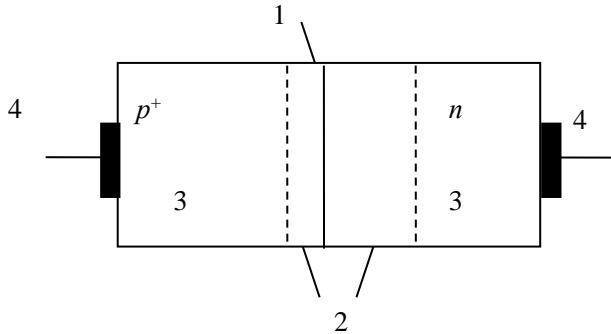


Рис. 2.1. Схематическое изображение идеального p - n -перехода

На рис. 2.2 условно показан кристалл, одна часть объема которого имеет дырочную электропроводность, а другая – электронную. До установления термодинамического равновесия между p - и n -областями и в отсутствии внешнего электрического поля в таком переходе протекают следующие физические процессы. Поскольку концентрация дырок в p -области гораздо выше их концентрации в n -области, то дырки из p -области диффундируют в n -область. Однако, как только дырки попадают в n -область, они начинают рекомбинировать с электронами, основными носителями зарядов в n -области и их концентрация по мере углубления быстро убывает. Аналогично электроны из n -области диффундируют в p -область. Такое движение зарядов создает диффузионный ток электронов и дырок. Плотность полного диффузионного тока, проходящего через границу раздела, определится суммой

$$j_{\vec{a}\vec{e}\vec{o}} = j_{\vec{a}\vec{e}\vec{o}n} + j_{\vec{a}\vec{e}\vec{o}p} = q(D_n \frac{dn}{dx} + D_p \frac{dp}{dx}).$$

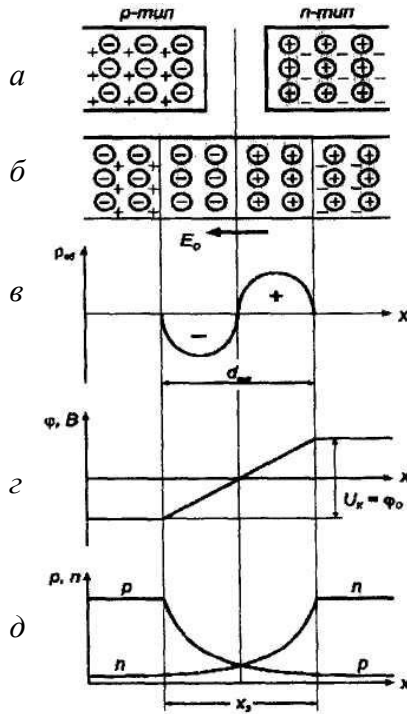


Рис. 2.2. Идеальный плоскостной p - n -переход: *а* - отдельные p - и n -полупроводники; *б* - схематическое изображение идеального плоскостного p - n -перехода; *в* - распределение плотности объемных зарядов; *г* - распределение потенциала; *д* - распределение электронов проводимости и дырок; +, - - ионы; «+», «-» - дырки и электроны

Встречная диффузия подвижных носителей заряда приводит к появлению в n -области нескомпенсированных положительных зарядов ионов донорной примеси, а в p -области – отрицательных зарядов ионов акцепторной примеси, связанной с кристаллической решеткой полупроводника (рис. 2.2, б).

Распределение объемной плотности указанных зарядов $\rho_{об}$ показано на рис. 2.2, в. Таким образом, на границе областей образуются два слоя зарядов, равных по величине, но противоположных по знаку. Положительный заряд обусловлен ионизованными донорами, отрицательный заряд – ионизованными акцепторами. Образуется **область пространственного заряда** (ОПЗ), состоящая из двух разноименно заряженных слоев. Образовавшаяся область пространственных зарядов и представляет собой *p-n*-переход. Его ширина обычно равна $d_{p-n} = 10^{-3} - 10^{-4}$ мм. Объемные (пространственные) заряды в переходе образуют электрическое поле, направленное от положительно заряженных доноров к отрицательно заряженным акцепторам, то есть от *n*-области к *p*-области и называемое диффузионным электрическим полем. Этому полю соответствует контактная разность потенциалов $U_k = \varphi_0$, зависящая от материала и уровня легирования. Например, U_k для германиевых *p-n*-переходов составляет (0,3 - 0,4) В, а для кремниевых (0,7 - 0,8) В. Так как электрическое поле неподвижных зарядов *p-n*-перехода при термодинамическом равновесном состоянии препятствует диффузии основных носителей заряда в соседнюю область, то считают, что между *p*- и *n*-областями устанавливается потенциальный барьер, φ_0 , распределение потенциала которого вдоль структуры *p-n*-перехода показано на рис. 2.2, г.

Основные носители заряда при встречной диффузии рекомбинируют в приконтактных областях *p-n*-перехода, что приводит к образованию в этом месте обедненного подвижными носителями заряда слоя, который обладает малой удельной проводимостью (как беспримесный или собственный полупроводник) и поэтому называется **обедненным** или **запирающим** слоем x_3 (рис. 2.2, д).

Диффузионное электрическое поле является тормозящим для основных и ускоряющим для неосновных носителей заряда. Электроны *p*-области и дырки *n*-области, то есть неос-

новые носители, совершая тепловое, движение попадают в пределы диффузионного электрического поля, увлекаются им и перебрасываются в противоположные области образуя ток дрейфа. Выведение носителей заряда из области полупроводника, где они являются неосновными, через электронно-дырочный переход, ускоряющим электрическим полем называется *экстракцией*. То есть в поле напряженностью E_0 возникают дрейфовые токи электронов и дырок

$$j_{\dot{a}p} = j_{\dot{a}pn} + j_{\dot{a}pp} = q\mu_n nE + q\mu_p pE.$$

Движение зарядов прекращаются, когда наступает динамическое равновесие и токи через переход уравниваются $j_{\dot{a}e\delta} + j_{\dot{a}p} = 0$. Дрейфовые токи направлены навстречу диффузионным и равны им. При отсутствии внешнего поля результирующий ток через p - n -переход для каждого типа носителей равен нулю.

На рис. 2.3 приведены зонные диаграммы, иллюстрирующие этапы формирования электронно-дырочного перехода.

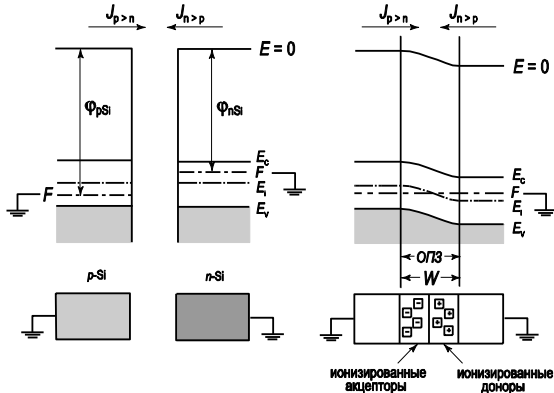


Рис. 2.3. Схема, иллюстрирующая образование p - n -перехода

Граница областей донорной и акцепторной примеси в полупроводнике получила название металлургического p - n -перехода. Границу, где уровень Ферми пересекает середину запрещенной зоны, называют физическим p - n -переходом.

Энергетическая диаграмма p - n -перехода в состоянии равновесия строится, начиная с уровня Ферми, единого для всей системы (рис. 2.4). В каждой области положение энергетических уровней не изменяется, поэтому после «соприкосновения» полупроводников p - и n -типа, когда уровень Ферми оказывается общим, энергетические зоны изгибаются в области перехода.

2.2.2. Контактная разность потенциалов

Важным параметром p - n -перехода является **контактная разность потенциалов** $U_k = U_0 = \varphi_0$. Можно рассчитать контактную разность потенциалов, используя энергетическую диаграмму p - n -перехода.

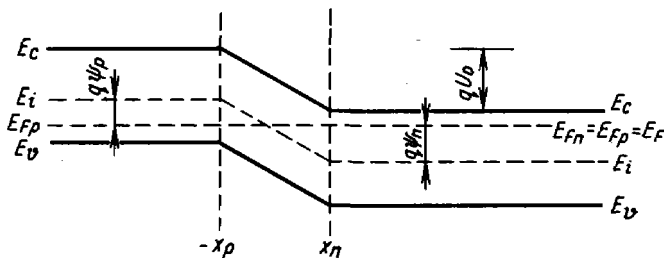


Рис. 2.4. Энергетическая диаграмма p - n -перехода в состоянии термодинамического равновесия

Из рис. 2.4 следует, что величина потенциального барьера

$$qU_0 = q\varphi_0 = (E_{ip} - E_{Fp}) - (E_{Fn} - E_{in}),$$

$$\varphi_0 = \frac{1}{q} [(E_{ip} - E_{Fp}) + (E_{Fn} - E_{in})] \quad (2.1)$$

Учитывая, что формулы, определяющие концентрации основных носителей через уровень Ферми данного полупроводника и уровень Ферми собственного полупроводника и концентрацию носителей в собственном полупроводнике имеют вид

$$p_{p0} = n_i \exp[(E_{ip} - E_{Fp})/kT]; \quad n_{p0} = n_i \exp[(E_{Fn} - E_{in})/kT] \quad (2.2)$$

и, что

$$p_{p0} = N_A; \quad n_{p0} = N_D; \quad n_{n0} \cdot p_{n0} = p_{p0} \cdot n_{p0} = (n_i)^2 \quad (2.3)$$

получаем, что контактная разность потенциалов или высота потенциального барьера

$$\varphi_0 = \frac{kT}{q} \ln \frac{n_{n0} p_{p0}}{n_i^2} = \varphi_0 \ln \frac{n_{n0} p_{p0}}{n_i^2} = \varphi_0 \ln \frac{N_A N_D}{n_i^2}. \quad (2.4)$$

2.2.3. Ширина p - n -перехода

Другим важным параметром p - n -перехода является его ширина, обозначенная $W = W_p + W_n$. Для нахождения W можно воспользоваться уравнением Пуассона. Уравнение Пуассона описывает связь электрического поля и потенциала в p - n -переходе. В одномерном приближении это уравнение имеет вид

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{\rho(x)}{\varepsilon_s \varepsilon_0}, \quad (2.5)$$

где $\psi(x)$ – зависимость потенциала от координаты; $\rho(x)$ – плотность объемного заряда; ε_s – диэлектрическая проницаемость полупроводника; ε_0 – диэлектрическая постоянная.

Для решения этого уравнения выберем начало координат в области металлургического p - n -перехода. При этом донорный полупроводник будет находиться в области $x > 0$ (в дальнейшем обозначим цифрой I), а акцепторный – в области $x < 0$ (в дальнейшем обозначим цифрой II).

Заряд в ОПЗ p - n -перехода для полупроводника n -типа обусловлен зарядом ионизованных доноров с плотностью N_D^+ , для полупроводника p -типа – зарядом ионизованных акцепторов с плотностью N_A^+ . Поэтому для области I $\rho(x) = qN_D^+$, для области II $\rho(x) = qN_A^+$. Будем решать уравнение Пуассона отдельно для областей I и II. После интегрирования уравнения Пуассона получаем для области I:

$$E(x) = -\frac{qN_D^+}{\varepsilon_s \varepsilon_0} (W_n - x), \quad (2.6)$$

для области II:

$$E(x) = -\frac{qN_A^+}{\varepsilon_s \varepsilon_0} (W_p + x). \quad (2.7)$$

Знак минус в выражениях (2.6), (2.7) указывает, что направление электрического поля противоположно направлению оси x .

Из соотношений (2.6) и (2.7) следует, что электрическое поле E максимально на металлургической границе p - n -перехода ($x = 0$), линейно спадает по области пространственного заряда и равно нулю на границах ОПЗ – квазинейтральный объем полупроводника ($x = W_n$; $x = W_p$).

Максимальная величина электрического поля E_{\max} будет равна

$$E_{\max} = \frac{qN_A W_p}{\varepsilon_s \varepsilon_0} = \frac{qN_D W_n}{\varepsilon_s \varepsilon_0}. \quad (2.8)$$

Для нахождения распределения потенциала (а, следовательно, и зависимости потенциальной энергии от координаты) проинтегрируем еще раз уравнение (2.7) при следующих граничных условиях: $x = W$, $\psi(W) = 0$. Получаем

$$\psi(x) = \frac{qN_A}{\varepsilon_s \varepsilon_0} \left(\frac{x^2}{2} + W_p x \right) + const, \quad x < 0. \quad (2.9)$$

Используя граничные условия $x = -W_p$; $\psi = \Delta\varphi_0$, находим константу интегрирования:

$$const = -\frac{qN_A}{\varepsilon_s \varepsilon_0} \left(\frac{W_p^2}{2} - W_p^2 \right) = \frac{qN_A}{\varepsilon_s \varepsilon_0} \frac{W_p^2}{2} + \Delta\varphi_0. \quad (2.10)$$

Подставляя полученные значения константы в соотношение (2.9), получаем для распределения потенциала $\psi(x)$ в области $x < 0$:

$$\psi(x) = \frac{qN_A}{\varepsilon_s \varepsilon_0} \left(x^2 + 2W_p x + \frac{W_p^2}{2} \right) = \frac{qN_A}{2\varepsilon_s \varepsilon_0} (x + W_p)^2 + \Delta\varphi_0. \quad (2.11)$$

Проводя аналогичное интегрирование для области $x > 0$, получаем

$$\psi(x) = -\frac{qN_D}{\varepsilon_s \varepsilon_0} \left(\frac{x^2}{2} - W_n x \right) + const, \quad x > 0. \quad (2.12)$$

Используя граничные условия $x = -W_n$, $\psi = 0$, для константы интегрирования в этой области получаем:

$$const = \frac{qN_D}{\varepsilon_s \varepsilon_0} \left(\frac{W_n^2}{2} - W_n^2 \right) = -\frac{qN_D}{\varepsilon_s \varepsilon_0} \frac{W_n^2}{2}. \quad (2.13)$$

Подставляя полученные значения константы в соотношение (2.12), получаем для распределения потенциала $\psi(x)$ в области $x > 0$:

$$\psi(x) = -\frac{qN_D}{2\varepsilon_s \varepsilon_0} (x^2 - 2W_n x + W_n^2) = -\frac{qN_D}{2\varepsilon_s \varepsilon_0} (x + W_n)^2. \quad (2.14)$$

Таким образом, закон изменения потенциала ψ в p -области (отсчет идет от уровня в квазинейтральной области):

$$\psi_1(x) = \frac{qN_A}{2\varepsilon_s \varepsilon_0} (x + W_p)^2, \quad x < 0, \quad (2.15)$$

и, наоборот, в n -области:

$$\psi_2(x) = -\frac{qN_D}{2\varepsilon_s \varepsilon_0} (x - W_n)^2, \quad x > 0. \quad (2.16)$$

На рис. 2.5 приведена диаграмма, иллюстрирующая распределение электрического поля и потенциала в p - n -переходе, рассчитанная по соотношениям (2.6), (2.7), (2.15) и (2.16).

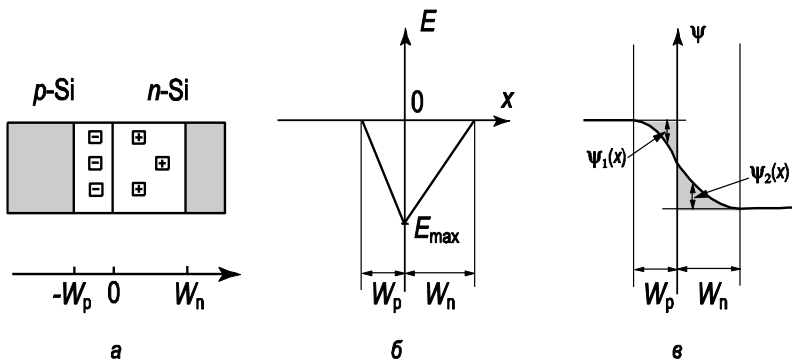


Рис. 2.5. Диаграмма, иллюстрирующая распределение электрического поля и потенциала в p - n -переходе: a - структура p - n -перехода; b - распределение электрического поля в ОПЗ; $в$ - распределение потенциала в ОПЗ

На металлургической границе p - n -перехода при $x = 0$ значение потенциала $\psi_1 + \psi_2 = \Delta\varphi_0 = \varphi_{n0} + \varphi_{p0}$, или

$$\Delta\varphi_0 = \frac{q}{2\varepsilon_s\varepsilon_0} (N_A W_p^2 + N_D W_n^2). \quad (2.17)$$

Согласно уравнению электронейтральности в замкнутых системах величины положительного и отрицательного заряда на единицу площади должны быть равны:

$$Q_D = Q_A; \quad qN_A W_p = qN_D W_n. \quad (2.18)$$

Следовательно,

$$W_n = \frac{N_A W_p}{N_D}. \quad (2.18, a)$$

Подставляем выражение (2.18, *a*) в (2.17), получаем

$$\begin{aligned}\Delta\varphi_0 &= \frac{q}{2\varepsilon_s\varepsilon_0} \left(N_A W_p^2 + N_A W_p \frac{N_A W_p}{N_D} \right) = \frac{q}{2\varepsilon_s\varepsilon_0} W_p^2 \left(N_A + \frac{N_A^2}{N_D} \right) = \\ &= \frac{q}{2\varepsilon_s\varepsilon_0} W_p^2 N_A^2 \left(\frac{1}{N_A} + \frac{1}{N_D} \right).\end{aligned}$$

Несложные преобразования позволяют получить выражение для ширины обедненных областей W_p и W_n в p - и n -областях соответственно:

$$W_p = \sqrt{\frac{2\varepsilon_s\varepsilon_0\Delta\varphi}{qN_A^2\left(\frac{1}{N_A} + \frac{1}{N_D}\right)}}; \quad W_n = \sqrt{\frac{2\varepsilon_s\varepsilon_0\Delta\varphi}{qN_D^2\left(\frac{1}{N_A} + \frac{1}{N_D}\right)}}. \quad (2.19)$$

Из предыдущих формул легко видеть, что с ростом легирования p -области ширина p - n -перехода W_p в акцепторной части полупроводника уменьшится.

Полная ширина p - n -перехода W , равная $W = W_p + W_n$, будет вычисляться по формуле

$$W = \sqrt{\frac{2\varepsilon_s\varepsilon_0\Delta\varphi_0}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}. \quad (2.20)$$

Для несимметричных p^+ - n -переходов (концентрация акцепторов существенно больше концентрации доноров) из соотношений (2.18) и (2.18, *a*) следует, что ширина обедненной области в полупроводнике p -типа будет существенно меньше, чем ширина обедненной области в полупроводнике n -типа:

$$N_A \gg N_D \rightarrow W_p \ll W_n .$$

Таким образом, вся обедненная область p^+n -перехода сосредоточена в области с низким значением концентрации легирующей примеси $W = W_n$.

Формула (2.20) определяет полную толщину резкого p - n -перехода. Реальные p - n -переходы бывают несимметричными, то есть $N_A \gg N_D$ или $N_D \gg N_A$, поэтому толщина резкого несимметричного p - n -перехода

$$W = \sqrt{\frac{2\varepsilon_s \varepsilon_0 \Delta\varphi_0}{qN}} , \quad (2.21)$$

где N – концентрация примеси в слабелегированной области.

Толщина плавного p - n -перехода с линейным распределением примеси

$$W = \sqrt[3]{\frac{12\varepsilon_s \varepsilon_0 \Delta\varphi_0}{qa}} , \quad (2.22)$$

где a – градиент концентрации примеси.

2.2.4. Прямое включение p - n -перехода

При использовании p - n -перехода в полупроводниковых приборах к нему подключается внешнее напряжение. Величина и полярность этого напряжения определяет электрический ток, проходящий через p - n -переход.

Пусть источник внешнего напряжения подключен положительным полюсом к полупроводнику p -типа (рис. 2.6, a). Та-

кое напряжение, у которого полярность совпадает с полярностью основных носителей, называется *прямым*.

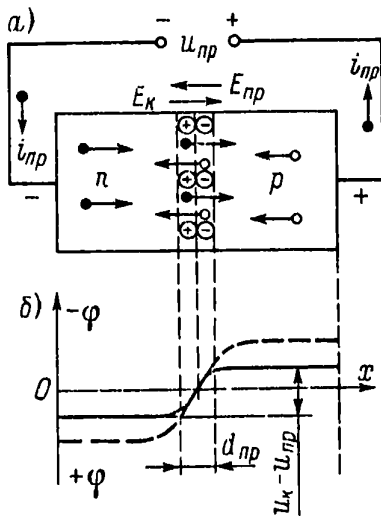


Рис.2.6. Электронно-дырочный переход при прямом напряжении

Действие прямого напряжения $u_{пр}$, вызывающее прямой ток $i_{пр}$ через переход, поясняется потенциальной диаграммой на рис. 2.6, б. (На этом и следующих рисунках потенциальная диаграмма изображена упрощенно. Для рассмотрения p - n -перехода процессы в остальных частях цепи не представляют интереса. Поэтому на диаграммах не показано изменение потенциала вдоль n - и p -областей, то есть их сопротивление принято равным нулю. Не показано также изменение потенциала в контактах областей n и p с электродами, к которым присоединены провода от источника напряжения.)

Электрическое поле, создаваемое в p - n -переходе прямым напряжением, действует навстречу полю контактной разности потенциалов. Это показано на рисунке векторами E_k и $E_{пр}$. Ре-

зультирующее поле становится слабее, и разность потенциалов в переходе уменьшается, то есть высота потенциального барьера понижается, возрастает диффузионный ток, так как большее число носителей может преодолеть пониженный барьер. Ток дрейфа при этом почти не изменяется, так как он зависит главным образом от числа неосновных носителей, попадающих за счет своих тепловых скоростей на p - n -переход из n - и p -областей. Если пренебречь падением напряжения на сопротивлении областей n и p , то напряжение на переходе можно считать равным $u_k - u_{пр}$. Для сравнения на рис. 2.6, б штриховой линией приведена потенциальная диаграмма при отсутствии внешнего напряжения.

Как известно, в этом случае токи $i_{диф}$ и $i_{др}$ равны и компенсируют друг друга. При прямом напряжении $i_{диф} > i_{др}$ и поэтому полный ток через переход, т. е. прямой ток, уже не равен нулю:

$$i_{пр} = i_{диф} - i_{др} > 0 \quad (2.23)$$

Если барьер значительно понижен, то $i_{диф} \gg i_{др}$ и можно считать, что $i_{пр} \cong i_{диф}$, т.е. прямой ток в переходе является чисто диффузионным.

Введение носителей заряда через пониженный под действием прямого напряжения потенциальный барьер в область, где эти носители являются неосновными, называется **инжекцией носителей заряда**. Слово «инъекция» означает «введение, впрыскивание». Применение термина «инъекция» необходимо для того, чтобы отличать данное явление от электронной эмиссии, в результате которой получают свободные электроны в вакууме или разреженном газе. Область полупроводникового прибора, из которой инжектируются носители, называется **эмиттерной областью** или **эмиттером**. А область, в которую инжектируются неосновные для этой области носители заряда, называется **базовой областью** или **базой**. Таким образом, ес-

ли рассматривать инжекцию электронов, то n -область является эмиттером, а p -область - базой. Для инжекции дырок, наоборот, эмиттером служит p -область, а базой - n -область.

Обычно концентрация примесей, а следовательно, и основных носителей в n - и p -областях весьма различна. Поэтому инжекция электронов из области с более высокой концентрацией основных носителей преобладает. Соответственно этому области и называют «эмиттер» и «база». Например, если $n_n \gg p_p$, то инжекция электронов из n -области в p -область значительно превосходит инжекцию дырок в обратном направлении. В данном случае эмиттером считают n -область, а базой - p -область, так как инжекцией дырок можно пренебречь.

При прямом напряжении не только понижается потенциальный барьер, но также уменьшается толщина запирающего слоя ($d_{np} < d$) и его сопротивление в прямом направлении становится малым (единицы — десятки Ом).

Поскольку высота барьера u_k при отсутствии внешнего напряжения составляет несколько десятых долей вольта, то для значительного понижения барьера и существенного уменьшения сопротивления запирающего слоя достаточно подвести к p - n -переходу такое же прямое напряжение (десятые доли вольта). Поэтому большой прямой ток можно получить при очень небольшом прямом напряжении.

Очевидно, что при некотором прямом напряжении можно вообще уничтожить потенциальный барьер в p - n -переходе. Тогда сопротивление перехода, то есть запирающего слоя, станет близким к нулю и им можно будет пренебречь. Прямой ток в этом случае возрастет и будет зависеть только от сопротивления n - и p -области. Теперь уже этими сопротивлениями пренебрегать нельзя, так как именно они остаются в цепи и определяют силу тока.

Рассмотрим еще характер прямого тока в разных частях цепи (рис. 2.6, а). Электроны из n -области движутся через переход в p -область, а навстречу им из p -области в n -область пе-

ремещаются дырки, то есть через переход протекают два тока: электронный и дырочный. Во внешних проводниках, конечно, движутся только электроны. Они перемещаются в направлении от минуса источника к n -области и компенсируют убыль электронов, диффундирующих через переход в p -область. А из p -области электроны уходят по направлению к плюсу источника, и тогда в этой области образуются новые дырки. Такой процесс происходит непрерывно, и, следовательно, непрерывно протекает прямой ток.

У левого края n -области электронный ток имеет наибольшее значение. По мере приближения к переходу этот ток уменьшается, так как все большее число электронов рекомбинирует с дырками, движущимися через переход навстречу электронам, а дырочный ток i_p , наоборот, увеличивается. Полный прямой ток $i_{пр}$ в любом сечении, конечно, один и тот же:

$$i_{пр} = i_n + i_p = \text{const.} \quad (2.24)$$

Это следует из основного закона последовательной электрической цепи: во всех частях такой цепи ток всегда одинаков.

Так как толщина перехода очень мала, и он обеднен носителями, то в нем рекомбинирует мало носителей и ток здесь не изменяется. А далее электроны, инжектированные в p -область, рекомбинируют с дырками. Поэтому по мере удаления от перехода вправо в p -области ток i_n продолжает уменьшаться, а ток i_p увеличивается. У правого края p -области ток i_n наименьший, а ток i_p наибольший. На рис. 2.7 показано изменение этих токов вдоль оси x для случая, когда ток i_n преобладает над током i_p , вследствие того, что $n_n > p_n$ и подвижность электронов больше подвижности дырок. Конечно, при прямом напряжении кроме диффузионного тока есть еще

ток дрейфа, вызванный движением неосновных носителей. Но если он очень мал, то его можно не принимать во внимание.

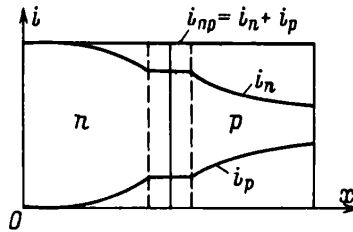


Рис. 2.7. Распределение электронного и дырочного тока в p - n -переходе

Энергетическая диаграмма p - n -перехода при прямом смещении (рис. 2.8) строится и в этом случае, начиная с уровня Ферми, но в p -области его следует опустить на величину приложенного напряжения U . В каждой из областей относительное положение энергетических уровней не изменится.

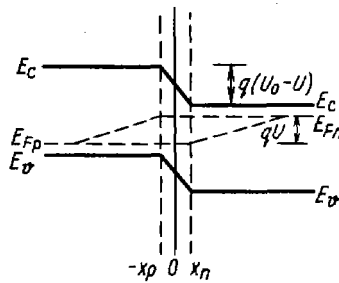


Рис. 2.8. Энергетическая диаграмма при прямом включении p - n -перехода

2.2.5. Уровень инжекции

Уровень инжекции – это отношение приращения концентрации неосновных носителей, полученного в результате

инжекции, к равновесной концентрации основных носителей заряда

$$Y = \frac{\Delta p_n}{n_{n0}} = \frac{\Delta n_p}{p_{p0}}, \quad (2.25)$$

где $\Delta p_n = p_n - p_{n0}$ и $\Delta n_p = n_p - n_{p0}$.

Сравнение приращения концентрации неосновных носителей с концентрацией основных носителей заряда объясняется следующим образом. Инжектированные неосновные носители заряда создают у границы перехода объемные заряды для компенсации, которых из областей полупроводника к границам перехода притекают основные носители заряда. То есть у границ перехода создаются избыточные концентрации не только неосновных, но и основных носителей зарядов $\Delta n_n \approx \Delta p_n$ в n -области и $\Delta p_p \approx \Delta n_p$ в p -области.

Уровень инжекции считают низким, а области называют квазинейтральными, если приращение концентрации инжектированных носителей много меньше концентрации основных носителей заряда

$$Y = \frac{\Delta p_n}{n_{n0}} \ll 1, Y = \frac{\Delta n_p}{p_{p0}} \ll 1,$$

то есть $\Delta p_n \ll n_{n0}$ для полупроводника n -типа и $\Delta n_p \ll p_{p0}$ для полупроводника p -типа.

Уровень инжекции считают средним, если приращение концентрации инжектированных носителей сравнимо с концентрацией основных носителей заряда

$$Y = \frac{\Delta p_n}{n_{n0}} \approx 1, Y = \frac{\Delta n_p}{p_{p0}} \approx 1,$$

то есть $\Delta p_n \approx \Delta n_n \approx n_{n0}$ для полупроводника n -типа и $\Delta n_p \approx \Delta p_p \approx p_{p0}$ для полупроводника p -типа.

Высоким уровнем инжекции будем считать такой, при котором приращение концентрации инжектированных носителей много больше концентрации основных носителей заряда

$$Y = \frac{\Delta p_n}{n_{n0}} \gg 1, Y = \frac{\Delta n_p}{p_{p0}} \gg 1,$$

то есть $\Delta p_n \gg n_{n0}$ для полупроводника n -типа и $\Delta n_p \gg p_{p0}$ для полупроводника p -типа.

В случае среднего и высокого уровня инжекции электронейтральность областей нарушается. Однако при рассмотрении физических процессов в p - n -переходе будем полагать уровень инжекции низким и области полупроводника за границами перехода электрически нейтральными.

2.2.6. Обратное включение p - n -перехода

Пусть источник внешнего напряжения подключен положительным полюсом к области n , а отрицательным — к области p (рис. 2.9, а). Под действием такого **обратного напряжения** $u_{обр}$ через переход протекает очень небольшой **обратный ток** $i_{обр}$, что объясняется следующим образом. Поле, создаваемое обратным напряжением, складывается с полем контактной разности потенциалов.

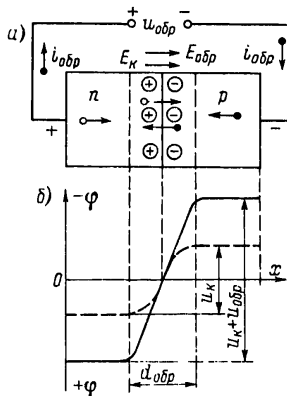


Рис. 2.9. Электронно-дырочный переход при обратном напряжении

На рис. 2.9, *a* это показывают одинаковые направления векторов E_k и $E_{обр}$. Результирующее поле усиливается, и высота потенциального барьера теперь равна $u_k + u_{обр}$ (рис. 2.9, *б*). Уже при небольшом повышении барьера диффузионное перемещение основных носителей через переход прекращается, то есть $i_{диф} = 0$, так как собственные скорости носителей недостаточны для преодоления барьера. А ток проводимости остается почти неизменным, поскольку он определяется главным образом числом неосновных носителей, попадающих на p - n -переход из n - и p -областей.

Выведение неосновных носителей через p - n -переход ускоряющим электрическим полем, созданным обратным напряжением, называют *экстракцией носителей заряда* (слово «экстракция» означает «выдергивание, извлечение»).

Таким образом, обратный ток $i_{обр}$ представляет собой ток проводимости, вызванный перемещением неосновных носителей. Обратный ток получается очень небольшим, так как неосновных носителей мало и, кроме того, сопротивление запирающего слоя при обратном напряжении очень велико. Действительно, при повышении обратного напряжения поле в месте перехода становится сильнее и под действием этого поля больше основных носителей «выталкивается» из пограничных слоев в глубь n - и p -областей. Поэтому с увеличением обратного напряжения увеличивается не только высота потенциального барьера, но и толщина запирающего слоя ($d_{обр} > d$). Этот слой еще сильнее обедняется носителями, и его сопротивление значительно возрастает, т. е. $R_{обр} \gg R_{np}$.

Уже при сравнительно небольшом обратном напряжении обратный ток становится практически постоянным. Это объясняется тем, что число неосновных носителей ограничено. С повышением температуры концентрация их возрастает,

и обратный ток увеличивается, а обратное сопротивление уменьшается.

Рассмотрим подробнее, как устанавливается обратный ток при включении обратного напряжения. Сначала возникает переходный процесс, связанный с движением основных носителей. Электроны в n -области движутся по направлению к положительному полюсу источника, то есть удаляются от p - n -перехода. А в p -области, удаляясь от p - n -перехода, движутся дырки. У отрицательного электрода они рекомбинируют с электронами, которые приходят из проводника, соединяющего этот электрод с отрицательным полюсом источника.

Поскольку из n -области уходят электроны, она заряжается положительно, так как в ней остаются положительно заряженные атомы донорной примеси. Подобно этому p -область заряжается отрицательно, так как ее дырки заполняются приходящими электронами и в ней остаются отрицательно заряженные атомы акцепторной примеси.

Рассмотренное движение основных носителей в противоположные стороны продолжается лишь малый промежуток времени. Такой кратковременный ток подобен зарядному току конденсатора. По обе стороны p - n -перехода возникают два разноименных объемных заряда, и вся система становится аналогичной заряженному конденсатору с диэлектриком, в котором имеется значительный ток утечки (его роль играет обратный ток). Но ток утечки конденсатора в соответствии с законом Ома пропорционален приложенному напряжению, а обратный ток p - n -перехода сравнительно мало зависит от напряжения.

Приведем энергетическую диаграмму p - n -перехода при обратном включении (рис. 2.10). Вначале изображают уровень Ферми для любой из двух нейтральных областей p - n -перехода и дополняют его соответствующими зоной проводимости и валентной зоной. После этого с учетом направления смещения строят квазиуровень Ферми для другой нейтральной области,

а затем соответствующие зону проводимости и валентную зону, причем ширина запрещенной зоны постоянна. При обратном смещении уровень Ферми в p -области следует поднять на величину приложенного напряжения U . И, наконец, зоны проводимости, валентные зоны и квазиуровни Ферми соединяют через обедненную область.

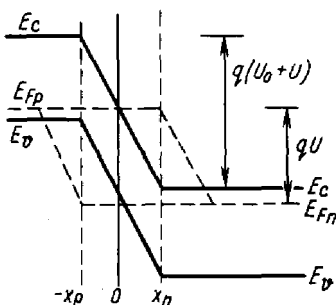


Рис. 2.10. Энергетическая диаграмма при обратном включении p - n -перехода

2.2.7. Теоретическая вольт-амперная характеристика p - n -перехода

Вольт-амперная характеристика представляет собой график зависимости тока во внешней цепи p - n -перехода от величины и полярности напряжения, прикладываемого к нему. Эта зависимость может быть получена экспериментально или рассчитана на основании уравнения вольт-амперной характеристике. Для облегчения расчетов обычно выбирают упрощенную модель реального p - n -перехода, то есть идеальный p - n -перехода.

p - n -переход считается идеальным при следующих условиях:

1) в обедненном слое нет генерации, рекомбинации и рассеяния носителей; носители проходят через обедненный

слой мгновенно, то есть токи носителей одного знака на обеих границах одинаковы;

2) вне обедненного слоя нет электрического поля, здесь носители движутся только вследствие диффузии; сопротивление нейтральных областей пренебрежимо мало по сравнению с сопротивлением обедненного слоя; уровень инжекции низкий;

3) границы p - n -перехода плоские, носители движутся только в направлении перпендикулярном этим границам, краевые эффекты не учитываются, то есть p - n -переход плоскопараллелен, бесконечной протяженности;

4) внешние области за границами p - n -перехода электрически нейтральны, заряд инжектированных неосновных носителей быстро нейтрализуется основными носителями, поступающими из внешней цепи. Толщина нейтральных областей много больше диффузионной длины неосновных носителей в этих областях.

Получим вольт-амперную характеристику p - n -перехода. Для этого запишем уравнение непрерывности в общем виде:

$$\frac{dp}{dt} = G - R - \frac{1}{q} \operatorname{div}(j). \quad (2.26)$$

Будем рассматривать стационарный случай $\frac{dp}{dt} = 0$.

Рассмотрим ток в квазинейтральном объеме полупроводника n -типа справа от обедненной области p - n перехода ($x > 0$). Темп генерации G в квазинейтральном объеме равен нулю: $G = 0$. Электрическое поле E тоже равно нулю: $E = 0$. Дрейфовая компонента тока также равна нулю: $I_E = 0$, следовательно, ток диффузионный $j = -qD \frac{dp}{dx}$. Темп рекомбинации R при малом уровне инжекции описывается соотношением

$$R = -\frac{p_n - p_{n0}}{\tau}. \quad (2.27)$$

Воспользуемся следующим соотношением, связывающим коэффициент диффузии, длину диффузии и время жизни неосновных носителей: $D\tau = L_p^2$.

С учетом отмеченных выше допущений уравнение непрерывности имеет вид

$$\frac{d^2 p_n}{dx^2} - \frac{p_n - p_{n0}}{L_p^2} = 0. \quad (2.28)$$

Граничные условия для диффузионного уравнения в p - n -переходе имеют вид:

$$\text{при } x = 0, p_n = p_{n0} e^{\beta V_G}; \text{ при } x \rightarrow \infty, p_n = p_{n0}. \quad (2.29)$$

Решение дифференциального уравнения (2.28) с граничными условиями (2.29) имеет вид:

$$p_n - p_{n0} = p_{n0} (e^{\beta V_G} - 1) e^{-\frac{x}{L_D}}. \quad (2.30)$$

Соотношение (2.30) описывает закон распределения инжектированных дырок в квазинейтральном объеме полупроводника n -типа для электронно-дырочного перехода (рис. 2.11). В токе p - n -перехода принимают участие все носители, пересекающие границу ОПЗ с квазинейтральным объемом p - n -перехода. Поскольку весь ток диффузионный, подставляя (2.30) в выражение для тока, получаем:

$$j_{pD} = -qD_p \left. \frac{dp_n}{dx} \right|_{x=0} = q \frac{D_p p_{n0}}{L_p} e^{\beta V_G}. \quad (2.31)$$

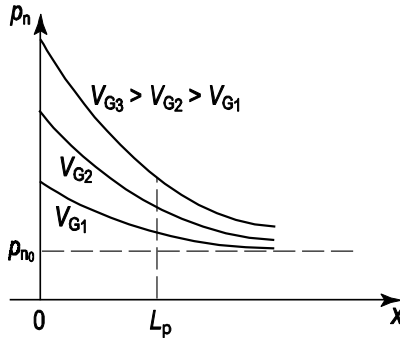


Рис. 2.11. Распределение неравновесных инжектированных из эмиттера носителей по квазинейтральному объему базы p - n -перехода

Соотношение (2.31) описывает диффузионную компоненту дырочного тока p - n -перехода, возникающую при инжекции неосновных носителей при прямом смещении. Для электронной компоненты тока p - n -перехода аналогично получаем:

$$j_{nD} = q \frac{D_n n_{p0}}{L_n} e^{\beta V_G}. \quad (2.31, a)$$

На рис. 2.12 изображены диффузионные токи в несимметричном p - n -переходе при прямом смещении.

При $V_G = 0$ дрейфовые и диффузионные компоненты уравновешивают друг друга. Следовательно,

$$j_{pE} = q \frac{D_p p_{n0}}{L_p}; \quad j_{nE} = q \frac{D_n n_{p0}}{L_n}.$$

Полный ток p - n -перехода является суммой всех четырех компонент тока p - n -перехода:

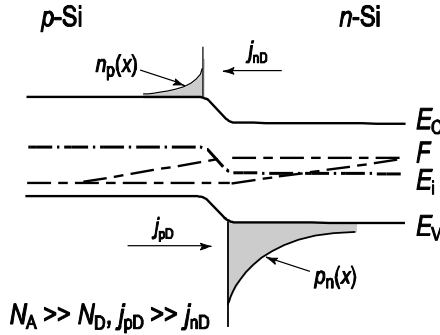


Рис. 2.12. Токи в несимметричном p - n -переходе при прямом смещении

$$j = \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) (e^{\beta U} - 1). \quad (2.32)$$

Выражение в скобках имеет физический смысл обратного тока p - n -перехода. Действительно, при отрицательных напряжениях $V_G < 0$ ток дрейфовый и обусловлен неосновными носителями. Все эти носители уходят из цилиндра длиной L_n со скоростью L_n/τ_p . Тогда для дрейфовой компоненты тока получаем:

$$j_n = \frac{qL_n n_{p0}}{\tau_n} = \frac{qL_n n_{p0}}{L_n^2 / D_n} = \frac{qD_n n_{p0}}{L_n}.$$

Нетрудно видеть, что это соотношение эквивалентно полученному ранее при анализе уравнения непрерывности.

Если требуется реализовать условие односторонней инжекции (например, только инжекции дырок), то из соотноше-

ния (2.32) следует, что нужно выбрать малое значение концентрации неосновных носителей n_{p0} в p -области. Отсюда следует, что полупроводник p -типа должен быть сильно легирован по сравнению с полупроводником n -типа: $N_A \gg N_D$. В этом случае в токе p - n -перехода будет доминировать дырочная компонента (рис. 2.12).

Таким образом, ВАХ p - n -перехода имеет вид:

$$J = J_s (e^{\beta V_G} - 1). \quad (2.33)$$

Плотность тока насыщения J_s равна:

$$J_s = \frac{qD_n n_{p0}}{L_n} + \frac{qD_p p_{n0}}{L_p} = \frac{qL_n n_{p0}}{\tau_n} + \frac{qL_p p_{n0}}{\tau_p}. \quad (2.34)$$

ВАХ p - n -перехода, описываемая соотношением (2.33), приведена на рис. 2.13.

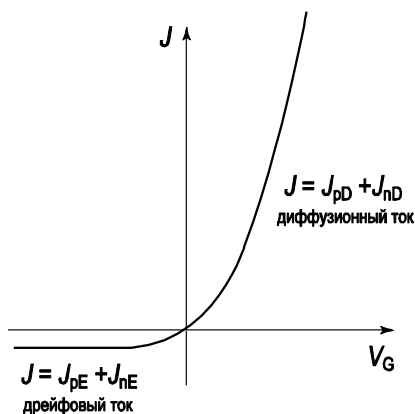


Рис. 2.13. Вольт-амперная характеристика идеального p - n -перехода

Как следует из соотношения (2.33) и рис. 2.13, вольт-амперная характеристика идеального p - n -перехода имеет ярко выраженный несимметричный вид. В области прямых напряжений ток p - n -перехода диффузионный и экспоненциально возрастает с ростом приложенного напряжения. В области отрицательных напряжений ток p - n -перехода – дрейфовый и не зависит от приложенного напряжения.

2.2.8. Вольт-амперная характеристика реального p - n -перехода

При выводе теоретической вольт-амперной характеристики учитывались только явления инжекции и экстракции носителей и их диффузия в нейтральных областях, прилегающих к p - n -переходу.

Реальная (экспериментальная) статическая вольт-амперная характеристика p - n -перехода отличается от теоретической из-за того, что при выводе теоретической вольт-амперной характеристики не учитывался ряд факторов, таких, как

- наличие омического сопротивления эмиттерной и базовой областей кристалла, а также контактов и выводов;
- протекание процессов генерации и рекомбинации носителей в переходе;
- наличие токов утечки по поверхности кристалла;
- возникновение различных видов пробоя перехода при больших обратных напряжениях и др.

Рассмотрим влияние этих факторов на вид реальной вольт-амперной характеристики p - n -перехода.

При прямом смещении отклонение реальной вольт-амперной характеристики от идеальной происходит по следующим причинам:

- а) рекомбинация носителей заряда в области перехода. В области перехода, как и в нейтральных областях полупровод-

ника происходит рекомбинация носителей. Электроны n -области, обладающие достаточной энергией, могут попасть в обедненный слой и рекомбинировать там с дырками, приходящими из p -области. При этом электроны уходят из n -области, а дырки из p -области. Вследствие такого движения носителей возникает дополнительный прямой ток, называемый **током рекомбинации**. Полный прямой ток p - n -перехода складывается из тока инжекции $I = I_s (e^{\frac{U}{\varphi_T}} - 1)$ и тока рекомбинации

Расчеты, подробности которых мы опускаем, показывают, что величина скорости объемной рекомбинации равна

$$V_{рек} \cong \frac{1}{2} \sigma V_t \cdot N_t n_i \cdot \exp\left(\frac{qU_{вн}}{2kT}\right), \quad (2.35)$$

где σ – эффективное сечение рекомбинации; V_t – тепловая скорость носителя заряда; N_t – концентрация примесных центров, на которых происходит рекомбинация; n_i – собственная концентрация носителей заряда в полупроводниковом материале, из которого изготовлен диод; $U_{вн}$ – внешнее смещение, приложенное к диоду.

В этом случае рекомбинационная составляющая тока равна

$$j_{\partial\partial e} = \int_0^w q \cdot V_{\partial\partial e} dx = qW \frac{\sigma}{2} V_t N_t \cdot n_i \exp\left(\frac{qU_{\partial\partial}}{2kT}\right), \quad (2.36)$$

где W – ширина области объемного заряда;

Следовательно, в реальном p - n -переходе прямой ток больше, чем в идеальном. Сравним ток инжекции и ток рекомбинации. Ток инжекции более резко зависит от напряже-

ния $I \approx \exp\left(\frac{U_{\text{ái}}}{\varphi_T}\right)$, в то время как ток рекомбинации

$I_{\text{дái}} \approx \exp\left(\frac{U_{\text{ái}}}{2\varphi_T}\right)$. Из-за наличия множителя $1/2$ в показателе

экспоненты ток рекомбинации с ростом прямого напряжения увеличивается медленнее, чем ток инжекции. Ток инжекции

пропорционален $I \approx n_i^2 \approx \exp\left(-\frac{\Delta E_g}{kT}\right)$, он более резко зависит

от температуры и ширины запрещенной зоны, чем ток реком-

бинации, который пропорционален $I_{\text{дái}} \approx n_i \approx \exp\left(-\frac{\Delta E_g}{2kT}\right)$.

Таким образом, ток рекомбинации является основной составляющей прямого тока при малых напряжениях в полупроводниках с широкой запрещенной зоной, при больших напряжениях основную часть тока составляет ток инжекции.

б) большой уровень тока (высокий уровень инжекции).

Создание большого уровня инжекции ($\Delta p_n \geq n_{n0}$ и $\Delta n_p \geq p_{p0}$) приводит к появлению электрического поля в квазинейтральной области и падению напряжения на сопротивлении этой области (базе), которое становится сравнимым с падением напряжения на сопротивлении области объемного заряда, при этом на вольт-амперной характеристике p - n -перехода появляется почти линейный участок.

Появление электрического поля в базе приводит к тому, что к диффузионному движению носителей в базе добавляется дрейфовая составляющая. В этом случае выражение для плотности тока имеет вид

$$j_p = -q \cdot 2D_p \frac{dp_n}{dx}; \quad (2.37)$$

в) падение напряжения на сопротивлении базовой области и переходном сопротивлении невыпрямляющих контактов приводит к тому, что не все напряжение, приложенное к внешним зажимам, приходится на слой объемного заряда. Можно показать, что, если $U_{вн} = U_{\Pi} + U_{Б}$, где $U_{вн}$ – внешнее смещение; U_{Π} – напряжение, приложенное непосредственно к p - n -переходу; $U_{Б}$ – напряжение на сопротивлении базы и контактах, то

$$U_{\Pi} = \frac{U_{вн}}{2} - \frac{1}{2} \frac{kT}{q} \ln\left(\frac{p_{n0}}{n_{n0}}\right), \quad (2.38)$$

где p_{n0} – концентрация дырок в базовой области в равновесном состоянии; n_{n0} – концентрация электронов в базовой области с учетом электрического поля в области базы и падения напряжения на распределенном сопротивлении и контактах.

Уравнение прямой ветви реальной вольт-амперной характеристики записывается в виде

$$I = S_{p-n} \frac{2D_p}{W_B} n_i \exp\left(\frac{qU_{вн}}{2kT}\right), \quad (2.39)$$

где W_B – ширина базовой области; S_{p-n} – площадь перехода; n_i – собственная концентрация носителей заряда в полупроводнике.

При прямых смещениях, превышающих величину контактной разности потенциалов, электронно-дырочный переход теряет свои нелинейные свойства.

На рис. 2.14 представлена реальная и идеальная вольт-амперные характеристики p - n -перехода при прямом смещении.

Одна из основных причин, влияющих на вид вольт-амперной характеристики при обратном смещении – наличие токов утечки в местах выхода электронно-дырочного перехода на поверхность кристалла.

На поверхности всегда имеются дефекты, которые приводят к возникновению заряда, а, следовательно, к появлению сопротивления утечки, шунтирующего p - n -переход.

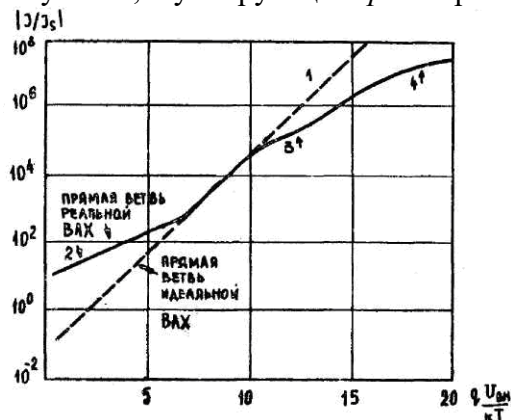


Рис. 2.14. Прямая ветвь вольт-амперной характеристики p - n -перехода: 1 - инжекция; 2 - рекомбинация носителей заряда в области объемного заряда; 3 - большой уровень тока (высокий уровень инжекции); 4 - падение напряжения на сопротивлении базовой области и переходном сопротивлении невыпрямляющих контактов

Поверхностные токи утечки в некоторых случаях во много раз превышают теоретически рассчитанную величину обратного тока. При обратном смещении, поданном на переход, область объемного заряда расширяется, а концентрация подвижных носителей заряда в области объемного заряда много меньше, чем в равновесном состоянии, $n \cdot p \leq n_i^2$. При таких условиях генерация носителей заряда не уравновешивается рекомбинацией, а скорость генерации V_G определяется

выражением $V_{\Gamma} \cong n_i / \tau_{эфф}$, где $\tau_{эфф}$ – эффективное время жизни. С учетом скорости генерации значение тока, обусловленного этим процессом, определяется как

$$j_{\Gamma} = \int_0^W q |V_{\Gamma}| dx \cong q |V_{\Gamma}| W = \frac{qn_i W}{\tau_{эфф}}, \quad (2.40)$$

где W – та же ширина области объемного заряда.

Поскольку ширина области пространственного заряда является функцией внешнего смещения, то и значение обратного тока диода также зависит от приложенного смещения:

$$I_{\Gamma} \sim (\varphi_{\kappa} + U_{вн})^{1/2} \text{ – для резкого перехода;}$$

$$I_{\Gamma} \sim (\varphi_{\kappa} + U_{вн})^{1/3} \text{ – для линейного перехода.}$$

Без учета токов утечки общий обратный ток имеет вид

$$j_{общ} = q \left(\frac{n_i^2}{N_D} \sqrt{\frac{D_p}{\tau_p}} + \frac{n_i^2}{N_A} \sqrt{\frac{D_n}{\tau_n}} \right) + \frac{qn_i W}{\tau_{эфф}}. \quad (2.41)$$

На рис. 2.15 представлена реальная (кривая А) и идеальная (кривая В) вольт-амперные характеристики p - n -перехода при обратном включении.

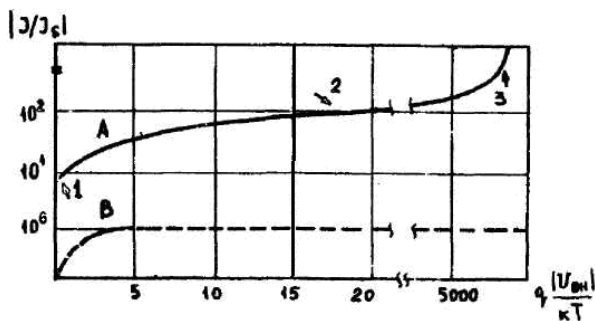


Рис. 2.15. Обратная ветвь вольт-амперной характеристики p - n -перехода: 1, 2 - генерация носителей заряда, токи утечки; 3 - пробой

Температурная зависимость вольт-амперной характеристики определяется экспоненциальным множителем $qU_{обн}/kT$ и током насыщения I_s , плотность которого определяется выражением (2.34).

Выражение (2.34) можно переписать в виде

$$I_s = S_{p-n} \left(L_p \frac{p_{n0}}{\tau_p} + L_n \frac{n_{p0}}{\tau_n} \right), \quad (2.42)$$

тогда становится более ясным физический смысл тока насыщения. Ток насыщения – это ток, вызванный генерацией неосновных носителей заряда со скоростью генерации p_{n0}/τ_p и n_{p0}/τ_n на расстояниях L_p и L_n от границы области объемного заряда.

Температурная зависимость тока насыщения определяется в основном температурной зависимостью собственной концентрации:

$$I_s = S_{p-n} q n_i^2 \left(\frac{D_p}{N_D L_p} + \frac{D_n}{N_A L_n} \right). \quad (2.43)$$

В этом выражении принято, что все атомы полностью ионизированы, то есть $N_D \approx N_D^+ \approx n_{n0}$; $N_A \approx N_A^+ \approx p_{p0}$.

Собственная концентрация носителей заряда является сложной функцией температуры:

$$\begin{aligned} n_i &= (N_C N_V)^{1/2} \exp\left(-\frac{E_C - E_V}{2kT}\right) = \\ &= 2 \left(\frac{2\pi k}{h^2} \right)^{3/2} (m_n^* \cdot m_p^*)^{3/4} T^{3/2} \exp\left(-\frac{\Delta E_g}{2kT}\right), \end{aligned} \quad (2.44)$$

где N_C и N_V – эффективные плотности уровней в зоне проводимости и в валентной зоне; m_n^* и m_p^* – эффективные массы электронов и дырок; h – постоянная Планка; ΔE_g – ширина запрещенной зоны.

Ширина запрещенной зоны полупроводника, из которого изготовлен переход, в свою очередь, также меняется с температурой:

$$\Delta E_g = \Delta E_{g0} - \beta T, \quad (2.45)$$

где E_{g0} – ширина запрещенной зоны при температуре абсолютного нуля; β – температурный коэффициент.

