

156

П.Е.Троян

ТВЕРДОТЕЛЬНАЯ ЭЛЕКТРОНИКА

Учебное пособие

Томск-2006

Троян П.Е.

Твердотельная электроника:

Учебное пособие. - Томск: Томский государственный университет систем управления и радиоэлектроники, 2006. - 321 с.

В учебном пособии рассмотрены физические основы твердотельной электроники, устройство, принцип действия, характеристики и параметры основных классов полупроводниковых приборов различного назначения, их эквивалентные схемы и модели, а также вопросы технологии изготовления и применения приборов.

Учебное пособие предназначено для студентов специальности 210104 «Микроэлектроника и твердотельная электроника», обучающихся по направлению "Электроника и микроэлектроника», и написано в соответствии с рабочей программой по дисциплине и требованиями ГОС ВПО по данному направлению.

©ТроянП.Е

2006

©Томский государственный университет систем управления и радиоэлектроники
2006

СОДЕРЖАНИЕ

1. Введение	7
1.1. Цели, задачи и предмет курса "Твердотельная электроника"	7
1.2. Краткая историческая справка	1
1.3. Современное состояние и перспективы развития электроники	11
2. Физические основы твердотельной электроники	13
2.1. Элементы зонной теории полупроводников	13
2.2. Параметры, характеризующие свойства полупроводниковых материалов	17
2.3. Фундаментальная система уравнений твердотельной электроники	23
2.4. Собственные, примесные и компенсированные полупроводники	28
2.5. Диапазон рабочих температур полупроводниковых приборов	34
2.6. Равновесные и неравновесные носители зарядов в полупроводниках. Основные и неосновные носители. Закон действующих масс	37
2.7. Полупроводники в электрическом поле	40
2.8. Генерация и рекомбинация носителей в полупроводниках	44
2.9. Уравнение электронейтральности	45
2.10. Явления на поверхности полупроводников	46
3. Контакты металл-полупроводник. Выпрямляющие и омические переходы на контакте металла с полупроводником	49
3.1. Введение	49
3.2. Энергетическая диаграмма выпрямляющего контакта металл-полупроводник	49
3.3. Принцип выпрямления тока на контакте металл-полупроводник по энергетическим диаграммам	52
3.4. Вольтамперная характеристика выпрямляющего контакта металл-полупроводник	55
3.5. Расчет напряженности поля и потенциала на кон-	

такте металл-полупроводник. Ширина области пространственного заряда	57
3.6. Диод Шоттки: структура, эквивалентная схема, параметры эквивалентной схемы. Модель диода Шоттки	60
3.7. Эффект Шоттки	64
3.8. Достоинства и недостатки диода Шоттки	68
3.9. Омические контакты и их параметры	69
4. Электронно-дырочные переходы	72
4.1. Виды электронно-дырочных переходов при контакте полупроводников	72
4.2. Механизм образования ЭДП. Определение ЭДП	72
4.3. Контактная разность потенциалов φ_k^o . Зависимость φ_k^o от температуры, ширины запрещенной зоны, концентрации легирующей примеси	74
4.4. Потоки носителей зарядов в ЭДП по энергетическим диаграммам. Односторонняя проводимость p-n перехода	78
4.5. Некоторые понятия и определения по ЭДП	82
4.6. Методы получения ЭДП	86
4.7. Расчет напряженности электрического поля и потенциала в ЭДП. Ширина ОПЗ для резкого и плавного переходов	87
4.8. Вольтамперная характеристика идеального ЭДП. Диоды с "толстой" и "тонкой" базами	93
4.9. Вольтамперная характеристика реального ЭДП	104
4.10. Явления в ЭДП при высоком уровне инжекции	108
4.11. Диффузионная и барьерная емкости ЭДП	114
4.12. Эквивалентная схема ЭДП. Параметры эквивалентной схемы	118
4.13. Полная проводимость p-n перехода. Зависимость параметров от частоты	120
4.14. Переходные процессы в ЭДП	128
4.15. Зависимость выпрямляющих свойств ЭДП от частоты	135
4.16. Пробой электронно-дырочного перехода	137
4.17. Зависимость параметров ЭДП от температуры	148

4.18. Зарядоуправляемая модель ЭДП	151
4.19. Гетеропереходы	153
5. Диоды на основе электронно-дырочных переходов	156
5.1. Классификация и маркировка диодов	156
5.2. Выпрямительные диоды. Выпрямительные столбы и блоки	160
5.3. Универсальные и импульсные диоды	161
5.4. СВЧ-диоды	164
5.5. Варикапы	174
5.6. Стабилитроны. Лавинно-пролетные диоды	177
5.7. Туннельные и обращенные диоды	190
5.8. Фотодиоды	208
5.9. Светодиоды	214
5.10. Оптопары	225
6. Биполярные транзисторы	229
6.1. Общие сведения о биполярных транзисторах (БТ)	229
6.2. Потоки носителей зарядов в БТ	231
6.3. Внутренние и внешние параметры БТ	233
6.4. Статические параметры	236
6.5. Явления в БТ при больших токах	239
6.6. Модуляция толщины базы коллекторным напряжением (эффект Эрли)	240
6.7. Пробой транзистора	243
6.8. Статические характеристики	245
6.9. Динамический режим работы	249
6.10. Усилительные свойства	251
6.11. Частотные параметры	253
6.12. Эквивалентная схема биполярного транзистора	257
6.13. Системы z , y и h -параметров	261
6.14. Модели биполярного транзистора	267
6.15. Некоторые разновидности биполярных транзисторов	269
6.16. Основные параметры биполярных транзисторов и их ориентировочные значения	276
6.17. Маркировка транзисторов	278
7. Тиристоры	278
8. Полевые транзисторы	284
8.1. Полевые транзисторы с управляемым р-п переходом	

дом	284
8.2. Дифференциальные параметры	287
8.3. Полевые транзисторы с управляющим переходом	
металл-полупроводник	290
8.4. Полевые транзисторы с изолированным затвором	291
8.5. Импульсный режим полевых транзисторов	294
9. Сенсоры, датчики, преобразователи	298
9.1. Датчики температуры	298
9.1.1. Терморезисторы	298
9.1.2. Диоды, как датчики температуры	302
9.1.3. Транзисторы, как датчики температуры	304
9.2. Полупроводниковые термоэлектрические преоб- разователи	305
9.3. Датчики магнитного поля. Преобразователи.	314
9.3.1. Магниторезисторы	316
9.3.2. Магнитодиоды	317
9.3.3. Биполярные магниторезисторы	319
Список литературы	321

ВВЕДЕНИЕ

1.1. Цели, задачи и предмет курса "Твердотельная электроника"

Курс "Твердотельная электроника" относится к циклу обще-профессиональных дисциплин, читаемых для студентов специальностей, обучающихся по направлению "Электроника и микроэлектроника". Для освоения дисциплины необходимы знания по курсу общей физики, высшей математики и электротехники. Поскольку из трех специальностей университета две ("Промышленная электроника" и "Электронные приборы и устройства") не изучают курса физики твердого тела, то в учебное пособие включен раздел "Физические основы твердотельной электроники", позволяющий дать студентам этих специальностей минимальные сведения по физике полупроводников, необходимые для усвоения дисциплины "Твердотельная электроника".

Учебное пособие может быть также полезным для студентов, обучающихся по другим специальностям университета радиоэлектронного профиля.

Электроника - это область науки и техники, связанная с изучением физических явлений, протекающих в вакууме, газах, жидкостях и твердых телах, технологией изготовления электронных приборов и их применением для создания электронных устройств различного назначения.

Так как в проводящих средах и приборах на их основе ток создается в основном за счет движения электронов, то в широком понимании все созданные приборы можно, в общем называть электронными. Однако в связи с различными средами, в которых перемещаются носители зарядов, различают следующие основные классы электронных приборов: электровакуумные, газоразрядные, твердотельные. Причем класс твердотельных приборов в основном представлен полупроводниковыми приборами.

Настоящее учебное пособие представляет собой краткий конспект лекций по дисциплине "Твердотельная электроника", который в ТУСУРе изучают студенты трех специальностей: "Промышленная электроника", "Электронные приборы и устройства", "Микроэлектроника и твердотельная электроника".

Предмет курса "Твердотельная электроника" ограничивается областью твердотельных приборов.

Твердотельная электроника - направление в электронике, охватывающее изучение физических явлений в твердых телах, лежащих в основе работы приборов, технологии изготовления и применения твердотельных приборов. По конструктивно-технологическим признакам твердотельные приборы подразделяются на дискретные и интегральные. Основным видом твердотельных приборов являются полупроводниковые приборы.

Полупроводниковые приборы - это приборы, принцип действия которых основан на специфических свойствах полупроводниковых материалов и структур на их основе и связан с движением заряженных частиц в твердом теле. По функциональному назначению полупроводниковые приборы делятся на усилительные, генераторные, переключательные и индикаторные. Основными классами полупроводниковых приборов являются классы диодов, транзисторов, тиристоров.

Современный этап развития электроники - микроэлектроника. Микроэлектроника - динамично развивающееся научно-техническое направление, базирующееся на достижениях в области физики твердого тела, технологии, микросхемотехники и системотехники.

Микроэлектроника - это научно-техническое направление в электронике, охватывающее проблемы исследования, проектирования и конструирования, а также изготовления и применения микроэлектронных изделий для создания электронных устройств различного назначения. Причем под микроэлектронными изделиями понимаются изделия с высокой степенью интеграции. Основными элементами современных электронных устройств являются интегральные микросхемы.

Электронное устройство - совокупность электрически связанных между собой пассивных (резисторы, конденсаторы, индуктивности) и активных (диоды, транзисторы, тиристоры) компонентов, создающих схему управления и исполнения, выполняющих определенную функцию. Причем пассивные и активные элементы могут иметь дискретное или интегральное исполнение.

Возможности создания новых устройств автоматики, вычислительной техники, телевидения, медицинской техники, бытовой электронной аппаратуры и т.д. зависят от наличия соответствующей элементной базы. В качестве основной элементной базы современных устройств электроники выступают полупроводниковые приборы и интегральные схемы, т.е. приборы твердотельной электроники.

Поскольку в государственном образовательном стандарте по направлению "Электроника и микроэлектроника" предусмотрено изучение дисциплины "Микроэлектроника", то в данном пособии вопросы, относящиеся к этому курсу, не рассматриваются. Основное содержание данного учебного пособия посвящено изучению физики работы полупроводниковых приборов, их параметров, эквивалентных схем и моделей, возможным областям применения приборов.

Целью изучения дисциплины "Твердотельная электроника" является приобретение знаний по физическим основам действия твердотельных приборов, их электрическим характеристикам в статическом и динамическом режимах, реакции приборов на внешние воздействия, представлению приборов в виде электрических моделей и методов экспериментального определения параметров моделей и приборов, а также выработка навыков в совершенствовании и углублении знаний по твердотельным приборам.

Задачей изучения курса является приобретение умений и навыков в расчете параметров твердотельных приборов, умения правильно выбрать прибор для построения электронной схемы с учетом поставленной задачи, обеспечить надежную эксплуатацию прибора с максимальным использованием его возможностей, а также приобретение практических навыков экспериментального определения параметров приборов и моделей.

Студенты, изучившие дисциплину "Твердотельная электроника", **должны знать и уметь:**

1. Физические основы твердотельной электроники.
2. Устройство, принцип действия, основные характеристики и параметры основных классов полупроводниковых приборов, их эквивалентные схемы и модели.

3. Уметь экспериментально определять параметры приборов и параметры электрических моделей приборов.

4. Иметь представления о современном состоянии элементной базы твердотельной электроники и тенденциях ее развития. Обладать навыками получения новых знаний по твердотельным приборам.

1.2. Краткая историческая справка

В целом развитие электроники как научно-технического направления прошло три этапа: вакуумная электроника, этапа дискретных полупроводниковых приборов и микроэлектроники. Поскольку предметом данного пособия является твердотельная электроника, то остановимся очень кратко на важнейших вехах развития данного этапа.

Несмотря на то, что до начала 50-х годов прошлого века господствующее положение в электронике занимала вакуумная электроника, работы по созданию твердотельных приборов восходят к середине 19 века. Впервые выпрямляющие свойства на контакте между металлами и сернистыми соединениями обнаружены в 1884г. А еще ранее свойства полупроводников начали изучать Фарадей (1833г.) и Беккерель (1839г.).

Как научно-техническое направление твердотельная электроника сформировалась в начале 20 века после создания основ электродинамики (Максвелл 1861-1873гг.), открытия фотопроводимости (Смит, 1873г.), односторонней проводимости контакта металл-полупроводник (Браун 1874г.), фотоэлектронной эмиссии (Герц, Столетов 1887-1905гг.), открытия электрона (Дж.Дж.Томсон 1897г.), создания электронной теории (Лоренц 1892-1909гг.).

В России первые шаги по использованию твердотельных приборов принадлежат Попову, который при изобретении радио применил порошковый когерер, в котором использованы нелинейные свойства зернистых систем, и Лосеву, который использовал отрицательное дифференциальное сопротивление на контакте металла с полупроводником для усиления и генерации электромагнитных колебаний. Значительный объем работ в области полупроводниковой электроники выполнен Иоффе и учениками его школы, а также Шоттки и Шокли.

Важнейшей вехой в развитии твердотельной электроники является изобретение в 1948г. Бардиным, Браттейном и Шокли биполярного транзистора, за что они были удостоены Нобелевской премии, и полевого транзистора с управляющим $p-n$ переходом (1952г. Шокли), а также полевого транзистора со структурой металл-диэлектрик-полупроводник (1963г. Нейман, Хофстейн).

Начиная с 1962г., когда появилась первая интегральная микросхема, в электронике начинается новый этап развития электроники - микроэлектроника, который характеризует современное состояние электроники.

Среди ученых России существенный вклад в развитие твердотельной электроники внесли Иоффе, Френкель, Тамм, Басов, Прохоров, Алферов, Федотов, Степаненко, Викулин, Стафеев, Пасынков и др.

1.3. Современное состояние и перспективы развития электроники

Современный этап развития электроники - этап микроэлектроники, характеризующийся созданием широкой номенклатуры и массовым выпуском интегральных микросхем различного назначения, запоминающих устройств, микропроцессорных комплексов, однокристалльных ЭВМ, заказных больших интегральных микросхем, программируемых логических матриц, специализированных интегральных схем интерфейсов локальных вычислительных сетей.

Однако этап схемотехнической интегральной электроники в силу ряда ограничений и недостатков начинает постепенно переходить к этапу функциональной нанoeлектроники. Скорее всего, будущее электроники - это нанoeлектроника на основе наноматериалов и нанотехнологий. Тем не менее, видимо, ближайшее десятилетие будет характеризоваться широким использованием интегральных схем, их совершенствованием и бурным развитием нанoeлектроники. В области микроэлектроники работы будут идти по пути увеличения степени интеграции, снижения потребляемой мощности, увеличения быстродействия и по-

вышения частотного диапазона, создания тонкопленочных микросхем с пленочными активными элементами.

Функциональная микроэлектроника - перспективное научно-техническое направление в электронике по созданию приборов и устройств для обработки, передачи и хранения информации на основе сред с динамическими неоднородностями, создаваемыми путем использования новых физически явлений в конденсированных средах. Основными разделами функциональной электроники являются: оптоэлектроника, акустоэлектроника, магнитоэлектроника, криоэлектроника, биоэлектроника.

Наноэлектроника - научно-техническое направление в электронике, связанное с исследованием физических явлений в наноматериалах и наноструктурах, технологии получения наноматериалов и наноприборов и области применения наноприборов.

2. ФИЗИЧЕСКИЕ ОСНОВЫ ТВЕРДОТЕЛЬНОЙ ЭЛЕКТРОНИКИ

2.1. Элементы зонной теории полупроводников

Наиболее распространенной моделью для описания свойств полупроводников является модель энергетических зон, т.е. зонная модель. Зонная теория полупроводников является сложной квантово-механической теорией и в полном объеме не может быть изложена в данном пособии.

Здесь мы рассмотрим лишь два фрагмента зонной теории полупроводников, необходимых для понимания работы твердотельных приборов.

1. Если взять отдельный атом, например, кремния, то его строение можно представить как ядро, где сосредоточен положительный заряд и расположенные вокруг ядра дискретные энергетические уровни, на которых располагаются электроны. При образовании твердого тела атомы сближаются друг с другом на очень малые расстояния (менее 0,5 нм). При этом плотность атомов в твердом теле достигает 10^{22} - 10^{23} см⁻³. В результате из-за близкого расположения атомов и большой их плотности дискретные энергетические уровни расщепляются в зону разрешенных энергий. Теоретически и экспериментально показано, что зоны разрешенных энергий разделены запрещенной зоной. Таким образом, первый фрагмент зонной теории сводится к тому, что полупроводник можно представить в виде модели энергетических зон, т.е. зонной диаграммы, представленной на рис.2.1.

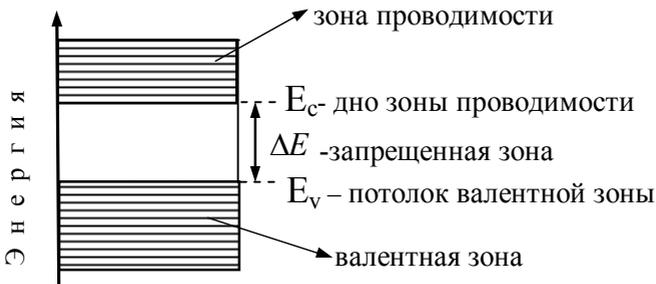


Рисунок 2.1. Зонная энергетическая диаграмма полупроводника

На диаграмме по вертикальной оси отложены значения энергии, которую может иметь электрон, по горизонтальной оси - координата. Верхняя часть энергетической диаграммы представляет собой зону проводимости, а нижний уровень этой зоны E_c называется дном зоны проводимости. Нижняя часть диаграммы - валентная зона, верхний уровень E_v которой называется потолком валентной зоны.

Запрещенная зона характеризуется шириной запрещенной зоны $\Delta E = E_c - E_v$. Для удобства будем изображать зонную диаграмму полупроводника в упрощенном варианте (рис.2.2), не обозначая осей координат и не штрихуя энергетические уровни в валентной зоне и зоне проводимости.

Надо иметь в виду, что количество энергетических уровней в валентной зоне и зоне проводимости одного порядка. Однако степень заполнения уровней электронами различна.

Так при $T = 0\text{ К}$ все уровни валентной зоны заполнены, а уровни зоны проводимости не заняты.

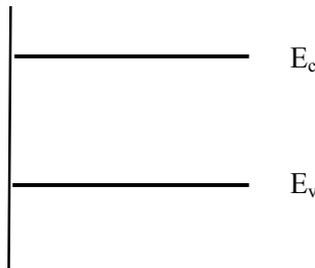


Рис. 2.2. Упрощенное изображение зонной диаграммы полупроводника

При этом электроны, находящиеся в валентной зоне, ток переносить не могут, т.к. для их перемещения нет свободных уровней. При $T > 0\text{ К}$ часть носителей может перейти из валентной зоны в зону проводимости. Поскольку количество электронов в зоне проводимости меньше числа уровней, то они являются свободными и могут перемещаться, создавая ток. Одновременно при переходе электрона из валентной зоны в зону

проводимости в валентной зоне образуется незанятый энергетический уровень. Этот незанятый уровень называют дыркой, которая имеет положительный заряд и может также перемещаться и переносить ток.

2. Второй фрагмент зонной теории, который необходим для понимания работы твердотельных приборов, заключается в том, чтобы вывести соотношения, показывающие, какое же количество электронов и дырок может находиться в разрешенных зонах в свободном состоянии, поскольку количество этих носителей определяет ток в приборах.

Для определения концентрации свободных электронов в зоне проводимости можно рассуждать следующим образом. Если известно количество уровней, расположенных в зоне проводимости, то, зная вероятность их заполнения, можно путем интегрирования по всей зоне проводимости определить количество свободных электронов n_0 , т.е.

$$n_0 = \int_{E_c}^{\infty} N_c f(E) dE, \quad (2.1)$$

где $f(E)$ - функция вероятности заполнения электроном уровня с энергией E ;

N_c - эффективная плотность энергетических состояний в зоне проводимости (количество уровней, отнесенных к единице объема). Цифра 2 в выражении (2.2) отражает, что на каждом уровне в соответствии с принципом Паули может находиться два электрона. Из квантовой теории полупроводников известно, что значение N_c определяется соотношением

$$N_c = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2}, \quad (2.2)$$

где m_n^* - эффективная масса электрона;

k, h - постоянные Больцмана и Планка соответственно;

T - температура.

Вероятность заполнения уровня с энергией E определяется соотношением Ферми-Дирака

$$f(E) = \frac{1}{\exp\left(\frac{E - E_F}{kT} + 1\right)}. \quad (2.3)$$

Для **невырожденных** полупроводников уровень Ферми всегда расположен так, что $E_c - E_F \gg kT$. Под невырожденным полупроводником будем понимать полупроводник, у которого уровень Ферми лежит в пределах запрещенной зоны. Причем большинство полупроводниковых приборов изготавливаются на основе невырожденных полупроводников.

Пренебрегая единицей в знаменателе выражения (2.3), получим функцию вероятности в виде (2.4), что говорит о том, что носители в невырожденных полупроводниках подчиняются статистике Максвелла-Больцмана

$$f(E) = \exp\left(-\frac{E - E_F}{kT}\right). \quad (2.4)$$

Тогда интегрирование выражения (2.1) дает формулу для расчета концентрации свободных электронов в зоне проводимости полупроводника

$$n_0 = N_c \exp\left(-\frac{E_c - E_F}{kT}\right). \quad (2.5)$$

Видно, что число свободных электронов (n_0) определяется кроме N_c температурой и положением уровня E_F относительно E_c : с ростом T n_0 увеличивается по экспоненте, а при увеличении $E_c - E_F$ по экспоненте уменьшается.

По аналогии с выражением (2.5) для электронов можно записать выражения для расчета концентрации свободных дырок в валентной зоне

$$p_0 = N_v \exp\left(-\frac{E_F - E_v}{kT}\right), \quad (2.6)$$

где N_v – эффективная плотность энергетических состояний в валентной зоне, определяемая как

$$N_v = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2}, \quad (2.7)$$

где m_p^* – эффективная масса дырки.

Следует иметь в виду, что N_c и N_v имеют численные значения одного порядка при заданной температуре, если значения m_n^* и m_p^* близки: $N_c, N_v \approx 10^{25} \text{ м}^{-3}$. Характер зависимости p_0 от T и $(E_F - E_v)$ такой же, что и для n_0 : с ростом T p_0 увеличивается, а с увеличением $(E_F - E_v)$ уменьшается по экспоненте.

2.2. Параметры, характеризующие свойства полупроводниковых материалов

Полупроводники - это широкий класс веществ, имеющих удельное сопротивление, лежащее между удельным сопротивлением металлов и диэлектриков, отличительной особенностью которых является очень сильная реакция на внешние воздействия (температуры, освещения, магнитных и электрических полей). Полупроводниковыми свойствами обладают 12 элементов периодической системы Менделеева (Si, Ge, Sn и др.). Их называют элементарными полупроводниками. Практическое применение из элементарных полупроводников нашли кремний и германий. Кроме элементарных полупроводников существуют полупроводниковые двойные, тройные и более сложные соединения. Среди них наибольшее применение получили двойные соединения арсенида и фосфида галлия ($GaAs, GaP$).

Для характеристики свойств полупроводниковых материалов используется более 50 параметров. Здесь остановимся на параметрах полупроводниковых материалов, необходимых для понимания работы твердотельных приборов. К их числу относятся следующие.

1. ρ -удельное сопротивление, [Ом·см]. Под удельным сопротивлением понимается сопротивление между противоположными гранями куба, вырезанного из полупроводника,

имеющего единичный размер грани (например, 1 сантиметр или 1 метр).

По определению $\rho_m < \rho_{n/n} < \rho_o$ (10^{-6} - 10^{-4}) $<$ (10^{-4} - 10^{10}) $<$ ($>10^{10}$ Ом·см).

Величина, обратная удельному сопротивлению, называется удельной электропроводностью или удельной проводимостью

$$\sigma \left[\frac{C_M}{c_M} \right]; \rho = \frac{1}{\sigma}.$$

2. ΔE – ширина запрещенной зоны, [эВ]. Под шириной запрещенной зоны понимается энергетический зазор между дном зоны проводимости E_c и потолком валентной зоны E_v . Ширина запрещенной зоны зависит от температуры.

$$\Delta E(T) = \Delta E(0) - \eta T, \quad (2.8)$$

где $\Delta E(0)$ – ширина запрещенной зоны при $T = 0$ К;

η – температурный коэффициент ширины запрещенной зоны. Для германия и кремния $\eta = 2,4 \cdot 10^{-4}$ эВ/К.

Уменьшение ΔE с ростом температуры обусловлено увеличением амплитуды колебаний узлов кристаллической решетки. Для наиболее широко используемых полупроводниковых материалов ΔE имеет следующие значения: $\Delta E_{Ge} = 0,72$ эВ, $\Delta E_{Si} = 1,12$ эВ и $\Delta E_{GaAs} = 1,43$ эВ при $T = 300$ К.

3. μ_n, μ_p – подвижность носителей, $\left[\frac{cM^2}{B \cdot c} \right]$. По определению

нию $\mu = \frac{v_{dp}}{E}$, т.е. подвижность – это скорость дрейфа носителей

при воздействии единичного электрического поля. Таким образом, с подвижностью носителей в полупроводниках ассоциируется скорость перемещения зарядов.

По поводу подвижности необходимо знать следующее:

- подвижность электронов во всех полупроводниках всегда больше подвижности дырок ($\mu_n > \mu_p$);

• подвижность электронов и дырок является функцией концентрации легирующей примеси и температуры ($\mu_{n,p} = f(N, T)$).

При увеличении концентрации примеси возрастает число рассеивающих центров и снижается скорость перемещения носителей зарядов. Существует эмпирическая формула, связывающая подвижность с концентрацией

$$\mu = \mu_1 + \frac{\mu_2}{1 + \left(\frac{N^+}{N}\right)^\alpha}, \quad (2.9)$$

где N^+ – количество рассеивающих центров, см^{-3} .

Значения параметров (формула 2.9) для кремния и германия представлены в табл.2.1.

Таблица 2.1 Параметры для расчета подвижности носителей

Материал	П а р а м е т р ы			
	μ_1 , $\text{см}^2/\text{В}\cdot\text{с}$	μ_2 , $\text{см}^2/\text{В}\cdot\text{с}$	N^0 , см^{-3}	α
Кремний:				
	электроны	65	1265	$8,5 \cdot 10^{16}$
дырки	48	447	$6,3 \cdot 10^{16}$	0,76
Германий:				
	электроны	50	3850	$8,1 \cdot 10^{16}$
дырки	42	1860	$1,4 \cdot 10^{17}$	0,43

По данной формуле рассчитаны графические зависимости $\mu_n, \mu_p = f(N^+)$, представленные на рис.2.3.

В диапазоне рабочих температур полупроводниковых приборов подвижность уменьшается с ростом температуры в основном за счет решеточного рассеяния, причем $\mu \sim T^{-3/2}$.

4. τ_n, τ_p – время жизни носителей, [с].

Если, например, под воздействием освещения в зоне проводимости полупроводника появятся избыточные электроны, то

их концентрация будет уменьшаться во времени за счет процессов рекомбинации по следующему закону

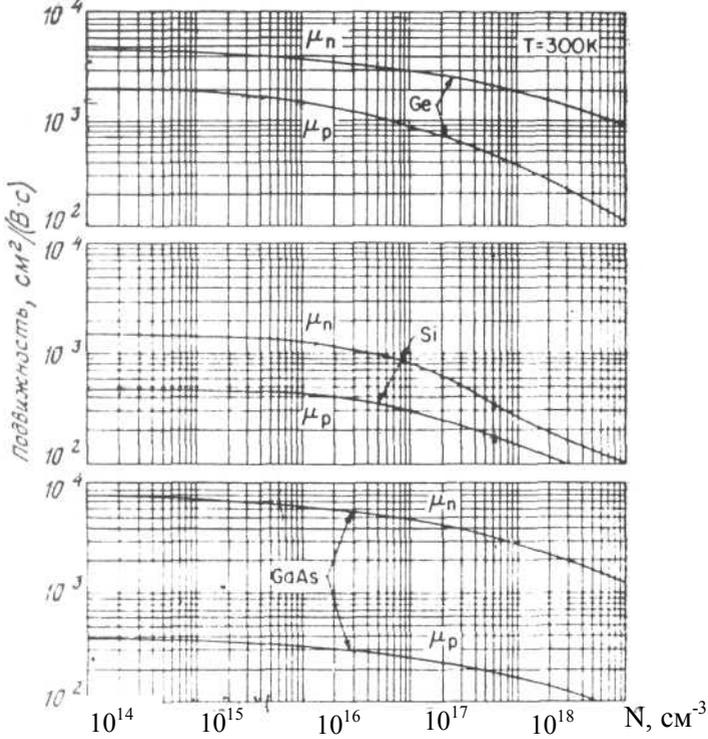


Рис. 2.3. Зависимость подвижности электронов и дырок в $GaAs$, Ge и Si от концентрации легирующей примеси

$$\Delta n(t) = \Delta n(0) \exp\left(-\frac{t}{\tau_n}\right) \quad (2.10)$$

Тогда под **временем жизни** τ_n надо понимать время, в течение которого начальная концентрация избыточных электронов $\Delta n(0)$ уменьшается в e раз. Время жизни зависит от концентрации легирующей примеси и количества дефектов в полупроводнике: с ростом концентрации примеси время жизни уменьшается, т.к. примеси выступают в роли центров рекомбинации.

В случае рекомбинационных ловушек одного типа время жизни носителей возрастает с увеличением температуры. Это связано с тем, что при увеличении температуры уровень Ферми смещается ближе к середине запрещенной зоны, и вероятность захвата электронов на ловушки с последующей рекомбинацией уменьшается, т.е. вызывает рост τ .

5. D_n, D_p – коэффициенты диффузии электронов и дырок,

$$\left[\frac{\text{см}^2}{\text{с}} \right].$$

Если в полупроводниковом материале существует градиент концентрации носителей, например, электронов, то возникает диффузионный поток Φ_n . Величина этого потока определяется градиентом концентрации

$$\Phi_n = -D_n \text{grad } n \quad (2.11)$$

Коэффициент пропорциональности между Φ_n и $\text{grad } n$ называется коэффициентом диффузии D_n . Численное значение коэффициента диффузии показывает число частиц, проходящих через единичную площадку (например, 1 см^2), расположенную перпендикулярно вектору потока, за единицу времени.

6. L_n, L_p – длина диффузионного смещения, [см].

Если в полупроводнике существует градиент концентрации, то носители перемещаются в сторону уменьшения концентрации. При этом за счет процессов рекомбинации концентрация избыточных носителей уменьшается не только во времени, но и в пространстве по экспоненциальному закону

$$\Delta n(x) = \Delta n(0) \exp - \frac{x}{L_n} \quad (2.12)$$

Тогда под *длиной диффузионного смещения* L_n понимается расстояние, на котором концентрация носителей уменьшается в e раз.

Типичные значения $L_{n,p}$ составляют 10^{-3} - 10^{-1} см.

Перечисленные выше параметры (μ, τ, L, D) связаны между собой следующими соотношениями

$$\left. \begin{aligned} D_n &= \frac{kT}{q} \mu_n \\ D_p &= \frac{kT}{q} \mu_p \end{aligned} \right\} \text{ - соотношение Эйнштейна; } \quad (2.13)$$

$$L_n = \sqrt{D_n \tau_n}, \quad L_p = \sqrt{D_p \tau_p} \quad (2.14)$$

Данные соотношения широко используются при решении многих задач твердотельной электроники.

Величина $\frac{kT}{q}$ называется тепловым потенциалом. Он равен 0,0258 В. при 300 К

7. n_i – собственная концентрация носителей, [см⁻³].

В собственном полупроводнике, зонная диаграмма которого приведена на рис. 2.4., свободные электроны и дырки могут появиться только за счет перехода определенной части электронов из валентной зоны в зону проводимости. При этом в валентной зоне образуется такое же количество свободных дырок. Количество свободных носителей зависит от температуры и ширины запрещенной зоны полупроводника.

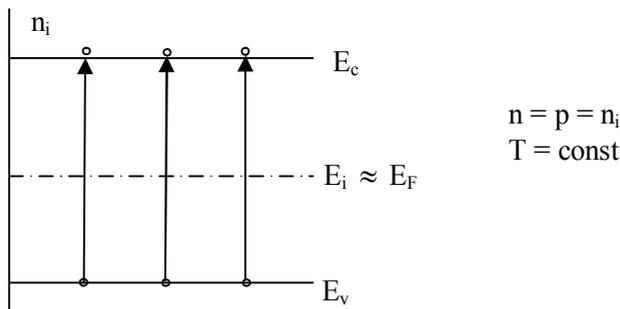


Рисунок 2.4. Зонная диаграмма собственного полупроводника

Таким образом, **собственная концентрация носителей** n_i – это концентрация свободных электронов и дырок в собственном полупроводнике при заданной температуре. Численные значения n_i составляют: для германия - $2,5 \cdot 10^{13} \text{ см}^{-3}$, для кремния - 10^{10} см^{-3} при $T=300 \text{ К}$.

8. E_{np} – электрическая прочность, [В/см].

Если образец полупроводника разместить между металлическими обкладками и подать на него изменяющееся пилообразное напряжение, то ток через структуру будет изменяться по закону Ома: $j = \sigma E$.

Однако при некоторой напряженности поля E в полупроводнике произойдет резкое возрастание тока j (Рис. 2.5). Явление резкого возрастания тока называют **пробоем**, а напряженность поля, при которой происходит пробой - **электрической прочностью**. Значения E_{np} для кремния и германия составляют приблизительно $8 \cdot 10^5 \text{ В/см}$ и $3 \cdot 10^5 \text{ В/см}$ соответственно.

9. ε – диэлектрическая проницаемость материала. Величина безразмерная.

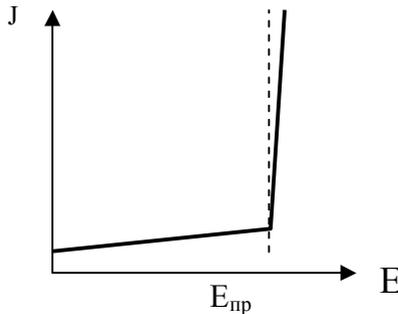


Рис. 2.5. Зависимость тока через полупроводник от напряженности электрического поля.

Характеризует материальную среду, в которой происходит перемещение носителей зарядов. $\varepsilon_{Si} = 12$, $\varepsilon_{Ge} = 16$, $\varepsilon_{GaAs} = 11$.

2.3. Фундаментальная система уравнений твердотельной электроники

Описание физических явлений и процессов в твердотельных приборах может быть осуществлено двояко. Во-первых, они могут быть рассмотрены на качественном уровне. С другой стороны, зачастую, качественного описания процессов бывает недостаточно и требуется строгое количественное описание явлений. Кроме того, расчет параметров твердотельных приборов требует наличия значительного количества расчетных формул и соотношений. Наконец, в некоторых случаях описание процессов и явлений, происходящих в приборах, возможно только путем математического анализа моделей. Аналитическое описание физических процессов в твердотельных приборах и получение различных расчетных соотношений базируется в основном на использовании уравнений полного тока, непрерывности и Пуассона, с использованием других известных соотношений. При рассмотрении данного вопроса важно уяснить физический смысл фундаментальных уравнений твердотельной электроники.

Уравнение полного тока. Ток - это направленное движение заряженных частиц. В полупроводниках ток может протекать за счет движения двух видов заряженных частиц - электронов и дырок. Для того, чтобы заряженные частицы могли перемещаться, необходимо наличие сил, которые бы заставили частицы направленно перемещаться. В твердотельных приборах направленное движение заряженных частиц осуществляется за счет действия сил электрического поля и градиента концентрации.

Ток, обусловленный движением заряженных частиц под действием сил электрического поля, называется **дрейфовым**. Величина дрейфового тока определяется количеством носителей и скоростью их перемещения

$$j_n^{dp} = qn\mu_n E; \quad j_p^{dp} = qp\mu_p E. \quad (2.15)$$

При этом надо иметь в виду, что по определению напряженность электрического поля $E = -\frac{d\varphi}{dx}$ и вектор напряженности всегда направлен от плюса к минусу. На рис.2.6 показаны направления движения электронов и дырок в электрическом поле и вектора дрейфовых токов.

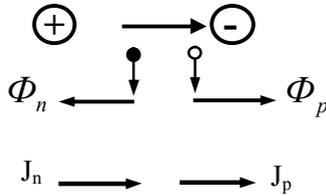


Рис.2.6. Вектора потоков Φ и токов I электронов и дырок под действием электрического поля.

Диффузионный ток - ток, обусловленный направленным движением электронов и дырок за счет градиента концентрации.

Если в какой-то части кристалла полупроводника большая концентрация электронов, а в другой части их мало, то возникает диффузионный поток электронов, создающий диффузионный ток. Величина диффузионного тока определяется градиентом концентрации

$$\begin{aligned} j_n^{diff} &= qD_n \text{grad } n, \\ j_p^{diff} &= -qD_p \text{grad } p. \end{aligned} \quad (2.16)$$

Знак "минус" в выражении для j_p^{diff} связан с тем, что вектора тока и градиента направлены в разные стороны. Рис.2.7 разъясняет это утверждение.

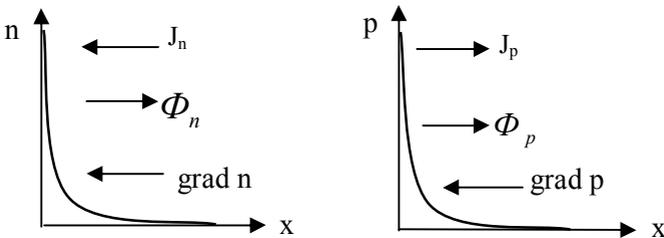


Рис. 2.7. Направление векторов потоков Φ , токов и градиента для диффузии электронов и дырок.

Таким образом, физический смысл уравнения полного тока заключается в следующем: ток в полупроводниковых приборах переносится электронами и дырками, а перемещаются они под действием сил электрического поля и градиента концентрации

$$j = j_n + j_p = j_n^{op} + j_n^{du\phi} + j_p^{op} + j_p^{du\phi},$$

$$j = qn\mu_n E + qp\mu_p E + qD_n grad n - qD_p grad p. \quad (2.17)$$

Уравнение (2.17) и является уравнением полного тока. Если быть строгим, то в выражение (2.17) надо добавить еще одну компоненту тока - ток смещения, определяемый как производная по времени от вектора электрической индукции

$$j_{cm} = \frac{d}{dt}(\epsilon\epsilon_0 E). \quad (2.18)$$

Причем, как увидим в дальнейшем, в большинстве полупроводниковых приборов на основе $p-n$ переходов ток по своей природе является диффузионным и для его нахождения необходимо уметь определять градиент концентрации

Как это делается, показано на рис.2.8.

$$grad n = \frac{n(x_1) - n(x_2)}{\Delta x} \quad (2.19)$$

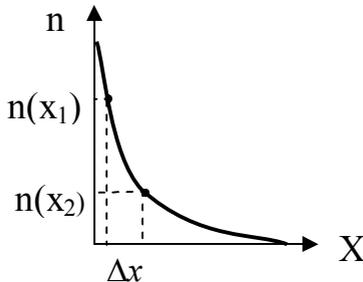


Рис. 2.8. Определение градиента концентрации электронов.

Уравнение Пуассона. Для решения многих задач в полупроводниковой электронике используется уравнение Пуассона. Для одномерного случая оно имеет следующий вид

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\epsilon\epsilon_0}, \quad (2.20)$$

где $\rho(x)$ — плотность заряда;

ϵ_0 — электрическая постоянная.

Физический смысл этого уравнения сводится к тому, что,

если известна плотность заряда $\rho(x)$, то значение напряженности электрического поля E и потенциала φ может быть определено в любой точке пространства путем решения уравнения Пуассона при заданных граничных условиях.

Уравнение непрерывности. Величина тока в полупроводниковых приборах определяется концентрацией носителей и градиентом концентрации. При этом концентрация носителей может изменяться во времени. Характер изменения концентрации носителей во времени позволяет описывать уравнение непрерывности. Это уравнение показывает, как и по каким причинам изменяется концентрация носителей во времени. Физический смысл уравнения можно представить следующей схемой.

$$\left[\begin{array}{l} \text{Изменение} \\ \text{концентрации} \\ \text{носителей} \\ \text{во времени} \end{array} \right] = \left[\begin{array}{l} \text{процессы} \\ \text{генерации} \end{array} \right] - \left[\begin{array}{l} \text{процессы} \\ \text{рекомбинации} \end{array} \right] - \left[\begin{array}{l} \text{растекание} \\ \text{тока} \end{array} \right]$$

То есть, за счет процессов генерации концентрация носителей во времени возрастает, а рекомбинация и растекание тока приводят к уменьшению концентрации носителей.

Математически это выглядит следующим образом.

$$\left. \begin{array}{l} \frac{dn}{dt} = G_n - R_n + \frac{1}{q} \operatorname{div} j_n \\ \frac{dp}{dt} = G_p - R_p - \frac{1}{q} \operatorname{div} j_p, \end{array} \right\} \quad (2.21)$$

где $G_n = \frac{n_0}{\tau_n}$; $G_p = \frac{p_0}{\tau_p}$ — скорости генерации;

$R_n = \frac{n}{\tau_n}$; $R_p = \frac{p}{\tau_p}$ — скорости рекомбинации;

$\operatorname{div} j_n = \frac{dj_n}{dx}$; $\operatorname{div} j_p = \frac{dj_p}{dx}$ — дивергенция токов.

Подставляя значения $G_n, G_p, R_n, R_p, \operatorname{div} j_n, \operatorname{div} j_p$ в выражение (2.21), получим наиболее распространенную форму записи

уравнения непрерывности

$$\left. \begin{aligned} \frac{dn}{dt} &= -\frac{n-n_0}{\tau_n} + \frac{1}{q} \operatorname{div} j_n \\ \frac{dp}{dt} &= -\frac{p-p_0}{\tau_p} - \frac{1}{q} \operatorname{div} j_p \end{aligned} \right\} \quad (2.22)$$

Если токи носят диффузионный характер, т.е. $j_n = qD_n \frac{dn}{dx}$

и $j_p = -qD_p \frac{dp}{dx}$, то уравнение (2.22) можно представить в виде

дифференциальных уравнений второго порядка

$$\left. \begin{aligned} \frac{dp}{dt} &= D_p \frac{d^2 p}{dx^2} - \frac{p-p_0}{\tau_p} \\ \frac{dn}{dt} &= D_n \frac{d^2 n}{dx^2} - \frac{n-n_0}{\tau_n} \end{aligned} \right\} \quad (2.23)$$

2.4. Собственные, примесные и компенсированные полупроводники

Собственный полупроводник - это полупроводник, не содержащий примеси, либо содержащий ее в таком количестве, которое не влияет на свойства полупроводника. В собственном полупроводнике свободные электроны могут появиться в зоне проводимости только за счет их перехода из валентной зоны. При этом появление свободного электрона в зоне проводимости сопровождается появлением незанятого энергетического уровня в валентной зоне - дырки. В таком полупроводнике количество свободных электронов равно количеству свободных дырок. Приравнивая выражения (2.5) и (2.6) для n и p и принимая

$$\frac{E_c - E_v}{2} = E_i, \text{ получим выражение, позволяющее оценивать по-}$$

ложение уровня Ферми в собственном полупроводнике

$$E_F = E_i + \frac{kT}{2} \ln \frac{N_v}{N_c} \quad (2.24)$$

или с учетом выражений для N_v и N_c

$$E_F = E_i + \frac{3kT}{4} \ln \frac{m_p^*}{m_n^*}. \quad (2.25)$$

Эти выражения показывают, что в собственном полупроводнике уровень Ферми лежит вблизи середины запрещенной зоны E_i и отклоняется от E_i в зависимости от соотношения эффективных масс электрона и дырки. При $m_n^* = m_p^*$ он лежит точно посередине запрещенной зоны.

Так как $n = p = n_i$, то можно записать с учетом того, что $n \cdot p = n_i^2$.

$$n_i^2 = N_c N_v \exp - \frac{\Delta E}{kT}, \quad (2.26)$$

где $\Delta E = E_c - E_v$.

Тогда
$$n_i = \sqrt{N_c N_v} \exp - \frac{\Delta E}{2kT}, \quad (2.27)$$

т.е. собственная концентрация носителей зависит от ширины запрещенной зоны полупроводника и температуры.

Подчеркивая сильную температурную зависимость n_i , можно записать выражение для n_i , подставив в выражение (2.27) значения N_c и N_v

$$n_i^2 = BT^3 \exp \left(- \frac{\Delta E}{kT} \right), \quad (2.28)$$

где B – константа, включающая в себя все параметры N_c и N_v , кроме температуры. Причем надо иметь в виду, что экспоненциальная зависимость более сильная, чем степенная, т.е. n_i^2 зависит от температуры по экспоненциальному закону.

Схема, поясняющая появление свободных электронов и дырок в собственном полупроводнике, представлена на рисунке 2.9.

Поскольку в собственном полупроводнике концентрации

электронов и дырок равны, то электропроводность такого полупроводника равна.

$$\sigma_i = q\mu_n n + q\mu_p p = q(\mu_n + \mu_p) n_i. \quad (2.29)$$

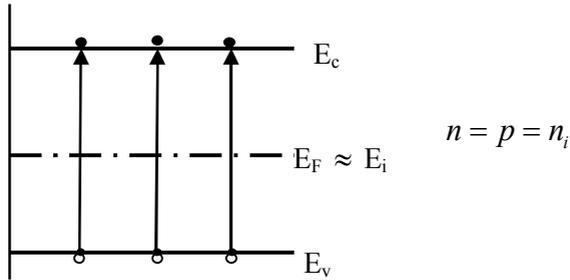


Рис. 2.9. Схема, поясняющая появление свободных электронов и дырок в собственном полупроводнике

Примесные полупроводники - это полупроводники, в которые введена примесь. В зависимости от того, какая примесь вводится в полупроводник, они подразделяются на электронные и дырочные. Рассмотрим, как образуется электронный полупроводник на примере основных материалов электроники - германия или кремния.

***n* – тип (электронный, донорный).** Если в беспримесный полупроводник ввести примесь с валентностью на единицу большую валентности основного вещества, т.е. элемент пятой группы таблицы Менделеева (поскольку германий и кремний четырехвалентные элементы), то происходит следующее. Четыре валентных электрона атома основного вещества объединяются с четырьмя валентными электронами атома примеси и создают устойчивую электронную оболочку. При этом один электрон атома примеси оказывается незадействованным в этой сильной связи, и он слабо связан с атомом примеси. При наличии внешних воздействий, например температуры, слабосвязанный электрон может оторваться от атома примеси и стать свободным. Такой электрон, свободно перемещаясь по кристаллу, создает электронный тип проводимости. Поэтому такие полупроводники называют электронными или полупроводниками *n* – типа проводимости. Акт отрыва слабосвязанного электрона от атома

примеси называется ионизацией примеси. Очевидно, что при условии полной ионизации примеси, т.е. когда от каждого атома примеси оторвался один электрон, число свободных электронов равно количеству введенной в полупроводник примеси: $n = N_q$.

Если электрон оторвался от атома примеси, то атом превращается в положительно заряженный ионизованный донор N_q^+ . Несмотря на то, что донор заряжен, участие в токопереносе он не принимает, так как находится либо в узлах кристаллической решетки, либо в междоузлиях и неподвижен. Обычно в качестве донорной примеси в германии и кремнии используется мышьяк и фосфор.

Донор - это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, занятый в невозбужденном состоянии электроном и способный в возбужденном состоянии отдать электрон в зону проводимости. Зонная диаграмма полупроводника n -типа представлена на рис. 2.10.

Важно, что в таких полупроводниках уровень Ферми лежит выше уровня E_i .

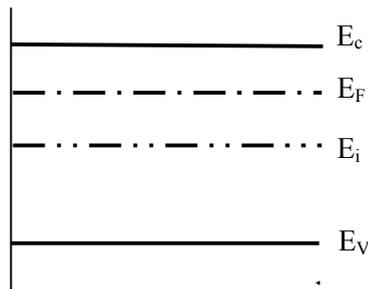


Рис. 2.10. Зонная диаграмма полупроводника n -типа.

Тогда, в соответствии с выражением (2.5), для электронов при постоянной температуре необходимо, чтобы зазор $E_c - E_F$ был меньше, чем в собственном.

Поскольку в собственном полупроводнике $E_c - E_F = \frac{\Delta E}{2} = E_i$, то из этого следует, что в полупроводнике n -типа уровень Ферми должен быть

ближе к E_c , т.е. лежать выше уровня E_i .

Количественно сдвиг уровня Ферми относительно E_i зависит от концентрации введенной легирующей примеси по отношению к собственному полупроводнику.

$$E_F = E_i + kT \ln \frac{N_q}{n_i}. \quad (2.30)$$

Выражение (2.30) показывает, что в электронном полупроводнике уровень Ферми лежит выше середины запрещенной зоны на величину, определяемую концентрацией введенной донорной примеси. Таким образом, положение уровня Ферми в полупроводнике задается концентрацией примеси.

Из выражения (2.5), полагая, что при полной ионизации примеси $n = N_q$, можно записать еще одно выражение для определения уровня Ферми в полупроводнике n – типа

$$E_F = E_c - kT \ln \frac{N_c}{N_q}. \quad (2.31)$$

Это выражение говорит о том, что в полупроводнике n – типа уровень Ферми лежит ниже уровня E_c на величину, определяемую отношением $\frac{N_c}{N_q}$, т.е. определяется концентрацией легирующей примеси.

Электропроводность электронного полупроводника задается формулой

$$\sigma_n = q\mu_n n. \quad (2.32)$$

p – тип (дырочный, акцепторный). Если в беспримесный полупроводник ввести примесь с валентностью на единицу меньше валентности основного вещества, т.е. элемент третьей группы, то четыре валентных электрона атома основного вещества притянут три валентных электрона атома примеси. Однако для образования устойчивой оболочки необходимо восемь электронов. Недостающий электрон притягивается с предпоследней электронной оболочки. Это приводит к образованию незанятого энергетического уровня - дырки. Дырка оказывается слабосвя-

занной с атомом примеси и может оторваться от него и стать свободной. Перемещаясь свободно по кристаллу, дырки создают дырочный тип проводимости. Такие полупроводники называют дырочными или полупроводниками p – типа проводимости. При условии полной ионизации акцепторной примеси число свободных дырок определяется концентрацией акцепторной примеси $p = N_a$. В ионизованном состоянии акцепторная примесь отрицательно заряжена (N_a^-), но, так же как и ионизованный донор, в токопереносе участия не принимает. Обычно в качестве акцепторной примеси выступают Ga, Jn, Al .

Акцептор - это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, свободный от электрона в невозбужденном состоянии и способный захватить электрон из валентной зоны в возбужденном состоянии.

Зонная энергетическая диаграмма полупроводника p – типа представлена на рис.2.11. В дырочных полупроводниках концентрация дырок больше, чем в собственных, поэтому в соответствии с выражением (2.6) уровень Ферми должен лежать ниже уровня E_i . Количественно сдвиг уровня E_F по отношению к E_i определяется концентрацией легирующей акцепторной примеси N_a

$$E_F = E_i - kT \ln \frac{N_a}{n_i} . \quad (2.33)$$

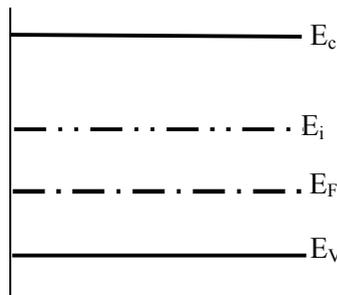


Рисунок 2.11. Зонная диаграмма полупроводника p – типа

Из выражения (2.36) для дырок, полагая, что при условии полной ионизации акцепторов $P = N_a$ можно записать

$$E_F = E_v + kT \ln \frac{N_v}{N_a} . \quad (2.34)$$

Выражения (2.33) и (2.34) говорят о том, что в полупроводниках p – типа уровень Ферми лежит ниже уровня E_i или выше уровня E_v на величину, определяемую концентрацией легирующей примеси N_a .

Электропроводность дырочного полупроводника определяется соотношением

$$\sigma_p = q\mu_p p . \quad (2.35)$$

Компенсированные полупроводники. Если в полупроводнике одновременно присутствуют донорная и акцепторная примесь, то такой полупроводник называют компенсированным. При этом тип проводимости полупроводника определяется наибольшей концентрацией легирующей примеси, а концентрация свободных основных носителей задается разностной концентрацией. Если $N_q > N_a$, то полупроводник обладает электронным характером проводимости и для него пригодны все расчетные формулы и соотношения, свойственные полупроводникам n – типа. При этом концентрация свободных электронов $n = N_q - N_a = N^1$,

где N^1 – разностная концентрация.

2.5. Диапазон рабочих температур полупроводниковых приборов

В этом разделе мы должны дать ответ на вопрос, почему полупроводниковые приборы принципиально не могут работать выше каких-то температур и ниже определенных температур.

Для этого необходимо рассмотреть зависимость электропроводности полупроводников от температуры. Она имеет вид, представленный на рис. 2.12, характерный для полупроводников p – и n – типов.

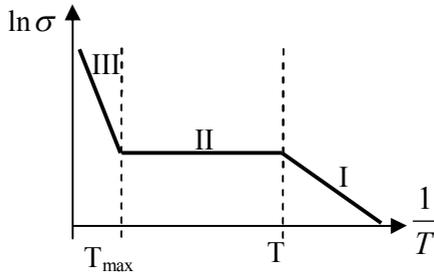


Рис. 2.12. Зависимость электропроводности полупроводников от температуры

Зависимость $\ln \sigma = f(1/T)$ имеет три участка. На участке I, называемом участком примесной проводимости, σ растет при увеличении температуры T , что связано с увеличением концентрации носителей, становящихся свободными в результате ионизации примеси (рис.2.13).

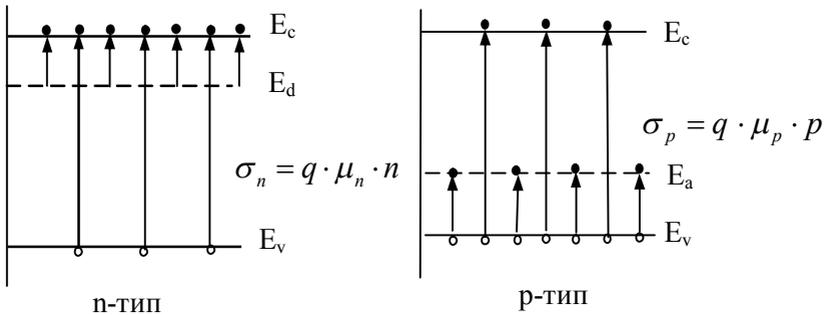


Рис. 2.13. Рисунок, показывающий изменение концентрации свободных электронов и дырок в примесных полупроводниках.

Поскольку $\sigma_n = q\mu_n n$ и $\sigma_p = q\mu_p p$, то, пренебрегая температурной зависимостью подвижности носителей, можно говорить о том, что рост электропроводности обусловлен увеличением концентрации носителей (электронов в n -полупроводнике и дырок в p -полупроводнике) за счет их поступления с примесных уровней.

При некоторой температуре T все уровни акцепторов в полупроводнике p -типа будут заняты электронами, поступаю-

щими на них из валентной зоны, а все донорные уровни в n – полупроводнике отдадут электроны в зону проводимости. Так что дальнейшее увеличение температуры не дает увеличения σ_n и σ_p , поскольку с примесных уровней все носители ушли, а переход электронов из валентной зоны в зону проводимости еще не возможен из-за малой их энергии. Эту область II называют участком истощения примеси.

При дальнейшем увеличении температуры ($T > T_{\max}$) энергии электронов становится достаточно, чтобы перейти из валентной зоны в зону проводимости с образованием свободного электрона и свободной дырки, что соответствует процессам, наблюдаемым в собственных полупроводниках. То есть при $T > T_{\max}$ полупроводники n – и p – типа проводимости становятся собственными.

Рис.2.14 показывает характер изменения положения уровня Ферми. Поэтому участок III называют участком собственной проводимости.

Поскольку абсолютное большинство полупроводниковых приборов работает на примесных полупроводниках, то их переход к собственной проводимости означает, что они становятся неработоспособными.

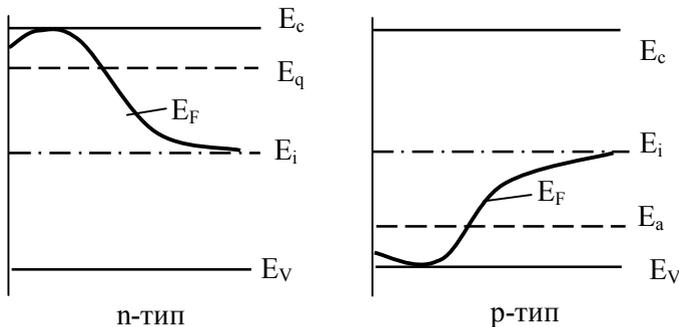


Рис. 2.14. Изменение положения уровня Ферми в примесных полупроводниках при изменении температуры.

Основные классы полупроводниковых приборов в качестве

основного элемента содержат один или несколько электронно-дырочных переходов. При $T > T_{\max}$ полупроводник n – типа становится собственным, и полупроводник p – типа становится также собственным, т.е. $p - n$ переход перестает существовать. А это означает, что при таких температурах не могут работать полупроводниковые приборы, содержащие $p - n$ переход, как основной элемент прибора.

T_{\max} – максимальная рабочая температура полупроводниковых приборов определяется температурой перехода полупроводников к участку собственной проводимости. Для приборов из кремния $T_{\max} \leq 150^\circ\text{C}$, а для приборов из германия $T_{\max} \leq 100^\circ\text{C}$. Разница в максимальной рабочей температуре определяется различием в ширине запрещенной зоны: чем больше ΔE , тем выше температура, при которой полупроводник становится собственным.

Теперь дадим ответ на вопрос, почему полупроводниковые приборы не могут работать ниже каких-то температур. Чтобы существовал $p - n$ переход, необходимо наличие контакта полупроводников с различным типом проводимости. Образование полупроводников n – и p – типов, рассмотренное ранее, предполагает наличие процесса ионизации примеси, т.е. отрыв слабосвязанных электронов и дырок от атомов примеси. В отсутствии других возмущающих факторов этот процесс происходит за счет температуры. Если температура будет ниже какой-то T_{\min} , то ионизация примеси не будет происходить, т.е. слои n – и p – типов образовываться не смогут, а это делает невозможным существование электронно-дырочного перехода.

T_{\min} – минимальная рабочая температура полупроводниковых приборов, определяемая невозможностью ионизации примесей. Для приборов из германия и кремния она составляет минус $(70-80)^\circ\text{C}$.

2.6. Равновесные и неравновесные носители зарядов в полупроводниках. Основные и неосновные носители.

Закон действующих масс

В любом полупроводнике присутствуют электроны и дырки. Так в полупроводнике n – типа есть электроны, которые называются основными носителями, и есть дырки, которые называются неосновными носителями заряда. В полупроводнике p – типа дырки являются основными носителями, а электроны – неосновными.

Основные носители – это носители, определяющие характер электропроводности полупроводника. Поскольку $\sigma_n = q\mu_n n$, а $\sigma_p = q\mu_p p$, то электроны в n – полупроводнике и дырки в p – полупроводнике являются основными носителями. Чтобы не возникало недоразумений, для обозначения основных и неосновных носителей используются следующие символы:

n_n – концентрация электронов в n – полупроводнике;

p_n – концентрация дырок в n – полупроводнике;

p_p – концентрация дырок в p – полупроводнике.

n_p – концентрация электронов в p – полупроводнике;

Таким образом, n_n и p_p – основные носители заряда, а n_p и p_n – неосновные носители заряда.

Соотношение между концентрациями основных и неосновных носителей в полупроводнике задается законом действующих масс, рассматриваемым ниже в данном разделе.

При постоянной температуре процессы генерации и рекомбинации носителей уравниваются, и полупроводник находится в состоянии термодинамического равновесия. Концентрации свободных носителей в полупроводнике, соответствующие термодинамическому равновесию, называются равновесными и обозначаются $n_{n0}, p_{p0}, n_{p0}, p_{n0}$.

При воздействии на полупроводник нетеплового внешнего энергетического фактора (освещение, сильное электрическое поле и т.д.) происходит дополнительная к тепловой генерация носителей заряда. Появляются неравновесные носители зарядов.

Их концентрации будем обозначать n_n, p_p, p_n, n_p .

Разница между неравновесной и равновесной концентрациями называется избыточной концентрацией и обозначается Δn и Δp . При этом

$$\Delta n = n_n - n_{n0};$$

$$\Delta p = p_p - p_{p0}.$$

Причем избыточная концентрация носителей в полупроводнике может возникнуть не только за счет нетепловой генерации, но и в результате различных процессов, таких как инжекция, экстракция и т.д.

В собственном полупроводнике концентрации электронов и дырок можно записать следующим образом

$$\begin{aligned} n &= N_c \exp\left(-\frac{E_c - E_F}{kT}\right) = N_c \exp\left(-\frac{E_c - E_i + E_i - E_F}{kT}\right) = \\ &= n_i \exp\left(-\frac{E_i - E_F}{kT}\right), \end{aligned} \quad (2.36)$$

так как $N_c \exp\left(-\frac{E_c - E_i}{kT}\right) = n_i$.

Аналогично

$$\begin{aligned} p &= N_v \exp\left(-\frac{E_F - E_v}{kT}\right) = N_v \exp\left(-\frac{E_F - E_i + E_i - E_v}{kT}\right) = \\ &= p_i \exp\left(-\frac{E_F - E_i}{kT}\right). \end{aligned} \quad (2.37)$$

так как $N_v \exp\left(-\frac{E_i - E_v}{kT}\right) = p_i$.

Тогда произведение $n \cdot p = n_i^2$. Это выражение и представляет собой закон действующих масс, который обычно записывается так

$$n \cdot p = n_i^2 = const. \quad (2.38)$$

Суть закона заключается в том, что для данного полупроводника при заданной температуре произведение концентрации

электронов на концентрацию дырок есть величина постоянная, равная квадрату собственной концентрации. Действительно, $n_i^2 = N_c N_v \exp\left(-\frac{\Delta E}{kT}\right)$ для данного полупроводника ($\Delta E = const$) при заданной температуре ($T = const$) есть константа.

Соотношение (2.38) выполняется и для примесных полупроводников. Так для полупроводника n – типа оно имеет вид

$$n_{n0} p_{n0} = n_i^2 = const ,$$

а для p – типа $p_{p0} n_{p0} = n_i^2 = const$. (2.39)

Причем закон действующих масс выполняется только для равновесных носителей.

2.7. Полупроводники в электрическом поле

Работа полупроводниковых приборов связана с воздействием на полупроводник или полупроводниковую структуру внешнего напряжения. При этом в указанных структурах возникают области, где полупроводник находится под воздействием электрического поля. При отсутствии электрических полей зонная энергетическая диаграмма выглядит так, что уровни энергетических зон проходят горизонтально (рис.2.15 а). Это говорит о том, что в любой точке объема полупроводника энергия одинакова. Напомним, что энергетическая диаграмма - это зависимость энергии носителей (вертикальная ось) от координаты (горизонтальная ось). Если электрон находится, например, на уровне E_c , то такой энергией он обладает и на поверхности и в объеме.

При помещении полупроводника в электрическое поле энергия носителей уже не будет постоянной, если электрон находится на каком-то уровне, поскольку на носители действует электрическое поле. Энергетическая диаграмма полупроводника, помещенного в электрическое поле, представлена на рис.2.15 б, в. Угол наклона зон зависит от напряженности поля, а направление наклона - от направления вектора напряженности. Важно, что энергия электронов вблизи электрода с отрицательным по-

тенциалом больше, чем у положительного.

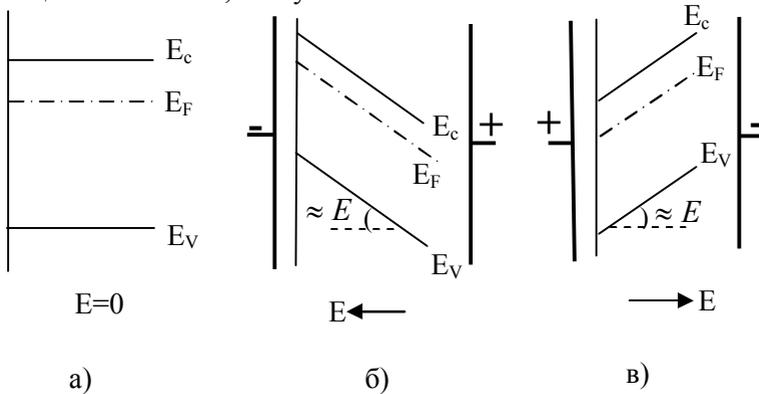


Рис. 2.15. Зонные диаграммы полупроводника n -типа вне поля (а), в электрических полях разного направления (б, в).

В дальнейшем при рассмотрении работы полупроводниковых приборов, ответ на вопрос, в какой области полупроводниковой структуры имеется электрическое поле, дает энергетическая диаграмма прибора: в области структуры, где имеется наклон зон, там присутствует электрическое поле. По углу наклона энергетических зон можно судить о величине поля, а по направлению наклона - о направлении вектора напряженности.

При работе полупроводниковых приборов в них часто возникают сильные электрические поля. Рассмотрим некоторые явления, возникающие в приборах в области сильных полей.

Ударная ионизация. Свободный электрон или дырка под действием сил электрического поля набирают энергию W

$$W = qE\lambda,$$

где λ – длина свободного пробега, которую надо понимать как расстояние между двумя актами взаимодействия.

Если энергия, приобретенная носителем в поле, превысит энергию ионизации атома примеси или собственного атома полупроводника W_i , то произойдет процесс ударной ионизации, в результате которого появятся дополнительные носители заряда. Критерий ударной ионизации $W > W_i$. Каждый вновь появившийся носитель может также набрать энергию, большую W_i , и

произвести акт ударной ионизации. В результате такого процесса может произойти лавинное размножение носителей зарядов.

Количественно процесс ударной ионизации характеризуется коэффициентом ударной ионизации α . Численное значение α показывает количество актов ионизации, произведенное единичным носителем на единичном пути. Этот параметр зависит от напряженности электрического поля E . Для расчетов обычно используют следующую аппроксимацию

$$\alpha = A/E^m, \quad (2.40)$$

где A – некоторый коэффициент;

m – показатель степени, имеющий значения от 5 до 8 в зависимости от материала.

Процесс лавинного размножения носителей количественно характеризуется параметром M – коэффициентом лавинного размножения, который может быть определен как отношение количества частиц на входе к их количеству на выходе.

Туннелирование. В области сильных электрических полей из-за большого наклона энергетических зон (рис.2.16) возможен процесс перехода электронов из валентной зоны в зону проводимости без изменения энергии. Этот процесс называется туннелированием. Вероятность туннелирования электронов из зоны проводимости в валентную зону такая же, но так как в зоне проводимости электронов меньше, то преобладает процесс туннельного перехода электронов из валентной зоны в зону проводимости.

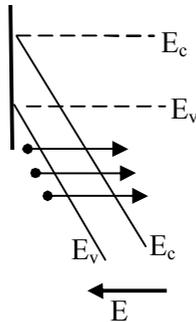


Рисунок 2.16. Туннелирование электронов из валентной зоны в зону проводимости при помещении полупроводника в сильное электрическое поле

Явление туннелирования имеет место при больших напряженностях электрического поля (10^5 - 10^6) В/см.

Насыщение скорости. В слабых электрических полях дрейфовая скорость носителей меньше, чем тепловые скорости. В сильных полях носители набирают дополнительную энергию (разогреваются) и дрейфовые скорости становятся соизмеримыми с тепловыми. При этом наблюдается изменение подвижности носителей.

В области высоких температур, когда процессы рассеяния определяются рассеянием на донорах, разогрев носителей приводит к увеличению числа столкновений носителей с атомами, и дрейфовая скорость перестает расти при увеличении напряженности электрического поля (рис.2.17 а), т.е. происходит насыщение скорости, что сопровождается уменьшением подвижности (рис.2.17 б). Это наблюдается в полях порядка 10^4 В/см.

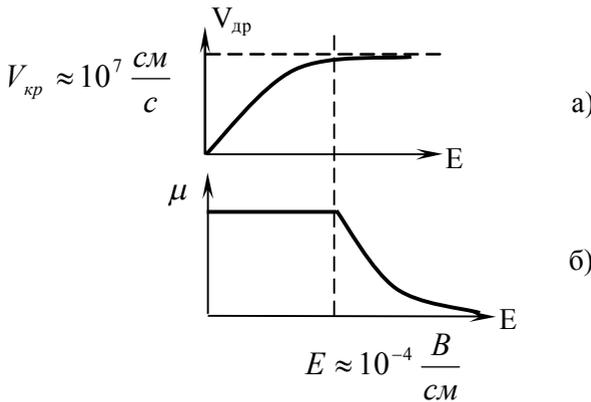


Рис. 2.17. Зависимость дрейфовой скорости (а) и подвижности (б) от напряженности электрического поля.

В области низких температур, когда главную роль играют процессы рассеяния на ионизованных примесях, подвижность возрастает при увеличении напряженности. Но происходит это только в области очень низких температур. При рассмотрении работы полупроводниковых приборов главную роль играют процессы, соответствующие области высоких температур, поскольку приборы работают в этом диапазоне температур.

2.8. Генерация и рекомбинация носителей в полупроводниках

Явления генерации и рекомбинации носителей играют важную роль в работе полупроводниковых приборов. Кроме того, эти явления лежат в основе работы целого класса приборов - приборов оптоэлектроники.

Генерация - образование свободных электронов и дырок. В зависимости от того, что является причиной появления свободных носителей, генерация может быть следующей: **термогенерация** - когда свободные носители возникают за счет теплового возбуждения; **фотогенерация** - свободные носители получают энергию от квантов света; **полевая** - возбуждение происходит за счет сил электрического поля. Поступление электрона в зону проводимости при генерации может происходить непосредственно из валентной зоны (межзонная генерация), либо через незаполненные или заполненные уровни ловушек (рис.2.18).

Появление свободного электрона в зоне проводимости сопровождается образованием свободной дырки.

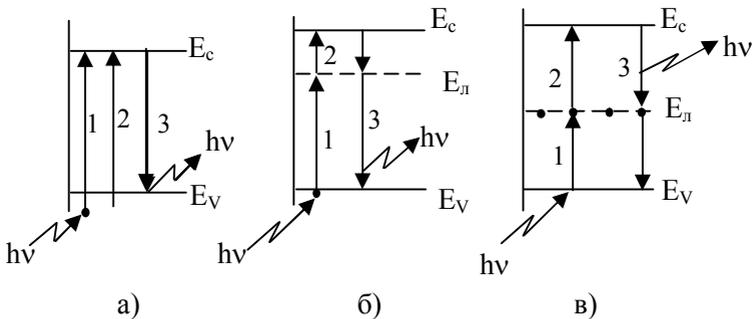


Рис. 2.18. Генерация (1) и рекомбинация (3) носителей из зоны в зону (а), через незаполненные уровни ловушек E_t (б) (2) или заполненные уровни (в).

. При генерации появляются избыточные носители заряда по сравнению с равновесными.

Рекомбинация - исчезновение свободных носителей заряда. По механизму процесса различают следующие виды рекомбинации. **Межзонная** рекомбинация - происходит при переходе свободного электрона из зоны проводимости в валентную зону

(процесс 3, рис.2.18,а). При этом исчезает как свободный электрон, так и свободная дырка. Однако этот процесс маловероятен, т.к. свободный электрон и свободная дырка должны оказаться одновременно в одно и то же время и в одном и том же месте в кристалле. Вероятность этого процесса $\sim 10^{-4}$.

Более вероятным является процесс рекомбинации через рекомбинационные уровни ловушек (процесс 3, рис. 2.18 б, в). При этом электрон может занять уровень ловушки и находиться там, пока не появится свободная дырка. Роль рекомбинационных ловушек могут выполнять примесные атомы или ионы и другие дефекты кристаллической решетки. Из этого следует, что время жизни носителей зависит от концентрации центров рекомбинации. Наиболее быстро процессы рекомбинации протекают на поверхности полупроводника, где из-за нарушения кристаллической решетки имеется большое количество центров рекомбинации.

В зависимости от того, как преобразуется энергия, выделяемая при рекомбинации, различают два вида рекомбинации:

излучательная - излучается квант видимого излучения (фотон);

безызлучательная - возбуждается квант тепловой энергии (фонон).

В состоянии термодинамического равновесия процессы генерации и рекомбинации взаимно уравновешены.

2.9. Уравнения электронейтральности

Достаточно часто у студентов формируется неверное представление о том, что полупроводник n – типа заряжен отрицательно, а p – типа - положительно. При этом логика рассуждений такова. Поскольку в полупроводнике n – типа основными носителями являются электроны, их в кристалле много, они заряжены отрицательно. Следовательно, кристалл полупроводника n – типа заряжен отрицательно, а p – типа - положительно. Данные рассуждения неверны и противоречат закону электронейтральности. Согласно этому закону в любом сечении полупроводника, как и в кристалле в целом, суммарный заряд равен нулю. Поскольку в полупроводнике имеются электроны и дыр-

ки, ионизованные доноры и акцепторы, то уравнение электронейтральности имеет следующий вид

$$p + N_d^+ = n + N_a^- . \quad (2.41)$$

Физически, если вернуться к вопросу об образовании полупроводников n – и p – типов, это понятно: образование свободного отрицательно заряженного электрона сопровождается появлением положительно заряженного ионизованного донора, т.е. суммарный заряд равен нулю.

2.10. Явления на поверхности полупроводников

На поверхности полупроводников разыгрывается значительное количество сложных физических явлений и процессов, влияющих или определяющих работу полупроводниковых приборов. В физике полупроводников физика поверхности занимает важное место. В данном разделе мы остановимся только на трех физических явлениях, используемых для создания приборов. Это явления **обеднения** и **обогащения поверхности основными носителями** и **инверсии поверхностной проводимости**. Рассмотрение этих явлений проведем на примере полупроводника с n – типом проводимости.

Представим себе, что вблизи поверхности полупроводника находится пластина, на которую подан отрицательный потенциал или на поверхности полупроводника произошла адсорбция отрицательных частиц. При этом электроны начнут отталкиваться от поверхности полупроводника и уходить вглубь. Произойдет явление **обеднения** поверхности полупроводника основными носителями. На зонной энергетической диаграмме явление обеднения отражается изгибом зон вверх на величину φ_s (рис.2.19 а). Это происходит в силу того, что уменьшение концентрации на поверхности требует в силу соотношения $n = N_c \exp\left(-\frac{E_c - E_F}{kT}\right)$, чтобы зазор между уровнями E_c и E_F на поверхности при постоянной температуре увеличился. Концентрация электронов на поверхности n_s будет определяться как

$$n_s = n_0 \exp\left(-\frac{\varphi_s}{kT}\right), \quad (2.42)$$

где n_0 – концентрация электронов в объеме.

Поскольку в обедненном слое концентрация носителей мала, то он обладает повышенным сопротивлением. Причем сопротивление этого слоя зависит от изгиба зон φ_s . Данное явление используется для создания выпрямляющих контактов.

Если на поверхности полупроводника или вблизи нее образуется положительный заряд, то электроны притягиваются к поверхности и, в соответствии с выражением (2.5), поверхность обогатится основными носителями. На зонной диаграмме это выразится изгибом зон вниз. Концентрация носителей на поверхности будет определяться как

$$n_s = n_0 \exp\left(\frac{\varphi_s}{kT}\right).$$

Поскольку вблизи поверхности концентрация электронов велика, то обогащенный слой имеет малое сопротивление, и это используется для создания омических контактов.

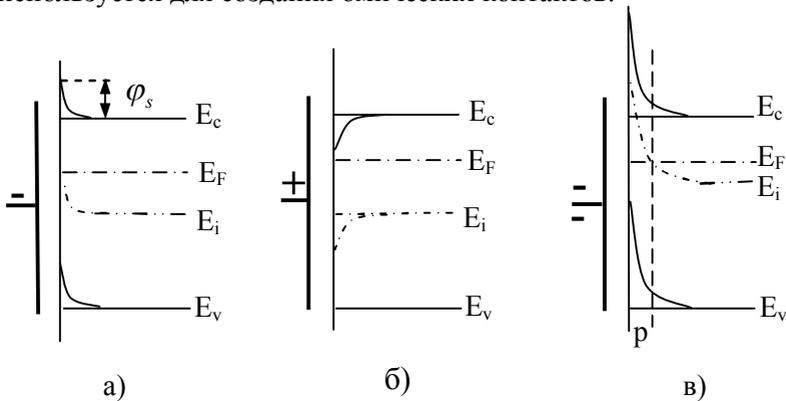


Рис. 2.19. Энергетические диаграммы полупроводника, поясняющие явления обеднения (а), обогащения (б) и инверсии (в).

При наличии у поверхности большего, чем в случае обеднения, отрицательного заряда, изгиб зон будет сильнее. При опре-

деленной величине заряда на поверхности произойдет явление ***инверсии поверхностной проводимости***, т.е. на поверхности полупроводника n – типа образуется слой с p – типом проводимости (рис.2.19 в). Это явление используется для создания широкого класса полевых приборов.

При рассмотрении этого вопроса надо иметь в виду, что уровень середины запрещенной зоны должен всегда находиться посередине запрещенной зоны в любом сечении полупроводника, а уровень Ферми E_F не должен не сдвигаться по запрещенной зоне, не изгибаться, так как его положение определяется концентрацией легирующей примеси. При изменении состояния на поверхности концентрация примеси, а, следовательно, и уровень E_F своего положения не изменяет.

3. КОНТАКТЫ МЕТАЛЛ-ПОЛУПРОВОДНИК. ВЫПРЯМЛЯЮЩИЕ И ОМИЧЕСКИЕ ПЕРЕХОДЫ НА КОНТАКТЕ МЕТАЛЛА С ПОЛУПРОВОДНИКОМ

3.1. Введение

Впервые в 1874г. была обнаружена зависимость сопротивления контакта металл-полупроводник от полярности приложенного напряжения. Затем в 1904г. были созданы кристаллические детекторы на основе прижимного контакта металлической проволоочки с природными полупроводниковыми кристаллами.

Детальное исследование свойств диэлектрических выпрямляющих контактов металл-полупроводник было проведено русским ученым Лосевым О.В. в 1919-1922гг. С помощью созданных им кристаллических детекторов с необычной ВАХ удалось получить детектирование, усиление и генерацию высокочастотных электромагнитных колебаний.

В 1938г. В. Шоттки предложил теорию явлений, сопровождающих выпрямление на контакте металла с полупроводником. В дальнейшем выпрямляющие контакты такого типа стали называться контактами с барьером Шоттки.

Особенно интенсивно работы по созданию приборов на основе контактов с барьером Шоттки проводились с 1950г. В результате этого разработано большое число приборов с разнообразными функциональными возможностями: для выпрямления переменного тока, ряд СВЧ-диодов, стабилитроны, импульсные диоды, фотодетекторы и солнечные батареи. Кроме того, контакты металл-полупроводник используются как затворы в полевых транзисторах и т.д.

Обязательной деталью любого полупроводникового прибора является омический контакт, который также представляет собой структуру металл-полупроводник. В данном разделе описаны свойства выпрямляющих и омических переходов на основе системы металл-полупроводник.

3.2. Энергетическая диаграмма выпрямляющего контакта металл-полупроводник

Рассмотрим механизм образования контакта металл - полупроводник n – типа проводимости при условии, что термодина-

мическая работа выхода электронов в вакуум из металла φ_M больше, чем термодинамическая работа выхода электронов из полупроводника φ_{II} . На рис.3.1 представлены зонные диаграммы металла и полупроводника в изолированном состоянии. Металл характеризуется термодинамической работой φ_M , а полупроводник термодинамической работой выхода φ_{II} , отсчитанной от уровня Ферми в полупроводнике и электронным средством χ , отсчитанным от уровня E_c до уровня вакуума.

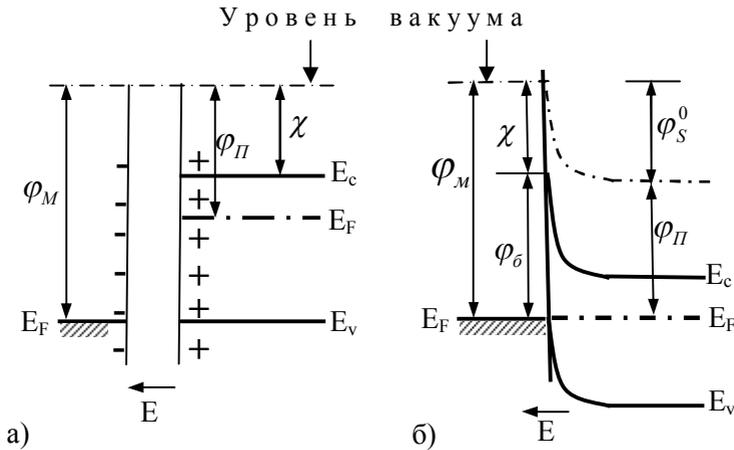


Рисунок 3.1. Зонная диаграмма металла и полупроводника n-типа в изолированном состоянии (а) и при плотном контакте (б)

При плотном соединении твердых тел, когда зазор между металлом и полупроводником менее 10^{-7} см, что имеет место, например, при напылении металла на поверхность полупроводника, между твердыми телами становится возможен обмен носителями зарядов.

В силу того, что $\varphi_M > \varphi_{II}$ электроны начинают переходить из полупроводника в металл. При этом приповерхностный слой полупроводника обедняется электронами.

На энергетической диаграмме (рис.3.1 б) это отражается появлением изгиба зон вверх. Электроны, перешедшие в металл,

создают на его поверхности отрицательный заряд, а в приповерхностном слое полупроводника нескомпенсированные ионизованные доноры формируют положительный заряд. В результате этого процесса между металлом и полупроводником возникает внутреннее электрическое поле E . Причем направление поля таково, что оно является тормозящим для электронов, переходящих из полупроводника в металл. Движение электронов из полупроводника в металл будет продолжаться до тех пор, пока напряженность внутреннего электрического поля E не достигнет величины, достаточной для того, чтобы прекратить поток электронов из полупроводника в металл, обусловленный разницей работ выхода. В этом случае система металл-полупроводник приходит к состоянию термодинамического равновесия, которое характеризуется единством уровня Ферми в металле и полупроводнике (рис.3.1. б). Для того, чтобы зонная энергетическая диаграмма (рис.3.1. б) была изображена правильно, необходимо иметь ввиду следующее.

Первое. При контактировании металла с полупроводником объемные свойства материалов не изменяются. Изменения происходят только в области поверхности полупроводника, где происходит обеднение основными носителями. Причем, для того, чтобы правильно отразить степень обеднения поверхности, необходимо от уровня вакуума на границе контакта отложить значение энергии электронного сродства χ и в эту точку провести уровень E_c полупроводника.