В общем виде температурная зависимость обратного тока диода с n - p^+ -переходом записывается в виде

$$I_s = S_{p-n} q \frac{L_p}{\tau_p} \cdot \frac{n_i^2}{N_D} \cong B \cdot \exp\left(-\frac{\Delta E_g}{kT}\right), \quad (2.46)$$

где B – некоторый коэффициент, который включает в себя все сомножители, по температурной зависимости более слабые, чем экспонента. Соотношение (2.46) позволяет по температурной зависимости тока насыщения оценить ширину запрещенной зоны материала, из которого изготовлен p - n -переход. Для такого определения необходимо построить зависимость I_s от $1/T$ и по тангенсу угла наклона полученной прямой найти ΔE_g . Температурная зависимость обратного тока позволяет оценить параметры материала только для германиевых диодов, в которых обратный ток определяется в основном током насыщения. В кремниевых приборах обратный ток обусловлен генерационной составляющей тока и поэтому

$$I_{обр} \approx I_G \approx S_{p-n} q \frac{n_i}{\tau_{эфф}} \cdot W. \quad (2.47)$$

Отношение тока генерации к току насыщения составляет для p - n -перехода

$$\frac{I_G}{I_s} = \frac{n_i W}{N_D L_p}. \quad (2.48)$$

Поэтому температурная зависимость обратного тока для кремниевого диода определяет не ширину запрещенной зоны, а рекомбинационный уровень энергии E_r . Значение энергии этого уровня может быть значительно меньше $\Delta E_g/2$.

Экспериментальная характеристика p - n -перехода существенно отличается от идеальной, построенной по закону $I = I_s (e^{\frac{U}{\phi_T}} - 1)$. Сравним теоретическую и реальную характеристику p - n -перехода (рис. 2.16).

Ток через реальный прямо включенный переход можно представить в виде ряда составляющих: ток инжекции; ток

генерации носителей в объеме полупроводника; ток рекомбинации носителей; канальный ток и ток утечки.

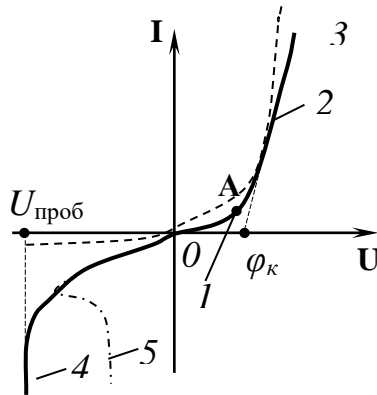


Рис. 2.16. Вольт-амперная характеристика p - n -перехода:

----- - идеальная; ——— - реальная

Вольт-амперную характеристику p - n -перехода при прямом включении можно условно разделить на 3 участка в соответствии с величиной протекающего тока:

1 - область малых токов. На этом участке реальная характеристика расположена ниже идеальной, рост тока замедлен из-за рекомбинации дырок и электронов в реальном p - n -перехода;

2 – область средних токов. На этом участке реальная и идеальная характеристики совпадают;

3 - область больших токов. На этом участке реальная характеристика расположена ниже идеальной, так как существенно увеличивается падение напряжения в объеме полупроводника.

Реальный обратный ток через p - n -переход намного превышает тепловой ток теоретической вольт-амперной характеристики и характеризуется суммой токов:

- тепловой ток;

- ток термогенерации в обратном смещенном переходе;
- ток утечки.

При большом обратном напряжении $U = U_{идoa}$ – напряжение пробоя, величина которого зависит от концентрации примесей в полупроводнике и температуры, наступает **пробой p - n -перехода**. Различают электрический пробой (участок 4) и тепловой пробой (участок 5).

2.2.9. Вольт-амперная характеристика p - n -перехода в полулогарифмических координатах

Построение прямой ветви вольт-амперной характеристики p - n -перехода в полулогарифмических координатах (рис. 2.17) может дать информацию относительно возможных механизмов протекания тока. Там, где вольт-амперная характеристика экспоненциальна, в данной системе координат получается прямолинейный отрезок. 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 $\exp(qU/2kT)$, что дает на графике отрезок прямой линии с угловым коэффициентом 1/2.

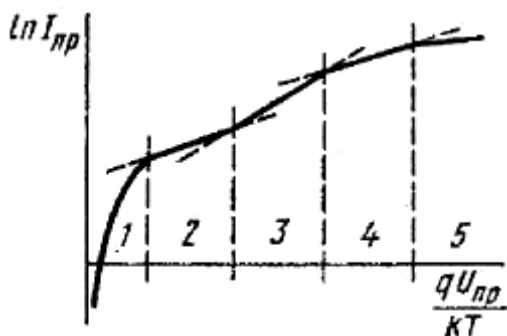


Рис. 2.17. Прямая ветвь вольт-амперной характеристики p - n -перехода: 1 – участок малых токов; 2 – участок с преобладанием рекомбинации в p - n -переходе; 3 – участок с преобладанием процесса инжекции; 4 – участок, связанный с модуляцией сопротивления базы при высоком уровне инжекции; 5 – участок, связанный с нарушением условий на границах p - n -перехода и на невыпрямляющем контакте с базой

При бóльших напряжениях преобладает ток, связанный с инжекцией носителей заряда и зависящий от напряжения как $\exp(qU/kT)$, 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 $\exp(qU/2kT)$ или на графике – отрезок прямой с угловым коэффициентом 1/2.

Наконец, при очень больших прямых токах вольт-амперная характеристика p - n -перехода становится неэкспоненциальной из-за нарушения условий на электронно-дырочном переходе и на невыпрямляющем контакте. Соответствующий отрезок на графике не будет прямой линией.

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 p - n -переходов участок, соответствующий рекомбинационному току, может сразу перейти в участок, соответствующий высокому уровню инжекции. Их границу обнаружить нельзя. Искажает также вид характеристик изменение времени жизни носителей, изменение температуры, наличие каналов поверхностной электропроводности и т.п.

2.2.10. Пробой p - n -перехода

При повышении обратного смещения, приложенного к переходу, можно обнаружить явление, при котором небольшое изменение напряжения приводит к резкому возрастанию тока. Если при этом величина тока не ограничена добавочным сопротивлением, то *p-n*-переход может быть нарушен. Это явление называется пробоем электронно-дырочного перехода. Резкое увеличение дифференциальной проводимости *p-n*-перехода при достижении обратным напряжением критического значения называется **пробоем** *p-n*-перехода. Различают три основных механизма пробоя:

1) лавинный, связанный с лавинным размножением носителей за счет ударной ионизации и за счет вырывания электронов из атомов сильным электрическим полем. Этот пробой характерен для *p-n*-переходов большой толщины, получающихся при сравнительно малой концентрации примесей в полупроводниках. Пробивное напряжение для лавинного пробоя составляет десятки или сотни вольт;

2) туннельный, обусловленный туннельным эффектом. Сущность последнего состоит в том, что при поле напряженностью более 10^5 В/см, действующем в *p-n*-переходе малой толщины, некоторые электроны проникают через переход без изменения своей энергии. Тонкие переходы, в которых возможен туннельный эффект, получают при высокой концентрации примесей. Напряжение, соответствующее туннельному пробую, обычно не превышает единиц вольт;

3) тепловой пробой возникает из-за нарушения устойчивости теплового режима *p-n*-перехода. Это означает, что количество теплоты, выделяющейся в переходе от нагрева его обратным током, превышает количество теплоты, отводимой от перехода. В результате температура перехода возрастает, сопротивление его уменьшается и ток увеличивается, что приводит к перегреву перехода и его тепловому разрушению.

Лавинный и туннельный пробой - два вида электрического пробоя, которые нередко сопутствуют друг другу. Электрический пробой является обратимым, то есть при этом пробое в переходе не происходит необратимых изменений (разрушения структуры вещества). Поэтому работа p - n -перехода в режиме электрического пробоя допустима. Тепловой пробой необратим, так как он сопровождается разрушением структуры вещества в месте p - n -перехода.

Лавинный пробой электронно-дырочного перехода обусловлен эффектом лавинного размножения носителей заряда в электрическом поле обратно-смещенного перехода. Лавинное размножение носителей при обратном напряжении на p -перехода происходит в результате ударной ионизации атомов полупроводника. Носители заряда получают энергию от электрического поля и расходуют ее на генерацию новых пар и взаимодействие с решеткой. При лавинном размножении происходит увеличение токов электронов и дырок в M раз, где M – коэффициент размножения. Условие лавинного пробоя для большинства электронно-дырочных переходов записывается в виде

$$\int_{x=W_n}^{x=W_p} \alpha(\bar{E}, x) dx = 1, \quad (2.49)$$

где W_n и W_p – границы области объемного заряда; $\alpha(\bar{E}, x)$ – коэффициент ударной ионизации, показывающий, какое количество носителей заряда образовалось на единицу пути в единицу времени.

Используя условия лавинного пробоя (2.49) и располагая зависимостью коэффициента ударной ионизации от поля $\alpha(\bar{E}, x)$, можно определить величину критического поля, при котором происходит пробой перехода, а затем и величину пробивного напряжения

$$U_{\text{проб}} = - \int_{W_n}^{W_p} \bar{E}(x) dx ; \quad (2.50)$$

Максимальное значение напряженности электрического поля в запирающем слое резкого перехода связано с обратным напряжением соотношением

$$E_{\text{max}} = \sqrt{2qN^*(U_{\text{id}} + \varphi_e) / \varepsilon_0 \varepsilon} , \quad (2.51)$$

где $N^* = \frac{N_a N_d}{N_a + N_d}$ - приведенная концентрации примеси.

Можно принять, что лавинный пробой происходит, когда максимальная напряженность электрического поля достигает критического значения $E_{\text{кр}}$. Из выражения (2.51) получаем

$$U_{\text{id}} = \frac{\varepsilon \varepsilon_0 E_{\text{кр}}^2}{2qN^*} \quad (2.52)$$

Напряжение, при котором происходит стремительное возрастание тока, называется напряжением лавинного пробоя $U_{\text{проб}}$.

Для резкого несимметричного p - n -перехода приведенная концентрация примеси в p - n -переходе равна концентрации примеси в базе. Следовательно, напряжение лавинного пробоя обратно пропорционально концентрации примеси в базе. Для увеличения напряжения пробоя уменьшают концентрацию легирующей примеси в базе p - n -перехода.

Выражение для определения напряжения пробоя резких несимметричных p - n -переходов имеет вид

$$U_{\text{ид.}} = 60 \cdot \left(\frac{\Delta E_g}{1,1} \right)^{1,5} \cdot \left(\frac{N}{10^{16}} \right)^{-0,75}, \quad (2.53)$$

где N – концентрация примеси в базе, см^{-3} ;

ΔE_g – ширина запрещенной зоны, эВ.

Для p - n -переходов с линейным распределением примеси

$$U_{\text{ид.}} = 60 \cdot \left(\frac{\Delta E_g}{1,1} \right)^{1,2} \cdot \left(\frac{a}{3 \cdot 10^{20}} \right)^{-0,4}, \quad (2.54)$$

где a – градиент концентрации примеси в p - n -переходе, см^{-4} .

Вольт-амперные характеристики p - n -перехода на участке лавинного пробоя имеют следующий вид (рис. 2.18): для лавинного пробоя характерен рост обратного тока при практически неизменном напряжении. Напряжение лавинного пробоя с повышением температуры увеличивается (рис. 2.18). С ростом температуры увеличивается число столкновений носителей с атомами решетки, а средняя длина свободного пробега уменьшается. Носители набирают меньшую скорость, интенсивность ударной ионизации уменьшается, и ток умножения носителей заряда падает. Для приобретения энергии достаточной для ударной ионизации, носители должны перемещаться в более сильном электрическом поле, то есть для поддержания лавинного процесса необходимо увеличить напряженность электрического поля, а значит и приложенное напряжение.

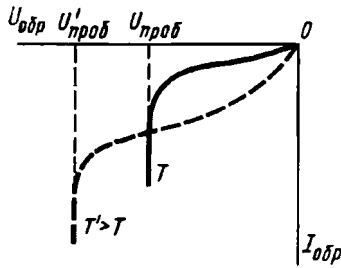


Рис. 2.18. Вольт-амперная характеристика $p-n$ -перехода при лавинном пробое для разных температур

Лавинный пробой реального $p-n$ -перехода имеет свою специфику, которая обусловлена искривлением фронта диффузии, возникающим при создании $p-n$ -перехода. Центральная часть $p-n$ -перехода, полученного методом локальной диффузии, (рис. 2.19) имеет плоскую геометрию (участок 1), боковые части – цилиндрическую (участок 2), а углы – сферическую (участок 3).

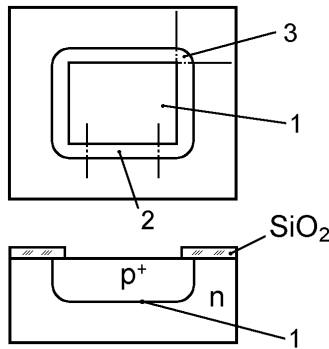


Рис. 2.19. Структура планарного $p-n$ -перехода

Отклонение формы $p-n$ -перехода от плоской части увеличивает напряженность электрического поля в нем при том же приложенном напряжении, что способствует снижению

пробивного напряжения. Поэтому расчет напряжения пробоя проводят по формулам для p - n -перехода цилиндрической и сферической геометрии, считая, что коллекторный p - n -переход резкий:

$$U_{\text{проб.}}^* = U_{\text{проб.пл.}} \{[(m + \gamma_0 + 1)\gamma_0^m]^{1/(m+1)} - \gamma_0\}, \quad (2.55)$$

где $U_{\text{проб. пл.}}$ - напряжение пробоя плоской части p - n -перехода определяемое по формуле (2.53);

$m = 1$ – для цилиндрического перехода, $m = 2$ – для сферического перехода;

$\gamma_0 = r_j / x_{d \text{ макс}}$, где r_j – радиус кривизны p - n -перехода, примерно равный глубине залегания коллекторного перехода $x_{\text{ко}}$; $x_{d \text{ макс}}$ – ширина обедненного слоя при пробое.

Туннельный пробой, или зиннеровский эффект, основан на механизме туннельного прохождения тока через электронно-дырочный переход. Квантомеханическое рассмотрение явления туннелирования через треугольный потенциальный барьер высотой ΔE_T и толщиной W позволяет найти вероятность прохождения электронов и выражение для плотности туннельного тока:

$$j_T = \frac{(2m^*)^{1/2} q^3 \bar{E}}{4\pi^2 \hbar^2 \Delta E_T^{1/2}} \cdot \exp\left[\frac{4(m^*)^{1/2} \Delta E_T^{3/2}}{3q\bar{E}\pi}\right]. \quad (2.56)$$

Туннельный пробой происходит в очень тонких p - n -переходах, что возможно при очень высокой концентрации примесей (порядка 10^{19} см^{-3}), когда ширина перехода становится малой (порядка $0,01 - 0,02 \text{ мкм}$) и при небольших значениях обратного напряжения (несколько вольт), когда возникает большой градиент электрического поля. Высокое значение напряженности электрического поля, воздействуя на атомы кристаллической решетки, повышает энергию валентных

электронов и приводит к их туннельному «просачиванию» сквозь «тонкий» энергетический барьер (рис. 2.20) из валентной зоны p -области в зону проводимости n -области. Причем «просачивание» происходит без изменения энергии носителей заряда.

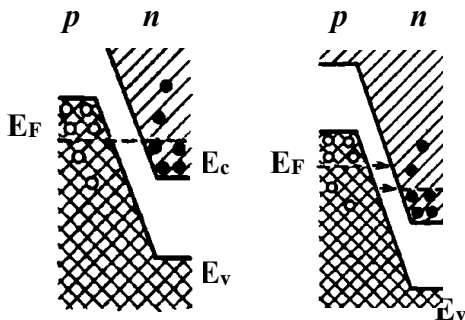


Рис. 2.20. Зонная диаграмма тонкого p - n -перехода:
 a – отсутствие напряжения; b – при обратном напряжении

Для туннельного пробоя также характерен резкий рост обратного тока при практически неизменном обратном напряжении. Возрастание туннельного тока в области пробоя происходит при некоторой критической напряженности электрического поля. Максимальная напряженность поля определяется соотношением

$$E_{\max} = \frac{qN_a N_d}{\varepsilon \varepsilon_0 (N_a + N_d)} W. \quad (2.57)$$

С учетом того, что толщина резкого p - n -перехода $W = \sqrt{\frac{2\varepsilon_s \varepsilon_0 (\varphi_0 + U_{i\ddot{a}\ddot{a}})}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}$, и $E_{\max} = E_{\text{кр}}$ напряжение туннельного пробоя

$$U_{\text{проб}} = \frac{\varepsilon_s \varepsilon_0 \hbar^2}{2q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right) - \varphi_0. \quad (2.58)$$

Следовательно, при туннельном пробое пробивное напряжение оказывается обратно пропорциональным концентрации примеси. Туннельный пробой может происходить только в *p-n*-переходах, изготовленных в полупроводниках с большой концентрацией примеси, так как для туннелирования необходимы малая толщина потенциального барьера и, следовательно, малая ширина перехода. При малой ширине перехода пробивные напряжения его оказываются также малыми. Поэтому контактной разностью потенциалов в выражении (2.58) пренебрегать не следует, так как она может быть сравнима с пробивным напряжением.

Величина напряжения пробоя зависит от температуры перехода (рис. 2.21), причем для туннельного механизма пробоя рост температуры уменьшает напряжение пробоя в отличие от лавинного. 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.

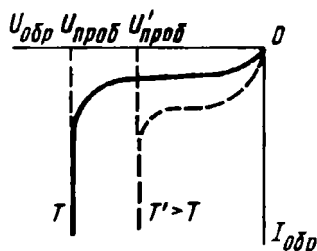


Рис. 2.21. Вольт-амперная характеристика p - n -перехода при туннельном пробое для разных температур

Тепловой пробой - это пробой, при котором увеличение тока происходит за счет разогрева перехода. Тепловой пробой имеет место в p - n -переходах с плохим теплоотводом, поскольку при этом в электронно-дырочном переходе выделяется больше тепла, чем отводится. Мощность рассеяния в переходе $P_{рас} = J_{обр} \cdot U_{обр}$, где $J_{обр}$ и $U_{обр}$ - соответственно обратный ток и напряжение. Отводимая от перехода мощность $P_{омс} = (T_{п} - T_{к})/R_T$, где $T_{п}$ - температура перехода; $T_{к}$ - температура корпуса; R_T - тепловое сопротивление на участке p - n -переход - корпус. Условие теплового пробоя: $\frac{dP_{рас}}{dT} > \frac{dP_{омс}}{dT}$.

Приближенные формулы для расчета предельных режимов работы полупроводникового прибора:

$$0,077U_{обр}J_{обр} > \frac{1}{R_T}; \quad (2.59)$$

$$J_{обр} < \frac{13}{U_{обр}R_T}; \quad (2.60)$$

Вольт-амперная характеристика p - n -перехода в области теплового пробоя имеет вид, приведенный на рис. 2.22.

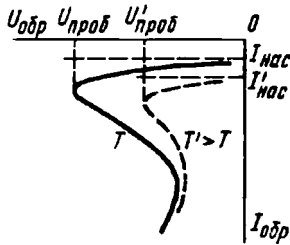


Рис. 2.22. Вольт-амперная характеристика p - n -перехода при тепловом пробое для разных температур

С ростом тока напряжение на p - n -переходе сначала возрастает, а затем начинает уменьшаться. На кривой может быть два экстремума максимум и минимум. Между ними находится участок с отрицательным дифференциальным сопротивлением, то есть с уменьшением напряжения ток увеличивается. Этот участок объясняется тем, что с увеличением выделяемой мощности увеличивается температура, а, следовательно, увеличивается ток и выделяемая мощность. С увеличением температуры напряжение пробоя уменьшается, так как увеличивается обратный ток насыщения.

Поверхностным пробоем p - n -перехода называют пробой перехода, который происходит в месте выхода перехода на поверхность кристалла и на пробивное напряжение которого оказывают влияния поверхностные состояния.

Если поверхностный заряд имеет знак, противоположный знаку основных носителей в базе, то на ее поверхности образуется обогащенный слой. Из-за возникновения обогащенного слоя толщина p - n -перехода у поверхности базы уменьшается. В связи с меньшей толщиной p - n -перехода у поверхности пробой его будет происходить именно там, а пробивное напряжение окажется тем меньше, чем больше суже-

ние перехода у поверхности. По своей природе поверхностный пробой может быть лавинным, туннельным или тепловым.

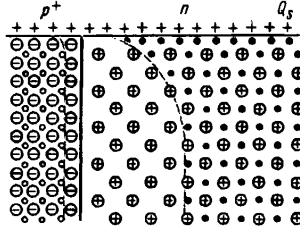


Рис. 2.23. Искажение границ p - n -перехода под влиянием поверхностного заряда

2.2.11. Емкость p - n -перехода

Любая система, в которой при изменении потенциала φ меняется электрический заряд Q , обладает емкостью. Величина емкости C определяется соотношением $C = \left| \frac{\partial Q}{\partial \varphi} \right|$.

Для p - n -перехода можно выделить два типа зарядов: заряд в области пространственного заряда ионизованных доноров и акцепторов Q_B и заряд инжектированных носителей в базу из эмиттера Q_p . При различных смещениях на p -переходе при расчете емкости будет доминировать тот или иной заряд. В связи с этим для емкости p - n -перехода выделяют барьерную емкость C_B и диффузионную емкость C_D .

Барьерная емкость C_B – это емкость p - n -перехода при обратном смещении $V_G < 0$, обусловленная изменением заряда ионизованных доноров в области пространственного заряда.

$$C_B = \frac{\partial Q_B}{\partial V_G}. \quad (2.61)$$

Величина заряда ионизованных доноров и акцепторов Q_B на единицу площади для несимметричного p - n -перехода равна

$$\begin{aligned} Q_B &= qN_D W = qN_D \sqrt{\frac{2\varepsilon_s \varepsilon_0 (\Delta\varphi_0 - V_G)}{qN_D}} = \\ &= \sqrt{qN_D \varepsilon_s \varepsilon_0 (\Delta\varphi_0 - V_G)}. \end{aligned} \quad (2.62)$$

Дифференцируя выражение (2.62), получаем

$$C_B = \frac{\sqrt{2qN_D \varepsilon_s \varepsilon_0}}{2\sqrt{\Delta\varphi_0 - V_G}} = \frac{\varepsilon_s \varepsilon_0}{W}. \quad (2.63)$$

Из уравнения (2.63) следует, что барьерная емкость C_B представляет собой емкость плоского конденсатора, расстояние между обкладками которого равно ширине области пространственного заряда W . Поскольку ширина ОПЗ зависит от приложенного напряжения V_G , то и барьерная емкость также зависит от приложенного напряжения. Численные оценки величины барьерной емкости показывают, что ее значение составляет десятки или сотни пикофарад.

Барьерную емкость резкого p - n -перехода можно записать в виде

$$C_B = S \sqrt{\frac{eqN_A N_D}{2 \cdot (U_K + |U_{обп}|) \cdot (N_A + N_D)}}. \quad (2.64)$$

В общем случае зависимость зарядной емкости от приложенного к p - n -переходу обратного напряжения выражается формулой

$$C_B = \frac{C_0}{\left(1 + \frac{|U_{обр}|}{U_\kappa}\right)^\gamma}. \quad (2.65)$$

где C_0 – емкость p - n -перехода при $U_{обр} = 0$; γ – коэффициент, зависящий от типа p - n -перехода (для резких p - n -переходов $\gamma = 1/2$, а для плавных $\gamma = 1/3$).

Зарядная емкость увеличивается с ростом N_A и N_D , а также с уменьшением обратного напряжения.

Диффузионная емкость C_D – это емкость p - n -перехода при прямом смещении $V_G > 0$, обусловленная изменением заряда инжектированных носителей в базу из эмиттера Q_p .

$$C_D = \frac{\partial Q_p}{\partial V_G}, \quad (2.66)$$

Заряд инжектированных носителей, например дырок в n -области, определяется как

$$Q_p = qS \int_{\delta_n}^{\infty} [p_n(x) - p_{n0}] dx = qL_p S p_{n0} \left[\exp\left(\frac{U_{np}}{\varphi_T}\right) - 1 \right]. \quad (2.67)$$

Тогда диффузионная емкость, обусловленная изменением общего заряда неравновесных дырок в n -области, определится по формуле

$$C_{Dp} = \frac{dQ_p}{dU_{np}} = \frac{qSL_p p_{n0}}{\varphi_T} \exp\left(\frac{U_{np}}{\varphi_T}\right). \quad (2.68)$$

Аналогично диффузионная емкости, обусловленная инжекцией электронов в p -область,

$$C_{Dn} = \frac{dQ_n}{dU_{np}} = \frac{qSL_n p_{p0}}{\varphi_T} \exp\left(\frac{U_{np}}{\varphi_T}\right). \quad (2.69)$$

Общая диффузионная емкость

$$C_D = C_{Dp} + C_{Dn} = \frac{qS}{\varphi_T} (L_n n_{p0} + L_p p_{n0}) \exp\left(\frac{U_{np}}{\varphi_T}\right). \quad (2.70)$$

Полная емкость p - n -перехода определяется суммой зарядной и диффузионной емкостей:

$$C_{p-n} = C_B + C_D. \quad (2.71)$$

Зависимость емкости p - n -перехода от напряжения называется вольт-фарадной характеристикой (рис. 2.24). При отрицательных напряжениях и в окрестности нуля преобладает барьерная емкость, при больших положительных – диффузионная.

Зависимость барьерной емкости C_B от приложенного обратного напряжения V_G используется для приборной реализации. Полупроводниковый диод, реализующий эту зависимость, называется *варикапом*.

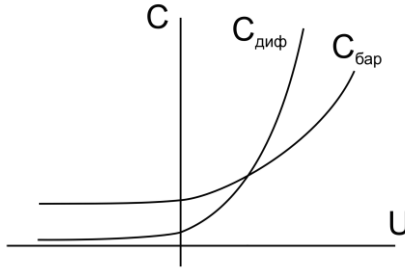


Рис. 2.24. Зависимость емкостей p - n -перехода от напряжения.

2.2.12. Переходные процессы в p - n -переходах

Переходные процессы в p - n -переходах связаны в основном с двумя явлениями, происходящими при быстром изменении напряжения на p - n -переходе или тока через p - n -переход.

Первое из них - это накопление неосновных носителей заряда в базе p - n -перехода при его прямом включении и их рассасывание при уменьшении напряжения. Так как электрическое поле в базе p - n -перехода 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 p - n -перехода в режиме переключения.

Второе явление, происходящее в p - n -переходах при их переключении, - это перезарядка барьерной емкости, которая также происходит не мгновенно и поэтому может влиять на свойства p - n -переходов.

При сравнительно больших плотностях прямого тока через p - n -переход (высокий уровень инжекции) существенно накопление неосновных носителей в базе, а перезарядка барьерной емкости является второстепенным процессом. При малых плотностях тока (низкий уровень инжекции) переходные

процессы в p - n -переходе определяются перезарядкой барьерной емкости, накопление же неосновных носителей заряда в базе практически не сказывается.

Временные зависимости напряжения и тока, характеризующие переходные процессы в p - n -переходе, зависят также от сопротивления внешней цепи, в которую включен p - n -переход. Поэтому рассмотрим четыре предельных случая переходных процессов в несимметричном p - n -переходе.

Переходные процессы при высоком уровне инжекции в режиме генератора напряжения. Рассмотрим процессы, происходящие в p - n -переходе при его включении в цепь с малым полным сопротивлением (в том числе и с малым сопротивлением источника питания) по сравнению с сопротивлением p - n -перехода.

При подаче на p - n -переход прямого напряжения ток через него 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. 2.25). Этот процесс ***модуляции сопротивления базы*** происходит не мгновенно, так как накопление дырок в базе связано с относительно медленным процессом диффузии их от p - n -перехода в глубь базы.

По мере накопления дырок и уменьшения сопротивления базы происходит перераспределение всего внешнего напряжения между сопротивлением базы и ОПЗ p - n -перехода; падение напряжения на базе уменьшается (рис. 2.25, б), а на ОПЗ увеличивается (рис. 2.25, в), вызывая увеличение уровня инжекции. 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.

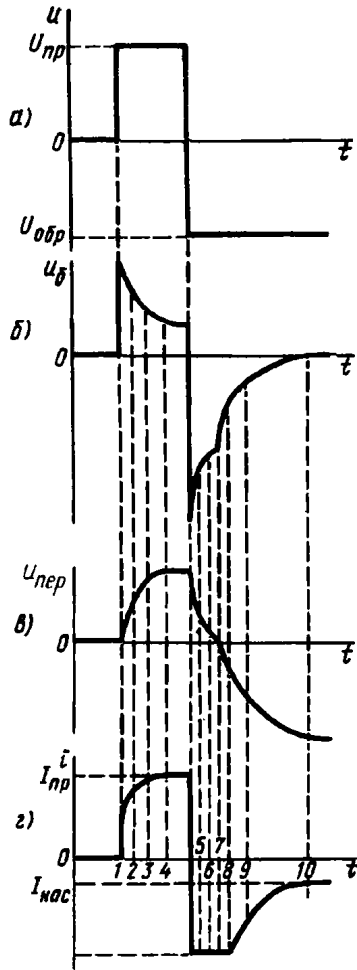


Рис. 2.25. Зависимость от времени напряжения на $p-n$ -переходе (а), напряжения на базе (б), напряжение на ОПЗ $p-n$ -перехода (в) и тока через $p-n$ -перехода (г) при работе на больших импульсах напряжения и тока в режиме генератора напряжения

О значении тока через p - n -переход можно судить по градиенту концентрации дырок в базе около p - n -перехода. Градиент концентрации дырок около p - n -перехода возрастает с увеличением напряжения на p - n -переходе при постоянном напряжении и при уменьшении напряжения на базе. Следует отметить, что сопротивление базы 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 p - n -переход и чем больше время жизни дырок. Кроме того, число накопленных дырок зависит от геометрии базы.

При переключении p - n -перехода с прямого напряжения на обратное в начальный момент наблюдается большой обратный ток, ограниченный в основном последовательным сопротивлением базы. Источник питания вместе с сопротивлением базы в это время можно считать генератором тока для p - n -перехода.

После переключения p - n -перехода 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 ОПЗ p - n -перехода сохраняется прямое падение напряжения (рис. 2.25, в).

После уменьшения концентрации дырок в базе около p - n -перехода до нуля обратный ток начинает уменьшаться со временем, о чем свидетельствует уменьшение градиента концентрации дырок около p - n -перехода. С течением времени все накопленные в базе дырки уходят через p - n -переход или

рекомбинируют в базе, в результате чего обратный ток уменьшается до стационарного значения тока насыщения (рис. 2.25, *з*). В это время заканчивается восстановление обратного сопротивления p - n -перехода.

Процесс рассасывания накопленных носителей происходит значительно медленнее процесса их накопления, поэтому именно процесс рассасывания и определяет частотные свойства большинства p - n -переходов.

Переходные процессы при высоком уровне инжекции в режиме генератора тока. Этот режим соответствует включению p - n -перехода в схему с большим сопротивлением, которое и определяет значение тока в цепи с p - n -переходом. Рассмотрим процессы, происходящие в p - n -переходе, при прохождении через него импульса прямого тока прямоугольной формы (рис. 2.26).

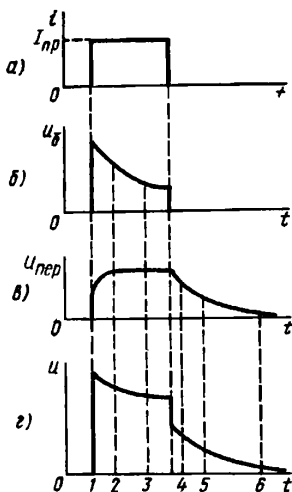


Рис. 2.26. Зависимость от времени тока на p - n -переходе (*а*), напряжения на базе (*б*), напряжения на ОПЗ перехода (*в*) и напряжения на p - n -переходе (*г*) при работе на больших импульсах тока в режиме генератора тока

В первый момент прохождения через p - n -переход импульса прямого тока на p - n -переходе падает относительно большое напряжение, которое в дальнейшем уменьшается из-за уменьшения сопротивления базовой области, связанного с накоплением неравновесных носителей в базе.

После окончания процесса накопления неосновных носителей в базе напряжения на ОПЗ, на базе и на p - n -переходе достигают установившихся значений. Распределение дырок в базе в это время также соответствует некоторому установившемуся состоянию.

В момент окончания импульса тока через p - n -переход, то есть в момент разрыва цепи исчезает падение напряжения на объемном сопротивлении базы диода (рис. 2.26, б). Концентрация дырок в базе около p - n -перехода мгновенно измениться не может. Поэтому напряжение на ОПЗ и соответственно на p - n -переходе после выключения тока уменьшается замедленно по мере рекомбинации неравновесных носителей в базе. Остаточное напряжение на p - n -переходе уменьшится до нуля после рекомбинации всех неравновесных носителей заряда в базовой области.

Переходные процессы при низком уровне инжекции в режиме генератора напряжения. При приложении к p - n -переходу малого прямого напряжения (рис. 2.27) эффект модуляции сопротивления базы из-за малого уровня инжекции пренебрежимо мал. Поэтому сопротивление диода в данном случае имеет емкостный характер. В первый момент напряжение на p - n -переходе близко к нулю, а ток через p - n -переход ограничен только сопротивлением базы (рис. 2.27, в). По мере заряда барьерной емкости напряжение на p - n -переходе и ток через p - n -переход стремятся к некоторым установившимся значениям, которые определяются активной составляющей сопротивления p - n -перехода. В момент переключения p - n -перехода напряжение на барьерной емкости не может измениться мгновенно, оно достигает установившегося значения через некоторое время. Ток через p - n -переход также зависит от времени, что характерно для емкостного сопротивления.

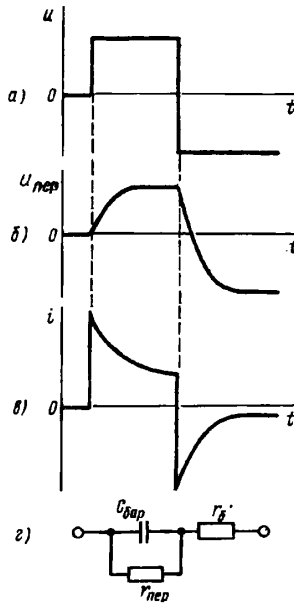


Рис. 2.27. Зависимость от времени напряжения на p - n -переходе (а), напряжения на ОПЗ перехода (б) и тока через p - n -переход (в) при малых импульсах напряжения в схеме с генератором напряжения, а также эквивалентная схема диода для малых сигналов (г)

Переходные процессы при низком уровне инжекции в режиме генератора тока. Осциллограммы импульса прямого тока малой амплитуды и падения напряжения на p - n -переходе при питании от генератора тока представлены на рис. 2.28. В первый момент пропускания импульса тока через p - n -переход весь ток состоит из емкостной составляющей. Поэтому напряжение на p - n -переходе в первый момент определяется падением напряжения на сопротивлении базы. По мере заряда барьерной емкости увеличивается и напряжение на p - n -переходе. При выключении p - n -перехода 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 уменьшается напряжение на емкости и остаточное напряжение на p - n -переходе.

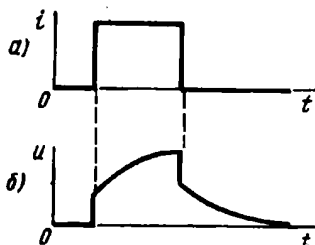


Рис. 2.28. Зависимость тока через p - n -переход (а) и напряжения на p - n -переходе (б) при работе на малых импульсах тока в схеме с генератором тока

2.2.13. Частотные свойства p - n -перехода

Наиболее часто диоды с p - n -переходом используют для выпрямления, детектирования, модуляции напряжения синусоидальной или почти синусоидальной формы.

Рассмотрим поведение p - n -перехода при воздействии на него синусоидального тока или напряжения различной частоты. Подадим на вход цепи, состоящей из последовательного соединения с p - n -переходом резистора R (рис. 2.29), напряжение $E(t) = U_{max} \sin \omega t$, где U_{max} - амплитуда напряжения; $\omega = 2\pi f$ - круговая частота; $f = 1/T$ — циклическая частота; T — период колебаний; t - время. Переходные процессы в p^+ - n -переходе протекают в течение времени порядка времени жизни дырок τ_p в n -базе. На низкой частоте, когда $\omega\tau_p \ll 1$ (или $T \gg \tau_p$), для каждого момента времени изменения синусоидального напряжения переходные процессы, связанные с рекомбинаци-

ей, успевают установиться. Такой режим называют **квазистационарным**. Форма тока через p - n -переход, как показано на рис. 2.30, *a*, оказывается несинусоидальной. Ток диода протекает практически только в первом полупериоде. Значение выпрямленного тока равно среднему значению тока за период T и составляет $I_{\text{вп.ср}} \approx U_{\text{max}}/(\pi R)$. На повышенной частоте, когда $\omega\tau_p \approx 1$ (или период $T \approx \tau_p$), выпрямляющие свойства p - n -перехода ухудшаются. При таких частотах дырки, инжектированные в n -базу за положительный полупериод, не успевают полностью рекомбинировать в базе. Во время отрицательного полупериода часть накопленных в n -базе дырок рекомбинирует, а часть уходит в p -область, создавая всплеск обратного тока (рис. 2.30, *б*), наибольшее мгновенное значение которого $I_{\text{обр,и}}$.

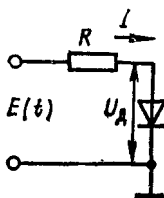


Рис. 2.29. Схема включения диода

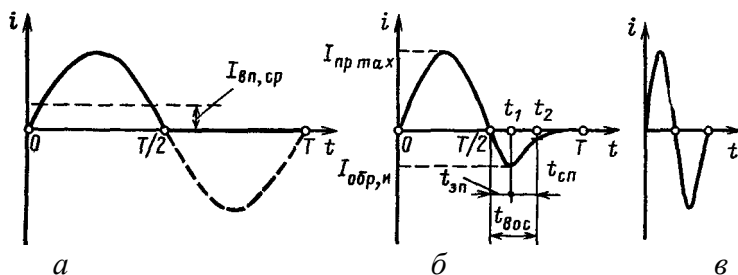


Рис. 2.30. Кривые тока через диод на низких (*a*), средних (*б*) и высоких (*в*) частотах

Интервал времени между моментом $T/2$ и моментом t_1 называется временем запаздывания обратного напряжения $t_{зп}$, так как в течение этого интервала полярность напряжения диода соответствует прямому направлению. Значение $t_{зп} \approx 0,5 \tau_p$. Далее обратный ток спадает, стремясь к своему стационарному значению. Условное окончание фазы спада устанавливают по заданному уровню обратного тока, например $0,25I_{обр.и}$. Время спада обратного тока $t_{сп} = t_1 - t_2 \approx (0,3 - 0,4) \tau_p$. Время обратного восстановления $t_{вос} = t_2 - T/2 = t_{зп} + t_{сп} \approx 0,5 \tau_p$.

На очень высоких частотах $\omega \tau_p \gg 1$ заряд дырок, введенных в n -базу за положительный полупериод, полностью выводится во внешнюю цепь за отрицательный полупериод и диод теряет выпрямительные свойства (рис. 2.30, в).

Частотная характеристика выпрямленного тока имеет спадающий с ростом частоты вид, как показано на рис. 2.31. Частотная зависимость выпрямленного тока имеет спадающий с ростом частоты вид.

Перезаряд барьерной емкости, затягивающий в основном фазу спада тока диода, зависит от значения R , поэтому с увеличением R частотный диапазон выпрямленных токов сужается.

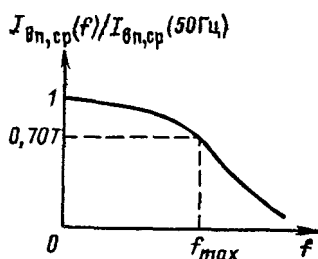


Рис. 2.31. Зависимость относительного значения выпрямленного тока от частоты переменного напряжения

2.2.14. Эквивалентные схемы p - n -перехода

Для расчета электронных схем, содержащих диоды с p - n -переходом используют эквивалентные схемы замещения. Полное математическое описание физических процессов в p - n -переходе достаточно сложно, поэтому p - n -переход заменяют упрощенной моделью, описываемой системой уравнений, или некоторой эквивалентной схемой, содержащей сопротивления, емкости, индуктивности, источники тока или напряжения. К эквивалентным схемам применимы известные методы расчета электрических цепей.

Следует различать два вида эквивалентных схем: для малых (расчет режимов на высоких частотах) и для больших (импульсный режим) сигналов.

Эквивалентные схемы для малых сигналов используются когда переменная составляющая сигнала значительно меньше постоянных токов и напряжений, поэтому переход в первом приближении можно считать линейным элементом

Эквивалентная схема при работе p - n -перехода на малом сигнале приведена на рис. 2.32, *а*. Сопротивление R_0 в этой схеме представляет собой суммарное сравнительно небольшое сопротивление n - и p -областей и контактов этих областей с выводами. Нелинейное сопротивление $R_{нл}$ при прямом напряжении равно $R_{пр}$, т.е. невелико, а обратном напряжении $R_{нл} = R_{обр}$, т.е. оно очень большое. Приведенная эквивалентная схема в различных частных случаях может быть упрощена. На низких частотах емкостное сопротивление очень велико и можно емкость не учитывать. Тогда при прямом напряжении в эквивалентной схеме остаются лишь сопротивления R_0 и $R_{пр}$ (рис. 2.32, *б*), а при обратном напряжении - только сопротивление $R_{обр}$, так как $R_0 \ll R_{обр}$ (рис. 2.32, *в*). На высоких частотах емкости имеют сравнительно небольшое сопротивление. Поэтому при прямом напряжении получается схема по рис.

2.32, *з* (если частота не очень высокая, то $C_{\text{диф}}$ практически не влияет), а при обратном остаются $R_{\text{обр}}$ и C_{δ} (рис. 2.32, *д*).

Следует иметь в виду, что существует еще емкость $C_{\text{в}}$ между выводами диода, которая может заметно шунтировать диод на очень высоких частотах. Она показана на рисунке штрихами. На СВЧ может также проявляться индуктивность выводов.

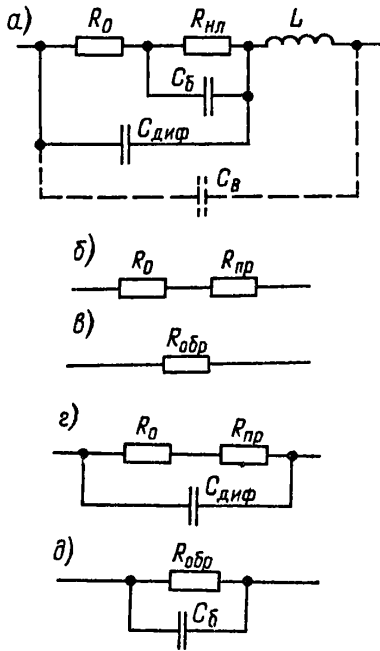


Рис. 2.32. Эквивалентная схемы *p-n*-перехода для малого сигнала

Эквивалентная схема для большого сигнала используется в тех случаях, когда постоянный сигнал по величине сравним с переменной составляющей. Эта схема содержит следующие элементы. Сопротивление $r_{\text{б}}$ представляет собой омическое сопротивление базы (высокоомной области структуры

диода), сопротивление $r_{ут}$ - сопротивление утечки реального диода. Сопротивление базы при большом сигнале - нелинейный элемент. Диод - нелинейный элемент, отражает накопление и рассасывание носителей заряда в базе $p-n$ -перехода. Динамические свойства диода отражаются барьерной $C_{бар}$ и диффузионной $C_{диф}$ емкостями. Барьерная емкость зависит от напряжения на $p-n$ -переходе. Обычно используют значение барьерной емкости для заданного диапазона изменения напряжения. Диффузионная емкость, отражающая накопление неосновных неравновесных носителей заряда, зависит от величины прямого тока. К этой схеме можно также добавить емкость $C_{в}$ между выводами диода, которая может заметно шунтировать диод на очень высоких частотах и индуктивность выводов, которые определяются также как и малосигнальной схеме.

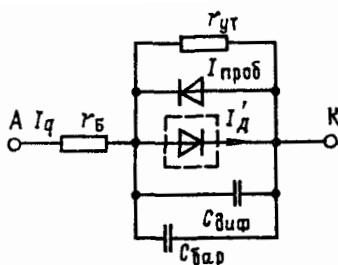


Рис. 2.33. Эквивалентная схемы $p-n$ -перехода для большого сигнала

2.2.15. Влияние температуры на свойства $p-n$ -перехода

Влияние температуры на обратный ток диода. С ростом температуры увеличивается скорость тепловой генерации электронно-дырочных пар во всех областях $p-n$ -перехода. Это приводит к резкому (по экспоненциальному закону) возрастанию с температурой концентрации неосновных носите-

лей в n - и p -областях перехода и, следовательно, к увеличению тока насыщения. Концентрация неосновных носителей заряда в полупроводнике зависит от температуры по закону:

$$p_{n0} = \frac{n_i^2}{n_{n0}} \approx e^{-\frac{E_g}{kT}}, n_{p0} = \frac{n_i^2}{p_{p0}} \approx e^{-\frac{E_g}{kT}}. \quad (2.72)$$

При выводе этих соотношений использовался закон действующих масс и то обстоятельство, что n_{n0} и p_{p0} не меняются с изменением температуры во всем интервале истощения примеси.

Время жизни неосновных носителей заряда и их диффузионные длины с изменением температуры меняются значительно слабее и их изменением можно пренебречь и считать, что ток насыщения

$$J_S \approx n_i^2 \approx e^{-\frac{E_g}{kT}}. \quad (2.73)$$

Другая компонента обратного тока – ток термогенерации в слое объемного заряда – также будет расти вследствие увеличения скорости тепловой генерации электронно-дырочных пар G_0

$$J_{\text{ген}} \Big|_{V=\text{const}} \approx G_0 \approx n_i^2 \approx e^{-\frac{E_g}{2kT}}. \quad (2.74)$$

Третья компонента обратного тока – ток утечки – также растет с ростом температуры, но значительно слабее. Ввиду того, что этот ток может быть обусловлен различными механизмами, из которых не все до конца понятны, то общий анализ зависимости $J_{\text{ут.}}(T)$ затруднен.

В общем случае зависимость обратного тока от температуры приведена на рис. 2.34. Очевидно, что при низких температурах будет преобладать ток утечки. Поскольку $J_{\text{ген}}$ в O_3

растет с температурой быстрее тока утечки, в каком то интервале температур он станет больше. Ток насыщения растет с температурой еще быстрее, так что он может стать основным при более высоких температурах.

Поскольку J_S и $J_{ген}$ в ОЗ в общем случае пропорциональны $\exp\left(-\frac{E_g}{mkT}\right)$, где m – коэффициент, соответствующий определенной компоненте обратного тока. Для диодов, у которых эти компоненты преобладают над токами утечки, снимая зависимость $J_{обр.}(T)$ при постоянном обратном смещении, можно по величине m оценить механизм протекания тока через диод в том или ином температурном интервале. Для этого необходимо построить в полулогарифмическом масштабе график зависимости $\ln J_{i\ddot{a}\delta} = f\left(\frac{1}{T}\right)$ и по наклону выбранного прямолинейного участка определить коэффициент m :

$$m = \frac{E_g}{k} \cdot \frac{1/T_1 - 1/T_2}{\ln J_{i\ddot{a}\delta}(\dot{O}_2) - \ln J_{i\ddot{a}\delta}(\dot{O}_1)}. \quad (2.75)$$

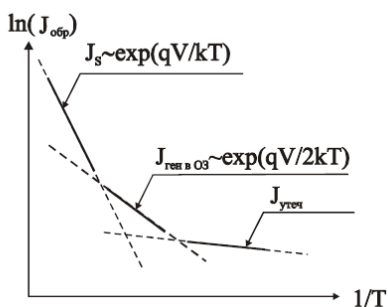


Рис. 2.34. Температурная зависимость обратного тока через диод

Влияние температуры на прямую ветвь ВАХ диода.

Прямой ток через диод создается диффузионными потоками основных носителей, преодолевающих энергетический барьер. С ростом температуры равновесная высота потенциального барьера уменьшается, поскольку уровень Ферми как в n -так и в p -области с повышением температуры стремится к середине запрещенной зоны полупроводника. С понижением потенциального барьера увеличиваются диффузионные потоки основных носителей, то есть увеличивается прямой ток перехода. Иначе говоря, при большей температуре p - n -перехода тот же прямой ток достигается при меньшем смещении. Следовательно, прямая ветвь ВАХ p - n -перехода, без учета сопротивления базы, с ростом температуры смещается влево, в сторону меньших напряжений (рис. 2.35, *а*). ВАХ базы (рис. 2.35, *б*) наоборот, сдвигается вправо, так как сопротивление базы с температурой растет из-за снижения подвижности свободных носителей заряда. Таким образом, результирующая вольт-амперная характеристика диода при прямом смещении сложным образом зависит от температуры. При малых токах характеристика смещается с ростом температуры влево, а при больших токах (где влияние базы сильнее) – вправо (рис. 2.35, *в*).

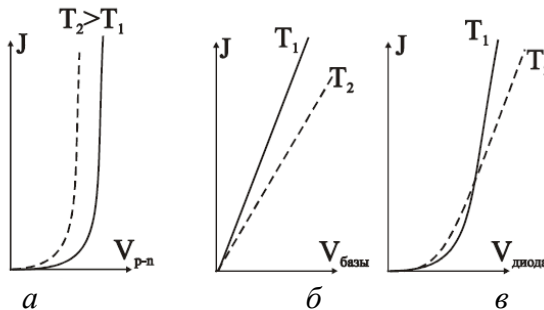


Рис. 2.35. Влияние температуры на прямую ветвь вольт-амперной характеристика p - n -перехода (*а*); базы (*б*); диода (*в*)

2.3. Разновидности электрических переходов

2.3.1. Гетеропереходы

Гетеропереходом называют контакт двух полупроводников различного вида и разного типа проводимости, например $p\text{Ge} - n\text{GaAs}$. Отличие гетеропереходов от обычного p - n -перехода заключается в том, что в обычных p - n -переходах используется один и тот же вид полупроводника, например $p\text{Si} - n\text{Si}$. Поскольку в гетеропереходах используются разные материалы, необходимо, чтобы у этих материалов с высокой точностью совпадали два параметра: температурный коэффициент расширения (ТКР) и постоянная решетки.

С учетом сказанного количество материалов для гетеропереходов ограничено. Наиболее распространенными из них являются германий Ge, арсенид галлия GaAs, фосфид индия InP, четырехкомпонентный раствор InGaAsP.

В зависимости от ширины запрещенной зоны E_g , электронного сродства χ и типа легирования узкозонной и широкозонной областей гетероперехода возможны различные комбинации E_g и χ . На рис. 2.36 показаны эти комбинации при условии равенства термодинамических работ выхода.

Для построения зонных диаграмм, детального анализа распределения электрического поля и потенциала в области пространственного заряда гетероперехода, а также величины и компонент электрического тока для гетеропереходов необходимо учитывать, что у различных полупроводников будут отличаться значения электронного сродства χ , ширины запрещенной зоны E_g и диэлектрической проницаемости ϵ_s .

С учетом этих факторов построим зонную диаграмму гетероперехода германий – арсенид галлия ($p\text{Ge} - n\text{GaAs}$). Значения параметров полупроводниковых материалов, выбран-

ных для расчета зонной диаграммы, приведены в таблице. Приведем в контакт германий p -Ge и арсенид галлия n -GaAs.

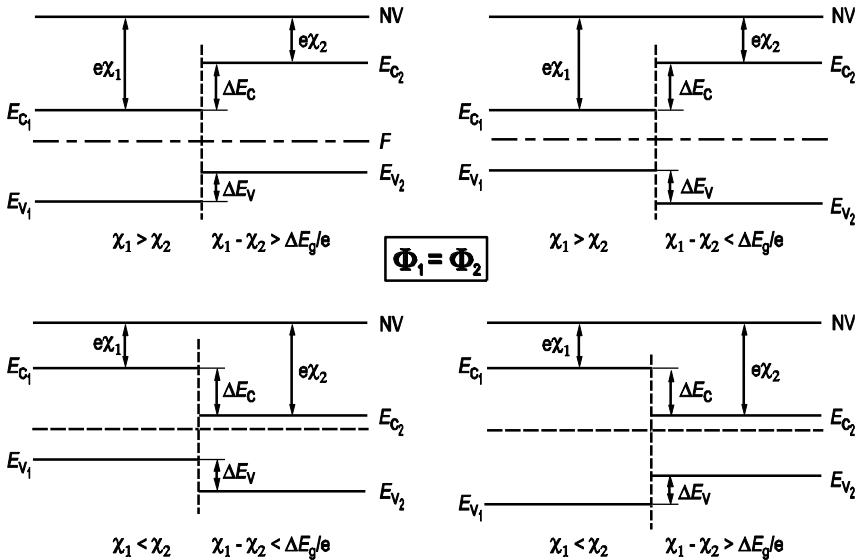


Рис. 2.36. Зонные диаграммы гетеропереходов при различных комбинациях E_g и χ в случае равенства термодинамических работ выхода $\Phi_1 = \Phi_2$

При построении зонной диаграммы гетероперехода учтем следующие факторы:

- уровень вакуума $E = 0$ непрерывен;
- электронное сродство в пределах одного сорта полупроводника χ_{Ge} и χ_{GaAs} постоянно;
- ширина запрещенной зоны E_g в пределах одного сорта полупроводника остается постоянной.

С учетом этого в процессе построения зонной диаграммы гетероперехода при сращивании дна зоны проводимости

E_C этих полупроводников на металлургической границе перехода на зонной диаграмме образуется «пик». Величина «пика» ΔE_C равна:

$$\Delta E_C = \chi_{Ge} - \chi_{GaAs}.$$

Параметры выбранных для расчета полупроводниковых материалов

Параметры материала	Обозначение	Германий (pGe)	Арсенид галлия (nGaAs)
Постоянная решетки, Å	a	5,654	5,658
Коэффициент линейного температурного расширения, $10^{-6} K^{-1}$	ТКР	5,9	6,0
Легирующая концентрация, $см^{-3}$	$N_{A,D}$	$3 \cdot 10^{16}$	10^{16}
Расстояние от уровня Ферми до зоны разрешенных энергий, эВ	W_0	0,14	0,17
Расстояние от уровня Ферми до середины запрещенной зоны, эВ	φ	0,21	0,55
Электронное средство, В	χ	4,05	4,07

При сшивании вершины валентной зоны E_V в области металлургического перехода получается «разрыв» ΔE_V . Величина «разрыва» равна:

$$\Delta E_V = -\chi_{Ge} - E_{g\text{ Ge}} + \chi_{GaAs} + E_{g\text{ GaAs}} = -\Delta E_C + (E_{g\text{ GaAs}} - E_{g\text{ Ge}}).$$

Из приведенных соотношений следует, что суммарная величина «пика» ΔE_C и «разрыва» ΔE_V составляет $\Delta E_C + \Delta E_V = (E_{g\text{ GaAs}} - E_{g\text{ Ge}})$.

На рис. 2.37 приведена построенная таким образом зонная диаграмма гетероперехода $pGe - nGaAs$.

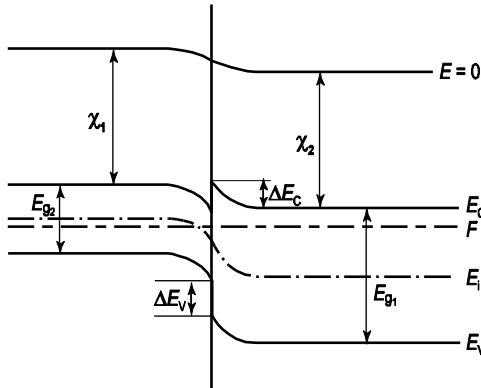


Рис. 2.37. Зонная диаграмма гетероперехода $pGe - nGaAs$ в равновесных условиях

Рассмотрим зонную диаграмму гетероперехода из этих же материалов (германия и арсенида галлия), но с другим типом проводимости — $pGaAs - nGe$ (рис. 2.38). Используем те же самые принципы при построении этой зонной диаграммы. Получаем, что в этом случае «разрыв» наблюдается в энерге-

тическом положении дна зоны проводимости и величина этого «разрыва» ΔE_C равна $\Delta E_C = \chi_{\text{Ge}} - \chi_{\text{GaAs}}$.

«Пик» наблюдается в области металлургического перехода для энергии вершины валентной зоны E_V . Величина «пика» ΔE_V равна

$$\Delta E_V = -\chi_{\text{Ge}} - E_{g \text{ Ge}} + \chi_{\text{GaAs}} + E_{g \text{ GaAs}} = -\Delta E_C + (E_{g \text{ GaAs}} - E_{g \text{ Ge}}).$$

Аналогичным образом можно построить зонные диаграммы для гетеропереходов при любых комбинациях уровней легирования, ширины запрещенной зоны и электронного сродства.

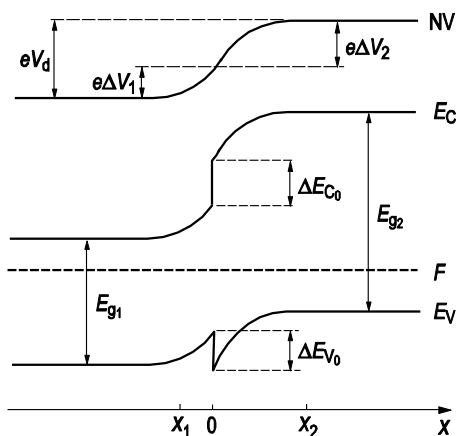


Рис. 2.38. Зонная диаграмма гетероперехода $n\text{Ge} - p\text{GaAs}$ в равновесных условиях

На рис. 2.39 приведены соответствующие зонные диаграммы для различных типов гетеропереходов. Обращает на себя внимание тот факт, что «пик» и «разрыв» для энергетических уровней E_V , E_C в области металлургического перехода могут наблюдаться в различных комбинациях.

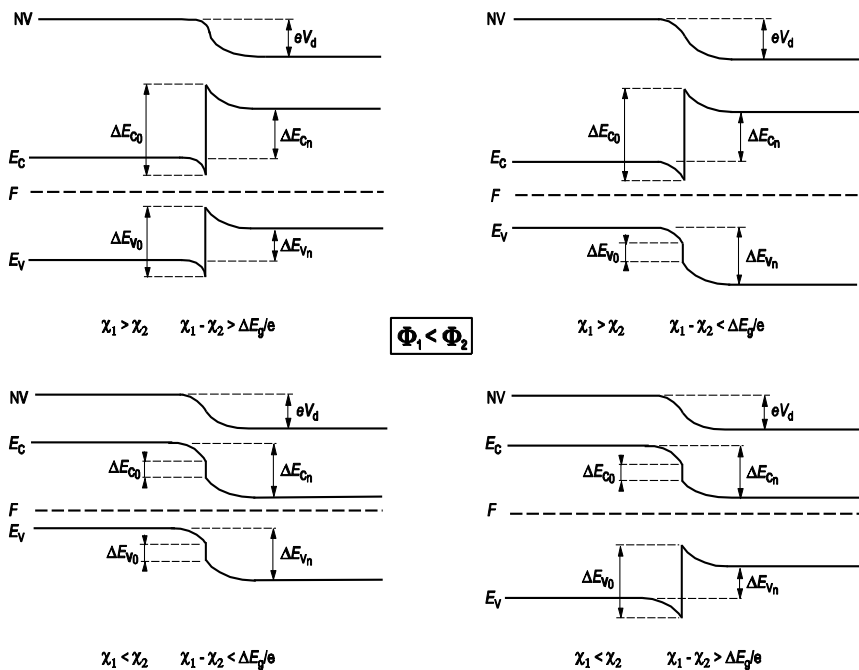


Рис. 2.39. Зонные диаграммы для различных типов гетеропереходов при условии, что термодинамическая работа выхода первого слоя меньше, чем второго ($\Phi_1 < \Phi_2$), и при различных комбинациях для электронного средства (пояснения на рисунках)

Распределение электрического поля и потенциала в области пространственного заряда для гетероперехода будет как и в случае p - n -перехода, но с различными значениями диэлектрических постоянных ϵ_s для левой и правой частей. Решение уравнения Пуассона в этом случае дает следующие выражения для электрического поля E , потенциала V и ширины обедненной области W_{1n} и W_{2p} при наличии внешнего напряжения:

$$E_{1\max} = \frac{qN_D W_{1n}}{\varepsilon_1 \varepsilon_0}; \quad E_{2\max} = \frac{qN_A W_{2p}}{\varepsilon_2 \varepsilon_0}, \quad (2.76)$$

$$V_{1n} = \frac{qN_D W_{1n}^2}{2\varepsilon_1 \varepsilon_0}; \quad V_{2p} = \frac{qN_A W_{2p}^2}{2\varepsilon_2 \varepsilon_0}, \quad (2.77)$$

$$W_{1n} = \sqrt{\frac{2\varepsilon_1 \varepsilon_2 \varepsilon_0 (\Delta\varphi_0 - V)}{qN_D^2 \left(\frac{\varepsilon_1}{N_A} + \frac{\varepsilon_2}{N_D} \right)}}; \quad W_{2p} = \sqrt{\frac{2\varepsilon_1 \varepsilon_2 \varepsilon_0 (\Delta\varphi_0 - V)}{qN_A^2 \left(\frac{\varepsilon_1}{N_A} + \frac{\varepsilon_2}{N_D} \right)}}. \quad (2.78)$$

Полная ширина области пространственного заряда гетероперехода W , равная $W = W_{1n} + W_{2p}$, будет описываться следующим уравнением:

$$W = \sqrt{\frac{2\varepsilon_1 \varepsilon_2 \varepsilon_0 (\Delta\varphi_0 - V)}{q} \left(\frac{1}{N_A \varepsilon_1} + \frac{1}{N_D \varepsilon_2} \right)}. \quad (2.79)$$

Высота потенциального барьера в гетеропереходе $\Delta\varphi_0$ будет определяться суммой потенциалов для каждой из областей гетероперехода:

$$\Delta\varphi_0 = V_{1n} + V_{2p}. \quad (2.80)$$

Функциональная зависимость электрического поля и потенциала в области пространственного заряда гетероперехода от координаты будет соответственно линейной и квадратичной, как и в случае p - n -перехода. Скачок электрического поля в гетеропереходе на металлургической границе обусловлен различными значениями диэлектрических постоянных ε_1 и ε_2 . В этом случае, согласно теореме Гаусса,

$$\varepsilon_1 E_{1\max} = \varepsilon_2 E_{2\max}. \quad (2.81)$$

На рис. 2.40 показаны распределения электрического поля и потенциала в области пространственного заряда гетероперехода.

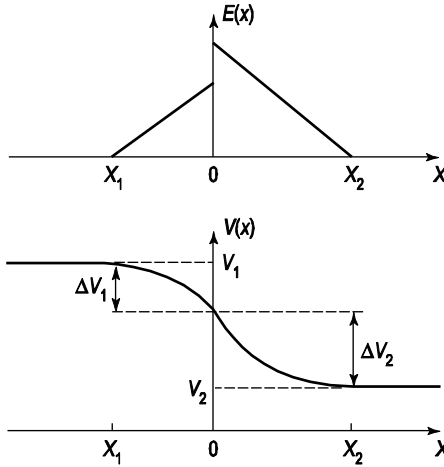


Рис. 2.40. Распределение электрического поля и потенциала в области пространственного заряда гетероперехода $n\text{Ge} - p\text{GaAs}$

Рассмотрим зонную диаграмму гетероперехода при приложении внешнего напряжения V . Как и в случае p - n -перехода, знак напряжения будет определяться знаком приложенного напряжения на p -область гетероперехода. На рис. 2.41 приведены зонные диаграммы при положительном и отрицательном напряжениях на гетеропереходе $n\text{Ge} - p\text{GaAs}$. Пунктиром на этих же зонных диаграммах изображены энергетические уровни в равновесных условиях $V = 0$.

Расчет вольт-амперных характеристик гетероперехода проводится исходя из баланса токов термоэлектронной эмиссии. Для вольт-амперной характеристики гетероперехода получаем следующую зависимость:

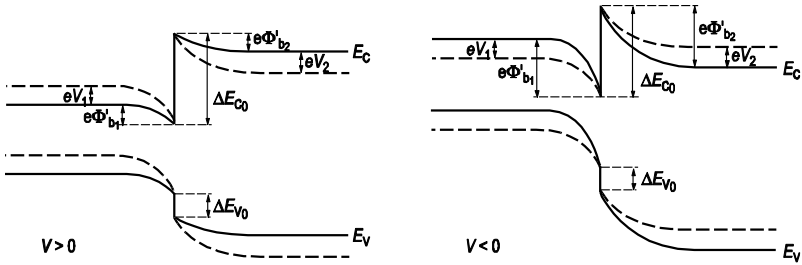


Рис. 2.41. Зонные диаграммы гетероперехода $n\text{Ge} - p\text{GaAs}$ при положительном $V > 0$ и отрицательном $V < 0$ напряжениях.

Пунктиром изображены энергетические уровни в равновесных условиях $V = 0$

$$J = J_s (e^{\beta V} - 1). \quad (2.82)$$

Для различных типов гетеропереходов экспоненциальная зависимость тока от напряжения в виде (2.82) сохраняется, выражение для тока J_s модифицируется.

Для гетеропереходов типа $p\text{Ge} - n\text{GaAs}$ легко реализовать одностороннюю инжекцию даже в случае одинакового уровня легирования в эмиттере $p\text{Ge}$ и базе $n\text{GaAs}$ гетероперехода $N_D = N_A$. Действительно, при прямом смещении отношение дырочной J_p и электронной J_n компонент инжекционного тока будет определяться отношением концентрации неосновных носителей:

$$\frac{J_p}{J_n} = \frac{qL_p p_n}{\tau_p} \bigg/ \frac{qL_n n_p}{\tau_n} \approx \frac{p_n}{n_p} = \frac{n_{i2}^2}{N_D} \bigg/ \frac{n_{i1}^2}{N_A} = \frac{n_{i2}^2}{n_{i1}^2}. \quad (2.83)$$

Поскольку арсенид галлия – более широкозонный полупроводник, чем германий, то собственная концентрация в арсениде галлия (n_{i2}) будет много меньше, чем в германии (n_{i1}), следовательно, дырочная компонента J_p инжекционного тока будет много меньше, чем электронная компонента J_n . Весь инжекционный ток в гетеропереходе pGe – nGaAs будет определяться электронной компонентой.

На зонной диаграмме гетеропереходов видно, что в области «пика» для электронов или дырок реализуется потенциальная яма. Расчеты электрического поля в этой области показывают, что его значение достигает величины $E \sim 10^6$ В/см. В этом случае электронный газ локализован в узкой пространственной области вблизи металлургической границы гетероперехода. Для описания такого состояния используют представление о двумерном электронном газе. Решение уравнения Шредингера свидетельствует о наличии энергетических уровней, существенно отстоящих друг от друга (рис. 2.42).

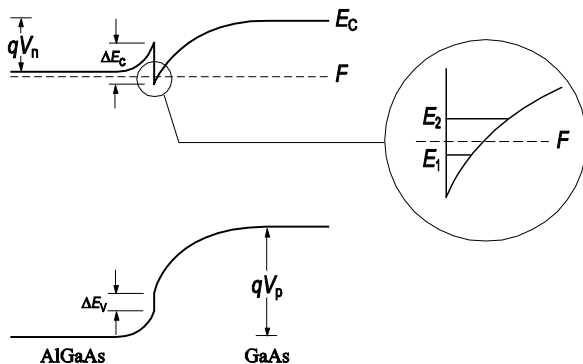


Рис. 2.42. Зонная диаграмма гетероперехода, иллюстрирующая двумерное квантование

Физические свойства двумерного электронного газа существенно отличаются от свойств трехмерного электронного газа. Для двумерного электронного газа меняется плотность квантовых состояний в разрешенных зонах, спектр акустических и оптических фононов, а следовательно, кинетические явления в двумерных системах (подвижность носителей, магнитосопротивление и эффект Холла). Экспериментальные исследования двумерного квантования вблизи металлургической границы гетероперехода позволили изучить и объяснить эти явления.

2.3.2. Контакт полупроводников с одним типом электропроводности

Контакт полупроводников с одним типом электропроводности, но с разной удельной проводимостью обозначают как p^+-p- или n^+-n- переход, причем знаком "+" отмечают область полупроводника с большей концентрацией примесей (рис. 2.43).

Вблизи такого контакта происходят процессы, аналогичные процессам в электронно-дырочном переходе, то есть 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. 2.43, в). Таким образом, на контакте двух полупроводников с одним типом электропроводности, но разной удельной проводимостью также образуется область объемного заряда, диффузионное поле и контактная разность потенциалов. В данном случае в слаболегированной области объемный заряд образован избыточной концентрацией основных носителей заряда.

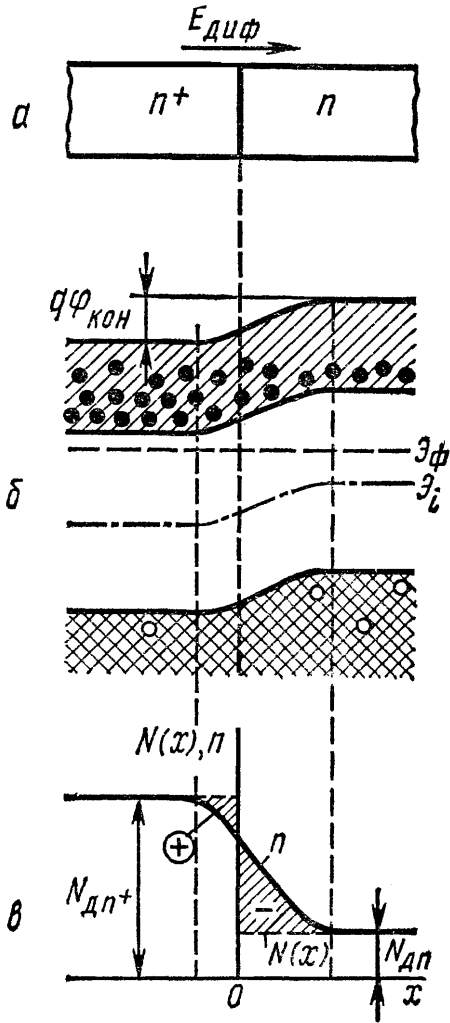


Рис. 2.43. Контакт полупроводников с одним типом электропроводности: *a* - структура контакта; *б* - энергетическая диаграмма контакта; *в* - зависимости концентрации примесей и концентрации носителей заряда от координаты

Контактная разность потенциалов для таких переходов определяется соотношениями, аналогичными (2.4), то есть для p^+ - p -перехода

$$\varphi_{\text{кон}} = \frac{kT}{q} \ln \frac{p_{p^+0}}{p_{po}} ; \quad (2.84)$$

для n^+ - n -перехода

$$\varphi_{\text{кон}} = \frac{kT}{q} \ln \frac{n_{n^+0}}{n_{no}} . \quad (2.85)$$

Распределение носителей заряда вблизи контакта полупроводников с одним типом электропроводности (рис. 2.43, в) показывает, что в данном случае отсутствует слой с меньшей концентрацией по сравнению с концентрацией носителей в слаболегированной области, то есть 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 одним типом электропроводности не обладает выпрямляющими свойствами и имеет малое сопротивление по сравнению с сопротивлением слаболегированной области.

Другой особенностью контакта двух полупроводников с одним типом электропроводности 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 сильно легированной n^+ -области инжекция дырок в высокоомную n -область также оказывается ничтожно малой.

Невыпрямляющие и неинжектирующие контакты широко используют в полупроводниковых приборах наряду с выпрямляющими и инжектирующими. Однако из-за существования на контакте полупроводников с одним типом электропроводности потенциального барьера для неосновных носителей заряда,двигающихся из слаболегированной области к контакту, может происходить накопление этих неосновных носителей вблизи контакта при определенной полярности внешнего напряжения (рис. 2.44). При приложении внешнего напряжения оно падает в основном в объеме слаболегированной области. Поэтому только там будет происходить изгиб энергетических уровней и зон, приводящий к образованию потенциальных ям для неосновных носителей заряда. Эффект накопления неосновных носителей заряда и их последующего рассасывания - эффект инерционный. Поэтому он может ухудшать быстродействие полупроводниковых приборов.

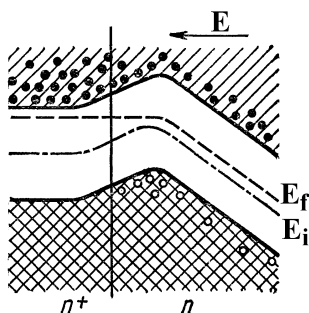


Рис. 2.44. Накопление неосновных носителей заряда (дырок) вблизи контакта двух полупроводников с электропроводностью n -типа при наличие внешнего электрического поля

Промежуточное положение между p^+-p - или n^+-n - и p - n -переходом занимают p - i - или n - i -переходы. Такие переходы образуются между двумя полупроводниками, один из которых имеет электронную или дырочную электропроводность, а другой - собственную.

На рис. 2.45 показаны энергетическая диаграмма и изменение концентраций на границе полупроводников с p - и i -областями. Вследствие разности концентраций носителей заряда в p - и i -областях происходит инжекция дырок из p -области в i -область и электронов из i -области в p -область.

В виду малой величины инжекционной составляющей электронного тока потенциальный барьер на границе перехода создается неподвижными отрицательными ионами акцепторов p -области и избыточными дырками i -области, диффундирующими в нее из p -области. Поскольку $p_{p0} \gg p_i$, глубина распространения запирающего слоя в i -области значительно больше, чем в p -области.

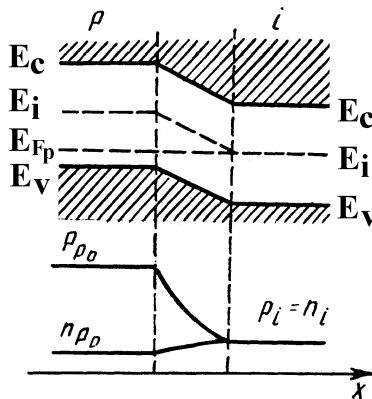


Рис. 2.45. Энергетическая диаграмма p - i -перехода и изменение концентраций на границе полупроводников с p - и i -областями

2.3.3. Контакт металл – полупроводник. Барьер Шоттки

В случае контакта металл - полупроводник возможны различные комбинации (p - и n -типы полупроводника) и соотношения термодинамических работ выхода из металла и полупроводника. В зависимости от этих соотношений в области контакта могут реализоваться три состояния. Первое состояние соответствует условию плоских зон в полупроводнике, в этом случае реализуется нейтральный контакт. Второе состояние соответствует условию обогащения приповерхностной области полупроводника (дырками в p -типе и электронами в n -типе), в этом случае реализуется *омический контакт*. И наконец, в третьем состоянии приповерхностная область полупроводника обеднена основными носителями, в этом случае в области контакта со стороны полупроводника формируется область пространственного заряда ионизованных доноров или акцепторов и реализуется *блокирующий контакт*, или *барьер Шоттки*.

В полупроводниковых приборах наибольшее применение получили блокирующие контакты металл – полупроводник, или барьеры Шоттки. Рассмотрим условие возникновения барьера Шоттки. Ток термоэлектронной эмиссии с поверхности любого твердого тела определяется уравнением Ричардсона:

$$j_T = AT^2 \exp\left(-\frac{\Phi}{kT}\right). \quad (2.86)$$

где A – постоянная Ричардсона; Φ - термодинамическая работа выхода это энергия Ферми с обратным знаком, $\Phi = -F$; T - температура; k - постоянная Больцмана.

Для контакта металл – полупроводник n -типа выберем условие, чтобы термодинамическая работа выхода из полу-

проводника $\Phi_{п/п}$ была меньше, чем термодинамическая работа выхода из металла Φ_{Me} . В этом случае согласно уравнению (2.86) ток термоэлектронной эмиссии с поверхности полупроводника $j_{п/п}$ будет больше, чем ток термоэлектронной эмиссии с поверхности металла:

$$\Phi_{Me} > \Phi_{п/п}; \quad j_{Me} < j_{п/п}.$$

При контакте таких материалов в начальный момент времени ток из полупроводника в металл будет превышать обратный ток из металла в полупроводник и в приповерхностных областях полупроводника и металла будут накапливаться объемные заряды – отрицательные в металле и положительные в полупроводнике. В области контакта возникнет электрическое поле, в результате чего произойдет изгиб энергетических зон. Вследствие эффекта поля термодинамическая работа выхода на поверхности полупроводника возрастет. Этот процесс будет проходить до тех пор, пока в области контакта не выравняются токи термоэлектронной эмиссии и соответственно значения термодинамических работ выхода на поверхности.

На рис. 2.46 показаны зонные диаграммы различных этапов формирования контакта металл – полупроводник. В условиях равновесия в области контакта токи термоэлектронной эмиссии выравнялись, вследствие эффекта поля возник потенциальный барьер, высота которого равна разности термодинамических работ выхода: $\varphi_k = \Phi_{Me} - \Phi_{п/п}$.

Для контакта металл – полупроводник p -типа выберем условие, чтобы термодинамическая работа выхода из полупроводника $\Phi_{п/п}$ была больше, чем термодинамическая работа выхода из металла Φ_{Me} . В этом случае ток термоэлектронной эмиссии с поверхности полупроводника $j_{п/п}$ будет меньше, чем ток термоэлектронной эмиссии с поверхности металла, согласно уравнению (2.86).

При контакте таких материалов в начальный момент времени ток из металла в полупроводник p -типа будет превышать обратный ток из полупроводника в металл и в приповерхностных областях полупроводника и металла будут накапливаться объемные заряды – положительные в металле и отрицательные в полупроводнике.

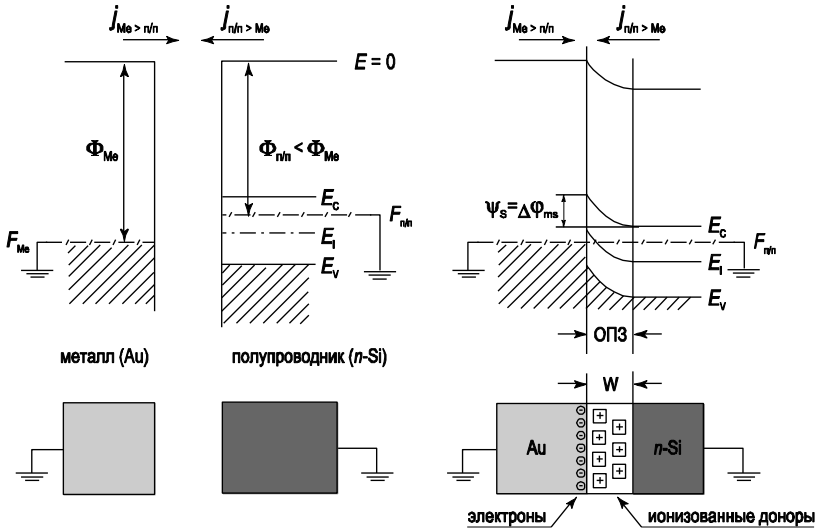


Рис. 2.46. Зонная диаграмма, иллюстрирующая образование барьера Шоттки

В дальнейшем картина перехода к равновесному состоянию и формирования потенциального барьера для контакта металл – полупроводник p -типа аналогична рассмотренной выше для контакта металл – полупроводник n -типа.

Рассмотрим, как меняется зонная диаграмма контакта металл – полупроводник при приложении внешнего напряжения V_G , знак которого соответствует знаку напряжения на металлическом электроде. Величина внешнего напряжения при

положительном знаке $V_G > 0$ не должна быть больше контактной разности потенциала $\Delta\varphi_{ms}$, при отрицательном напряжении $V_G < 0$ она ограничивается только электрическим пробоем структуры. На рис. 2.47 приведены соответствующие зонные диаграммы при положительном и отрицательном напряжениях на металлическом электроде барьеров Шоттки. Из рисунка видно, что роль внешнего напряжения в барьере Шоттки сводится только к регулированию высоты потенциального барьера и величины электрического поля в ОПЗ полупроводника.

Рассмотрим, как меняются электрическое поле и потенциал в области пространственного заряда контакта металл – полупроводник в виде барьера Шоттки. Для определенности будем рассматривать полупроводник n -типа. За знак приложенного напряжения примем знак напряжения, приложенного к металлическому электроду, полупроводниковый электрод считаем заземленным.

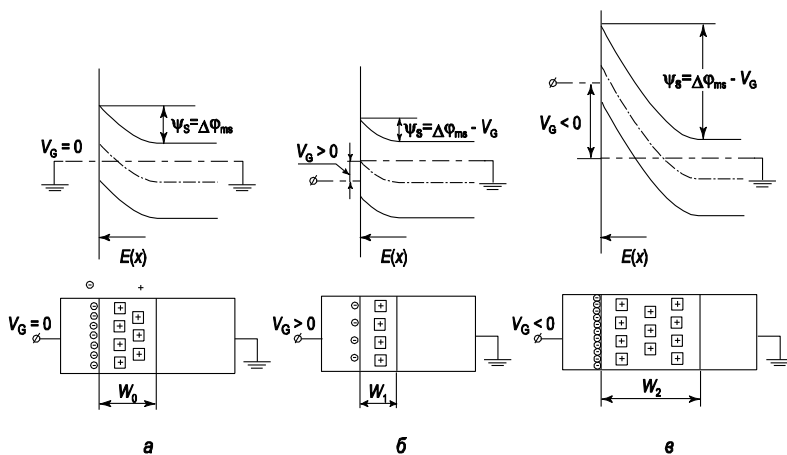


Рис. 2.47. Зонная диаграмма барьера Шоттки при различных напряжениях на затворе:
 а - $V_G = 0$; б - $V_G > 0$, прямое смещение; в - $V_G < 0$, обратное смещение

Вне зависимости от полярности напряжения для барьерных структур все внешнее напряжение будет приложено к области пространственного заряда, поскольку в этой области концентрация свободных носителей существенно меньше, чем в других областях барьера Шоттки.

Связь электрического поля и потенциала для любых материалов с пространственно распределенным объемным зарядом описывается уравнением Пуассона. В одномерном приближении это уравнение имеет вид

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{\rho(x)}{\varepsilon_s \varepsilon_0}, \quad (2.87)$$

где $\psi(x)$ – зависимость потенциала от координаты, $\rho(x)$ – плотность объемного заряда, ε_s – диэлектрическая проницаемость полупроводника, ε_0 – диэлектрическая постоянная.

Заряд в области пространственного заряда барьера Шоттки для полупроводника n -типа обусловлен зарядом ионизованных доноров с плотностью N_D^+ . Поэтому

$$\rho(x) = qN_D^+. \quad (2.88)$$

При интегрировании уравнения Пуассона учтем, что величина электрического поля $E(x) = -\nabla \psi$:

$$\frac{d}{dx} \frac{d\psi}{dx} = -\frac{\rho(x)}{\varepsilon_s \varepsilon_0}, \quad (2.89)$$

или

$$\frac{dE}{dx} = -\frac{qN_D^+}{\varepsilon_s \varepsilon_0}. \quad (2.90)$$

Проведем интегрирование уравнения (2.90). Выберем константу интегрирования из расчета, что при $x = W$ электрическое поле E равно нулю:

$$E(x) = \frac{qN_D^+}{\varepsilon_s \varepsilon_0} (W - x). \quad (2.91)$$

Из соотношения (2.91) следует, что электрическое поле E максимально на границе металл – полупроводник ($x = 0$), линейно спадает по области пространственного заряда и равно нулю на границе ОПЗ – квазинейтральный объем полупроводника ($x = W$).

Для нахождения распределения потенциала (а, следовательно, и зависимости потенциальной энергии от координаты) проинтегрируем еще раз уравнение (2.91) при следующих граничных условиях: $x = W$, $\psi(W) = 0$. Получаем

$$\psi(x) = qN_D \frac{(W - x)^2}{2\varepsilon_s \varepsilon_0}. \quad (2.92)$$

Максимальное значение потенциала реализуется при $x = 0$ и составляет

$$\psi_{\max} = \psi_s - V_G = \Delta\varphi_{\text{ms}} - V_G, \text{ при } \Delta\varphi_{\text{ms}} = \Phi_{\text{Me}} - \Phi_{\text{п/п}}. \quad (2.93)$$

В этом случае можно рассчитать значение ширины обедненной области W , подставляя соотношение (2.93) в (2.92):

$$W = \sqrt{\frac{2\varepsilon_s \varepsilon_0 (\Delta\varphi_{\text{ms}} - V_G)}{qN_D}}. \quad (2.94)$$

Соотношение (2.94) является очень важным для барьерных структур. Это уравнение является универсальным и описывает зависимость ширины обедненной области W от приложенного напряжения V_G и концентрации легирующей примеси N_D для большинства барьерных структур. На рис. 2.48 приведена диаграмма, иллюстрирующая распределение электрического поля и потенциала в барьере Шоттки при обратном смещении, рассчитанных на основании соотношений (2.91) и (2.92).

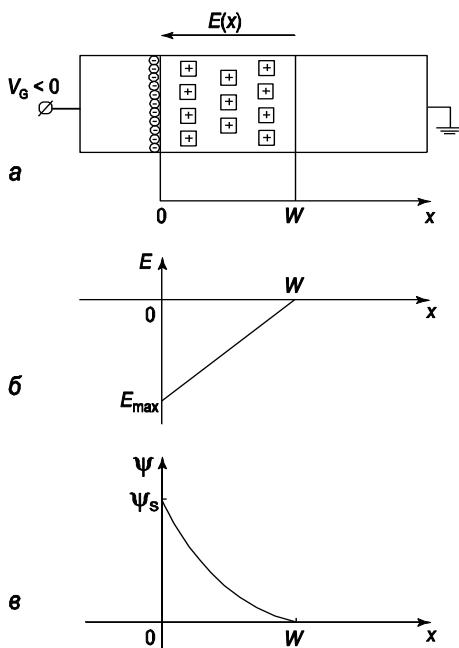


Рис. 2.48. Диаграмма, иллюстрирующая распределение электрического поля и потенциала в барьере Шоттки: а - структура барьера Шоттки при обратном смещении; б - распределение электрического поля в ОПЗ;

ϕ - распределение потенциала в ОПЗ

Вольт-амперная характеристика барьера Шоттки будет иметь вид

$$J = J_0 (e^{\beta V_G} - 1); \quad J_0 = \frac{1}{4} q n_s v_0. \quad (2.95)$$

где v_0 – тепловая скорость электронов, $v_0 = \left(\frac{8kT}{\pi m^*} \right)^{1/2}$; n_s – поверхностная концентрация в полупроводнике на границе с металлом $n_s = n_0 e^{-\beta \Delta \phi_{ms}}$; n_0 – равновесная концентрация основных носителей в полупроводнике, $n_0 = \left(\frac{2\pi m^* kT}{h^2} \right)^{3/2} e^{\frac{E_c - F}{kT}}$.

На рис. 2.49 приведена вольт-амперная характеристика барьера Шоттки.

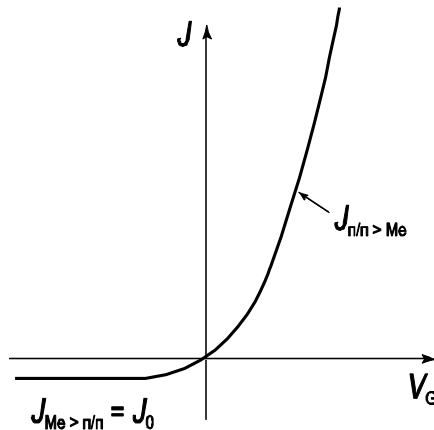


Рис. 2.49. Вольт-амперная характеристика барьера Шоттки

Вольт-амперная характеристика барьера Шоттки имеет ярко выраженный несимметричный вид. В области прямых смещений ток экспоненциально сильно растёт с ростом приложенного напряжения. В области обратных смещений ток от напряжения не зависит. В обоих случаях, при прямом и обратном смещении, ток в барьере Шоттки обусловлен основными носителями – электронами. По этой причине диоды на основе барьера Шоттки являются быстродействующими приборами, поскольку в них отсутствуют рекомбинационные и диффузионные процессы. Несимметричность вольт-амперной характеристики барьера Шоттки – типичная для барьерных структур. Зависимость тока от напряжения в таких структурах обусловлена изменением числа носителей, принимающих участие в процессах зарядопереноса. Роль внешнего напряжения заключается в изменении числа электронов, переходящих из одной части барьерной структуры в другую.

2.3.4. Омические контакты

Обязательным элементом каждого прибора является так называемый *омический контакт*. Он обеспечивает соединение полупроводникового кристалла с внешним выводом. Омический контакт это обычно контакт металл - полупроводник. В случае контакта металла с полупроводником n - типа при условии, что работа выхода металла меньше работы выхода полупроводника ($\Phi_{Me} < \Phi_{nn}$), или же в случае контакта металла с полупроводником p - типа, когда $\Phi_{Me} > \Phi_{nn}$, вблизи границы образуется слой с повышенной концентрацией основных носителей. Такой переход, обладающий повышенной по сравнению с объемом полупроводника удельной проводимостью, называют омическим, так как он не обладает униполярными свойствами. При подключении прямого и обратного напряжения изменяется лишь степень обогащения основными носителями приконтактного слоя.

Для омических контактов характерны следующие особенности:

- близкая к линейной вольт-амперная характеристика. Вольт-амперная характеристика омического контакта должна подчиняться закону Ома $I = U/R$ вне зависимости от полярности приложенного напряжения. Контакт, имеющий линейную вольт-амперную характеристику, не выпрямляет протекающий через него переменный ток, то есть является невыпрямляющим. Таким образом, омический контакт должен иметь линейную и симметричную вольт-амперную характеристику;

- омический контакт должен быть неинжектирующим, то есть через омический контакт должна отсутствовать инжекция неосновных носителей заряда в прилегающую область полупроводника, накопление неосновных носителей в омическом переходе или вблизи него;

- омические контакты должны иметь минимальное электрическое сопротивление, не зависящее от направления и значения тока в рабочем диапазоне токов;

- металл перехода должен обладать высокой электро- и теплопроводностью и температурным коэффициентом расширения (ТКР), близким к ТКР полупроводника;

- омические контакты должны иметь стабильные электрические и механические свойства.

Обычно для создания омического контакта используют следующие металлы: свинец, олово и их сплавы, золото, алюминий.

Создание омических контактов представляет собой довольно сложную задачу. Одним из способов улучшения свойств омического контакта является применение в качестве полупроводника n^+-n или p^+-p -структуры, то есть в качестве омического контакта необходимо использовать структуру $m-n^+-n$ или $m-p^+-p$, где через m обозначен слой металла. Типичная структура омического контакта приведена на рис. 2.50.

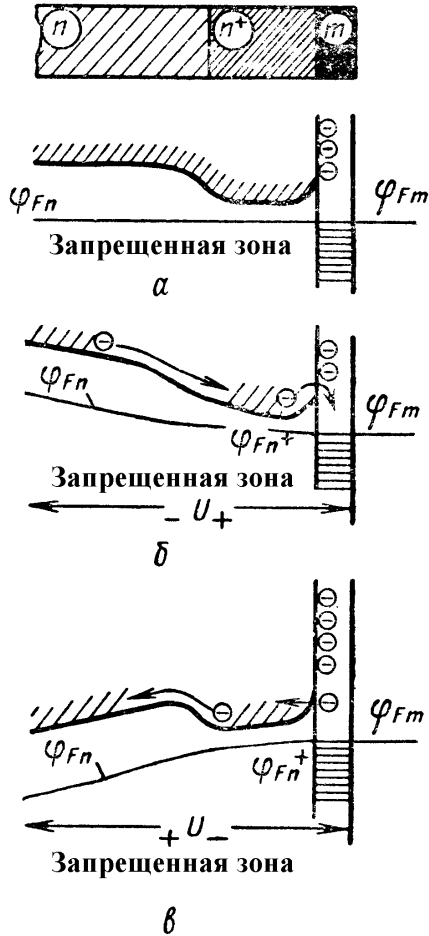


Рис. 2.50. Зонные диаграммы омического контакта:
 а - равновесное состояние; б - обратное напряжение;
 в - прямое напряжение

Эта структура состоит из двух переходов: $m-n^+$ и n^+-n , n^+ -слой сильно легирован (порядка 10^{19} см^{-3}). Поскольку степень легирования n^+ -слоя велика и полупроводник вырожден, то толщина ОПЗ между металлом и n^+ -полупроводником

чрезвычайно мала (2 - 3 нм) и носители заряда беспрепятственно преодолевают барьер за счет туннельного эффекта. n^+ -слой имеет удельное сопротивление во много раз меньше, чем удельное сопротивление объема полупроводника. Так как приконтактная область будет характеризоваться высокой концентрацией примесей, инжекция неосновных носителей в ней будет ослаблена.

Оба перехода $m-n^+$ и n^+-n не являются инжектирующими и не обладают вентильными свойствами. В целом структура $m-n^+-n$ ведет себя как омическое сопротивление при любой полярности напряжения. Рассмотрим механизм протекания токов. Пусть напряжение приложено минусом к n -области, плюсом к металлу. Тогда потенциалы слоев n^+ и n повысятся, высота барьера n^+-n увеличится, а высота барьера $m-n^+$ уменьшится (рис. 2.50, б). Электроны из слоя n будут свободно переходить в слой n^+ независимо от высоты барьера n^+-n , а понижение барьера $m-n^+$ обеспечит переход электронов из n^+ -слоя в m -слой. Пусть теперь напряжение приложено плюсом к n -области. При этом потенциалы n^+ и n слоев понизятся и высота барьера n^+-n сделается меньше; соответственно электроны n^+ -слоя смогут переходить в n -слой. Барьер $m-n^+$ повысится, но так как он очень тонкий, то электроны слоя m будут проходить его за счет туннельного эффекта, как показано на рис. 2.50, в.

Таким образом, важнейшим свойством омического контакта является его двухсторонняя проводимость.

Другое важное свойство связано с ничтожным временем жизни носителей в n^+ -слое, поскольку он сильно легирован и имеет малое удельное сопротивление. Интенсивная рекомбинация в n^+ -слое и отсутствие инжекции делают повышение концентрации в области омического контакта редким явлением. Поэтому считают, что концентрация электронов и дырок на омическом контакте имеют равновесные значения.

2.3.5. Явления на поверхности полупроводников

Физические процессы на поверхности полупроводника оказывают большое влияние на электрические характеристики и параметры полупроводниковых приборов и интегральных схем.

Рассмотрим зонную диаграмму приповерхностной области полупроводников в равновесных условиях. Рассмотрим, как будет меняться концентрация свободных носителей в приповерхностной области полупроводника, когда вблизи этой поверхности создается электрическое поле. Для примера будем считать, что электрическое поле создается заряженной металлической плоскостью с поверхностной плотностью зарядов σ . Поскольку силовые линии электрического поля должны быть замкнуты, то на поверхности полупроводника возникает равный по величине, но противоположный по знаку электрический заряд. В зависимости от знака заряда на металлической плоскости (положительной или отрицательной) экранирующий это поле заряд в приповерхностной области полупроводника также будет различных знаков. На рис. 2.51 приведены ситуации положительно и отрицательно заряженной плоскости.

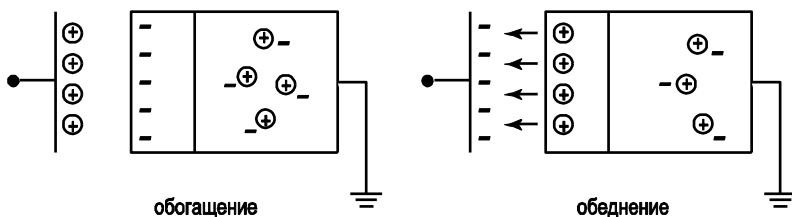


Рис. 2.51. Изменение концентрации свободных носителей в приповерхностной области полупроводника при наличии вблизи поверхности заряженной металлической плоскости

Случай, когда в приповерхностной области возрастает концентрация свободных носителей, носит название *обогащение*, а когда в приповерхностной области уменьшается концентрация свободных носителей – *обеднение*.

Изменение концентрации свободных носителей в приповерхностной области полупроводника под действием внешнего электрического поля получило название *эффекта поля*.

При наличии внешнего поля приповерхностная область в полупроводнике не будет электронейтральной. Заряд, возникший в этой области, обычно называется пространственным зарядом, а сама область – областью пространственного заряда. Наличие электрического поля $E(z)$ в ОПЗ меняет величину потенциальной энергии электрона. Если поле направлено от поверхности вглубь полупроводника, то электроны в этом случае будут иметь минимальную энергию у поверхности, что соответствует наличию потенциальной ямы для электронов.

Изменение потенциальной энергии электронов

$$\Delta U = U(z) - U(\infty) = \int_{\infty}^z E(z) dz, \quad (2.96)$$

где $U(\infty)$ – потенциальная энергия электронов в квазинейтральном объеме полупроводника.

Поскольку на дне зоны проводимости кинетическая энергия электронов равна нулю ($E = \frac{\hbar^2 k^2}{2m^*}$), то изменение потенциальной энергии по координате должно точно так же изменить энергетическое положение дна зоны проводимости E_C , а соответственно и вершины валентной зоны E_V . На зонных диаграммах это выражается в изгибе энергетических зон.

Величина разности потенциалов между квазинейтральным объемом и произвольной точкой ОПЗ получила название *электростатического потенциала*

$$\psi = \frac{1}{q} \int_{\infty}^z E(z) dz. \quad (2.97)$$

Значение электростатического потенциала на поверхности полупроводника называется **поверхностным потенциалом** и обозначается символом ψ_s .

Знак поверхностного потенциала ψ_s соответствует знаку заряда на металлическом электроде, вызывающего изгиб энергетических зон.

При $\psi_s > 0$ зоны изогнуты вниз, при $\psi_s < 0$ зоны изогнуты вверх (рис. 2.52).

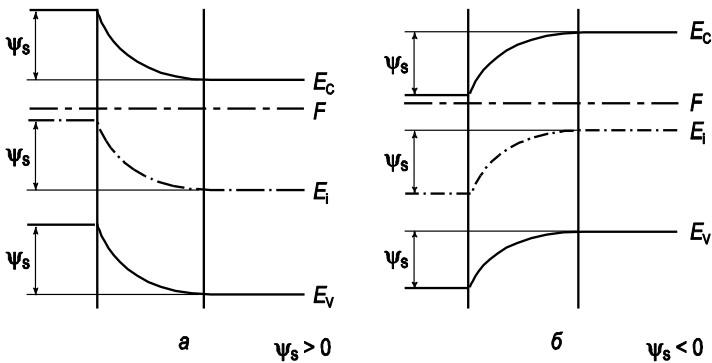


Рис. 2.52. Энергетические зоны на поверхности полупроводника *n*-типа:

а - в случае обеднения; *б* - в случае обогащения

Величины n_s и p_s – концентрации электронов и дырок на поверхности – носят названия **поверхностных концентраций**

Количественной характеристикой эффекта поля, характеризующей глубину проникновения поля в полупроводник, является дебаевская длина экранирования.

$$L_D = \sqrt{\frac{kT\varepsilon_s\varepsilon_0}{q^2n_0}} = \sqrt{\frac{\varepsilon_s\varepsilon_0}{qN_D} \frac{kT}{q}} \quad (2.98)$$

Таким образом, по физическому смыслу дебаевская длина экранирования L_D соответствует среднему расстоянию, на которое проникает электрическое поле в полупроводник при малых уровнях возмущения.

Кроме рассмотренных явлений, электропроводность приповерхностного слоя полупроводника может изменяться под действием электрического поля, возникающего за счет напряжения, прикладываемого к металлу и полупроводнику, разделенных диэлектриком. Если предположить, что до включения напряжения поверхностные состояния на границе полупроводника и диэлектрика отсутствуют, то электропроводности приповерхностного слоя и объема полупроводника будут одинаковыми.

При включении напряжения между металлом и полупроводником возникает электрическое поле и на поверхности металла и приповерхностном слое полупроводника, как на пластинах конденсатора, накапливаются заряды. Электропроводность приповерхностного слоя полупроводника меняется. Зависимость электропроводности приповерхностного слоя полупроводника от напряжения принято называть эффектом поля.

3. ФИЗИКА СТРУКТУРЫ МЕТАЛЛ - ДИЭЛЕКТРИК - ПОЛУПРОВОДНИК

3.1. Идеальная структура металл – диэлектрик – полупроводник

Структуры металл – диэлектрик – полупроводник, или сокращенно МДП-структуры, широким интересом к изучению их физических свойств обязаны появлению планарной технологии и развитию нового класса полупроводниковых приборов, работающих на основе эффекта поля, таких как приборы с зарядовой связью, полевые транзисторы с изолированным затвором, репрограммируемые элементы памяти с плавающим затвором и т.п. МДП-структуры позволяют анализировать основные процессы, протекающие в такого рода приборах, и являются чрезвычайно удобными объектами исследования. Устройство МДП-структуры следует из ее названия.

МДП-структура представляет собой монокристаллическую пластину полупроводника, называемую *подложкой*, закрытую с планарной стороны *диэлектриком*. Металлический электрод, нанесенный на диэлектрик, носит название *затвора*, а сам диэлектрик называется *подзатворным*. На обратную непланарную сторону полупроводниковой пластины наносится металлический электрод, называющийся омическим контактом. Довольно часто в качестве диэлектрика в МДП-структурах используют окислы, поэтому вместо МДП употребляется название МОП-структура.

Итак, МДП-структура, приведенная на рис. 3.1, состоит из затвора, подзатворного диэлектрика, полупроводниковой подложки и омического контакта.

Рассмотрим зонную энергетическую диаграмму МДП-структуры при равновесных условиях. Согласно правилу построения зонных диаграмм необходимо, чтобы в системе при отсутствии приложенного напряжения:

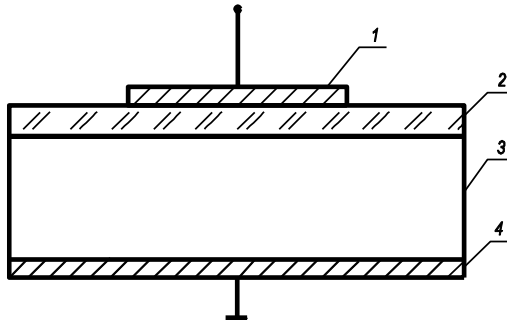


Рис. 3.1. Устройство МДП-структуры:
 1 – затвор, 2 – подзатворный диэлектрик;
 3 – полупроводниковая подложка;
 4 – омический контакт

- а) уровень вакуума был непрерывен;
- б) электронное сродство диэлектрика и полупроводника в каждой точке было постоянно;
- в) уровень Ферми был одинаков.

На рис. 3.2 приведена построенная таким образом зонная диаграмма для *идеальной МДП-структуры*. Под идеальной МДП-структурой будем понимать такую систему металл – диэлектрик – полупроводник, когда:

- отсутствуют поверхностные состояния на границе раздела полупроводник – диэлектрик;
- термодинамические работы выхода металла затвора и полупроводника подложки равны между собой;
- отсутствуют заряженные центры в объеме подзатворного диэлектрика;
- сопротивление подзатворного диэлектрика бесконечно велико, так что сквозной ток через него отсутствует при любых напряжениях на затворе.

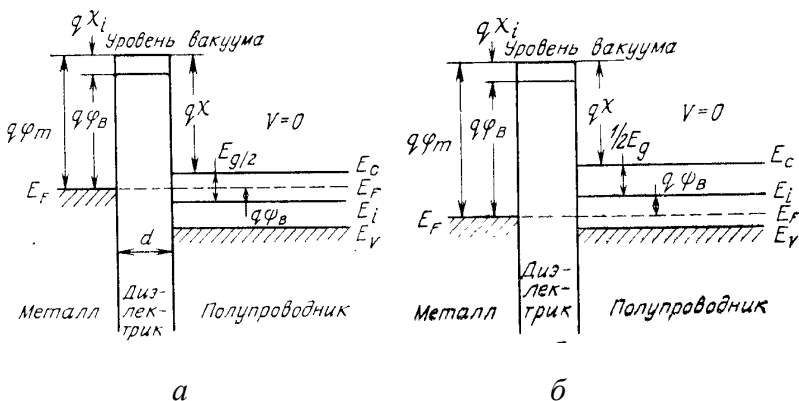


Рис. 3.2. Зонные диаграммы идеальных МДП-структур при напряжении на затворе $V = 0$: *а* – полупроводник *n*-типа; *б* - полупроводник *p*-типа

МДП-структуры, близкие к идеальным, получают, используя «хлорную» технологию термического выращивания диоксида кремния на кремнии, причем для *n*-Si в качестве материала затвора используется алюминий, а для *p*-Si - золото.

МДП-структуры, в которых нарушается одно из вышеперечисленных требований, получили название реальных МДП-структур.

Когда к идеальной МДП-структуре приложено напряжение того или другого знака, на полупроводниковой поверхности могут возникнуть три основные ситуации (рис. 3.3). Рассмотрим их сначала для МДП-структуры с полупроводником *p*-типа. Если к металлическому электроду структуры приложено отрицательное напряжение ($V < 0$), край валентной зоны у границы с диэлектриком изгибается вверх и приближается к уровню Ферми (рис. 3.3, *а*). Поскольку в идеальной МДП-структуре сквозной ток равен нулю, уровень Ферми в полупроводнике остается постоянным. Так как концентрация дырок экспоненциально зависит от разности энергий ($E_F - E_V$),

такой изгиб зон приводит к увеличению числа основных носителей (дырок) у поверхности полупроводника. Этот режим называется **режимом обогащения (аккумуляции)**.

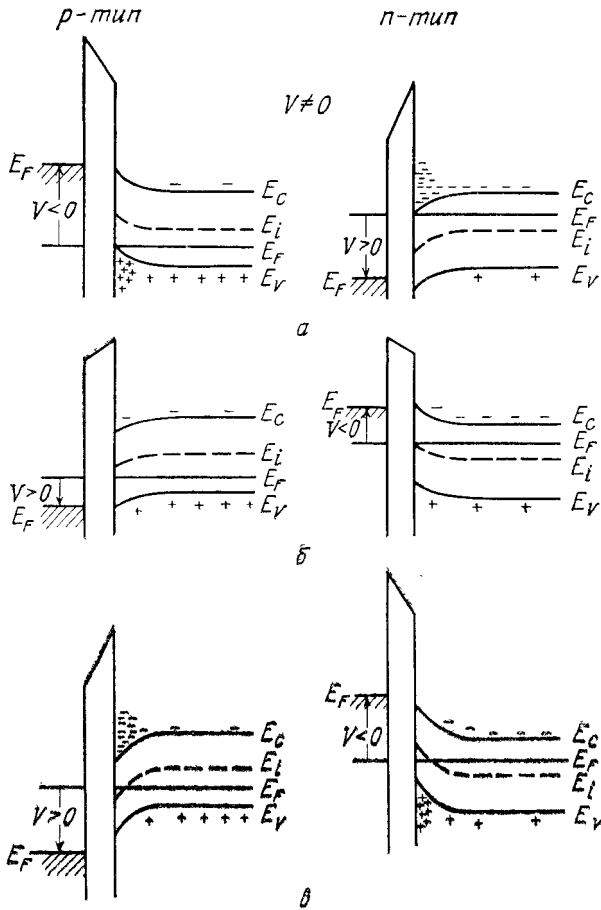


Рис. 3.3. Зонные диаграммы идеальных МДП-структур при напряжении на затворе $V \neq 0$: а – режим аккумуляции; б – режим обеднения; в – режим инверсии

Если к МДП-структуре приложено не слишком большое положительное напряжение ($V > 0$), зоны изгибаются в обратном направлении и приповерхностная область полупроводника обедняется основными носителями (рис. 3.3, б). Этот режим называют *режимом обеднения* или *истощения* поверхности. При больших положительных напряжениях зоны изгибаются вниз настолько сильно, что вблизи поверхности происходит пересечение уровня Ферми E_F с собственным уровнем E_i . В этом случае (рис. 3.3, в) концентрация неосновных носителей (электронов) у поверхности превосходит концентрацию основных носителей (дырок). Эта ситуация называется *режимом инверсии*. Аналогичное рассмотрение можно провести и для МДП-структуры с полупроводником n -типа. Указанные режимы осуществляются при напряжении противоположной полярности.

3.2. Область пространственного заряда в равновесных условиях

Рассмотрим изменение энергетического спектра свободных носителей заряда в приповерхностной области полупроводника под действием внешнего электрического поля. При этом будем считать, что на поверхности полупроводника энергетический спектр при отсутствии внешнего поля точно такой же, как и в объеме, то есть поверхность полупроводника является идеальной, без поверхностных состояний.

Из-за малой концентрации свободных носителей заряда в объеме полупроводника возможно проникновение электрического поля вглубь полупроводника на большие, по сравнению с межатомными, расстояния. Проникшее электрическое поле перераспределяет свободные носители заряда. Это явление получило название *эффекта поля*. Таким образом, *эффект поля* – это изменение концентрации свободных носите-

лей в приповерхностной области полупроводника под действием внешнего электрического поля.

Поскольку заряд свободных носителей или ионизованных доноров пространственно распределен в приповерхностной области полупроводника и эта область не является электронейтральной, она получила название область пространственного заряда.

Отметим, что в случае реализации эффекта поля источником внешнего электрического поля могут быть заряды на металлических пластинах вблизи поверхности полупроводника, заряды на границе и в объеме диэлектрического покрытия и т.д.

Наличие электрического поля $E(z)$ в ОПЗ меняет величину потенциальной энергии электрона в этой области. Если электрическое поле направлено от поверхности вглубь полупроводника, то электроны будут иметь минимальную энергию в этом поле вблизи поверхности, где для них энергетическое положение соответствует наличию потенциальной ямы. Поскольку на дне зоны проводимости кинетическая энергия электронов равна нулю, изменение потенциальной энергии по координате должно изменить точно так же ход дна зоны проводимости, а соответственно и вершины валентной зоны. Этот эффект изображен на зонных диаграммах, приведенных на рис. 3.3, 3.4, и получил название *изгиба энергетических зон*. Величина разности потенциалов между квазинейтральным объемом и произвольной точкой ОПЗ получила название *электростатического потенциала*

$$\psi = \frac{1}{q} \int_{\infty}^z E(z) dz . \quad (3.1)$$

Значение электростатического потенциала на поверхности полупроводника называется *поверхностным потенциа-*

лом и обозначается ψ_s . На зонной диаграмме (рис. 3.4) величина ψ_s отрицательна.