Второе. Значение термодинамической работы выхода полупроводника в объеме остается также неизменным. Тогда, откладывая значение φ_{II} в полупроводнике от уровня Ферми, получим уровень вакуума для полупроводника. Из-за наличия внутреннего электрического поля на контакте появляется барьер φ_s^o , называемый поверхностным потенциалом.

Из рис.3.1. б видно, что на контакте металл-полупроводник образуются два барьера. Барьер $\varphi_\delta = \varphi_M - \chi$ называют барьером Шоттки и барьер $\varphi_s^o = \varphi_M - \varphi_{II}$. Эти барьеры определяют величины потоков электронов из металла в полупроводник

(φ_{σ}) и из полупроводника в металл (φ_s^o). Причем барьер φ_{σ} в идеальном случае остается постоянным, тогда как барьер φ_s^o может изменяться при приложении напряжения к структуре. Область на контакте, где в полупроводнике понижена концентрация электронов, называется областью пространственного заряда (ОПЗ).

3.3. Принцип выпрямления на контакте металл-полупроводник по энергетическим диаграммам

Принцип действия многих полупроводниковых приборов, имеющих различное функциональное назначение, основан на эффекте односторонней проводимости выпрямляющих переходов. Односторонняя проводимость перехода заключается в том, что при приложении к выпрямляющему переходу напряжения одной полярности переход обладает высокой проводимостью, а при смене полярности сопротивление перехода резко возрастает.

В тридцатые годы XX века немецким ученым Шоттки была разработана теория выпрямляющего контакта металл-полупроводник. Рассмотрим принцип односторонней проводимости на таком контакте по энергетическим диаграммам. Для этого необходимо рассмотреть потоки носителей зарядов на контакте металл-полупроводник при отсутствии напряжения, а затем при приложении прямого и обратного напряжений (рис.3.2). Под прямым напряжением будем понимать такое напряжение, когда отрицательный потенциал приложен к полупроводнику n – типа.

Для понимания сути вопроса необходимо иметь в виду, что в невырожденных полупроводниках носители подчиняются статистике Максвелла-Больцмана. В соответствии с этим в полупроводнике и металле есть носители, имеющие высокие энергии, но их мало, и есть носители с малой энергией.

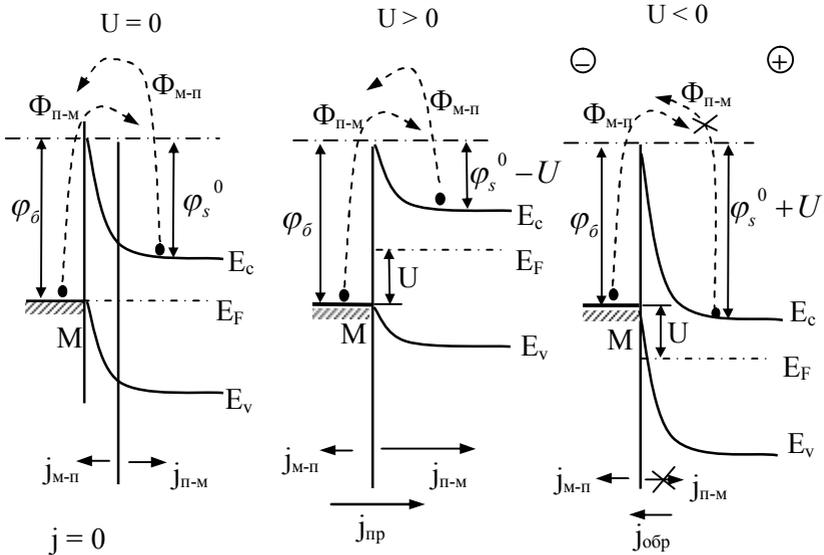


Рисунок 3.2. Зонные энергетические диаграммы контакта металл – полупроводник в состоянии термодинамического равновесия (а), при приложении прямого (б) и обратного (в) напряжения

На рис.3.3 представлена кривая распределения носителей по энергиям.

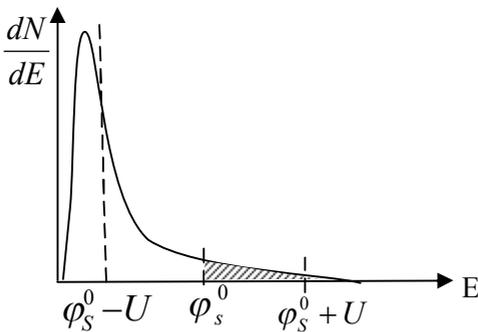


Рисунок 3.3. Распределение электронов по энергиям в соответствии со статистикой Максвелла-Больцмана

В состоянии термодинамического равновесия ($U = 0$) наблюдается поток носителей $\Phi_{M-\Pi}$ из металла в полупроводник через барьер φ_{δ} , т.е. происходит надбарьерная эмиссия Шоттки. Поток носителей при этом определяется температурой и величиной барьера φ_{δ} .

Процессы туннелирования электронов сквозь барьер учитывать не будем. Поскольку в полупроводнике есть носители, энергия которых больше φ_s^o , то они переходят из полупроводника в металл, создавая поток $\Phi_{\Pi-M}$.

Величина этого потока зависит от барьера φ_s^o и температуры. В состоянии равновесия ($U = 0$) потоки $\Phi_{M-\Pi}$ и $\Phi_{\Pi-M}$ равны и суммарный ток через структуру М-П равен нулю. При приложении прямого смещения ($U > 0$) энергия электронов в полупроводнике возрастает, т.е. уровень Ферми в полупроводнике поднимается относительно E_F в металле на величину приложенного напряжения. При этом величина барьера для электронов, переходящих из полупроводника в металл, уменьшится на величину приложенного напряжения и станет равной $\varphi_s^o - U$. Уменьшение барьера приведет к экспоненциальному росту потока $\Phi_{\Pi-M}$. Поскольку величина барьера φ_{δ} при этом остается неизменной, то поток $\Phi_{M-\Pi}$ не изменится. В результате при приложении прямого смещения (минус на полупроводнике n -типа) ток через структуру М-П будет определяться в основном потоком электронов $\Phi_{\Pi-M}$.

При приложении обратного смещения ($U < 0$), когда минус от источника питания подключен к металлу, уровень Ферми в полупроводнике опустится по отношению к E_F в металле на величину приложенного напряжения U . Величина потенциального барьера для потока электронов $\Phi_{\Pi-M}$ станет равной $\varphi_s^o + U$, и поток электронов фактически уменьшится до нуля.

Величина тока через контакт М-П при этом определяется потоком электронов $\Phi_{П-М}$.

Таким образом, прямой ток через контакт М-П определяется потоком электронов из полупроводника в металл и зависит от температуры и напряжения. Обратный ток через контакт определяется потоком электронов из металла в полупроводник, зависящим от величины барьера φ_0 и температуры. Если барьер φ_0 достаточно большой, то обратный ток мал по величине и не зависит в идеальном случае от обратного напряжения, т.к. величина барьера φ_0 постоянна.

В результате рассмотрения на качественном уровне вопроса о выпрямляющих свойствах перехода М-П показано, что такие структуры обладают эффектом односторонней проводимости и имеют диодную вольтамперную характеристику (рис.3.4).

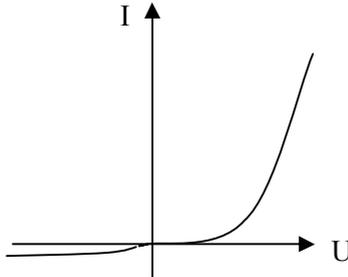


Рисунок 3.4. Вид вольтамперной характеристики идеального контакта М-П

3.4. Вольтамперная характеристика выпрямляющего контакта металл-полупроводник

Вольтамперная характеристика (ВАХ) это аналитическая или графическая зависимость тока от напряжения. Для выпрямляющего контакта М-П качественный вид ВАХ описан в предыдущем параграфе. В данном разделе выведем аналитическое выражение для ВАХ. Для аналитического описания ВАХ обычно используются две теории: диодная и диффузионная. Их различие состоит в соотношении между шириной ОПЗ L и длиной свободного пробега электронов l . При $l \gg L$ реализуется ди-

одная теория, когда рассеянием носителей в ОПЗ пренебрегают. Для случая $l \ll L$ имеет место диффузионная теория выпрямления. В данном случае рассмотрен вывод ВАХ исходя из диодной теории. При этом падением напряжения на сопротивлении объема полупроводника пренебрегают.

Для нахождения выражения тока, протекающего через барьер М-П, необходимо исходить из того, что он определяется надбарьерными потоками носителей Φ_{M-P} и Φ_{P-M} , т.е.

$$j = j_{P-M} - j_{M-P} \quad (3.1)$$

Токи являются термоэлектронными по своей природе. При этом величины токов равны

$$\left. \begin{aligned} j_{M-P} &= A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right) \\ j_{P-M} &= A^* T^2 \exp\left(\frac{-\varphi_{\delta}}{kT}\right) \exp\left(\frac{qU}{kT}\right) \end{aligned} \right\} \quad (3.2)$$

где A^* – постоянная Ричардсона для полупроводников.

Она составляет $(1,11 \div 1,19)A$ для германия и $(2,1 \div 2,2)A$ для кремния, где A – постоянная Ричардсона для термоэлектронной эмиссии в вакуум, равная $120 \frac{A}{K^2 c m^2}$.

Тогда в соответствии с (3.1)

$$j = A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right) \left(\exp \frac{qU}{kT} - 1 \right). \quad (3.3)$$

Поскольку величина $A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right)$ определяет ток насыщения контакта (обратный ток)

$$j_s = A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right), \quad (3.4)$$

то
$$j = j_s \left(\exp \frac{qU}{kT} - 1 \right) \quad (3.5)$$

Проанализируем ВАХ контакта М-П для четырех случаев.

1) При $U > 0$ и $|U| \ll \frac{kT}{q}$, разлагая экспоненту в ряд и ограничиваясь вторым членом разложения, получим

$$j \approx j_s \frac{qU}{kT}, \quad (3.6)$$

т.е. в области малых прямых смещений ток является линейной функцией напряжения.

2) При $U > 0$ и $|U| \gg \frac{kT}{q}$, если $U > 2,3 \frac{kT}{q}$, то $\exp \frac{qU}{kT} \gg 1$ и ВАХ примет вид $j = j_s \exp \frac{qU}{kT}$ (3.7).

Т.е. ток зависит от напряжения по экспоненте.

3) Для $U < 0$ и $|U| \ll \frac{kT}{q}$ вновь получаем линейную зависимость тока от напряжения, как и в (3.6).

4) Для $U < 0$ и $|U| \gg \frac{kT}{q}$, если $U \leq -2,3 \frac{kT}{q}$, то $\exp \frac{qU}{kT} \ll 1$ и $j = -j_s$, (3.8)

т.е. в идеальном контакте М-П при обратном напряжении протекает ток насыщения j_s , противоположный току в прямом направлении. Внешний вид ВАХ идеального контакта М-П представлен на рис.3.4.

3.5. Расчет напряженности поля и потенциала на контакте металл-полупроводник. Ширина области пространственного заряда

Для нахождения напряженности электрического поля E и потенциала φ в области пространственного заряда (ОПЗ) контакта М-П необходимо решить уравнение Пуассона. Решение уравнения проводится для одномерного случая, полагая, что площадь контакта велика и рассматривается контакт металла с полупроводником n – типа проводимости (рис.3.5). Причем из-

гиб зон $\varphi_s^0 > 2,3 kT$, что дает право пренебрегать наличием свободных электронов в ОПЗ.

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\varepsilon\varepsilon_0}, \quad (3.9)$$

где $\rho(x)$ – плотность заряда в ОПЗ;

ε – диэлектрическая проницаемость полупроводника;

ε_0 – электрическая постоянная.

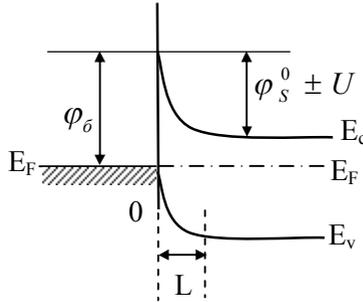


Рис. 3.5. Энергетическая диаграмма контакта М-П для $U = 0$

Плотность заряда $\rho(x)$ в ОПЗ для однородного легированного полупроводника n – типа с учетом принятых допущений равна $\rho(x) = qN_d$ и уравнение (3.9) принимает следующий вид

$$\frac{d^2\varphi}{dx^2} = -\frac{qN_d}{\varepsilon\varepsilon_0}. \quad (3.10)$$

Граничные условия для решения уравнения (3.10) следующие:

$$\begin{aligned} \varphi(x) &= \varphi_s^0 \pm U && \text{при } x = 0; \\ \varphi(x) &= 0 \text{ и } \frac{d\varphi}{dx} = 0 && \text{при } x = L. \end{aligned} \quad (3.11)$$

За начало отсчета энергии принят уровень E_c . Значение потенциала φ измеряется в вольтах. Интегрируя уравнение (3.10), получим

$$\int_0^{d\varphi/dx} d\left(\frac{d\varphi}{dx}\right) = \int_x^L -\frac{qN_\delta}{\varepsilon\varepsilon_0} dx; \quad \frac{d\varphi}{dx} = -\frac{qN_\delta}{\varepsilon\varepsilon_0}(L-x). \quad (3.12)$$

С учетом того, что $E = -\frac{d\varphi}{dx}$, это означает, что

$$E(x) = \frac{qN_\delta}{\varepsilon\varepsilon_0}(L-x). \quad (3.13)$$

Таким образом, напряженность электрического поля в ОПЗ контакта М-П изменяется по линейному закону: она максимальна при $x = 0$ и равна нулю при $x = L$ (рис.3.6).

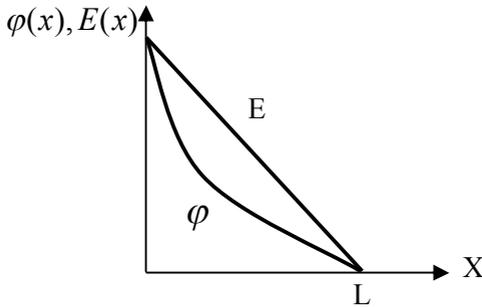


Рисунок 3.6. Ход напряженности электрического поля E и потенциала φ в ОПЗ контакта М-П

Проводя интегрирование уравнения (3.12), получим выражение для $\varphi(x)$:

$$\varphi(x) = \frac{qN_\delta}{\varepsilon\varepsilon_0}(L-x)^2. \quad (3.14)$$

То есть потенциал $\varphi(x)$ является квадратичной функцией координаты x . Подставляя граничное условие (3.11), получим

$$\text{выражение для ширины ОПЗ: } L = \sqrt{\frac{2\varepsilon\varepsilon_0}{qN_\delta}(\varphi_s^o \pm U)}. \quad (3.15)$$

Видно, что L зависит от концентрации легирующей примеси в полупроводнике N_δ и напряжения U . При приложении прямого смещения ширина ОПЗ уменьшается, при обратном - увеличивается.

3.6. Диод Шоттки: структура, эквивалентная схема, параметры эквивалентной схемы. Модель диода Шоттки

Детальное исследование теории выпрямления на контакте М-П проведено немецким ученым Шоттки. Диоды на основе контакта М-П называются диодами Шоттки. Структуры диодов Шоттки представлены на рис.3.7.

По определению диод Шоттки - это полупроводниковый прибор на основе контакта М-П, принцип действия которого основан на явлении термоэлектронной эмиссии.

Для представленных на рис.3.7 структур имеются два контакта М-П. Выпрямляющим контактом является контакт $M1$ - полупроводник n - типа. Контакт $M2$ - полупроводник n^+ - типа - омические контакты, свойства которых описаны ниже. Система, состоящая из $M2 - n^+ - n$ - слоев, называется базой диода. Слой n^+ - типа вводится для уменьшения сопротивления базы.

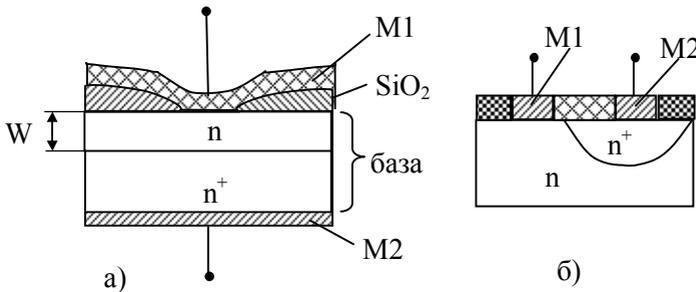


Рисунок 3.7. Структура дискретного (а) и интегрального диодов Шоттки (б)

Эквивалентная схема - это совокупность электрически связанных простейших элементов электрической цепи (резистор, конденсатор, индуктивность, генераторы тока и напряжения), причем каждый из элементов отражает тот или иной процесс, происходящий в приборе.

Для диода Шоттки эквивалентная схема представлена на рис.3.8. Для любого прибора должны быть записаны выражения, позволяющие по измеряемым величинам рассчитывать значения параметров эквивалентной схемы.

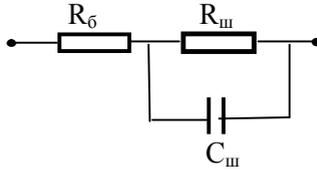


Рис. 3.8. Эквивалентная схема диода Шоттки

$R_б$ – омическое сопротивление базы. Данный элемент отражает падение напряжения на базе при протекании тока через прибор. Численное значение $R_б$ при простейшей геометрии диода определяется по следующей формуле

$$R_б = \rho_б \frac{W}{S}, \quad (3.16)$$

где $\rho_б$ – удельное сопротивление полупроводника n – типа;
 W – толщина слоя n – типа;
 S – площадь контакта.

Предполагается, что сопротивление системы $M2-n^+$ мало и на величину $R_б$ мало влияет.

В некоторых учебниках сопротивление $R_б$ обозначают через R_s и называют сопротивлением растекания. Для точечного контакта сопротивление растекания можно определять по формуле

$$R_б = \rho_б \frac{1}{\pi a}, \quad (3.17)$$

где a – радиус точечного контакта.

Эффект односторонней проводимости диода Шоттки отражен на эквивалентной схеме дифференциальным сопротивлением $R_{ш}$. По определению $R_{ш} = \frac{dU}{dI}$. Продифференцируем выражение для тока через диод по напряжению

$$\frac{dI}{dU} = \frac{qI_s}{kT} \exp \frac{qU}{kT}. \quad (3.18)$$

Тогда, с учетом того, что $I = I_s \left(\exp \frac{qU}{kT} - 1 \right)$, данное выражение можно представить как $I + I_s = I_s \exp \frac{qU}{kT}$ и, взяв обратную величину от (3.18), получим

$$R_{ш} = \frac{kT}{q(I + I_s)}. \quad (3.19)$$

Видно, что при прямых включениях диода, когда через него течет значительный ток, $R_{ш}$ мало, а при обратных - $R_{ш}$ велико.

Диод Шоттки обладает емкостными свойствами. На эквивалентной схеме емкостные свойства отражены емкостью $C_{ш}$.

$C_{ш}$ - это емкость плоского конденсатора, одной из обкладок которой является металл, а второй (воображаемой) обкладкой является изменяющаяся граница ОПЗ (рис.3.9). Функцию диэлектрика в таком плоском конденсаторе выполняет ОПЗ. Тогда

$$C_{ш} = \frac{\varepsilon \varepsilon_0 S}{L} \text{ или с учетом (3.15)}$$

$$C_{ш} = S \sqrt{\frac{\varepsilon \varepsilon_0 q N_d}{2(\varphi_s^o \pm U)}}. \quad (3.20)$$

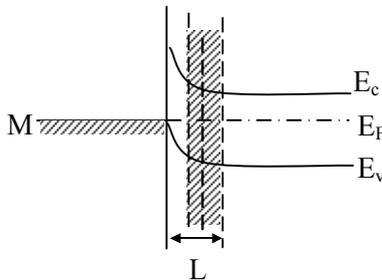


Рисунок 3.9. Структура контакта М-П, поясняющая смысл $C_{ш}$

Видно, что $C_{ш}$ зависит от концентрации легирующей примеси в полупроводнике и приложенного напряжения. При прямых смещениях ширина ОПЗ уменьшается, что вызывает увеличение

$C_{ш}$, а при обратных напряжениях L увеличивается и $C_{ш}$ уменьшается. Емкость $C_{ш}$ отражает наличие токов смещения на контакте М-П.

Полная величина переменного тока через диод \tilde{I}_d равна сумме переменного тока, связанного с движением зарядов через ОПЗ \tilde{I} и тока смещения $\tilde{I}_{см}$. Таким образом, для протекания переменного тока через диод Шоттки существуют два параллельных канала: через $R_{ш}$ и $C_{ш}$. Причем и сопротивление $R_{ш}$ и емкость $C_{ш}$ зависят от величины и полярности подаваемого на диод напряжения. При обратном смещении диода $R_{ш}$ резко возрастает, и основная часть переменного тока протекает через $C_{ш}$, а эквивалентную схему для переменного сигнала можно представить как последовательное соединение сопротивления R_0 и емкости $C_{ш}$.

Модель диода Шоттки. Для автоматизированного моделирования и расчета электронных схем с использованием диода Шоттки используется модель диода в системе "Самрис". Любая модель предполагает совокупность схемы замещения прибора и набор математических выражений. Обычно в качестве схемы замещения используются физические эквивалентные схемы. Такие модели относятся к числу электрических функциональных моделей.

$$\left. \begin{aligned} I_{ш} &= C_{ш} \frac{dU}{dt} + \frac{U_{ш}}{R_{ш}} + I'_{ш} \\ I'_{ш} &= I_s \left(\exp \frac{dU_m}{mkT} - 1 \right) \end{aligned} \right\} \quad (3.21)$$

$$\left. \begin{aligned} C_{ш} &= C_{ш}^o \left(1 + \frac{U_{ш}}{\varphi_s^o} \right)^{-\frac{1}{2}} \\ C_{ш}^o &= S \sqrt{\frac{\varepsilon \varepsilon_0 q N_{\delta}}{2 \varphi_s^o}} \\ R_{\delta} &= \rho_{\delta} \frac{W}{S} \\ U_B &= I_{ш} R_{\delta} \end{aligned} \right\} \quad (3.21)$$

где m – коэффициент неидеальности диода Шоттки. Определяется экспериментально.

Таким образом, схема замещения (рис.3.10) и система уравнений (3.21) представляют собой модель диода Шоттки.

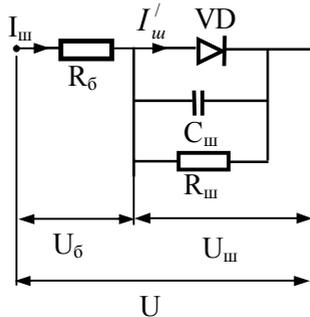


Рис. 3.10. Схема замещения для модели диода Шоттки

3.7. Эффект Шоттки

Эффект Шоттки - это понижение высоты потенциального барьера Шоттки φ_{δ} на контакте М-П, возникающее из-за наличия сильного электрического поля и сил зеркального изображения.

Сущность эффекта проще всего рассмотреть на примере контакта металл-вакуум, а затем применить полученные выводы для контакта М-П, заменив в выражениях диэлектрическую проницаемость вакуума $\varepsilon = 1$ на диэлектрическую проницаемость полупроводника. Изначально на контакте металл-вакуум

существует прямоугольный потенциальный барьер $\varphi_{\bar{o}}$. Из-за наличия электрического поля у поверхности металла прямоугольный барьер OAB превращается в треугольный OAC . Учет сил зеркального изображения (при выходе электрона из металла в вакуум в металле на таком же расстоянии от поверхности образуется его зеркальное изображение, т.е. положительный заряд), определяемый законом Кулона, преобразует прямоугольный барьер OAB к виду ODB (рис.3.11).

Совместное действие сил электрического поля и учет сил зеркального изображения приводит к тому, что прямоугольный барьер OAB превращается в колоколообразный OFC . В данном случае для выхода электрона из металла в вакуум необходимо преодолеть барьер не $\varphi_{\bar{o}}$, а $\varphi_{\bar{o}} - \Delta\varphi_{\bar{o}}$.

Оценим величину $\Delta\varphi_{\bar{o}}$.

$$\varphi_{\bar{o}} = [\varphi_{zu}(x) + \varphi_E(x)] = -\frac{q^2}{16\pi\epsilon\epsilon_0 x} - qEx. \quad (3.22)$$

$$\Delta\varphi_{\bar{o}} = 2qEx_m$$

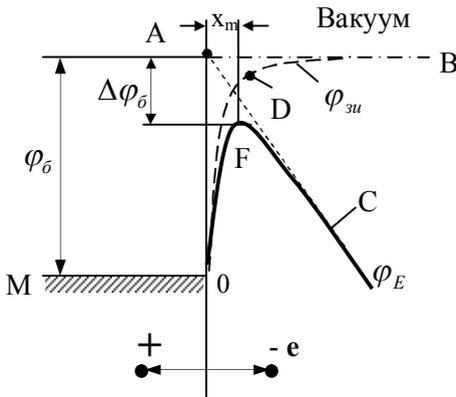


Рис. 3.11. Контакт металла с вакуумом

$\varphi_{\bar{o}}$ — высота прямоугольного барьера; $\Delta\varphi_{\bar{o}}$ — понижение высоты барьера φ ;

φ_{zu} — потенциальный барьер с учетом сил зеркального изображения;

φ_E — потенциальный барьер с учетом сил электрического поля

Необходимо найти x_m из условия, что при $x = x_m$ $\frac{d\varphi_{\delta}}{dx} = 0$.

Тогда, дифференцируя (3.22) и приравнявая $\frac{d\varphi_{\delta}}{dx} = 0$, получим

$$x_m = \sqrt{\frac{q}{16\pi\varepsilon\varepsilon_0 E}} \quad (3.23)$$

$$\text{и } \Delta\varphi_{\delta} = 2Ex_m = \sqrt{\frac{q^3 E}{4\pi\varepsilon\varepsilon_0}}. \quad (3.24)$$

Значение напряженности поля E не является постоянным и зависит от x . Однако, учитывая, что значение x_m невелико, можно взять значение E постоянным, равным максимальной напряженности на контакте М-П, т.е. при $x = 0$

$$E_{(x=0)} = \frac{qN_{\delta}}{\varepsilon\varepsilon_0} L = \frac{qN_{\delta}}{\varepsilon\varepsilon_0} \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_s^o \pm U)}{qN_{\delta}}} = \sqrt{\frac{2qN_{\delta}(\varphi_s^o \pm U)}{\varepsilon\varepsilon_0}}. \quad (3.25)$$

Выражение для $\Delta\varphi_{\delta}$ будет иметь вид

$$\Delta\varphi_{\delta} = \beta_u (\varphi_s^o \pm U)^{1/4}, \quad (3.26)$$

$$\text{где } \beta_u = \left[\frac{q^3 N_{\delta}}{8\pi^2 (\varepsilon\varepsilon_0)^3} \right]^{1/4}.$$

С учетом эффекта Шоттки ВАХ имеет вид

$$J = A^* T^2 \exp\left(-\frac{\varphi_{\delta} - \Delta\varphi_{\delta}}{kT}\right) \cdot \left(\exp\frac{qU}{kT} - 1\right) \quad (3.27)$$

При прямом смещении, когда $U \geq 2,3kT/q$

$$j = A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right) \cdot \exp\left(\frac{U + \Delta\varphi_{\delta}}{kT}\right). \quad (3.28)$$

С увеличением прямого напряжения $\Delta\varphi_{\delta} = \beta_u (\varphi_s^o - U)^{1/4}$ уменьшается и рост тока с напряжением происходит медленнее, чем дает теория. Это отклонение от теории можно учесть, записав ВАХ в следующем виде

$$j = A^* T^2 \exp\left(-\frac{\varphi_{\delta}}{kT}\right) \cdot \exp\left(\frac{qU}{nkT}\right),$$

(3.29)

где n – коэффициент неидеальности диода Шоттки. Его определяют по углу наклона ВАХ, построенной в координатах

$$\ln j = f(U) \text{ как } \frac{d(\ln j)}{dU} = \frac{q}{nkT} \text{ (рис.3.12).}$$

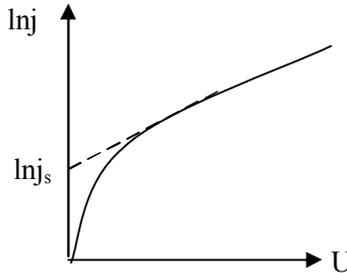


Рис. 3.12. Зависимость $\ln j$ от U для диодов Шоттки

Экстраполируя отрезок прямой на рис.3.12 до $U = 0$, получают значение $\ln j_s$

$$\ln j_s = \ln(A^* T^2) - \frac{\varphi_{\delta}}{kT}. \quad (3.30)$$

Отсюда можно определить высоту барьера φ_{δ}

$$\varphi_{\delta} = kT \ln\left(\frac{A^* T^2}{j_s}\right). \quad (3.31)$$

При обратном напряжении, когда $\Delta\varphi_{\delta} = \beta_{in}(\varphi_s^o + U)^{-1/4}$ и уменьшение барьера $\Delta\varphi_{\delta}$ существенно, это приводит к росту тока в обратном направлении. Таким образом, эффект Шоттки проявляется и при прямом, и при обратном напряжениях в изменении тока по сравнению с идеальной ВАХ (рис.3.13).

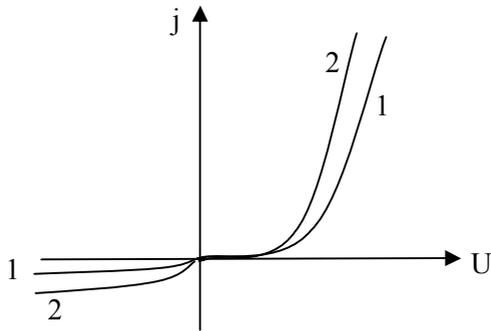


Рисунок 3.13. ВАХ идеального диода (1), и с учетом эффекта Шоттки (2)

3.8. Достоинства и недостатки диодов Шоттки

Диоды Шоттки используются для изготовления дискретных приборов СВЧ-диапазона, как импульсные диоды, а также в интегральных схемах. Мощные (силовые) диоды изготавливаются обычно на основе n – кремния и имеют рабочие токи до сотен ампер и очень высокое быстродействие по сравнению с диодами на основе p – n переходов.

Главное достоинство диодов Шоттки - высокое быстродействие, что связано с отсутствием эффекта накопления неосновных носителей в базе диода и не требуется времени для их расщепления, что определяет длительность переходных процессов, а также более низкие рабочие напряжения при прямом смещении, особенно для больших токов.

К недостаткам диодов Шоттки следует отнести низкие обратные допустимые напряжения и большие, чем в диодах на основе p – n переходов, обратные токи. Низкие пробивные напряжения связаны с эффектом усиления напряженности электрического поля на краях контакта М-П (рис.3.14). Этот эффект устраняется, если в слое SiO_2 окно имеет пологие края (как показано на рис.3.7 а) или применением охранного p – n перехода (рис.3.15). Более высокие обратные токи обусловлены малой высотой реальных барьеров Шоттки, а также наличием туннелирования сквозь барьер при больших обратных напряжениях.

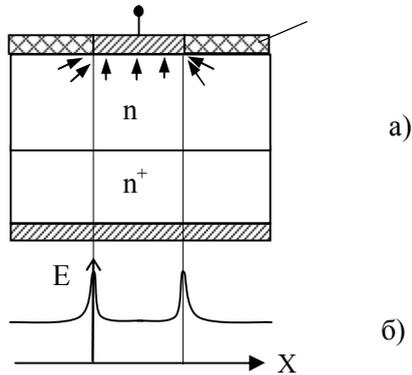


Рис. 3.14. Структура диода Шоттки (а) и распределение напряженности электрического поля на контакте (б)

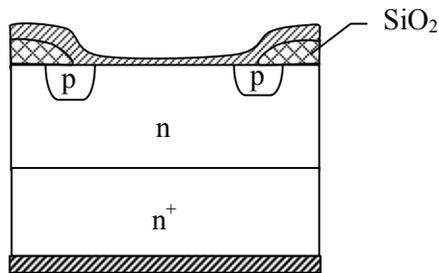


Рисунок 3.15. Конструкция диода Шоттки с охранным p-n переходом

3.9. Омические контакты и их параметры

Основное назначение омических переходов - электрическое соединение полупроводника с токопроводящими частями полупроводникового прибора. Омические переходы играют важную роль в работе полупроводниковых, особенно интегральных, приборов, поскольку их число больше, чем выпрямляющих. Теория омических переходов разработана недостаточно хорошо, и поэтому их формирование базируется на экспериментальных данных.

Основные требования, предъявляемые к омическим переходам, следующие:

- 1) отсутствие инжекции из контакта в полупроводник;

- 2) минимальное падение напряжения на переходе;
- 3) линейность и симметричность ВАХ;
- 4) высокая скорость рекомбинации.

Количественно омические переходы характеризуются следующими параметрами.

1. **Скорость рекомбинации**. Определяется как отношение потока носителей через переход φ_n к избыточной концентрации носителей у перехода $n_{zp} - n_0$ (n_{zp} – концентрация носителей у

перехода, n_0 – равновесная концентрация)
$$S = \frac{\varphi}{n_{zp} - n_0} .$$

2. **Сопротивление омического перехода**. Оценивается как отношение падения напряжения на переходе ΔU к току через переход I .

$$R = \frac{\Delta U}{I} .$$

Падение напряжения ΔU определяется как разница между падением потенциала в полупроводнике, измеренное методом зонда и падением напряжения между полупроводником и контактом. Чем меньше сопротивление R , тем лучше переход.

3. **Коэффициент выпрямления** K_B . Определяется как отношение прямого тока через переход I^+ к обратному I^- при равных значениях прикладываемого напряжения.

$$K_B = \frac{I^+}{I^-} | U = const .$$

Идеальный омический переход соответствует $K_B = 1$.

4. **Коэффициент нелинейности** K_H . Оценивается как отношение статического сопротивления R_{cm} к дифференциальному сопротивлению R_o при заданном токе I .

$$K_H = \frac{R_{cm}}{R_o} | I = const$$

Идеальный омический переход должен иметь прямолинейную и симметричную ВАХ (рис.3.16).

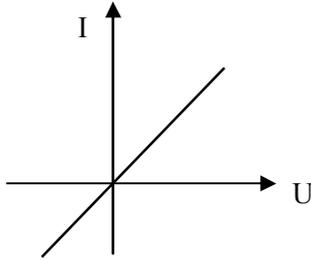


Рисунок 3.16. ВАХ идеального омического перехода

Реальные омические переходы обычно имеют более сложную структуру (рис.3.17).

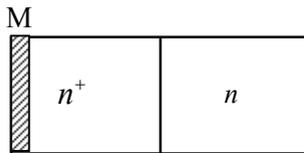


Рисунок 3.17. Структура реального омического перехода

4. ЭЛЕКТРОННО-ДЫРОЧНЫЕ ПЕРЕХОДЫ

4.1. Виды электрических переходов при контакте полупроводников

В твердотельной электронике используются следующие варианты контактов полупроводниковых слоев:

контакт полупроводников с одним типом проводимости, но с различным уровнем легирования ($n^+ - n, p^+ - p$),

контакт полупроводников с различным типом проводимости ($p - n, p^+ - n, n^+ - p$). При этом контактирующие слои могут иметь либо одинаковую ширину запрещенной зоны, либо изготавливаться на основе полупроводников с различной шириной запрещенной зоны. В зависимости от того, какие слои образуют электрический переход, их подразделяют на *гомопереходы и гетеропереходы*.

Гомопереход - это электрический переход на контакте полупроводников с одинаковой шириной запрещенной зоны.

Гетеропереход - это электрический переход на контакте полупроводников с различной шириной запрещенной зоны.

В свою очередь гомо- и гетеропереходы делятся на изотипные ($n^+ - n, p^+ - p$) и анизотипные ($p - n, p^+ - n, n^+ - p$).

Основой современных твердотельных приборов являются анизотипные гомопереходы. Их теория хорошо разработана. Поэтому вполне логично первоначально рассмотреть свойства таких переходов, а затем на базе этих знаний ознакомиться с теорией и применением гетеропереходов.

Таким образом, данный раздел учебного пособия посвящен в основном изучению анизотипных гомопереходов.

4.2. Механизм образования электронно-дырочного перехода. Определение электронно-дырочного перехода

Представим, что два полупроводника с различным типом проводимости приведены в тесное соприкосновение так, что между слоями р- и n-типа проводимости становится возможным обмен носителями зарядов (рис.4.1). В силу того, что в р-полупроводнике концентрация дырок p_{p0} много больше, чем

концентрация дырок в n-слое p_{n0} , возникает процесс диффузии дырок из полупроводника р-типа в слой n-полупроводника. По аналогичной причине ($n_{n0} \gg n_{p0}$) электроны из полупроводника n-типа проводимости диффундируют в р-полупроводник. В приконтактной области образуются нескомпенсированные слои с положительно заряженными донорами (в слое n-типа) и отрицательно заряженными акцепторами (в слое р-типа). Возникшее диффузионное движение основных носителей должно было продолжаться до исчезновения градиента концентрации. Однако возникшее в области контакта электрическое поле, которое называется диффузионным или более точно полем контактной разности потенциалов, направлено так, что и для электронов и для дырок оно является тормозящим. Когда величина внутреннего электрического поля E станет достаточно большой, диффузионные потоки электронов и дырок прекратятся и структура придет в состояние термодинамического равновесия.

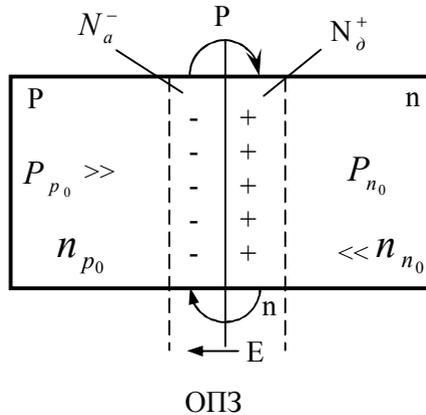


Рис. 4.1. К механизму образования электронно-дырочного перехода

Электронно-дырочный переход (ЭДП) - это переходная область на контакте полупроводников с различным типом проводимости, отличительной особенностью которой является наличие внутреннего электрического поля.

4.3. Контактная разность потенциалов φ_k^o . Зависимость φ_k^o от температуры, ширины запрещенной зоны, концентрации легирующей примеси

Возникающее в переходном слое на контакте полупроводников с различным типом проводимости электрическое поле является тормозящим для основных носителей. Из-за этого для основных носителей возникает потенциальный барьер, величина которого определяется напряженностью электрического поля в р-п переходе. Лучше всего потенциальный барьер для основных носителей зарядов показать, используя зонную энергетическую диаграмму электронно-дырочного перехода.

Поскольку в дальнейшем достаточно часто используются зонные диаграммы р-п переходов, то необходимо записать порядок изображения таких диаграмм. Он следующий.

Изображаются границы ОПЗ.

Задается положение уровня Ферми в полупроводниках. При этом надо помнить, что в состоянии термодинамического равновесия уровень Ферми в полупроводниках р- и п-типов находится на одной горизонтальной линии. При приложении напряжения происходит смещение уровней Ферми относительно друг друга на величину приложенного напряжения. Причем уровень Ферми в полупроводнике, к которому приложен отрицательный потенциал, лежит выше, чем в полупроводнике, находящемся под положительным смещением.

Изображается зонная диаграмма полупроводников р- и п-типов вне зоны ОПЗ.

Проводятся линии, соединяющие уровни E_c и E_v полупроводников р- и п-типов. Эти линии находятся в ОПЗ р-п перехода.

На рис.4.2 изображена зонная энергетическая диаграмма р-п перехода в состоянии термодинамического равновесия. Барьер φ_k^o для основных носителей заряда называется контактной разностью потенциалов. Величина этого барьера определяет потоки основных носителей зарядов в р-п переходе. Величина барьера φ_k^o определяется напряженностью внутреннего электрического поля E , поскольку именно это поле препятствует перемещению

основных носителей: электронов из n-области в p-область и дырок из p-области в n-область. Электрическое поле существует только в ОПЗ.

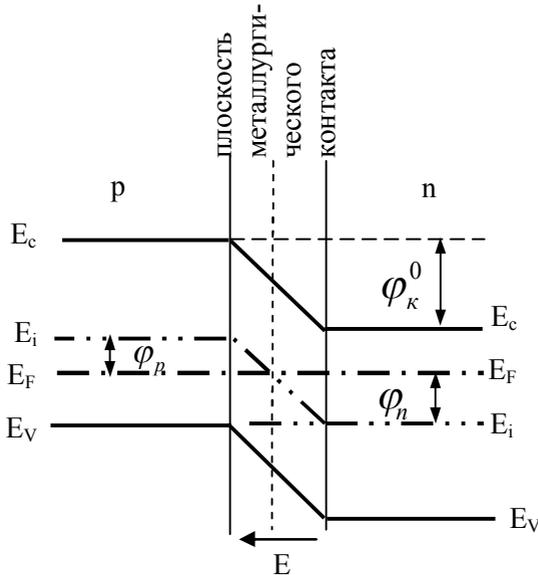


Рис. 4.2. Зонная диаграмма электронно-дырочного перехода в состоянии термодинамического равновесия

Плоскость, где происходит смена типа проводимости, называется плоскостью металлургического контакта.

Запишем выражение для расчета φ_k^o . Из рис.4.2 видно, что

$\varphi_k^o = \varphi_p + \varphi_n$. Значения φ_p и φ_n определяются следующим образом: $\varphi_n = E_F - E_i$; $\varphi_p = E_i - E_F$.