Выразим концентрацию электронов n и дырок p в ОПЗ через электростатический потенциал ψ . В квазинейтральном объеме в невырожденном случае концентрация электронов и дырок имеет вид

$$\begin{aligned} n_0 &= n_i e^{\beta\varphi_0}; \\ p_0 &= p_i e^{-\beta\varphi_0}, \end{aligned} \quad (3.2)$$

где $\beta = \frac{q}{kT}$; φ_0 – расстояние от уровня Ферми до середины запрещенной зоны в квазинейтральном объеме. Величины n и p в ОПЗ вычисляются по формулам

$$\begin{aligned} n &= n_i e^{\beta\varphi} = n_i e^{\beta(\psi+\varphi_0)} = n_0 e^{\beta\psi}; \\ p &= p_i e^{-\beta\varphi} = p_i e^{-\beta(\psi+\varphi_0)} = p_0 e^{-\beta\psi}. \end{aligned} \quad (3.3)$$

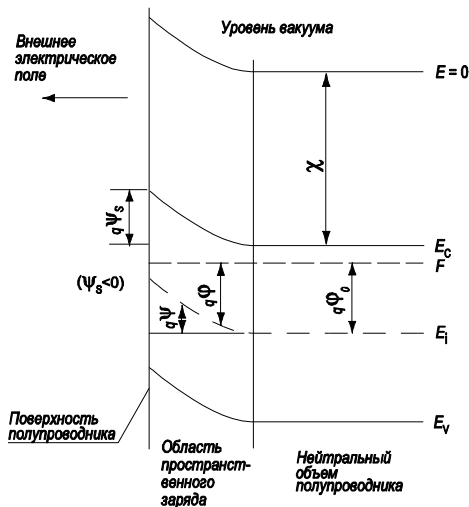


Рис. 3.4. Зонная диаграмма приповерхностной области полупроводника n -типа

Величины концентраций электронов n_s и дырок p_s на поверхности носят название **поверхностной концентрации** и имеют значения

$$\begin{aligned} n_s &= n_0 e^{\beta\psi_s}, \\ p_s &= p_0 e^{-\beta\psi_s}. \end{aligned} \quad (3.4)$$

В зависимости от направления и величины внешнего электрического поля, типа полупроводниковой подложки различают 4 различных состояния поверхности полупроводника: обогащение, обеднение, слабая инверсия и сильная инверсия. Все эти ситуации отражены на рис. 3.5 для полупроводника n -типа.

Обогащение – состояние поверхности полупроводника, когда поверхностная концентрация основных носителей больше, чем концентрация основных носителей в нейтральном объеме (см. рис. 3.5, *a*). В режиме обогащения для полупроводника n -типа: $n_s > n_0$, зоны изогнуты вниз, $\psi_s > 0$; для полупроводника p -типа: $p_s > p_0$, зоны изогнуты вверх, $\psi_s < 0$.

Обеднение – состояние поверхности полупроводника, когда поверхностная концентрация основных носителей меньше, чем концентрация основных носителей в квазинейтральном объеме, но больше, чем поверхностная концентрация неосновных носителей (см. рис. 3.5, *б*). В режиме обеднения для полупроводника n -типа: $p_s < n_s < n_0$, зоны изогнуты вверх, $\psi_s < 0$, $0 < |\psi_s| < \varphi_0$; для полупроводника p -типа: $n_s < p_s < p_0$, зоны изогнуты вниз, $\psi_s > 0$, $0 < \psi_s < \varphi_0$.

Переход от состояния обогащения к состоянию обеднения происходит при значении поверхностного потенциала $\psi_s = 0$, получившем название **потенциала «плоских» зон**. При этом концентрации основных и неосновных носителей на поверхности и в объеме совпадают.

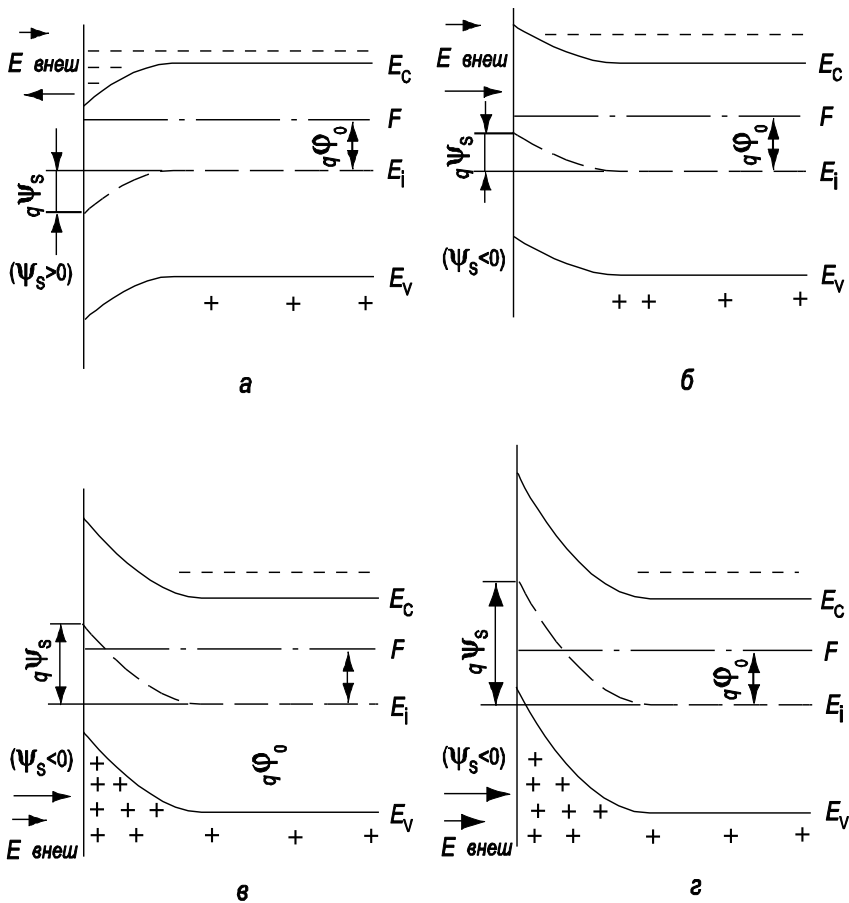


Рис. 3.5. Зонная диаграмма приповерхностной области полупроводника n -типа при различных состояниях поверхности: a - обогащение; $б$ - обеднение; $в$ - слабая инверсия; $г$ - сильная инверсия

Слабая инверсия – состояние поверхности полупроводника, когда поверхностная концентрация неосновных носите-

лей больше, чем поверхностная концентрация основных, но меньше, чем концентрация основных носителей в квазинейтральном объеме (см. рис. 3.5, в). В режиме слабой инверсии для полупроводника n -типа: $n_s < p_s < n_0$, зоны изогнуты вверх, $\psi_s < 0$, $\varphi_0 < |\psi_s| < 2\varphi_0$; для полупроводника p -типа: $p_s < n_s < p_0$, зоны изогнуты вниз, $\psi_s > 0$, $\varphi_0 < \psi_s < 2\varphi_0$.

Переход от области обеднения к области слабой инверсии происходит при значении поверхностного потенциала $|\psi_s| = \varphi_0$, соответствующем состоянию поверхности с собственной проводимостью

$$n_s = p_s = n_i.$$

Сильная инверсия – состояние поверхности полупроводника, когда поверхностная концентрация неосновных носителей больше, чем концентрация основных носителей в квазинейтральном объеме (см. рис. 3.5, г). В режиме сильной инверсии для полупроводника n -типа: $p_s > n_0$, зоны изогнуты вверх, $\psi_s < 0$, $|\psi_s| > 2\varphi_0$; для полупроводника p -типа: $n_s > p_0$, зоны изогнуты вниз, $\psi_s > 0$, $\psi_s > 2\varphi_0$.

Переход от области слабой инверсии к области сильной инверсии происходит при значении поверхностного потенциала $\psi_s = 2\varphi_0$, получившем название «**порогового**» **потенциала**. При этом концентрация неосновных носителей на поверхности равна концентрации основных носителей в объеме полупроводника.

Та область в ОПЗ, где суммарная концентрация свободных носителей электронов и дырок меньше, чем концентрация ионизированной примеси, называется **областью обеднения**. Область в ОПЗ, где концентрация свободных неосновных носителей больше, чем основных, получила название **инверсионного канала**.

3.3. Приповерхностная область пространственного заряда

Одной из основных задач при анализе области пространственного заряда полупроводника является нахождение связи между электростатическим потенциалом $\psi(z)$, с одной стороны, и величинами заряда в области пространственного заряда Q_s , емкости ОПЗ C_s – с другой. Нахождение этой связи основано на решении уравнения Пуассона для ОПЗ.

Уравнение Пуассона для полупроводника p -типа

$$\frac{d^2\psi}{dz^2} = \frac{\rho(z)}{\varepsilon_0\varepsilon_s}. \quad (3.5)$$

Величина $\rho(z)$ в общем случае, когда отсутствует ограничение на малость возмущения, принимает значение

$$\rho(z) = q(N_D^+ - N_A^- + p - n). \quad (3.6)$$

В квазинейтральном объеме, где условие электронейтральности выполняется, $\rho(z) = 0$.

Тогда

$$N_D^+ - N_A^- = n_0 - p_0. \quad (3.7)$$

Поскольку, как было показано в (3.2 – 3.4),

$$\begin{aligned} n &= n_0 e^{\beta\varphi}, & p &= p_0 e^{-\beta\varphi}, \\ n_0 &= n_i e^{-\beta\varphi_0}, & p_0 &= p_i e^{\beta\varphi_0}, \end{aligned}$$

для $\rho(z)$ в ОПЗ имеем

$$\rho(z) = -qp_0 [e^{-2\beta\varphi_0} (e^{\beta\psi} - 1) - e^{-\beta\psi} - 1]. \quad (3.8)$$

Подставляя (3.8) в (3.5), имеем для нахождения $\psi(z)$ дифференциальное уравнение

$$\frac{d^2\psi}{dz^2} = \frac{qp_0}{\varepsilon_s \varepsilon_0} [e^{-2\beta\varphi_0} (e^{\beta\psi} - 1) - e^{-\beta\psi} - 1]. \quad (3.9)$$

Домножим выражение для дебаевской длины экранирования, которое представлено в главе 2 формулой (2.98), слева и справа на величину $\frac{d\psi}{dz}$. Тогда левая часть выражения принимает вид

$$\frac{d\psi}{dz} \cdot \frac{d^2\psi}{dz^2} = \frac{1}{2} \frac{d}{dz} \left(\frac{d\psi}{dz} \right)^2. \quad (3.10)$$

Следовательно,

$$d \left(\frac{d\psi}{dz} \right)^2 = \frac{qp_0}{\varepsilon_s \varepsilon_0} [e^{-2\beta\varphi_0} (e^{\beta\psi} - 1) - e^{-\beta\psi} - 1] d\psi. \quad (3.11)$$

Проинтегрировав (3.11) от бесконечности до некоторой точки ОПЗ, получаем

$$\left(\frac{d\psi}{dz} \right)^2 = \frac{1}{2} \frac{qp_0}{\varepsilon_s \varepsilon_0} \frac{1}{\beta} [(e^{-\beta\psi} + \beta\psi - 1) + e^{-2\beta\varphi_0} (e^{\beta\psi} - \beta\psi - 1)]. \quad (3.12)$$

Воспользовавшись определением дебаевской длины экранирования L_D (2.98), а также соотношением $E(z) = -\frac{d\psi}{dz}$, получаем

$$E^2 = \left(\frac{kT}{q} \right)^2 \frac{1}{2L_D^2} [(e^{-\beta\psi} + \beta\psi - 1) + e^{-2\beta\varphi_0} (e^{\beta\psi} - \beta\psi - 1)]. \quad (3.13)$$

Обозначим

$$F(\psi, \varphi_0) \equiv [(e^{-\beta\psi} + \beta\psi - 1) + e^{-2\beta\varphi_0} (e^{\beta\psi} - \beta\psi - 1)]^{1/2}. \quad (3.14)$$

Из (3.13) и (3.14) имеем

$$E = -\frac{d\psi}{dz} = \pm \frac{kT}{q} \frac{1}{\sqrt{2}L_D} F(\psi, \varphi_0). \quad (3.15)$$

Соотношение (3.15) называется *первым интегралом уравнения Пуассона*.

Знак электрического поля выбирается в зависимости от знака поверхностного потенциала. Если $\psi_s > 0$ (обеднение основными носителями или инверсия), поле направлено вглубь полупроводника по оси z и положительно. При $\psi_s < 0$ поле E направлено против оси z и отрицательно.

Величина электрического поля на поверхности

$$E_s = \pm \frac{kT}{q} \frac{1}{\sqrt{2}L_D} F(\psi_s, \varphi_0). \quad (3.16)$$

Поскольку согласно теореме Гаусса величина электрического поля на поверхности E_s связана определенным образом с плотностью пространственного заряда на единицу площади Q_{sc} , имеем

$$Q_{sc} = \varepsilon_s \varepsilon_0 E_s = \pm \frac{\sqrt{2} \varepsilon_s \varepsilon_0 kT}{qL_D} F(\psi_s, \varphi_0). \quad (3.17)$$

Выражение (3.17) для заряда в ОПЗ справедливо для любых значений поверхностного потенциала. Однако использование его для конкретных случаев довольно затруднено в силу громоздкости функции $F(\psi, \varphi_0)$ в виде (3.14). На рис. 3.6 приведено значение заряда в ОПЗ Q_{sc} как функции поверхностного потенциала ψ_s , рассчитанное для конкретного случая.

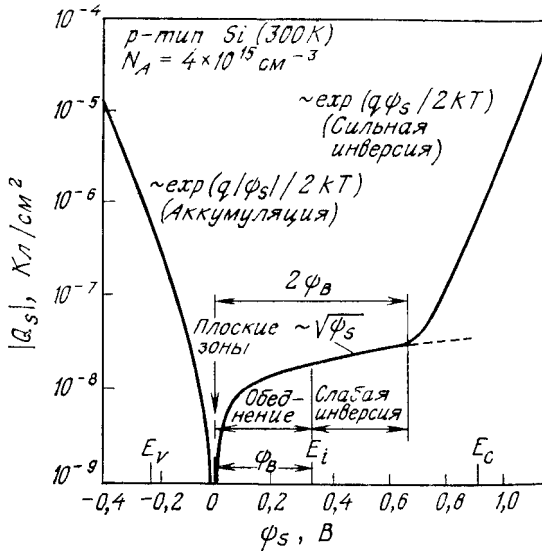


Рис. 3.6. Зависимость заряда в ОПЗ от поверхностного потенциала ψ_s , рассчитанная для кремния p -типа

3.4. Распределение плотности пространственного заряда, электрического поля и потенциала в идеальной МДП-структуре в режиме сильной инверсии

При смещениях, соответствующих режимам обеднения и инверсии, приложенное напряжение отталкивает дырки от поверхности, образуя обедненный слой, и индуцирует электро-

ны, образующие инверсный слой. Зонная диаграмма идеальной МДП-структуры, с полупроводником p -типа в режиме сильной инверсии имеет вид, приведенный на рис. 3.7, *a*. На рис. 3.7, *б* приведено распределение заряда в структуре. Для обеспечения электронейтральности структуры заряд на ее металлическом электроде Q_M должен быть равен сумме электронного заряда в инверсном слое $Q_{inv} = Q_n$ и заряда ионизированных акцепторов в области обеднения полупроводника:

$$Q_M = Q_{inv} + qN_aW = Q_{inv} + Q_B = Q_S. \quad (3.18)$$

Все заряды отнесены к единице площади границы раздела, W – толщина обедненного слоя, Q_S – полная поверхностная плотность заряда в полупроводнике.

Распределение электрического поля (рис. 3.7, *б*) потенциала (рис. 3.7, *г*) соответствуют первому и второму интегралу уравнения Пуассона. Поле на границе между диэлектриком и полупроводником имеет разрыв, вследствие различных значений диэлектрической проницаемости граничащих материалов оно уменьшается от $E_i = Q_S/\varepsilon_i$ до E_s . Напряженность поля в инверсном слое резко уменьшается по мере удаления от поверхности

$$E(x) = E_s \exp\left(-\frac{x}{L_D}\right). \quad (3.19)$$

В области обеднения электрическое поле между подложкой и затвором однородное и описывается уравнением Пуассона

$$\frac{d^2\varphi}{dx^2} = \frac{\rho}{\varepsilon_0\varepsilon_s} \quad (3.20)$$

При режиме обеднения в полупроводнике возникает

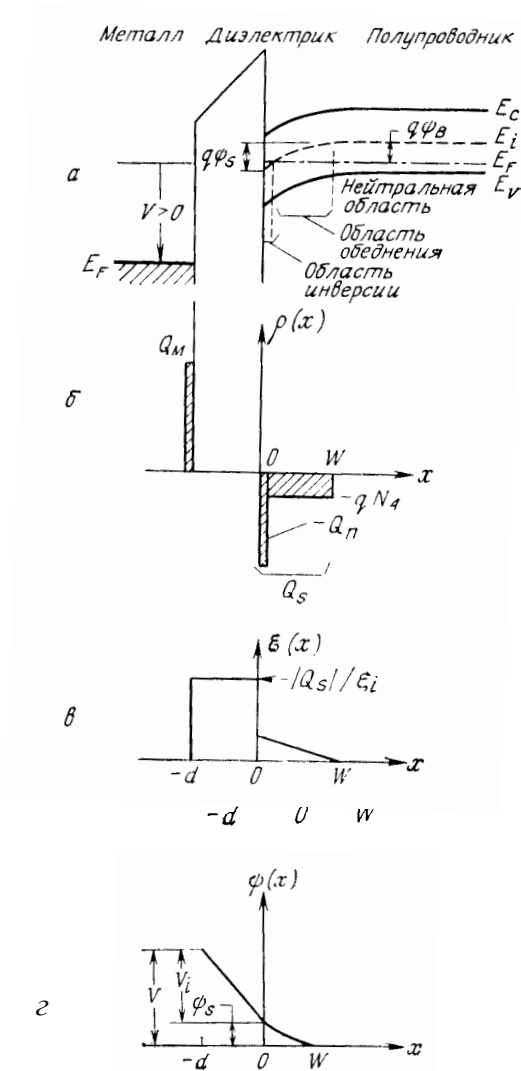


Рис. 3.7. Зонная диаграмма идеальной МДП-структуры (а), и распределение зарядов (в условиях инверсии) (б), электрического поля (в) и потенциала (г)

ОПЗ, образованная отрицательными ионами акцепторной примеси, толщиной W . Плотность заряда в ОПЗ, отнесенная к единице объема $\rho = -qN_a$, уравнение Пуассона принимает вид

$$\frac{d^2\varphi}{dx^2} = \frac{qN_a}{\varepsilon_0\varepsilon_s} \quad (3.21)$$

Интегрируя (3.21) с граничными условиями $\varphi = 0$ и $d\varphi/dx = 0$, при $x = W$, получаем

$$E(x) = -\frac{d\varphi}{dx} = \frac{qN_a(W-x)}{\varepsilon_0\varepsilon_s} = \frac{qN_aW}{\varepsilon_0\varepsilon_s} \left(1 - \frac{x}{W}\right) = E_s \left(1 - \frac{x}{W}\right) \quad (3.22)$$

В отсутствии разности работ выхода приложенное напряжение делится между полупроводником и диэлектриком, то есть $V = V_i + \varphi(x)$, где V_i – падение напряжения на слое диэлектрика; $\varphi(x)$ – падение напряжения в ОПЗ полупроводника. Решение уравнения Пуассона дает

$$\varphi(x) = \frac{qN_a(W-x)^2}{2\varepsilon_0\varepsilon_s} = \frac{qN_aW^2}{2\varepsilon_0\varepsilon_s} \left(1 - \frac{x}{W}\right)^2 = \varphi_s \left(1 - \frac{x}{W}\right)^2, \quad (3.23)$$

где φ_s – поверхностный потенциал на границе раздела полупроводник - диэлектрик.

Падение напряжения на слое диэлектрика

$$V_i = E_i \cdot d = \frac{Q_M \cdot d}{\varepsilon_i \varepsilon_0}, \quad (3.24)$$

Но с другой стороны сумма зарядов взятых на единицу площади на управляющем электроде (металле) Q_M и индуди-

рованного пространственного заряда Q_s равна нулю $Q_M + Q_s = 0$. Удельная емкость диэлектрика $C_i = \varepsilon_i \varepsilon_0 / d$, тогда

$$V_i = \frac{Q_M}{C_i} = -\frac{Q_s}{C_i} \quad (3.25)$$

Таким образом, напряжение на металлическом электроде (затворе)

$$V_i = -\frac{Q_s}{C_i} + \varphi_s. \quad (3.26)$$

Определим пороговое напряжение считая, что в инверсном слое отсутствуют подвижные носители заряда

$$V_{\text{пд}} = \varphi_s - \frac{Q_A}{C_i} = 2\varphi_0 - \frac{Q_A}{C_i}. \quad (3.27)$$

Пороговое напряжение – это напряжение на затворе при котором начинается режим сильной инверсии.

3.5. Вольт-фарадные характеристики идеальной МДП-структуры

3.5.1. Емкость области пространственного заряда

Поскольку полный заряд в ОПЗ Q_{sc} зависит от величины поверхностного потенциала ψ_s , то область пространственного заряда обладает определенной емкостью C_{sc} .

Величина C_{sc} , как следует из соотношения (3.17), будет равна

$$C_{sc} \equiv \frac{\partial Q_{sc}}{\partial \psi_s} = \frac{\varepsilon_s \varepsilon_0}{\sqrt{2} L_D} \frac{\left[(1 - e^{-\beta \psi_s}) + e^{-2\beta \varphi_0} (e^{\beta \psi_s} - 1) \right]}{F(\psi_s, \varphi_0)}. \quad (3.28)$$

Для того чтобы получить выражения для емкости ОПЗ в различных случаях (обеднение, обогащение, инверсия), можно либо непосредственно воспользоваться (3.28). Напомним, что рассматривается полупроводник p -типа.

Область обогащения ($\psi_s < 0$). Емкость ОПЗ C_{sc} обусловлена емкостью свободных дырок C_p :

$$C_{sc} = C_p = \frac{\varepsilon_s \varepsilon_0}{L_D} e^{-\frac{\beta \psi_s}{2}}. \quad (3.29)$$

Область обеднения и слабой инверсии ($2\varphi_0 > \psi_s > 0$). Емкость ОПЗ C_{sc} обусловлена емкостью области ионизованных акцепторов C_B :

$$C_{sc} = C_B = \frac{\varepsilon_s \varepsilon_0 q N_A}{\sqrt{2 \left(\psi_s - \frac{kT}{q} \right)}} = \frac{\varepsilon_s \varepsilon_0}{W}. \quad (3.30)$$

Из соотношения (3.30) следует, что емкость C_{sc} в области обеднения слабо зависит от поверхностного потенциала ψ_s , убывая с ростом последнего. Минимальное значение емкости C_{sc} достигается вблизи порогового значения поверхностного потенциала.

Емкость ОПЗ в области обеднения и слабой инверсии эквивалентна емкости плоского конденсатора, заполненного диэлектриком с относительной диэлектрической проницаемостью ε_s , пластины которого находятся друг от друга на расстоянии W равном ширине ОПЗ.

Плоские зоны ($\psi_s = 0$). Соотношения (3.29) и (3.30) несправедливы при $\psi_s \rightarrow 0$, то есть в области плоских зон у поверхности полупроводника. Непосредственная подстановка $\psi_s = 0$ в выражение (3.29) приводит к неопределенности типа «ноль делить на ноль».

Для расчета емкости плоских зон C_{FB} необходимо провести разложение экспоненты в (3.29) в ряд и после предельных переходов имеем

$$C_{\text{sc}} = C_{\text{FB}} = \frac{\varepsilon_s \varepsilon_0}{L_D} = \sqrt{\frac{\varepsilon_s \varepsilon_0 q N_A}{kT/q}}. \quad (3.31)$$

Емкость ОПЗ в плоских зонах эквивалентна емкости плоского конденсатора с обкладками, удаленными на дебаевскую длину экранирования.

Область сильной инверсии ($\psi_s > 2\varphi_0$). Емкость ОПЗ C_{sc} обусловлена емкостью свободных электронов C_n в инверсионном слое и при достаточно больших значениях поверхностного потенциала $\beta(\psi_s - 2\varphi_0) \geq 7$ будет равна

$$C_n = \frac{\varepsilon_s \varepsilon_0}{\sqrt{2}L_D} e^{\frac{\beta(\psi_s - 2\varphi_0)}{2}}. \quad (3.32)$$

Из анализа (3.29) и (3.32) следует, что емкости свободных носителей в обогащении и сильной инверсии экспоненциально зависят от поверхностного потенциала ψ_s и имеют одинаковые значения, если величину поверхностного потенциала отсчитывать для инверсии от порогового значения $\psi_s = 2\varphi_0$.

На рис. 3.8 приведен график зависимости емкости ОПЗ C_{sc} от величины поверхностного потенциала ψ_s , рассчитанной по соотношениям (3.29 – 3.32).

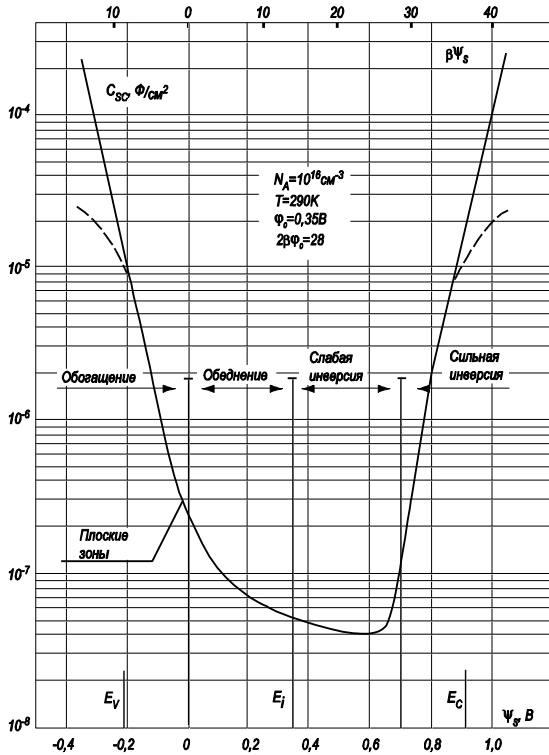


Рис. 3.8. Зависимость емкости области пространственного заряда C_{sc} от поверхностного потенциала, рассчитанная в классическом (сплошная линия) и вырожденном (пунктирная линия) случае

3.5.2. Емкость МДП-структур

Одним из наиболее распространенных методов изучения свойств структур металл – диэлектрик – полупроводник является метод, основанный на анализе зависимости емкости МДП-структуры $C_{МДП}$ от напряжения на затворе V_G , так называемый метод вольт-фардных характеристик (ВФХ) или $C-V$

метод. Для использования этого метода рассмотрим подробно теорию емкости МДП-структур.

Емкость МДП-структуры представляет собой последовательное соединение емкости диэлектрика C_i и приповерхностной емкости полупроводника C_s

$$\frac{1}{C_{i\bar{i}\bar{i}}} = \frac{1}{\tilde{N}_i} + \frac{1}{C_s}, \quad C_{i\bar{i}\bar{i}} = \frac{C_i C_s}{C_i + C_s}. \quad (3.33)$$

График зависимости емкости МДП-структуры от напряжения на затворе, так называемая вольт-фарадная характеристика МДП-структуры, приведен на рис. 3.9.

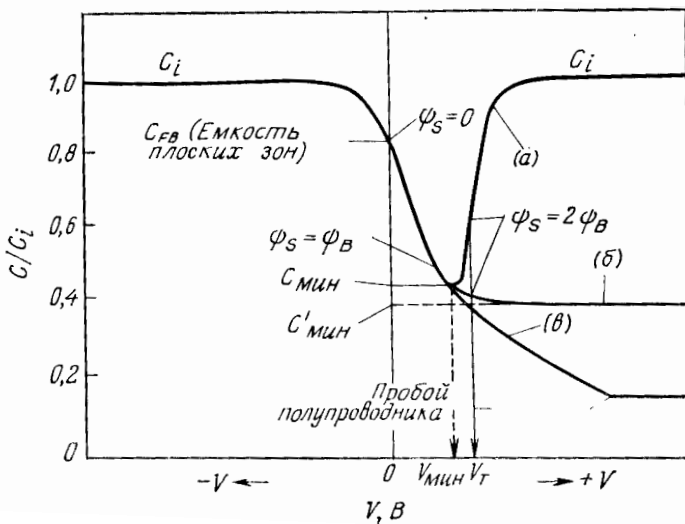


Рис. 3.9. Вольт-фарадные характеристики идеальной МДП-структуры: *а* – в области низких частот; *б* – в области высоких частот; *в* – в режиме глубокого обеднения

Если $V < 0$ и полупроводниковая подложка имеет p -тип проводимости, то около поверхности полупроводника накапливаются дырки (режим аккумуляции, обогащения). В этом режиме дифференциальная емкость полупроводника существенно больше емкости диэлектрика, поэтому полная емкость структуры близка к величине емкости диэлектрика:

$$C_i = \frac{\varepsilon_0 \varepsilon_i}{d}. \quad (3.34)$$

Когда напряжение на затворе меняется в направлении напряжения плоских зон, поверхностное обогащение исчезает и вследствие роста длины Дебая у поверхности емкость начинает уменьшаться.

При переходе МДП-структуры в режим обеднения ($V > 0$) дырки удаляются от поверхности, образуется ОПЗ и начинает сказываться влияние приповерхностной емкости C_s , значение которой зависит от напряжения на затворе. Обедненная область действует как добавочный слой диэлектрика. В этом режиме полная емкость МДП-структуры состоит из последовательно соединенных емкостей диэлектрика и приповерхностной обедненной области полупроводника

$$C_{i\bar{i}\bar{i}} = \left(\frac{1}{\tilde{N}_i} + \frac{W}{\varepsilon_s \varepsilon_0} \right)^{-1}, \quad (3.35)$$

где W – ширина приповерхностного обедненного слоя, которая зависит как от напряжения на затворе, так и концентрации примеси.

Из выражения (3.35) следует, что с увеличением ширины обедненной области полная емкость МДП-структуры уменьшается.

При превышении V значения порогового напряжения в МДП-структуре происходит инверсия проводимости приповерхностного слоя: поверхностная концентрация электронов в инверсионном слое растет экспоненциально с напряжением, а поверхностный потенциал увеличивается пропорционально квадрату толщины обедненной области. После того как значение W достигнет максимальной величины, дальнейшее приращение положительного заряда на затворе будет компенсироваться возрастанием концентрации электронов в канале. Появление избыточных электронов обеспечивается достаточно медленной генерацией электронно-дырочных пар в ОПЗ. Поэтому, если к постоянному напряжению на затворе V добавляется малое переменное напряжение dV с высокой частотой, то концентрация носителей в инверсионном слое (электронов) не успевает изменяться с частотой переменного напряжения и емкость МДП-структуры остается постоянной (рис. 3.9, кривая б).

Если же частота изменений V низкая, то изменение концентрации носителей в инверсионном слое (электронов) успевает следовать за изменением напряжения. Дифференциальная емкость инверсионного слоя значительно превышает емкость диэлектрика, поэтому полная емкость МДП-структуры в режиме инверсии резко возрастает, снова приближаясь к емкости диэлектрика (рис. 3.9, кривая а).

Рассмотрим теперь, как меняется емкость МДП-структуры при подаче на затвор импульса напряжения. Пусть в исходном состоянии напряжение на затворе равнялось нулю. Если теперь к затвору «скачком» прикладывается положительное напряжение превышающее пороговое, то дырки удаляются от поверхности полупроводника и образуется ОПЗ. Однако инверсия проводимости не происходит, так как инерционный процесс генерации электронов не успевает за изменением напряжения. Вследствие того, что инверсионный слой отсутствует, положительный заряд на металлическом затворе

может быть скомпенсирован только ионами акцепторов в ОПЗ. Поэтому толщина ОПЗ увеличивается. Удельная емкость структуры в этом режиме по аналогии с (3.35) равна:

$$C_{iAI} = \left(\frac{1}{\tilde{N}_i} + \frac{W_{\max} + \Delta W}{\varepsilon_s \varepsilon_0} \right)^{-1}, \quad (3.36)$$

Фактически в выражении (3.36) отражено сохранение режима обеднения и при напряжениях на затворе, больших порогового. Толщина обедненной области под увеличивается на ΔW , а значение емкости C_s и общей емкости МДП-структуры уменьшается (рис. 3.9, кривая ν), при этом зависимость емкости затвора от напряжения на затворе аналогична зависимости барьерной емкости обратно-смещенного p - n -перехода от обратного напряжения.

3.6. Компоненты заряда в реальном диоксиде кремния и их влияние на ВФХ МДП-структуры

Из всех МДП-структур наиболее важными являются структуры металл - SiO_2 - Si (МОП). МОП-структура создается в результате высокотемпературного окисления Si. Слой SiO_2 , возникающий при термическом окислении кремния, следует рассматривать как некоторый переходный слой с переменным химическим составом. Предполагается, что на границе монокристаллического кремния находится моноатомный слой нестехиометрического SiO_x ($1 < x < 2$; стехиометрическому диоксиду кремния соответствует $x = 2$), представляющий собой неполностью окисленный кремний. Затем следует промежуточный слой SiO_2 с большими внутренними механическими напряжениями толщиной 10 - 40 Å, который переходит в обычный ненапряженный стехиометрический аморфный SiO_2 (рис. 3.10).

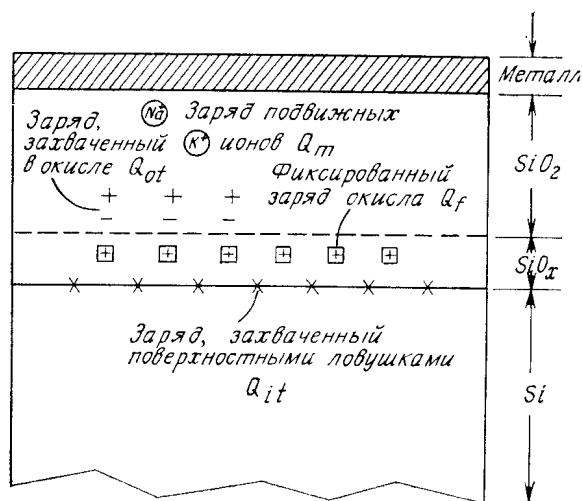


Рис. 3.10. Классификация зарядов, присутствующих в термически окисленном кремнии

Отличие характеристик реальных МОП-структур от соответствующих зависимостей идеальных МДП-структур обусловлено рядом факторов, к их числу относятся:

- 1) возникновение на границе раздела SiO₂ - Si поверхностных ловушек и зарядов в оксиде;
- 2) различие работ выхода алюминия и кремния;
- 3) образование подвижных ионов, которые могут внедряться в кристаллическую решетку SiO₂, в ходе процесса оксидирования.

В настоящее время принята следующая классификация этих зарядов и ловушек (рис. 3.10):

1. Заряд, захваченный поверхностными ловушками Q_{it} , представляющий собой заряд электронных состояний, которые локализованы на границе раздела Si - SiO₂ и энергия которых лежит в глубине запрещенной зоны полупроводника.

Эти поверхностные состояния, называемые также быстрыми (а иногда пограничными), могут достаточно быстро перезаряжаться, обмениваясь электронами (дырками) с кремнием. Поверхностные состояния, вероятно, обусловлены избыточными атомами кремния (трехвалентным кремнием), избыточным кислородом или примесными атомами.

2. Фиксированный заряд окисла Q_f , расположенный на границе раздела или в непосредственной близости от нее. Величина этого заряда остается практически постоянной во всей области электрических полей, характерных для рабочего диапазона напряжений на МОП-структурах.

3. Заряд, захваченный в окисле Q_{ot} . Этот заряд возникает, например, при рентгеновском облучении структур или инжекции горячих электронов в диэлектрик. Соответствующие ловушки более или менее равномерно распределены по толщине слоя окисла.

4. Заряд подвижных ионов Q_m (например, ионов натрия, калия, то есть щелочных металлов, которые легко абсорбируется диоксидом кремния), который может перемещаться в слое окисла при стрессовых термополевых нагрузках МДП-структур.

Величину всех этих зарядов обычно относят к единице площади границы раздела, то есть измеряют в единицах Кл·см². Вместо Q_i часто используют соответствующие поверхностные плотности, которые обозначают символом N_i с теми же индексами ($N_i = Q_i/q$ - число зарядов на 1 см²). Поскольку энергетические уровни состояний, захватывающих поверхностный заряд Q_{it} , непрерывно распределены в запрещенной зоне полупроводника, полезной характеристикой является энергетическая плотность поверхностных состояний

$$D_{it} = \frac{1}{q} \frac{dQ_{it}}{dE}, \quad [\text{число зарядов}/(\text{см}^2 \cdot \text{эВ})]. \quad (3.37)$$

Поверхностные состояния и заряд в оксиде влияют на форму ВФХ, сдвигая и растягивая ее вдоль оси напряжений (рис. 3.11). Положение реальных ВФХ характеризуется так называемым сдвигом напряжений плоских зон V_{FB} , определяемым по отношению к ВФХ идеальной МДП-структуры.

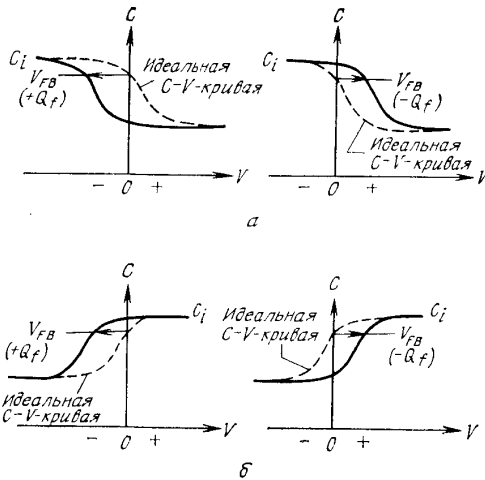


Рис. 3.11. Сдвиг ВФХ вдоль оси напряжений, обусловленный положительным или отрицательным фиксированным зарядом оксида: *а* – для полупроводника *p*-типа; *б* – для полупроводника *n*-типа

3.7. Распределение плотности пространственного заряда, электрического поля и потенциала в реальной МДП-структуре в режиме сильной инверсии

На рис. 3.12 приведена зонная диаграмма и распределение заряда электрического поля и потенциала в МОП-структуре, содержащей как фиксированный заряд, так и заряд, захваченный в окисле. Из сравнения рис. 3.12 и 3.7 следует,

что в первом случае те же значения поверхностного потенциала достигаются при меньших напряжениях смещения.

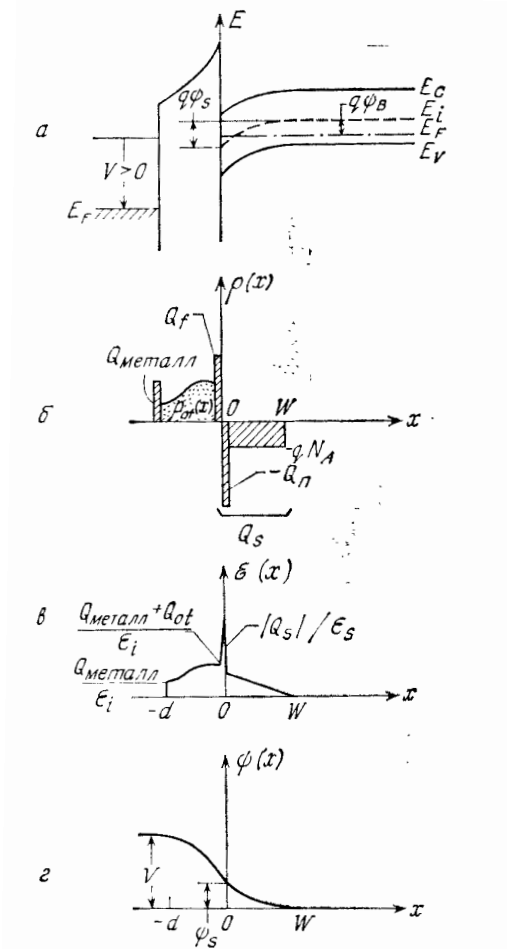


Рис. 3.12. МДП-структура с фиксированным и захваченным в окисле зарядами: *а* – зонная диаграмма; *б* – распределение заряда; *в* – электрическое поле; *г* – потенциала

Рассмотрим более подробно связь между напряжением на затворе V_G МДП-структуры и поверхностным потенциалом ψ_s . Все приложенное напряжение V_G к МДП-структуре делится между диэлектриком и полупроводником, причем очевидно, что падение напряжения в полупроводнике равняется поверхностному потенциалу ψ_s .

Таким образом,

$$V_G = V_{ox} + \psi_s. \quad (3.38)$$

Из (3.38) и анализа зонных энергетических диаграмм на рис. 3.3 следует, что знак поверхностного потенциала ψ_s в действительности соответствует знаку напряжения на затворе V_G . Действительно, положительное напряжение на затворе идеальной МДП-структуры вызывает изгиб зон вниз у полупроводников n - и p -типа, что соответствует положительным значениям поверхностного потенциала. Отрицательное напряжение V_G вызывает изгиб зон вверх у поверхности полупроводника, что соответствует отрицательному значению поверхностного потенциала ψ_s .

Из условия электронейтральности следует, что заряд на металлическом электроде Q_M должен быть равен суммарному заряду в ОПЗ Q_{sc} , заряду поверхностных состояний на границе раздела полупроводник-диэлектрик Q_{ss} и встроенному заряду в диэлектрик вблизи границы раздела Q_{ox} . Тогда

$$-Q_M = Q_{sc} + Q_{ss} + Q_{ox}. \quad (3.39)$$

Согласно определению геометрической емкости диэлектрика C_{ox} ,

$$C_{ox} = \frac{Q_M}{V_{ox}}, \quad (3.40)$$

отсюда

$$V_{\text{ox}} = \frac{Q_{\text{M}}}{C_{\text{ox}}} = -\frac{Q_{\text{sc}}}{C_{\text{ox}}} - \frac{Q_{\text{ss}}}{C_{\text{ox}}} - \frac{Q_{\text{ox}}}{C_{\text{ox}}}. \quad (3.41)$$

Учитывая, что между металлом и полупроводником существует разность термодинамических работ выхода $\Delta\varphi_{\text{ms}}$, получаем

$$V_{\text{G}} = \Delta\varphi_{\text{ms}} + \psi_{\text{s}} - \frac{Q_{\text{sc}}}{C_{\text{ox}}} - \frac{Q_{\text{ss}}}{C_{\text{ox}}} - \frac{Q_{\text{ox}}}{C_{\text{ox}}}. \quad (3.42)$$

Из соотношения (3.42) следует, что если $V_{\text{G}} > 0$, то $\psi_{\text{s}} > 0$, величины $Q_{\text{sc}} < 0$, $Q_{\text{ss}} < 0$, то есть падение напряжения на диэлектрик $V_{\text{ox}} > 0$. Аналогично будет соотношение знаков и при $V_{\text{G}} < 0$. Поскольку

$$Q_{\text{ss}} = -qN_{\text{ss}}(\psi_{\text{s}} - \varphi_0), \quad (3.43)$$

То подставив (3.43) в (3.42), имеем

$$V_{\text{G}} = \Delta\varphi_{\text{ms}} - \frac{Q_{\text{ox}}}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}}\varphi_0 + \psi_{\text{s}} - \frac{Q_{\text{sc}}}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}}\psi_{\text{s}}. \quad (3.44)$$

Введем понятие напряжения плоских зон V_{FB} (Flat Band) для реальной МДП-структуры. **Напряжением плоских зон** V_{FB} называется напряжение на затворе реальной МДП-структуры, соответствующее значению поверхностного потенциала в полупроводнике, равному нулю:

$$V_{\text{FB}} \equiv V_{\text{G}}(\psi_{\text{s}} = 0). \quad (3.45)$$

С учетом определения (3.45) из (3.44) следует

$$V_{\text{FB}} = \Delta\varphi_{\text{ms}} - \frac{Q_{\text{ox}}}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}} \varphi_0. \quad (3.46)$$

Таким образом, связь между напряжением на затворе V_G и поверхностным потенциалом ψ_s с учетом (3.46) задается в виде

$$V_G = V_{\text{FB}} + \psi_s + \frac{qN_{\text{ss}}}{C_{\text{ox}}} \psi_s - \frac{Q_{\text{sc}}}{C_{\text{ox}}}. \quad (3.47)$$

Введем *пороговое напряжение* V_T как напряжение на затворе V_G , когда в равновесных условиях поверхностный потенциал ψ_s равен пороговому значению $2\varphi_0$:

$$V_T \equiv V_{G(\psi_s=2\varphi_0)}. \quad (3.48)$$

Из (3.46) - (3.48) следует, что

$$V_T = \Delta\varphi_{\text{ms}} + 2\varphi_0 - \frac{Q_{\text{ox}}}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}} 2\varphi_0 - \frac{Q_B}{C_{\text{ox}}}, \quad (3.49)$$

или с учетом определения V_{FB} :

$$V_T = V_{\text{FB}} + 2\varphi_0 - \frac{Q_B}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}} 2\varphi_0. \quad (3.50)$$

Из (3.50) следует, что если отсчитывать пороговое напряжение V_T от напряжения плоских зон V_{FB} , то оно будет состоять из падения напряжения в полупроводнике $2\varphi_0$ и падения напряжения на подзатворном диэлектрике за счет заряда ионизованных акцепторов и заряда в поверхностных состояниях.

3.8. Определение параметров МДП-структур на основе анализа C - V характеристик

Анализ вольт-фарадных характеристик позволяет получить обширную информацию об основных параметрах МДП-структур: типе проводимости полупроводниковой подложки (n - или p -тип); концентрации легирующей примеси в подложке и законе ее распределения в приповерхностной области полупроводника; величине и знаке встроенного в диэлектрик МДП-структуры заряда; толщине подзатворного окисла; плотности поверхностных состояний на границе раздела полупроводник – диэлектрик. Рассмотрим более подробно эти вопросы.

Определение типа проводимости полупроводниковой подложки. Для определения типа проводимости подложки воспользуемся высокочастотной вольт-фарадной характеристикой. Как следует из вида высокочастотной C - V кривой при обогащении основными носителями, емкость МДП-структуры максимальна и определяется емкостью диэлектрика. В инверсии же емкость МДП-структуры минимальна. Таким образом, если максимум емкости C - V кривой лежит в более положительных напряжениях, чем минимум, то подложка изготовлена из полупроводника n -типа, если же максимум C - V кривой находится в более отрицательных напряжениях, то подложка изготовлена из полупроводника p -типа. На рис. 3.13 приведены для примера высокочастотные ВФХ на n - и p -типах подложки.

Определение толщины подзатворного диэлектрика. Поскольку, как было показано ранее, в обогащении емкость МДП-структуры определяется только геометрической емкостью диэлектрика $C_{\text{ок}}$, то

$$C = C_{\text{ок}} = \frac{\varepsilon_{\text{ок}} \varepsilon}{d_{\text{ок}}}, \quad (3.51)$$

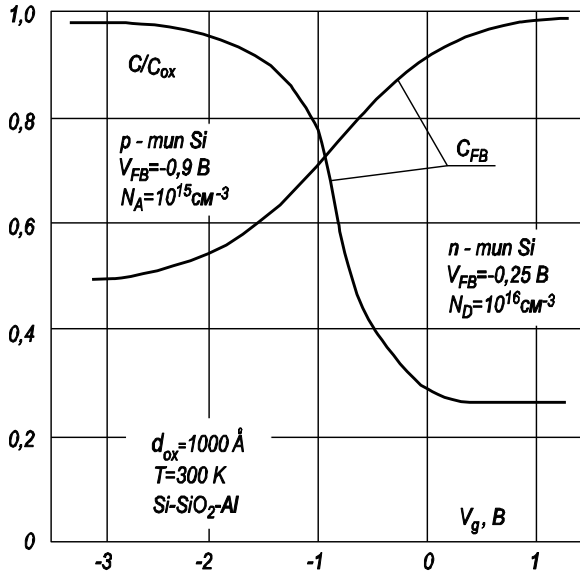


Рис. 3.13. Высоочастотные ВАХ МДП-структур, изготовленных на полупроводниковых подложках n - и p -типа

где ε_{ox} – относительная диэлектрическая проницаемость оксида.

Отсюда следует, что

$$d_{\text{ox}} = \frac{\varepsilon \varepsilon_{\text{ox}}}{C_{\text{ox}}}, \quad (3.52)$$

где C_{ox} – удельная емкость подзатворного диэлектрика, то есть емкость на единицу площади.

Для подстановки в (3.52) экспериментальных значений необходимо сначала пронормировать емкость, то есть разделить экспериментальное значение емкости на площадь S МДП-структуры. При напряжениях на затворе $V_G - V_{\text{FB}} \approx (2 - 3)$ В практически для всех МДП-структур полная емкость C

только на 2 - 3 % отличается от емкости диэлектрика. Исключение составляют структуры со сверхтонким окислом $d_{\text{ox}} < 100 \text{ \AA}$, у которых в этой области V_G становится существенным квантование в ОПЗ, и это отличие может достигать 10 %.

Определение величины и профиля концентрации легирующей примеси. Для определения величины концентрации легирующей примеси воспользуемся следующим свойством высокочастотных $C - V$ характеристик МДП-структур: их емкость в области инверсии достигает минимальной величины C_{min} и определяется только емкостью области ионизованных доноров C_B и емкостью диэлектрика C_{ox} . При этом

$$\frac{1}{C_{\text{min}}} = \frac{1}{C_{\text{ox}}} + \frac{1}{C_B}. \quad (3.53)$$

Используя для емкости окисла C_{ox} выражение (3.51) и для емкости области ионизованных акцепторов из (3.30), получаем

$$C_{\text{min}} = \frac{\varepsilon_{\text{ox}} \varepsilon_0}{d_{\text{ox}} + \frac{\varepsilon_s - W}{\varepsilon_{\text{ox}}}}. \quad (3.54)$$

Выражение (3.54) совместно с (3.30) для емкости ОПЗ ионизованных акцепторов приводит к выражению для концентрации легирующей примеси:

$$N_A = \frac{2 \left(2\varphi_0 - \frac{kT}{q} \right)}{\varepsilon_s \varepsilon_0 q} \cdot \left(\frac{\frac{C_{\text{ox}}}{C_{\text{min}}} - 1}{\frac{C_{\text{ox}}}{C_{\text{min}}}} \right)^{-2}. \quad (3.55)$$

На рис. 3.14 приведена номограмма зависимости нормированной величины емкости C_{\min}/C_{ox} от толщины d_{ox} для систем Si - SiO₂ с концентрацией легирующей примеси N_A в качестве параметра.

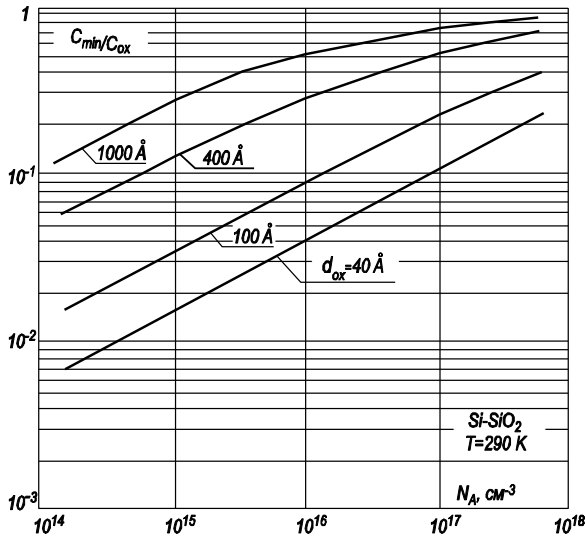


Рис. 3.14. Зависимость нормированной величины емкости C_{\min}/C_{ox} в минимуме высокочастотной ВАХ от толщины подзатворного диэлектрика d_{ox} при различных величинах концентрации легирующей примеси для кремниевых МДП-структур

Из рис. 3.14 видно, что чем меньше толщина диэлектрика и ниже концентрация легирующей примеси, тем больше перепад емкости от минимального до максимального значений наблюдается на ВФХ. Для определения профиля концентрации N_A от расстояния вглубь полупроводника z воспользуемся высокочастотной C - V кривой, снятой в области неравновесного обеднения. Неравновесное обеднение возможно ре-

лизовать в том случае, когда период напряжения развертки меньше постоянной τ генерационного времени неосновных носителей в ОПЗ. В этом случае величина поверхностного потенциала может быть больше $\psi_s > 2\varphi_0$, а ширина ОПЗ соответственно больше, чем ширина ОПЗ в равновесном случае. Возьмем также МДП-структуру с достаточно тонким окислом, таким, чтобы падением напряжения на окисле V_{ox} можно было бы пренебречь по сравнению с величиной поверхностного потенциала, т.е. $V_{ox} \ll \psi_s$; $V_G \approx \psi_s$. В этом случае тангенс угла наклона зависимости $C^{-2} = f(V_G)$ определит величину концентрации N_A :

$$N_A(z) = \frac{2}{q\epsilon_s\epsilon_0} \left[\frac{d(C^{-2})}{dV_G} \right]^{-1} \quad (3.56)$$

Значение координаты z , которой соответствует рассчитанная величина N_A , определяется при подстановке значения $\psi_s = V_G$ в выражение для ширины ОПЗ:

$$z = \sqrt{\frac{2\epsilon_s\epsilon_0 V_G}{qN_A}}. \quad (3.57)$$

В предельном случае, когда толщина окисла $d_{ox} \rightarrow 0$, эту величину используют, измеряя неравновесную емкость как емкость барьеров Шоттки при обратном смещении.

Определение величины и знака встроенного заряда.

Для определения величины и знака встроенного в диэлектрик МДП-структуры заряда обычно пользуются высокочастотным методом ВФХ. Для этого, зная толщину подзатворного диэлектрика d_{ox} , концентрацию легирующей примеси N_A и работу выхода материала затвора, рассчитывают согласно (3.33) и (3.37) теоретическое значение емкости плоских зон C_{FB} МДП-

структуры и напряжения плоских зон $V_{\text{FB}} = \Delta\varphi_{\text{ms}}$. Поскольку экспериментальная C - V кривая высокочастотная, т.е. $C_{\text{ss}} \rightarrow 0$, то, проводя сечение $C = \text{const} = C_{\text{FB}}$ (теор.), мы получаем при пересечении этой кривой с экспериментальной ВФХ напряжение, соответствующее $\psi_s = 0$, т.е. экспериментальное напряжение плоских зон $V_{\text{FB}}(\text{эксп.})$. При этом, согласно (3.46),

$$V_{\text{FB}}(\text{эксп.}) - V_{\text{FB}}(\text{теор.}) = -\frac{Q_{\text{ox}}}{C_{\text{ox}}} + \frac{qN_{\text{ss}}}{C_{\text{ox}}}\varphi_0. \quad (3.58)$$

Если $Q_{\text{ox}}, Q_{\text{ss}} > 0$, то $V_{\text{FB}}(\text{эксп.}) > V_{\text{FB}}(\text{теор.})$, и наоборот, если $Q_{\text{ox}}, Q_{\text{ss}} < 0$, то $V_{\text{FB}}(\text{эксп.}) < V_{\text{FB}}(\text{теор.})$.

Таким образом, знак и величина суммарного заряда в плоских зонах определяются соотношением (3.58) однозначно. Для вычленения заряда в поверхностных состояниях воспользуемся тем, что он обусловлен основными носителями (p -тип, $Q_{\text{ss}}(\psi_s = 0) > 0$ и n -тип, $Q_{\text{ss}}(\psi_s = 0) < 0$), захваченными на поверхностные состояния. Зная величину N_{ss} , можно рассчитать величину заряда в поверхностных состояниях Q_{ss} и таким образом из (3.46) определить величину и знак встроенного в диэлектрик заряда Q_{ox} .

4. ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

Полупроводниковым диодом называют нелинейный электронный прибор с двумя выводами. В зависимости от внутренней структуры, типа, количества и уровня легирования внутренних элементов диода и вольт-амперной характеристики свойства полупроводниковых диодов бывают различными.

Полупроводниковые диоды могут отличаться друг от друга, например, по следующим признакам:

а) по применяемым исходным материалам (германиевые, кремниевые, на арсениде галлия и др.);

б) по структуре перехода (точечные и плоскостные);

в) по технологии изготовления (сплавные, диффузионные, эпитаксиальные и др.);

г) по функциональному назначению (выпрямительные, детекторные, импульсные и т. д.);

д) по частотному диапазону (низкочастотные, высокочастотные и сверхвысокочастотные);

е) по мощности рассеяния (маломощные, средней мощности и мощные);

ж) по конструктивному оформлению (стеклянные, металлические, металлокерамические, пластмассовые).

4.1. Методы изготовления полупроводниковых диодов

В *точечных диодах* p - n -переходы образуются точечно-контактным способом. Например, к пластине германия, напыленной с помощью олова на кристаллодержатель, подводят и прижимают заостренную иглу из бериллиевой бронзы (рис. 4.1). Диаметр острия порядка 20 – 50 мкм. В месте соприкосновения иглы с полупроводником образуется выпрямляющий переход. Для улучшения его свойств через контакт

иглы с германием пропускают мощные короткие импульсы тока, при этом конец иглы сплавляется с полупроводником, обеспечивая стабильность и механическую прочность контакта. Одновременно при повышенной температуре медь диффундирует внутрь германия, образуя под контактной иглой полусферическую область p -типа, так как медь в германии является акцепторной примесью. Для увеличения концентрации акцепторной примеси в p -области зачастую на конец иглы перед прижатием наносят индий или алюминий.

Таким образом, p - n -переход образуется в результате диффузии примеси из иглы и возникновения под иглой p -области в германии n -типа. Точечные диоды имеют очень маленькую емкость, так как площадь p - n -перехода небольшая, поэтому их используют главным образом при изготовлении диодов высокой и сверхвысокой частоты.

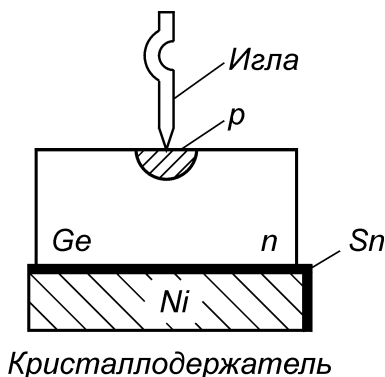


Рис. 4.1. Структура точечного диода

В *плоскостных диодах*, у которых линейные размеры перехода, определяющие его площадь, значительно больше толщины, p - n -переходы получают методами сплавления, диффузии, эпитаксии, ионной имплантации.

Сплавные переходы (отсюда и название сплавные диоды) получают вплавлением примеси в монокристалл полупроводника. Например, небольшая таблетка алюминия помещается на поверхность пластины кремния n -типа, имеющего ориентацию $\langle 111 \rangle$. Затем пластину с таблеткой нагревают до температуры несколько ниже температуры плавления алюминия, но достаточной для образования эвтектики Al–Si (приблизительно $580\text{ }^{\circ}\text{C}$), в результате чего образуется небольшая область расплава Al–Si. При остывании на границе расплава кристаллизуется тонкий слой кремния, сильно легированный алюминием, т.е. слой с проводимостью p -типа (p^+) на кристалле n -типа (рис. 4.2).

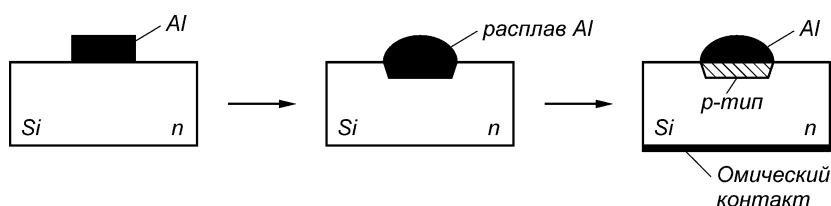


Рис. 4.2. Схема изготовления сплавного перехода

Верхняя часть алюминиевой таблетки используется в качестве омического контакта к области p -типа. На обратную поверхность кристалла напыляют сплав золото-сурьма, содержащий около 0,1 % сурьмы, и впаляют его при температуре около $400\text{ }^{\circ}\text{C}$ для создания невыпрямляющего омического контакта с кремнием n -типа.

Сплавные p - n -переходы получают несимметричными, т.е. $p_p \gg n_n$, поэтому у такого диода дырочная составляющая тока диффузии значительно больше электронной составляющей:

$$I_{\text{диф}} = I_p + I_n \approx I_p. \quad (4.1)$$

Область p - n -перехода с большой концентрацией основных носителей заряда называют *эмиттером*, а область с меньшей концентрацией основных носителей, в которую инжектируются неосновные носители для этой области носителей заряда, – *базой*.

Сплавные резкие переходы имеют значительно большую площадь, чем точечные, соответственно их емкость много больше.

При изготовлении плоскостного диффузионного перехода применяют различные методы, в основе каждого из них – диффузия примесного вещества (донорного или акцепторного) в исходную полупроводниковую пластину p -или n -типа (соответственно) при больших температурах (более 1000 °С). При этом концентрация введенной в поверхностный слой примеси уменьшается с глубиной, поэтому p - n -переход получается плавным.

Глубина диффузии примеси в кристалле зависит от температуры и времени проведения диффузии, поэтому ее легко контролировать.

С целью получения меньшего разброса параметров при изготовлении диодов в едином технологическом цикле методом дополнительного травления уменьшают площадь диффузионной области, т. е. создают так называемый меза-переход (в переводе с испанского выступ, столик) (рис. 4.3, *а*).

Планарные переходы получили свое название потому, что p - n -переходы диодных структур (также это относится и к транзисторным структурам) и контакты ко всем областям расположены на одной плоскости полупроводникового кристалла. Схематично последовательность операций при получении планарных переходов показана на рис. 4.3, *б*. Нагревая пластину кремния в потоке кислорода, получают на ее поверхно-

сти слой диоксида кремния необходимой толщины (обычно в пределах 0,5 – 1,2 мкм).

Затем методом фотолитографии селективно удаляют оксидный слой и в свободные от SiO_2 «окна» проводят диффузию примеси (например, атомов бора), получают p - n -переход. Применяя вновь метод фотолитографии, травлением удаляют SiO_2 с участков кремния n - и p -типов для создания омических контактов к этим областям методом напыления алюминия.

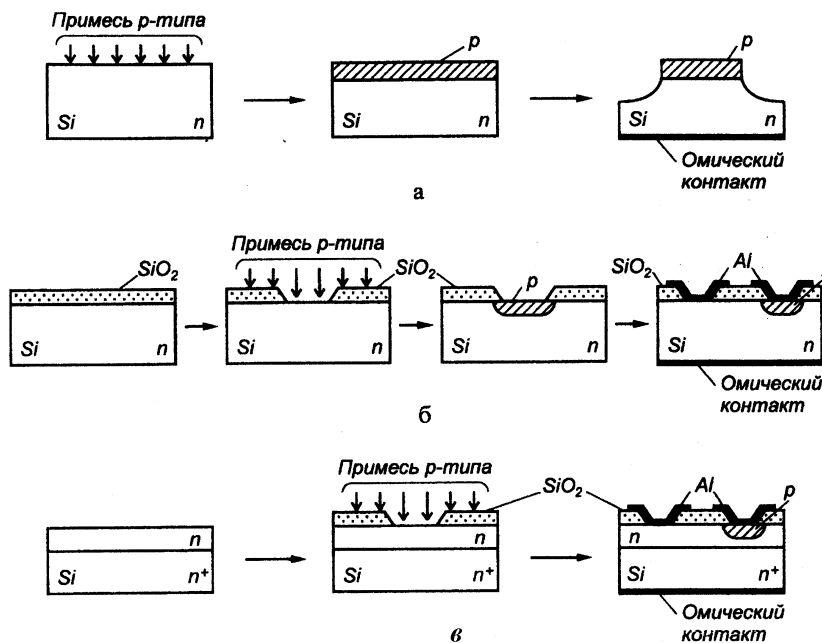


Рис. 4.3. Основные диффузионные методы изготовления p - n -переходов: а - диффузионный меза-переход; б - диффузионный планарный переход; в - диффузионный планарный переход на эпитаксиальной подложке

Эпитаксиальные структуры обычно используется в планарной технологии для уменьшения последовательного сопротивления. На поверхности сильно легированного низкоомного кремния выращивается слабо легированный высокоомный эпитаксиальный слой. И далее технология получения диффузионного перехода (рис. 4.3, в) аналогична предыдущему методу.

Методом эпитаксии можно получить плоскостный диод непосредственно без использования процесса диффузии. На полупроводниковую пластину кремния p -типа наращивают кристаллический слой, называемый эпитаксиальным с донорной примесью, в результате чего получают резкий p - n -переход. Наращивание эпитаксиального слоя проводится из паровой фазы на поверхность монокристаллического полупроводника, при этом кристаллографическая решетка эпитаксиальной пленки продолжает ориентацию решетки исходной пластины – подложки.

Эпитаксиальные диоды обычно имеют малое падение напряжения в открытом состоянии и высокое пробивное напряжение.

4.2. Выпрямительные диоды

Выпрямительными называют диоды, предназначенные для выпрямления переменного тока. Основу выпрямительного диода составляет обычный электронно-дырочный переход. Вольт-амперная характеристика такого диода, приведенная на рис. 4.4, а, имеет ярко выраженную нелинейность и описывается уравнением

$$J = J_s (e^{\beta V_G} - 1). \quad (4.2)$$

В прямом смещении ток диода инжекционный, большой по величине и представляет собой диффузионную компоненту

тока основных носителей. При обратном смещении ток диода маленький по величине и представляет собой дрейфовую компоненту тока неосновных носителей. В состоянии равновесия суммарный ток, обусловленный диффузионными и дрейфовыми токами электронов и дырок, равен нулю:

$$j_{pE} - j_{nD} + j_{nE} - j_{pD} = 0. \quad (4.3)$$

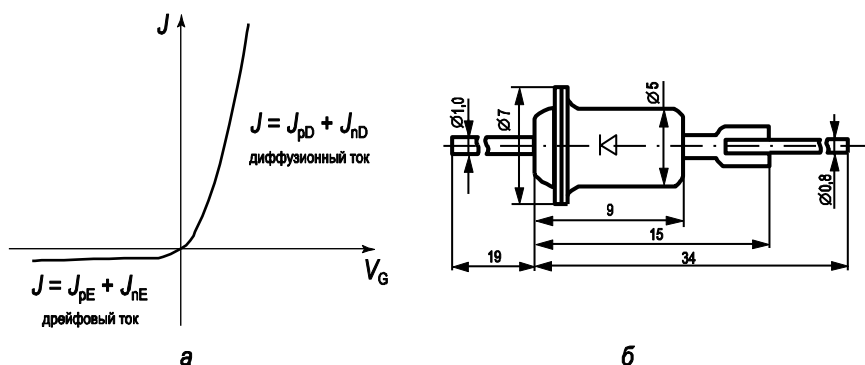


Рис. 4.4. Выпрямительный диод:
 а – вольт-амперная характеристика;
 б - конструкция корпуса

В качестве полупроводниковых материалов для выпрямительных диодов используют германий и кремний. В выпрямительных диодах используются плоскостные несимметричные $p-n$ -переходы, причем для германиевых диодов они создаются главным образом сплавным способом, а для кремниевых – сплавным и диффузионным. Для выпрямительных диодов характерно, что они имеют малые сопротивления в проводящем состоянии, поэтому позволяют пропускать большие токи. Так как площадь $p-n$ -перехода большая, то барьер-

ная емкость этих диодов будет большой, достигая значений десятков пикофарад.

Вольт-амперные характеристики германиевых и кремниевых диодов одинаковой конструкции различаются между собой. На рис. 4.5 для сравнения показаны характеристики германиевого (рис. 4.5, а) и кремниевого (рис. 4.5, б) диодов, имеющих одинаковую конструкцию и предназначенных для работы в одном и том же диапазоне токов и напряжений. Поскольку ширина запрещенной зоны у кремния больше, чем у германия, то обратный ток кремниевых диодов значительно меньше. Кроме того, обратная ветвь характеристики кремниевых диодов не имеет явно выраженного участка насыщения, что обусловлено генерацией носителей зарядов в *p-n*-переходе и токами утечки по поверхности кристалла.

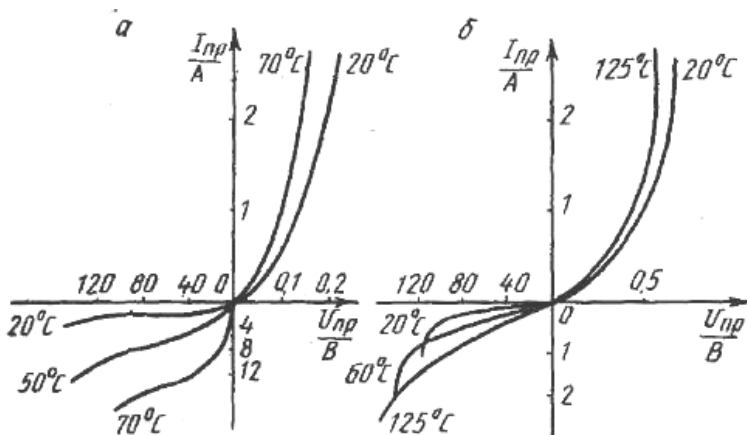


Рис. 4.5. Вольт-амперные характеристики выпрямительных диодов

Вследствие большого обратного тока у германиевых диодов наступает тепловой пробой, приводящий к разрушению кристалла. У кремниевых диодов из-за малого обратного тока

вероятность теплового пробоя мала, и у них возникает электрический пробой.

Поскольку прямой ток диода определяется из уравнения $I_{np} = I_s[\exp(U_{np}/\varphi T) - 1]$, вследствие меньшего значения обратного тока кремниевого диода его прямой ток, равный току германиевого диода, достигается при большем значении прямого напряжения. Поэтому мощность, рассеиваемая при одинаковых токах, в германиевых диодах меньше, чем в кремниевых.

На характеристики диодов существенное влияние оказывает температура окружающей среды. С ростом температуры становится интенсивнее генерация носителей зарядов и увеличиваются обратный и прямой токи диода.

Для приближенной оценки можно считать, что с увеличением температуры на 10 °С обратный ток германиевых диодов возрастает в два, а кремниевых – в два с половиной раза. Однако вследствие того, что при комнатной температуре обратный ток у германиевого диода значительно больше, чем у кремниевого, абсолютное значение приращения обратного тока у германиевого диода с ростом температуры оказывается в несколько раз больше, чем у кремниевого. Это приводит к увеличению потребляемой диодом мощности, его разогреву и уменьшению напряжения теплового пробоя. У кремниевых диодов из-за малого обратного тока вероятность теплового пробоя мала, и у них вначале возникает электрический пробой.

Пробой кремниевых диодов определяется процессами лавинного умножения носителей зарядов при ионизации атомов кристаллической решетки. С повышением температуры увеличивается тепловое рассеивание подвижных носителей зарядов и уменьшается длина их свободного пробега. Для того, чтобы электрон на меньшем пути приобрел энергию, достаточную для ионизации, необходимо увеличение ускоряющего поля, что достигается при большем обратном напряже-

нии. Это объясняет увеличение пробивного напряжения кремниевых диодов с ростом температуры.

Полупроводниковые выпрямительные диоды обычно характеризуются следующими основными параметрами:

- постоянный обратный ток $I_{обр}$ (мкА; мА) при некоторой величине постоянного обратного напряжения $U_{обр}$;
- постоянное прямое напряжение при постоянном прямом токе $I_{пр}$ (мА; А).

При серийном выпуске выпрямительных диодов в технических условиях на них указываются средние значения параметров: $I_{обр. ср.}$ при $U_{обр. ампл. макс.}$ и $U_{пр. ср.}$ при $I_{пр. ср. макс.}$.

В технических условиях на выпрямительные диоды приводятся предельно допустимые электрические режимы эксплуатации: максимально-допустимое амплитудное значение обратного напряжения $U_{обр. макс.}$ (В), средний прямой ток $I_{пр. ср.}$ (А).

Для анализа приборных характеристик выпрямительного диода важными также являются такие дифференциальные параметры, как коэффициент выпрямления, характеристические сопротивления и емкости диода в зависимости от выбора рабочей точки.

Различают два вида характеристического сопротивления диодов: дифференциальное сопротивление r_D и сопротивление по постоянному току R_D .

Дифференциальное сопротивление определяется как

$$r_D = \frac{dU}{dI} = \left[\frac{dI}{dU} \right]^{-1} = \beta j_s e^{\beta V} + j_s - j_s = [\beta(I + I_s)]^{-1} = \frac{kT/q}{I + I_s}. \quad (4.4)$$

На прямом участке вольт-амперной характеристики диода дифференциальное сопротивление r_D невелико и составляет значение несколько Ом. Действительно, при значении прямо-

го тока диода $I = 25$ мА и значении теплового потенциала $kT/q = 25$ мВ величина дифференциального сопротивления r_D будет равна $r_D = 1$ Ом. На обратном участке вольт-амперной характеристики диода дифференциальное сопротивление r_D стремится к бесконечности, поскольку в идеальных диодах при обратном смещении ток не зависит от напряжения.

Сопротивление по постоянному току R_D определяется как отношение приложенного напряжения V_G к протекающему току I через диод:

$$R_D = \frac{U}{I} = \frac{U}{I_0(e^{\beta U} - 1)}. \quad (4.5)$$

На прямом участке вольт-амперной характеристики сопротивление по постоянному току больше, чем дифференциальное сопротивление $R_D > r_D$, а на обратном участке – меньше $R_D < r_D$.

В точке вблизи нулевого значения напряжения $V_G \ll kT/q$ значения сопротивления по постоянному току и дифференциального сопротивления совпадают. Действительно, разложив экспоненту в ряд в соотношении (4.5), получаем

$$R_D = \frac{kT}{q} \frac{1}{I_0} = r_D. \quad (4.6)$$

Используя характерное значение для обратного тока диода $I_0 = 25$ мкА, получаем величину сопротивления диода в нулевой точке $R_{D0} = r_{D0} = 1$ кОм. На рис. 4.6, а приведена зависимость дифференциального сопротивления диода ГД402 от величины тока при прямом смещении.

С учетом полученных дифференциальных параметров можно построить эквивалентную малосигнальную схему диода для низких частот (рис. 4.6, в). В этом случае наряду с уже

описанными элементами – дифференциальным сопротивлением (рис. 4.6, *а*) и емкостями диода (рис. 4.6, *б*) - необходимо учесть омическое сопротивление квазинейтрального объема базы ($r_{об}$) диода. Сопротивление квазинейтрального объема эмиттера можно не учитывать, поскольку в диодах эмиттер обычно легирован существенно более сильно, чем база.

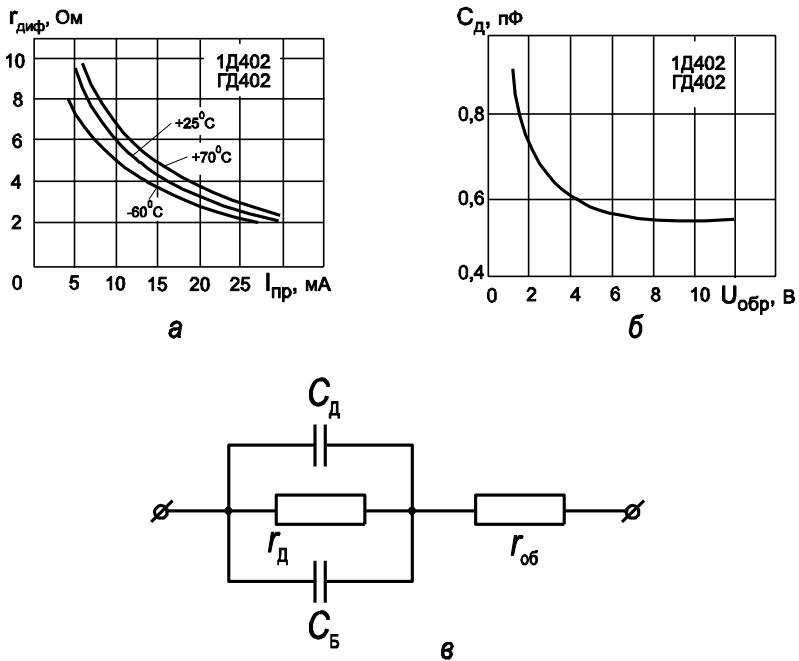


Рис. 4.6. Приборные характеристики и эквивалентная малосигнальная схема для выпрямительных диодов:
а - зависимость дифференциального сопротивления диода от величины тока при прямом смещении;
б - зависимость емкости диода от обратного напряжения;
в - эквивалентная малосигнальная схема диода для низких частот

На рис. 4.7 приведена схема, иллюстрирующая выпрямление переменного тока в диоде.

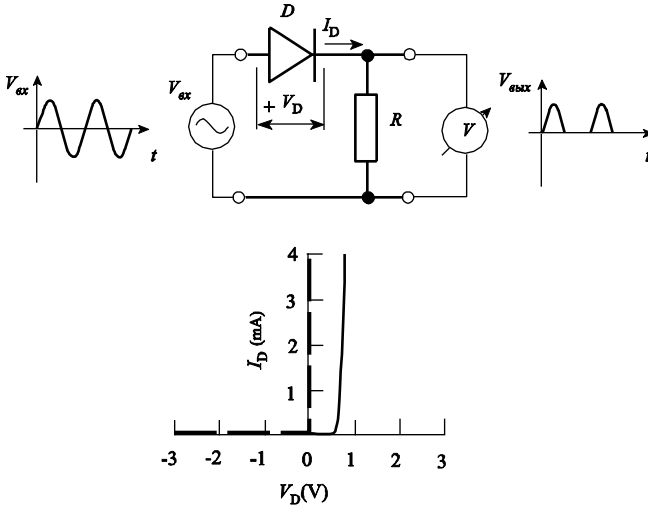


Рис. 4.7. Схема, иллюстрирующая выпрямление переменного тока с помощью диода

Рассмотрим, каков будет коэффициент выпрямления идеального диода на основе p - n -перехода. Для этого рассчитаем по уравнению (4.2) коэффициент выпрямления K как отношение прямого тока к обратному току диода. Получаем

$$K = \frac{J^+}{J^-} = \frac{e^{\beta V_G} - 1}{e^{-\beta V_G} - 1}. \quad (4.7)$$

Как следует из соотношения (4.7), при значениях переменного напряжения, модуль которого V_G меньше, чем тепловой потенциал kT/q , полупроводниковый диод не выпрямляет переменный ток. Коэффициент выпрямления достигает приемлемых величин при значениях V_G по крайней мере в 4 раза

больших, чем тепловой потенциал kT/q , что при комнатной температуре $T = 300$ К соответствует значению напряжения $V_G = \pm 0,1$ В.

4.3. Варикапы

Ранее указывалось, что ширина p - n -перехода и его емкость зависят от приложенного к нему напряжения.

Варикап – это полупроводниковый диод, который используется в качестве электрически управляемой емкости.

В варикапах используется свойство барьерной емкости обратно-смещенного p - n -перехода изменять свою величину в зависимости от приложенного к нему напряжения.

Барьерная емкость p - n -перехода равна емкости плоского конденсатора:

$$C_{\delta} = \frac{\varepsilon\varepsilon_0 S}{X_n}, \quad (4.8)$$

где S – площадь p - n -перехода; X_n – ширина области объемных зарядов.

С увеличением величины обратного напряжения $U_{обр}$ на p - n -переходе его барьерная емкость C_{δ} уменьшается. Характер изменения C_{δ} в зависимости от приложенного к p - n -переходу напряжения $U_{обр}$ показан на рис. 4.8. Функциональная зависимость емкости варикапа от напряжения определяется профилем легирования базы варикапа. В случае однородного легирования емкость обратно пропорциональна корню из приложенного напряжения. Задавая профиль легирования в базе варикапа $N_D(x)$, можно получить различные зависимости емкости варикапа от напряжения $C(U)$ – линейно убывающие, экспоненциально убывающие. Емкость варикапа меняется в широких пределах и его значение при обратном приложенном напряжении U определяют из выражения

$$C_{\sigma}(U) = C_{\sigma}(0) \left[\frac{U_{\kappa}}{U_{\kappa H} + U} \right]^{1/n}, \quad (4.9)$$

где $C_{\sigma}(0)$ – емкость при $U = 0$; φ_{κ} – значение контактной разности потенциалов равное

$$\varphi_{\kappa} = \frac{kT}{e} \ln \frac{n_p n_n}{n_i^2}, \quad (4.10)$$

$n = 2$ для резких переходов и $n = 3$ для плавных переходов.

Варикапы изготавливают на основе германия, кремния, арсенида галлия.

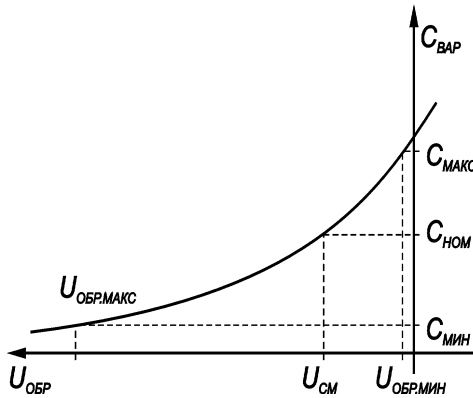


Рис. 4.8. Зависимость емкости варикапа от напряжения смещения

Основные параметры варикапа: номинальная (начальная) емкость $C_{\text{ном}}$; добротность Q_{σ} , коэффициент перекрытия по емкости K_C и температурный коэффициент емкости (ТКЕ) $\alpha_{C_{\sigma}}$.

Номинальная емкость варикапа $C_{ном}$ – барьерная емкость p - n -перехода при заданном напряжении смещения и составляет от долей пФ до сотен пФ.

Коэффициент перекрытия по емкости K_c – отношение емкости варикапа при двух заданных значениях обратных напряжений:

$$K_c = \frac{C_{макс}}{C_{мин}}. \quad (4.11)$$

$C_{макс}$ ограничивается емкостью при $U = 0$, т.е. $C(0)$. $C_{мин}$ ограничивается обратным допустимым напряжением. K_c позволяет определить величину изменения емкости в диапазоне рабочих напряжений от $U_{мин}$ до $U_{макс}$ ($U_{макс}$ по абсолютной величине может достигать 200 В). Характерные значения K_c составляют 2 – 20, причем если плавные и резкие p - n -переходы имеют обычно $K_c \leq 4$, то в случае сверхрезких p - n -переходов K_c имеют большие значения.

Добротность Q_v – отношение реактивного сопротивления варикапа на заданной частоте переменного сигнала X_c к сопротивлению потерь при заданном значении емкости или обратного напряжения. Q_v измеряют обычно при тех же обратных напряжениях, что и емкость варикапа. Как правило, варикапы работают в диапазоне высоких и сверхвысоких частот ($f > 1$ МГц), для которых

$$Q_v = \frac{1}{2\pi f C r}, \quad (4.12)$$

где r – последовательное (по отношению к C) сопротивление диода, включающее сопротивление потерь в объеме кристалла полупроводника, сопротивление контакта и элементов конструкции. В настоящее время достигнуты значения $Q_v > 500$

на

$f = 50$ МГц при $C = 70$ пФ.

Температурный коэффициент емкости (ТКЕ) $\alpha_{C\varepsilon}$ – отношение относительного изменения емкости к вызывающему его абсолютному изменению температуры окружающей среды, т.е., другими словами относительное изменение емкости при изменении температуры окружающей среды на 1 градус:

$$\alpha_{C\varepsilon} = \frac{\Delta C}{C\Delta T}. \quad (4.13)$$

Для кремниевых варикапов с резким p - n -переходом ТКЕ имеет значение порядка $5 \cdot 10^{-4}$ град $^{-1}$ при $|U| = 4$ В. С усилением зависимости емкости варикапа от напряжения, а также при понижении $U_{мин}$ ТКЕ возрастает и при $U = 1$ В может достигать $5 \cdot 10^{-3}$ град $^{-1}$.

В радиоэлектронных устройствах свойство нелинейного изменения емкости варикапа используют для получения параметрического усиления, умножения частоты и т.д., а возможность электрического управления емкостью – для дистанционной и безынерционной перестройки резонансной частоты колебательного контура.

4.4. Стабилитроны

Стабилитроном называется полупроводниковый диод, предназначенный для стабилизации уровня напряжения в схеме. Для этого используются приборы, у которых на вольт-амперной характеристике имеется участок со слабой зависимостью напряжения от проходящего тока. Такой участок обуславливается электрическим пробоем p - n -перехода при включении диода в обратном направлении. ВАХ стабилитрона имеет вид, представленный на рис. 4.9, а. Подобной вольт-

амперной характеристикой обладают сплавные диоды с базой, изготовленной из низкоомного полупроводникового материала. При этом напряженность электрического поля в p - n -переходе стабилитрона значительно выше, чем у обычных диодов. При относительно небольших обратных напряжениях в p - n -переходе возникает сильное электрическое поле, вызывающее его электрический пробой. Напряжение, при котором происходит лавинный пробой, зависит от удельного сопротивления полупроводникового материала. С ростом удельного сопротивления напряжение лавинного пробоя увеличивается.

Так как кремниевые диоды имеют меньшее значение обратного тока, обладают большей устойчивостью к тепловому пробую, чем германиевые диоды, поэтому в качестве стабилитронов применяют только кремниевые диоды.

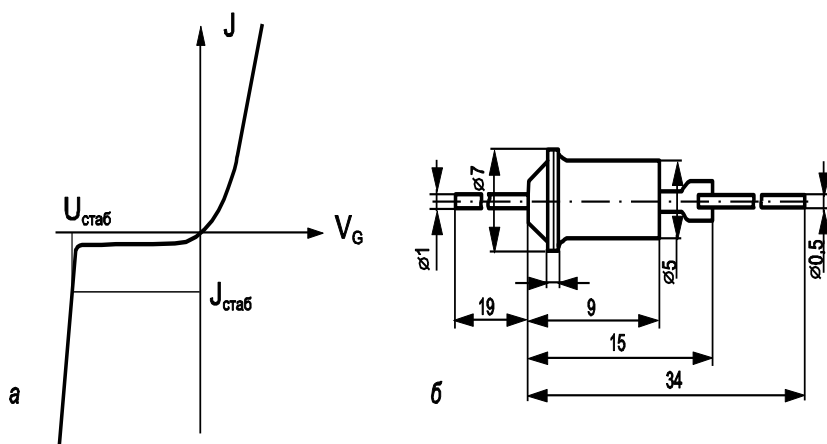


Рис. 4.9. Вольт-амперная характеристика (а) и конструкция корпуса (б) стабилитрона

Основными электрическими параметрами стабилитрона являются (в скобках даны их типовые значения):

- **напряжение стабилизации** U_{cm} – падение напряжения на стабилитроне при протекании заданного тока стабилизации (несколько вольт – сотни вольт);

- **ток стабилизации** I_{cm} – номинальное значение тока, протекающего через стабилитрон, определяющее напряжение стабилизации (несколько мА – несколько А);

- **разброс величины напряжения стабилизации** ΔU_{cm} при протекании заданного тока стабилизации (несколько единиц процента);

- **дифференциальное** или **динамическое сопротивление** r_{cm} , которое определяется при заданном значении тока стабилизации на участке пробоя как

$$r_{cm} = \frac{dU_{cm}}{dI_{cm}}. \quad (4.14)$$

Дифференциальное сопротивление определяет наклон обратной ветви вольт-амперной характеристики стабилитрона. Величина r_{cm} колеблется в пределах от 1 до 1000 Ом;

- **температурный коэффициент напряжения стабилизации** TKU_{cm} – относительное изменение напряжения стабилизации $\Delta U_{cm} / U_{cm}$ к изменению температуры окружающей среды на ΔT , т.е. при изменении температуры окружающей среды на один градус при постоянном значении тока стабилизации (плюс - минус сотые – тысячные доли процента на градус):

$$TKU_{cm} = \frac{\Delta U_{cm}}{U_{cm} \Delta T} 100, [\%/град]. \quad (4.15)$$

Кроме того, для стабилитронов существуют следующие предельно-допустимые параметры:

- максимально-допустимый ток стабилизации в диапазоне температур $I_{ст. макс}$ (десятки мА – единицы А);
- минимальный ток стабилизации в диапазоне температур $I_{ст. мин}$ (от 1 – 3 мА);
- максимально-допустимая мощность в диапазоне окружающей температуры $P_{макс}$ (от нескольких милливатт до нескольких ватт);
- полное тепловое сопротивление стабилитрона R_{Tn-c} , т.е. тепловое сопротивление переход–среда:

$$R_{Tn-c} = \frac{T_c - T_n}{P_{макс}} [^{\circ}\text{C}/\text{Вт}], \quad (4.16)$$

где T_c , T_n – температура среды и перехода соответственно.

Разброс по R_{Tn-c} составляет от десятков до сотен градусов на ватт.

При достижении напряжения на стабилитроне, называемого напряжением стабилизации $U_{стаб}$, ток через стабилитрон резко возрастает. Дифференциальное сопротивление $r_{ст}$ идеального стабилитрона на этом участке ВАХ стремится к 0, в реальных приборах величина $r_{ст}$ составляет значение $r_{ст} \approx 2 - 50 \text{ Ом}$.

Основное назначение стабилитрона – стабилизация напряжения на нагрузке, при изменяющемся напряжении во внешней цепи. В связи с этим последовательно со стабилитроном включают нагрузочное сопротивление, демпфирующее изменение внешнего напряжения. Поэтому стабилитрон называют также опорным диодом.

Напряжение стабилизации $U_{стаб}$ зависит от физического механизма, обуславливающего резкую зависимость тока от напряжения. Различают два физических механизма, ответственных за такую зависимость тока от напряжения, – лавинный и туннельный пробой p - n -перехода.

Для стабилитронов с туннельным механизмом пробоя напряжение стабилизации $U_{\text{стаб}}$ невелико и составляет величину менее 5 вольт: $U_{\text{стаб}} < 5 \text{ В}$. Для стабилитронов с лавинным механизмом пробоя напряжение стабилизации обычно имеет большие значения и составляет величину более 8 вольт: $U_{\text{стаб}} > 8 \text{ В}$.

На рис. 4.10 представлена простейшая схема стабилизации напряжения постоянного тока на стабилитроне. По этой схеме осуществляется стабилизация напряжения как при изменении входного напряжения, так и при изменении величины сопротивления нагрузки.

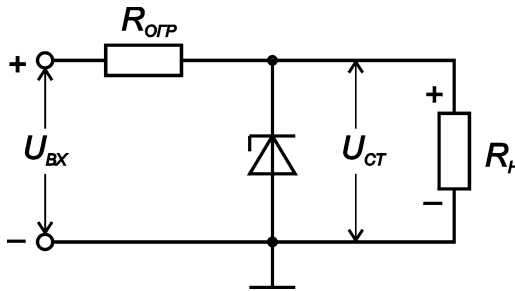


Рис. 4.10. Схема включения полупроводникового стабилитрона в схему стабилизации напряжения на нагрузке

При возрастании входного напряжения увеличивается ток стабилитрона, а следовательно, и обратный ток I_o и падение напряжения на ограничительном сопротивлении $R_{огр}$. Приращение напряжений $\Delta U_{\text{вх}}$ и $\Delta I_o R_{огр}$ взаимно компенсируются, поэтому напряжение стабилизации остается прежним.

Величина ограничительного сопротивления может быть определена по формуле

$$R_{огр} = \frac{U_{\text{вх}} - U_{\text{см}}}{I_{\text{см}} + I_{\text{н}}}, \quad (4.17)$$

где I_n – ток, протекающий через сопротивление нагрузки.

Прямая ветвь вольт-амперной характеристики кремниевых диодов, в том числе стабилитронов, имеет крутой подъем, поэтому для стабилизации малых напряжений (порядка 0,8 – 1,5 В) можно использовать кремниевые диоды, включенные в прямом направлении. Такие приборы называют **стабисторами**.

Отечественная промышленность выпускает кремниевые стабилитроны на напряжение стабилизации от 5 до 300 В при токе стабилизации от 0,1 мА до 2 А и при рассеиваемой мощности от 0,15 до 50 Вт. На рис. 4.11 приведены зависимости дифференциальных параметров различных стабилитронов от прямого тока и температуры.

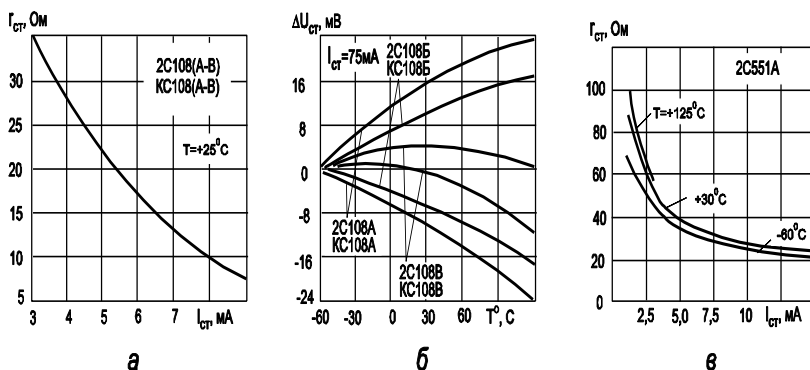


Рис. 4.11. Дифференциальные параметры различных стабилитронов: а - зависимость дифференциального сопротивления от прямого тока; б - зависимость изменения напряжения стабилизации от температуры; в - зависимость дифференциального сопротивления от прямого тока

Как следует из приведенных данных, значение дифференциального сопротивления для стабилитронов обратно пропорционально току стабилизации и составляет десятки Ом

при рабочих параметрах токов. Точность значения напряжения стабилизации составляет десятки милливольт в стандартном температурном диапазоне.

4.5. Туннельный и обращенный диоды

Туннельным диодом называют полупроводниковый диод на основе p^+-n^+ -перехода с сильнолегированными областями, на прямом участке вольт-амперной характеристики которого наблюдается n -образная зависимость тока от напряжения. На рис. 4.12 приведена вольт-амперная характеристика типичного туннельного диода при прямом смещении.

Проанализируем особенности вольт-амперной характеристики туннельного диода. Для этого рассмотрим p^+-n^+ -переход, образованный двумя вырожденными полупроводниками.

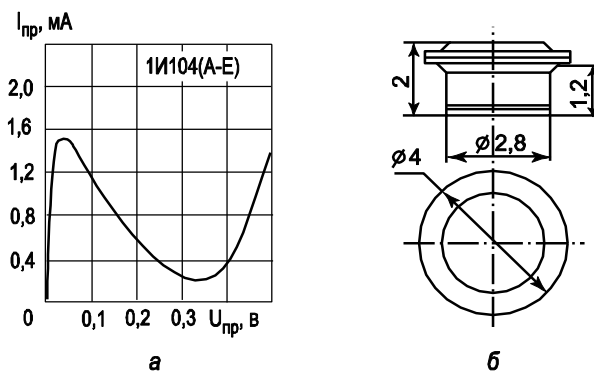


Рис. 4.12. Туннельный диод 1И104:
a – вольт-амперная характеристика при прямом смещении;
б - конструкция туннельного диода

Если концентрация доноров N_D и акцепторов N_A в эмиттере и базе диода будет порядка 10^{20} см^{-3} , то концентрация ос-

новых носителей будет много больше эффективной плотности состояний в разрешенных зонах $p_{p0}, n_{n0} \gg N_C, N_V$. В этом случае уровень Ферми будет находиться в разрешенных зонах p^+ и n^+ полупроводников.

В полупроводнике n^+ -типа все состояния в зоне проводимости вплоть до уровня Ферми заняты электронами, а в полупроводнике p^+ -типа – дырками. Зонная диаграмма p^+-n^+ -перехода, образованного двумя вырожденными полупроводниками, приведена на рис. 4.13.

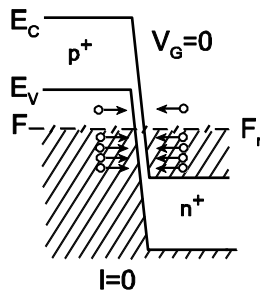


Рис. 4.13. Зонная диаграмма p^+-n^+ -перехода в состоянии равновесия

С позиции анализа токов для диффузионного тока (прямого) имеет место большая высота потенциального барьера. Чтобы получить типичные значения прямого тока, нужно приложить большое прямое напряжение (больше или примерно равное половине ширины запрещенной зоны $E_g/2$). В выражении для дрейфового тока (обратного) концентрация неосновных носителей $\left(p_{n0} = \frac{n_i^2}{N_D} \right)$ мала и поэтому обратный ток тоже будет мал.

Геометрическая ширина вырожденного p - n -перехода сравнима с дебройлевской длиной волны электрона. В этом

случае в вырожденном $p^+ - n^+$ -переходе можно ожидать проявления квантово-механических эффектов, одним из которых является туннелирование через потенциальный барьер. При узком барьере вероятность туннельного просачивания через барьер отлична от нуля.

Рассмотрим более подробно туннельные переходы в вырожденных $p^+ - n^+$ -переходах при различных напряжениях. На рис. 4.14 показана зонная диаграмма туннельного диода при обратном смещении.

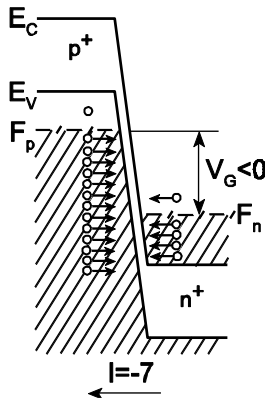


Рис. 4.14. Зонная диаграмма туннельного диода при обратном смещении

При обратном напряжении ток в диоде обусловлен туннельным переходом электронов из валентной зоны p^+ -области на свободные места в зоне проводимости n^+ -области. Поскольку концентрация электронов и число мест велики, то туннельный ток резко возрастает с ростом обратного напряжения. Такое поведение вольт-амперных характеристик резко отличает туннельный диод от обычного выпрямительного диода.

При прямом напряжении ток в диоде обусловлен туннельным переходом электронов из зоны проводимости n^+ -области на свободные места в валентной зоне p^+ -области. Поскольку туннельные переходы происходят без рассеяния, то есть с сохранением энергии туннелирующей частицы, то на зонной диаграмме эти процессы будут отражены прямыми горизонтальными линиями. На рис. 4.15 показаны зонные диаграммы туннельного диода при прямом смещении, соответствующие различным участкам прямой ветви вольт-амперной характеристики.

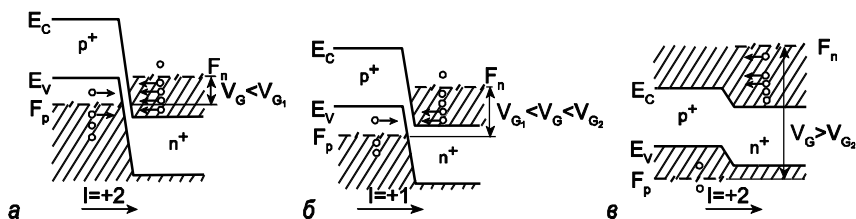


Рис. 4.15. Зонные диаграммы туннельного диода при прямом смещении:
а - участок 1; *б* - участок 2; *в* - участок 3

На участке 1 при небольшом прямом напряжении напротив электронов зоны проводимости начинают появляться свободные места в валентной зоне при той же самой энергии. По мере роста напряжения число свободных мест возрастает и ток растет с ростом напряжения. Туннельный ток достигает максимума, когда все свободные места в валентной зоне оказываются по энергии напротив энергетических уровней, занятых электронами в зоне проводимости (участок 2). Затем по мере роста прямого напряжения число этих свободных мест начинает уменьшаться, поскольку по энергии напротив уровней, занятых электронами в зоне проводимости, оказываются состояния в запрещенной зоне (энергетические уровни в иде-

альных полупроводниках в запрещенной зоне отсутствуют). На участке 3 туннельный ток уменьшается с ростом напряжения и превращается в ноль, когда запрещенная зона p^+ -полупроводника будет находиться по энергии напротив уровней, занятых электронами в зоне проводимости.

При дальнейшем росте прямого напряжения появляется компонента обычного диффузионного тока p - n -перехода.

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. 4.16. Уровень Ферми при такой средней концентрации примесей может быть расположен на потолке валентной зоны p -области и на дне зоны проводимости n -области диода, то есть потолок валентной зоны p -области и дно зоны проводимости n -области при нулевом смещении на диоде находятся на одной высоте по энергетической диаграмме.

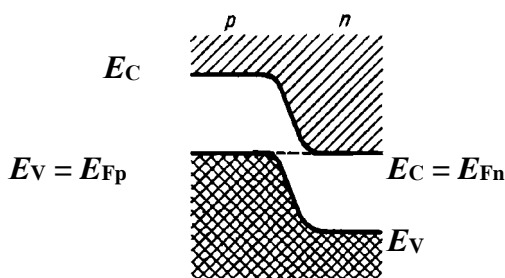


Рис. 4.16. Зонная диаграмма обращенного диода в состоянии равновесия

На рис. 4.17 приведена вольт-амперная характеристика обращенного диода. Обратная ветвь ВАХ обращенного диода аналогична обратной ветви ВАХ 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).

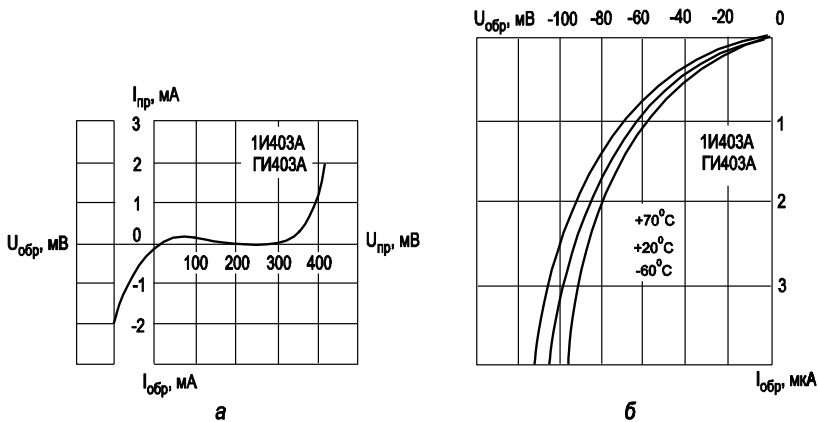


Рис. 4.17. Вольт-амперная характеристика германиевого обращенного диода ГИ403: a - полная ВАХ; b - обратный участок ВАХ при разных температурах

Что касается тока при прямом смещении, то туннельная компонента ВАХ будет полностью отсутствовать в связи с тем, что нет полностью заполненных состояний в зоне проводимости. Поэтому при прямом смещении в таких диодах до напряжений, больше или равных половине ширины запрещенной зоны, ток будет отсутствовать. Прямая ветвь ВАХ обращенного диода аналогична прямой ветви ВАХ 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 p - n -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.

Таким образом, обращенные диоды обладают выпрямляющим эффектом, но пропускное (проводящее) направление у них соответствует 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 p - n -переходу областях обращенные диоды оказываются мало чувствительными к воздействиям проникающей радиации.

Таким образом, **обращенный диод** – это туннельный диод без участка с отрицательным дифференциальным сопротивлением. Высокая нелинейность вольт-амперной характеристики при малых напряжениях вблизи нуля (порядка микровольт) позволяет использовать этот диод для детектирования слабых сигналов в СВЧ-диапазоне.

4.6. Высокочастотные и сверхвысокочастотные диоды

Сверхвысокочастотными называют полупроводниковые диоды, используемые для преобразования, детектирования, усиления, умножения, генерирования и управления уровнем мощности сигналов сантиметрового и миллиметрового диапазона волн.

На высоких частотах (свыше 30 МГц) реактивности начинают сильно сказываться, особенно барьерная емкость, которая при закрытом диоде обуславливает паразитное прохождение переменной составляющей сигнала. Радикальный путь уменьшения барьерной емкости состоит в уменьшении площади p - n -перехода. Первым конструктивным вариантом на этом пути являются *точечные формовочные диоды* (рис. 4.18).

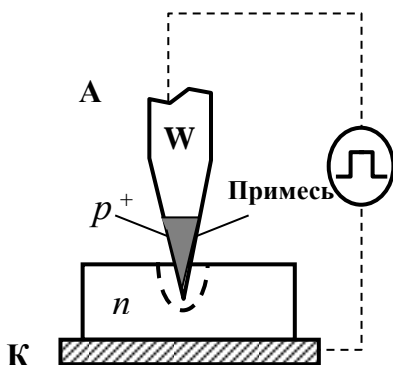


Рис. 4.18. Точечный формовочный диод

Базой диода является монокристалл германия или кремния n -типа. С одной стороны, база имеет омический контакт с держателем, а с другой – к ней прижимается вольфрамовая игла с небольшим количеством примеси на конце. Для образования перехода применяют электроформовку – пропускание короткого импульса тока. Под действием импульса тока место контакта разогревается, и примесь диффундирует в базу, образуя область эмиттера микроскопических размеров. Точечные диоды могут пропускать выпрямленный ток до нескольких десятков миллиампер и иметь обратные напряжения до 150 В. Общая емкость таких диодов не превышает 1 пФ, что позволяет использовать их на частотах до 300 - 600 МГц.

Для работы в диапазоне СВЧ применяются бесформовочные точечные диоды (рис. 4.19). В таких диодах используется переход металл – полупроводник (переход Шоттки) между вольфрамовым эмиттером и полупроводниковой базой диода. Острозаточенная игла-пружина создает переход с диаметром не менее 30 мкм, что обеспечивает уменьшение барьерной емкости до сотых долей пикофарады.

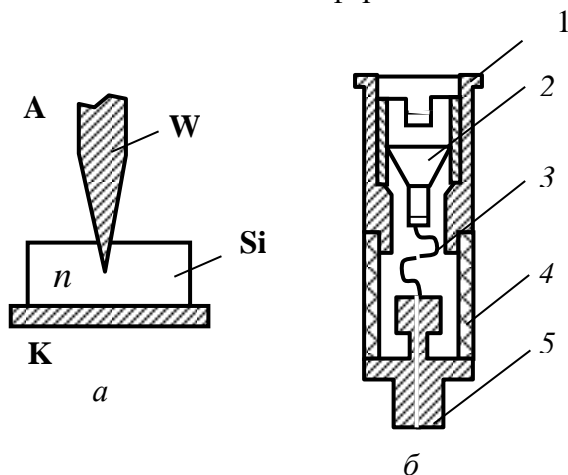


Рис. 4.19. Точечный бесформовочный диод: *а* – точечный переход металл – полупроводник; *б* – конструкция СВЧ-диода: 1 – верхний вывод; 2 – настроечный штифт; 3 – вольфрамовая игла – пружина; 4 – керамический изолятор; 5 – кристаллодержатель с кристаллом – нижний вывод

Диапазон рабочих частот бесформовочных диодов достигает десятков гигагерц. Так как это СВЧ-диапазон, то конструкция корпуса выбирается 30 мкм, что обеспечивает уменьшение барьерной емкости до сотых долей пикофарады. Диапазон рабочих частот бесформовочных диодов достигает десятков гигагерц. Так как это СВЧ-диапазон, то конструкция

корпуса выбирается с расчетом минимизации индуктивности выводов и обеспечения удобства сочленения с коаксиальными кабелями, волноводами и полосковыми линиями. Существенный недостаток таких диодов – чувствительность к механическим и электрическим нагрузкам. Даже кратковременно рассеиваемая мощность в таких диодах не должна превышать 100 мВт при обратных напряжениях 3 - 5 В. Для защиты от внешних излучений точечные диоды СВЧ хранятся и транспортируются в специальных металлических защитных колпачках.

4.7. Импульсные диоды

Импульсные диоды – полупроводниковые диоды, используемые в качестве ключевых элементов в схемах при воздействии импульсов малой длительности (микросекунды, доли микросекунд).

Одно из важнейших требований к ним – надежная работа в цепях с высокой частотой, например до 500 МГц. Импульсные диоды обладают высоким быстродействием, т.е. малым временем восстановления высокого обратного сопротивления при изменении полярности приложенного напряжения с прямой на обратную. Улучшение временных параметров достигается снижением емкости перехода и сокращением времени жизни носителей заряда. Этому требованию удовлетворяют диоды точечного типа. Могут быть использованы и диоды с *p-n*-переходом, изготовленным по диффузионной технологии. В цепях с напряжением до нескольких киловольт применяются выпрямительные столбы, состоящие из последовательно соединенных выпрямительных элементов, объединенных в одном корпусе.

Условия работы импульсных диодов обычно соответствуют высокому уровню инжекции, то есть относительно высоким токам. Поэтому свойства и параметры импульсных ди-

одов определяются переходными процессами, рассмотренными в главе 2.

При переключении диода с прямого направления на обратное в начальный момент времени через диод идет большой обратный ток, ограниченный в основном объемным сопротивлением базы (при идеальном генераторе напряжения). С течением времени накопленные в базе неосновные носители заряда рекомбинируют или уходят из базы через p - n -переход, после чего обратный ток уменьшается до своего стационарного значения (рис. 4.20, *а*). Переходный процесс, в течение которого обратное сопротивление полупроводникового диода восстанавливается до постоянного значения после быстрого переключения с прямого направления на обратное, называют восстановлением обратного сопротивления диода.