С учетом выражений (3.5 и 3.8) значения $(E_F - E_i)_n$ и $(E_i - E_F)_p$ выглядят следующим образом:

$$\left. \begin{aligned} \varphi_n = (E_F - E_i)_n = kT \ln \frac{N_d}{n_i} \end{aligned} \right\} \quad (4.1)$$

$$\varphi_p = (E_i - E_F)_p = kT \ln \frac{N_a}{n_i}$$

Тогда выражение для φ_k^o примет следующий вид

$$\varphi_k^o = kT \ln \frac{N_a N_d}{n_i^2}$$

(4.2.)

Из выражения (4.2) можно получить еще ряд соотношений, позволяющих производить расчет φ_k^o . Поскольку при полной ионизации примеси $n_{n0} = N_d$, а $p_{p0} = N_a$, тогда выражение

(4.2) принимает следующий вид

$$\varphi_k^o = kT \ln \frac{n_{n0} p_{p0}}{n_i^2}.$$

(4.3)

С учетом закона действующих масс для полупроводников n -типа ($n_{n0} p_{n0} = n_i^2$) и p -типа ($p_{p0} n_{p0} = n_i^2$) выражение (4.2) принимает вид

$$\varphi_k^o = kT \ln \frac{p_{p0}}{p_{n0}} \quad \varphi_k^o = kT \ln \frac{n_{n0}}{n_{p0}} \quad (4.4)$$

Наконец, учитывая, что значения n_i^2 можно определять, как

$n_i^2 = N_c N_v \exp\left(-\frac{\Delta E}{kT}\right)$, формула (4.2) представляется в виде

$$\varphi_k^o = \Delta E - kT \ln \frac{N_c N_v}{N_d N_a}. \quad (4.5)$$

Из анализа выражений (4.2)-(4.5) можно сделать следующие выводы.

Высота потенциального барьера φ_k^o зависит от концентрации легирующей примеси в полупроводниках p - и n -типов. Действительно, при изменении концентрации примеси в полупроводнике изменяется положение уровня Ферми. Для невырожденного

n – полупроводника значение φ_n может изменяться практически от нуля до $\Delta E/2$. В полупроводнике p -типа от нуля до $\Delta E/2$. Из чего следует, что максимальное значение φ_k^0 для $p-n$ переходов на основе невырожденных полупроводников не может быть больше ширины запрещенной зоны полупроводника ΔE .

Значение барьера φ_k^0 зависит от температуры. Учитывая, что с увеличением температуры уровень Ферми приближается к середине запрещенной зоны, значения φ_n и φ_p будут уменьшаться.

Следовательно, величина барьера φ_k^0 также уменьшается с ростом температуры. Это является одной из причин очень сильной температурной зависимости параметров полупроводниковых приборов на основе $p-n$ переходов.

Высота потенциального барьера φ_k^0 определяется шириной запрещенной зоны полупроводника, из которого изготовлены $p-n$ переходы: чем больше ΔE , тем выше барьер φ_k^0 . Типичные значения φ_k^0 для переходов из германия составляют $(0,3 \div 0,5)$ В, для кремниевых – $(0,6 \div 0,8)$ В. На рис.4.3 представлена зависимость φ_k^0 для резких $p-n$ переходов из различных материалов от концентрации примеси в слаболегированной области $p-n$ перехода.

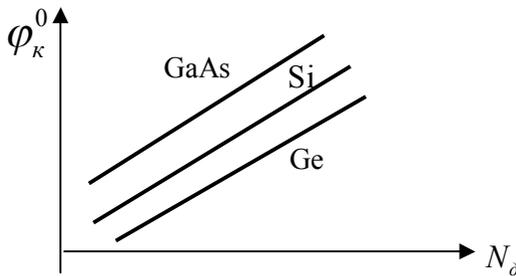


Рисунок 4.3. Зависимость φ_k^0 от концентрации примеси в слабом

4.4. Поток носителей зарядов в ЭДП по энергетическим диаграммам. Односторонняя проводимость $p-n$ перехода

Для понимания основного свойства электронно-дырочного перехода - явления односторонней проводимости, необходимо рассмотреть движение носителей заряда в переходе. Лучше всего это сделать с использованием зонных энергетических диаграмм р-п перехода. Рассмотрим движение носителей зарядов для состояния термодинамического равновесия, когда напряжение на переходе равно нулю ($U = 0$), при приложении прямого ($U > 0$) и обратного смещения ($U < 0$). Под прямым смещением на р-п переходе понимается такое смещение, когда отрицательный потенциал от источника питания прикладывается к полупроводнику n-типа.

На рис.4.4 представлены диаграммы для $U = 0$, $U > 0$ и $U < 0$.

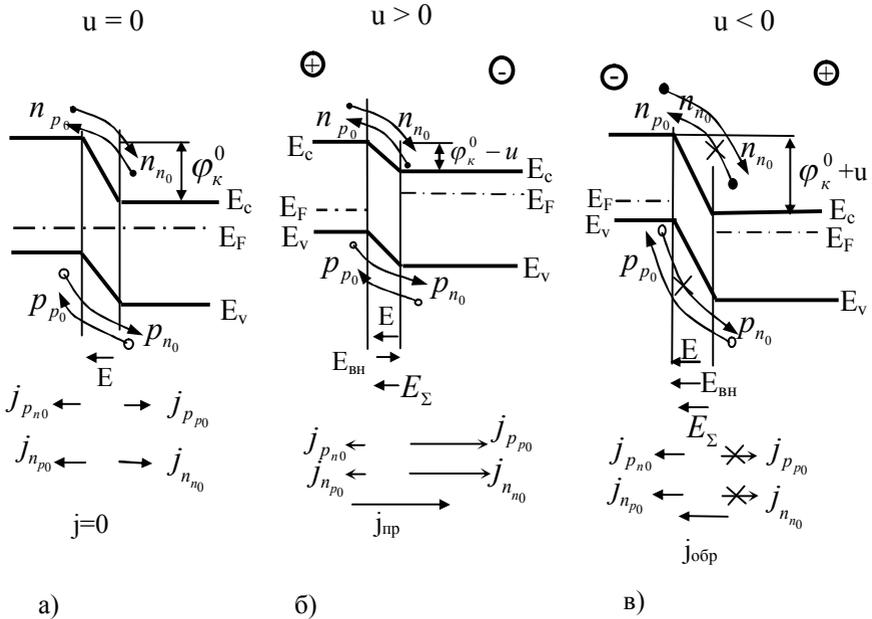


Рисунок 4.4. Зонные энергетические диаграммы р-п перехода в состоянии термодинамического равновесия (а), при прямом (б) и обратном смещении (в)

Для правильного понимания сущности работы ЭДП необходимо при рассмотрении вопроса помнить и руководствоваться уравнением полного тока: токи в полупроводнике образуются за счет движения электронов и дырок, а перемещаться они могут за счет сил электрического поля и градиента концентрации носителей.

При $U = 0$ под действием сил внутреннего электрического поля неосновные носители n_{p0} , p_{n0} втягиваются в ОПЗ перехода и, перемещаясь, создают токи неосновных носителей $j_{n_{p0}}$ и $j_{p_{n0}}$. Эти токи малы, поскольку концентрации неосновных носителей невелики. Для основных носителей (n_{n0}, p_{p0}) внутреннее поле является тормозящим. Оно создало барьер φ_k^o для этих носителей. И все же, поскольку носители в невырожденных полупроводниках подчиняются статистике Больцмана, то некоторая незначительная часть носителей способна преодолеть барьер φ_k^o и создать токи основных носителей $j_{p_{p0}}$ и $j_{n_{n0}}$. Рис.4.5 иллюстрирует распределение носителей по энергиям по статистике Максвелла-Больцмана. Ток основных носителей при $U = 0$ создается носителями, имеющими энергию $W > \varphi_k^o$. Поскольку таких носителей мало, то и токи $j_{p_{p0}}$ и $j_{n_{n0}}$ невелики.

В состоянии термодинамического равновесия токи основных носителей и токи неосновных носителей равны и противоположно направлены. Суммарный ток через $p-n$ переход равен нулю.

При приложении прямого смещения ($U > 0$), поскольку вектора внутреннего поля E и поля от внешнего источника $E_{вн}$ противоположно направлены, суммарное поле в переходе уменьшается, что приводит к понижению барьера φ_k^o на величину приложенного напряжения.

В результате диффузионные потоки основных носителей резко возрастают, давая значительные токи $j_{n_{n0}}$ и $j_{p_{p0}}$.

В этом случае через барьер $(\varphi_k^0 - U)$ может перейти значитель-

ное количество носителей (заштрихованная область справа от $(\varphi_k^0 - U)$ на рис.4.5).

Так как функция распределения Максвелла-Больцмана экспоненциальна, то это означает, что при понижении барьера φ_k^0 потоки основных носителей возрастают по экспоненте. Таким образом, при приложении прямого смещения ток через р-п переход определяется движением основных носителей заряда за счет их диффузии.

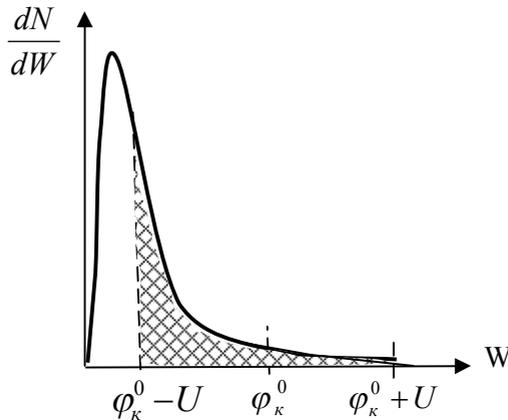


Рис. 4.5. Распределение носителей заряда в полупроводнике, приходящееся на единичный интервал энергий $\frac{dN}{dW}$ от энергии W по статистике Максвелла-Больцмана.

Он достаточно большой по величине и по экспоненте изменяется от приложенного напряжения.

С увеличением прямого смещения уменьшается суммарная напряженность электрического поля в р-п переходе. При этом уменьшается глубина проникновения поля в области полупро-

водника, прилегающего к металлургическому контакту. Поэтому уменьшается ширина ОПЗ р-п перехода.

При приложении обратного напряжения ($U < 0$) высота потенциального барьера φ_k^o увеличивается, так как суммарное электрическое поле возрастает. Количество основных носителей, способных преодолеть барьер ($\varphi_k^o + U$), по экспоненте убывает и становится ничтожно малым. Можно считать, что ток основных носителей при $U < 0$ падает до нуля. Для неосновных носителей зарядов барьера не существует, электрическое поле E_z велико и оно втягивает неосновные носители в ОПЗ р-п перехода, создавая обратный ток. Этот ток мал по величине и в идеальном р-п переходе не зависит от напряжения.

Здесь необходимо дать разъяснение относительно природы обратного тока. Поскольку первопричиной движения неосновных носителей является электрическое поле, то можно полагать, что этот ток является дрейфовым, и он должен определяться напряженностью электрического поля.

Однако обратный ток в идеальном р-п переходе от напряжения не зависит. Это говорит о том, что обратный ток не является дрейфовым. Тогда остается полагать, что ток и основных и неосновных носителей заряда в р-п переходе является диффузионным по своей природе. Действительно, если за счет электрического поля, например, электроны из р-полупроводника втягиваются полем в ОПЗ, то у границы ОПЗ число электронов n_{p0} становится меньше, чем в объеме полупроводника р-типа. Это значит, что электроны из объема р-полупроводника начнут диффундировать к ОПЗ р-п перехода, создавая диффузионный ток.

Таким образом, показано, что р-п переход обладает эффектом односторонней проводимости и имеет диодную характеристику, изображенную на рис.4.6. Эффект односторонней проводимости заключается в том, что р-п переход хорошо проводит ток в прямом направлении и имеет малый ток при обратном напряжении (см. рис. 4.6.).

При увеличении обратного напряжения электрическое поле E_{Σ} возрастает.

Это приводит к увеличению глубины проникновения поля в области полупроводников, т.е. с увеличением обратного напряжения толщина р-п перехода увеличивается.

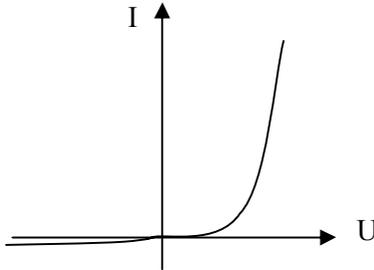


Рисунок 4.6. Качественный вид ВАХ электронно-дырочного перехода

4.5. Некоторые понятия и определения по электронно-дырочному переходу

Инжекция неосновных носителей заряда - это явление перехода основных носителей заряда через ОПЗ р-п перехода в область, где они становятся неосновными. Например, при прямом смещении электроны из n-области, где они являются основными носителями, переходят в р-область, где они становятся неосновными носителями. Поскольку при обратном смещении основные носители не могут преодолеть барьер $(\varphi_k^o + U)$ (см. рис.4.4 в), то явление инжекции при обратном смещении отсутствует. Можно говорить о том, что, поскольку прямой ток в р-п переходе связан с движением основных носителей, прямой ток обусловлен инжекцией неосновных носителей заряда.

Экстракция неосновных носителей заряда - это явление втягивания неосновных носителей заряда электрическим полем электронно-дырочного перехода. Так как при обратном смещении р - n перехода инжекции нет, то, следовательно, обратный ток в переходе обусловлен явлением экстракции.

Симметричный и несимметричный электронно-дырочный переход. По соотношению концентрации легирующей примеси в р-и n-областях р-п переходы делятся на симмет-

ричные и несимметричные. Если $N_a = N_d$, то ширина ОПЗ, лежащая в p - и n -областях, одинакова, и такой переход называется симметричным (рис.4.7 а). В нем $d_p = d_n$ и токи в переходе в равной степени обусловлены движением электронов и дырок (рис.4.7 а, б).

Если, например, $N_a \gg N_d$, то ОПЗ будет преимущественно располагаться в n -области перехода, поскольку глубина проникновения поля в полупроводник определяется концентрацией носителей.

При $N_a \gg N_d$ $p_{p0} \gg n_{n0}$ и поле в p -полупроводник проникает на меньшую глубину, чем в n -полупроводник (рис.4.8 а). При этом изменяется и характер протекания тока. Поскольку $p_{p0} \gg n_{n0}$, то потоком электронов из n -области в p -область можно пренебречь, так как они дают токи одного направления (рис.4.8 б). Прямой ток в таком несимметричном переходе обусловлен движением в основном дырок p_{p0} . Поскольку в полупроводнике p -типа концентрация примеси велика, то, следовательно, велико и значение p_{p0} .

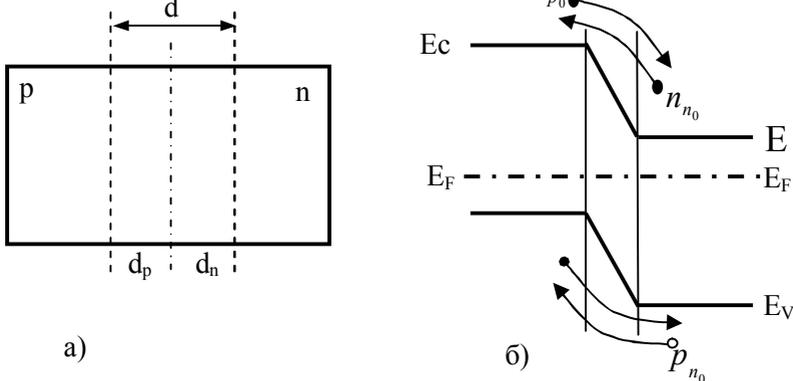
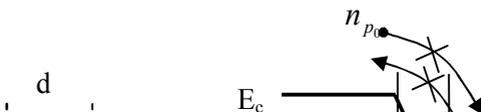


Рисунок 4.7. Симметричный p - n переход

А это означает, что концентрация неосновных носителей n_{p0} ничтожно мала и их движением также можно пренебрегать.



Таким образом в $p^+ - n$ переходе и прямой и обратные токи обусловлены движением только дырок (рис.4.8 б).

Для случая, когда $N_o \gg N_a$, ситуация представлена на рис.4.9. В данном случае $d_p \gg d_n$ и ток в $n^+ - p$ переходе в силу указанных выше причин переносится в основном электронами (рис.4.9 а,б).

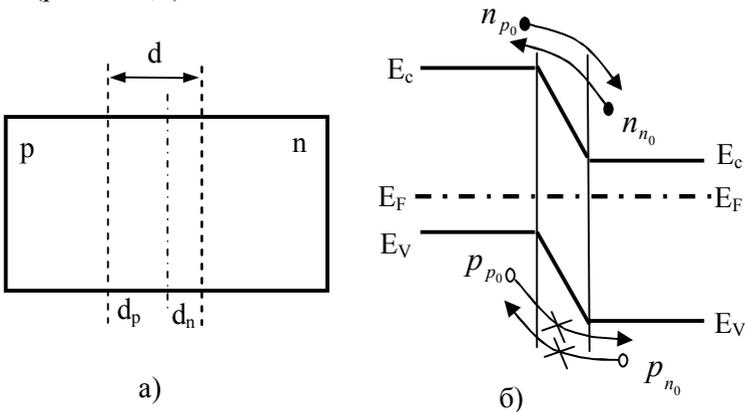


Рис. 4.9. Несимметричный $n^+ - p$ переход.

В полупроводниковой электронике обычно используются несимметричные $p^+ - n$ или $n^+ - p$ переходы.

Эмиттер - это сильнолегированная область $p-n$ перехода, откуда идет инжекция неосновных носителей. Для $p^+ - n$ перехода эмиттер - это p -область, для $n^+ - p$ - это n -область.

База - это слаболегированная область $p-n$ перехода, куда идет инжекция неосновных носителей.

Плавный и резкий электронно-дырочный переходы. По характеру распределения примеси в $p-n$ переходе все переходы делятся на резкие и плавные.

Резкий $p-n$ переход - это переход, в котором размер области, где происходит смена типа проводимости d' , много меньше ширины ОПЗ d , т.е. $d' \ll d$ (рис.4.10).

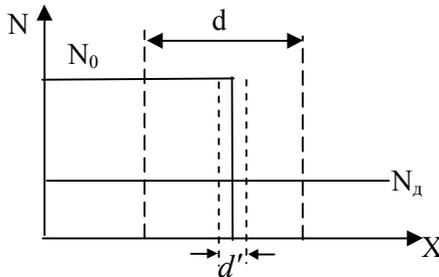


Рисунок 4.10. Резкий $p-n$ переход

Плавный $p-n$ переход - это переход, в котором ширина области, где изменяется тип проводимости d' , сравнима с шириной ОПЗ d , т.е. $d' \approx d$ (рис.4.11).

Вид перехода (резкий или плавный) определяется технологией изготовления $p-n$ перехода.

Точечные и плоскостные электронно-дырочные переходы. По геометрическим размерам (площадь перехода) $p-n$ переходы подразделяются на точечные и плоскостные.

Точечный - это переход, у которого линейные размеры, определяющие площадь, много меньше, чем характеристическая длина.

Плоскостной - это переход, у которого линейные размеры, определяющие площадь, много больше характеристической длины.

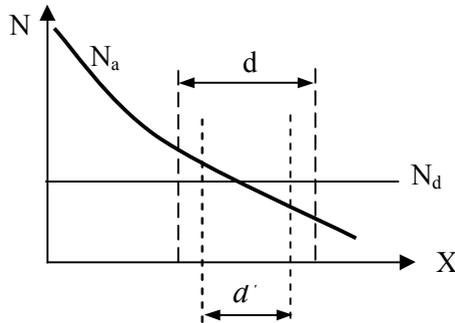


Рисунок 4.11. Плавный p-n переход

Под характеристической длиной понимается либо длина диффузионного смещения, либо толщина базы, в зависимости от того, какова толщина базы по отношению к длине диффузионного смещения.

Если толщина базы $W \ll L$, то характеристической длиной является толщина базы W .

4.6. Методы получения электронно-дырочного перехода

Для получения электронно-дырочных переходов наиболее часто используются четыре метода: сплавление, диффузия, эпитаксия, ионное легирование. Рассмотрим кратко каждый из методов.

Сплавление. Для получения сплавного p-n перехода на поверхности полупроводника располагают навеску. Например, на поверхности полупроводника n-типа располагают навеску из алюминия или галлия. Графитовую кассету с большим числом кристаллов полупроводника с навесками помещают в печь для сплавления и нагревают до температур выше температуры плавления навески. В расплавленном состоянии навеска растворяет в

себе полупроводник. После охлаждения происходит рекристаллизация полупроводникового кристалла с примесью расплавленной навески с образованием полупроводникового слоя с другим, р- типом проводимости. Образуется сплавной электронно-дырочный переход с резкой сменой типа проводимости (резкий р-п переход).

Диффузия. При создании диффузионных р-п переходов используют диффузию в полупроводник примеси, находящейся в газообразном, жидком или твердом состоянии. Например, на поверхности полупроводника п-типа располагают навеску из галлия (элемент третьей группы). При помещении структур в диффузионную печь и нагревании до температуры ниже, чем температура плавления навески, происходит диффузия атомов навески в полупроводник. После охлаждения в кристалле полупроводника формируется р-п переход с плавным распределением примеси (см. рис.4.11), т.е. плавный р-п переход. При жидкофазной и газофазной диффузии в жидкости или газе находятся элементы третьей группы, которые в процессе диффузионного нагрева проникают в слой полупроводника, изменяя тип проводимости на противоположный.

Эпитаксия. При производстве полупроводниковых приборов широко используют эпитаксиальное наращивание - наращивание монокристаллических слоев полупроводника на поверхности монокристаллической подложки того же полупроводника. В зависимости от использованной примеси можно получить монокристаллический слой полупроводника с другим типом проводимости, чем у подложки, т.е. создать резкий эпитаксиальный р-п переход.

Ионное легирование. Суть метода состоит в бомбардировке поверхности полупроводника ионами элементов третьей или пятой группы системы Менделеева. Ионы получают и ускоряют в специальных установках до энергий $150 \div 200$ кэВ. Имея высокую энергию, ионы внедряются в полупроводник и распределяются вглубь образца по закону функции ошибок, создавая плавный р-п переход. Поскольку возможно получение ионов третьей и пятой групп, то р-п переходы можно получать путем бомбардировки полупроводников р- и п-типов проводимости.

4.7. Расчет напряженности электрического поля и потенциала в ЭДП. Ширина ОПЗ для резкого и плавного переходов

Для нахождения напряженности электрического поля и потенциала в $p-n$ переходе необходимо решить уравнение Пуассона

$$\frac{d^2\varphi}{dx^2} + \frac{d^2\varphi}{dy^2} + \frac{d^2\varphi}{dz^2} = -\frac{\rho(x, y, z)}{\varepsilon\varepsilon_0}.$$

(4.6)

Решение этого уравнения проведем при следующих допущениях.

Рассматриваем одномерную задачу.

Полагаем, что в ОПЗ подвижных носителей заряда нет.

Все поле сосредоточено в ОПЗ. Прилегающие к переходу области полупроводников являются электрически нейтральными.

Такие допущения соответствуют обратному и малым прямым напряжениям на переходе.

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

$$\frac{d^2\varphi}{dx^2} = -\frac{qN(x)}{\varepsilon\varepsilon_0},$$

(4.7)

где $N(x) = N_d - N_a$ – приведенная концентрация.

Решение уравнения (4.7) проведем при следующих граничных условиях

$$\left. \begin{array}{l} x = -d_p; \quad \varphi(-d_p) = 0; \quad \frac{d\varphi}{dx} = 0 \\ x = -d_n; \quad \varphi(d_n) = \varphi_k^o \pm U; \quad \frac{d\varphi}{dx} = 0 \end{array} \right\} \quad (4.8)$$

Поскольку необходимо найти распределение напряженности и потенциала для резкого и плавного переходов, то сначала найдем решение уравнения (4.7) в общем виде, а затем общее реше-

ние применим для частных случаев (резкого и плавного переходов).

Уравнение (4.7) можно переписать в следующем виде

$$\frac{d}{dx} \left(\frac{d\varphi}{dx} \right) = -\frac{qN(x)}{\varepsilon\varepsilon_0}, \text{ тогда } d \left(\frac{d\varphi}{dx} \right) = -\frac{qN(x)}{\varepsilon\varepsilon_0} dx.$$

(4.9)

Для упрощения записи будем пользоваться определенными интегралами с переменными пределами интегрирования.

Интегрируя (4.9), получим выражение для напряженности поля $E(x)$

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

т.е.
$$E(x) = \frac{q}{\varepsilon\varepsilon_0} \int_{-d_p}^x N(x) dx.$$

(4.10)

Для нахождения выражения для расчета потенциала $\varphi(x)$ необходимо проинтегрировать выражение (4.10)

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

(4.11)

Интегрируя (4.11) по частям, получим следующее выражение

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

(4.12)

С учетом граничных условий (4.8) выражение (4.12) принимает следующий вид

$$\varphi_k^o \pm U = \frac{q}{\varepsilon\varepsilon_0} \int_{-d_p}^{d_n} xN(x) dx$$

(4.13)

с учетом того, что равенство нулю интеграла $\int_{-d_p}^{d_n} N(x)dx = 0$ отражает условие электронейтральности в ОПЗ р-п перехода.

Таким образом, выражения для расчета напряженности электрического поля и потенциала в р-п переходе в общем виде найдены. Теперь применим эти общие выражения для расчета линейного плавного и резкого переходов, поскольку в этом случае функции распределения примеси легко задаются аналитически.

Резкий электронно-дырочный переход. Для резкого р-п перехода, исходя из его определения, можно записать

$$\left. \begin{aligned} N(x) &= -N_a \text{ при } -d_p < x < 0 \\ N(x) &= N_d \text{ при } 0 < x < d_n. \end{aligned} \right\} \quad (4.14)$$

Для области полупроводника р-типа ($-d_p < x < 0$) получим выражение для расчета напряженности путем интегрирования (4.10) с учетом условий (4.14)

$$\left(\frac{d\varphi}{dx}\right)_p = \frac{dN_a}{\varepsilon\varepsilon_0}(x + d_p). \quad (4.15)$$

Выражение (4.15) показывает, что в р-области р-п перехода напряженность поля является линейной функцией координаты. При $x = -d_p$ поле равно нулю, а при $x = 0$ напряженность максимальна. Учитывая, что в резком переходе подинтегральная функция $N(x)$ в выражении (4.10) имеет разрыв, целесообразно записать

$$\left(\frac{d\varphi}{dx}\right)_n = -\frac{q}{\varepsilon\varepsilon_0} \left[\int_{-d_p}^0 N(x)dx + \int_0^x N(x)dx \right] = \frac{q}{\varepsilon\varepsilon_0} (N_a d_p - N_d x) \quad (4.16)$$

то есть в области п-полупроводника напряженность также является линейной функцией координаты, обращаясь в ноль при

$x = d_n$, т.к. $N_a d_p - N_d d_n = 0$ – условие электронейтральности ОПЗ p-n перехода.

Для нахождения потенциала $\varphi(x)$ требуется проинтегрировать выражения (4.15) и (4.16). Тогда

$$\varphi(x)_p = \frac{q}{\varepsilon\varepsilon_0} N_a \int_{-d_p}^x (x + d_p) dx = \frac{q}{2\varepsilon\varepsilon_0} N_a (x + d_p)^2 \quad (4.17)$$

Потенциал φ в области p-полупроводника является квадратичной функцией координаты. При $x = -d_p$ $\varphi(x)_p = 0$. Для области p-n перехода запишем по аналогии с (4.16)

$$\varphi(x)_n = \int_{-d_p}^0 \left(\frac{d\varphi}{dx} \right) dx + \int_0^x \left(\frac{d\varphi}{dx} \right)_n dx = \frac{q}{2\varepsilon\varepsilon_0} N_a d_p^2 + \frac{q}{\varepsilon\varepsilon_0} N_a d_p x - \frac{q}{2\varepsilon\varepsilon_0} N_d x^2, \quad (4.18)$$

т.е. $\varphi(x)_n$ – квадратичная функция координаты. При $x = d_n$ $\varphi(x)_n = \varphi_k^0 \pm U$.

Для определения ширины ОПЗ p-n перехода будем исходить из того, что общая ширина $d = d_p + d_n$ и по условию электронейтральности $N_a d_p = N_d d_n$. Тогда можно записать, что

$$\begin{aligned} d &= d_p + \frac{N_a}{N_d} d_p = \frac{N_a + N_d}{N_d} d_p; & d_p &= \frac{N_d}{N_a + N_d} d \\ d &= d_n + \frac{N_d}{N_a} d_n = \frac{N_a + N_d}{N_a} d_n; & d_n &= \frac{N_a}{N_a + N_d} d. \end{aligned} \quad (4.19)$$

Для получения выражения для ширины перехода необходимо записать полное падение потенциала после интегрирования (4.13) с учетом условий (4.14) для резкого перехода

$$\varphi_k^0 \pm U = \frac{q}{2\varepsilon\varepsilon_0} (N_a d_p^2 + N_d d_n^2). \quad (4.20)$$

После подстановки (4.19) в (4.20) получим

$$\varphi_k^o \pm U = \frac{qN_a N_d}{2\varepsilon\varepsilon_0(N_a + N_d)} d^2, \text{ отсюда}$$

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \cdot \frac{N_a + N_d}{N_d N_a} \cdot (\varphi_k^o \pm U)}.$$

(4.21)

Таким образом, ширина $p-n$ перехода определяется концентрацией легирующей примеси и зависит от напряжения на переходе: при приложении прямого смещения ширина уменьшается, при обратном - увеличивается.

Для несимметричных переходов, когда либо $N_a \gg N_d$, либо $N_d \gg N_a$, выражение (4.21) принимает более простой вид

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0}{qN_d} \cdot (\varphi_k^o \pm U)},$$

(4.22)

где N_d – концентрация примеси в базе.

Из выражений (4.21) и (4.22) следует, что прямые напряжения, которые можно подавать на $p-n$ переход, не могут быть больше φ_k^o , поскольку при $U > \varphi_k^o$ разница $(\varphi_k^o - U)$ становится отрицательной. При $U \geq \varphi_k^o$ $p-n$ переход исчезает.

На рис.4.12 приведены графики распределения E, φ и N от координаты для резкого электронно-дырочного перехода. Пунктирными линиями показано распределение плотности заряда с учетом подвижных носителей в $p-n$ переходе. Эти заряды несколько изменяют ход напряженности и потенциала, но не меняют характера зависимости.

Плавный электронно-дырочный переход с линейным распределением концентрации легирующей примеси. Для расчета плавного $p-n$ перехода примем, что $N(x) = ax$, т.е. примеси распределены по объему полупроводника по линейному закону,

a – градиент концентрации примесей, который можно считать постоянным при малой ширине ОПЗ.

Решение уравнения (4.10) с учетом выражения для $N(x) = ax$ дает следующее

$$\frac{d\varphi}{dx} = -\frac{q}{\varepsilon\varepsilon_0} \int_{-d_p}^x N(x)dx = -\frac{q}{\varepsilon\varepsilon_0} \int_{-d_p}^x axdx = -\frac{qa}{2\varepsilon\varepsilon_0}(x^2 - d_p^2).$$

(4.23)

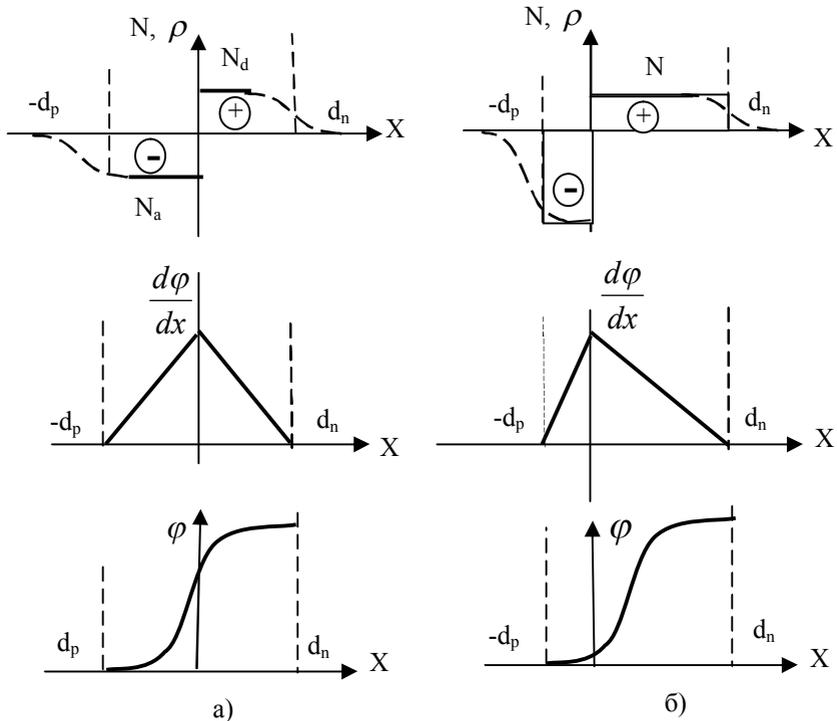


Рисунок 4.12. Распределение концентрации легирующей примеси N , плотности объемного заряда ρ , градиента концентрации $\frac{d\varphi}{dx}$, и потенциала φ в резком симметричном (а) и несимметричном (б) р-п переходах

Видно, что напряженность электрического поля является квадратичной параболой.

Интегрирование выражения (4.23) дает ход потенциала

$$\varphi(x) = -\frac{qa}{2\varepsilon\varepsilon_0} - d \int_{-d_p}^x (x^2 - d_p^2) dx = -\frac{qa}{\varepsilon\varepsilon_0} \left(\frac{x^3}{6} - \frac{d_p^2 x}{2} - \frac{d_p^3}{3} \right), \quad (4.24)$$

т.е. $\varphi(x)$ – кубическая парабола.

Для симметричного плавного перехода $d_p = d_n = \frac{d}{2}$. Тогда решение уравнения (4.13) с учетом характера перехода дает

$$\varphi_k^o \pm U = \frac{qa}{12\varepsilon\varepsilon_0} d^3, \text{ откуда}$$

$$d = \sqrt[3]{\frac{12\varepsilon\varepsilon_0}{qa} (\varphi_k^o \pm U)}. \quad (4.25)$$

Видно, что ширина $p-n$ перехода зависит от градиента концентрации и приложенного к переходу напряжения. При этом очевидно, что ширина ОПЗ в плавном переходе в меньшей степени изменяется при изменении напряжения.

4.8. Вольтамперная характеристика идеального ЭДП.

Диоды с "толстой" и "тонкой" базами

Расчет ВАХ $p-n$ перехода проведем при следующих допущениях:

Рассматриваем одномерную модель;

Электрическое поле сосредоточено только в ОПЗ $p-n$ перехода, т.е. токи, протекающие в переходе, малы и не вызывают существенного падения напряжения на сопротивлении базы;

Омические переходы идеальны, т.е. около них в полупроводнике нет избыточной концентрации неосновных носителей;

Явления на поверхности полупроводника несущественны; В ОПЗ перехода не наблюдается процессов генерации и рекомбинации неравновесных носителей:

Рекомбинация неосновных носителей в базе происходит по линейному закону, т.е. скорость рекомбинации пропорциональна избыточной концентрации неосновных носителей.

При таких допущениях описываемая зависимость тока от напряжения соответствует ВАХ идеального $p-n$ перехода. Надо иметь в виду, что при выводе ВАХ такого идеального перехода учтены только явления инжекции и экстракции неосновных носителей. При выводе ВАХ надо исходить из того, что токи, текущие в $p-n$ переходе, являются диффузионными по своей природе, и для их нахождения необходимо знать градиент концентрации неосновных носителей.

На рис.4.13 представлены распределения концентрации неосновных носителей в симметричном $p-n$ переходе. Симметрия перехода позволяет найти выражение для расчета тока, например, дырок, а затем по аналогии записать выражение тока для электронов.

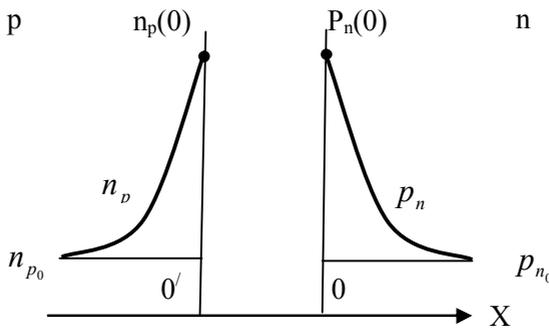


Рис. 4.13. Распределение концентрации неосновных носителей при прямом смещении $p-n$ перехода.

Для нахождения градиента концентрации важно знать концентрацию носителей на границе перехода (при $x = 0$ и $x = 0'$); $n_p(0)$ и $p_n(0)$. В соответствии со статистикой Максвелла-Больцмана эти выражения имеют следующий вид

$$\left. \begin{aligned} p_n(0) &= p_{n0} \exp \frac{qU}{kT} \\ n_p(0) &= n_{p0} \exp \frac{qU}{kT} \end{aligned} \right\} \quad (4.26)$$

т.е. концентрация неосновных носителей на границе определяется равновесной концентрацией неосновных носителей и приложенным к $p-n$ переходу напряжением. Концентрации $n_p(0)$ и $p_n(0)$ называются граничными.

Выведем ВАХ $p-n$ перехода при приложении к нему напряжения, имеющего постоянную и малую переменную составляющие

$$U = U + U_m \exp j\omega t.$$

Условие малости переменной составляющей означает, что

$$U_m \ll \frac{kT}{q}, \text{ т.е. амплитуда переменного сигнала не превышает}$$

10 мВ.

Существует несколько способов вывода ВАХ $p-n$ перехода, однако наиболее строго вывод базируется на решении уравнения непрерывности. Рассмотрим дырочную составляющую

$$\text{тока } p-n \text{ перехода } \frac{\partial p_n}{\partial t} = -\frac{1}{q} \operatorname{div} j_p - R_p + G_p.$$

При принятых допущениях $G_p - R_p = \frac{p_n - p_{n0}}{\tau_p}$, а

$$j_p = -qD_p \frac{\partial p_n}{\partial x}, \quad \frac{\partial p_n}{\partial t} = D_p \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p} \quad (4.27)$$

Решение уравнения (4.27) проведем с учетом следующих обстоятельств. Так при $x = 0$

$$p_n(0) = p_{n0} \exp \frac{q}{kT} (U + \dot{U}_m \exp j\omega t) = p_{n0} \exp \frac{qU}{kT} \exp \left(\frac{q\dot{U}_m}{kT} \exp j\omega t \right).$$

Поскольку аргумент второй экспоненты мал, можно разложить

ее в ряд, взяв два первых члена разложения ($\exp y \approx 1 + y + \dots$). В результате получим, что при $x = 0$

$$p_n(0) = p_{n0} \exp \frac{qU}{kT} + p_{n0} \left(\exp \frac{qU}{kT} \right) \frac{q\dot{U}_m}{kT} \exp j\omega t. \quad (4.28)$$

Таким образом, концентрация неосновных носителей заряда около $p-n$ перехода имеет постоянную и переменную составляющие. Частота изменения переменной составляющей та же, что и частота приложенного переменного напряжения.

Для преобразования дифференциального уравнения (4.27) выберем форму решения в виде суммы постоянной и переменной составляющих концентрации. Тогда решение (4.27) должно иметь вид

$$p_n(x, t) = p_{n0} + \Delta p_n(x) + \dot{p}_n \exp j\omega t. \quad (4.29)$$

где $p_n(x, t)$ — полная концентрация неосновных носителей заряда;

$\Delta p_n(x)$ — постоянная составляющая избыточной концентрации неосновных носителей, равная $(p_n - p_{n0})$, зависящая только от координаты;

$\dot{p}_n(x) \exp j\omega t$ — переменная составляющая избыточной концентрации неосновных носителей заряда, зависящая от координаты и времени.

После подстановки (4.29) в (4.27) получим

$$\begin{aligned} \dot{p}_n(x) j\omega \exp j\omega t = D_p \frac{\partial^2 [\Delta p_n(x)]}{\partial x^2} + D_p \frac{\partial^2 \dot{p}_n(x)}{\partial x^2} \exp j\omega t - \\ - \frac{\Delta p_n(x)}{\tau_p} - \frac{\dot{p}_n(x)}{\tau_p} \exp j\omega t. \end{aligned} \quad (4.30)$$

В уравнении (4.30) есть слагаемые, зависящие и не зависящие от времени. Уравнение (4.30) справедливо только в том случае, если алгебраические суммы не зависящих и зависящих от времени

составляющих отдельно равны нулю.

Для постоянной составляющей из уравнения (4.30) получим

$$\frac{d^2[\Delta p_n(x)]}{dx^2} = \frac{\Delta p_n(x)}{L_p^2}, \quad (4.31)$$

которое необходимо решать при следующих граничных условиях для избыточной концентрации Δp_n .

$$\left. \begin{aligned} x = 0 \quad \Delta p_n(0) &= p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \\ x = W_n \quad \Delta p_n(W_n) &= 0. \end{aligned} \right\} \quad (4.32)$$

Смысл параметра W_n ясен из рис.4.14.

Для переменной составляющей избыточной концентрации

$$\frac{d^2[\dot{p}_n(x)]}{dx^2} = \frac{\dot{p}_n(x)}{\dot{L}_p^2}, \quad (4.33)$$

где $\dot{L}_p = \frac{L_p}{\sqrt{1 + j\omega\tau_p}}$ — комплексная длина диффузионного смещения.

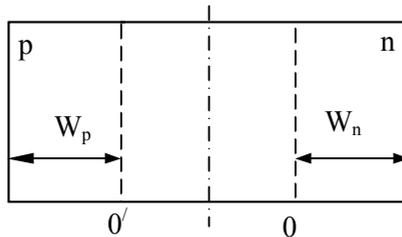


Рис. 4.14. Симметричный одномерный $p - n$ переход, принятый для расчета

Решение проводится при граничных условиях

$$\left. \begin{aligned} x = 0 \quad \dot{p}_n(0) &= p_{n0} \left(\exp \frac{qU}{kT} \right) \frac{q\dot{U}_m}{kT} \\ x = W_n \quad \dot{p}_n(W_n) &= 0. \end{aligned} \right\} \quad (4.34)$$

Полученные уравнения для постоянной и переменной составляющих имеют сходный вид. Поэтому дальше проведем решение для постоянной составляющей. Для переменной составляющей решение аналогично.

Решение дифференциального уравнения (4.31) удобно искать с использованием гиперболических функций в виде

$$\Delta p_n(x) = A_1 ch \frac{x}{L_p} + A_2 sh \frac{x}{L_p}, \quad (4.35)$$

поскольку это упрощает поиск произвольных постоянных, если заданы граничные условия, причем одно из них нулевое.

Подставляя в уравнение (4.35) значение $x = 0$ и граничное условие (4.32), получим

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

При использовании второго граничного условия при $x = W_n$ получим значение коэффициента A_2 .

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

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

Окончательный вид уравнения (4.35) примет вид

$$\Delta p_n(x) = p_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \left(ch \frac{x}{L_p} - cth \frac{W_n}{L_p} sh \frac{x}{L_p} \right). \quad (4.36)$$

Зная характер распределения избыточных носителей, и учитывая, что ток дырок является диффузионным, после дифференцирования (нахождения градиента концентрации) получим.

$$j_p(x) = -\frac{qD_p p_{n0}}{L_p} \left(sh \frac{x}{L_p} - cth \frac{W_n}{L_p} ch \frac{x}{L_p} \right) \left(\exp \frac{qU}{kT} - 1 \right). \quad (4.37)$$

Для $x = 0$ плотность тока дырок определяется следующим выражением

$$j_p(0) = \frac{qD_p p_{n0}}{L_p} \operatorname{cth} \frac{W_n}{L_p} \left(\exp\left(\frac{qU}{kT}\right) - 1 \right). \quad (4.38)$$

По аналогии для плотности тока электронов при $x = 0'$ имеем

$$j_n(0') = \frac{qD_n n_{p0}}{L_n} \operatorname{cth} \frac{W_p}{L_p} \left(\exp\left(\frac{qU}{kT}\right) - 1 \right). \quad (4.39)$$

Суммируя плотности токов электронов и дырок в одном сечении, получаем

$$j = \left(\frac{qD_p p_{n0}}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{qD_n n_{p0}}{L_n} \operatorname{cth} \frac{W_p}{L_n} \right) \left(\exp \frac{qU}{kT} - 1 \right). \quad (4.40)$$

Обозначим через j_s плотность тока насыщения, который практически не зависит от напряжения

$$j_s = \frac{qD_p p_{n0}}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{qD_n n_{p0}}{L_n} \operatorname{cth} \frac{W_p}{L_p}. \quad (4.41)$$

Таким образом, ВАХ $p-n$ перехода записывается следующим образом

$$j = j_s \left(\exp \frac{qU}{kT} - 1 \right). \quad (4.42)$$

Аналитический вид ВАХ $p-n$ перехода полностью совпадает с ВАХ диода Шоттки, поэтому анализ характеристики аналогичен: при малых прямых и обратных напряжениях, когда

$U \ll \frac{kT}{q}$ – ВАХ представляет линейную зависимость тока от напряжения $j \sim \frac{q}{kT}U$; при больших обратных напряжениях $j = -j_s$; при больших прямых - $j \sim \exp \frac{qU}{kT}$.

Несколько замечаний по поводу тока насыщения.

Выражение для j_s , представленное формулой (4.41), может быть преобразовано путём замены концентрации неосновных носителей $n_{p0} = \frac{n_i^2}{P_{p0}}$ и $P_{n0} = \frac{n_i^2}{n_{n0}}$. Тогда выражение для j_s примет вид

$$j_s = n_i^2 \left(\frac{qD_p}{L_p n_{n0}} + \frac{qD_n}{L_n P_{p0}} \right).$$

(4.43)

Из (4.43) следует, что ток насыщения в р-п перехода является очень сильно зависящим от температуры ($n_i^2 = N_c N_v \exp - \frac{\Delta E}{kT}$) и ширины запрещённой зоны полупроводника: с ростом температуры j_s по экспоненте возрастает.

2. Поскольку $j_s \sim n_i^2$, то это означает, что для р-п перехода из германия ($n_i = 2.5 * 10^{13} \text{ см}^{-3}$) и кремния ($n_i = 10^{10} \text{ см}^{-3}$) токи могут различаться на шесть порядков. Причём токи насыщения в германиевых диодах больше, чем в кремниевых.

3. Формула для j_s (4.43) может быть также записана в форме (4.44) путём домножения слагаемых в скобке на L_p и L_n .

$$j_s = \frac{qD_p P_{n0} L_p}{\tau_p} + \frac{qD_n n_{p0} L_n}{\tau_n}$$

(4.44)

Выражения для ВАХ $p-n$ перехода (4.40) в таком виде практически не используется, поскольку в зависимости от соотношения между длиной диффузного смещения L и толщиной базы W все переходы делятся на переходы с толстой и тонкой базами.

Диод с толстой базой - это диод, у которого толщина базы превышает длину диффузного смещения ($W \gg L$). В этом случае аргументы гиперболических котангенсов в соотношении (4.40) значительно превышают единицу, а сами гиперболические котангенсы близки к единице: $cth \frac{W_n}{L_p} \approx 1$ и $cth \frac{W_p}{L_p} \approx 1$

Воспользовавшись формулами Эйлера

$$shy = \frac{e^y - e^{-y}}{2} \text{ и } chy = \frac{e^y + e^{-y}}{2}$$

запишем распределение концентрации неосновных носителей в диоде с толстой базой:

$$\Delta P_n(x) = P_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \exp - \frac{X}{L_p} \quad (4.45)$$

т.е. избыточные неосновные носители распределяются по экспоненциальному закону в объёме полупроводника. Вид распределения концентрации неосновных носителей в диоде с толстой базой при различных напряжениях на переходе показан на рис.

4.15. Тогда с учётом того, что $cth \frac{W}{L} \approx 1$, выражение для ВАХ принимает следующий вид:

$$j = \left(\frac{qD_p P_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) \left(\exp \frac{qU}{kT} - 1 \right) \quad (4.46)$$

где ток насыщения равен

$$j_s = \frac{qD_p P_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n}$$

(4.47)

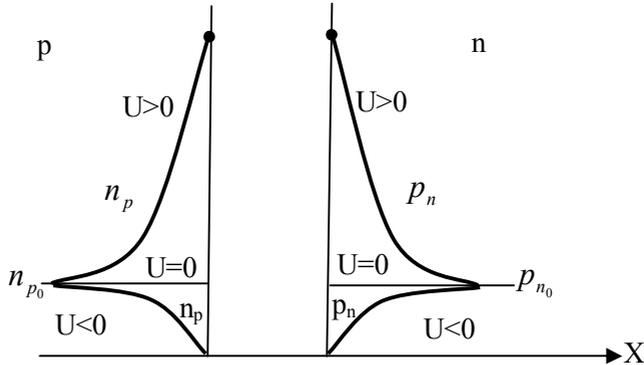


Рисунок 4.15. Распределение концентрации неосновных носителей в диоде с толстой базой при различных напряжениях

Диод с тонкой базой - это диод, толщина базы которого значительно меньше диффузионной длины неосновных носителей ($W_n \ll L_p$). В этом случае аргументы гиперболических функций в соотношении (4.40) малы. Разлагая гиперболические котангенсы в ряд и ограничиваясь первым членом разложения ($cth x \approx \frac{1}{x}$; $shx \approx x$ и $chx \approx 1$), получим выражение для распределения концентрации неосновных носителей в диоде с тонкой базой

$$\Delta P_n(x) = P_{n0} \left(\exp \frac{qU}{kT} - 1 \right) \left(1 - \frac{x}{W_n} \right) \quad (4.48)$$

Из (4.48) видно, что в диоде с тонкой базой распределение концентрации неосновных носителей подчиняется линейному закону (рис.4.16).

Плотность тока для диода с тонкой базой определяется выражением (4.49).

$$j = \left(\frac{qD_p P_{n0}}{W_n} + \frac{qD_n n_{p0}}{W_p} \right) \left(\exp \frac{qU}{kT} - 1 \right). \quad (4.49)$$

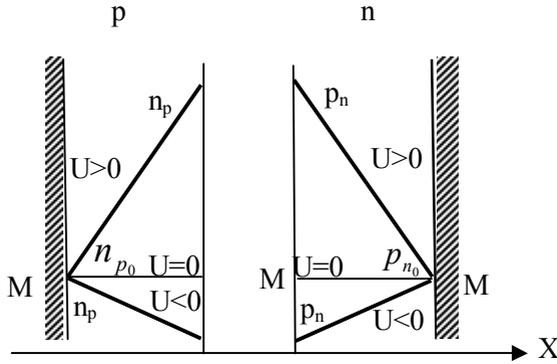


Рисунок 4.16 . Распределение концентрации неосновных носителей заряда в диоде с тонкой базой при различных напряжениях

В этом случае обратный ток насыщения определяется как

$$j_s = \frac{qD_p P_{n0}}{W_n} + \frac{qD_n n_{p0}}{W_p} \tag{4.50}$$

Физически выражение (4.49) понятно: так как ток неосновных носителей является диффузионным, то для его нахождения градиент концентрации определяется как $\frac{p_n(0) - p_{n0}}{W_n}$ и $\frac{n_p(0) - n_{p0}}{W_p}$.

Отличие формул (4.46) и (4.49) состоит в различии характеристического параметра: в диоде с толстой базой - это длина диффузного смещения, в диоде с тонкой базой - толщина базы W .

Необходимо отметить, что обратная ветвь ВАХ p-n-переходов для диодов с толстой и тонкой базой отличается (рис.4.17)

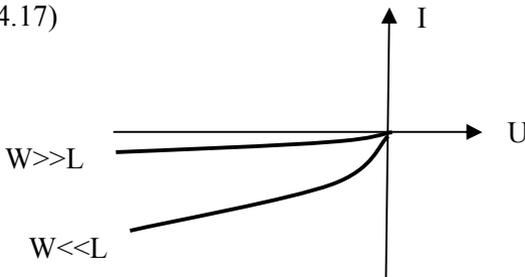


Рисунок 4.17. Обратная ветвь ВАХ p-n-перехода для диодов с толстой и тонкой базой

Это связано с тем, что при увеличении обратного напряжения увеличивается ширина ОПЗ и соответственно уменьшается ширина базы W , что в соответствии с (4.50) вызывает увеличение обратного тока.

Фактически в диодах с тонкой базой участка насыщения на обратной ветви ВАХ не наблюдается (рис.4.17).

На рис.4.18 представлены графики распределения концентрации неосновных носителей в диодах с толстой и тонкой базами, поясняющие вид ВАХ таких диодов (4.17).

Поскольку обратный ток является диффузионным по своей природе, то он определяется градиентом концентрации неосновных носителей. В случае диода с толстой базой градиент концентрации (касательная к распределению $P_n(x)$) не изменяется при увеличении напряжения.

Для диодов с тонкой базой увеличение напряжения вызывает возрастание градиента концентрации, т.к. граница омического контакта неподвижна.

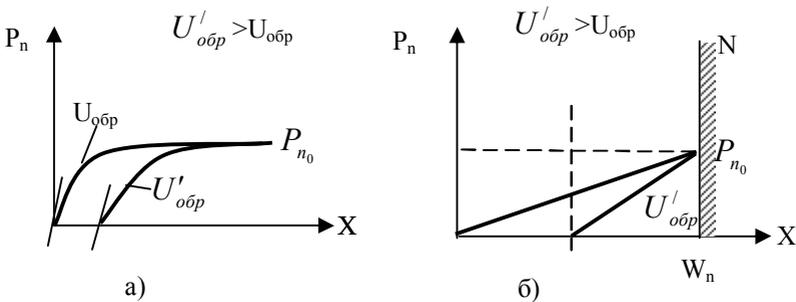


Рисунок 4.18. Распределение концентрации неосновных носителей в диоде с толстой (а) и тонкой (б) базами при увеличении обратного напряжения

При этом рост напряжения увеличивает ширину ОПЗ и уменьшает толщину базы, создавая рост градиента концентрации.

4.9. Вольт - амперная характеристика реального электронно-дырочного перехода.

Вывод ВАХ, приведенной в предыдущем разделе, проведен для случая, когда токи в $p - n$ переходе определяются явлениями инжекции и экстракции. В реальных условиях работы полупроводниковых приборов на величину тока через $p - n$ переход оказывают влияние явление генерации носителей в ОПЗ, рекомбинации, явления на поверхности полупроводника, влажность, загрязнения поверхности и т.д. Рассмотрим влияние указанных факторов на ВАХ реального $p - n$ перехода.

Рассмотрение проведем для прямой и обратной ветви ВАХ (рис.4.19) Весь участок прямой ветви ВАХ можно разделить на три участка: область малых напряжений (I), область средних напряжений (II) и область больших токов (III).

В области малых прямых напряжений, когда уменьшение барьера φ_k^0 мало, вблизи $p - n$ перехода скапливается большое количество разноименных зарядов и возможен процесс их рекомбинации (рис.4.20).

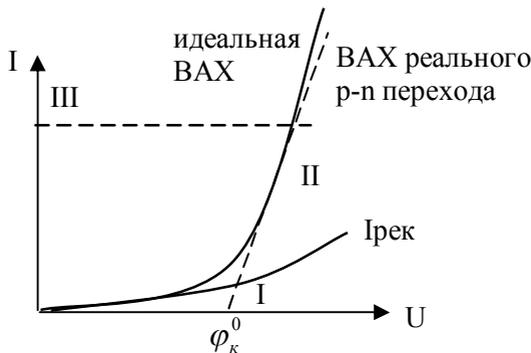


Рисунок 4.19. ВАХ идеального и реального $p - n$ перехода

Составляющую тока, связанную с процессом рекомбинации носителей в $p-n$ переходе называют рекомбинационным током. При больших напряжениях процесс рекомбинации также имеет место, но при этом ток, обусловленный инжекцией неосновных носителей заряда, значительно больше рекомбинационного.

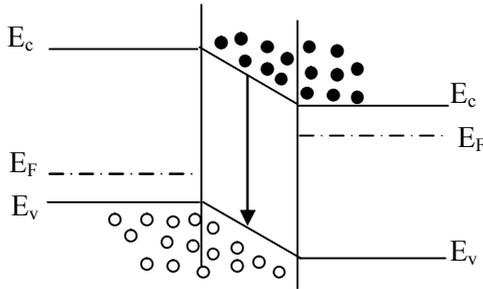


Рисунок 4.20. Процесс рекомбинации носителей в области малых напряжений

Поэтому рекомбинационный ток играет существенную роль только при малых напряжениях (рис.4.19).

Плотность рекомбинационного тока в симметричном $p-n$ переходе определяется выражением (4.51)

$$j_{pek} = \frac{qn_i d}{\tau} \cdot \frac{kT}{q(\varphi_k^0 - U)} \left[\exp \frac{qU}{2kT} - \exp \left(-\frac{qU}{2kT} \right) \right] \quad (4.51)$$

Наиболее важное значение рекомбинационный ток имеет для $p-n$ переходов на основе полупроводников с большой шириной запрещенной зоны, поскольку такие переходы имеют большое значение φ_k^0 , и инжекция носителей в области малых напряжений затруднена.

В области средних напряжений (участок II) ток в реальном $p-n$ переходе практически полностью определяется инжекцией и совпадает с расчетным.

На участке больших токов (участок III, рис.4.19) наблюдаются два отклонения реальной ВАХ от идеальной: значения токов в реальном $p-n$ переходе всегда меньше расчетных по формуле (4.40) и ток зависит не по экспоненте, а становится линейной функцией напряжения.

Первое обусловлено тем, что в реальном $p-n$ переходе часть прикладываемого напряжения падает на базе диода (рис. 4.21) и уменьшает фактическое напряжение, падающее на $p-n$ переходе, которое определяет величину тока.

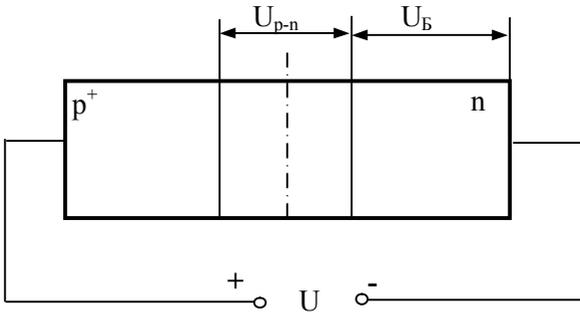


Рисунок 4.21. Падение напряжения в реальном $p-n$ переходе

Отклонение от экспоненциальной зависимости наблюдается, когда прикладываемое напряжение становится соизмеримым с величиной φ_k^0 и потенциальный барьер исчезает.

Обратная ветвь.

Обратный ток в реальном электронно-дырочном переходе состоит не только из тока насыщения, обусловленного экстракцией

неосновных носителей (рис.4.22). Можно записать, что

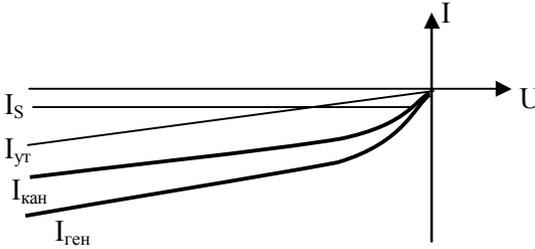
$$I_{обр} = I_S + I_{yt} + I_{кан} + I_{ген},$$

где: I_S - ток насыщения, обусловленный экстракцией неосновных носителей;

$I_{ут}$ - ток утечки по поверхности полупроводника;

$I_{кан}$ - канальный ток;

$I_{ген}$ - ток тепловой генерации носителей в ОПЗ $p-n$ перехода.



Ток утечки возникает из-за загрязнения поверхности полупроводника и может существенно влиять на обратную ветвь ВАХ при достаточно больших обратных напряжениях.

Учитывая, что абсолютное большинство приборов корпусировано, либо поверхность полупроводника защищена другим способом, ток утечки может быть сведен к минимуму.

Величина тока утечки зависит от степени загрязнения и по линейному закону зависит от напряжения.

Канальный ток возникает в результате адсорбции электроположительных или электроотрицательных частиц на поверхности полупроводника. При адсорбции заряженных частиц возможна инверсия поверхностной проводимости. Так, например, при адсорбции электроположительных частиц в слое p -типа на поверхности образуется канал с n -типом проводимости. Возникающий канал шунтирует $p-n$ переход увеличивая обратный ток. Этот ток также может быть сведен к минимуму путем герметизации в газонаполненный корпус. Кроме возникновения канала проводимости на величину обратного тока оказывают влияние процессы генерации и рекомбинации носителей на поверхности. Процессы генерации увеличивают обратные токи. Поскольку поверхность полупроводника - это сильно нарушенная кристаллическая решетка, то процессы рекомбинации на поверхности протекают наиболее интенсивно, вызывая увеличение обратного тока.

Ток тепловой генерации. Процессы тепловой генерации носителей происходят и в слоях полупроводника, прилегающих к ОПЗ р-п- перехода, и в ОПЗ. Причем в ОПЗ эти процессы протекают наиболее интенсивно. Образующиеся при генерации электроны и дырки (рис.4.23) “растаскиваются” электрическим полем перехода в разные стороны, создавая ток $I_{ген}$. Величина тока тепловой генерации определяется выражением

$$I_{ген} = S \frac{qn_i d}{\tau}, \quad (4.52)$$

где S - площадь р-п-перехода.

Так как ширина р-п-перехода увеличивается с ростом обратного напряжения, то $I_{ген}$ зависит от обратного напряжения, увеличиваясь с ростом последнего пропорционально \sqrt{U} .

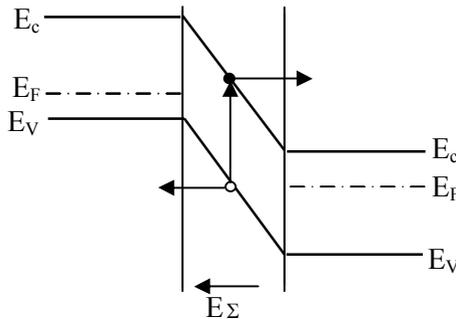


Рисунок 4.23. Иллюстрация возникновения генерационного тока

Таким образом, обратный ток р-п-перехода складывается как минимум из двух компонент (при условии, что токи $I_{ут}$ и $I_{кан}$ уменьшены до нуля): тока насыщения I_S и тока тепловой генерации $I_{ген}$. Конкретно, величина обратного тока определяется либо током I_S , либо током $I_{ген}$. Так в приборах на основе кремния основная компонента обратного тока - это ток тепловой генерации, тогда как в германиевых приборах - ток насыщения. Это связано с различием в величине тока насыще-

ния: у германиевых приборов ток насыщения значительно больше чем у кремниевых приборов.

4.10. Явления в электронно-дырочном переходе при высоком уровне инжекции.

Прежде всего введем понятие высокого уровня инжекции. Если граничная концентрация инжектированных носителей p_n в несимметричном $p^+ - n$ переходе мала по сравнению с концентрацией основных носителей n_{n0} ($p_n \ll n_{n0}$), то это соответствует низкому уровню инжекции (рис.4.24).

Если $p_n \gg n_{n0}$, то это соответствует высокому уровню инжекции. При высоком уровне инжекции в р-п-переходе возникают следующие явления:

1. В базе р-п перехода возникает внутреннее электрическое поле;

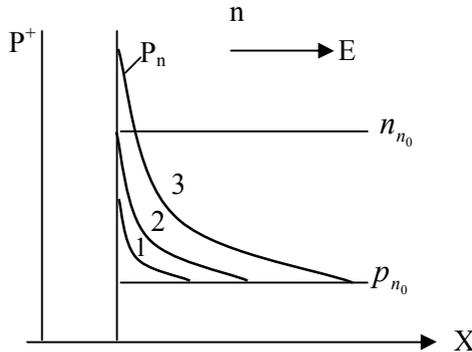


Рисунок 4.24. Распределение концентрации инжектированных носителей в базе диода при низком (1), среднем (2) и высоком (3) уровне инжекции

Увеличивается концентрация основных носителей в базе на границе р-п- перехода;

Уменьшается сопротивление базы;

Изменяется характер протекания тока.

Рассмотрим подробнее каждое из указанных явлений.

Инжекция значительного количества дырок в базу нарушает условие электронейтральности базы. Для того чтобы восстановить электронейтральность электроны начинают подтягиваться к границе р-п-перехода. Поскольку база легирована однородно, то градиент концентрации электронов отсутствует. Тогда необходима сила, которая заставляет электроны перемещаться к границе ОПЗ р-п-перехода, т.е. необходимо возникновение электрического поля E в базе р-п-перехода. Напряженность этого внутреннего поля можно определить, полагая равным нулю ток электронов в $p^+ - n$ переходе. Тогда из уравнения тока

$$j_n = q\mu_n n_n E + qD_n \frac{dn_n}{dx} = 0 \quad \text{можно записать, заменяя } \mu_n \text{ на}$$

$$\frac{q}{kT} D_n \text{ и сокращая } D_n,$$

$$E = -\frac{KT}{q} \cdot \frac{1}{n_n} \cdot \frac{dn_n}{dx}$$

(4.53)

Необходимо отметить, что условие электронейтральности в базе выполняется, если $p_n - p_{no} \approx n_n - n_{no}$ или, с учетом того, что $p_n \gg p_{no}$

$$p_n \approx n_{no} - N_q$$

(4.54)

Причем надо иметь в виду, что равенство (4.54) является приближительным, так как для поддержания неравномерного распределения основных носителей в базе необходимо электрическое поле, которое и возникает из-за неполной компенсации заряда неосновных носителей вблизи р-п-перехода. Кроме электрического поля, возникающего из-за неполной компенсации, существует еще одна компонента, связанная с падением напряжения на базе диода при протекании тока. Выражение напряженности, определенное по (4.53) включает обе составляющие и с учетом условия электронейтральности (4.54) принимает вид (4.55)

$$E = -\frac{KT}{q} \cdot \frac{1}{p_n + N_q} \cdot \frac{dp_n}{dx} \quad (4.55)$$

Эффект уменьшения сопротивления базы при высоком уровне инжекции очевиден: при высоком уровне инжекции концентрация неосновных носителей, особенно в $p-n$ переходах с тонкой базой, высока и сопротивление слаболегированного слоя уменьшается. Кроме этого надо отметить, что высокий уровень инжекции вызывает изменение времени жизни и подвижности носителей, и условий на омическом контакте.

Наконец, при высоком уровне инжекции изменяется характер протекания тока: кроме диффузионной составляющей появляется дрейфовая составляющая, обусловленная наличием внутреннего поля в базе. Запишем выражения для ВАХ $p^+ - n$ перехода для случая, когда $W \ll L$, при высоком уровне инжекции. Так как в $p^+ - n$ переходе электронным током можно пренебречь, тогда ток, обусловленный диффузией и дрейфом дырок равен:

$$j = j_p = q\mu_p p_n E - qD_p \frac{dp_n}{dx}, \quad (4.56)$$

Заменяя μ_p на $\frac{q}{KT} D_p$ и подставляя выражение для E (4.55) получим

$$j_p = -qD_p \left(1 + \frac{p_n}{p_n + N_q} \right) \frac{dp_n}{dx}, \quad (4.57)$$

С учетом соотношения $p_n \gg N_q$ (n_{no}) уравнение (4.57) принимает вид

$$j_p = -q2D_p \frac{dp_n}{dx}, \quad (4.58)$$

Это уравнение напоминает вид уравнения диффузионного тока для дырок, в котором коэффициент диффузии удвоен за счет наличия

электрического поля. Поскольку $p_n \approx n_n - N_q$, то $\frac{dp_n}{dx} \approx \frac{dn_n}{dx}$ и уравнение (4.58) можно переписать в виде

$$j_p = -q2D_p \frac{dn_n}{dx} \quad (4.59)$$

Решение уравнения (4.58) известно и имеет выражение, записанное по аналогии с уравнением (4.49)

$$j = j_p = \frac{q2D_p p_{n0}}{W_n} \left(\exp \frac{qU_{p-n}}{kT} - 1 \right), \quad (4.60)$$

В данном случае в выражении, стоящем под экспонентой, стоит U_{p-n} - падение напряжения на ОПЗ р-п-перехода, поскольку при высоком уровне инжекции (большие токи) значительная часть напряжения падает на базе р-п- перехода. А величину тока через р-п переход определяет падение напряжения на самом переходе U_{p-n} . Для того чтобы его определить, необходимо исходить из того, что $U = U_{p-n} + U_B$, и для нахождения U_{p-n} надо рассчитать падение напряжения на базе U_B . Оно находится следующим образом

$$U_B = \int_0^{U_B} dU_B = - \int_{W_n}^0 E dx = \frac{kT}{q} \int_{W_n}^0 \frac{1}{h_n} \cdot \frac{dn_n}{dx} dx = \frac{kT}{q} \int_{n_{n0}}^{n_n^{(0)}} \frac{dn_n}{n_n} = \frac{kT}{q} \ln \frac{n_n^{(0)}}{n_n} \quad (4.61)$$

Условия электронейтральности позволяют допустить, что $n_n(0) \approx p_n(0)$ и тогда

$$n_n(0) \approx p_n(0) = p_{n0} \exp \frac{qU_{p-n}}{kT}. \quad (4.62)$$

После подстановки в (4.62) выражения (4.61) получим

$$U_B = \frac{kT}{q} \ln \frac{p_{no}}{n_{no}} \exp \frac{qU_{p-n}}{kT}. \text{ Домножая числитель и знаменатель}$$

на n_{no} и учтя закон действующих масс, получим выражение

$$U_B = \frac{kT}{q} \ln \left(\frac{n_i}{n_{no}} \right)^2 + U_{p-n} = 2 \frac{kT}{q} \ln \frac{n_i}{n_{no}} + 2U_{p-n}.$$

Тогда с учетом того, что $U = U_{p-n} + U_B$, получим

$$U_{p-n} = \frac{U}{2} - \frac{kT}{q} \ln \frac{n_i}{n_{no}} = \frac{U}{2} + \frac{kT}{q} \ln \frac{n_{no}}{n_i}, \quad (4.63)$$

Подставляя (4.63) в (4.60) имеем

$$j = \frac{q2D_p p_{no}}{W_n} \cdot \left(\exp \frac{qU}{2kT} \cdot \frac{n_{no}}{n_i} - 1 \right).$$

Пренебрегая единицей в скобках, так как $n_{no} \gg n_i$ и $\exp \frac{qU}{2kT} \gg 1$ запишем окончательное выражение для ВАХ при высоком уровне инжекции.

$$j = \frac{q2D_p n_i}{W_n} \exp \frac{qU}{2kT}, \quad (4.64)$$

Выражение для ВАХ (4.64) напоминает вид ВАХ р-n-перехода с тонкой базой, где вместо коэффициента диффузии D стоит удвоенный коэффициент, а вместо напряжения U в экспоненте записано $\frac{U}{2}$, отражающее факт падения напряжения

в среднем пополам между ОПЗ и базой.

Несколько замечаний по поводу применимости формулы (4.64). Нижняя граница определяется исходя из непрерывности хода ВАХ при переходе от низкого к высокому уровню инжекции. Приравнявая выражения (4.40) и (4.64), получим:

$$\frac{qD_p p_{no}}{W_n} \left(\exp \frac{qU}{2kT} - 1 \right) = \frac{q2D_p n_i}{W_n} \exp \frac{qU}{2kT},$$

(4.65)

Пренебрегая в области больших токов единицей в левой части, запишем $\frac{2n_i}{p_{no}} = \exp \frac{qU}{2kT}$. Подставляя этот результат в любую

часть выражения (4.65) получим значение нижней границы тока, выше которого надо пользоваться формулой (4.64).

$$j_H = \frac{q2D_p n_i \cdot 2n_i}{W_n \cdot p_{no}} = \frac{4qD_p n_i^2}{W_n \cdot p_{no}} = \frac{4qD_p N_a}{W_n},$$

(4.66)

При подстановке типичных параметров в формулу (4.66) получаем, что $j_H \approx 0,1 \frac{A}{\text{см}^2}$.

Верхнюю границу применимости формулы (4.64) можно получить исходя из неравенства $U_{p-n} < \varphi_k^0$, которое можно переписать, используя (4.63) и (4.3)

$$\frac{U}{2} - \frac{kT}{q} \ln \frac{n_i}{n_{no}} < \frac{kT}{q} \ln \frac{n_{no} p_{po}}{n_i^2} \quad \text{или}$$

$$\frac{U}{2} < \frac{kT}{q} \ln \frac{p_{po}}{n_i}$$

и тогда, подставляя $\frac{U}{2}$ в выражение (4.64)

$$j_B < \frac{2qD_p n_i}{W_n} \cdot \frac{p_{po}}{n_i} < \frac{2qD_p p_{po}}{W_n} \quad \text{или} \quad j_B = \frac{2qD_p N_a}{W_n},$$

где N_a - концентрация акцепторов в эмиттере. Порядок величины $j_B \approx 100 A / \text{см}^2$.

Таким образом, можно говорить о том, что практически во всем диапазоне рабочих напряжений р-п-переход работает в условиях высокого уровня инжекции.

В р-п-переходах с толстой базой, когда $W \gg L$, сопротивление базы можно считать не зависящим от тока, т.е. пренебрегать модуляцией сопротивления базы. Тогда ВАХ р-п-перехода может быть представлена в виде:

$$I = I_s \left[\exp \frac{q(U - IR_\sigma)}{kT} - 1 \right],$$

где R_σ - сопротивление базы р-п-перехода

Если сопротивление R_σ известно, то рассчитать ВАХ можно, принимая в качестве независимой переменной значение тока I .

При больших токах, когда напряжение на переходе приближа-

ется к контактной разности потенциалов φ_k^0 и как бы стабилизи-

руется, ВАХ р-п-перехода с толстой базой можно записать так

$$U = \varphi_k^0 + IR_\sigma.$$

Это говорит о том, что при $U \approx \varphi_k^0$ ВАХ становится близкой к линейной, а не экспоненциальной, поскольку экспоненциальный характер ВАХ связан с экспоненциальной зависимостью концентрации инжектированных носителей от напряжения.

4.11. Диффузионная и барьерная ёмкости ЭДП

При подключении р-п-перехода к переменному источнику напряжения через него течет емкостной ток, что свидетельствует об емкостных свойствах р-п-перехода. Для описания емкостных свойств перехода вводится понятие диффузионной и барьерной емкостей.

Барьерная емкость - это емкость плоского конденсатора, мнимыми обкладками которого являются изменяющиеся границы ОПЗ, а функцию диэлектрика выполняет сама ОПЗ (рис. 4.25).

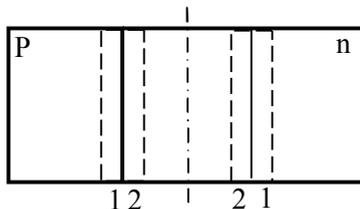


Рисунок 4.25. К понятию барьерной емкости р-п-перехода

Барьерная емкость есть проявление токов смещения в $p-n$ переходе. Она определяется той долей тока, которая не связана с движением носителей заряда через $p-n$ переход. Действительно при приложении, например, прямого смещения границы ОПЗ смещаются так, что ширина его уменьшается, а при обратном напряжении - увеличивается.

При этом границы из плоскости 1 (соответствующие $U = 0$) смещаются в положения, соответствующие плоскости 2, (соответствующие $U > 0$), за очень короткий промежуток времени (10^{-14} - 10^{-13} с). Смещение носителей и вызывает появление всплеска тока, который по своей природе является током смещения. Эти изменяющиеся границы ОПЗ принимаются за мнимые обкладки плоского конденсатора. Поскольку в ОПЗ подвижных носителей в идеальном случае нет, то ОПЗ можно считать диэлектрическим.

Значение барьерной емкости $C_{\bar{\sigma}}$ таким образом можно считать по формуле емкости плоского конденсатора

$$C_{\bar{\sigma}} = \frac{\varepsilon\varepsilon_0 S}{d}, \quad (4.67)$$

где S - площадь $p-n$ перехода.

Подставляя в (4.67) выражение для ширины ОПЗ d , получим выражение для расчета $C_{\bar{\sigma}}$ для резкого симметричного перехода

$$C_{\bar{\sigma}} = S \sqrt{\frac{q\varepsilon\varepsilon_0 N_a N_d}{2(N_a + N_d) \cdot (\varphi_k^0 \pm U)}}, \quad (4.68)$$

Для резкого несимметричного $p-n$ перехода

$$C_{\sigma} = S \sqrt{\frac{q \varepsilon \varepsilon_0 N_{\sigma}}{2(\varphi_k^0 \pm U)}}, \quad (4.69)$$

где N_{σ} - концентрация примеси в базе р-п-перехода.

Для плавного $p-n$ перехода с линейным распределением примеси

$$C_{\sigma} = S \sqrt{\frac{qa(\varepsilon \varepsilon_0)^2}{12(\varphi_k^0 \pm U)}}, \quad (4.70)$$

Зависимость емкости от напряжения называется вольт-фардовой характеристикой (ВФХ) (рис. 4.26).

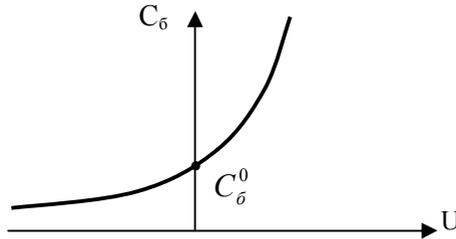
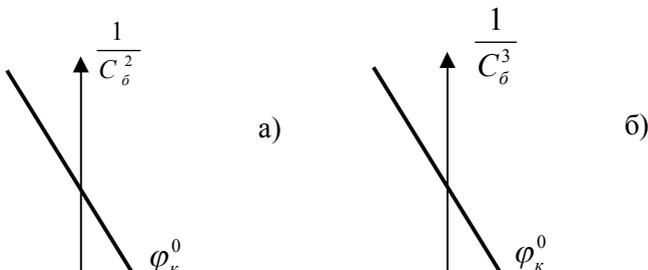


Рисунок 4.26. Вольт-фардовая характеристика р-п перехода

Барьерная емкость проявляется и при прямом, и при обратном напряжениях. Типичные значения барьерной емкости составляют десятки -сотни пикофард для плоскостных р-п-переход и менее десяти пикофард для точечных переходов.

Перестроенные в координатах $\frac{1}{C_{\sigma}^2} = f(U)$ и $\frac{1}{C_{\sigma}^3} = f(U)$

ВФХ позволяют определить по углу наклона прямых концентрацию примеси в базе резкого перехода или градиент концентрации в плавном переходе, а также значение φ_k^0 (рис. 4.27).



Более сложной для восприятия является диффузионная емкость. Для её понимания рассмотрим распределение концентрации инжектированных в n - базу носителей в $p^+ - n$ переходе (рис.4.28).

При $U = 0$ $p_n = p_{n0}$

При увеличении напряжения до U_1 и U_2 возрастает граничная концентрация $p_n^1(0)$ и $p_n^2(0)$, и общая концентрация носителей в базе, т.е. происходит изменение заряда в базе $p - n$ перехода при изменении напряжения на переходе. Можно записать, что

$$\Delta Q = C_D \Delta U,$$

где ΔQ - изменение заряда инжектированных дырок в базе.

ΔU - изменение напряжения на переходе.

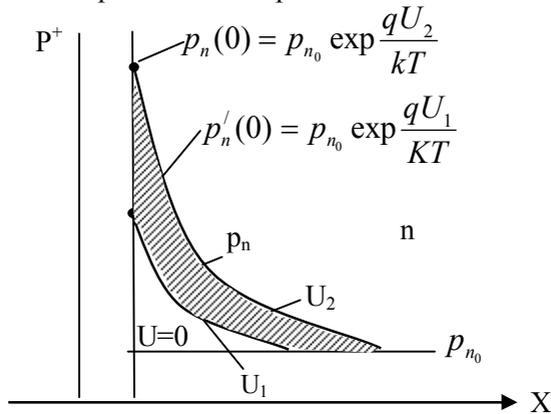


Рисунок 4.28. К понятию диффузионной емкости $p - n$ перехода

Тогда коэффициент пропорциональности между ΔQ и ΔU , и есть диффузионная емкость, поскольку в любом случае емкость есть $C = \frac{dQ}{dU}$.

Можно дать следующее определение диффузионной емкости.

Диффузионная емкость - это ёмкость, отражающая изменения заряда инжектированных носителей в базе при изменении напряжения на переходе. Причём диффузионная емкость не связана с протеканием токов смещения, что существенно отличает её от барьерной емкости. Можно говорить о том, что наличие инжектированного заряда в базе $p-n$ перехода ассоциируется с диффузионной емкостью. Поскольку при обратном смещении явление инжекции отсутствует, то это означает, что диффузионная емкость имеет место только при прямых смещениях $p-n$ перехода. Численные значения диффузионной емкости составляют сотни - тысячи пикофарад. Говоря об емкостных свойствах диода, необходимо иметь в виду, что они определяются барьерной емкостью при обратных напряжениях и диффузионной емкостью при прямых смещениях. В силу того, что абсолютное значение $C_D \gg C_\delta$, поэтому при прямых смещениях главную роль играет диффузионная емкость.

Эквивалентная схема ЭДП. Параметры эквивалентной схемы.

Электрическая эквивалентная схема $p-n$ перехода понятна из рис.4.29.

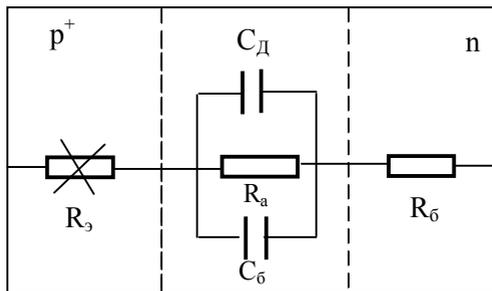


Рисунок 4.29. К эквивалентной схеме $p-n$ перехода

Для протекания тока через переход сопротивления создают область эмиттерного полупроводника, ОПЗ перехода и база. Так как эмиттер сильно легирован, то падение напряжения на нем можно не учитывать.

Область пространственного заряда представлена емкостями C_D , C_σ и сопротивлением R_a , а падение напряжения на базе отражается включением в эквивалентную схему резистора R_σ (рис.4.30).

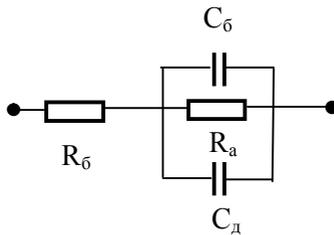


Рисунок 4.30. Эквивалентная схема p-n перехода

R_σ - омическое сопротивление базы. Оно зависит от удельного сопротивления полупроводника и геометрических размеров перехода. Для простейшей конструкции плоскостного p-n перехода определяется как $R_\sigma = \rho_\sigma \frac{W}{S}$.

Для точечного p-n перехода $R_\sigma = \frac{\rho_\sigma}{2\pi a}$, где a - радиус контакта.

Значение сопротивление базы может лежать в пределах долей Ома ÷ десятка Ом .

Барьерная емкость в $p-n$ переходе рассчитывается по формулам (4.68 - 4.70).

Диффузионная емкость может быть рассчитана по формулам

$$C_D = \frac{q}{kT} (I + I_S) \frac{W^2}{2D_\delta} \quad \text{- для диода с тонкой базой,} \quad (4.71)$$

$$C_D = \frac{q}{2kT} (I + I_S) \tau_\delta \quad \text{- для диода с толстой базой,} \quad (4.72)$$

где τ_δ , D_δ - время жизни и коэффициент диффузии неосновных носителей в базе.