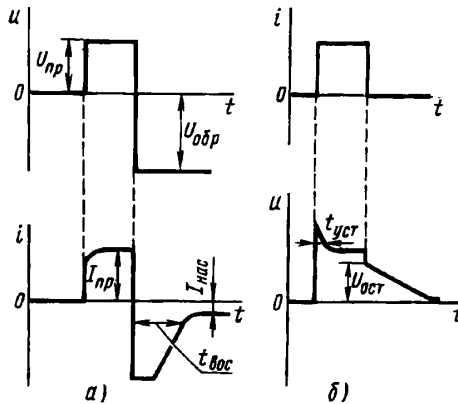


Рис. 4.20. Осциллограммы токов и напряжений импульсного диода при его работе в схемах с генератором напряжения (*а*) и с генератором тока (*б*)

При пропуске импульса тока в прямом направлении через диод наблюдается выброс напряжения в первый момент

после включения (рис. 4.20, б). Вызвано это повышенным падением напряжения до тех пор, пока не произойдет накопление неосновных носителей заряда в базе диода в результате инжекции и не уменьшится объемное сопротивление базы. Переходный процесс, в течение которого прямое сопротивление полупроводникового диода устанавливается до постоянного значения после быстрого включения диода в прямом направлении, называют установлением прямого сопротивления диода.

Специфическими параметрами импульсных диодов являются время установления прямого напряжения диода и время восстановления обратного сопротивления.

Время установления прямого напряжения диода $t_{уст}$ — это интервал времени от момента подачи импульса прямого напряжения на диод до момента достижения заданного значения прямого тока.

Время восстановления обратного сопротивления $t_{вос}$ — это интервал времени от момента прохождения тока через ноль после переключения диода с заданного прямого тока в состояние заданного обратного напряжения до момента достижения обратным током заданного низкого значения.

По времени восстановления импульсные диоды подразделяют на три группы:

- высокого быстродействия ($t_{вос} \leq 10$ нс);
- среднего быстродействия (10 нс $\leq t_{вос} \leq 100$ нс);
- низкого быстродействия ($t_{вос} > 100$ нс).

Импульсные диоды характеризуются также величинами **максимального прямого и обратного импульсного тока, максимального импульсного прямого сопротивления.**

Время, в течение которого обратный ток постоянен, называют **временем среза.**

Для импульсных диодов **время среза** $t_{ср}$ и **время восстановления обратного сопротивления диода** являются важными параметрами. Для уменьшения их значения существуют

несколько способов. Во-первых, можно уменьшать время жизни неравновесных носителей в базе диода за счет введения глубоких рекомбинационных центров в квазинейтральном объеме базы. Во-вторых, можно делать базу диода тонкой для того, чтобы неравновесные носители рекомбинировали на тыльной стороне базы.

Другим способом уменьшения времени восстановления обратного сопротивления является использование базы с неравномерной концентрацией примеси. У таких диодов концентрация примесей в базе при приближении к p - n -переходу уменьшается, поэтому неравномерной оказывается и концентрация основных носителей - электронов. За счет этого электроны диффундируют в сторону p - n -перехода, оставляя вдали от него нескомпенсированный заряд положительных ионов. Это приводит к возникновению электрического поля в базе, направленного в сторону перехода. Под действием этого поля дырки, инжектированные в базу при включении диода в прямом направлении, концентрируются (накапливаются) у границы p - n -перехода. Поэтому такие диоды называют **диодами с накоплением заряда**.

5. БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

5.1. Структура и основные режимы работы

В 1948 г. американские ученые Дж. Бардин и В. Браттейн создали полупроводниковый триод, или транзистор. Это событие имело громадное значение для развития полупроводниковой электроники. Транзисторы могут работать при значительно меньших напряжениях, чем ламповые триоды, и не являются простыми заменителями последних: их можно использовать не только для усиления и генерации переменного тока, но и в качестве ключевых элементов.

Транзистором называется полупроводниковый прибор с двумя электронно-дырочными переходами, предназначенный для усиления и генерирования электрических сигналов. Определение «**биполярный**» указывает на то, что работа транзистора связана с процессами, в которых принимают участие носители заряда двух сортов (электроны и дырки).

Биполярный транзистор (БТ) состоит из трех областей монокристаллического полупроводника с разным типом проводимости: **эмиттера**, **базы** и **коллектора** (рис. 5.1).

Переход, который образуется на границе эмиттер – база, называется эмиттерным, а на границе база – коллектор – коллекторным. В зависимости от типа проводимости крайних слоев различают транзисторы *p-n-p* и *n-p-n*.

Область транзистора, расположенная между переходами, называется **базой** (Б). Примыкающие к базе области чаще всего делают неодинаковыми. Одну из них изготавливают так, чтобы из нее наиболее эффективно происходила инжекция в базу, а другую – так, чтобы соответствующий переход наилучшим образом осуществлял экстракцию инжестированных носителей из базы.

Область транзистора, основным назначением которой является инжекция носителей в базу, называют эмиттером (Э), а соответствующий переход – эмиттерным.

Область, основным назначением которой является экстракция носителей из базы, называют коллектором (К), а переход – коллекторным.

Условные обозначения обоих типов транзисторов, рабочие полярности напряжений и направления токов показаны на рис. 5.1.

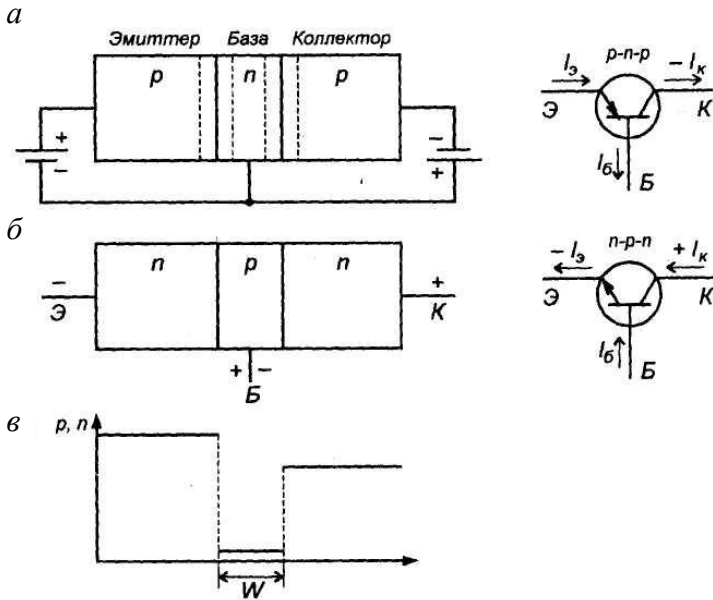


Рис. 5.1. Схематическое изображение биполярного плоскостного транзистора и его условное изображение: а - p-n-p-типа; б - n-p-n-типа; в – распределение концентраций основных носителей заряда вдоль структуры транзистора в равновесном состоянии; W - толщина базы

По технологии изготовления транзисторы делятся на *сплавные*, *планарные*, а также *диффузионно-сплавные*, *мезапланарные* и *эпитаксиально-планарные* (рис. 5.2).

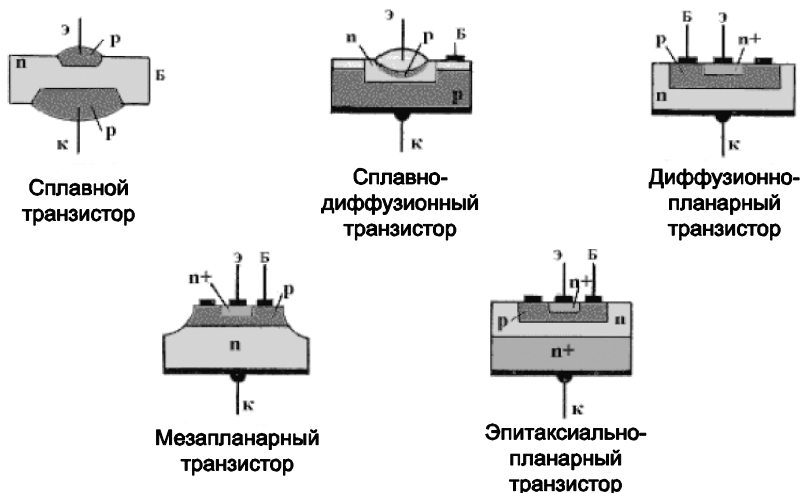


Рис. 5.2. Разновидности транзисторов по технологии изготовления

Конструктивно биполярные транзисторы оформляются в металлических, пластмассовых или керамических корпусах (рис. 5.3).

Каждый из переходов транзистора можно включить либо в прямом, либо в обратном направлении. В зависимости от этого различают три режима работы транзистора:

- режим *отсечки* – оба *p-n*-перехода закрыты, при этом через транзистор обычно идет сравнительно небольшой ток;
- режим *насыщения* – оба *p-n*-перехода открыты;
- *активный* режим – один из *p-n*-переходов открыт, а другой закрыт.

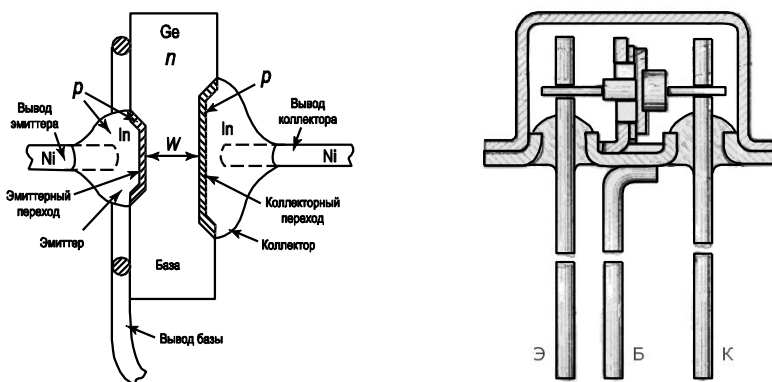


Рис. 5.3. Конструктивное оформление биполярного транзистора

В режиме отсечки и режиме насыщения управление транзистором невозможно. В активном режиме такое управление осуществляется наиболее эффективно, причем транзистор может выполнять функции активного элемента электрической схемы.

Если на эмиттерном переходе напряжение прямое, а на коллекторном переходе – обратное, то включение транзистора считают *нормальным*, при противоположной полярности – *инверсным*.

По характеру движения носителей тока в базе различают диффузионные и дрейфовые биполярные транзисторы.

Основные характеристики транзистора определяются в первую очередь процессами, происходящими в базе. В зависимости от распределения примесей в базе может присутствовать 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 называют *дрейфовым*,

если же поле в базе отсутствует – бездрейфовым (*диффузионным*).

5.2. Схемы включения транзистора

Различают три схемы включения транзистора в зависимости от того, какой из электродов транзистора является общим для входного и выходного сигналов (рис. 5.4): с общей базой (ОБ); с общим эмиттером (ОЭ) и с общим коллектором (ОК). Первый закон Кирхгофа применительно к транзистору дает равенство

$$I_{\varepsilon} = I_{\kappa} + I_{\theta}, \quad (5.1)$$

то есть ток эмиттера в транзисторе распределяется между базой и коллектором.

В этих схемах источники постоянного напряжения и резисторы нагрузки обеспечивают режимы работы транзисторов по постоянному току.

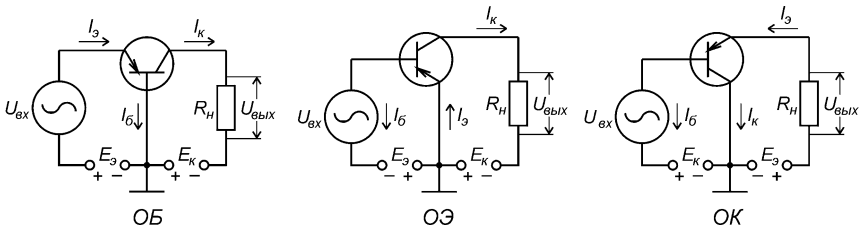


Рис. 5.4. Основные схемы включения транзисторов

Входные сигналы переменного тока создаются источниками $U_{вх}$. Они изменяют ток эмиттера транзистора, а соответственно и ток коллектора, и ток базы согласно уравнению (5.1), то есть

$$I_{\vartheta} + \Delta I_{\vartheta} = I_{\kappa} + \Delta I_{\kappa} + I_{\bar{\sigma}} + \Delta I_{\bar{\sigma}}. \quad (5.2)$$

Если из равенства (5.2) вычесть равенство (5.1), получим

$$\Delta I_{\bar{\sigma}} = \Delta I_{\kappa} + \Delta I_{\bar{\sigma}}. \quad (5.3)$$

Тогда

$$\Delta I_{\bar{\sigma}} = \Delta I_{\vartheta} - \Delta I_{\kappa} = \Delta I_{\vartheta} - \alpha \Delta I_{\vartheta} = \Delta I_{\vartheta} (1 - \alpha). \quad (5.3,а)$$

Транзистор в любой схеме применения характеризуется *коэффициентом прямой передачи по току*, который определяет отношение приращения выходного тока к вызвавшему его приращению входного тока при постоянном напряжении в выходной цепи, то есть

$$K = \frac{\Delta I_{\text{вых}}}{\Delta I_{\text{вх}}} \text{ при } U_{\text{вых}} = \text{const}. \quad (5.4)$$

Для схем с ОБ входным током является ток эмиттера, выходным – ток коллектора. Так как эмиттерный переход находится в открытом состоянии, то входное сопротивление схемы – сопротивление эмиттер–база ($R_{\text{вх}} = U_{\text{вх}} / I_{\text{вх}}$) будет малым (единицы – десятки Ом).

Тогда для этой схемы включения транзистора коэффициент прямой передачи по току

$$K_I = \frac{\Delta I_{\kappa}}{\Delta I_{\vartheta}} = \frac{\alpha \Delta I_{\vartheta}}{\Delta I_{\vartheta}} = \alpha < 1. \quad (5.5)$$

Следовательно, схема с общей базой не обладает усиленными свойствами.

Коэффициент усиления по напряжению на малом сигнале для этой схемы можно определить как

$$K_U = \frac{\Delta U_{\text{вых}}}{\Delta U_{\text{вх}}} = \frac{\Delta I_{\kappa} R_H}{\Delta I_{\text{э}} r_{\text{э}}} = \frac{\alpha I_{\text{э}} R_H}{\Delta I_{\text{э}} r_{\text{э}}} = \alpha \frac{R_H}{r_{\text{э}}}, \quad (5.6)$$

где $r_{\text{э}}$ – дифференциальное сопротивление перехода эмиттер–база. Так как отношение $R_H / r_{\text{э}}$ значительно больше входного сопротивления, то схема с ОБ может усиливать входное напряжение.

Коэффициент усиления по мощности

$$K_P = \frac{\Delta P_{\text{вых}}}{\Delta P_{\text{вх}}} = \frac{(\Delta I_{\kappa})^2 R_H}{(\Delta I_{\text{э}})^2 r_{\text{э}}} = \alpha^2 \frac{R_H}{r_{\text{э}}}. \quad (5.7)$$

Так как отношение $R_H / r_{\text{э}} > \alpha^2$, то можно утверждать, что схема с ОБ обладает некоторым усилением по мощности. Например, теоретические расчеты показали, что для схемы с ОБ при $\alpha = 0,98$ коэффициент усиления по напряжению K_U может быть равен 196, а по мощности $K_P = 192$.

Для схемы с ОЭ входным током является базовый ток, а выходным током коллекторный. Входное сопротивление в этой схеме будет примерно на два порядка больше, чем в схеме с ОБ:

$$R_{\text{вх}} = \frac{U_{\text{бэ}}}{I_{\text{б}}} = \frac{U_{\text{бэ}}}{I_{\text{э}}(1-\alpha)} = \frac{r_{\text{э}}}{1-\alpha}. \quad (5.8)$$

Коэффициент прямой передачи по току в схеме с ОЭ равен:

$$K_I = \frac{\Delta I_{\kappa}}{\Delta I_{\sigma}} = \frac{\alpha \Delta I_o}{\Delta I_o (1-\alpha)} = \frac{\alpha}{1-\alpha} = \beta \gg 1, \quad (5.9)$$

то есть эта схема обладает усилением по току. Из этой формулы видно, что чем больше будет коэффициент α , тем больше будет величина β . Например, при $\alpha = 0,96$ получим $\beta = 24$, при $\alpha = 0,99$ имеем $\beta = 99$.

Коэффициент усиления по напряжению

$$K_U = \frac{\Delta U_{\text{вых}}}{\Delta U_{\text{вх}}} = \frac{\Delta U_{\text{эк}}}{\Delta I_{\sigma} r_{\sigma}} = \frac{\Delta I_{\kappa} R_{\text{н}}}{\Delta I_{\sigma} r_{\sigma}} = \frac{\alpha \Delta I_{\sigma} R_{\text{н}} (1-\alpha)}{(1-\alpha) \Delta I_{\sigma} r_{\sigma}} = \alpha \frac{R_{\text{н}}}{r_{\sigma}} \quad (5.10)$$

такой же, как и в схеме с общей базой.

Так как схема с ОЭ обладает усилением по току и по напряжению, то величина коэффициента усиления по мощности будет наибольшей

$$K_P = K_I K_U = \frac{\alpha^2 R_{\text{н}}}{(1-\alpha) r_{\sigma}} = \frac{K_{P(\text{об})}}{1-\alpha}, \quad (5.11)$$

где $K_{P(\text{об})}$ – коэффициент усиления по мощности в схеме с ОБ.

Для схемы с ОК входным током является ток базы, а выходным током служит ток эмиттера. Тогда коэффициент усиления по току $K_I = \Delta I_{\sigma} / [\Delta I_{\sigma} (1-\alpha)] = 1/(1-\alpha)$ будет наибольший. Входное сопротивление транзисторов в схеме с ОК

$$R_{\text{вх}} = \frac{U_{\text{бк}}}{I_{\sigma}} = \frac{I_{\sigma} (R_{\text{н}} + r_{\sigma})}{I_{\sigma} (1-\alpha)} = \frac{R_{\text{н}} + r_{\sigma}}{1-\alpha} \quad (5.12)$$

значительно превышает входное сопротивление транзисторов в схемах с ОБ и ОЭ. Коэффициент усиления по напряжению равен

$$K_U = \frac{\Delta U_{\text{вых}}}{\Delta U_{\text{вх}}} = \frac{\Delta I_{\text{э}} R_H}{R_{\text{вх}} \Delta I_{\text{к}}} = \frac{\Delta I_{\text{э}} R_H (1 - \alpha)}{\Delta I_{\text{э}} (1 - \alpha) (R_H + r_{\text{э}})} = \frac{R_H}{R_H + r_{\text{э}}}. \quad (5.13)$$

Так как $R_H \gg r_{\text{э}}$, то в схеме с ОК $K_U \approx 1$.

Коэффициент усиления по мощности в схеме с ОК равен $K_P = K_I K_U \approx 1/(1 - \alpha)$, т.е. приблизительно равен коэффициенту усиления по току. Обобщенные данные по параметрам для различных схем включения транзистора представлены в табл. 5.1.

Таблица 5.1

Параметры для схем включения транзистора

Параметр	Обозначение	Схема включения транзистора		
		ОБ	ОЭ	ОК
Усиление по току	K_I	менее 1	(10 -250)	(10 -250)
Усиление по напряжению	K_U	Среднее (30 - 300)	Высокое (50 - 2000)	менее 1
Усиление по мощности	K_P	Среднее (30-300)	Высокое более 1000	Среднее (10-250)
Входное сопротивление	$R_{\text{вх}}$	Низкое (20 - 120) Ом	Среднее (150 Ом -1,5кОм)	Высокое (10 - 500 кОм)
Выходное сопротивление	$R_{\text{вых}}$	Высокое (1 - 1,5)	Среднее (10 - 100)	Низкое (10 - 100)

ление		МОм)	кОм)	Ом)
-------	--	------	------	-----

Практические схемы включения транзистора отличаются от приведенного на рис. 5.4 добавлением радиокомпонентов (резисторов, конденсаторов, индуктивностей и др.) для задания рабочего режима, нагрузки, разделения постоянной и переменной составляющих.

Таким образом, для всех схем включения транзистора общим является усиление по мощности.

Необходимо обратить внимание на то, что при обрыве базового вывода при подключенных к транзистору источниках питания, коллекторный ток в зависимости от конструкции транзистора (в первую очередь от толщины базы) может резко возрасти, что приведет к выходу транзистора из строя. Поэтому при работе с транзисторами рекомендуется базовый вывод всегда подключать к схеме первым и отключать последним.

5.3. Основные физические процессы в биполярных транзисторах

В рабочем режиме биполярного транзистора протекают следующие физические процессы: *инжекция, диффузия, рекомбинация и экстракция*.

Рассмотрим транзистор *p-n-p*-типа в нормальном активном режиме, то есть эмиттерный переход включен в прямом направлении, а коллекторный - в обратном, в схеме с ОБ.

Из-за наличия двух *p-n*-переходов биполярный транзистор может производить усиление по току, напряжению и мощности. Дырки (основные носители заряда в эмиттерной области *p-n-p*-транзистора) инжектируются прямо смещенным эмиттерным переходом в *n*-область базы, а электроны из области базы одновременно проходят через переход в *p*-область. Эмиттерная *p*-область легирована обычно сильнее, концентрация акцепторов в ней $N_{a3} = 10^{19} - 10^{20} \text{ см}^{-3}$, а концентрация

доноров в базе $N_{\text{дб}} = 10^{14} - 10^{15} \text{ см}^{-3}$, т.е. инжекция электронов из n -области в p -область много меньше, чем инжекция дырок из p -области в n -область. Дырки, попавшие в базу, оказываются в избытке у эмиттерного перехода и начинают свое движение (диффузию) вдоль базы туда, где их концентрация меньше, то есть к обратному смещенному коллекторному переходу. Далее дырки втягиваются полем обратного смещенного коллекторного перехода в противоположную область (коллектор) – экстракция зарядов – и выходят через коллекторный вывод. В базе также происходит рекомбинация инжектированных дырок с электронами (основными носителями). Для уменьшения потерь дырок в базе необходимо, чтобы ее толщина была как можно меньше. Если база тонкая, то практически все дырки дойдут до коллекторного перехода (98 – 99 % дырок), т.е. дырочный ток коллектора составляет 0,98 – 0,99 дырочного тока эмиттера $I_{\text{рк}} = (0,98 - 0,99) I_{\text{рз}}$, а 1 – 2 % дырок теряется в базе за счет рекомбинации.

Процесс переноса инжектированных носителей через базу – диффузионный. Характерное расстояние, на которое неравновесные носители распространяются от области возмущения, – диффузионная длина L_p . Поэтому если необходимо, чтобы инжектированные носители достигли коллекторного перехода, длина базы W должна быть меньше диффузионной длины L_p . Условие $W < L_p$ является необходимым для реализации транзисторного эффекта – управления током во вторичной цепи через изменение тока в первичной цепи.

В процессе диффузии через базу инжектированные основные носители рекомбинируют с основными носителями в базе. Для восполнения прорекомбинировавших основных носителей в базе через внешний контакт должно подойти такое же количество носителей. Таким образом, ток базы – это рекомбинационный ток.

Продиффундировавшие через базу без рекомбинации носители попадают в электрическое поле обратного смещенного

коллекторного p - n -перехода и экстрагируются из базы в коллектор. Таким образом, в БТ реализуются четыре физических процесса:

- инжекция из эмиттера в базу;
- диффузия через базу;
- рекомбинация в базе;
- экстракция из базы в коллектор.

Эти процессы для одного типа носителей схематически показаны на рис. 5.5, а, б.

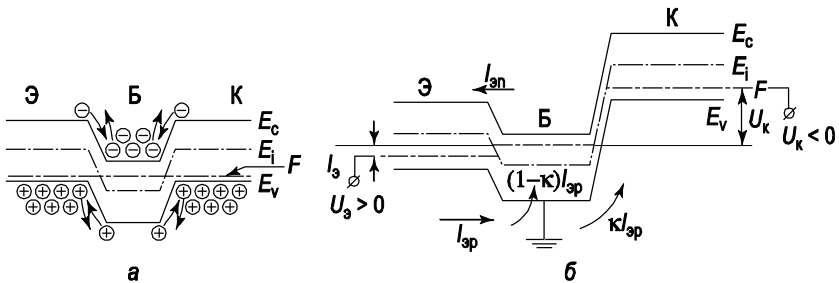


Рис. 5.5. Зонная диаграмма биполярного транзистора:
а - в равновесном состоянии; б - в активном режиме

Если бы эмиттерный и коллекторный переходы находились на большом расстоянии друг от друга, то есть толщина базы W была бы значительно больше диффузионной длины неосновных носителей в базе, то носители, инжектируемые эмиттером, не доходили бы до коллектора, так как рекомбинировали бы в базе. В этом случае каждый из переходов можно рассматривать в отдельности, не учитывая их взаимодействия, причем вольт-амперная характеристика эмиттерного перехода представляла бы прямую ветвь характеристики диода, а коллекторного перехода – обратную ветвь.

Основная особенность биполярного транзистора заключается во взаимном влиянии переходов друг на друга. В бипо-

лярных плоскостных транзисторах для эффективного влияния эмиттерного перехода на коллекторный необходимо выполнение следующих требований:

- толщина базы транзистора W должна быть много меньше диффузионной длины инжектируемых в нее носителей L_δ , то есть $W = 1,5 - 25 \text{ мкм} < L_\delta$;

- концентрация основных носителей в базе должна быть много меньше концентрации основных носителей в области эмиттера;

- концентрация основных носителей в области коллектора должна быть несколько меньшей, чем в области эмиттера;

- площадь коллекторного перехода должна быть в несколько раз больше площади эмиттерного перехода.

На рис. 5.5, *а* показана зонная диаграмма биполярного транзистора в схеме с общей базой в условиях равновесия. Значками (+) и (-) на этой диаграмме указаны основные и неосновные носители.

Для биполярного транзистора в схеме с общей базой активный режим (на эмиттерном переходе – прямое напряжение, на коллекторном – обратное) является основным.

Рассмотрим компоненты токов в эмиттерном и коллекторном переходах (рис. 5.5, *б*). Для любого p - n -перехода ток J определяется суммой электронного J_n и дырочного J_p тока, а они в свою очередь имеют дрейфовую и диффузионную составляющие:

$$\begin{aligned}
 J &= J_p + J_n = J_{pD} + J_{pE} + J_{nD} + J_{nE} = \\
 &= \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) (e^{\beta V_G} - 1)
 \end{aligned}
 \tag{5.14}$$

При приложении к эмиттерному переходу прямого напряжения $U_3 > 0$ в биполярном транзисторе p - n - p происходит инжекция дырок из эмиттера в базу $I_{эp}$ и электронов из базы в эмиттер $I_{эn}$. Ввиду того, что эмиттер легирован намного

сильнее базы, ток инжектированных дырок $I_{эp}$ будет значительно превышать ток электронов $I_{эn}$. Инжектированные в базу дырки в результате диффузии будут перемещаться к коллекторному переходу, и если ширина базы W много меньше диффузионной длины L_p , почти все дырки дойдут до коллектора и электрическим полем коллекторного перехода будут переброшены в p -область коллектора. Возникающий вследствие этого коллекторный ток лишь немного меньше тока дырок, инжектированных эмиттером.

Вольт-амперные характеристики БТ в активном режиме ($U_k < 0, |U_k| \gg 0$):

$$I_э = I_к + I_б, \quad (5.15)$$

где $I_э$ – ток в цепи эмиттера, $I_к$ – ток в цепи коллектора, $I_б$ – ток на базовом выводе.

В активном режиме к эмиттеру приложено прямое напряжение и через переход течет эмиттерный ток $I_э$, имеющий две компоненты:

$$I_э = I_{эp} + I_{эn}, \quad (5.16)$$

где $I_{эp}$ – ток инжекции дырок из эмиттера в базу, $I_{эn}$ – ток инжектированных электронов из базы в эмиттер. Величина «полезной» дырочной компоненты равняется $I_{эp} = \gamma \cdot I_э$, где γ – эффективность эмиттера. Величина дырочного эмиттерного тока, без рекомбинации дошедшая до коллектора, равняется $\gamma\chi I_э$, где χ – коэффициент переноса неосновных носителей через базу.

Ток базы $I_б$ транзистора будет состоять из трех компонент, включающих электронный ток в эмиттерном переходе $I_{эn} = (1 - \gamma) \cdot I_э$, рекомбинационный ток в базе $(1 - \chi)\gamma I_э$ и тепловой ток коллектора $I_{к0}$.

Тепловой ток коллектора $I_{к0}$ имеет две составляющие:

$$I_{к0} = I_0 + I_g, \quad (5.17)$$

где I_0 – тепловой ток, I_g – ток генерации.

На рис. 5.6 приведена схема биполярного транзистора в активном режиме, иллюстрирующая компоненты тока в схеме с общей базой.

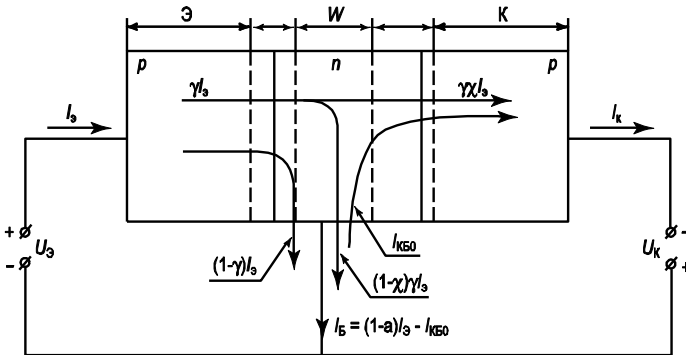


Рис. 5.6. Схема, иллюстрирующая компоненты тока в биполярном транзисторе в схеме с общей базой

В схеме с ОЭ входной цепью является цепь базы. Так как ток базы существенно меньше тока эмиттера, то можно получить усиление по току. Изменяя базовый ток, меняем количество основных носителей в области базы, то есть заряд базы и, следовательно, понижается потенциальный барьер между эмиттером и базой. Это вызывает инжекцию неосновных носителей заряда из эмиттера. Большинство инжектированных носителей доходит до коллекторного перехода, изменяя его ток. Основные носители заряда, введенные в базу из вывода базы, могут либо рекомбинировать и исчезнуть, либо быть инжектированы в эмиттер, что мало вероятно. Таким образом, на одной основной носитель заряда, вошедший в базу,

приходится много неосновных носителей заряда прошедших от эмиттера к коллектору. В этом заключается усиление по току в схеме с ОЭ.

5.4. Модуляция сопротивления базы

Как известно, ширина p - n -перехода зависит от напряжения на нем. В активном режиме эмиттерный переход смещен в прямом направлении, его ширина мала и изменение этой ширины при изменении $U_{ЭБ}$ незначительно. Коллекторный переход, смещенный в обратном направлении, имеет сравнительно большую ширину, и изменения ее при изменениях напряжения $U_{КБ}$ оказываются практически равными изменениям толщины базы.

Изменение ширины базы, под действием изменения напряжения смещения коллекторного перехода $U_{КБ}$ называется *модуляцией ширины базы* или *эффектом Эрли* (рис. 5.7).

Изменение ширины базы существенно влияет на физические процессы в базе.

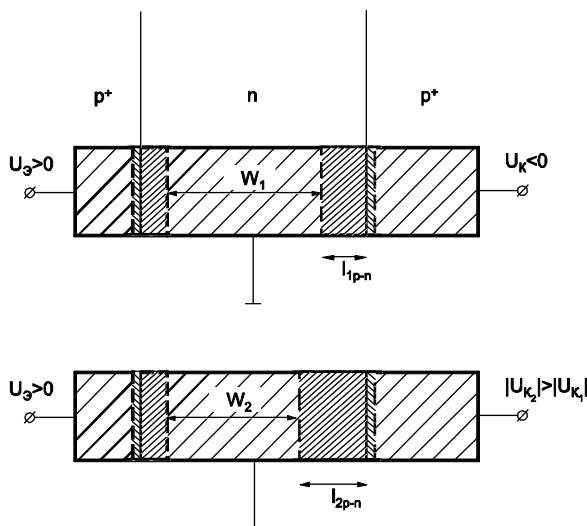


Рис. 5.7. Эффект Эрли – эффект модуляции ширины базы биполярного транзистора

При постоянном токе эмиттера $J_3 = const$ и соответственно постоянном градиенте концентрации инжектированных носителей $dp/dx = const$ с увеличением напряжения на коллекторе $U_{КБ}$ увеличивается ширина обедненной области l_{p-n} коллекторного перехода, что вызывает уменьшение ширины квазинейтрального объема базы W . Это, в свою очередь, влечет за собой уменьшение концентрации инжектированных носителей $p_n(0)$ на границе эмиттерного перехода (так как градиент dp/dx должен оставаться постоянным) (рис. 5.8). Поскольку концентрация инжектированных дырок на границе эмиттерного перехода $p_n(0) = p_0 \cdot \exp(\beta U_{ЭБ})$ определяется напряжением на эмиттере, то ее уменьшение возможно только при уменьшении напряжения $U_{ЭБ}$ на эмиттере. Таким образом, если поставлено условие: $J_3 = const$, $dp/dx = const$, то при увеличении коллекторного напряжения $U_{КБ}$ должно происходить уменьшение эмиттерного напряжения $U_{ЭБ}$.

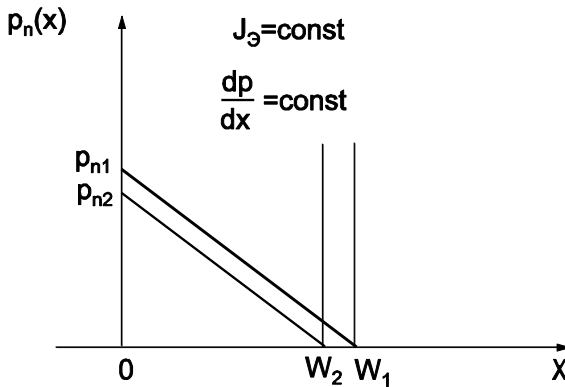


Рис. 5.8. Влияние эффекта модуляции ширины базы БТ на концентрацию неосновных носителей на границе эмиттер – база

При постоянном напряжении на эмиттерном переходе $U_{ЭБ} = const$ с увеличением $U_{КБ}$, то есть при уменьшении ширины квазинейтрального объема базы W , концентрация дырок у эмиттерного перехода остается неизменной $p_n(0) = const$. Тогда изменяется градиент концентрации инжектированных носителей dp/dx , что эквивалентно изменению тока эмиттера $J_э$, в данном случае он увеличивается.

Следовательно, если одна из входных величин ($J_э$ или $U_{ЭБ}$) задана, то вторая оказывается функцией коллекторного напряжения. Такое влияние называется *внутренней обратной связью по напряжению*.

Коэффициент обратной связи по напряжению в биполярном транзисторе в схеме с общей базой показывает, как изменится напряжение на эмиттерном переходе при единичном изменении напряжения на коллекторном переходе при условии, что ток эмиттера поддерживается постоянным

$$\mu_{\dot{y}\dot{e}} = \left. \frac{dU_{\dot{e}}}{dU_{\dot{y}}} \right|_{J_{\dot{y}}=const} .$$

5.5. Статические характеристики биполярных транзисторов

С точки зрения теории электрических цепей транзистор, имеющий три вывода, представляет собой трехполюсник. При включении трехполюсника один вывод является входным, другой – выходным, а третий вывод является общим. Поэтому трехполюсник можно представить как четырехполюсник (рис. 5.9), обозначив входной ток и входное напряжение индексом «1», а выходные ток и напряжение индексом «2». Одним семейством статических характеристик устанавливается

связь между тремя из четырех величин. Одна из этих величин выбирается независимой переменной (аргументом), и величина ее откладывается по оси абсцисс. Вторая величина будет зависимой переменной (функцией), величина которой откладывается по оси ординат. Из оставшихся двух величин одну выбирают постоянной, другую оставляют свободной, т.е. не учитывают. Ряд кривых, полученных для различных постоянных значений третьей величины, называемой параметром, и будет представлять собой семейство статических вольт-амперных характеристик.

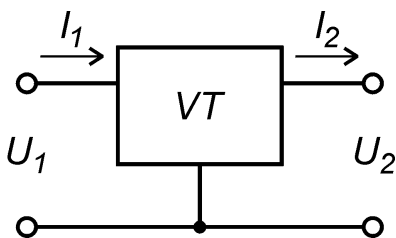


Рис. 5.9. Схема трехполюсника, включенного как четырехполюсник

Наиболее распространенными системами статических вольт-амперных характеристик биполярных транзисторов являются:

- семейство входных характеристик $I_1 = f(U_1)$ при $U_2 = const$;
- семейство выходных характеристик $I_2 = f(U_2)$ при $I_1 = const$.

Входные характеристики транзистора $p-n-p$ -типа в схеме с ОБ определяются зависимостью тока эмиттера от напряжения между эмиттером и базой при постоянном напряжении между коллектором и базой $I_3 = f(U_{эб})$ при $U_{кб} = const$ (рис. 5.10, а).

При напряжении $U_{кб} = 0$ вольт-амперная характеристика аналогична характеристике полупроводникового диода при прямом включении. Эмиттерный ток экспоненциально возрастает с увеличением напряжения между эмиттером и базой. Увеличение напряжения коллектор – база (в абсолютном значении) и повышение температуры смещает характеристики в сторону оси ординат (то есть влево).

Выходные характеристики транзисторов $p-n-p$ -типа в схеме с ОБ, называемые также коллекторными, определяются зависимостью тока коллектора от напряжения между коллектором и базой при постоянном токе эмиттера $I_k = f(U_{кб})$ при $I_э = const$ (рис. 5.10, б). При токе $I_э = 0$ через коллекторный переход течет обратный ток $I_{кб0}$, величина которого слабо зависит от коллекторного напряжения.

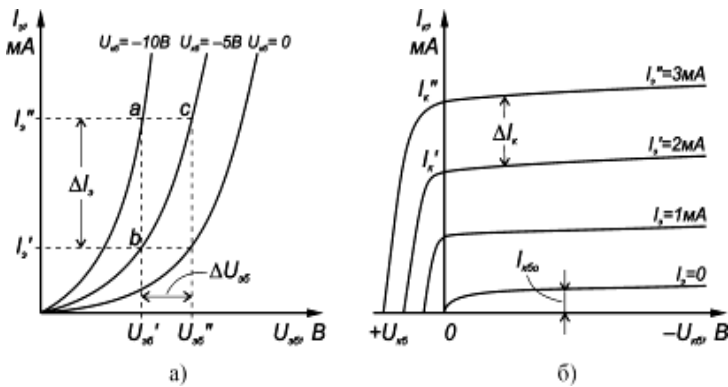


Рис. 5.10. Семейство входных (а) и выходных (б) характеристик $p-n-p$ -транзистора в схеме с ОБ

С ростом эмиттерного тока растет коллекторный ток, что соответствует принципу действия транзистора. Общий характер этих зависимостей аналогичен обратной ветви вольт-амперной характеристики диода, так как коллекторный переход включен в обратном направлении.

На рис. 5.11, *a* представлены семейства входных характеристик транзистора *p-n-p*-типа в схеме с ОЭ $I_b = f(U_{эб})$ при $U_{кэ} = const$. Входным током транзистора при этой схеме включения является ток базы, характеристика которого при напряжении $U_{кэ} = 0$ идет из начала координат и представляет собой суммарный ток эмиттерного и коллекторного переходов, соединенных параллельно и подключенных к источнику в прямом направлении.

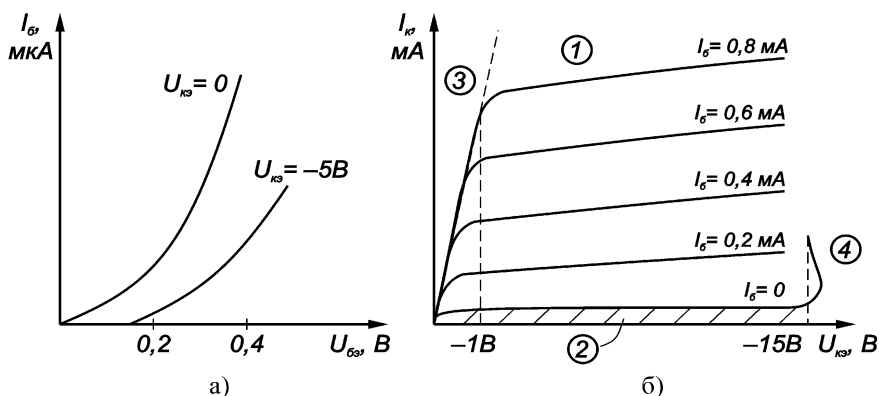


Рис. 5.11. Семейство входных (*a*) и выходных (*б*) характеристик *p-n-p*-транзистора в схеме с ОЭ

При увеличении по абсолютной величине напряжения коллектор - эмиттер входные характеристики сдвигаются вправо, так как с ростом напряжения вероятность рекомбинаций неосновных носителей в области базы уменьшается, что приводит к уменьшению тока базы и росту коллекторного тока.

В справочниках обычно приводятся входные характеристики при $U_{кэ} = 0$ и $U_{кэ} = -5$ В. Все остальные характеристики, снятые при $|U_{кэ}| > 1$ В, незначительно отличаются от характеристики, снятой при $U_{кэ} = -5$ В.

Коллекторные, т.е. выходные, характеристики транзистора в схеме с ОЭ $I_k = f(U_{кэ})$ при $I_b = const$ приведены на рис. 5.11, б. Выходные характеристики в схеме с ОЭ имеют больший наклон, чем в схеме с ОБ, то есть наблюдается большая зависимость тока коллектора от напряжения на коллекторе. С ростом базового тока коллекторные характеристики располагаются выше, так как увеличение базового тока происходит вследствие увеличения тока эмиттера, следовательно, и тока базы. Выходные характеристики транзистора при выборе в качестве параметра тока I_b можно описать с помощью уравнения

$$I_k = \beta I_b + I_{кэо},$$

где $I_{кэо}$ – начальный («сквозной») ток, который протекает через структуру при токе $I_b = 0$, т.е. при обрыве в цепи базы.

Так как $I_b = (1 - \alpha) I_э - I_{кбo}$, то при $I_b = 0$ имеем $I_э = I_k = I_{кбo}$, поэтому получим $I_{кэо} = I_{кбo} / (1 - \alpha) \approx \beta I_{кбo}$.

Ток $I_{кэо}$ в β раз больше тока $I_{кбo}$, т.е. тепловой ток в схеме с ОЭ значительно больше, чем в схеме с ОБ, что является недостатком схемы с ОЭ по сравнению со схемой с ОБ.

5.6. Эквивалентная схема биполярного транзистора

Эквивалентная схема – эта схема, в которой реальные процессы в нелинейных устройствах заменены на набор активных (источники тока и напряжения) и пассивных (резисторы, емкости) элементов, адекватно описывающих взаимосвязь входных и выходных параметров. На основе рассмотренных характеристик представим эквивалентную схему транзистора при включении по схеме с общей базой и общим эмиттером в следующем виде (рис. 5.12, 5.13). Приведенная эквивалентная схема справедлива для рассмотрения статических характеристик биполярного транзистора, а также для рассмотрения этих

характеристик в области низких частот. Эта схема называется Т-образной эквивалентной схемой, отражает основные физические процессы, происходящие в транзисторе, и удобна для их анализа.

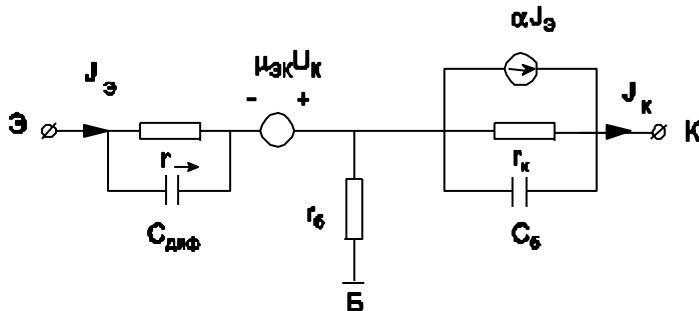


Рис. 5.12. Эквивалентная схема биполярного транзистора при включении по схеме с общей базой

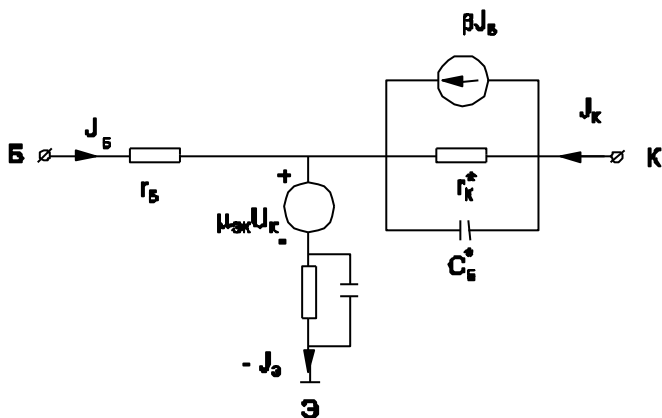


Рис. 5.13. Эквивалентная схема биполярного транзистора при включении по схеме с общим эмиттером

На эквивалентных схемах изображены основные пассивные элементы (сопротивления $r_э$, $r_κ$, $r_б$, емкости коллек-

торного C_B и эмиттерного $C_{\text{диф}}$ переходов), активные элементы (генераторы тока $\alpha J_э$, βJ_B в коллекторной цепи, источник ЭДС $\mu_{эк} U_K$ в эмиттерной цепи, отражающий обратную связь между эмиттером и коллектором).

Основные параметры эквивалентной схемы транзистора выражаются через конструктивно-технологические параметры следующим образом:

$$r_э = \frac{kT}{q} \frac{1}{J_э}; \quad r_K = \sqrt{\frac{2qN_D}{\varepsilon_s \varepsilon_0}} \cdot \frac{L_p^2}{W} \cdot \frac{\sqrt{|U_K|}}{\gamma \cdot J_э}; \quad \mu_{эк} = -\sqrt{\frac{\varepsilon_s \varepsilon_0}{2qN_{DB}}} \frac{kT/q}{W \sqrt{|U_K|}};$$

$$\alpha = \frac{dJ_{эп}}{dJ_э} \frac{dJ_K}{dJ_{эп}} = \gamma \cdot \chi; \quad \gamma = \frac{dJ_{эп}}{dJ_э} = 1 - \frac{N_{DB}}{N_{AЭ}};$$

$$\chi = \frac{dJ_K}{dJ_{эп}} = 1 - \frac{1}{2} \frac{W^2}{L^2}.$$

Величины коэффициентов $\square \alpha, r_э, r_K, \mu_{эк}$ для биполярного транзистора лежат в пределах: $\alpha = 0,95 - 0,995$, $r_э = 1 - 10$ Ом, $r_K = 10 - 10^6$ Ом, $\mu_{эк} = 10^{-3} \div 10^{-5}$.

Для биполярного транзистора при включении по схеме с общим эмиттером эквивалентная схема выглядит аналогично.

Основные параметры эквивалентной схемы имеют тот же вид, что и в схеме с общей базой, кроме C_K^* и r_K^* , равных: $C_K^* = C_K(\beta + 1)$, $r_K^* = r_K(\beta + 1)$.

Формулы Эберса – Молла являются универсальными соотношениями, которые описывают характеристики биполярных транзисторов во всех режимах работы.

Для такого рассмотрения представим БТ в виде эквивалентной схемы, приведенной на рис. 5.14. Два диода и два источника тока включены навстречу друг другу. Источники токов управляют токами диодов, сами диоды считаются идеальными.

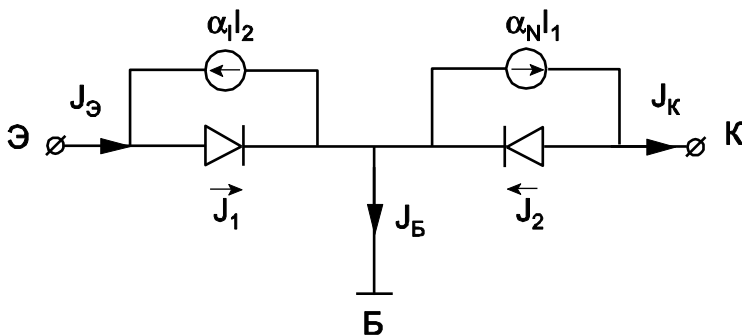


Рис. 5.14. Эквивалентная схема биполярных транзисторов во всех режимах работы

При нормальном включении через эмиттерный p -переход течет ток I_1 , через коллекторный переход течет ток $\alpha_N I_1$ – меньший чем I_1 , вследствие рекомбинации части инжектированных носителей в базе. На рис. 5.14 этот процесс изображен как генератор тока $\alpha_N I_1$, где α_N – коэффициент передачи эмиттерного тока. При инверсном включении транзистора прямому коллекторному току I_2 будет соответствовать эмиттерный ток $\alpha_1 I_2$, где α_1 – коэффициент инверсии. Таким образом, токи эмиттера $J_э$ и коллектора $J_к$ в общем случае состоят из инжектируемого (I_1 или I_2) и экстрагируемого ($\alpha_N I_1$ или $\alpha_1 I_2$) токов:

$$\begin{aligned} J_э &= I_1 - \alpha_1 I_2, \\ J_к &= \alpha_N I_1 - I_2. \end{aligned} \quad (5.18)$$

Величины токов I_1 и I_2 выражаются для p - n -переходов стандартным способом:

$$\begin{aligned} I_1 &= I_{\varepsilon 0}' (\exp(\beta U_{\varepsilon}) - 1), \\ I_2 &= I_{\kappa 0}' (\exp(\beta U_{\kappa}) - 1), \end{aligned} \quad (5.19)$$

где $I_{\varepsilon 0}'$ и $I_{\kappa 0}'$ – тепловые (обратные) токи p - n -переходов. Отметим, что токи $I_{\varepsilon 0}'$ и $I_{\kappa 0}'$ отличаются от обратных токов эмиттера $I_{\varepsilon 0}$ и коллектора $I_{\kappa 0}$ биполярного транзистора.

Оборвем цепь эмиттера ($J_{\varepsilon} = 0$) и подадим на коллекторный переход большое запирающее напряжение U_{κ} . Ток, протекающий в цепи коллектора при этих условиях, будем называть тепловым током коллектора $I_{\kappa 0}$. Поскольку $J_{\varepsilon} = 0$, из (5.18) следует, что $I_1 = \alpha_1 I_2$, а из (5.19) $I_2 = -I_{\kappa 0}'$, поскольку $U \gg kT/q$.

Полагая $J_{\kappa} = I_{\kappa 0}$, получаем в этом случае:

$$\begin{aligned} J_{\kappa} &= \alpha_N \alpha_1 I_2 - I_2 = I_2 (\alpha_N \alpha_1 - 1) = I_{\kappa 0}' (1 - \alpha_N \alpha_1) = I_{\kappa 0}, \\ I_{\kappa 0}' &= \frac{I_{\kappa 0}}{1 - \alpha_N \alpha_1}. \end{aligned} \quad (5.20)$$

Отличие обратных токов $I_{\kappa 0}$ и $I_{\kappa 0}'$ состоит в том, что ток $I_{\kappa 0}$ определяется при короткозамкнутой цепи эмиттера ($U_{\varepsilon} = 0$), а ток $I_{\kappa 0}'$ - при разомкнутой ($J_{\varepsilon} = 0$).

Обозначим ток эмиттера при большом отрицательном смещении и разомкнутой цепи коллектора через $I_{\varepsilon 0}'$ – тепловой ток эмиттера:

$$I_{\varepsilon 0}' = \frac{I_{\varepsilon 0}}{1 - \alpha_N \alpha_1}. \quad (5.21)$$

Величины теплового эмиттерного и коллекторного токов значительно меньше, чем соответствующие тепловые токи диодов.

Подставляя (5.19) в (5.18), получаем:

$$J_э = I_{э0}' (\exp(\beta U_э) - 1) - \alpha_1 I_{к0}' (\exp(\beta U_к) - 1),$$

$$J_к = \alpha_N I_{э0}' (\exp(\beta U_э) - 1) - I_{к0}' (\exp(\beta U_к) - 1), \quad (5.22)$$

$$J_б = (1 - \alpha_N) I_{э0}' (\exp(\beta U_э) - 1) + (1 - \alpha_1) I_{к0}' (\exp(\beta U_к) - 1),$$

где $J_б$ – ток базы, равный разности токов эмиттера $J_э$ и коллектора $J_к$.

Формулы (5.22) получили название формул Эберса – Молла и полезны для анализа статических характеристик биполярного транзистора при любых сочетаниях знаков токов и напряжений.

При измерении теплового тока коллектора $I_{к0}$ дырки как неосновные носители уходят из базы в коллектор: $J_к = J_б$ ($J_э = 0$). При этом поток дырок из базы в эмиттер не уравновешен и их переходит из эмиттера в базу больше, чем в равновесных условиях. Это вызовет накопление избыточного положительного заряда в базе и увеличение потенциального барьера на переходе эмиттер – база, что, в конце концов, скомпенсирует дырочные токи.

Таким образом, необходимо отметить, что при изменении теплового тока коллектора эмиттер будет заряжаться отрицательно по отношению к базе.

5.7. Дифференциальные параметры биполярных транзисторов в схеме с общей базой

Основными величинами, характеризующими параметры биполярного транзистора, являются коэффициент передачи тока эмиттера α , сопротивление эмиттерного ($r_э$) и коллекторного ($r_к$) переходов, а также коэффициент обратной связи эмиттер – коллектор $\mu_{эк}$.

Дифференциальным коэффициентом передачи тока эмиттера называется отношение приращения тока коллектора к вызвавшему его приращению тока эмиттера при постоянном напряжении на коллекторе:

$$\alpha = \left. \frac{dJ_к}{dJ_э} \right|_{U_к=const}. \quad (5.23)$$

Коэффициент передачи эмиттерного тока α характеризует изменение коллекторного тока $J_к$ при вызвавшем его изменении эмиттерного тока $J_э$. Ток коллектора обусловлен дырками, дошедшими от эмиттерного перехода до коллекторного. Поэтому важны доля дырок, дошедших до коллекторного перехода и нерекombинировавших в базе, и доля дырочного тока в эмиттерном токе.

Для коэффициента передачи α можно записать:

$$\alpha = \frac{dJ_к}{dJ_э} = \frac{dJ_{эп}}{dJ_э} \frac{dJ_к}{dJ_{эп}} = \gamma \cdot \chi, \quad (5.24)$$

где $\gamma = \frac{dJ_{эп}}{dJ_э}$ – коэффициент инжекции, или эффективность

эмиттера, $\chi = \frac{dJ_к}{dJ_{эп}}$ – коэффициент переноса.

Таким образом, γ – доля полезного дырочного тока в полном токе эмиттера $J_э$, а коэффициент χ показывает долю

эмиттерного дырочного тока, без рекомбинации дошедшего до коллекторного перехода.