Происхождение формул (4.71) и (4.72) поясняется в разделе 4.13.

R_a - дифференциальное сопротивление р-п перехода, т.е. сопротивление перехода малому переменному сигналу при постоянном смещении. Определяется R_a следующим образом:

$$R_a = \frac{dU}{dI} = \frac{kT}{q(I + I_S)}.$$

Величина дифференциального сопротивления велика при обратном смещении и составляет от десятых долей до сотен Ом при прямом смещении.

Кроме того, для характеристики диода вводится понятие сопротивления постоянному току R_0 , определяемое как

$$R_0 = \frac{U}{I}.$$

В виде, представленном на рис.4.30 эквивалентная схема используется редко. Для обратных и прямых напряжений эквивалентные схемы представлены на рис.4.31.

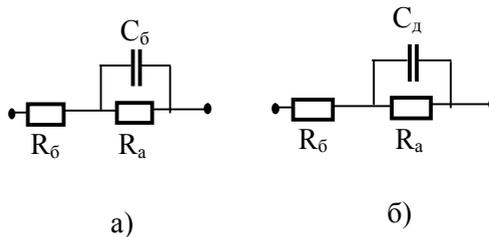


Рисунок 4.31. Эквивалентные схемы р – п перехода для обратного (а) и прямого смещения (б)

4.13. Полная проводимость $p - n$ перехода. Зависимость параметров от частоты.

В связи с аналогией дифференциальных уравнений для распределения постоянной и переменной составляющих концентрации неосновных носителей (4.31) и (4.33), и граничных условий (4.32) и (4.34) можно сразу по аналогии с постоянной составляющей записать выражение для переменного тока через $p - n$ переход. Для этого необходимо сделать следующие замены: длина диффузионного смещения L является комплексным параметром и представляется в виде

$$\dot{L} = \frac{2}{\sqrt{1 + j\omega\tau}}, \quad (4.73)$$

и выражение $\left(\exp \frac{qU}{kT} - 1 \right)$ заменяется на $\left(\exp \frac{qU}{kT} \right) \frac{q\dot{U}_m}{kT}$.

В результате получим выражение для переменного тока

$$\dot{I}_m = \frac{Sq\dot{U}_m}{kT} \left(\frac{qD_p p_{no}}{\dot{L}_p} \operatorname{cth} \frac{W_n}{\dot{L}_p} + \frac{qD_n n_{po}}{\dot{L}_p} \operatorname{cth} \frac{N_p}{\dot{L}_p} \right) \exp \frac{qU}{kT} \quad (4.74)$$

Из выражения (4.74) видно, что переменный ток в переходе связан с переменным напряжением линейным законом, если это напряжение мало. Отсюда следует, что для описания свойств $p - n$ перехо-

да можно воспользоваться обычным приемом электротехники -

ввести полную проводимость для переменного тока: $\dot{Y} = \frac{\dot{I}_m}{\dot{U}_m}$

$$\dot{Y} = \frac{\dot{I}_m}{\dot{U}_m} = \frac{Sq}{kT} \left(\frac{qD_p p_{n0}}{\dot{L}_p} \operatorname{cth} \frac{W_n}{\dot{L}_p} + \frac{qD_n n_{p0}}{\dot{L}_n} \operatorname{cth} \frac{W_p}{\dot{L}_n} \right) \exp \frac{qU}{kT}$$

(4.75)

Выражение для полной проводимости $p-n$ перехода при параллельной схеме замещения можно записать в виде

$$\dot{Y} = \frac{1}{R_a} + j\omega C_D,$$

(4.76)

Анализ выражения (4.75) в общем случае достаточно сложен, поэтому обычно ограничивается рассмотрением диодов с толстой или тонкой базами для области низких и высоких частот.

$W \gg L$, низкие частоты.

При $W \gg L$ $\operatorname{cth} \left(\frac{W}{L} \sqrt{1 + j\omega\tau} \right) \approx 1$ и длину диффузии можно

для низких частот ($\omega\tau \ll 1$) представить в виде разложения в биномиальный ряд выражения $\sqrt{1 + j\omega\tau}$. Тогда

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

Подставляя выражение для \dot{L} в (4.75) и проведя группирование слагаемых получим

$$\dot{Y} = \frac{Sq}{kT} \left[\left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) + \frac{j\omega}{2} (q p_{n0} L_p + q n_{p0} L_n) \right] \exp \frac{qU}{kT}$$

(4.77)

В соответствии с (4.76)

$$\frac{1}{R_a} = \frac{qS}{kT} \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) \exp \frac{qU}{kT} \quad \text{или} \quad R_a = \frac{kT}{q(I + I_s)},$$

поскольку выражение в скобках – это ток насыщения в диоде с толстой базой. Тогда, учитывая, что $I = I_s \left(\exp \frac{qU}{kT} - 1 \right)$ можно

представить как $I + I_S = I_S \exp \frac{qU}{kT}$. Таким образом в диоде с толстой базой в области низких частот сопротивление перехода переменному току определяется как $R_a = \frac{kT}{q(I + I_S)}$ и не зависит от частоты переменного сигнала.

Выражение в скобках для мнимой части уравнения полной проводимости – это тоже ток насыщения в диоде с толстой базой, поскольку выражение $qp_{n_0}L_p + qn_{p_0}L_n$ можно преобразовать следующим образом

$$\frac{qp_{n_0}L_pL_p}{L_p} + \frac{qn_{p_0}L_nL_n}{L_n} = \frac{qp_{n_0}D_p\tau_p}{L_p} + \frac{qD_n n_p \tau_n}{L_n}.$$

Это уравнение при условии $\tau_p = \tau_n = \tau$ преобразуется к виду

$$\tau \left(\frac{qD_p p_{n_0}}{L_p} + \frac{qD_n n_{p_0}}{L_n} \right). \text{ Выражение в скобках – ток насыщения.}$$

Тогда в соответствии с (4.76)

$$C_D = \frac{q}{2kT} (I + I_S) \tau, \quad (4.78)$$

Таким образом, диффузионная емкость для диода с толстой базой в области низких частот от частоты не зависит и определяется током через диод и временем жизни носителей в базе.

2. $W \gg L$, высокие частоты. В этом случае критерий высоких частот $\omega\tau \gg 1$ и $\operatorname{cth}\left(\frac{W}{L}\sqrt{1+j\omega\tau}\right) \approx 1$. Для того, чтобы преобразовать выражение (4.73) удобно воспользоваться формулами Муавра:

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

Тогда $\frac{1}{\dot{L}} \approx \frac{1}{L} \sqrt{\frac{\omega\tau}{2}} + j \frac{1}{L} \sqrt{\frac{\omega\tau}{2}}$. Подставляя это значение \dot{L} в выражение (4.75) и разделяя действительную и мнимую части, получим:

$$\dot{Y} = \frac{qS}{kT} \left[\begin{array}{l} \sqrt{\frac{\omega}{2}} \left(\frac{qD_p p_{n0} \sqrt{\tau_p}}{L_p} + \frac{qD_n n_p \sqrt{\tau_n}}{L_n} \right) + \\ j \sqrt{\frac{\omega}{2}} \left(\frac{qD_p p_{p0} \sqrt{\tau_p}}{L_p} + \frac{qD_n n_{p0} \sqrt{\tau_n}}{L_n} \right) \end{array} \right] \exp \frac{qU}{kT} \quad (4.79)$$

Из (4.79) с учетом (4.76) можно записать R_a и C_D при условии, что $\tau_n = \tau_p = \tau$ или $p_{n0} \gg n_{p0}$, либо $n_{p0} \gg p_{n0}$,

$$\left. \begin{array}{l} R_a = \frac{kT}{q(I + I_S)} \sqrt{\frac{2}{\omega\tau}} \\ C_D = \frac{q}{kT} (I + I_S) \sqrt{\frac{\tau}{2\omega}} \end{array} \right\} \quad (4.80)$$

Видно, что в диоде с толстой базой в области высоких частот и R_a и C_D зависят от частоты. Выражения (4.80) можно переписать в следующем виде:

$$\left. \begin{array}{l} R_a^{BЧ} = R_a^{HЧ} \cdot \sqrt{\frac{2}{\omega\tau}} \\ C_D^{BЧ} = C_D^{HЧ} \sqrt{\frac{2}{\omega\tau}} \end{array} \right\} \quad (4.81)$$

3. $W \ll L$, низкие частоты. Для области низких частот критерий низких частот $\left| \frac{W}{L} \sqrt{1 + j\omega\tau} \right| \ll 1$.

Разлагая в ряд гиперболический котангенс, и беря два первых члена разложения, имеем

$$cth \frac{W}{L} \approx \frac{L}{W} + \frac{W}{3L} = \frac{L}{W} \left(1 + \frac{W^2}{3L^2} \right) \text{ или}$$

$$cth \frac{W}{L} = \frac{L}{W} \left(1 + \frac{W^2}{3L^2} + j\omega\tau \frac{W^2}{3L^2} \right) \approx \frac{L}{W} \left(1 + j\omega\tau \frac{W^2}{3L^2} \right),$$

(4.82)

В полученном выражении членом $\frac{W^2}{3L^2}$ пренебрегаем, поскольку

он при $W \ll L$ мал. Третий член разложения $j\omega\tau \frac{W^2}{3L^2}$ так же мал, но его нельзя исключить, поскольку при этом исчезнет мнимая часть. Подставляя (4.82) в (4.75) и разделяя действительную, и мнимую части, получим:

$$\dot{Y} = \frac{Sq}{kT} \left[\left(\frac{qD_p p_{no}}{W_n} + \frac{qD_n n_{po}}{W_p} \right) + j \frac{\omega}{3} (q p_{no} W_n + q n_{po} W_p) \right] \exp \frac{qU}{kT}$$

(4.83)

Из (4.83) запишем:

$$R_a = \frac{kT}{q(I + I_s)} \text{ и } C_D = \frac{q}{kT} (I + I_s) \frac{W^2}{3D},$$

(4.84)

т.е. для диода с тонкой базой в области низких частот R_a и C_D от частоты не зависят. При этом надо иметь виду, что диффузионная емкость в данном случае определяется не временем жизни носителей в базе, а величиной $\frac{W^2}{3D}$, которая, как будет показано далее, определяет время пролета носителей через базу.

4. $W \ll L$, высокие частоты. Аналогично диоду с толстой базой можно показать, что для диода с тонкой базой в области высоких частот выражения для R_a и C_D имеют следующий вид:

$$R_a^{BЧ} = R_a^{HЧ} \cdot \sqrt{\frac{2}{\omega\tau}} \quad (4.85)$$

$$C_D^{BЧ} = C_D^{HЧ} \cdot \sqrt{\frac{2}{\omega\tau}}$$

т.е. R_a и C_D уменьшаются с частотой. На рис.4.32 представлены графики зависимости R_a и C_D от частоты.

Из рисунков видно, что диффузионная емкость в диоде с толстой базой больше чем с тонкой и переход к области высоких частот, когда R_a и C_D начинают зависеть от частоты, выше для диодов с тонкой базой.

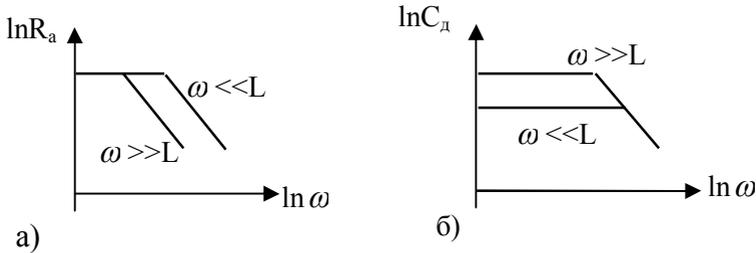


Рисунок 4.32. Зависимость дифференциального сопротивления (а) и диффузионной емкости (б) от частоты

Оценим время пролета инжектированных носителей в базе, на примере $p^+ - n$ перехода.

Величина тока дырок может быть определена как $j_p = qp_n v_p = qp_n \frac{dx}{dt}$, откуда $dt = \frac{qp_n}{j_p} dx$ и $t_{прол} = \int_0^{W_n} \frac{qp_n}{j_p} dx$.

С учетом того, что $j_p = \frac{qD_p p_{no}}{W_n} \left(\exp \frac{qU}{kT} - 1 \right)$ и

$p_n \approx p_{no} \left(\exp \frac{qU}{kT} - 1 \right) \left(1 - \frac{x}{W_n} \right)$, имеем

$$t_{прол} = q \int_0^{W_n} \frac{W_n - x}{qD_p} dx = \frac{W_n^2}{2D_p} :$$

В общем случае (для $p^+ - n$ и $n^+ - p$ переходов) можно записать

$$t_{прол} = \frac{W^2}{2D_{\sigma}},$$

где W - толщина базы;

D_{σ} - коэффициент диффузии неосновных носителей в базе.

Проведенный анализ полной проводимости и зависимости параметров от частоты имеет важное практическое значение. Рассмотрим эквивалентную схему $p - n$ перехода для переменного сигнала (рис. 4.33).

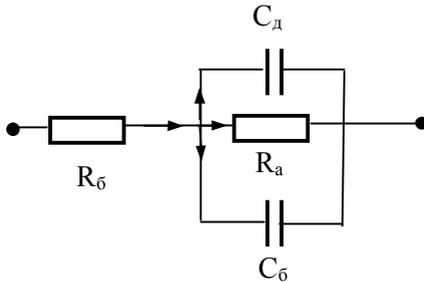


Рисунок 4.33. К пояснению эффекта шунтирования $p - n$ перехода барьерной емкостью.

При увеличении частоты действующего сигнала, в силу зависимости параметров от частоты, будут изменяться сопро-

тивления элементов эквивалентной схемы. Причем $x_{C\bar{\sigma}} \sim \frac{1}{\omega}$;

$$x_{C_D} \sim \frac{1}{\sqrt{\omega}};$$

$$x_{R_a} \sim \frac{1}{\sqrt{\omega}} \text{ т.е. с увеличением частоты сопротивление } x_{C\bar{\sigma}}$$

уменьшается наиболее сильно. Это приводит к тому, что при определенных частотах основная часть тока начнет протекать через барьерную емкость.

Можно говорить о том, что в области высоких частот барьерная емкость шунтирует $p-n$ переход. Частоту, на которой это происходит можно определить, приравняв сопротивление базы сопротивлению барьерной емкости.

$$R_{\bar{\sigma}} = \frac{1}{\omega C_{\bar{\sigma}}}; \quad \omega_{\text{пред}} = \frac{1}{R_{\bar{\sigma}} C_{\bar{\sigma}}} \quad \text{или} \quad f_{\text{пред}} = \frac{1}{2\pi R_{\bar{\sigma}} C_{\bar{\sigma}}} \quad (4.86)$$

Предельную частоту таким образом можно определить как частоту, на которой $p-n$ переход теряет выпрямляющие свойства. Естественно стремление разработчиков повысить этот важный параметр. На первый взгляд вопрос решается достаточно просто: необходимо уменьшить площадь перехода. Однако произведение $R_{\bar{\sigma}}$ и $C_{\bar{\sigma}}$ от площади для плоскостных диодов не зависит. Поэтому наиболее простой путь увеличения $f_{\text{пред}}$ - уменьшение сопротивления $R_{\bar{\sigma}}$ путем использования структуры, представленной на рис.4.34.

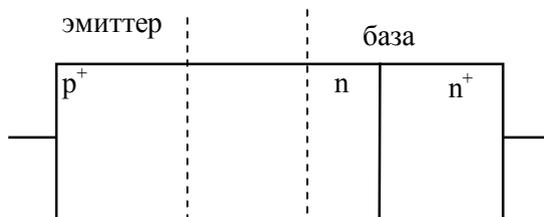


Рисунок 4.34. Структура базы реального диода.

Включение сильно легированного n^+ уменьшает сопротивление R_6 .

Для точечных диодов уменьшение площади $p-n$ перехода позволяет увеличить предельную частоту.

В заключении данного раздела необходимо еще раз подчеркнуть важность рассмотренного явления - шунтирования $p-n$ перехода барьерной емкостью, поскольку именно это явление в значительной степени определяет частотные свойства приборов на основе $p-n$ переходов.

4.14. Переходные процессы в ЭДП

Электронно-дырочный переход и приборы на его основе являются инерционным элементом по отношению к быстрым изменениям тока и напряжения. Связано это с тем, что при изменении напряжения изменяется ширина ОПЗ, т.е. идет процесс зарядки-разрядки барьерной емкости. Изменение напряжения вызывает также изменение зарядов инжектированных носителей, т.е. заряд-разряд диффузионной емкости. Протекание этих процессов наряду с наличием проводимости, характеризуемой выражением (4.42), приводит к тому, что $p-n$ переход имеет эквивалентную схему (рис.4.30), включающую параллельное соединение диффузионной и барьерной емкостей, которые и обуславливают инерционность процессов в $p-n$ переходе при воздействиях изменяющихся во времени сигналов. Причем надо иметь в виду, что при сравнительно больших прямых токах процессы накопления и рассасывания носителей заряда в базе (заряд-разряд C_D) играют более важную роль, чем перезарядка барьерной емкости. При малых плотностях тока переходные процессы определяются перезарядкой барьерной емкости.

Временные зависимости напряжения и тока, характеризующие переходные процессы в $p-n$ переходе, зависят также от сопротивления внешней цепи, в которую включен переход. Принципиально отличающимися можно выделить четыре предельных случая переходных процессов в $p-n$ переходе: про-

цессы при больших напряжениях и токах, когда диод включен в режиме с генератором тока и генератором напряжения и аналогичные процессы для малых напряжений и токов.

Необходимо отметить, что хотя для указанных четырех случаев осциллограммы токов и напряжений существенно отличаются, физические процессы, обуславливающие переходные процессы в $p-n$ переходе, во всех случаях одни и те же: накопление и рассасывание неосновных носителей заряда в базе, заряд-разряд барьерной и диффузионной емкостей. Поскольку наиболее часто диоды работают при воздействии импульсов большой амплитуды, то важно подробно рассмотреть переходные процессы, соответствующие уровню больших напряжений и токов. Процессы при воздействии малых напряжений рассмотрим кратко. Рассмотрим переходные процессы при включении и отключении диода, и переключения диода с прямой полярности на обратную.

Процессы при больших напряжениях и токах.

1. Режим работы в схеме с генератором напряжения.

В режиме с генератором напряжения диод включен так, что выполняя соотношение $R_D \gg R$, где R_D - сопротивление диода, R - полное сопротивление остальной цепи, включая сопротивление источника питания. Соответствующие осциллограммы токов и напряжений представлены на рис.4.35.

При подаче прямого напряжения на переход ток через него устанавливается не сразу, так как с течением времени происходит накопление в базе инжектированных дырок (рассматриваем $p^+ - n$ переход).

Учитывая, что ток через $p-n$ переход определяется градиентом концентрации, то рост тока во времени обуславливается увеличением последнего, что иллюстрирует рис.4.36. Увеличение граничной концентрации носителей свидетельствует об увеличении прямого смещения на $p-n$ переходе. Когда процесс инжекции уравнивается процессом рекомбинации в базе диода возникает установившееся распределение концентрации неос-

новых носителей (t_3) и ток в p-n переходе устанавливается на определенном уровне, равном I_{np} . Напряжение на переходе U_{p-n} так же перестает изменяться.

Падение напряжения на базе, определяемое как $U_B = IR_\sigma$, в начальной момент велико, а затем спадает до стационарного значения.

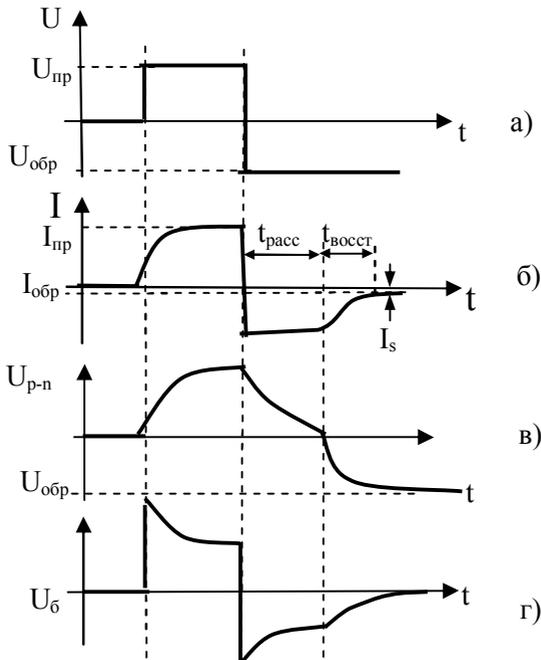


Рисунок 4.35. Изменение во времени питающего напряжения (а), тока через диод (б), падения напряжения на ОПЗ перехода (в) и базе диода (г) при работе на больших импульсах от генератора напряжения

Спад падения напряжения обусловлен уменьшением сопротивления базы R_b , из-за высокой концентрации носителей, возникающей по причине высокого уровня инжекции. Таким образом можно говорить о том, что переходные процессы при включении диода связаны с процессом накопления носителей заряда в базе и зарядом диффузионной емкости. Барьерная емкость при этом также заряжается, но ее влияние проявляется в меньшей мере, так как рассматривается прямое включение и высокий уровень инжекции, когда главную роль играет диффузионная емкость.

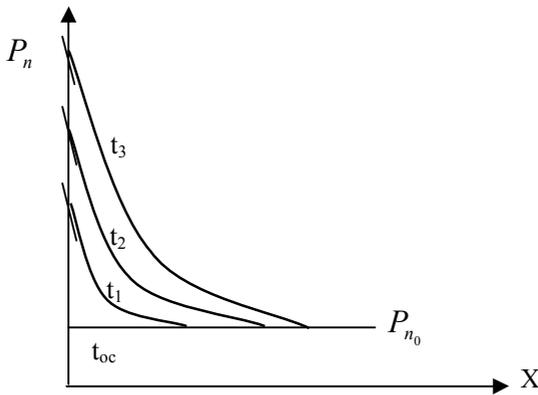


Рисунок 4.36. Распределение концентрации инжектированных дырок в базе p^+ - n перехода в различные промежутки времени

Рассмотрим процесс переключения p - n перехода с прямой полярности на обратную (рис. 4.37). В момент времени $t = t_3$ напряжение скачком меняется с прямого на обратное. Зная, что основное свойство p - n перехода — это явление односторонней проводимости, можно было бы ожидать, что ток через переход должен установиться на уровне малого обратного тока $I_{обр}$. И это действительно будет так, но не сразу, а после того, как произойдут переходные процессы.

В момент времени $t = t_3$ в базе перехода существует стационарное распределение концентрации неосновных носителей, создающих определенный градиент концентрации. При пере-

ключении носители заряда из базы мгновенно исчезнуть не могут, а, следовательно, не исчезнет и градиент концентрации неосновных носителей, задающий величину тока через р-п-переход. Однако направление вектора градиента при этом поменяется на противоположный, что обеспечит протекание тока в обратном направлении (рис.4.37. (t_3)). Поскольку

$p-n$ переход в момент t_3 и далее до t_8 смещен в прямом направлении то он имеет малое сопротивление и величина тока через переход определяется приложенным напряжением и сопротивлением базы диода (внутренне сопротивление генератора напряжения мало).

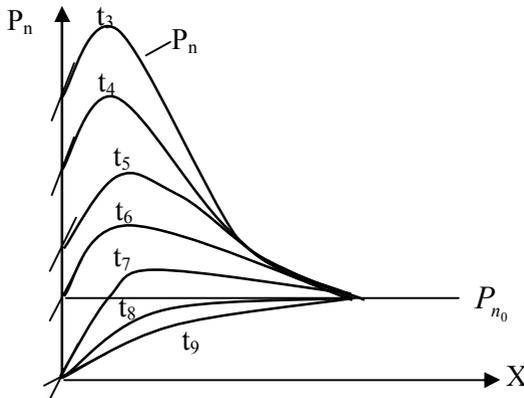


Рисунок 4.37. Распределение концентрации неосновных носителей в базе р-п перехода в различные моменты времени после переключения с прямой полярности на обратную

По мере рассасывания носителей (рис.4.37) напряжение на переходе остается прямым до $t-t_8$. При этом ток через р-п переход в обратном направлении может немного уменьшаться из-за увеличения сопротивления базы в результате уменьшения концентрации носителей в базе. Основным механизмом рассасывания неосновных носителей заряда является процесс их рекомбинации. Процесс рекомбинации определяется временем жизни носителей. Поэтому по промежутку времени t_{pacc} (рис.4.35.б) можно оценивать время жизни неосновных носите-

лей в базе р-п перехода. Процесс рассасывания накопленных носителей происходит значительно медленнее процесса накопления, поэтому процесс рассасывания и определяет частотные свойства диодов на основе р-п перехода.

Ход осциллограммы падения напряжения на базе U_{σ} , понятен с учетом того, что $U_{\sigma} = IR_{\sigma}$. После того как концентрация неосновных носителей на границе р-п перехода станет равна равновесной (рис.4.37. t_6) напряжение на переходе уменьшится до нуля. В дальнейшем начнется увеличение обратного напряжения до $U_{\sigma} = U_{обр}$ за счет перезарядки барьерной емкости. При обратном напряжении на р-п переходе протекает процесс экстракции неосновных носителей (рис.4.37. t_8, t_9). Обратный ток начинает уменьшаться и устанавливается на уровне обратного тока $I_{обр} = I_S$.

Таким образом, можно говорить о том, что переходные процессы при переключении диода с прямой на обратную полярность определяются процессом рассасывания неосновных носителей из базы диода, разрядом диффузионной и перезарядной барьерной емкостей. Длительность переходных процессов при этом контролируется в основном явлением рекомбинации и определяется временем жизни неосновных носителей в базе диода.

Важными параметрами переходных процессов являются время рассасывания и время восстановления (рис.4.35.б). Сумма этих времен дает важнейший параметр импульсных диодов - время восстановления обратного сопротивления.

Время восстановления обратного сопротивления $t_{восст.обр.}$ - это время от момента переключения диода с прямой полярности на обратную до момента, когда через диод установится ток, равный току насыщения.

2. Режим работы диода в схеме с генератором тока.

Режим генератора реализуется, когда последовательно с диодом включен резистор с большим сопротивлением так, что $R \gg R_\delta$. Рассмотрим процесс включения и выключения диода импульсом тока (рис.4.38). В режиме генератора ток в цепи диода определяется сопротивлением внешней цепи.

Падение напряжение на базе в момент подачи импульса тока велико, а затем спадает из-за уменьшения сопротивления базы, связанного с накоплением неосновных носителей в базе (рис. 4.38, б).

Падение напряжения на переходе U_{p-n} возрастает во времени до установившегося значения, поскольку в базе идет процесс накопления неосновных носителей (4.38.в). Ход зависимости $U = f(t)$ понятен, если учесть что $U = U_{p-n} + U_B$, то есть физические процессы в р-п переходе при включении те же самые: накопление неосновных носителей в базе р-п перехода и заряд диффузионной емкости.

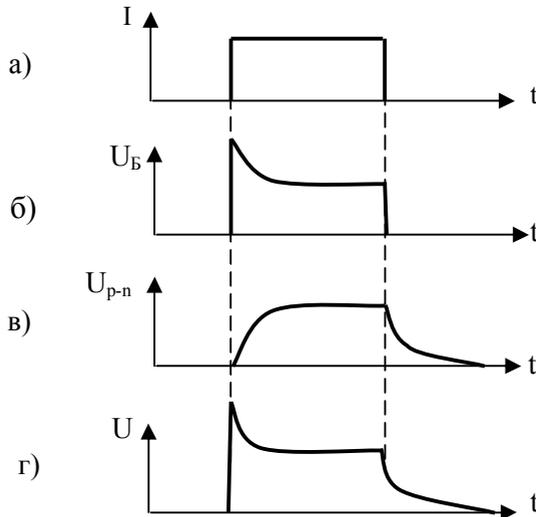


Рисунок 4.38. Осциллограммы тока (а), падения напряжения на базе (б), на ОПЗ р-п перехода (в) и полное падение напряжения (г)

После окончания импульса тока происходит процесс рассасывания неосновных носителей, накопившихся в базе. При этом напряжение на переходе U_{p-n} уменьшается до нуля за время, определяемое временем жизни неосновных носителей.

По длительности спада напряжения можно оценить время жизни неосновных носителей в базе диода.

При воздействии на р-п переход малых напряжений и токов наблюдаются особенности, связанные с тем, что отсутствует эффект модуляции сопротивления базы и главную роль в переходных процессах играет не диффузионная, а барьерная емкость. При этом длительность переходных процессов также определяется явлениями накопления и рассасывания неосновных носителей в базе диода.

4.15. Зависимость выпрямляющих свойств ЭДП от частоты.

Одно из основных свойств р-п перехода - явление односторонней проводимости, позволяющее преобразовать переменный сигнал в постоянный. Рассмотрим как изменяются выпрямляющие свойства перехода при увеличении частоты на примере синусоидального сигнала. Для этого оценим выпрямление в области низких (НЧ), средних (СЧ) и высоких частот (ВЧ). Критерием деления на указанные области частот является соотношение между временем жизни τ и временем воздействия напряжения на

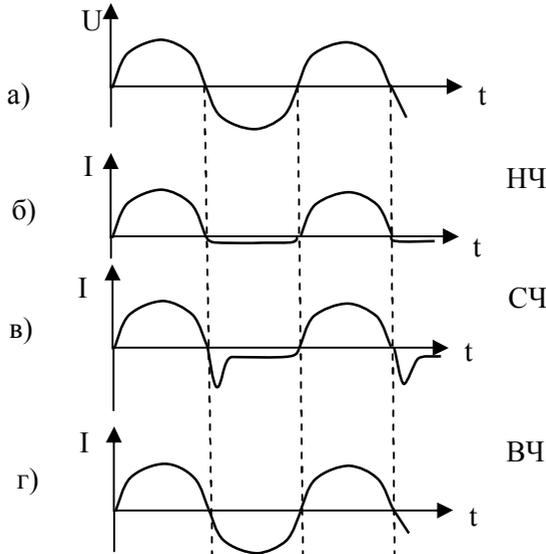
переход $\frac{1}{\omega}$. Если, $\frac{1}{\omega} \gg \tau$ т.е. $\omega\tau \ll 1$ то реализуется сигнал

низкой частоты. При $\frac{1}{\omega} \ll \tau$, т.е. $\omega\tau \gg 1$ - высокие частоты.

Соотношение $\frac{1}{\omega} \approx \tau$ ($\omega\tau \approx 1$) соответствует области средних частот.

В области низких частот сигнал изменяется во времени медленно по сравнению с τ и процессы рассасывания носителей успевают происходить. В этом случае имеем классический р-п

переход, хорошо проводящий в прямом направлении и имеющий малый ток в обратном (рис. 4.39б).



Для области средних частот за отрицательный полупериод синусоиды носители уже не успевают рассасываться, и на осциллограмме тока появляется всплеск тока (рис. 4.39, в).

В области высоких частот, когда время воздействия напряжения много меньше времени жизни неосновных носителей, процессы рекомбинации не успевают пройти и p-n-переход теряет свои выпрямляющие свойства: т.е. ток в обратном направлении равен току в прямом направлении (рис. 4.39, г).

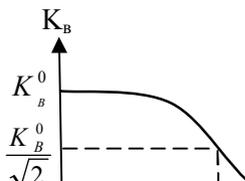
Введем понятие коэффициента выпрямления тока k_g , равно-

$$\text{го } k_g = \frac{I_{cp}^+}{I_{cp}^-},$$

где I_{cp}^+ - среднее за период значение тока в прямом направлении

I_{cp}^- - среднее за период значение тока в обратном направлении.

На рис. 4.40 представлена зависимость коэффициента выпрямления k_g от частоты.



Видно, что, начиная с некоторой частоты коэффициент выпрямления уменьшается. Взятый на уровне 0,707 от низкочастотного значения k_g^0 дает значение максимальной частоты диода f_{max} . При этом надо иметь в виду, что предельная частота $f_{пред}$ и максимальная f_{max} - это не одно и то же.

4.16. Пробой электронно-дырочного перехода

Пробой – это явление резкого возрастания обратного тока через р-п переход при подаче на него обратного напряжения. По физической сущности явлений, наблюдаемых при пробое, различают лавинный, туннельный и тепловой механизмы пробоя. Напряжение, при котором происходит пробой, называется пробивным напряжением.

Лавинный пробой

Механизм лавинного пробоя обусловлен процессом ударной ионизации. Если носитель заряда, например, электрон, оказывается в ОПЗ р-п перехода, то при обратном напряжении на него действуют достаточно большие силы электрического поля. Под действием этих сил электрон на длине свободного пробега λ набирает энергию W , равную $qE\lambda$. Если энергия, приобретенная электроном в электрическом поле, больше потенциала ионизации атома W_i , то произойдет процесс ударной ионизации: под действием электрона из нейтрального атома выбивается элек-

трон, т.е. из валентной зоны в зону проводимости генерируется свободный электрон. Таким образом, в результате многочисленных актов ударной ионизации увеличивается число свободных носителей, что и вызывает лавинообразное нарастание тока – пробой р-п перехода (рис. 4.41).

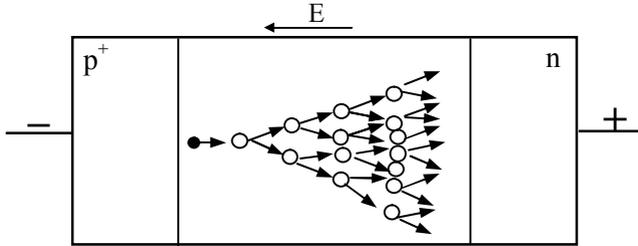


Рисунок 4.41. К механизму лавинного пробоя р-п перехода

Пробой по описанному механизму называется лавинным пробоем.

Для характеристики лавинного пробоя используются два параметра: коэффициент ударной ионизации α и коэффициент лавинного размножения M . Коэффициент лавинного размножения можно оценивать как отношение тока данных носителей (например, электронов), выходящих из перехода, к току этих же носителей, входящих в р-п переход. Коэффициент ударной ионизации α показывает количество актов ионизации, производимых носителем на длине пути d . (d – ширина ОПЗ). Обычно полагают, что коэффициент ударной ионизации α и лавинного размножения M одинаковы для электронов и дырок. Взаимосвязь между M и α находится путем решения уравнения непрерывности и дает следующее соотношение

$$1 - \frac{1}{M} = \int_0^d \alpha dx \quad (4.87)$$

При пробое $M \rightarrow \infty$ и тогда выражение $\int_0^d \alpha dx = 1$ - считается критерием лавинного пробоя.

Оценим величину напряжения пробоя при лавинном механизме, поскольку для пользователей приборов важен именно этот параметр.

Выражение для напряженности в резком $p^+ - n$ переходе

можно записать $E = \frac{qN_a}{\varepsilon\varepsilon_0}(d - x)$, а зависимость α от E задать в

виде $\alpha = A|E|^m$, где A – константа; m – параметр, зависящий от материала полупроводника (изменяется от 5 до 8). Тогда с учетом (4.87)

$$1 - \frac{1}{M} = \int_0^d \alpha dx = A \left(\frac{qN_a}{\varepsilon\varepsilon_0} \right)^m \int_0^d (d - x)^m dx = A \left(\frac{qN_a}{\varepsilon\varepsilon_0} \right)^m \frac{d^{m+1}}{m+1} \quad (4.88)$$

Подставляя в (4.88) выражение для толщины резкого $p-n$ перехода (4.22), в котором пренебрегаем φ_k^0 , т.к. $\varphi_k^0 \ll |U|$ получим:

$$1 - \frac{1}{M} = \frac{A}{m+1} \left(\frac{qN_a}{\varepsilon\varepsilon_0} \right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_a} \right)^{\frac{m+1}{2}} |U|^{\frac{m+1}{2}} \quad (4.89)$$

Полагая при пробое $U = U_{проб}$ и $M \rightarrow \infty$, запишем и (4.89)

$$1 = \frac{A}{m+1} \left(\frac{qN_a}{\varepsilon\varepsilon_0} \right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_a} \right)^{\frac{m+1}{2}} U_{проб}^{\frac{m+1}{2}} \quad (4.90)$$

Разделив (4.89) на (4.90) получим

$$1 - \frac{1}{M} = \left(\frac{U}{U_{проб}} \right)^{\frac{m+1}{2}} \quad (4.91)$$

$$\text{Из (4.91) имеем } M = \frac{1}{1 - \left(\frac{U}{U_{проб}} \right)^{\frac{m+1}{2}}} \text{ или } M = \frac{1}{1 - \left(\frac{U}{U_{проб}} \right)^s}, \quad (4.92)$$

где $\epsilon = \frac{m+1}{2}$ - коэффициент, различный для диодов из разных материалов ($\epsilon = 2 \dots 6$). При малых обратных напряжениях и $U \rightarrow 0$

$M \rightarrow 1$, а при $U \rightarrow U_{\text{проб}}$ - $M \rightarrow \infty$.

Из выражения (4.90) пробивное напряжение при лавинном пробое равно

$$U_{\text{проб}}^{\frac{m+1}{2}} = \frac{m+1}{A} \left(\frac{1}{2}\right)^{\frac{m+1}{2}} \left(\frac{\epsilon \epsilon_0}{q}\right)^{m-\frac{m+1}{2}} \left(\frac{1}{N_a}\right)^{m-\frac{m+1}{2}} \text{ или}$$

$$U_{\text{проб}} = \left[\frac{1}{2} \left(\frac{m+1}{A}\right)^{\frac{2}{m+1}} \left(\frac{\epsilon \epsilon_0}{q}\right)^{\frac{m-1}{m+1}} \right] \frac{1}{N_a^{\frac{m-1}{m+1}}}$$

(4.93)

Выражение в квадратных скобках включает в себя постоянные величины, обозначив его через B запишем (4.93) в виде

$$U_{\text{проб}} = \frac{B}{N_0^k},$$

где N - концентрация легирующей примеси в слаболегирован-

ной области р-п перехода; $k = \frac{m-1}{m+1}$.

Полученная теоретическая зависимость (4.94) имеет хорошее экспериментальное подтверждение (рис. 4.42)

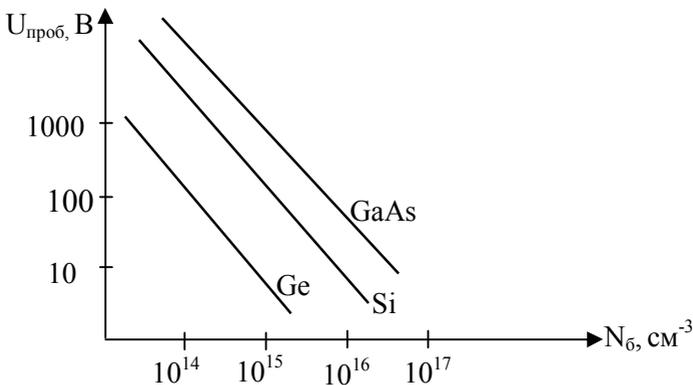


Рисунок 4.42. Зависимость пробивного напряжения от концентрации легирующей примеси в базе р-п перехода для раз-

Переходя от концентрации легирующей примеси к удельному сопротивлению базы можно записать $U_{проб} = B\rho^\alpha$.

Для кремниевых p-n переходов напряжения пробоя с численными параметрами коэффициентов β и α следующие:

$$p^+ - n \quad U_{проб} = 96\rho^{0,78}$$

$$n^+ - p \quad U_{проб} = 48\rho^{0,78}.$$

Для германиевых диодов:

$$p^+ - n \quad U_{проб} = 100\rho^{0,8}$$

$$n^+ - p \quad U_{проб} = 55\rho^{0,8}.$$

Значения B различны из-за отличий в подвижности носителей.

В инженерной практике часто используются полуэмпирические соотношения для напряжения пробоя:

$$\text{Резкий переход} \quad U_{проб} = 60\left(\frac{\Delta E}{1,1}\right)^{1,5} \left(\frac{10^{16}}{N_\delta}\right)^{0,75}$$

(4.95)

$$\text{Плавный переход} \quad U_{проб} = 60\left(\frac{\Delta E}{1,1}\right)^{1,2} \left(\frac{10^{20}}{a}\right)^{0,4}$$

(4.96)

Таким образом в резком электронно-дырочном переходе напряжение пробоя определяется концентрацией легирующей примеси в базе, а в плавном – градиентом концентрации, поскольку эти величины определяют ширину ОПЗ p-n перехода. Упрощенно данное заключение можно пояснить так: $U_{проб} = E_{проб} \cdot d$, т.е. ширина ОПЗ определяет $U_{проб}$, а она зависит и однозначно определяется концентрацией примеси в базе - N_δ .

Для лавинного механизма пробоя характерно увеличение напряжения пробоя с ростом температуры (рис. 4.43). Это связано с тем, что лавинный пробой возможен при ударной ионизации, которая наблюдается, если $W \gg W_i \geq qE\lambda$. С ростом температуры длина свободного пробега уменьшается и, для реализации условия $W > W_i$, требуются большие электрические поля. Необходимо отметить очень важную особенность лавинного пробоя: пробой носит локальный характер (рис.4.44). Т.е. формирование лавины происходит не по всей площади перехода, а в локальных участках с наибольшей напряженностью электрического поля, обусловленной не идеальностью границы ОПЗ.