Поскольку коэффициент передачи α определяется произведением коэффициентов инжекции γ и переноса χ как $\alpha = \gamma\chi$, то у сплавных транзисторов, где ширина базы составляет $W = 10 - 20$ мкм, в коэффициенте передачи α главную роль играет коэффициент переноса χ . У диффузионных транзисторов ширина базы равняется $W = (1 - 2)$ мкм и главную роль в коэффициенте передачи α играет коэффициент инжекции γ .

Сопротивление эмиттерного перехода определяется как отношение изменения напряжения на эмиттере к изменению тока эмиттера при постоянном токе коллектора:

$$r_{\text{э}} = \left. \frac{dU_{\text{э}}}{dJ_{\text{э}}} \right|_{J_{\text{к}}=\text{const}} . \quad (5.25)$$

Сопротивление коллекторного перехода определяется как отношение изменения напряжения на коллекторе к изменению тока коллектора при постоянном токе эмиттера:

$$r_{\text{к}} = \left. \frac{dU_{\text{к}}}{dJ_{\text{к}}} \right|_{J_{\text{э}}=\text{const}} . \quad (5.26)$$

Коэффициентом обратной связи называется отношение приращения напряжения на эмиттере к приращению напряжения на коллекторе при постоянном токе через эмиттер:

$$\mu_{\text{эк}} = \left. \frac{dU_{\text{к}}}{dU_{\text{э}}} \right|_{J_{\text{э}}=\text{const}} . \quad (5.27)$$

5.8. Дифференциальные параметры биполярных транзисторов в схеме с общим эмиттером

Схема включения биполярного транзистора с общим эмиттером приведена на рис. 5.15.

Характеристики транзистора в этом режиме будут отличаться от характеристик в режиме с общей базой. В транзисторе, включенном по схеме с общим эмиттером, имеет место усиление не только по напряжению, но и по току. Входными параметрами для схемы с общим эмиттером будут ток базы J_B и напряжение на коллекторе U_K , а выходными характеристиками будут ток коллектора J_K и напряжение на эмиттере $U_Э$.

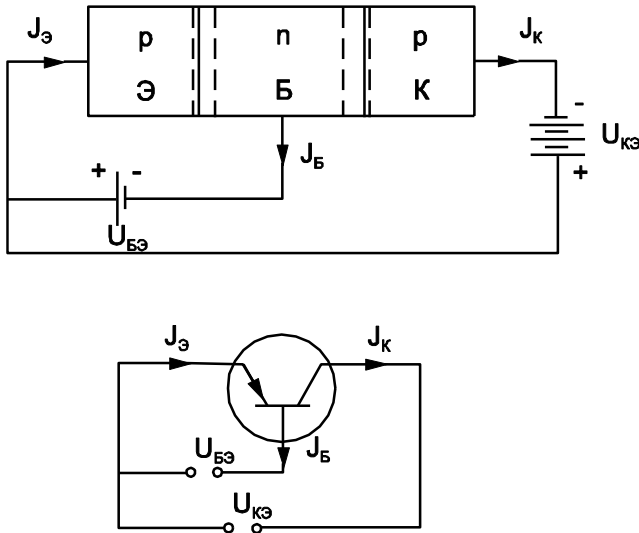


Рис. 5.15. Схема включения транзистора с общим эмиттером

Ранее при анализе биполярного транзистора в схеме с общей базой была получена связь между током коллектора и током эмиттера в следующем виде:

$$J_{\kappa} = \alpha J_{\text{э}} + I_{\kappa 0}. \quad (5.28)$$

В схеме с общим эмиттером (в соответствии с первым законом Кирхгофа) $J_{\text{э}} = J_{\text{б}} + J_{\kappa}$.

$$J_{\kappa} = \alpha(J_{\kappa} + J_{\text{б}}) + I_{\kappa 0} + \frac{U_{\kappa}}{r_{\kappa}}; \quad J_{\kappa} - \alpha J_{\kappa} = I_{\kappa 0} + \alpha J_{\text{б}} + \frac{U_{\kappa}}{r_{\kappa}}, \quad (5.28')$$

после перегруппирования сомножителей получаем:

$$J_{\kappa} = \frac{\alpha}{1-\alpha} J_{\text{б}} + \frac{I_{\kappa 0}}{1-\alpha} + \frac{U_{\kappa}}{(1-\alpha)r_{\kappa}}. \quad (5.29)$$

Коэффициент $\frac{\alpha}{1-\alpha}$ перед сомножителем $J_{\text{б}}$ показывает, как изменяется ток коллектора J_{κ} при единичном изменении тока базы $J_{\text{б}}$. Он называется коэффициентом усиления по току биполярного транзистора в схеме с общим эмиттером. Обозначим этот коэффициент β .

$$\beta = \frac{\alpha}{1-\alpha}. \quad (5.30)$$

Поскольку величина коэффициента передачи α близка к единице ($\alpha < 1$), то из уравнения (5.30) следует, что коэффициент усиления β будет существенно больше единицы ($\beta \gg 1$). При значениях коэффициента передачи $\alpha = 0,98 - 0,99$ коэффициент усиления будет лежать в диапазоне $\beta = 50 - 100$.

5.9. Малосигнальные параметры биполярного транзистора

Биполярный транзистор в схемотехнических приложениях представляют как четырехполюсник и рассчитывают его параметры для такой схемы. Для транзистора как четырехполюсника характерны два значения тока I_1 и I_2 и два значения напряжения U_1 и U_2 (рис. 5.16).

В зависимости от того, какие из этих параметров выбраны в качестве входных, а какие в качестве выходных, можно построить три системы формальных параметров транзистора как четырехполюсника. Это системы z -параметров, y -параметров и h -параметров. Рассмотрим их более подробно, используя линейное приближение.

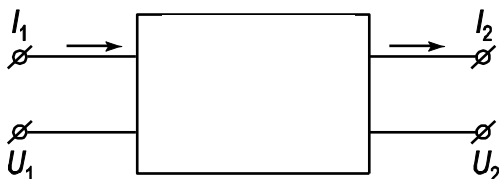


Рис. 5.16. Схема четырехполюсника

5.9.1. Система z -параметров

Зададим в качестве входных параметров биполярного транзистора как четырехполюсника токи I_1 и I_2 , а напряжения U_1 и U_2 будем определять как функции этих токов. Тогда связь напряжений и токов в линейном приближении будет иметь вид:

$$U_1 = z_{11}I_1 + z_{12}I_2; \tag{5.31}$$

$$U_2 = z_{21}I_1 + z_{22}I_2.$$

Коэффициенты z_{ik} в этих уравнениях определяются следующим образом:

- входное и выходное сопротивления

$$z_{11} = \left. \frac{U_1}{I_1} \right|_{I_2=0} \quad \text{и} \quad z_{22} = \left. \frac{U_2}{I_2} \right|_{I_1=0} ;$$

- сопротивления обратной и прямой передач

$$z_{12} = \left. \frac{U_1}{I_2} \right|_{I_1=0} \quad \text{и} \quad z_{21} = \left. \frac{U_2}{I_1} \right|_{I_2=0} .$$

Измерения z -параметров осуществляются в режиме холостого хода на входе ($I_1 = 0$) и выходе ($I_2 = 0$). Реализовать режим разомкнутого входа $I_1 = 0$ для биполярного транзистора достаточно просто (сопротивление эмиттерного перехода составляет всего десятки Ом и поэтому размыкающее сопротивление в цепи эмиттера в несколько кОм уже позволяет считать $I_1 = 0$). Реализовать режим разомкнутого выхода $I_2 = 0$ для биполярного транзистора сложно (сопротивление коллекторного перехода равняется десяткам МОм и размыкающее сопротивление в цепи коллектора в силу этого должно быть порядка ГОм).

5.9.2. Система u -параметров

Зададим в качестве входных параметров биполярного транзистора как четырехполюсника напряжения U_1 и U_2 , а токи I_1 и I_2 будем определять как функции этих напряжений. Тогда связь токов и напряжений в линейном приближении будет иметь вид:

$$I_1 = y_{11}U_1 + y_{12}U_2 ;$$

(5.32)

$$I_2 = y_{21}U_1 + y_{22}U_2.$$

Коэффициенты в уравнениях имеют размерность проводимости и определяются следующим образом:

- входная и выходная проводимости

$$y_{11} = \left. \frac{I_1}{U_1} \right|_{U_2=0} \quad \text{и} \quad y_{22} = \left. \frac{I_2}{U_2} \right|_{U_1=0};$$

- проводимости обратной и прямой передач

$$y_{12} = \left. \frac{I_1}{U_2} \right|_{U_1=0} \quad \text{и} \quad y_{21} = \left. \frac{I_2}{U_1} \right|_{U_2=0}.$$

Измерение y -параметров происходит в режиме короткого замыкания на входе ($U_1 = 0$) и выходе ($U_2 = 0$). Реализовать режим короткого замыкания на входе ($U_1 = 0$) для биполярного транзистора достаточно сложно (сопротивление эмиттерного перехода составляет всего десятки Ом и поэтому замыкающее сопротивление в цепи эмиттера должно составлять доли Ома, что достаточно сложно). Реализовать режим короткого замыкания на выходе $U_2 = 0$ для биполярного транзистора просто (сопротивление коллекторного перехода равняется десяткам МОм и замыкающие сопротивления в цепи коллектора могут быть даже сотни Ом).

5.9.3. Система h -параметров

Система h -параметров используется как комбинированная система из двух предыдущих, причем из соображений удобства измерения параметров биполярного транзистора выбирается режим короткого замыкания на выходе ($U_2 = 0$) и

режим холостого хода на входе ($I_1 = 0$). Поэтому для системы h -параметров в качестве входных параметров задаются ток I_1 и напряжение U_2 , а в качестве выходных параметров рассчитываются ток I_2 и напряжение U_1 , при этом система, описывающая связь входных I_1, U_2 и выходных I_2, U_1 параметров, выглядит следующим образом:

$$U_1 = h_{11}I_1 + h_{12}U_2 ; \tag{5.33}$$

$$I_2 = h_{21}I_1 + h_{22}U_2 .$$

Значения коэффициентов в уравнении для h -параметров определяются следующим образом:

– входное сопротивление при коротком замыкании на

выходе $h_{11} = \left. \frac{U_1}{I_1} \right|_{U_2=0} ;$

– выходная проводимость при холостом ходе во входной

цепи $h_{22} = \left. \frac{I_2}{U_2} \right|_{I_1=0} ;$

– коэффициент обратной связи при холостом ходе во

входной цепи $h_{12} = \left. \frac{U_1}{U_2} \right|_{I_1=0} ;$

– коэффициент передачи тока при коротком замыкании

на выходе $h_{21} = \left. \frac{I_2}{I_1} \right|_{U_2=0} .$

Эквивалентная схема четырехполюсника с h -параметрами приведена на рис. 5.17 а, б. Из этой схемы легко увидеть, что режим короткого замыкания на выходе или холостого хода на входе позволяет измерить тот или иной h -параметр.

Для различных схем включения биполярного транзистора (схема с общей базой, общим эмиттером и общим коллектором) h -параметры связаны друг с другом. В табл. 5.2 приведены эти связи, позволяющие рассчитывать h -параметры для схемы включения с общей базой, если известны эти параметры для схемы с общим эмиттером.

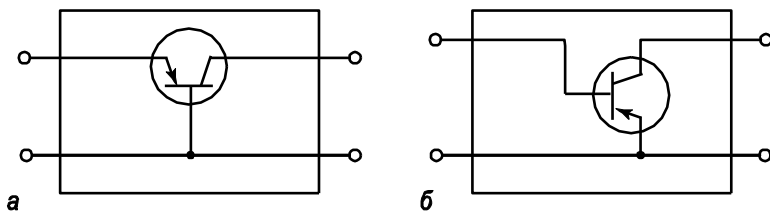


Рис. 5.17. Эквивалентная схема четырехполюсника:
 а - биполярный транзистор в схеме с общей базой;
 б - биполярный транзистор в схеме с общим эмиттером

Таблица 5.2

Связи между h -параметрами в схемах включения транзистора с общей базой и общим эмиттером

$h_{11б}$	$h_{11б} \approx \frac{h_{11э}}{1 + h_{21э}}$
$h_{12б}$	$h_{12б} \approx \frac{h_{11э}h_{22э} - h_{12э}(1 + h_{21э})}{1 + h_{21э}}$
$h_{21б}$	$h_{21б} \approx \frac{h_{21э}}{1 + h_{21э}}$
$h_{22б}$	

$$h_{22\delta} \approx \frac{h_{22\epsilon}}{1 + h_{21\epsilon}}$$

Дифференциальные параметры биполярных транзисторов зависят от режимов их работы. Для схемы с общим эмиттером наибольшее влияние испытывает коэффициент усиления эмиттерного тока $h_{21\epsilon}$ в зависимости от тока эмиттера. На рис. 5.18 приведена эта зависимость для транзисторов КТ215 различных типономиналов. В области малых токов (микро-мощный режим) коэффициент усиления уменьшается вследствие влияния рекомбинационной компоненты в эмиттерном переходе, а в области больших токов (режим высокого уровня инжекции) – коэффициент усиления уменьшается вследствие уменьшения коэффициента диффузии.

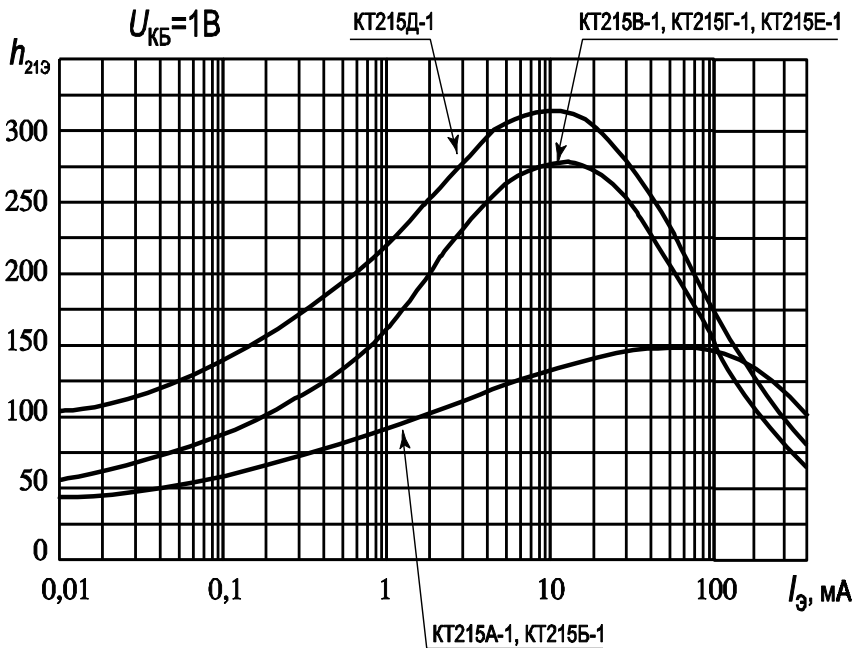


Рис. 5.18. Зависимость коэффициента h_{213} для различных транзисторов марки КТ215 от эмиттерного тока I_3

5.10. Частотные и импульсные свойства транзисторов

Динамические свойства биполярных транзисторов определяются переходной и частотной характеристиками.

На **частотные свойства** транзисторов большое влияние оказывают емкости $p-n$ -переходов. С увеличением частоты емкостное сопротивление уменьшается, и шунтирующее действие емкостей возрастает. Поэтому Т-образная эквивалентная схема транзистора на высоких частотах, кроме чисто

активных сопротивлений $r_{э}$, $r_{к}$, и $r_{б}$, содержит емкости $C_{э}$ и $C_{к}$, шунтирующие эмиттерный и коллекторный переходы. Особенно вредное влияние на работу транзистора оказывает емкость $C_{к}$, так как на высоких частотах емкостное сопротивление $1/\omega_{в}C_{к}$ оказывается значительно меньше, чем сопротивление $r_{к}$ и коллекторный переход теряет свои основные свойства. В данном случае влияние емкости $C_{к}$; аналогично влиянию емкости, шунтирующей p - n -переход в плоскостном полупроводниковом диоде.

Второй причиной ухудшения работы транзистора на высоких частотах является отставание по фазе переменного тока коллектора от переменного тока эмиттера. Это обусловлено инерционностью процесса прохождения носителей заряда через базу от эмиттерного перехода к коллекторному, а также инерционностью процессов накопления и рассасывания зарядов в базе.

Время пролета носителей через базу $\tau_{пр}$ у обычных транзисторов составляет приблизительно 0,1 мкс. Конечно, это время очень мало, но на частотах порядка единиц - десятков мегагерц становится заметным некоторый сдвиг фаз между переменными составляющими токов $I_{э}$ и $I_{к}$. Это приводит к увеличению переменного тока базы и, как следствие, к снижению коэффициента усиления по току. Это явление иллюстрируется векторными диаграммами, приведенными на рис. 5.19.

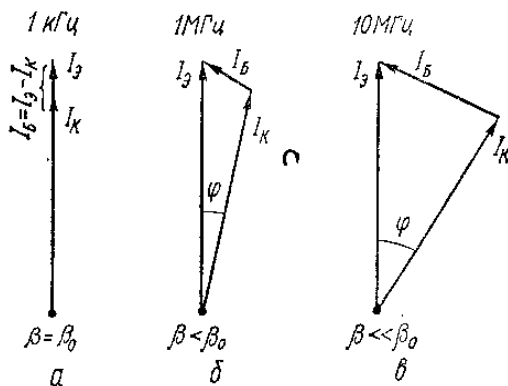


Рис. 5.19. Векторные диаграммы токов транзистора на разных частотах

Первая из них соответствует относительно низкой частоте, на которой все токи практически совпадают по фазе, а коэффициент β имеет наибольшую величину $\beta = \beta_0$. На более высокой частоте запаздывание тока I_K на время $\tau_{пр}$ относительно тока I_B ведет к появлению заметного сдвига фаз φ между этими токами. Теперь ток базы I_B равен не алгебраической (как на рис. 5.19, а), а геометрической разности токов I_B и I_K , вследствие чего он заметно увеличивается (рис. 5.19, б). На еще более высокой частоте коэффициент β становится еще меньше вследствие увеличения угла сдвига фаз φ и тока I_B (рис. 5.19, в).

Необходимо отметить, что с увеличением частоты коэффициент β уменьшается значительно сильнее, чем α . Коэффициент α снижается лишь вследствие влияния емкости C_K , а на величину β влияет, кроме этого, еще и сдвиг фаз между I_B и I_K . Следовательно, схема с общей базой имеет лучшие частотные свойства, чем схема с общим эмиттером.

Для определения коэффициентов усиления по току на частоте f могут быть использованы формулы

$$\alpha = \frac{\alpha_0}{\sqrt{1 + \left(\frac{f}{f_\alpha}\right)^2}}, \quad \beta = \frac{\beta_0}{\sqrt{1 + \left(\frac{f}{f_\beta}\right)^2}}. \quad (5.34)$$

где α_0 и β_0 - коэффициенты усиления по току при частоте $f = 0$; f_α и f_β - предельные частоты транзистора в схемах с общей базой и общим эмиттером соответственно.

Предельная частота усиления в схеме с ОБ - это частота, на которой модуль коэффициента передача тока эмиттера уменьшается в $\sqrt{2}$ раз.

Предельная частота усиления транзистора в схеме с ОЭ - это частота, на которой модуль коэффициента усиления тока базы уменьшается в $\sqrt{2}$ раз.

Граничная частота в схеме с ОЭ - это частота, на которой коэффициент усиления тока базы уменьшается до единицы.

Максимальная частота генерации - это частота, на которой коэффициент усиления по мощности становится равным единице.

Частотная характеристика определяет прохождение комплексных переменных токов через транзистор.

Амплитудно-частотная характеристика (АЧХ) представляет собой зависимость коэффициента передача тока от частоты (рис. 5.20).

Для расширения частотного диапазона транзисторов необходимо увеличивать скорость перемещения неосновных носителей зарядов через базу, уменьшать толщину слоя базы и коллекторную емкость. При выполнении этих условий транзисторы (например, дрейфовые, планарные) могут успешно работать на частотах порядка десятков и сотен мегагерц.

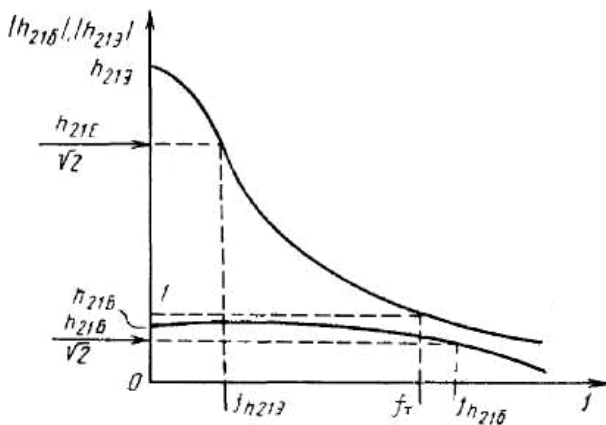


Рис. 5.20. Амплитудно-частотная характеристика

В заключение построим эквивалентную схему биполярного транзистора на высоких частотах для схемы с общей базой (рис. 5.21).

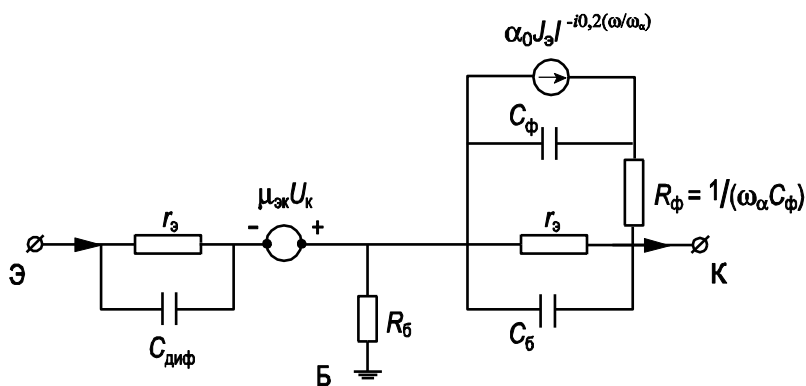


Рис. 5.21. Эквивалентная схема биполярного транзистора на высоких частотах для схемы с общей базой

На приведенной эквивалентной схеме основные параметры элементов в эмиттерной, базовой и коллекторной цепи такие же, как и для эквивалентной схемы при малых частотах. Различие этих двух схем проявляется в коллекторной цепи, где частотная зависимость коэффициента передачи $\alpha(\omega)$ изображена в виде фазосдвигающей RC -цепочки C_ϕ и R_ϕ в коллекторной цепи.

Процесс распространения инжектированных в базу неосновных носителей заряда от эмиттерного до коллекторного перехода идет диффузионным путем. Этот процесс достаточно медленный, и инжектированные из эмиттера носители достигнут коллектора не ранее чем за время диффузии носителей через базу, определяемое как $\tau_D = \frac{W}{v} \sim \frac{W \cdot L_p}{D}$. Такое запаздывание приведет к сдвигу фаз между током в эмиттерной и коллекторной цепях.

Переходная характеристика – это реакция системы на ступенчатый сигнал единичной амплитуды.

Процессы включения и выключения транзисторного ключа показаны на рис. 5.22. При включении транзистора (*a*) в его базу подается прямоугольный импульс тока с крутым фронтом. Ток коллектора достигает установившегося значения не сразу после подачи тока в базу. Имеется некоторое время задержки $t_{\text{зад}}$, спустя которое появляется ток в коллекторе. Затем ток в коллекторе плавно нарастает и после времени $t_{\text{нар}}$ достигает установившегося значения $I_{\text{к вкл}}$, таким образом время включения транзистора $t_{\text{вкл}} = t_{\text{зад}} + t_{\text{нар}}$.

При выключении транзистора (*б*) на его базу подается обратное напряжение, в результате чего ток базы меняет свое направление и становится равным $I_{\text{б вык}}$. Пока происходит рассасывание неосновных носителей заряда в базе, этот ток не меняет своего значения. Это время называется временем рассасывания $t_{\text{рас}}$. После окончания процесса рассасывания про-

исходит спад тока базы, который продолжается в течении времени $t_{\text{сп}}$. Таким образом время выключения транзистора равно $t_{\text{выкл}} = t_{\text{рас}} + t_{\text{сп}}$. Следует отметить, что при выключении транзистора, несмотря на изменение направления ток базы, транзистор в течение времени $t_{\text{рас}}$ остается включенным и коллекторный ток не меняет своего значения. Спад тока коллектора начнется одновременно со спадом тока базы и заканчиваются они практически одновременно.

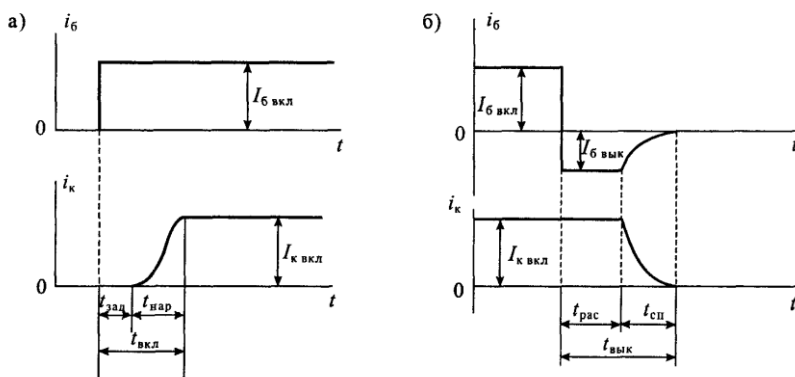


Рис. 5.22. Процессы при включении транзистора (а) и выключении (б)

6. ТИРИСТОРЫ

6.1. Структура и принцип действия

Тиристор – это полупроводниковый прибор с тремя и более p - n -переходами, вольт-амперная характеристика которого имеет участок с отрицательным дифференциальным сопротивлением, и который используется для переключения.

Структура тиристора показана на рис. 6.1. Тиристор представляет собой четырехслойный p_1 - n_1 - p_2 - n_2 прибор, содержащий три последовательно соединенных p - n -перехода (Π_1 , Π_2 и Π_3). Обе внешние области называют эмиттерами (\mathcal{E}_1 , \mathcal{E}_2), а внутренние области – базами (\mathcal{B}_1 , \mathcal{B}_2) тиристора (см. рис. 6.1, а). Переходы Π_1 и Π_2 называются эмиттерными, переход Π_3 – коллекторный переход.

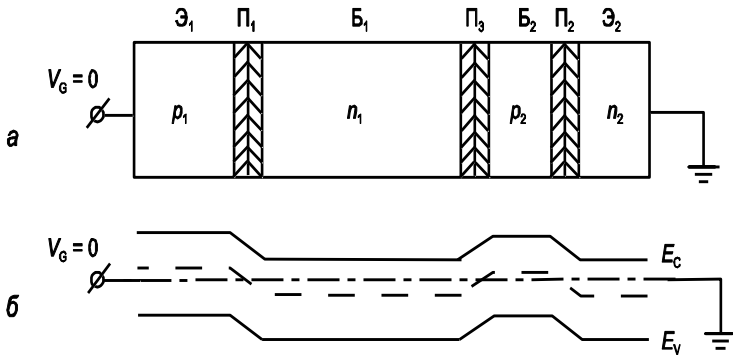


Рис. 6.1. Схема диодного тиристора:
а - структура диодного тиристора; б - зонная диаграмма

Тиристоры с двумя электродами называются диодными тиристорами или **динисторами**; с тремя электродами – триодными тиристорами или **тринисторами**, а чаще всего просто тиристорами.

Простейшие динисторы, запираемые в обратном направлении, обычно изготавливают из кремния и содержат четыре чередующихся p - и n -области. Область p_1 , в которую попадает ток из внешней цепи, называют **анодом**, область n_2 – **катодом**; области n_1 и p_2 – **базами**. Если на динистор подать напряжение $V_G = U$ плюсом на область p_1 (то есть на анод), минусом на область n_2 (то есть на катод), то переходы Π_1 и Π_2 , называемые эмиттерными, окажутся открытыми, а переход Π_3 (коллекторный) – закрытым. Практически все напряжение будет приложено к закрытому переходу Π_3 . И ток в цепи будет равен обратному току коллекторного перехода Π_3 , то есть будет иметь небольшое значение как и для любого кремниевого диода в запирающем направлении. При достаточно большом напряжении ($U = U_{пер}$) в переходе Π_3 возникает сильное электрическое поле, под действием которого начинается ударная ионизация, вызывающая лавинообразный рост тока и пробой перехода Π_3 . Сопротивление пробитого перехода Π_3 становится очень малым, поэтому напряжение на переходе Π_3 скачком снижается до значения $U \approx U_{отк} = (0,5 - 1,5)$ В, и ток в цепи динистора теперь определяется практически величиной сопротивления нагрузки и допустимой рассеиваемой мощностью динистора, т.е. динистор переходит в открытое состояние. Ток резко растет с увеличением напряжения U и может достигать очень больших значений.

Вольт-амперная характеристика (ВАХ) диодного тиристора, приведенная на рис. 6.2, имеет несколько различных участков. Прямое смещение тиристора соответствует положительному напряжению U , подаваемому на первый p_1 -эмиттер тиристора. Участок характеристики между точками O и A соответствует закрытому состоянию с высоким сопротивлением. В этом случае основная часть напряжения U падает на коллекторном переходе Π_3 , который смещен в обратном направлении. Эмиттерные переходы Π_1 и Π_2 включены в прямом направлении. Первый участок ВАХ тиристора аналогичен обратной ветви ВАХ p - n -перехода.

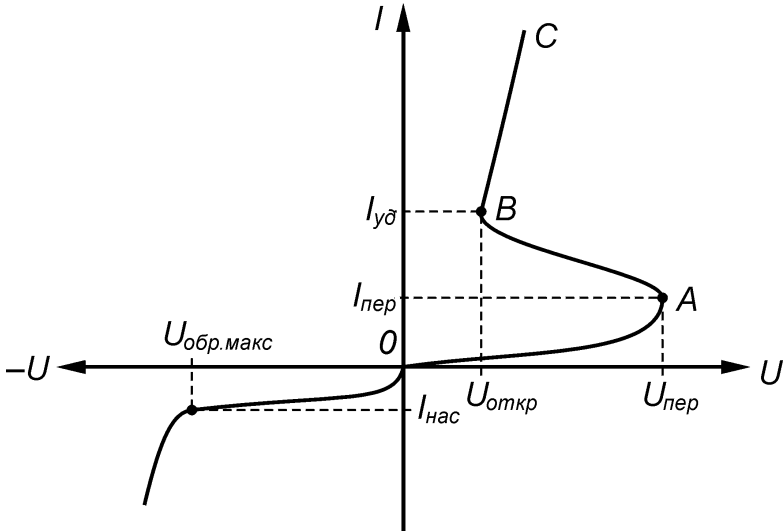


Рис. 6.2. Вольт-амперная характеристика динистора

При достижении напряжения U , называемого напряжением включения или переключения $U_{пер}$, или тока I , называемого током включения или переключения $I_{пер}$, ВАХ тиристора переходит на участок между точками B и C , соответствующий открытому состоянию (низкое сопротивление). Между точками A и B находится переходный участок характеристики с отрицательным дифференциальным сопротивлением, не наблюдаемый на статических ВАХ тиристора. Ток, соответствующий точке B , называют током удержания $I_{уд}$. Точка A соответствует прямому включению динистора, а величина $U_{пер}$ называется напряжением переключения.

Таким образом, динистор имеет две устойчивые области: область OA с координатами от $U = 0$ и $I = 0$ до $I < I_{пер}$ и $U < U_{пер}$ и область BC : от $I > I_{уд}$ и $U > U_{откр}$. Область, лежащая между точками A и B , является неустойчивой.

Прекратить ток через динистор можно либо снятием напряжения U , либо уменьшением тока I до значения $I < I_{уд}$. В

последнем случае динистор скачком переходит в режим работы, соответствующий участку OA , т.е. в режим закрытого состояния. При изменении полярности приложенного напряжения переходы P_1 и P_2 будут закрытыми и через динистор будет течь ток малой величины – ток насыщения $I_{нас}$ (микроамперы). Величина напряжения переключения зависит только от параметров областей p_1, n_1, p_2, n_2 и температуры окружающей среды (с ростом температуры $U_{пер}$ уменьшается).

На рис. 6.3 показана схема триодного тиристора с управляющими электродами и характеристики тиристора. Управляющий электрод может быть подключен к любой из баз (B_1, B_2) тиристора, как показано на рис. 6.3, *a*. Использование той или иной базы приводит лишь к изменению полярности источника управляющего напряжения, которая должна быть такой, чтобы облегчить включение тиристора. На управляющий электрод базы подается напряжение такой полярности, чтобы прилегающий к этой базе эмиттерный переход был включен в прямом направлении. Это приводит к росту тока через эмиттерный переход и снижению $U_{пер}$. При подаче в цепь управляющего электрода (на область p_2 , рис. 6.3, *a*) тока управления I_y ток через переход P_3 увеличивается. Дополнительная инжекция носителей заряда через p - n -переход приведет к увеличению тока коллектора на величину αI_y (где α – коэффициент передачи эмиттерного тока перехода P_3).

Увеличение тока через запертый коллекторный переход P_2 приблизительно равно увеличению приложенного напряжения, так как в любом из этих случаев увеличивается вероятность лавинного размножения носителей заряда. Поэтому с ростом I_y растет α и напряжение переключения тиристора $U_{пер}$ уменьшается (рис. 6.3, *d*); тиристор переходит из закрытого состояния в открытое при меньших напряжениях $U_{пер}$. После переключения тиристора в открытое состояние управляющий электрод теряет свои управляющие свойства.

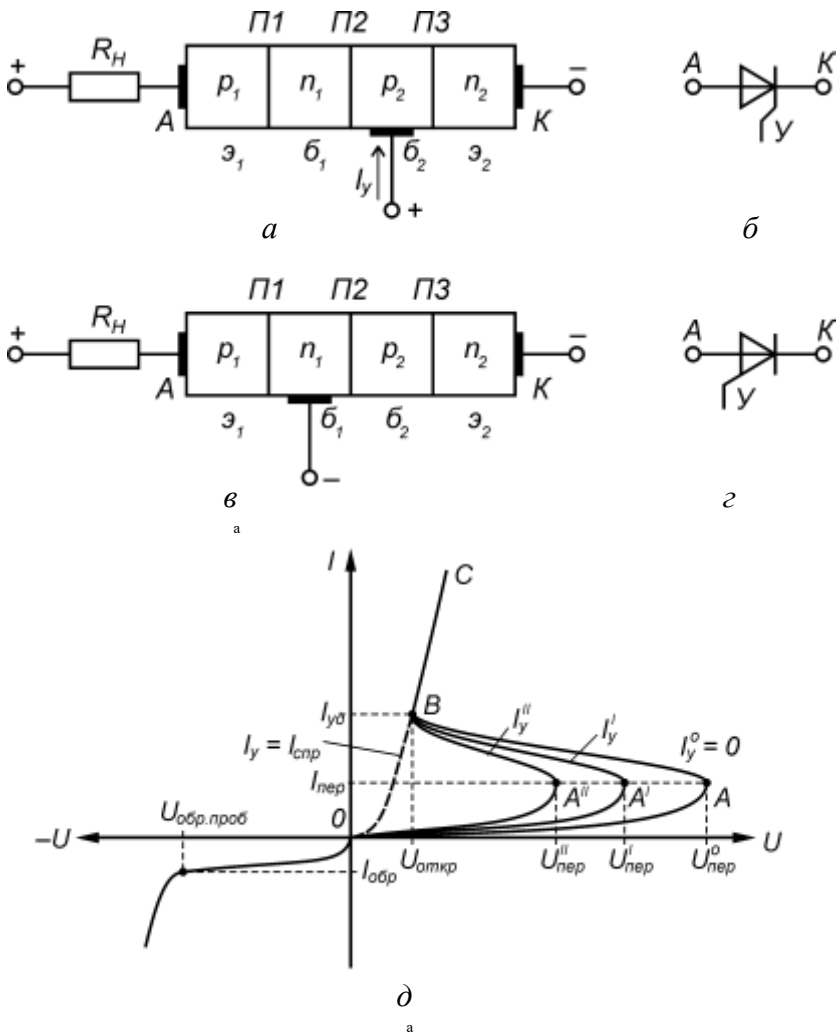


Рис. 6.3. Структурные изображения тиристоров и их условные обозначения с управлением по катоду (а, б) и по аноду (в, г). Вольтамперная характеристика тиристора (д). $I''_y > I'_y > I^o_y$

Для перевода тиристора в закрытое состояние необходимо или уменьшить напряжение на аноде до величины, при которой ток тиристора станет меньше тока включения, или подать на управляющий электрод напряжение обратной полярности. При $I_y = 0$ характеристика тиристора имеет такой же вид, как и у диода. Минимальный ток в открытом состоянии (при токе управляющего электрода $I_y = 0$) называют током удержания $I_{уд}$.

При достаточно больших значениях тока $I_{упр}$ ВАХ тиристора вырождается в прямую ветвь ВАХ диода. Критическое значение тока $I_{упр}$, при котором на ВАХ тиристора исчезает участок с отрицательным дифференциальным сопротивлением и тринистор включается, минуя запертое состояние, называется током спрямления $I_{спр}$.

Таким образом, наличие $I_{упр}$ принципиально не меняет существа процессов, определяющих вид ВАХ тиристора, но меняет значения параметров: напряжение переключения и ток переключения.

Тиристоры применяются в регулируемых источниках питания (например, в электроприводе), генераторах мощных импульсов, в линиях передачи электроэнергии постоянного тока, в системах автоматического управления и т.п.

Многослойные тиристорные структуры обычно создают в кремниевых монокристаллических пластинах введением в кремний примесей n - и p - типов. При этом в основном используют диффузионную и сплавную технологию или их комбинацию. Типичный профиль легирующей примеси в диффузионно-сплавном приборе показан на рис. 6.4. В качестве исходного материала выбрана подложка n -типа. Диффузией с обеих сторон подложки одновременно создают слои p_1 и p_2 . На заключительной стадии путем сплавления (или диффузии) с одной стороны подложки создают слой n_2 . Структура полученного тиристора имеет вид $p_1^+ - n_1 - p_2 - n_2^+$.

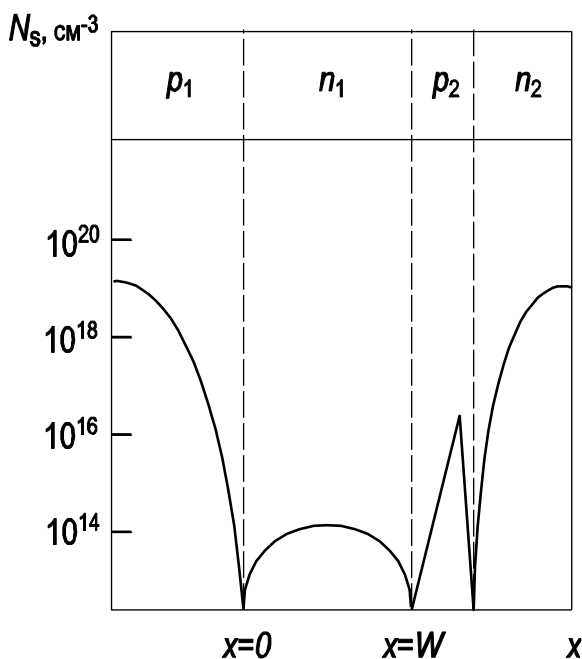


Рис. 6.4. Профиль концентрации легирующей примеси (N_s) в эмиттерах и базах тиристора

Материал и геометрию структуры тиристора выбирают исходя из его назначения. Для обеспечения высокого быстродействия тиристорov необходимо малая ширина базы и площадь переходов. Обеспечение высокого напряжения требует использование исходных материалов с высоким удельным сопротивлением. В силовых тиристорах, то есть мощных тиристорах, поперечные линейные размеры полупроводниковых структур равны 4 – 10 мм и более. Для повышения допустимой скорости нарастания тока при включении в силовом тиристорe используют так называемый разветвленный управляющий электрод, например, в форме снежинки, елочки и т.п. Существенное значение для силовых тиристорov имеет одно-

родность его электрофизических параметров по площади, что обуславливает высокие требования к качеству исходного кремния и технологии изготовления.

На рис. 6.5 представлены структуры тиристоров, изготовленных методом двойной двусторонней диффузии и методом планарной диффузии. Структура тиристора, полученного методом двойной двусторонней диффузии, используется для изготовления быстродействующих тиристоров, так как получается тонкая база, эмиттерные области (А и К) являются сильнолегированными диффузионными слоями, что обуславливает низкое сопротивление тиристора в открытом состоянии. Центральный переход – это переход с большим градиентом концентрации между диффузионным слоем и слоем исходного материала с малой концентрацией примесей, т.е. с высоким удельным сопротивлением, что обеспечивает высокое пробивное напряжение, а, следовательно, повышенное напряжение переключения.

При изготовлении планарных структур также используют метод двойной диффузии примесей.

Кристалл с готовой $p-n-p-n$ -структурой припаивают к кристаллодержателю и герметизируют в корпус обычно штырьевой или таблетной конструкции.

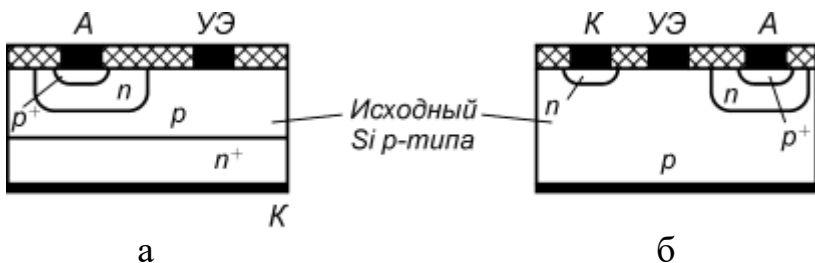


Рис. 6.5. Структуры тиристоров, полученные методом двойной двухсторонней диффузии (а) и методом планарной диффузии (б). УЭ – управляющий электрод

Управляющие тиристоры используются для коммутирования высоких значений токов, напряжений и мощностей. Поэтому корпуса тиристоров как правило являются достаточно массивными и в ряде случаев снабжены радиаторами для улучшения теплоотвода.

6.2. Основные параметры тиристоров

Преимуществом тиристоров перед транзисторами является, во-первых, низкое сопротивление включенного прибора, и, во-вторых, коэффициент усиления по мощности при включении тиристора достигает величины 10^5 .

Тиристоры характеризуются следующими основными электрическими параметрами:

Напряжение переключения $U_{пер}$ – прямое напряжение, при котором тиристор переходит из закрытого в открытое состояние.

Ток переключения $I_{пер}$ (иногда называют током включения $I_{вкл}$) – прямой ток, при котором тиристор переходит в открытое состояние, зачастую указывают при $I_{упр} = 0$.

Обратное напряжение $U_{обр}$ – напряжение, при котором тиристор может работать длительное время.

Обратный ток $I_{обр}$ – значение обратного тока при $U_{обр}$.

Напряжение в открытом состоянии $U_{отк}$ (иногда называют остаточным напряжением $U_{ост}$) – падение напряжения на тиристоре в открытом состоянии ($U_{отк} = U_{ост} = 1 - 3$ В).

Постоянное отпирающее напряжение на управляющем электроде тиристора $U_{y.от}$ – напряжение, подаваемое на управляющий электрод при постоянном прямом напряжении в закрытом состоянии, при котором тиристор переключается в открытое состояние.

Постоянный отпирающий ток управления $I_{y.от}$ – наименьший ток управляющего электрода, необходимый для

переключения тиристора в открытое состояние (имеет величину от единиц до десятков мА).

Ток удержания $I_{уд}$ – удерживающий ток тиристора – минимальное значение тока тиристора в открытом состоянии.

Ток спрямления $I_{спр}$ – отпирающий ток управления, когда S-образная прямая вольтамперная характеристика тиристора переходит в прямую ветвь диода (на рис. 6.3, δ ток показан пунктиром).

Ток в открытом состоянии I_{oc} – максимальное значение тока открытого состояния, ограничиваемое конструкцией и допустимой мощностью тиристора.

Время включения $t_{вкл}$ – время с момента подачи отпирающего импульса, когда напряжение на тиристоре начинает уменьшаться от уровня 0,9, до момента, когда напряжение уменьшится до уровня 0,1 своего максимального значения.

Время выключения $t_{выкл}$ – время, в течение которого тиристор из открытого состояния переходит в закрытое состояние.

В настоящее время разрабатываемые тиристоры различаются по мощности: маломощные, рассчитанные на ток до 100 мА, средней мощности – на ток до 20 А и мощные – на ток от 20 А до 10 кА. Тиристоры изготавливают на напряжения от нескольких вольт до 10 кВ и выше; скорость нарастания прямого тока и напряжения в них достигает 10^9 А/с и 10^9 В/с; время включения составляет от десятых долей микросекунд (у маломощных тиристоров) до десятков микросекунд (у мощных); время выключения – соответственно от единиц до сотен микросекунд; коэффициент полезного действия достигает 99 %.

6.3. Феноменологическое описание ВАХ динистора

Для объяснения ВАХ динистора используют двухтранзисторную модель. Из рис. 6.6 следует, что тиристор можно

рассматривать как соединение p - n - p транзистора с n - p - n транзистором, причем коллектор каждого из них соединен с базой другого. Центральный переход действует как коллектор дырок, инжектируемых переходом Π_1 , и как коллектор электронов, инжектируемых переходом Π_2 .

Взаимосвязь между токами эмиттера $I_э$, коллектора $I_к$ и статическим коэффициентом усиления по току α_1 p_1 - n_1 - p_2 транзистора и α_2 n_2 - p_1 - n_1 транзистора следующая.

Пусть I_{Π_1} – ток через переход Π_1 . Тогда часть тока I_{Π_1} , дошедшая до коллекторного перехода Π_3 $I_{\Pi_1 \rightarrow \Pi_3}$, будет равна:

$$I_{\Pi_1 \rightarrow \Pi_3} = \alpha_1 I_{\Pi_1}. \quad (6.1)$$

Если I_{Π_3} – ток через переход Π_2 , аналогично:

$$I_{\Pi_2 \rightarrow \Pi_3} = \alpha_2 I_{\Pi_3}. \quad (6.2)$$

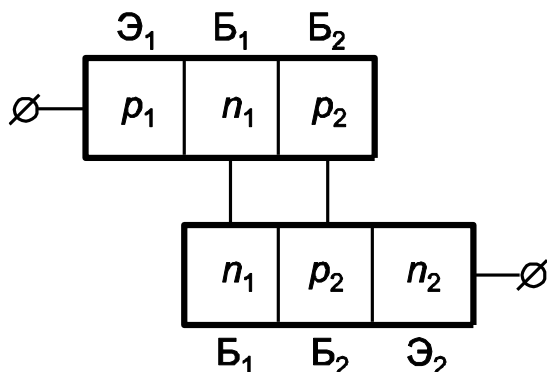


Рис. 6.6. Двухтранзисторная модель диодного тиристора

Учтем еще один фактор – лавинное умножение в переходе Π_3 через коэффициент лавинного умножения M . Тогда суммарный ток I_{Π_3} через переход Π_3 будет равен:

$$I_{\Pi_3} = M(\alpha_1 I_{\Pi_1} + \alpha_2 I_{\Pi_2} + I_{\text{к0}}), \quad (6.3)$$

где $I_{\text{к0}}$ – обратный ток перехода Π_3 (генерационный и тепловой).

В стационарном случае токи через переходы Π_1 , Π_2 , и Π_3 равны, тогда

$$I = M(\alpha_1 I + \alpha_2 I + I_{\text{к0}}), \quad (6.4)$$

откуда

$$I = \frac{MI_{\text{к0}}}{1 - M\alpha}; \quad I = \frac{MI_{\text{к0}}}{1 - M(\alpha_1 + \alpha_2)}, \quad (6.5)$$

где $\alpha = \alpha_1 + \alpha_2$ – суммарный коэффициент передачи тока первого ($p_1-n_1-p_2$) и второго ($n_2-p_2-n_1$) транзисторов.

Выражение (6.5) в неявном виде описывает ВАХ диодного тиристора на «закрытом» участке, поскольку коэффициенты M и α зависят от приложенного напряжения V_G . По мере роста α и M с ростом V_G , когда значение $M(\alpha_1 + \alpha_2)$ станет равно 1, из уравнения (6.5) следует, что ток I устремится к ∞ . Это условие и есть условие переключения тиристора из состояния «закрыто» в состояние «открыто».

Напряжение переключения $U_{\text{пер}}$ составляет у тиристорov от 20 - 50 В до 1000 - 2000 В, а ток переключения $I_{\text{пер}}$ – от долей микроампера до единиц миллиампера (в зависимости от площади).

Таким образом, в состоянии «закрыто» тиристор должен характеризоваться малыми значениями α и M , а в состоянии «открыто» – большими значениями коэффициентов α и M .

В закрытом состоянии (α – малы) все приложенное напряжение падает на коллекторном переходе Π_3 и ток тиристора – это ток обратного смещенного p - n -перехода. Энергетическая диаграмма тиристора в состоянии равновесия приведена ранее на рис. 6.1, а в режиме прямого смещения («+» на слое p_1) в закрытом состоянии представлена на рис. 6.7.

Если полярность напряжения между анодом и катодом сменить на обратную, то переходы Π_1 и Π_2 будут смещены в обратном направлении, а Π_3 – в прямом. ВАХ тиристора в этом случае будет обычная ВАХ двух обратносмещенных p - n -переходов.

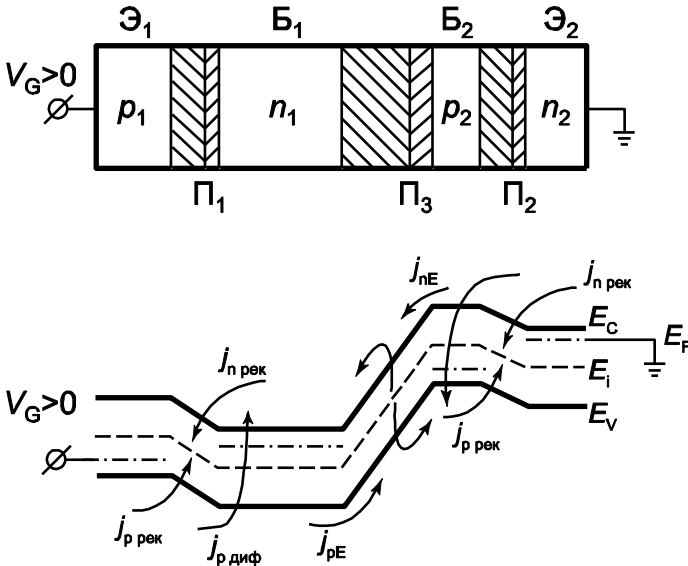


Рис. 6.7. Зонная диаграмма и токи в тиристоре в закрытом состоянии

В открытом состоянии (α – велики) все три перехода смещены в прямом направлении. Это происходит вследствие накопления объемных зарядов в базах n_2 , p_2 тиристора.

Действительно, при больших значениях коэффициента передачи α_2 электроны, инжектированные из n_2 -эмиттера в p_2 -базу, диффундируют к p - n -переходу коллектора Π_3 , проходят его и попадают в n_1 -базу. Дальнейшему прохождению электронов по тиристорной структуре препятствует потенциальный барьер эмиттерного перехода Π_1 . Поэтому часть электронов, оказавшись в потенциальной яме n_1 -базы, образует отрицательный избыточный заряд.

Инжектированные дырки из эмиттера p_1 в базу n_1 диффундируют к p - n -переходу коллектора Π_3 , проходят через него и попадают в базу p_2 . Дальнейшему их продвижению препятствует потенциальный барьер эмиттерного перехода Π_2 . Следовательно, в базе p_2 происходит накопление избыточного положительного заряда. В результате накопления избыточного положительного заряда в базе p_2 и отрицательного заряда в базе n_1 переход Π_3 смещается в прямом направлении, происходит резкое увеличение тока и одновременное уменьшение падения напряжения на тиристоре.

Зонная диаграмма тиристора в открытом состоянии имеет вид, приведенный на рис. 6.8, когда на всех p - n -переходах прямое смещение, на Π_1 и Π_2 за счет внешнего напряжения, и на Π_3 за счет объемных зарядов в базах B_1 и B_2 .

Величина падения напряжения в прямом участке ВАХ определяется прямым напряжением на трех прямосмещенных p - n -переходах и имеет величину порядка 1 - 2 В.

Таким образом, тиристор имеет два устойчивых состояния: малый ток, большое напряжение, высокое сопротивление и большой ток, малое напряжение, малое сопротивление. Переход тиристора из «закрытого» в «открытое» состояние связан с накоплением объемного заряда в базах B_1 и B_2 из-за ро-

ста значения коэффициента передачи эмиттерного тока α и коэффициента умножения M .

То есть рост α , M с ростом тока и напряжения в тиристоре является причиной перехода тиристора из состояния «закрытого» в состояние «открытого».

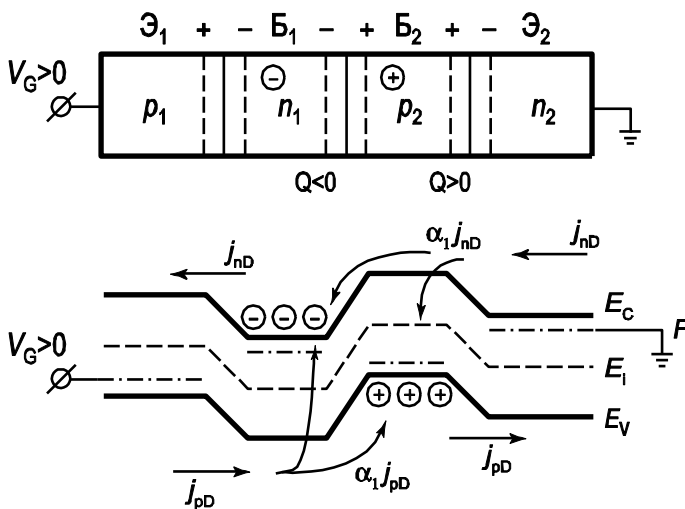


Рис. 6.8. Зонная диаграмма и токи тиристора в открытом состоянии

В открытом состоянии тиристор находится до тех пор, пока за счет проходящего тока поддерживаются избыточные заряды в базах, необходимые для понижения высоты потенциального барьера коллекторного перехода до величины, соответствующей прямому его включению. Если же ток уменьшить до значения $I_{уд}$, то в результате рекомбинации избыточные заряды в базах уменьшатся, p - n -переход коллектора окажется включенным в обратном направлении, произойдет перераспределение падений напряжений на p - n -переходах,

уменьшатся коэффициенты передачи α и тиристор перейдет в закрытое состояние.

Таким образом, тиристор в области прямых смещений (прямое включение) является бистабильным элементом, способным переключаться из закрытого состояния с высоким сопротивлением и малым током в открытое состояние с низким сопротивлением и большим током, и наоборот.

6.4. Способы включения и выключения тиристор

Переключение тиристора из закрытого состояния в открытое осуществляется с помощью введения тем или иным способом в одну или обе его базовые области неравновесных носителей заряда (то есть накопление избыточного отрицательного заряда электронов в n_1 -базе, и избыточного положительного заряда дырок в p_2 -базе) достаточного для выполнения условий включения.