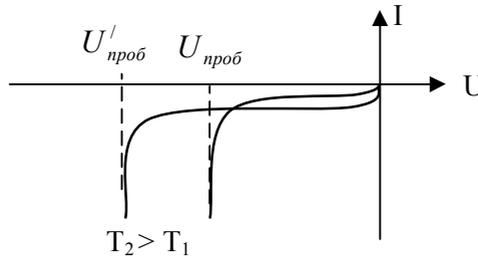


Рисунок 4.43. ВАХ p-n перехода при лавинном пробое для различных температур

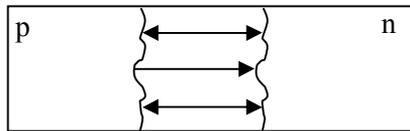


Рисунок 4.44. Пояснения к локальности процесса лавинного пробоя

Туннельный пробой

Туннельным пробоем называется электрический пробой перехода, вызванный квантово - механическим туннелированием носителей заряда через ОПЗ p-n перехода без изменения энергии. Условия, при которых возможно туннелирование определяются напряженностью электрического поля в переходе. На

рис.4.45 представлена зонная диаграмма р-п перехода при обратном смещении. Вероятность туннелирования определяется шириной ОПЗ, напряженностью электрического поля в р-п переходе и шириной запрещенной зоны. Чтобы туннелирование стало возможным необходимо, чтобы ширина ОПЗ была малой. Это возможно, если р-п переход изготовлен из сильно легированных полупроводников. При приложении обратного смещения из-за большого наклона зон становится возможным переход электронов, лежащих в диапазоне энергий ΔE в р-полупроводнике, из валентной зоны р-полупроводника в зону проводимости n-полупроводника.

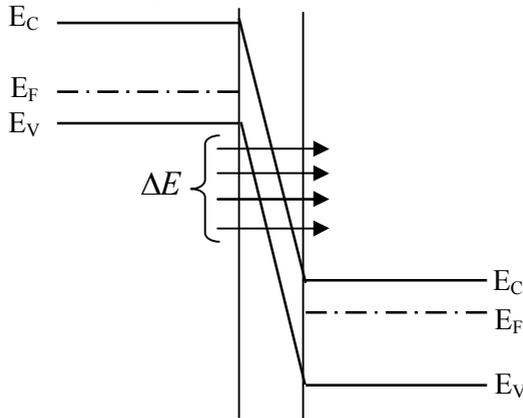


Рисунок 4.45. Зонная диаграмма при обратном напряжении для р-п перехода на основе сильнолегированных полупроводников

Вероятность туннелирования резко возрастает при увеличении напряженности электрического поля. Это вызывает сильный рост тока - туннельный пробой. Вероятность туннелирования зависит также от ширины запрещенной зоны полупроводника: чем меньше ΔE , тем более вероятно туннелирование.

Рассмотрим расчет пробивного напряжения для туннельного пробоя симметричного резкого р-п перехода. Пробой наступает тогда, когда максимальная напряженность поля E_{\max} достигает критической величины $E_{кр}$, достаточной для туннелирования. В

резком p-n переходе максимальная напряженность соответствует металлургической границе p-n перехода, т.е. при $x = 0$.

$$E_{\max} = \frac{qN_a}{\varepsilon\varepsilon_0} d_p = \frac{qN_a}{\varepsilon\varepsilon_0} \frac{N_q}{N_a + N_d} d$$

Подставляя выражение для d в E_{\max} получим

$$E_{\max}^2 = \frac{2q}{\varepsilon\varepsilon_0} \frac{N_a N_d}{N_a + N_d} (\varphi_k^0 - U)$$

При $E_{\max} = E_{кр}$ наступает туннельный пробой

$$U_{проб} = \frac{\varepsilon\varepsilon_0 E_{кр}^2}{2q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) - \varphi_k^0 \quad (4.97)$$

Из (4.97) видно, что и в случае туннельного пробоя напряжение пробоя определяется шириной ОПЗ p-n перехода, т.е. зависит от концентрации легирующей примеси: чем больше концентрация, тем меньше ширина ОПЗ и меньше напряжение пробоя. Величина пробивного напряжения для сильнолегированных p-n переходов невелика, составляя несколько вольт.

Выражение (4.97) можно, пренебрегая величиной φ_k^0 , переписать в следующем виде:

$$U_{проб} = C_p \rho_p + C_n \rho_n,$$

где C_p, C_n - постоянные величины включающие $\frac{\varepsilon\varepsilon_0 E_{кр}^2}{\mu_{p,n}}$,

ρ_p, ρ_n - удельные сопротивления p-n областей в Ом·см.

Для кремниевых и германиевых p-n переходов выражения принимают следующий вид:

$$U_{проб} = 200\rho_n + 73\rho_p \text{ - для кремниевых переходов}$$

$$U_{проб} = 190\rho_n + 94\rho_p \text{ - для германиевых переходов.}$$

Температурный коэффициент напряжения пробоя при туннельном механизме отрицательный, т.е. с ростом температуры напряжение пробоя уменьшается. Это обусловлено уменьшением ширины запрещенной зоны и ширины барьера для туннели-

рующих электронов, что вызывает увеличение вероятности туннелирования и уменьшение $U_{проб}$.

Как и лавинный пробой туннельный пробой также развивается в локальных участках р-п перехода.

Тепловой пробой

Тепловой пробой – это пробой, развитие которого обусловлено выделением в выпрямляющем переходе тепла вследствие прохождения тока через переход. Рассмотрим физический механизм теплового пробоя. Если на переход подать обратное напряжение, то через него будет протекать небольшой обратный ток и в переходе выделяется некоторая мощность

$$P_{выд} = U_{обр} I_{обр},$$

Выделяющаяся мощность вызывает увеличение температуры, если отводимая мощность меньше выделяющейся

$$P_{отв} = \frac{T - T_{окр}}{R_T},$$

где T - температура р-п перехода;

$T_{окр}$ - температура окружающей среды;

R_T - тепловое сопротивление (поток мощности при единичном градиенте температуры).

Нагрев р-п перехода вызывает увеличение обратного тока и увеличение выделяющейся мощности, что вновь ведет к росту температуры и возрастанию обратного тока, т.е. наблюдается процесс с положительной обратной связью: протекание обратного тока вы-

зывает его возрастание. Схематично процесс пробоя при соотноше-

нии $P_{выд} \geq P_{отв}$ можно представить следующей схемой

$$U_{обр} I_{обр} = P_{выд} \rightarrow T \rightarrow I_{обр} \rightarrow P_{выд} \rightarrow T \rightarrow I_{обр} \rightarrow P_{выд} \rightarrow \dots$$

Описанный процесс протекает во времени достаточно быстро, т.к. через короткий промежуток времени (10^{-7} - 10^{-6} с) обратный ток через р-п переход резко возрастает – происходит пробой. В случае наличия положительной обратной связи прибор должен

обладать ВАХ с участком отрицательного дифференциального сопротивления (рис. 4.46). Вид ВАХ при тепловом пробое показан на рис. 4.46.

Критерием теплового пробоя является соотношение

$$P_{\text{выд}} \geq P_{\text{отв}} \quad \text{или} \quad U_{\text{обр}} I_{\text{обр}} \geq \frac{T - T_{\text{окр}}}{R_T}$$

(4.98)

Выведем выражение для расчета напряжения пробоя по тепловому механизму. Для учета температурной зависимости обратного тока введем температурный коэффициент обратного тока α .

$$\alpha = \frac{1}{I_{\text{обр}}} \frac{dI_{\text{обр}}}{dT}; \quad \alpha \cdot dT = \frac{dI_{\text{обр}}}{I_{\text{обр}}}$$

(4.99)

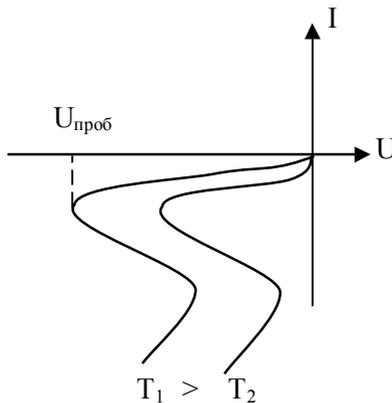


Рисунок 4.46. ВАХ р-n перехода при тепловом пробое

Интегрируя (4.99) в пределах температур от $T = T_{\text{окр}}$ - температуры окружающей среды, до $T = T$ - температуры р-n перехода при

пробое и для тока: $I_{\text{обр}} = I_{\text{окр}}$ - обратный ток при температуре окружающей среды до $I_{\text{обр}}$ при пробое получим:

$$\int_{T_{окр}}^T \alpha \cdot dT = \int_{I_{окр}}^{I_{обр}} \frac{dI_{обр}}{I_{обр}}; \quad \alpha(T - T_{окр}) = \ln \frac{I_{обр}}{I_{окр}}, \quad \text{что можно}$$

представить в виде

$$I_{обр} = I_{окр} \exp[\alpha(T - T_{окр})] \quad (4.100)$$

Выражение (4.100) отражает экспоненциальный рост тока при увеличении температуры. Подставляя (4.100) в уравнение теплового баланса (4.99) получим:

$$U_{обр} = \frac{T - T_{окр}}{R_T I_{окр}} \exp[-\alpha(T - T_{окр})] \quad (4.101)$$

При тепловом пробое $\frac{dU_{обр}}{dT} = 0$. Тогда, дифференцируя (4.101)

по температуре и приравнявая производную нулю, получим $1 - \alpha(T - T_{окр}) = 0$, откуда следует

$$T = T_{окр} + \frac{1}{\alpha} \quad (4.102).$$

Из экспериментальных данных известно, что значение температурного коэффициента обратного тока α обычно составляет 10^4 K^{-1} .

То есть температура при пробое р-п перехода не превышает температуру окружающей среды более чем на десять градусов. Следовательно тепловой пробой не вызывает значительного увеличения температуры р-п перехода, что несколько необычно, поскольку тепловой пробой предполагает развитие высоких температур.

При подстановке (4.102) в (4.101) и полагая $U_{обр} = U_{проб}$ по-

лучим

$$U_{проб} = \frac{1}{\alpha R_T I_{окр}} e$$

$$(4.103)$$

Таким образом напряжение пробоя р-п перехода при тепловом механизме зависит от конструктивного оформления прибора (R_T), температурного коэффициента обратного тока и обратного тока при заданной температуре окружающей среды. Можно отметить, что в германиевых приборах обратные токи больше, чем в кремниевых и в них преимущественно развивается тепловой пробой.

С увеличением температуры напряжение теплового пробоя уменьшается (рис.4,46), что обусловлено возрастанием обратного тока, соответствующего температуре окружающей среды и ухудшением условий теплоотвода.

Полученный результат, выраженный формулой (4.102), на первый взгляд, как уже отмечалось, кажется неожиданным. Действительно, при развитии теплового пробоя можно было бы ожидать достижения в момент пробоя достаточно высоких температур р-п перехода. Тем более этот результат кажется необычным еще и по той причине, что в ряде случаев при тепловом пробое наблюдается разрушение р-п переходов в результате плавления полупроводника или проявления необратимых физико-химических процессов в полупроводнике за счет высоких температур. Не совсем очевиден также тот факт, что время формирования теплового пробоя невелико и составляет $10^{-7} \div 10^{-6}$ с. Учитывая инерционность тепловых процессов, можно было бы ожидать значительно большего времени до развития пробоя.

Однако возникшие вопросы достаточно легко разрешимы, если исходить из шнуровой модели теплового пробоя. Аналогично тому, как это имеет место и при лавинном и при туннельном пробоях, пробой по тепловому механизму также носит локальный характер. При этом процессы, описанные в начале раздела, развиваются не по всей площади перехода, а в шнурах или каналах повышенной проводимости.

Возникновение шнуров в реальном р-п переходе может быть вызвано статической флуктуацией плотности обратного тока по площади перехода, статическим разбросом ширины ОПЗ, дефектами в р-п переходе. Действительно, если в каком-то месте р-п перехода плотность обратного тока оказалась несколько

выше, чем в других частях, то мощность, выделяемая в этом месте, возрастает, и процессы с положительной обратной связью, характерные для теплового пробоя, развиваются лавинообразно в этом участке р-п перехода, образуя канал повышенной проводимости.

Диаметр шнура, возникающего при тепловом пробое, может составлять всего несколько микрометров, а длина его определяется шириной ОПЗ, т.е. составляет десятки микрометров. Таким образом, при выделении даже малой мощности, удельная плотность мощности, выделяемая в объеме шнура, достигает значительной величины. Естественно, что локальная температура в шнуре также значительна и не соответствует соотношению (4.102). Можно говорить том, что соотношение (4.102) дает оценку интегральной температуры р-п перехода, тогда как локальная температура в шнурах значительно превышает рассчитанную по (4.102) и может, если не ограничивать обратный ток, достигать температуры плавления полупроводника или таких температур, при которых возникают необратимые физико-химические процессы в полупроводнике. Естественно, что подобные процессы делают тепловой пробой процессом необратимым. С позиций шнуровой модели пробоя становятся понятными и легко объяснимыми наблюдаемые малые времена формирования теплового пробоя: тепловые процессы в малом объеме шнура протекают достаточно быстро.

В заключении необходимо отметить, что напряжение пробоя по тепловому механизму зависит от очень многих факторов: условий окружающей среды (давления, влажности, т.д.), конструктивного оформления прибора и др.

4.17. Зависимость параметров ЭДП от температуры

Поскольку электрофизические свойства полупроводников зависят от температуры, следовательно, и параметры полупроводниковых приборов будут зависеть от температуры. Температурно-зависимыми параметрами являются: φ^{k^0} , j_s , n_i , D , τ , L , P_{n_0} , n_{p_0} .

Поскольку наиболее важной является зависимость $j_s = f(T)$, то сначала подробно остановимся на этом параметре.

1. $j_s = f(T)$. В выражении для

$$j_s = \frac{q \cdot D_p \cdot p_{n_0}}{L_p} + \frac{q \cdot D_n \cdot n_{p_0}}{L_n}$$

наибольшей зависимостью от температуры обладают p_{n_0} и n_{p_0} , т.к. они зависят от n_i^2 , поскольку в диапазоне рабочих температур $n_{n_0} = N_g$, а $p_{p_0} = N_a$.

Т.к. $n_i^2 = BT^3 \cdot \exp\left(-\frac{\Delta E}{kT}\right)$, то и p_{n_0} и n_{p_0} зависят от температуры.

Можно записать, что

$$I_s(T) = I_s \exp\left(-\frac{\Delta E}{kT}\right),$$

(4.104)

где I_s - некоторый, практически постоянный ток, соответствующий низкой температуре.

Тогда для некоторой заданной температуры T_0

$$I_s(T_0) = I_s \exp\left(-\frac{\Delta E}{kT_0}\right)$$

(4.105)

Поделив (4.104) на (4.105) получим

$$\frac{I_s(T)}{I_s(T_0)} = \exp\left(\frac{\Delta E(T - T_0)}{kTT_0}\right)$$

Или окончательно имеем

$$I_s(T) = I_s(T_0) \cdot \exp(\alpha \cdot \Delta T),$$

(4.106)

где $\alpha = \frac{\Delta E}{kTT_0}$ и $\Delta T = T - T_0$.

Из (4.106) следует, что j_s тем сильнее увеличивается с ростом температуры, чем больше ΔE . Так для $T = 300K$ у кремния $\alpha^{Si} = 0,16K^{-1}$, а у германия $\alpha^{Ge} = 0,1K^{-1}$.

Иногда в формуле (4.106) от основания e переходят к основанию 2 и тогда формула (4.106) принимает вид

$$I_s(T) = I_s(T_0) \cdot 2^{\frac{\Delta T}{T_y}}, \quad (4.107)$$

где T_y - температура удвоения тока j_s , определяемая как $T_y = \frac{\ln 2}{\alpha}$, и составляет $T_y^{Ge} = 7K$ и $T_y^{Si} = 4,5K$. На практике T_y составляет обычно $10K$ и это вызывает огромный рост j_s . Так у германиевого перехода при изменении температуры от $20^{\circ}C$ до $80^{\circ}C$ j_s возрастает в 430 раз. При больших температурах ток j_s увеличивается настолько, что диод теряет работоспособность.

В этом вопросе нельзя впасть в заблуждение. Дело в том, что на практике обратные токи у кремниевых приборов увеличиваются с ростом температуры медленнее, чем у германиевых, хотя предыдущие рассуждения говорят об обратном. Но дело в том, что у Si приборов $I_{обр} \approx I_{ген}$, а $I_{ген}$ пропорционален не n_i^2 , n_i . Поэтому T_y у кремниевых приборов вдвое больше, а α вдвое меньше, чем для I_S и составляет $T_y^{Si} = 9K$, а $\alpha^{Si} \approx 0,08K^{-1}$.

2. Прямой ток – с ростом температуры изменяется как

$I_{np} = I_s(T) \cdot (\exp \frac{qU}{kT} - 1) = I_s \cdot \exp[-\frac{\Delta E - qU}{kT}]$, т.е. растет с ростом температуры, но скорость роста I_{np} уменьшается при большом U - (рис. 4.47).

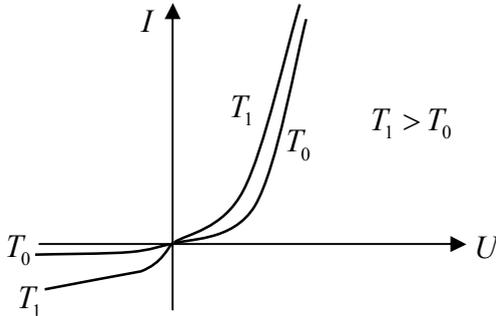


Рисунок 4.47. ВАХ ЭДП при различных температурах

3. $\varphi_k^0 = \Delta E - kT \cdot \ln \frac{N_c N_v}{N_a N_d}$. Температурный коэффициент

φ_k - $TK\varphi_k \approx 2 \frac{mB}{K}$. Уменьшение φ_k приводит к уменьшению ширины $d_{онз}$ и росту $C_{\bar{\sigma}}$. Уменьшение φ_k обусловлено тем, что с ростом температуры уровень Ферми в р- и п- полупроводниках стремится к E_i .

4. τ_n, τ_p - с ростом температуры τ увеличивается по эмпирическому закону $\tau_p(T) = \tau_p(T_0) (\frac{T}{T_0})^r$, где $r = 3,5$ для кремниевых приборов и $r = 2,2$ - для германиевых.

Так при изменении температуры от -60°C до $+80^{\circ}\text{C}$ (диапазон рабочих температур) τ_p^{Ge} увеличивается в 3 раза, а в кремнии - в 6 раз. Поэтому с ростом температуры скорость переходных процессов возрастает (увеличивается время восстановления обратного сопротивления), а $f_{пред}$ - уменьшается.

5. D и L зависят от температуры следующим образом:

$$D = D_0 \cdot \exp\left(-\frac{\Delta E_a}{kT}\right) \text{ и т.к. } L = \sqrt{D\tau}, \text{ то за счет роста } D \text{ и } \tau$$

L растет с ростом температуры.

4.18. Зарядоуправляемая модель ЭДП. Метод заряда

При быстром изменении напряжения на ЭДП ($T \approx T_\sigma$) значения токов определяются процессом накопления и рассасывания в базе ЭДП. ВАХ при этом называют *динамической*.

Для $p^+ - n$ перехода при инжекции дырок в базу из омического контакта поступает примерно такое же количество электронов, так что в целом база остается электронейтральной.

Суммарный заряд в базе равен $Q_D + Q_P \approx Q_n$,

где Q_D - заряд ионизированных доноров.

$$\text{Тогда } \frac{dQ_D}{dt} + \frac{dQ_P}{dt} = \frac{dQ_n}{dt}$$

(4.108)

При этом $\frac{dQ_n}{dt}$ обусловлен током электронов из контакта (i) и рекомбинацией носителей, т.е.

$$\frac{dQ_n}{dt} = i - \frac{Q_P}{\tau_p}$$

(4.109)

$$\frac{dQ_D}{dt} = \frac{dQ_D}{dU} \frac{dU}{dt} = C_\sigma \frac{dU}{dt}$$

(4.110)

Подставляя (4.109) и (4.110) в (4.108), получим

$$C_\sigma \frac{dU}{dt} + \frac{dQ_P}{dt} + \frac{Q_P}{\tau_p} = i \quad - \text{ это уравнение метода заряда}$$

(4.111).

Оно широко используется для анализа переходных процессов.

Динамическая ВАХ ЭДП в этой связи имеет вид

$$i = I(U) + C \frac{dU}{dt},$$

где $I(U)$ - статическая ВАХ, $C = C_{\bar{\sigma}} + C_D$.

Зная C , с помощью этого выражения можно определять динамические ВАХ - $i = f(U, \frac{dU}{dt})$ при различных $\frac{dU}{dt}$. Статические и динамические характеристики сильно отличаются.

Для описания переходных процессов в полупроводниковых приборах, в том числе диодах, широко используется метод заряда. Из рассмотрения сути переходных процессов следует, что изменение тока в диоде при включении и переключении определяется процессами накопления и рассасывания заряда в базе. Т.е. можно записать без учета изменения заряда в ОПЗ ($\frac{dQ_D}{dt} = 0$) уравнение заряда (4.111) в следующем виде

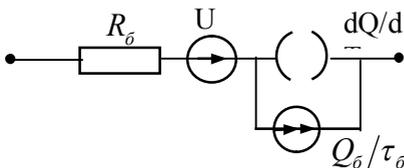
$$I = \frac{dQ_{\bar{\sigma}}}{dt} + \frac{Q_{\bar{\sigma}}}{\tau_{\bar{\sigma}}} \quad - \text{это основное уравнение метода заряда,} \quad (4.112)$$

где $\frac{dQ_{\bar{\sigma}}}{dt}$ - отражает процесс накопления зарядов в базе,

а $\frac{Q_{\bar{\sigma}}}{\tau_{\bar{\sigma}}}$ - отражает процесс рассасывания носителей из базы за счет рекомбинации.

Более понятно, если это уравнение записать так $\frac{dQ_{\bar{\sigma}}}{dt} = I - \frac{Q_{\bar{\sigma}}}{\tau_{\bar{\sigma}}}$.

Уравнению (4.112) можно сопоставить схему замещения ЭДП.



где, элемент $\left(\int \right) - \frac{dQ}{dt}$ - это конденсатор большой емкости, отражает процесс накопления носителей в базе и называется накопителем.

 - $\frac{Q_0}{\tau_0}$ - генератор тока, отражающий процесс рассасывания носителей из базы за счет рекомбинации.

Решение уравнения (4.112) дает вид зависимости $I = f(t)$

$$I = I_0 \exp\left(-\frac{t}{\tau}\right) \text{ - при отключении ЭДП,}$$

$$I = I_0 \left[1 - \exp\left(-\frac{t}{\tau_0}\right) \right] \text{ - при включении ЭДП,}$$

где I_0 - ток в установившемся режиме работы. Зависимость заряда в базе от времени при включении имеет вид

$$Q_0(t) = I \cdot \tau \left[1 - \exp\left(-\frac{t}{\tau_0}\right) \right],$$

где τ_0 - время жизни носителей заряда.

4.19. Гетеропереходы

Гетеропереходом называется электрический переход на контакте полупроводников с различной шириной запрещенной зоны. В зависимости от контактирующих полупроводников гетеропереходы делятся на изотипные ($n^+ - n, p^+ - p$) и анизотипные ($n^+ - p, p^+ - n$).

Электрический переход металл-полупроводник является частным случаем гетероперехода. Примерами гетеропереходов могут быть переходы германий-кремний, германий-арсенид галлия, арсенид галлия-фосфид галлия и т.д.

Для получения гетеропереходов с минимальным числом дефектов на границе раздела кристаллическая решетка одного по-

лупроводника должна с минимальными нарушениями переходить в кристаллическую решетку другого.

Основное отличие свойств гетероперехода от гомоперехода хорошо видно при рассмотрении зонной диаграммы p-n перехода (рис. 4.48). Правила изображения зонных диаграмм гетеропереходов не отличаются от изображения диаграмм гомопереходов. На рис. 4.48 представлены зонные диаграммы для случая контактирования узкозонного полупроводника n-типа и широкозонного полупроводника p-типа (рис. 4.48а) и контакт широкозонного полупроводника n-типа и узкозонного полупроводника p-типа (рис. 4.48,б).

Из рис. 4.48 видно, что высота потенциального барьера для электронов и дырок в гетеропереходе значительно отличаются. Так для контакта, изображенного на рис.4.48а, для электронов потенциальный барьер велик, что препятствует переходу электронов из n в p область, а на рис. 4.48б наоборот, высокий потенциальный барьер препятствует инжекции дырок. В любом случае в гетеропереходе инжекция носителей происходит из области полупроводника с широкой запрещенной зоной в узкозонный полупроводник.

Гетеропереходы на основе контактов полупроводников с одним типом проводимости ($n^+ - n, p^+ - p$) обладают эффектом односторонней проводимости без инжекции неосновных носителей. Значения параметров ΔE и ΔE_V определяются выражением (4.113).

$$\begin{aligned} \Delta E_C &= \chi_1 - \chi_2 \\ \Delta E_V &= \Delta E_2 - \Delta E_1 - \Delta E_C, \end{aligned} \quad (4.113)$$

где χ_1, χ_2 -электронное сродство полупроводников;

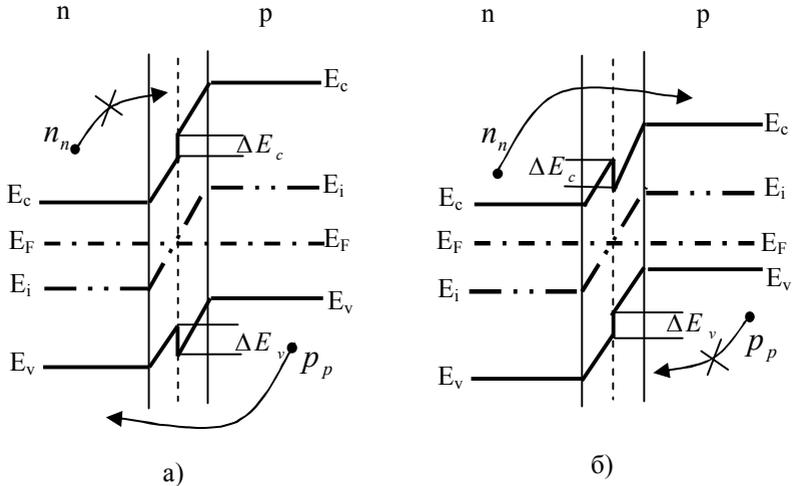


Рисунок 4.48. Энергетические диаграммы гетеропереходов
 а) узкозонный n-типа – широкозонный p-типа; б) широкозонный n-типа
 –узкозонный p-типа

ΔE_1 , ΔE_2 - ширина запрещенной зоны полупроводников.

Таким образом, основное отличие гетероперехода от гомоперехода заключается в том, что токи в таком переходе обусловлены движением носителей одного знака. Можно показать, что отношение тока электронов к току дырок в гетеропереходе определяется выражением (4.114).

$$\frac{I_p}{I_n} = \exp\left(-\frac{\Delta E_n - \Delta E_p}{kT}\right) \quad (4.114)$$

Тогда для гетероперехода, например, n-Ge – p-Si это отношение равно $\frac{I_p}{I_n} = e^{-16}$, что практически равно нулю. Тогда как в сим-

метричном гомопереходе отношение $\frac{I_p}{I_n} = 1$, т. е. гетеропереход

обладает односторонней инжекцией. Более детальное описание свойств гетероперехода можно получить из учебного пособия

В.И.Гамана «Физика полупроводниковых приборов» - Томск:
изд-во НТЛ, 2000.-426 с.

5. ДИОДЫ НА ОСНОВЕ ЭЛЕКТРОННО–ДЫРОЧНЫХ ПЕРЕХОДОВ (ЭДП)

5.1. Классификация и маркировка диодов

Полупроводниковый диод – это полупроводниковый прибор, содержащий один выпрямляющий переход и два вывода. По определению любой диод содержит, например, один ЭДП, а выполняемые диодами функции весьма разнообразны. Поэтому при рассмотрении диодов различного функционального назначения прежде всего необходимо указать какой физический эффект в р – n переходе используется для создания диода.

Любой полупроводниковый прибор характеризуется набором параметров. Все параметры делятся на предельнодопустимые и характеризующие. Предельно допустимыми являются параметры, превышение которых приводит к выходу прибора из рабочего состояния. Остальные параметры относятся к категории характеризующих.

Классификацию диодов можно проводить по ряду признаков:

- по материалам, используемым при изготовлении прибора – германий, кремний, арсенид галлия;

- по мощности: маломощные, средней мощности, мощные;

- по частоте: низкочастотные, на средние частоты, высокочастотные, сверхвысокочастотные;

- по технологии изготовления: сплавные, диффузионные, эпитаксиальные и т. д.;

- по геометрическим размерам: точечные, сплавные;

- по функциональному назначению: усилительные, генераторные, переключательные, индикаторные.

В соответствии с ГОСТ 10862 – 72 для маркировки диодов используют шестизначный буквенно–цифровой код.

Первый элемент маркировки - это буквы Г, К, А или цифры 1, 2, 3 - показывает материал, из которого изготовлен прибор (Г – германий, К – кремний, А – арсенид галлия – для приборов общего назначения и 1 –германий, 2 – кремний, 3 – арсенид

галлия - для приборов специального назначения). Отличие в приборах общего и специального назначения заключается в условиях заводской приемки.

Второй элемент маркировки – это буква, которая определяет функциональное назначение (подкласс) прибора:

- Д – диоды выпрямительные, универсальные, импульсные;
- Ц – выпрямительные столбы и блоки;
- А – сверхвысокочастотные диоды;
- В – варикапы;
- С – стабилитроны и стабисторы;
- И – туннельные и обращенные диоды;
- Г – генераторы шума;
- Ф – фотодиоды;
- Л – светоизлучающие диоды;
- О – Оптопары (оптроны).

Третий элемент маркировки – цифра, обозначающая назначение и некоторые технические характеристики прибора (табл. 5.1.)

Таблица 5.1. Третий элемент маркировки полупроводниковых диодов

Полупроводниковые диоды	Обозначения
Выпрямительные:	
а) малой мощности (со средним значением прямого тока не более 0,3А)	1
б) средней мощности (со средним значением прямого тока более 0,3А, но не более 10А)	2
Универсальные (с рабочей частотой не более 1000 МГц)	4
Импульсные:	
а) со временем восстановления обратного сопротивления более 150 нс	5
б) со временем восстановления обратного	6

сопротивления более 30 нс, но не более 150 нс в) со временем восстановления обратного сопротивления более 5 нс, но не более 30 нс г) со временем восстановления обратного сопротивления не менее 1нс и не более 5нс д) со временем восстановления обратного сопротивления менее 1нс	7 8 9
Выпрямительные столбы и блоки: а) столбы малой мощности (со средним значением прямого тока не более 0,3А) б) столбы средней мощности (со средним значением прямого тока более 0,3А, но не бо- лее 10А) в) блоки малой мощности (со средним зна- чением прямого тока не более 0,3А) г) блоки средней мощности (со средним значением прямого тока более 0,3А, но не более 10А)	1 2 3 4
Сверхвысокочастотные: а) смесительные б) детекторные в) параметрические г) регулирующие (переключательные, ог- раничительные и модуляторные) д) умножительные е) генераторные	1 2 4 5 6 7
Варикапы: а) подстроечные б) умножительные (варакторы)	1 2
Туннельные и обращенные: а) усилительные б) генераторные в) переключательные г) обращенные	1 2 3 4

Излучающие:	
инфракрасного диапазона	1
видимого диапазона (светодиоды):	
а) с яркостью не более 500кд/м ²	3
б) с яркостью более 500кд/м ²	4

Четвертый и пятый элементы маркировки – две цифры (от 01 до 99), обозначающие порядковый номер разработки.

Шестой элемент маркировки – буква, обозначающая параметрическую группу приборов.

Например. КД215А – диод полупроводниковый выпрямительный, общего применения, кремниевый, средней мощности, номер разработки 15, группы А; 2Д504Б – полупроводниковый импульсный диод из кремния, специального назначения, с временем восстановления обратного сопротивления более 150 нс, номер разработки 04, группа Б.

Наряду с маркировкой приборов по ГОСТ 10862-72 выпускаются и используются диоды, маркируемые по ГОСТ 5461-59. В соответствии с этим ГОСТ для обозначения диодов используются три элемента:

первый – буква Д, обозначающая, что прибор относится к классу диодов;

второй – цифры, указывающие материал и назначение прибора;

третий – буква, означающая параметрическую группу.

Цифры второго элемента маркировки означают следующее:

1 – 100 – точечные германиевые диоды;

101 – 200 – точечные кремниевые диоды;

201 – 300 – плоскостные кремниевые;

301 – 400 – плоскостные германиевые;

401 – 500 – смесительные;

501 – 600 – умножительные;

601 – 700 – детекторные;

701 – 749 – параметрические германиевые;

750 – 800 – параметрические кремниевые;

801 – 900 – стабилитроны;

901 – 950 – варикапы;

951 – 1000 – туннельные;

Диоды СВЧ
диапазона

1000 – 1100 – выпрямительные столбы.

Для маркировки мощных диодов ГОСТ 10862-79 не используется.

5.2. Выпрямительные диоды. Выпрямительные столбы и блоки

Выпрямительный диод – это полупроводниковый диод, предназначенный для преобразования переменного сигнала в постоянный в диапазоне частот от 50 Гц до 50 кГц.

Физической основой работы выпрямительных диодов является эффект односторонней проводимости ЭДП.

К числу предельных параметров выпрямительных диодов относятся следующие:

1) Допустимое обратное напряжение $U_{обр. доп.}$.

Поскольку выпрямительные диоды работают на эффекте односторонней проводимости, то рабочая точка не должна попадать на участок пробоя на ВАХ. С целью обеспечения надежной работы диодов обычно $U_{обр. доп.}$ выбирается на уровне $0,5 \div 0,8 U_{пробоя}$.

1) Обратный ток при заданном обратном напряжении $I_{обр. доп.}$. Обычно значение $I_{обр. доп.}$ определяется при $U_{обр. доп.}$.

3) Допустимый прямой ток $I_{пр. доп.}$.

Значение $I_{пр. доп.}$ определяется исходя из значений предельной плотности тока через р-п переход. Для приборов из кремния

предельная плотность тока $j_{пред.}^{Si}$ составляет $200 \frac{A}{см^2}$, а для

приборов из германия - $100 \frac{A}{см^2}$. Значение $I_{пр. доп.}$ должно быть

таким, чтобы плотность этого была бы не выше указанных для германия и кремния.

4) Допустимое прямое напряжение $U_{пр. доп.}$.

Это напряжение определяется как напряжение, соответствующее предельно допустимому прямому току.

5) Максимальная рабочая температура T_{max} .

6) Максимальная мощность P_{max} .

К числу характеризующих параметров для выпрямительных диодов относятся: значения диффузионной и барьерной емко-

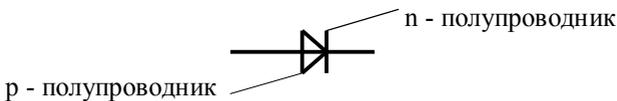
стей, постоянная времени диода, значение предельной и максимальной частот, среднее значение выпрямительного тока и т.д.

По значению прямого допустимого тока выпрямительные диоды делятся на: маломощные - $< 0,3 \text{ А}$;
 средней мощности – $0,3 \div 10 \text{ А}$;
 мощные (силовые) - $> 10 \text{ А}$.

Диоды, работающие на повышенных частотах (до 100 кГц) называют частотными. Для защиты р-п перехода от внешних воздействий и обеспечения хорошего теплоотвода кристалл монтируется в корпус. Диоды с малой мощностью обычно оформляются в пластмассовом корпусе с гибкими выводами. Диоды средней мощности – в металлическом или в металлокерамическом корпусе. Для мощных диодов кристалл размещается на массивном медном основании. Отличительные особенности выпрямительных арсенид-галлиевых диодов заключаются в более высоких рабочих температурах (до 250°C) и повышенном значении предельной частоты (до 1 МГц).

Выпрямительные столбы – это совокупность соединенных последовательно или параллельно - последовательно (с целью увеличения токовой нагрузки) выпрямительных диодов, размещенных в пластмассовом корпусе, и предназначенных для преобразования переменного сигнала в постоянный при амплитуде переменного сигнала больше, чем $U_{\text{обр. доп}}$ для отдельного диода. По мощности они делятся на маломощные ($I_{\text{пр. доп.}} < 0,3 \text{ А}$) и средней мощности ($0,3 \div 10 \text{ А}$).

Выпрямительные блоки – это полупроводниковый блок, собранный из выпрямительных диодов, соединенных по определенной электрической схеме и оформленных в единую конструкцию, выполняющий функцию преобразования переменного сигнала в постоянный. Выпрямительные блоки также делятся на маломощные и средней мощности. Например, выпрямительный блок КЦ401Б - выпрямительный мост, удвоитель напряжения; КЦ402А - Н – однофазный мост; КЦ403А-Н – два однофазных моста. Условно-графическое обозначение выпрямительных диодов и выпрямительных столбов



5.3. Универсальные и импульсные диоды

Импульсными диодами - называют диоды, работающие в импульсном режиме и имеющие малое время восстановления обратного сопротивления. Импульсные диоды работают в быстродействующих импульсных схемах с временем переключения менее 1 мкс. Физической основой работы импульсных диодов является эффект односторонней проводимости ЭДП. При столь коротких импульсах необходимо учитывать инерционность процессов выключения и принимать конструктивно-технологические меры, направленные на снижение C_{σ} и τ_{σ} .

Параметры импульсных диодов.

1) $t_{\text{вос.}}$ - время восстановления обратного сопротивления. Это время от момента подачи обратного напряжения на диод до момента установления обратным током заданного низкого значения (обычно равного току насыщения). По значению $t_{\text{вос.}}$ импульсные диоды делятся на пять групп. Соответственно в маркировке диодов это отражается третьей цифрой:

Третья цифра в маркировке:	$t_{\text{вос.}}$, нс
9	< 1
8	1 ÷ 5
7	5 ÷ 30
6	30 ÷ 150
5	> 150

2) $Q_{\text{пк}}$ – заряд переключения - часть накопленного заряда, вытекающая во внешнюю цепь при изменении направления тока с прямого на обратное.

3) $C_{\text{д}}$ – общая емкость – емкость, измеренная между выводами диода при заданных напряжении смещения и частоте.

4) $t_{\text{уст}}$ - время установления прямого напряжения – время от момента подачи импульса прямого тока до момента установления напряжения на диоде, равного 1,2 от U_0 (рис. 5.1).

5) $U_{\text{обр.доп.}}^{\text{умп}}$ - обратное допустимое напряжение в импульсе.

- 6) $I_{np.\dot{d}on.}^{имп}$ - допустимый прямой ток в импульсе.
 7) $U_{np.\dot{d}on.}^{имп}$ - допустимое прямое напряжение в импульсе.
 8) $I_{восст.\max}$ - максимальный ток восстановления.

На схемах импульсные диоды обозначаются так же, как и выпрямительные.

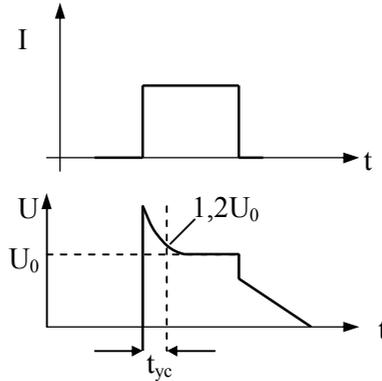


Рисунок 5.1. Осциллограммы тока и напряжения при работе импульсного диода в схеме с генератором тока

Планарная технология получения импульсных диодов дает возможность достаточно просто формировать в одном кристалле много диодных структур – диодные сборки или матрицы импульсных диодов.

Импульсные диоды широко применяются в схемах различного назначения, в том числе и схемах логических элементов. На рис. 5.2. представлена схема логического элемента, реализующего функцию «И» на три входа.

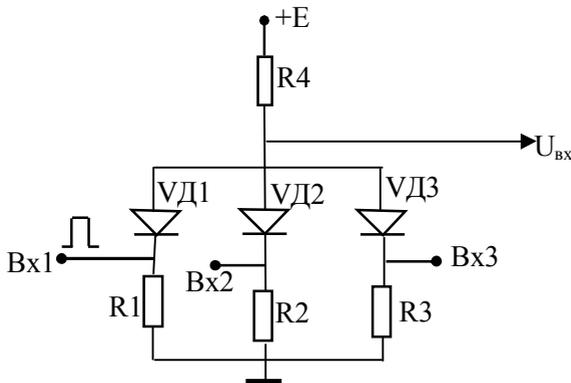


Рисунок 5.2. Схема логического элемента «И» на три вывода

Сигнал на выходе схемы появится только тогда, когда имеются сигналы на каждом из трех входов. При отсутствии сигнала диоды смещены в прямом направлении, имеют малое сопротивление и все напряжение от источника питания падает на R4, т.е. сигнал на выходе равен нулю.

При подаче импульса положительной полярности на все входы диоды переходят в закрытое состояние, они имеют высокое сопротивление, и напряжение от источника питания падает в основном на диодах, что вызывает появление на выходе сигнала, равного логической «1». Причем, это произойдет, если сигнал поступает на все три входа.

С целью повышения быстродействия импульсных диодов (уменьшения $t_{\text{восст.}}$) обычно базу диода легируют атомами золота, что приводит к уменьшению времени жизни носителей в базе - τ_{δ} . Второй путь уменьшения $t_{\text{восст.}}$ - формирование неоднородного легирования базы, сопровождающееся появлением электрического поля в базе (диоды с накоплением заряда или диоды с резким восстановлением обратного сопротивления). Создающееся за счет неоднородного легирования электрическое поле в базе не дает возможности инжектированным носителям диффундировать вглубь базы. Поэтому при переключении диода с прямого на обратное напряжение вытягивание неосновных носителей из базы происходит значительно быстрее, что уменьшает $t_{\text{восст.}}$.