Способами включения тиристора из закрытого состояния в открытое являются:

- повышение напряжения, приложенного к тиристор
- вплоть до величины напряжения переключения – включение по аноду;
- включение с помощью тока управления (подача сигнала на управляющий электрод);
- включение тиристора под действием теплоты (нагрев тиристорной структуры);
- включение тиристора под действием света;
- включение тиристора путем быстрого увеличения анодного напряжения на нем (переключение по dU/dt).

Практическое применение нашли включение с помощью тока управления и включение тиристора под действием света.

Сущность всех способов выключения заключается в следующем: при работе в открытом состоянии в базах тиристора накапливается большой избыточный заряд неравновес-

ных носителей, чтобы тиристор выключился необходимо, чтобы накопленный заряд в базах уменьшился до значения не превышающего критический заряд включения

Выключение тиристора может быть осуществлено следующими способами:

- путем уменьшения прямого тока до нуля или смены полярности анодного напряжения (выключение по основной цепи);

- подача на управляющий электрод обратного тока;

- смена полярности анодного напряжения с одновременной подачей в p_2 – базу обратного тока (комбинированное выключение);

- путем снятия прямого напряжения или путем разрыва цепи

Наиболее распространен способ выключение по основной цепи.

7. ПОЛЕВЫЕ ТРАНЗИСТОРЫ И ПРИБОРЫ С ЗАРЯДОВОЙ СВЯЗЬЮ

Во многих современных электронных цепях используются транзисторы, ток носителей которых течет по так называемому каналу, образованному внутри кремниевого кристалла. Этим током можно управлять, прикладывая электрическое поле. Такие приборы называют полевыми транзисторами (в англоязычной литературе применяют сокращение FET - Field Effect Transistor). В настоящее время эти транзисторы играют весьма важную роль, являясь элементами ИС, которые содержат на одном кристалле от тысяч до сотен тысяч полупроводниковых приборов. В свою очередь, на базе таких ИС создают компьютеры, микропроцессорные системы, устройства обработки сигналов и т.д.

Полевой транзистор (ПТ)– это полупроводниковый прибор, в котором изменение тока происходит под действием перпендикулярного току электрического поля, создаваемого входным сигналом.

Протекание электрического тока в полевых транзисторах обусловлено носителями заряда только одного знака (электронами или дырками), поэтому такие транзисторы называют также **униполярными** в отличие от биполярных.

По физическим эффектам, лежащим в основе управления носителями заряда, полевые транзисторы бывают трех видов: с управляющим *p-n*-переходом, с управляющим переходом металл полупроводник и со структурой металл–диэлектрик–полупроводник (МДП-транзисторы). Приборы этого типа в зарубежной литературе носят разные названия, в том числе MOSFET (Metal - Oxide - Semiconductor Field - Effect Transistor, то есть полевой транзистор металл - окисел - полупроводник), IGFET (Insulated-Gate Field - Effect Transistor, то есть полевой транзистор с изолированным затвором), MISFET (Metal - Insulator - Semiconductor Field - Effect Transistor).

sistor, то есть полевой транзистор металл - изолятор - полупроводник), MOST (Metal - Oxide - Semiconductor Transistor, то есть полевой транзистор металл - окисел - полупроводник), JFET (Junction Field - Effected Transistor, то есть полевой транзистор с управляющим *p-n*-переходом), MESFET (Metal - Semiconductor Field - Effected Transistor, то есть полевой транзистор с управляющим переходом металл - полупроводник). В полевых транзисторах в качестве полупроводникового материала используют в основном кремний и арсенид галлия, в качестве металлов: алюминий, молибден, золото; в качестве диэлектрика оксид кремния SiO_2 в МОП-транзисторах или сложные структуры, например $\text{SiO}_2\text{-Al}_2\text{O}_3$, $\text{SiO}_2\text{-Si}_3\text{N}_4$ в МДП-транзисторах.

Полевые транзисторы проще биполярных по структуре, кроме того, им присущ ряд ценных качеств:

- производство этих приборов проще, они имеют меньшие габариты и можно добиться более высокой степени интеграции ИС;

- потребляемая ими мощность меньше, чем у биполярных транзисторов (мощность, потребляемая МОП - транзисторами, составляет единицы нановатт, в то время как биполярные транзисторы потребляют единицы милливатт);

- применение полевых транзисторов улучшает экономические показатели изделий;

- характерной особенностью полевых транзисторов является высокое входное сопротивление (свыше 10 МОм) и высокий коэффициент усиления по напряжению;

- на базе полевых транзисторов легко создавать запоминающие устройства, работающие за счет накопления зарядов малыми внутренними емкостями;

- надежность полевых транзисторов выше надежности биполярных.

В области высоких токов полевые транзисторы имеют отрицательный температурный коэффициент, то есть ток в

данных приборах уменьшается с повышением температуры, благодаря этому возникает более однородное распределение температуры по площади прибора и снижается вероятность развития теплового или вторичного пробоя, характерного для биполярных транзисторов. Поскольку полевые транзисторы являются униполярными приборами, они не чувствительны к эффектам накопления неосновных носителей, и поэтому имеют более высокие граничные частоты и скорости переключения.

7.1. Полевой транзистор с управляющим *p-n*-переходом

*Полевой транзистор с управляющим *p-n*-переходом* – полупроводниковый прибор, в котором ток основных носителей заряда управляется поперечным электрическим полем обратнo смещенного *p-n*-перехода (или переходов).

Простейший полевой транзистор с управляющим *p-n*-переходом состоит из полупроводниковой пластины, имеющей электропроводность определенного типа, от концов которой сделаны два вывода – электроды стока (С) и истока (И), с одним (рис. 7.1, *а*) или двумя *p-n*-переходами (рис 7.1, *б*) от которых сделан третий вывод – затвор (З).

Электрод, от которого начинают движение основные носители заряда в канале, называют *истоком*, а электрод, к которому движутся основные носители заряда, называют *стоком*. Условные обозначения даны на рис. 7.1, *з, д*, а одна из структур выпускаемых промышленностью полевых транзисторов на рис. 7.1, *в*. Действие этих транзисторов основано на зависимости толщины области пространственного заряда (ОПЗ) *p-n*-перехода от приложенного к нему напряжения.

При включении между истоком и стоком транзистора источника напряжения E_1 (или $E_{СИ}$) по каналу от истока к стоку потечет ток основных носителей (в данном случае – электронов), величина которых определяется приложенным

напряжением и сопротивлением канала. Если на затвор транзистора подать напряжение E_2 (или $E_{3И}$) так, чтобы p - n -переход (или переходы) оказались смещенными в обратном направлении, то переход (переходы) расширяясь, уменьшают ширину канала W . Уменьшение поперечного сечения канала приводит к увеличению его сопротивления и поэтому к уменьшению протекающего по каналу тока.

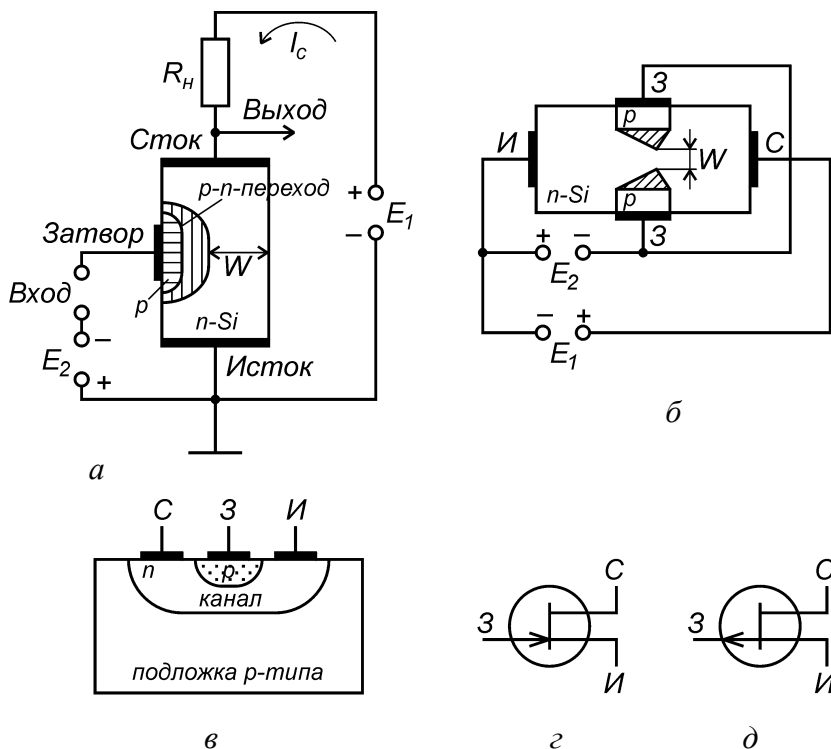


Рис. 7.1. Упрощенная структура полевого транзистора с одним управляющим p - n -переходом (а); с двумя управляющими p - n -переходами (б); типовая структура (в) и условные обозначения транзистора, имеющего канал n -типа (г) и p -типа (д)

При приложении напряжения $U_{ЗИ}$ или V_{GS} к затвору ПТ, обеспечивающего обратное смещение p - n -перехода ($U_{ЗИ} < 0$), происходит расширение обедненной области p - n -перехода в полупроводниковую подложку, поскольку затвор легирован существенно сильнее, чем подложка ($N_D \gg N_A$). При этом уменьшается поперечное сечение канала и, следовательно, увеличивается его сопротивление. Приложенное напряжение исток-сток $U_{СИ}$ или V_{DS} вызовет ток в цепи канала полевого транзистора. Знак напряжения $U_{СИ}$ необходимо выбирать таким образом, чтобы оно также вызывало обратное смещение затворного p - n -перехода, то есть было бы противоположно по знаку напряжению $U_{ЗИ}$. Таким образом, полевой транзистор с затвором в виде p - n -перехода представляет собой сопротивление, величина которого регулируется внешним напряжением.

При возрастании обратного смещения на затворе обедненный слой (слои) p - n -перехода все больше расширяются, ток стока I_C постепенно уменьшается и при достаточно высоком напряжении $U_{ЗИ}$ происходит смыкание области объемного заряда с краем кристалла (для полевого транзистора с одним p - n -переходом) или смыкание областей объемного заряда (для полевого транзистора с двумя p - n -переходами). Напряжение, при котором канал полностью перекрывается и ток через него практически прекращается, называют **напряжением отсечки** $U_{отс}$. При этом напряжении остаются лишь токи утечки.

Зависимость тока стока I_C от напряжения стока $U_{СИ}$ при постоянном напряжении на затворе ($U_{ЗИ} = const$) определяется стоковой характеристикой (рис. 7.2). Ширина p - n -перехода зависит также от тока, протекающего через канал.

Если затвор закорочен с истоком ($U_{ЗИ} = 0$), а на сток подается положительное относительно истока напряжение, то падение напряжения в канале при протекании по нему потока электронов приводит к тому, что ширина p - n -перехода увеличивается в направлении от истока к стоку, а ширина токопро-

водящего канала W соответственно уменьшается, то есть ширина обедненного слоя в канале не одинакова по длине канала.

При увеличении напряжения стока $U_{СИ}$ (при $U_{ЗИ} = 0$) увеличивается поток электронов к стоку, то есть происходит рост тока стока I_C . Но это в свою очередь приводит к увеличению падения напряжения в канале, расширению обедненной носителями области $p-n$ -переходов, дальнейшему сужению канала и росту его сопротивления. В итоге рост тока I_C с увеличением напряжения $U_{СИ}$ вначале замедляется, а затем практически прекращается. Дальнейшее повышение напряжения на стоке уже не приводит к росту тока I_C , так как одновременно с повышением $U_{СИ}$ увеличивается и сопротивление канала, наступает своеобразное равновесие, при котором увеличение $U_{СИ}$ и рост тока I_C вызывают дальнейшее сужение канала и соответственно уменьшение тока I_C . В результате этот ток остается практически постоянным. Напряжение, при котором наступает режим насыщения, называется **напряжением насыщения**.

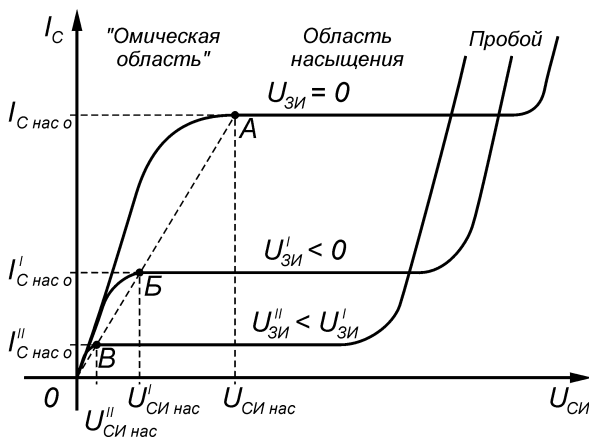


Рис. 7.2. Выходные характеристики полевого транзистора с управляющим $p-n$ -переходом

Оно, как видно из рис. 7.2, меняется при изменении напряжения на затворе. Область, в которой ток I_C существенно зависит от напряжения на стоке $U_{СИ}$ (слева от пунктирной линии АБВ, рис. 7.2), называется *омической областью*; правая область, в которой ток I_C остается практически неизменным – *областью насыщения*.

Характеристики передачи полевого транзистора представляют собой зависимости тока стока от напряжения на затворе при различных постоянных напряжениях на стоке (рис. 7.3).

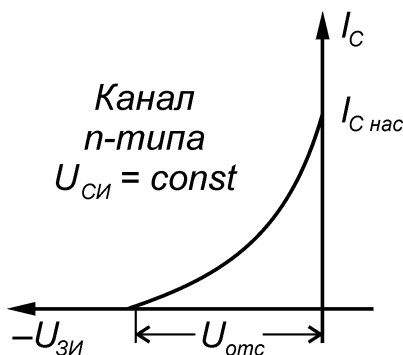


Рис. 7.3. Статическая характеристика передачи полевого транзистора с управляющим p - n -переходом

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 можно пренебречь в связи с малым изменением

ем тока стока в пологой части выходных статических характеристик.

Одним из основных параметров полевого транзистора с управляющим *p-n*-переходом является **крутизна передаточной характеристики**, определяющая эффективность изменения тока в цепи стока при изменении напряжения на затворе.

$$S = \left. \frac{dI_c}{dU_{3И}} \right|_{U_{сИ} = \text{const}} \quad (7.1)$$

Численно крутизна зависит от напряжения на затворе. С увеличением напряжения на затворе $|U_{3И}|$ ток стока и крутизна характеристики уменьшается. Максимальные значения крутизны характеристики при $U_{3И} = 0$ в реальных транзисторах равны порядку 20 мА/В.

К параметрам таких транзисторов относятся также **напряжение отсечки** $U_{отс}$; **ток насыщения стока** $I_{C_{нас\ o}}$ при $U_{3И} = 0$; **емкости** $C_{СИ}$ и $C_{3И}$; **сопротивление транзистора в омической области** R_o , определяемое наклоном стоковой характеристики; **внутренне сопротивление**

$$R_i = dU_c / dI_c, \quad (7.2)$$

определяемое в области насыщения при различных напряжениях на затворе $U_{3И}$; а также **граничная частота**.

Величины $I_{C_{нас\ o}}$ и $U_{отс}$ определяются геометрией транзистора, параметрами полупроводникового материала и распределением примесей в управляющем *p-n*-переходе.

Так как подвижность электронов выше подвижности дырок, то частотные свойства транзистора с *n*-каналом лучше, чем с *p*-каналом. По этой же причине транзисторы с *n*-

каналом имеют большой ток стока I_C и большую крутизну передаточной характеристики S .

Схемы включения полевых транзисторов определяются по электроду, общему для входной и выходной цепей. Различают схемы включения: с общим истоком (ОИ), с общим стоком (ОС) и с общим затвором (ОЗ). Схема включения с общим истоком (соответствует схеме включения с общим эмиттером для биполярных транзисторов) характеризуется наибольшим усилением по мощности и напряжению по сравнению с другими схемами включения.

Исходя из принципа действия и структуры полевого транзистора с управляющим p - n -переходом, можно составить его эквивалентную схему для низких частот (рис. 7.4).

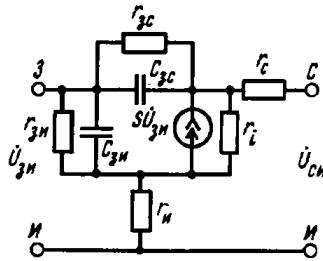


Рис. 7.4. Физическая эквивалентная схема полевого транзистора с управляющим p - n -переходом

Сопротивления r_c и $r_и$ 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 $r_и$. Общее для входной и

выходной цепей сопротивление $r_{и}$ 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 $C_{зи}$ и $C_{зс}$, сопротивления $r_{зи}$ и $r_{зс}$ 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 входному напряжению $U_{зи}$, коэффициентом пропорциональности является крутизна характеристики S .

7.2. МДП-транзисторы

Ко второй группе полевых транзисторов относятся приборы с изолированным затвором, так называемые МДП-транзисторы. У МДП-транзисторов в отличие от полевых транзисторов с управляющим p - n -переходом металлический затвор изолирован от полупроводника слоем диэлектрика и имеется дополнительный вывод от кристалла, на котором выполнен прибор, называемый подложкой (рис. 7.5).

В этих транзисторах управляющий электрод, т.е. затвор, отделен от канала тонким слоем диэлектрика (0,05 – 0,20 мкм) и допускает приложения напряжения любой полярности. Если в качестве диэлектрика применен оксид кремния, то эти транзисторы называются МОП-транзисторы (то есть со структурой металл–оксид–полупроводник).

МДП-транзисторы по способу образования канала подразделяются на транзисторы со встроенным каналом (канал создается при изготовлении) и с индуцированным каналом

(канал возникает под действием напряжения, приложенного к управляющим электродам). Важными параметрами МДП-транзистора являются длина канала L (обычно 0,1 – 5 мкм), ширина канала z (на рис. 7.5, б), толщина диэлектрика (изолятора) d с диэлектрической постоянной ε_i и проводимость канала.

Существенным преимуществом МДП-транзисторов является высокое входное сопротивление, достигающее значений более 10^4 МОм вместо более 10 МОм у транзисторов с управляющим p - n -переходом.

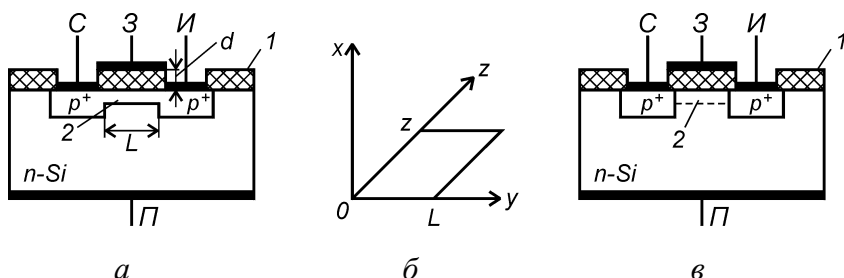


Рис. 7.5. МДП-транзистор с каналом p -типа:

- a – планарный транзистор со встроенным каналом;
- $б$ – геометрические размеры канала;
- $в$ – планарный транзистор с индуцированным каналом;
- 1 – диэлектрик; 2 – канал

Физической основой работы полевого транзистора со структурой металл – диэлектрик – полупроводник (МДП) является эффект поля. Напомним, что эффект поля состоит в том, что под действием внешнего электрического поля изменяется концентрация свободных носителей заряда в приповерхностной области полупроводника.

В полевых приборах со структурой МДП внешнее поле обусловлено приложенным напряжением на металлический электрод – затвор. В зависимости от знака и величины прило-

женного напряжения присутствуют четыре состояния области пространственного заряда (ОПЗ) полупроводника – обогащение, обеднение, слабая и сильная инверсия.

Полевые транзисторы в активном режиме могут работать только в области слабой или сильной инверсии, то есть в том случае, когда инверсионный канал между истоком и стоком отделен от объема подложки слоем обеднения. На рис. 7.6 приведена топология МДП - транзистора, где этот факт наглядно виден.

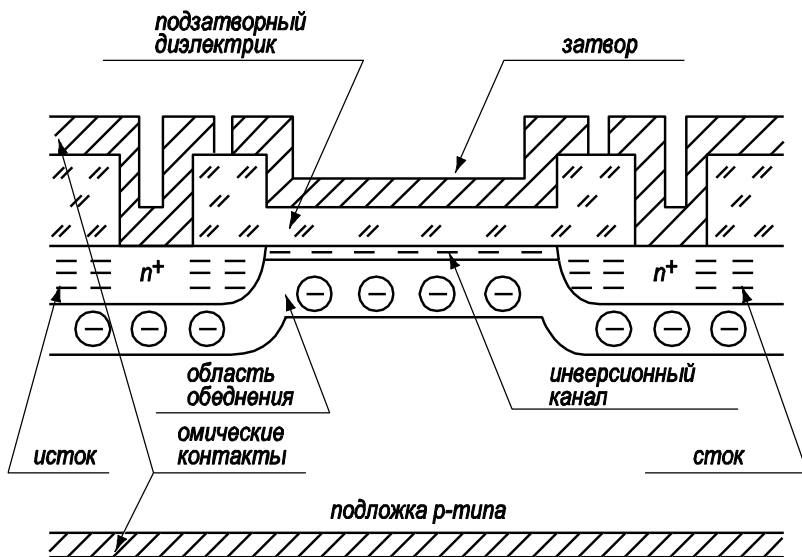


Рис. 7.6. Полевой транзистор со структурой металл – диэлектрик – полупроводник

В области инверсии концентрация неосновных носителей заряда в канале выше, чем концентрация основных носителей в объеме полупроводника. Изменяя величину напряжения на затворе, можно менять концентрацию свободных носителей в инверсионном канале и тем самым модулировать со-

противление канала. Источник напряжения в стоковой цепи вызовет изменяющийся в соответствии с изменением сопротивления канала ток стока и тем самым будет реализован эффект усиления. Таким образом, МДП-транзистор является сопротивлением, регулируемым внешним напряжением. К нему даже в большей степени, чем к биполярным приборам, подходит историческое название «транзистор», так как слово «transistor» образовано от двух английских слов – «transfer» и «resistor», что переводится как «преобразующий сопротивление».

МДП-транзисторы могут быть с p - и n -каналами. Условные графические обозначения МДП-транзисторов показаны на рис. 7.7. Стрелка, направленная во внутрь круга, указывает на n -тип канала; стрелка из круга - на канал p -типа.

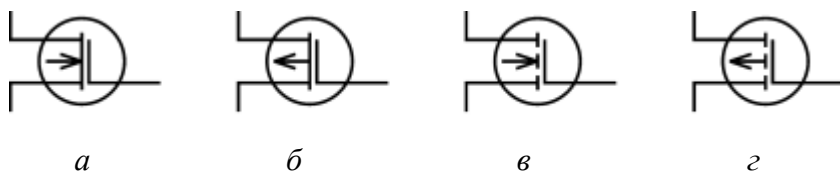


Рис. 7.7. Условные графические обозначения МДП-транзисторов со встроенным каналом (а, б) и с индуцированным каналом (в, г)

МДП-транзисторы характеризуются в основном теми же параметрами, что и полевые транзисторы с управляющим p - n -переходом. Основные параметры МДП-транзисторов и их ориентировочные значения следующие:

а) **крутизна характеристики.** Крутизна переходной характеристики S характеризуется изменением тока стока при единичном увеличении напряжения на затворе при постоянном напряжении на стоке и определяется как

$$S = dI_C / dU_{3И} \text{ при } U_{СИ} = const. \quad (7.3)$$

Обычно $S = 0,1 - 500 \text{ мА/В}$;

б) **внутреннее сопротивление** R_i характеризует изменение напряжения в выходной цепи, необходимое для единичного увеличения тока стока при неизменном напряжении на затворе и определяется как

$$R_i = \frac{dU_{\text{СЭ}}}{dI_C}, \quad U_{\text{СЭ}} = const; \quad (7.4)$$

в) коэффициент усиления μ характеризуется изменением напряжения в выходной цепи при единичном изменении напряжения во входной и неизменном токе стока и определяется как

$$\mu = \frac{dU_{\text{НЭ}}}{dU_{\text{СЭ}}}, \quad I_{\text{Н}} = const. \quad (7.5)$$

г) **начальный ток стока** $I_{C \text{ нач}}$ – ток стока при нулевом напряжении $U_{3И}$ у транзисторов со встроенным каналом $I_{C \text{ нач}} = 0,1 - 100 \text{ мА}$; с индуцированным каналом $I_{C \text{ нач}} = 0,01 - 0,5 \text{ мкА}$;

д) **напряжение отсечки** $U_{3И \text{ отс}} = 0,2 - 10 \text{ В}$;

е) **пороговое напряжение** $U_{3И \text{ пор}} = 1 - 6 \text{ В}$;

ж) **сопротивление сток–исток в открытом состоянии** $R_{СИ \text{ отк}} = 2 - 300 \text{ Ом}$;

з) **максимальный ток стока** $I_{C \text{ max}} = 10 \text{ мА} - 0,3 \text{ А}$;

и) **максимальная частота усиления** f_{max} – частота, на которой коэффициент усиления по мощности K_P равен единице; f_{max} – десятки – сотни МГц.

7.2.1. МДП-транзисторы со встроенным каналом

Устройство транзистора со встроенным p -каналом схематично представлено на рис. 7.5, *a*. Основой является пластина слаболегированного кремния с электропроводностью n -типа. Области стока и истока обладают электропроводностью p^+ -типа и соединены узкой слаболегированной областью кремния с электропроводностью p -типа, являющейся встроенным каналом. Затвор представляет собой металлический слой, изолированный от канала тонким диэлектриком толщиной d . В транзисторе со встроенным каналом происходит расширение или сужение имеющегося канала в результате приложения отрицательного или положительного напряжения соответственно. При положительном напряжении на затворе дырки оттесняются из области канала в объем полупроводника. Канал обедняется носителями заряда, сужается к стоку. При подаче на затвор отрицательного напряжения происходит обогащения дырками объема канала, он расширяется и его проводимость возрастает. Таким образом, изменяя напряжение на затворе, можно изменять проводимость канала, то есть ток, проходящий через этот канал.

В транзисторах со встроенным каналом ток в цепи стока будет протекать и при нулевом напряжении на затворе (рис. 7.8).

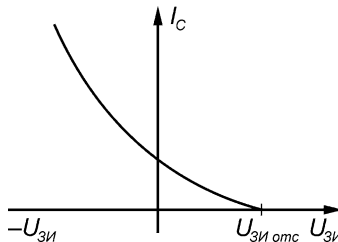


Рис. 7.8. Стокзатворная характеристика МДП-транзистора со встроенным каналом

Для прекращения его необходимо к затвору приложить положительное напряжение (для транзистора с каналом p -типа), равное напряжению $U_{зи\text{ отс}}$. При приложении отрицательного напряжения канал расширяется и ток увеличивается.

Выходные характеристики МДП-транзисторов аналогичны характеристикам полевых транзисторов с управляющим p - n -переходом (рис. 7.9). В них можно выделить крутую (I) и пологую области (II), а также область пробоя (III). В области I транзистор может работать как электрически управляемое сопротивление. Пологая область II обычно используется при построении усилительных каскадов. При увеличении напряжения $U_{СИ}$ ширина канала уменьшается вследствие подъема на нем напряжения и изменения результирующего электрического поля. Это особенно сильно проявляется в той части канала, которая находится вблизи стока. При напряжении $U_{СИ\text{ нас}}$ канал вблизи стока становится настолько узким, что наступает динамическое равновесие, когда увеличение напряжения $U_{СИ}$ вызывает уменьшение ширины канала и повышение его сопротивления. В итоге значение тока I_C достигает насыщения при увеличении напряжения $U_{СИ}$.

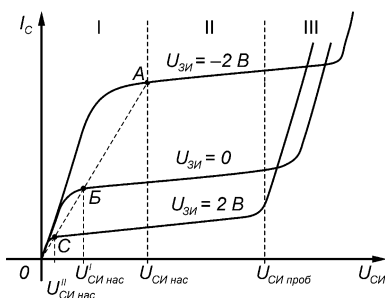


Рис. 7.9. Выходные характеристики МДП-транзистора со встроенным каналом p -типа: I – крутая, или «омическая», область; II – пологая или насыщенная область; III – область пробоя

7.2.2. МДП-транзисторы с индуцированным каналом

Структура МДП - транзистора с индуцированным p -каналом показана на рис. 7.5, *в*. В качестве исходного материала транзистора использован кремний n -типа. Чаще всего диэлектриком является оксид кремния SiO_2 . В оксиде кремния всегда содержатся положительно заряженные ионы натрия, калия, водорода, осаждаемые на поверхность кремния в процессе выполнения технологических операций.

Положительный заряд в пленке оксида кремния наводит (индуцирует) у поверхности n -области слой, в котором концентрация электронов выше, чем в объеме n -области. При подаче на затвор отрицательного напряжения $U_{3И}$ электроны поверхностного слоя отталкиваются в глубь полупроводника, а дырки движутся к поверхности. Приповерхностный слой приобретает дырочную электропроводность, то есть появляется тонкий инверсный слой, соединяющий сток с истоком, который играет роль канала. Толщина индуцированного канала составляет всего 1 – 5 нм. При приложении напряжения между истоком и стоком дырки, перемещаясь по каналу, создают ток стока. Изменяя напряжение на затворе, можно расширять или сужать канал, увеличивая или уменьшая ток стока.

Напряжение на затворе, при котором индуцируется канал, называют **пороговым напряжением** $U_{3И\text{ пор}}$ (рис. 7.10, *а*).

Вид выходных характеристик этих транзисторов (рис. 7.10, *б*) аналогичен выходным характеристикам МДП-транзисторов со встроенным p -каналом. Разница состоит в том, что напряжение затвор–исток всегда отрицательное, то есть $U_{3И} < 0$. Наблюдаются те же три области: омическая, насыщения и пробоя.

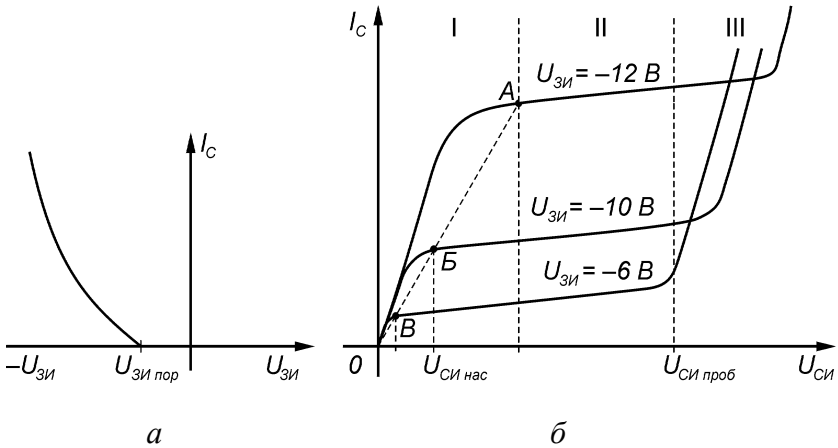


Рис. 7.10. Стокозатворные (а) и выходные (б) характеристики МДП-транзисторов с индуцированным p -каналом

7.3. Эффект смещения подложки

Подложка, на которой изготавливают МДП-транзистор, имеет противоположный каналу тип проводимости. Рассмотрим, как меняются характеристики МДП - транзистора при приложении напряжения между истоком и подложкой. Отметим, что приложенное напряжение между истоком и подложкой при условии наличия инверсионного канала падает на обедненную область индуцированного p - n -перехода.

В этом случае при прямом его смещении будут наблюдаться значительные токи, соответствующие прямым токам p - n -перехода. Эти токи попадут в стоковую цепь, и транзистор работать не будет. Поэтому используется только напряжение подложки, соответствующее обратному смещению индуцированного и истокового p - n -перехода. По полярности это будет напряжение подложки противоположного знака по сравнению с напряжением стока. При приложении напряжения канал-

подложка происходит расширение ОПЗ и увеличение заряда ионизированных акцепторов:

$$Q_B = \sqrt{2q\varepsilon_s\varepsilon_0N_A(\psi_{s0} + V_{SS})}. \quad (7.6)$$

Поскольку напряжение на затворе V_{GS} постоянно, то постоянен и заряд на затворе МДП-транзистора Q_m . Следовательно, из уравнения электронейтральности вытекает, что если заряд акцепторов в слое обеднения Q_B вырос, заряд электронов в канале Q_n должен уменьшиться. С этой точки зрения подложка выступает как второй затвор МДП-транзистора, поскольку регулирует также сопротивление инверсионного канала между истоком и стоком.

При возрастании заряда акцепторов в слое обеднения возрастет и пороговое напряжение транзистора V_T . Изменение порогового напряжения будет равно:

$$\Delta V_T = \frac{\Delta Q_B}{C_{ox}} = \sqrt{\frac{2\varepsilon_s\varepsilon_0N_A}{C_{ox}^2}} \left[\sqrt{\psi_{s0} + V_{SS}} - \sqrt{\psi_{s0}} \right]. \quad (7.7)$$

Поскольку смещение подложки приводит только к изменению порогового напряжения V_T , то переходные характеристики МДП-транзистора при различных напряжениях подложки V_{SS} смещаются параллельно друг другу. На рис. 7.11 показан эффект влияния смещения подложки на переходные характеристики.

7.4. Эквивалентная схема МДП-транзистора

Исходя из общефизических соображений, МДП-транзистор можно изобразить в виде эквивалентной схемы, представленной на рис. 7.12.

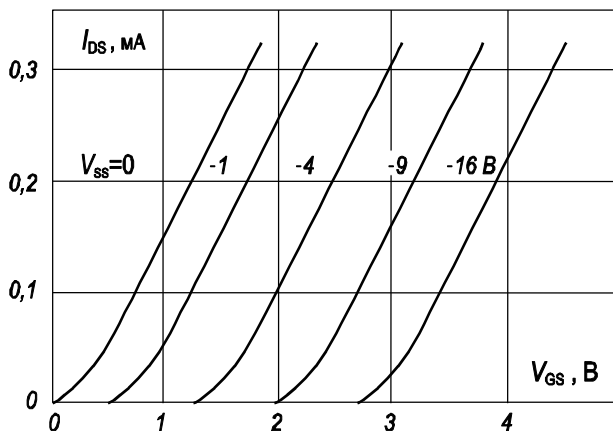


Рис. 7.11. Влияние напряжения смещения канал-подложка V_{SS} на переходные характеристики транзистора

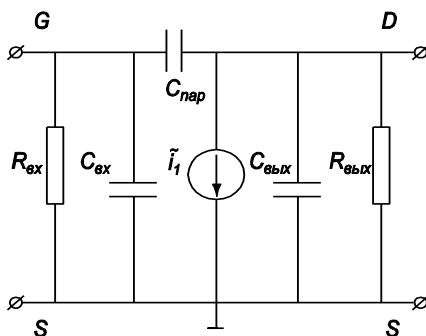


Рис. 7.12. Простейшая эквивалентная схема МДП-транзистора

В этой схеме $R_{вх}$ обусловлено сопротивлением подзатворного диэлектрика, входная емкость $C_{вх}$ – емкостью подзатворного диэлектрика и емкостью перекрытия затвор-исток. Паразитная емкость $C_{пар}$ обусловлена емкостью перекрытий затвор-сток. Выходное сопротивление $R_{вых}$ равно сопротивлению канала транзистора и сопротивлению легированных об-

ластей истока и стока. Выходная емкость $C_{\text{вых}}$ определяется емкостью p - n -перехода стока. Генератор тока \tilde{i}_1 передает эффект усиления в МДП-транзисторе.

7.5. Подпороговые характеристики МДП-транзистора

Для получения статических параметров МДП-транзисторов часто используют метод, предусматривающий соединение затвора и стока друг с другом и измерения зависимости тока стока от стокового напряжения. Соответствующая схема включения имеет вид, приведенный на рис. 7.13. В этой схеме транзистор всегда находится в режиме насыщения.

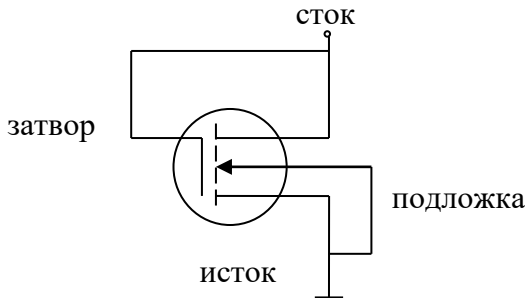


Рис. 7.13. Схема включения МДП-транзистора, в которой затвор и сток соединены друг с другом

Уравнение вольт-амперной характеристики МДП-транзистора с изолированным затвором в линейной области при $0 \leq U_{\text{СИ}} < U_{\text{СИ насыщ}}$ имеет вид

$$I_C = \frac{z\mu_n C_i}{L} \left[(U_{\text{с}} - U_{\text{н0}}) \cdot U_{\text{н}} - \frac{1}{2} U_{\text{сн}}^2 \right], \quad (7.8)$$

где z – ширина канала; L – длина канала; C_i – емкость диэлектрика; μ_n – подвижность носителей в канале.

В режиме насыщения при $U_{СИ\ насыщ} = U_3 - U_{пор}$ уравнение вольт-амперной характеристики МДП-транзистора с изолированным затвором принимает вид

$$I_{СИ\ насыщ} = \frac{z\mu_n C_i}{L} \frac{(U_3 - U_{пор})^2}{2}. \quad (7.9)$$

Согласно уравнению (7.9) график зависимости $(I_C)^{1/2}$ от U_3 должен быть линейным (сплошная линия 1, рис. 7.14). Пороговое напряжение легко найти, как точку пересечения с осью напряжений. В идеальном случае ниже порогового напряжения канала нет. В реальности этот график имеет вид, соответствующий пунктирной линии 2 на рис. 7.14. Такой вид графика связан с токами, протекающими в транзисторе при напряжениях меньше пороговых.

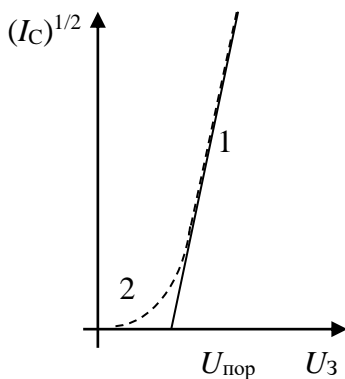


Рис. 7.14. График зависимости $(I_C)^{1/2}$ от напряжения на затворе U_3

Одно из наиболее удобных и полезных упрощений, принятых в простой теории МДП-транзистора, состоит в том, что пока его затворное напряжение меньше порогового, весь заряд свободных носителей в канале считается равным нулю, а когда затворное напряжение больше порогового, заряд обедненной области под затвором считается постоянным. Такое приближение удобно потому, что концентрации свободных носителей заряда в канале экспоненциально зависят от напряжения канала, тогда как плотность неподвижного заряда пропорциональна корню из этого напряжения. Однако вблизи порогового напряжения простая теория недостаточно точно описывает ток прибора, что подтверждается при тщательном измерении его характеристик. Главный момент здесь - это некоторый конечный (ненулевой) ток при затворных напряжениях, меньших порогового, который называется *предпороговым или подпороговым током*, то есть если напряжение на затворе МОП-транзистора ниже порогового и граница с окислом находится в условиях слабой инверсии, соответствующий ток стока называют подпороговым (или предпороговым) током.

Подпороговые токи обусловлены инверсным зарядом, который существует при затворных напряжениях меньше порогового. В интервале напряжений $U_{\text{инв}} < U_3 < U_{\text{пор}}$, где $U_{\text{инв}}$ - напряжение инверсии, при котором концентрация носителей в канале равна собственной концентрации носителей n_i , существует малый ток канала. Указанный интервал напряжений называется подпороговой областью, поверхностная концентрация носителей в нем изменяется от собственной концентрации носителей n_i до концентрации основных носителей в объеме. В условиях слабой инверсии основной компонентой тока канала является диффузионная составляющая.

Обращает на себя внимание тот факт, что в области слабой инверсии зависимость тока стока I_C или I_{DS} от напряжения на затворе U_3 или V_{GS} - экспоненциальная функция. На рис.

7.15 приведены переходные характеристики транзистора в этой области.

Подпороговая область характеристик особенно важна для МОП-транзисторов, предназначенных для работы в низковольтных полупроводниковых устройствах с малым энергопотреблением (цифровые логические схемы, устройства памяти), поскольку этот режим соответствует закрытому состоянию МОП-транзистора, а также описывает процесс переключения из закрытого состояния в открытое

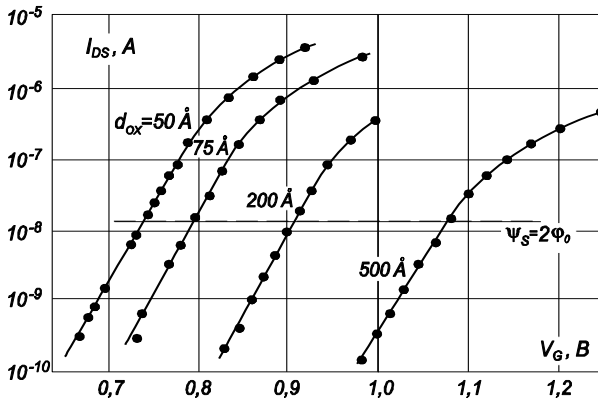


Рис. 7.15. Зависимость тока стока I_{DS} от напряжения на затворе V_G в предпороговой области для МДП-транзисторов с разной толщиной подзатворного диэлектрика.

Напряжение исток-сток $V_{DS} = 0,025 \text{ В}$

7.6. Приборы с зарядовой связью

Новым типом полевых полупроводниковых приборов, работающих в динамическом режиме, являются приборы с зарядовой связью (ПЗС). На рис. 7.16 приведена схема, поясняющая устройство и основные физические принципы работы ПЗС. Приборы с зарядовой связью представляют собой ли-

нейку или матрицу последовательно расположенных МДП - структур. Величина зазора между соседними МДП - структурами невелика и составляет 1 - 2 мкм. ПЗС - элементы служат для преобразования оптического излучения в электрические сигналы и передачи информации от одного элемента электронной схемы к другому.

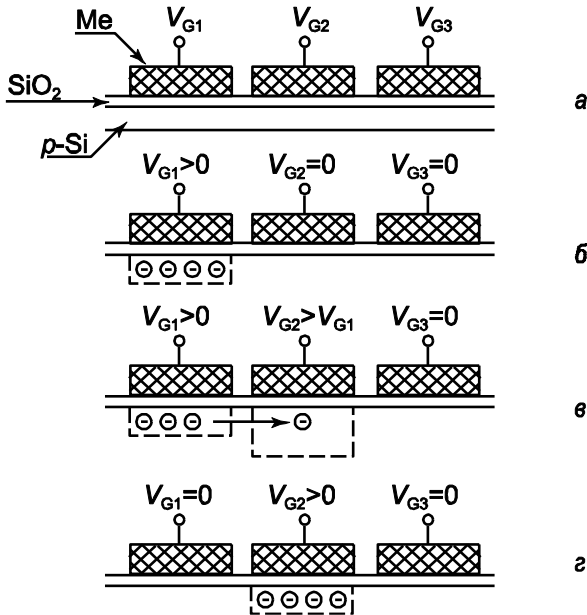


Рис. 7.16. Устройство и принцип работы приборов с зарядовой связью

Рассмотрим принцип работы ПЗС. При подаче обедняющего импульса напряжения V_{G1} на затвор 1-го элемента в ОПЗ полупроводника образуется неравновесный слой обеднения. Для электронов в полупроводнике p -типа это соответствует формированию под затвором 1-го элемента потенциальной ямы. Известно, что неравновесное состояние сохраня-

ется в период времени t порядка времени генерационно-рекомбинационных процессов $\tau_{\text{ген}}$. Поэтому все остальные процессы в ПЗС-элементах должны проходить за времена меньше $\tau_{\text{ген}}$. Пусть в момент времени $t_1 \gg \tau_{\text{ген}}$ в ОПЗ под затвор 1-го элемента инжектирован каким-либо образом информационный заряд электронов (рис. 7.16, б). Теперь в момент времени $t_2 > t_1$, но $t_2 \ll \tau_{\text{ген}}$ на затвор 2-го ПЗС элемента подадим напряжение $V_{G2} > V_{G1}$, способствующее формированию более глубокой потенциальной ямы для электронов под затвором 2-го элемента. Вследствие диффузии и дрейфа возникнет поток электронов из ОПЗ под 1-м элементом в ОПЗ под вторым элементом, как показано на рис. 7.16, в. Когда весь информационный заряд перетечет в ОПЗ 2-го ПЗС-элемента, напряжение на затворе V_{G1} снимается, а на затворе V_{G2} уменьшается до значения, равного V_{G1} (см. рис. 7.16, г). Произошла передача информационного заряда. Затем цикл повторяется и заряд передается дальше в ОПЗ 3-го ПЗС-элемента. Для того, чтобы приборы с зарядовой связью эффективно функционировали, необходимо, чтобы время передачи $t_{\text{пер}}$ от одного элемента к другому было много меньше времени генерационно-рекомбинационных процессов ($t_{\text{пер}} \ll \tau_{\text{ген}}$). Не должно быть потерь информационного заряда в ОПЗ вследствие захвата на поверхностные состояния, в связи с чем требуются МДП - структуры с низкой плотностью поверхностных состояний ($N_{\text{ss}} \approx 10^{10} \text{ см}^{-2} \cdot \text{эВ}^{-1}$).

ЗАКЛЮЧЕНИЕ

Дисциплина «Физические основы электроники» является базовым курсом при подготовке студентов направления подготовки бакалавров 210100 «Электроника и наноэлектроника», профиля «Микроэлектроника и твердотельная электроника». Не умаляя полезности существующих учебников и учебных пособий по данной дисциплине, настоящая книга может стать фундаментом для начального образования по твердотельной электронике, своеобразным мостиком для освоения новых областей электроники, таких как микроэлектроника, наноэлектроника, микросхемотехника и функциональная электроника.

В учебном пособии рассматриваются физические основы твердотельной электроники, связанные с контактными явлениями в электронно-дырочных переходах, барьерах Шоттки и гетеропереходах, а также основы физики поверхности полупроводников и МДП-структур, устройство, характеристики, параметры, система обозначений и простейшие схемы применения полупроводниковых диодов, биполярных и полевых транзисторов, а также приборов с отрицательным дифференциальным сопротивлением - тиристоров.

В пособии использованы соответствующие действующим ГОСТам терминология, система обозначений и условные графические изображения приборов твердотельной электроники.

Систематизированное рассмотрение явлений, лежащих в основе работы приборов твердотельной электроники, и основных типов полупроводниковых приборов облегчит понимание и восприятие материала лекций по дисциплине «Физические основы электроники», читаемых студентам.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Бонч-Бруевич В.Л. Физика полупроводников/ В.Л. Бонч-Бруевич, С. Г. Калашников. М.: Наука, 1977. 672 с.
2. Бормонтов Е.Н. Физика и метрология МДП-структур: учеб. пособие/ Е.Н. Бормонтов. Воронеж: ВГУ, 1997. 184 с.
3. Булычев А.Л. Электронные приборы/ А.Л. Булычев, В.А. Прохоренко. Минск: Вышэйш. шк., 1987. 316 с.
4. Дулин В.Н. Электронные приборы/ В.Н. Дулин. М.: Энергия, 1977. 424 с.
5. Гуртов В.А. Твердотельная электроника: учеб. пособие/ В.А. Гуртов. Петрозаводск, 2004. 312 с.
6. Зи С. Физика полупроводниковых приборов: в 2-х кн.: пер. с англ./ С. Зи. М.: Мир, 1984.
7. Кремниевые планарные транзисторы/ под ред. Я.А. Федотова. М.: Советское радио, 1973. 336 с.
8. Крутякова М.Г. Полупроводниковые приборы и основы их проектирования/ М.Г. Крутякова, Н.А.Чарыков, В.В. Юдин. М.: Радио и связь, 1983. 352 с.
9. Маллер Р. Элементы интегральных схем /Р. Маллер, Т. Кейминс. М.: Мир, 1989. 630 с.
10. Морозова И.Г. Физика электронных приборов/ И.Г. Морозова. М.: Атомиздат, 1980. 392 с.
11. Пасынков В.В. Полупроводниковые приборы: учебник для вузов/ В.В. Пасынков, Л.К. Чиркин. СПб.: Лань, 2003. 480 с.
12. Полупроводниковые приборы. Диоды, тиристоры, оптоэлектронные приборы: справочник/ под ред. Н.А. Горюнова. М.: Энергоиздат, 1987. 743 с.
13. Полупроводниковые приборы. Транзисторы: справочник/ под. ред. Н.А. Горюнова. М.: Энергоиздат, 1986. 904 с.
14. Росадо Л. Физическая электроника и микроэлектроника/ Л. Росадо. М.: Высш. шк., 1991. 351 с.

15. Степаненко И.П. Основы теории транзисторов и транзисторных схем/ И.П. Степаненко. М.: Энергия, 1977. 671 с.
16. Тугов Н.М. Полупроводниковые приборы/ Н.М. Тугов, Б.А. Глебов, Н.А. Чарыков. М.: Энергоатомиздат, 1990. 576 с.
17. Федотов Я.Л. Основы физики полупроводниковых приборов/ Я.Л. Федотов. М.: Советское радио, 1969. 592 с.
18. Шалимова К.В. Физика полупроводников/ К.В. Шалимова. М.: Энергия, 1976. 416 с.
19. Щука А.А. Электроника: учеб. пособие/ А.А. Щука; под ред. А.С. Ситова. СПб: БХВ-Петербург, 2005. 800 с.
20. Электронные приборы/ под ред. Г.Г. Шишкина. М.: Энергоатомиздат, 1989. 496 с.

ОГЛАВЛЕНИЕ

Введение	3
1. Физические основы работы приборов твердотельной электроники	5
1.1. Зонная структура полупроводников	5
1.2. Собственные и примесные полупроводники	7
1.3. Статистика электронов и дырок в полупроводниках. Концентрация носителей заряда и положение уровня Ферми	8
1.4. Концентрация электронов и дырок в собственном полупроводнике	12
1.5. Концентрация электронов и дырок в примесном полупроводнике	13
1.6. Определение положения уровня Ферми	15
1.7. Проводимость полупроводников	16
1.8. Токи в полупроводниках	17
1.9. Неравновесные носители. Генерация и рекомбинация носителей	19
1.10. Уравнение непрерывности	21
1.11. Электрические поля в кристаллах	25
2. Контактные явления	27
2.1. Разновидности электрических переходов и контактов	27
2.2. Электронно-дырочный переход	28
2.2.1. p - n -переход в состоянии термодинамического равновесия	29
2.2.2. Контактная разность потенциалов	34
2.2.3. Ширина p - n -перехода	35
2.2.4. Прямое включение p - n -перехода	41
2.2.5. Уровень инжекции	46
2.2.6. Обратное включение p - n -перехода	48
2.2.7. Теоретическая вольт-амперная характеристика p - n -перехода	51

2.2.8. Вольт-амперная характеристика реального p - n -перехода	57
2.2.9. Вольт-амперная характеристика p - n -перехода в полулогарифмических координатах	66
2.2.10. Пробой p - n -перехода	68
2.2.11. Емкость p - n -перехода	78
2.2.12. Переходные процессы в p - n -переходах	82
2.2.13. Частотные свойства p - n -перехода	89
2.2.14. Эквивалентные схемы p - n -перехода	92
2.2.15. Влияние температуры на свойства p - n -перехода	94
2.3. Разновидности электрических переходов	98
2.3.1. Гетеропереходы	98
2.3.2. Контакт полупроводников с одним ти- пом электропроводности	108
2.3.3. Контакт металл – полупроводник. Барьер Шоттки	113
2.3.4. Омические контакты	121
2.3.5. Явления на поверхности полупроводников	125
3. Физика структуры металл-диэлектрик-полупро- водник	129
3.1. Идеальная структура металл – диэлектрик – полупроводник	129
3.2. Область пространственного заряда в равно- весных условиях	133
3.3. Приповерхностная область пространственно- го заряда	139
3.4. Распределение плотности пространственного заряда, электрического поля и потенциала в идеальной МДП-структуре в режиме сильной инверсии	142

3.5. Вольт-фарадные характеристики идеальной МДП-структуры	146
3.5.1. Емкость области пространственного заряда	146
3.5.2. Емкость МДП-структур	149
3.6. Компоненты заряда в реальном диоксиде кремния и их влияние на ВФХ МДП-структуры	153
3.7. Распределение плотности пространственного заряда, электрического поля и потенциала в реальной МДП-структуре в режиме сильной инверсии	156
3.8. Определение параметров МДП-структур на основе анализа $C-V$ характеристик	161
4. Полупроводниковые диоды	167
4.1. Методы изготовления полупроводниковых диодов	167
4.2. Выпрямительные диоды	172
4.3. Варикапы	180
4.4. Стабилитроны	183
4.5. Туннельный и обращенный диоды	189
4.6. Высокочастотные и сверхвысокочастотные диоды	195
4.7. Импульсные диоды	198
5. Биполярные транзисторы	202
5.1. Структура и основные режимы работы	202
5.2. Схемы включения транзистора	206
5.3. Основные физические процессы в биполярных транзисторах	211
5.4. Модуляция сопротивления базы	217
5.5. Статические характеристики биполярных транзисторов	219
5.6. Эквивалентная схема биполярного транзистора	223

5.7. Дифференциальные параметры биполярных транзисторов в схеме с общей базой	228
5.8. Дифференциальные параметры биполярных транзисторов в схеме с общим эмиттером	230
5.9. Малосигнальные параметры биполярного транзистора	232
5.9.1. Система z - параметров	233
5.9.2. Система y - параметров	234
5.9.3. Система h - параметров	235
5.10. Частотные и импульсные свойства транзисторов	239
6. Тиристоры	245
6.1. Структура и принцип действия	245
6.2. Основные параметры тиристоров	253
6.3. Феноменологическое описание ВАХ диодистора	254
6.4. Способы включения и выключения тиристоров	260
7. Полевые транзисторы и приборы с зарядовой связью	262
7.1. Полевой транзистор с управляющим p - n -переходом	264
7.2. МДП-транзисторы	271
7.2.1. МДП-транзисторы со встроенным каналом	276
7.2.2. МДП-транзисторы с индуцированным каналом	278
7.3. Эффект смещения подложки	279
7.4. Эквивалентная схема МДП-транзистора	280
7.5. Подпороговые характеристики МДП-транзистора	282
7.6. Приборы с зарядовой связью	285
Заключение	288
Библиографический список	289

Учебное издание

Свистова Тамара Витальевна

ПРИБОРЫ ТВЕРДОТЕЛЬНОЙ
ЭЛЕКТРОНИКИ

В авторской редакции

Подписано к изданию 12. 11.2012.
Объем данных 11,9 Мб.

ФГБОУ ВПО «Воронежский государственный технический
университет»
394026 Воронеж, Московский просп., 14