Универсальные (высокочастотные) диоды

Это приборы универсального назначения, которые могут использоваться для выпрямления, детектирования, в импульсных схемах на частотах до 600 МГц. На схемах обозначаются также как выпрямительные и импульсные. Третий элемент маркировки – цифра 4. Например, КД407А, КД409А.

5.4. Сверхвысокочастотные диоды

Сверхвысокочастотный полупроводниковый диод (СВЧ-диод) – это полупроводниковый диод, предназначенный для преобразования и обработки сверхвысокочастотного сигнала.

Полупроводниковые СВЧ-диоды уже длительное время применяют в различной радиоэлектронной аппаратуре и измерительной технике СВЧ-диапазона, т.е. на частотах более 300 МГц. Вначале СВЧ – диоды использовали для детектирования и смещения сигналов. Для этих целей применяли точечные диоды, выпрямляющий электрический переход в которых возникал между кристаллом полупроводника и прижимным металлическим электродом в виде заостренной пружинки. Созданные в последнее время новые типы СВЧ-диодов практически целиком заменили точечные детекторные и смесительные диоды. Они дают возможность решать задачи генерации и усиления электромагнитных колебаний СВЧ диапазона, умножения частоты, модуляции, регулирования, ограничения сигналов и т.п. Все СВЧ-диоды делятся на смесительные, детекторные, параметрические, переключательные, умножительные, генераторные.

1) Смесительные диоды

Смесительный полупроводниковый диод — это полупроводниковый диод, предназначенный для преобразования высокочастотных сигналов в сигнал промежуточной частоты.

К смесительному диоду подводится сигнал и напряжение от специального генератора — гетеродина. В связи с нелинейностью ВАХ диода происходит образование сигнала разностной (промежуточной) частоты. Дальнейшее усиление входного сигнала осуществляется на этой промежуточной частоте, которая должна быть выше частот, соответствующих низкочастотным шумам, обратно пропорциональным частоте.

Основным параметром смесительных диодов, определяющим эффективность преобразования входных сигналов высокой частоты в сигналы промежуточной частоты, является параметр $L_{\text{прб}}$, называемый **потери преобразования смесительного диода** и равный отношению мощности СВЧ-сигнала на входе одной камеры к мощности сигнала промежуточной частоты, выделяемой в нагрузке смесительного диода в рабочем режиме.

$$L_{\text{прб}}[\text{дБ}] = 10 \lg \frac{P_{\text{СВЧ}}}{P_{\text{пр.ч}}}.$$

В большинстве приемных устройств СВЧ-диапазона отсутствуют усилители перед смесителем. Поэтому чувствительность всего приемного устройства, возможность различить полезный сигнал на фоне шумов зависят от уровня шумов смесительного диода. Уровень шумов смесительного диода (и других приборов) оценивают *шумовым отношением* $n_{\text{ш}}$ — отношением номинальной мощности шумов диода в рабочем режиме к номинальной мощности тепловых шумов соответствующего активного сопротивления при той же температуре и одинаковой полосе частот.

Другим параметром, характеризующим шумы смесительного диода и других приборов и систем, является *коэффициент шума* — отношение мощности шумов на выходе к той ее части, которая вызвана тепловыми шумами источника сигнала:

$$F = \frac{\left(\frac{P_{\text{сигн}}/P_{\text{ш}}}{\right)_{\text{вх}}}{\left(\frac{P_{\text{сигн}}/P_{\text{ш}}}{\right)_{\text{вых}}} = \frac{P_{\text{ш.вых}}}{P_{\text{ш.вх}} (P_{\text{сигн.вых}}/P_{\text{сигн.вх}})}$$

Обобщенным параметром приемного устройства, в смесителе которого использован диод с определенными потерями преобразования и шумовым отношением, является *нормированный коэффициент шума* — значение коэффициента шума приемного устройства со смесительным диодом на входе при коэффициенте шума усилителя промежуточной частоты $F_{\text{упч}}$, равном 1,5 дБ:

$$F_{\text{норм.}[дБ]} = L_{\text{прб}[дБ]} + 10 \lg(n_{\text{ш}} + F_{\text{упч}} - 1)$$

Одним из вспомогательных параметров смесительных диодов служит *выпрямленный ток* $I_{\text{ен}}$ — постоянная составляющая тока, протекающая в выходной цепи диода в рабочем режиме. Этот параметр используется для контроля исправности смесительного диода и гетеродина приемника, от которого на смесительный диод подается определенная мощность СВЧ-колебаний с определенной длиной волны.

Другим вспомогательным параметром является *коэффициент стоячей волны по напряжению* СВЧ-диода $K_{\text{стУ}}$ — коэффициент стоячей волны по напряжению в передающей линии СВЧ, когда она нагружена на определенную диодную камеру с СВЧ-

диодом в рабочем режиме. Чем лучше согласовано входное сопротивление камеры (с диодом) с волновым сопротивлением тракта, тем меньше коэффициент стоячей волны по напряжению и потери принимаемого сигнала.

Основные требования к исходному полупроводниковому материалу, структуре СВЧ - диодов и причины этих требований можно сформулировать следующим образом.

Условиями получения необходимых частотных свойств СВЧ-диодов являются отсутствие инжекции неосновных носителей заряда через выпрямляющий электрический переход в базу, малое время жизни неосновных носителей в базе и малое значение постоянной времени перезаряда барьерной емкости, т. е. малые барьерная емкость выпрямляющего перехода и сопротивление базы, которое должно быть малым и для уменьшения потерь мощности в диоде.

Пробивное напряжение, хотя оно и не является параметром СВЧ-диодов, должно быть большое. Во-первых, это необходимо для предотвращения детектирования на обратной ветви ВАХ тока от гетеродина, который выдает переменное напряжение с относительно большой амплитудой. Во-вторых, это необходимо для повышения надежности СВЧ-диола, так как на вход радиоприемных устройств могут попадать посторонние радиоимпульсы большой мощности. Для обеспечения достаточных значений пробивного напряжения, а также для уменьшения барьерной емкости концентрация примеси в базе около выпрямляющего электрического перехода должна быть малой, что противоречит требованию малости сопротивления базы.

Для повышения надежности СВЧ-диодов пробой их должен быть лавинным, а не тепловым, который происходит с неизбежным шнурованием тока.

Отсюда следует, что исходный полупроводниковый материал для СВЧ-диодов должен иметь большую ширину запрещенной зоны, малое время жизни неосновных носителей заряда и большую подвижность основных носителей заряда, т. е. при заданной концентрации примеси он должен иметь меньшее удельное сопротивление. Таким материалом, в частности, является арсенид галлия.

В качестве смесительных СВЧ-диодов наиболее широкое распространение в последнее время получили диоды Шоттки. Выпрямляющий электрический переход этих диодов формируют путем напыления металла на поверхность высокоомного эпитаксиального слоя арсенида галлия с использованием фотолитографии.

Размеры выпрямляющего перехода зависят от частоты, на которой должен работать диод. Для очень высоких частот (в десятки и сотни гигагерц, что соответствует миллиметровому диапазону длин волн) с целью уменьшения барьерной емкости необходимо уменьшение диаметра перехода Шоттки до 2...3 мкм. Такие размеры оказываются предельными для обычной фотолитографии, в частности, из-за подтравливания защитного слоя диоксида под маской фоторезиста при химическом травлении окон в слое диоксида. Для устранения этого явления используют методы ионно-плазменного травления.

Кроме того, при очень высоких частотах сказывается скин-эффект в металлическом контакте, образующем переход Шоттки. Поэтому с целью увеличения отношения периметра перехода к его площади необходимо создавать кольцевые, полосковые, крестообразные или эллиптические переходы. Для сохранения малой площади перехода, имеющего сложную конфигурацию, ширина кольца, полосок и т. д. должна быть мала (около 1 мкм). Метод фотолитографии при этом оказывается неприемлемым. Здесь используют методы рентгенолитографии и электроннолучевые методы литографии, которые имеют значительно большую разрешающую способность.

Для удобства включения в соответствующие элементы и цепи СВЧ (например, волноводные и коаксиальные линии передач) СВЧ-диоды оформляют в корпусах различной конструкции (рис. 5.3).

Типы корпусов СВЧ-диодов, их габаритные и присоединительные размеры (как и многих других полупроводниковых приборов) соответствуют ГОСТ 18472—82 (СТСЭВ 1818—79).

Диоды с патронной конструкцией корпуса, состоящего из керамической втулки и латунных фланцев или ниппелей (рис. 5.3, а), предназначены для использования в деци- и сантиметровом диапазонах длин волн, т. е. до частот около 12 ГГц.

Диоды с коаксиальной конструкцией корпуса (рис. 5.3,б) применяют в диапазоне коротких сантиметровых волн до частот около 30 ГГц. В миллиметровом диапазоне волн используют преимущественно диоды волноводной конструкции, т. е. волноводные вставки, которые и являются своеобразными корпусами СВЧ-диодов. Для полосковых линий и для интегральных СВЧ-микросхем используют либо диоды в миниатюрных корпусах, либо бескорпусные СВЧ-диоды, поверхность полупроводникового кристалла которых защищена только пленкой диоксида.

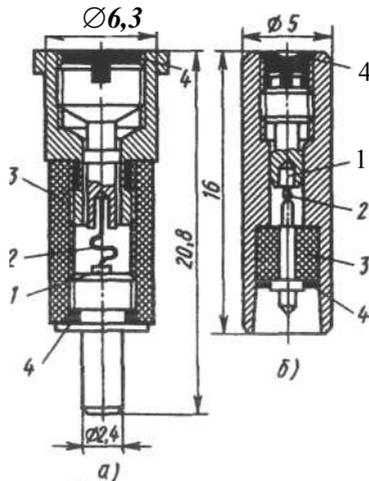


Рисунок 5.3. Конструкции некоторых СВЧ-диодов патронного (а) и коаксиального типов (б):

1 — кристалл полупроводника, 2 — контактная пружина, 3 — керамическая втулка, 4 — герметизирующая заливка

Конструкция корпуса СВЧ-диода может существенно влиять на его частотные свойства. Для уменьшения этого влияния электрическая емкость корпуса и индуктивность внешних и внутренних выводов должны быть минимальными. Металлические детали СВЧ-диодов обычно покрывают тонким слоем серебра или золота, обеспечивая тем самым минимальное их сопротивление, надежность контакта с внешними цепями и защиту от коррозии.

2) Детекторные диоды

Детекторный полупроводниковый диод – это полупроводниковый диод, предназначенный для детектирования сигнала.

При детектировании используется выпрямляющее свойство диода для выделения из модулированных по амплитуде ВЧ- или СВЧ-колебаний сигнала более низкой частоты, который потом поступает на вход усилителя (рис. 5.4). Одним из основных параметров детекторных СВЧ-диодов является *чувствительность по току* β_i — отношение приращения выпрямленного тока при заданной нагрузке в выходной цепи диода к мощности СВЧ-сигнала, подводимой к входу диодной камеры с детекторным диодом в рабочем режиме и вызвавшей это приращение.

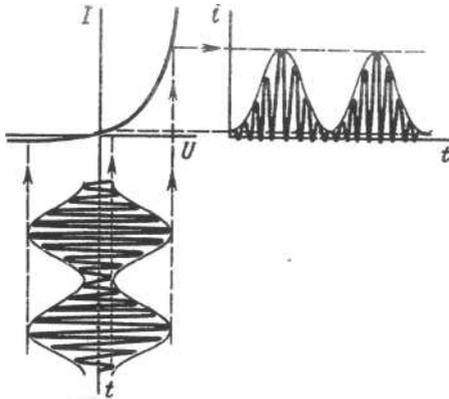


Рисунок 5.4. Графики, поясняющие детектирование высокочастотных колебаний — выделение из модулированных по амплитуде высокочастотных колебаний сигнала более низкой частоты

Чувствительность по току детекторного диода зависит от постоянного прямого тока смещения (на рис. 5.4. приведено графическое пояснение детектирования без постоянного смещения). Наибольшие значения чувствительности по току обычно получают при прямом токе смещения в несколько десятков микроампер, но при выборе тока смещения необходимо учитывать его влияние и на другие параметры.

Обобщенным параметром детекторного диода, учитывающим различные свойства диода и следующего за ним усилителя (ви-

деоусилителя), является коэффициентом качества детекторного диода, который характеризует чувствительность приемного устройства с детекторным диодом и определяется по формуле

$$M = \frac{\beta_i r_{\text{диф}}}{\sqrt{n_{\text{ш}} r_{\text{диф}} + r_{\text{ш}}}},$$

где $r_{\text{диф}}$ — дифференциальное сопротивление диода при определенном положительном смещении;

$n_{\text{ш}}$ — шумовое отношение СВЧ-диода;

$r_{\text{ш}}$ — эквивалентное шумовое сопротивление видеоусилителя, которое обычно принимают при расчетах равным 1 кОм.

Лучшие детекторные СВЧ-диоды имеют коэффициент качества более $100 \text{ Вт}^{-1/2}$. К таким диодам можно отнести, например, диоды Шоттки с планарно-эпитаксиальной структурой на основе арсенида галлия АА204А...АА204В, предназначенные для детектирования в сантиметровом диапазоне длин волн.

3) Переключательные диоды

Переключательный полупроводниковый диод — это полупроводниковый диод, предназначенный для применения в устройствах управления уровнем сверхвысокочастотной мощности.

Принцип действия переключательного диода основан на большом различии полного сопротивления СВЧ-сигналу при прямом постоянном токе через диод и при обратном постоянном напряжении на диоде. Именно поэтому СВЧ-тракт (волноводная, коаксиальная или полосковая линия), следующий за переключательным устройством с диодом, может быть либо открыт, либо закрыт для СВЧ-сигнала. Например, в радиолокационных станциях с фазированными решетками, содержащими тысячи идентичных антенных элементов, переключательные диоды должны обеспечить подачу мощного СВЧ-импульса на каждый элемент в определенные моменты времени. При этом мощные импульсы передатчика не должны попадать в канал чувствительного приемника.

Отсюда ясны основные требования к переключательным СВЧ-диодам. Они должны с минимальными потерями пропус-

вать СВЧ-мощность в состоянии пропускания и не пропускать — в состоянии запираания, обладать большой допустимой мощностью рассеяния, большим пробивным напряжением, малой собственной емкостью и достаточно большой скоростью переключения.

Обобщенным параметром переключающего диода является *критическая частота* $f_{кр}$, которая характеризует эффективность переключающего диода и определяется по формуле

$$f_{кр} = \frac{1}{2\pi C_{стр} \sqrt{r_{пр} r_{обр}}},$$

где $C_{стр}$ — емкость структуры; $r_{пр}$ — прямое сопротивление потерь (активная составляющая полного сопротивления диода) при определенном прямом токе смещения;

$r_{обр}$ — обратное сопротивление потерь при определенном обратном напряжении смещения.

Для увеличения допустимой мощности рассеяния диода необходимо увеличивать площадь выпрямляющего электрического перехода, что влечет за собой увеличение барьерной емкости. Поэтому большинство переключающих СВЧ-диодов имеет р-і-п структуру, толщина р-п перехода которой существенно увеличена из-за наличия между р- и п-областями слоя высокоомного полупроводника с собственной электропроводностью (рис. 5.5).

Практически р-і-п структуру для переключающих СВЧ-диодов формируют на исходном кристалле кремния с проводимостью, близкой к собственной, т.е. либо с небольшой концентрацией акцепторов (π -слой), либо с небольшой концентрацией доноров (ν -слой). Энергетическая диаграмма, распределение примесей, плотность объемного заряда и электрического поля в р-і-п- и р- π -п структурах показаны на рис. 5.5. Методы формирования этих структур различны: сплавление и диффузия примесей, эпитаксиальное наращивание, ионное легирование.

Диоды с р-і-п структурой отличаются меньшей барьерной емкостью, которая к тому же очень слабо зависит от напряжения

(особенно при больших концентрациях примесей в р- и п-областях). Практическая независимость емкости структуры от напряжения оказывается важным свойством переключательных диодов, так как изменение емкости с напряжением может вызвать дополнительные частотные искажения полезного сигнала.

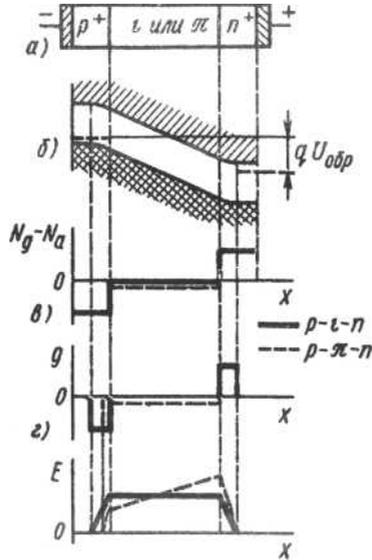


Рисунок 5.5. Диод с р-і-п структурой (а), энергетическая диаграмма (б), распределение примесей (в), плотности объемного заряда (г) и напряженности электрического поля (д)

Пробивное напряжение диодов с р-і-п структурой достигает нескольких сотен вольт, что существенно превышает пробивное напряжение диодов с обычным р-п переходом и с таким же уровнем легирования прилегающих областей.

Для переключательных СВЧ-диодов некоторых марок (2А523А-4 и др.) максимально допустимая мощность, которую может рассеять диод в непрерывном режиме, равна 20 Вт. Такие диоды представляют собой бескорпусные приборы с жесткими выводами — кристаллодержателями — и защитным покрытием. Диаметр их 2 мм, длина 3,6 мм.

Переключательный СВЧ-диод может работать при последовательном и при параллельном включении с линией передачи. В параллельной схеме при прямом смещении диод имеет небольшое сопротивление, шунтирующее линию, и большая часть СВЧ-мощности отражается обратно. Таким образом, при параллельной схеме для переключения СВЧ-тракта используют разницу в отражении, а не в поглощении. В самом диоде при этом поглощается незначительная часть падающей на него СВЧ-мощности, что позволяет относительно маломощному прибору управлять десятками и сотнями киловатт импульсной СВЧ-мощности.

Недостатком переключательных СВЧ-диодов с р-і-n структурой является инерционность процесса рассасывания носителей заряда (электронов и дырок) из і-слоя при переключении диода с прямого на обратное, так как толщина і-слоя может составлять несколько десятков микрометров, а скорость движения носителей заряда ограничена.

Значительно большую скорость переключения можно получить при использовании диодов Шоттки, изготовленных на основе арсенида галлия. Однако уровень переключаемой СВЧ-мощности при этом на несколько порядков ниже, чем при применении переключательных СВЧ-диодов с р-і-n структурой.

5.5. Варикапы

Варикап – это полупроводниковый прибор, представляющий из себя электрически управляемую емкость. Физической основой работы варикапа является зависимость ширины области пространственного заряда, а, следовательно, и барьерной емкости р-n перехода, от напряжения. Поскольку при прямом смещении ЭДП имеет малое сопротивление и потребляет значительную мощность, то для работы варикапов используется только обратная ветвь вольт-фарадной характеристики (ВФХ) (рис. 5.6).

По этой же причине для применения в качестве электрически управляемого конденсатора не используется диффузионная емкость ЭДП и ее зависимость от напряжения.

Параметры варикапа:

1. C_n – номинальная емкость,
2. C_1 – емкость варикапа при $U=0$.

3. C_2 – емкость варикапа при $U=U_{обр,доп}$,
4. $k_c = \frac{C_1}{C_2}$ - коэффициент перекрытия по емкости,
5. Q – добротность варикапа.

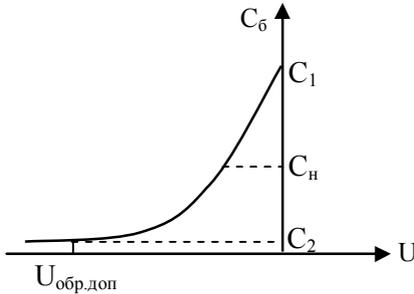


Рисунок 5.6. ВФХ варикапа

Добротность определяет степень энергетических потерь в ЭДП и определяется как величина, равная отношению реактивной мощности $P_{реакт} = \frac{\omega C_б U}{2}$ к мощности потерь $P_{пот}$. Мощность потерь имеет две составляющие: потери в токоведущих элементах P_r и потери в р-п- переходе $P_{р-п}$.

$$P_r = \frac{R_б I^2}{2} = \frac{1}{2} (\omega C_б U)^2 R_б, \quad P_{р-п} = \frac{U^2}{2R_a}.$$

Отсюда добротность варикапа равна

$$Q = \frac{1}{\omega C_б R_б + \frac{1}{\omega C_б R_a}} \quad (5.1.)$$

Как видно из (5.1) в область низких частот, когда мал первый член в знаменателе, добротность равна

$$Q_{нч} = \omega C_б R_a \quad (5.2)$$

На высоких частотах, когда пренебрежимо мал второй член,

добротность определяется как $Q_{\text{оч}} = \frac{1}{\omega C_{\sigma} R_{\sigma}}$ (5.3)

Максимальная добротность, определяемая условием $\frac{dQ}{d\omega} = 0$,

равна $Q_{\text{max}} = \frac{1}{2} \sqrt{\frac{R_a}{R_{\sigma}}}$ (5.4)

Частота, соответствующая максимальной добротности, равна

$$\omega_{Q_{\text{max}}} = \frac{1}{C_{\sigma} \sqrt{R_{\sigma} R_a}} \quad (5.5)$$

6. $\Delta f = f_{\text{max}} - f_{\text{min}}$ - диапазон частот варикапа (рис 5.7).

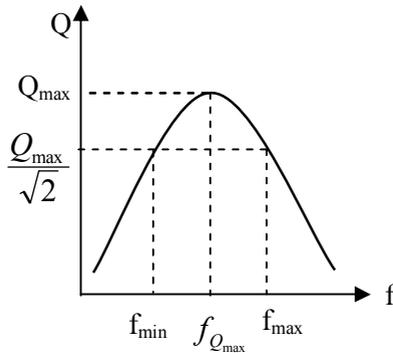


Рисунок 5.7. Зависимость добротности варикапа от частоты

Использовать варикап рекомендуется в диапазоне частот Δf , где добротность максимальна, а потери минимальны. Обозначая $Q_{\text{min}} = \frac{Q_{\text{max}}}{\sqrt{2}}$ можно записать выражение для определе-

ния f_{max} и f_{min} .

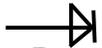
$$\left. \begin{aligned} f_{\text{max}} = \frac{1}{Q_{\text{min}} C_{\sigma} R_{\sigma}} \end{aligned} \right\} \quad (5.6)$$

$$f_{\min} = \frac{Q_{\min}}{C_{\sigma} R_{\sigma}}$$

$$7. TKC = \frac{1}{C_n} \frac{\Delta C_n}{\Delta T} - \text{температурный коэффициент емкости.}$$

$$8. TKQ = \frac{1}{Q_{\max}} \frac{\Delta Q_{\max}}{\Delta T} - \text{температурный коэффициент}$$

добротности.



—  — - условно графическое обозначение варикапа .

Варикапы по функциональному назначению делятся на умножительные и подстроечные.

5.6. Стабилитроны. Лавинно-пролетные диоды

Стабилитроном называют полупроводниковый диод на обратной ветви ВАХ которого имеется участок, где малому изменению напряжения соответствует большой диапазон изменения тока. Физической основой работы стабилитрона является явление электрического (лавинного или тунельного) пробоя. Действительно при электрическом пробое ВАХ на обратной ветви (кривая 1 рис.5.8.) содержит участок, где малому изменению напряжения ΔU соответствует большой диапазон изменения тока ΔI . Это обстоятельство является основой работы стабилитрона.

Поскольку для диодов на основе германия более вероятен тепловой пробой, ВАХ которого имеет вид кривой 2 на рис. 5.9, то на основе германия стабилитроны не изготавливаются. Все стабилитроны производятся на основе кремния.

Параметры стабилитрона (рис. 5.9).

1) U_{cm} - напряжение стабилизации. Напряжение стабилизации однозначно определяется напряжением пробоя р-п перехода.

2) I_{cm}^{\min} - минимальный ток стабилитрона. Это ток на участке перехода от тока насыщения к току пробоя. Обычно величина этого тока составляет 50 – 100 мкА.

3) I_{cm}^{\max} - максимальный ток стабилитрона. Величина этого тока ограничивается предельной плотностью тока и допустимой мощностью рассеивания. Так что $I_{cm}^{\max} = \frac{P_{\max}}{U_{cm}}$.

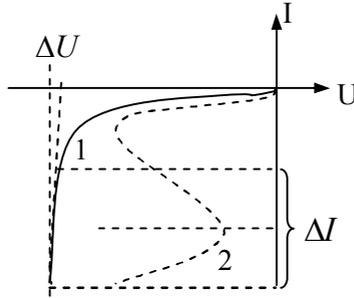


Рисунок 5.8. Обратная ветвь ВАХ диода при электрическом пробое

4) I_{cm}^{cp} -среднее значение тока стабилитрона, как среднearифметическое между I_{cm}^{\max} и I_{cm}^{\min} .

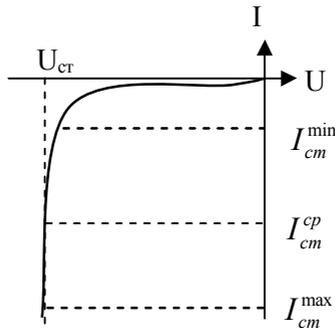


Рисунок 5.9. ВАХ стабилитрона

5) ТКН – температурный коэффициент напряжения стабилизации.

Численное значение ТКН определяется как относительное изменение напряжения стабилизации при изменении температуры

$$TKH = \frac{1}{U_{cm}} \frac{\Delta U_{cm}}{\Delta T}.$$

В зависимости от механизма пробоя ТКН может быть положительным (при лавинном пробое) и отрицательным (при туннельном пробое) (рис. 5.10).

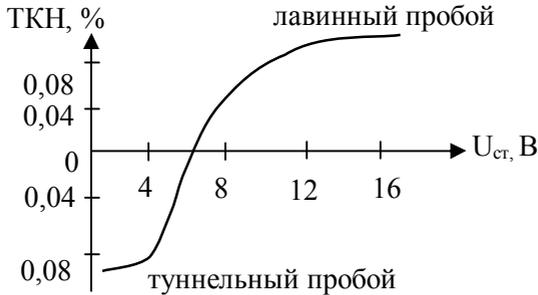


Рисунок 5.10. Зависимость знака ТКН от напряжения стабилизации

6) R_{cm} – статическое сопротивление, определяемое как

$$R_{cm} = \frac{U_{cm}}{I_{cp}}.$$

7) R_{∂} – динамическое (дифференциальное) сопротивление,

определяемое как $R_{\partial} = \frac{\Delta U_{cm}}{\Delta I_{cm}}.$

8) Q – коэффициент качества стабилитрона. Определяется

как $Q = \frac{R_{cm}}{R_{\partial}}.$

Чем больше значение Q , тем лучше стабилитрон выполняет функцию стабилизации напряжения.

 условно-графическое обозначение стабилитрона

Стабилитрон включается в схему, показанную на рисунке 5.11.

Как будет видно из дальнейшего рассмотрения, без ограничительного резистора стабилитрон функцию стабилизации напряжения на нагрузке выполнять не может.

При работе стабилитрона в схеме стабилизации напряжения возможны два варианта:

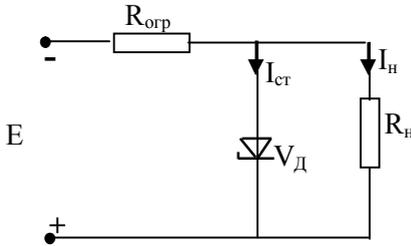


Рисунок 5.11. Схема включения стабилитрона

1) Изменяется входной сигнал на величину ΔE , а необходимо поддерживать неизменным напряжение на нагрузке U_n и ток через нагрузку I_n

2) Входной сигнал не изменяется ($E = \text{const}$), но изменяются параметры нагрузки (R_n). При этом стабилитрон должен поддерживать неизменным напряжение на нагрузке U_n .

Рассмотрим первый случай. В соответствии с уравнением Кирхгофа можно записать

$$E = U_{ст} + R_{огр} (I_{ст} + I_n) \quad (5.7)$$

При увеличении входного сигнала на величину ΔE в силу того, что напряжение $U_{ст}$ изменяется незначительно, то за счет увеличения тока через стабилитрон на величину $\Delta I_{ст}$ скачек входного сигнала погасится на ограничительном резисторе $R_{огр}$.

Так что
$$\Delta E = U_{ст} + R_{огр} (\Delta I_{ст} + I_n) \quad (5.8)$$

Отсюда видна необходимость обязательного наличия в схеме $R_{огр}$. Однако, несмотря на то, что скачек входного сигнала гасится на ограничительном резисторе, главная роль в поддержании неизменного напряжения на нагрузке принадлежит ста-

билитрону. Именно благодаря наличию на ВАХ стабилитрона участка, где малому изменению напряжения соответствует большой диапазон изменения тока, скачек входного сигнала ΔE гасится на $R_{оэп}$ и на нагрузку не влияет.

Величина ограничительного резистора в данном случае определяется соотношением

$$R_{оэп} = \frac{\Delta E - U_{cm}}{\Delta I_{cm} + I_n}, \quad (5.9)$$

$$\text{где } \Delta E = \frac{E_{\max} + E_{\min}}{2}; \quad \Delta I_{cm} = \frac{I_{cm}^{\max} + I_{cm}^{\min}}{2}.$$

Конечно, изменение входного сигнала вызывает изменение напряжения на стабилитроне и, следовательно, на нагрузке. Оценим величину $\Delta U_n = \Delta U_{cm}$. Можно записать, что

$$\Delta E = \Delta U_{cm} + R_{оэп} \left(\frac{\Delta U_{cm}}{R_0} + \frac{\Delta U_{cm}}{R_n} \right)$$

$$\text{Тогда } \Delta U_{cm} = \frac{\Delta E}{1 + \frac{R_{оэп}}{R_0} + \frac{R_{оэп}}{R_n}} \quad (5.10)$$

Из (5.10) видно, что при $\frac{R_{оэп}}{R_0} \gg 1$ $\Delta U_{cm} \ll \Delta E$, т.е. изменение напряжения на стабилитроне всегда много меньше, чем изменение входного сигнала. Таким образом, при выборе $R_{оэп}$ надо всегда выполнять условие: $R_{оэп} \gg R_0$.

Во втором случае входной сигнал не изменяется. Тогда при изменении R_n изменяется ток нагрузки I_n , и, чтобы поддержать неизменным U_{cm} , необходимо, чтобы ток через стабилитрон изменялся в противоположную сторону. Т.е. если I_n растет, то I_{cm} должен уменьшаться на такую же величину. Тогда при $E = const$ U_{cm} остается неизменным.

$$E = U_{cm} + R_{озр} (\Delta I_{cm} + \Delta I_n)$$

Видим, что и во втором случае стабилизация напряжения на нагрузке достигается благодаря способности стабилитрона изменять в широких пределах значение тока. Напряжение же на нем остается практически неизменным.

Значение $R_{озр}$ в этом случае выбирается исходя из следующего выражения:

$$R_{озр} = \frac{E - U_{cm}}{\Delta I_{cm} + \Delta I_n}, \quad (5.11)$$

где $\Delta I_{cm} = \frac{I_{cm}^{max} + I_{cm}^{min}}{2}$, $\Delta I_n = \frac{I_n^{max} + I_n^{min}}{2}$,

$$I_n^{max} = \frac{U_{cm}}{R_n^{min}}, \quad I_n^{min} = \frac{U_{cm}}{R_n^{max}}.$$

Для оценки изменения напряжения на нагрузке в этом случае используется формула

$$\Delta U_{cm} = \frac{R_{озр} \cdot \Delta I_n}{1 + \frac{R_{озр}}{R_o}} \quad (5.12)$$

Для получения большего, чем номинал прибора напряжения U_{cm} стабилитроны включают последовательно, с подключением параллельно стабилитронам выравнивающих резисторов. Для получения меньшего, чем номинал U_{cm} стабилитрон включают последовательно с резистором. Параллельное включение стабилитронов не рекомендуется из-за большого разброса параметров. Однако на практике широко используется многокаскадное включение стабилитронов (рис. 5.12).

При этом коэффициент стабилизации k_{cm} определяемый как

$$k_{cm} = \frac{\Delta E / E}{\Delta U_{cm} / U_{cm}}, \quad \text{будет определяться произведением}$$

коэффициентов стабилизации каждого каскада.

Для компенсации температурного дрейфа напряжения стабилизации прибегают к включению последовательно со

стабили-троном терморезистора, имеющего обратный температурный коэффициент.

Для стабилитронов с положительным ТКН для этой цели обычно используют дополнительный диод, включаемый последовательно со стабилитроном. Такие стабилитроны называют термокомпенсированными.

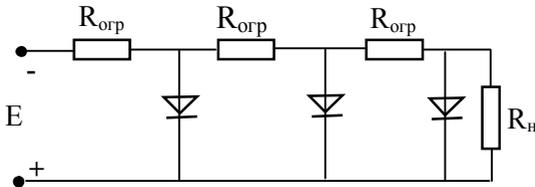


Рисунок 5.12. Многокаскадное включение стабилитронов

Для стабилитронов с положительным ТКН для этой цели обычно используют дополнительный диод, включаемый последовательно со стабилитроном. Такие стабилитроны называют термокомпенсированными.

Для стабилизации напряжений разной полярности используется встречное последовательное включение стабилитронов или выпускаемые промышленностью двуханодные стабилитроны (——).

Стабистор – это полупроводниковый диод, напряжение на котором в области прямого смещения слабо зависит от тока в заданном диапазоне и предназначен для стабилизации малых напряжений (<1 В). Для увеличения напряжения стабилизации применяют последовательное соединение стабисторов. Стабисторы имеют отрицательный ТКН. Основное отличие p-n перехода стабисторов – достаточно высокая степень легирования базы диода.

Стабилитроны и стабисторы имеют очень широкое практическое применение в схемах стабилизации напряжения, формирователей и ограничителей импульсов, схемах защиты и т.д.

По уровню мощности стабилитроны делятся на:

- маломощные < 0,3 Вт;
- средней мощности – 0,3 ÷ 5 Вт;
- мощные > 5 Вт,

а по значению напряжения стабилизации на напряжения до 9 В, от 10 до 99 В, и от 100 до 200 В.

Шумовые диоды

Полупроводниковый шумовой диод — это полупроводниковый прибор, являющийся источником шума с заданной спектральной плотностью в определенном диапазоне частот.

В начальной стадии лавинного пробоя процесс ударной ионизации оказывается неустойчивым: ударная ионизация возникает, срывается, возникает вновь в тех местах р-п перехода, где оказывается в данный момент достаточная напряженность электрического поля. Результатом случайной неравномерности генерации новых носителей заряда при ударной ионизации являются шумы, которые характерны для определенного диапазона токов. При работе таких, например, приборов, как стабилитроны, шумы — явление вредное. Именно поэтому диапазон токов, соответствующий шумам, исключают из диапазона рабочих токов стабилитронов. Однако для различных измерений в радиотехнике нужны генераторы шумовых напряжений.

Таким образом, в качестве генератора шумовых напряжений можно использовать диод в диапазоне обратных токов от минимального $I_{проб. \min}$ до максимального $I_{проб. \max}$ пробивного тока, где наблюдается наибольшая интенсивность электрических флуктуаций. Так, для шумовых диодов КГ401А...КГ401В этот диапазон соответствует значениям токов 10 мкА... 1 мА.

Основными параметрами шумовых диодов являются *спектральная плотность шума* $S_{ш}$ — эффективное значение напряжения шума, отнесенное к 1 Гц, при заданном токе пробоя, а также *границная частота равномерности спектра* $f_{гр}$ — наибольшая частота спектра, при которой удовлетворяется в сторону отрицательного отклонения заданное требование по неравномерности спектральной плотности шума (при заданном токе пробоя).

Одним из справочных параметров шумовых диодов является *средний температурный коэффициент спектральной плотности шума* $TKS_{ш}$ — отношение относительного изменения спектральной

ной плотности шума в заданном диапазоне рабочих температур к абсолютному изменению температуры окружающей среды при постоянном токе:

$$TKS_{ш} = \frac{1}{S_{ш.ср}} \frac{\Delta S}{\Delta T} \Big|_{I_{проб}} = const .$$

Обратный ток, предшествующий лавинному пробое, и пробивное напряжение при лавинном пробое увеличиваются с ростом температуры. В результате участок ВАХ, соответствующий наибольшей интенсивности шумов, смещается с изменением температуры в область больших токов и напряжений. Поэтому знак и значение температурного коэффициента спектральной плотности шума могут быть различными при разных постоянных токах, при которых производят измерение $TKS_{ш}$ шумового диода.

Лавинно-пролетные диоды

Лавинно-пролетный диод – это полупроводниковый диод, работающий в режиме лавинного размножения носителей заряда при обратном смещении электрического перехода и предназначенный для генерации сверхвысокочастотных колебаний.

Генерация электромагнитных СВЧ-колебаний может возникать в диодах с различной структурой. В качестве примера рассмотрим процессы, происходящие в структуре p^+-n-n^+ при обратном напряжении, имеющем постоянную и переменную составляющие. Когда суммарное напряжение превышает пробивное, начинается ударная ионизация - лавинный пробой. Пары электрон-дырка, генерируемые в узкой части $p-n$ перехода вблизи металлургической границы, где напряженность электрического поля достаточна для ударной ионизации, разделяются полем (рис. 5.13). Ток, вызванный движением новых носителей заряда, проходит до тех пор, пока эти носители не выйдут из $p-n$ перехода. За время пролета носителей заряда через переход (в нашем примере – электронов) напряжение на диоде может успеть уменьшиться, если частота переменной составляющей будет большой. Таким образом, из-за конечного времени пролета носителей появляется фазовый сдвиг между проходящим через диод током и прило-

женным к этому диоду переменным напряжением высокой частоты.

Фазовый сдвиг между током и напряжением определяется не только временем пролета, но и инерционностью процесса развития лавины при ударной ионизации. Действительно, момент времени приобретения носителем заряда достаточной для ионизации энергии, вероятнее всего, не совпадает с моментом столкновения этого носителя с одним из атомов полупроводника, т. е. с моментом ионизации.

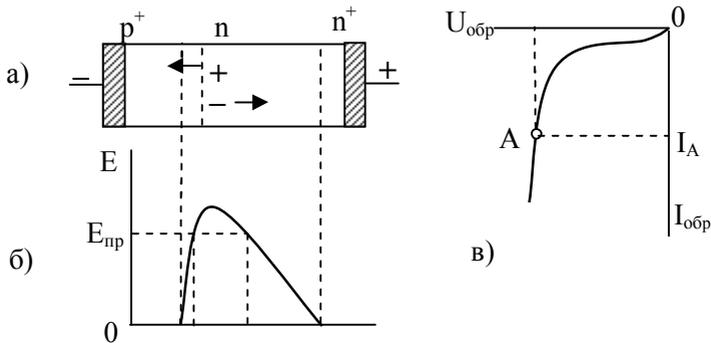


Рисунок 5.13. Структура лавинно-пролетного диода (а), распределение напряженности электрического поля по структуре (б) и положение рабочей точки (постоянного смещения) на ВАХ (в)

Кроме того, для приобретения добавочной энергии также необходимо некоторое время.

Допустим, что время пролета вместе с временем, определяемым инерционностью ударной ионизации, равно половине периода колебаний некоторой частоты переменного напряжения (рис.5.14,а). В этом случае переменный ток через диод будет отставать на полпериода от вызвавшего его переменного напряжения. Нарастание напряжения все время будет сопровождаться уменьшением тока, а уменьшение напряжения, наоборот, ростом тока. Это свидетельствует о том, что для данной частоты переменного напряжения в течение всего периода колебаний выполняется условие отрицательного дифференциального сопротивления.

При понижении частоты переменного напряжения (при увеличении периода колебаний) ток будет отставать от напряжения на угол, меньший 180° , так как время пролета и инерционность ударной ионизации остаются теми же. Когда с понижением частоты переменного напряжения фазовый сдвиг между током и напряжением будет составлять четверть периода, условия отрицательного дифференциального сопротивления будут выполняться на протяжении только половины периода, чередуясь через каждые четверть периода с условиями положительного дифференциального сопротивления (рис. 5.14,б). В этом предельном случае в среднем за период лавинно-пролетный диод не будет обладать отрицательным дифференциальным сопротивлением.

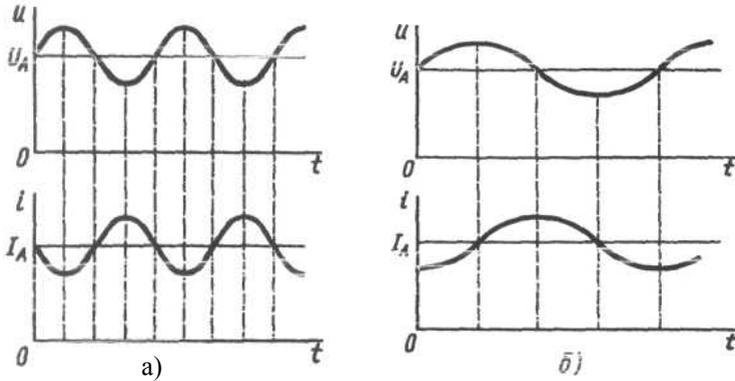


Рисунок 5.14. Зависимости напряжений и токов, иллюстрирующие появление отрицательного дифференциального сопротивления лавинно-пролетного диода:

а) при сдвиге фаз 180° , возникающем на больших частотах переменной составляющей напряжения, б) при сдвиге фаз 90° , соответствующем меньшим частотам переменной составляющей напряжения и отсутствию отрицательного дифференциального сопротивления в среднем за период

Аналогично с повышением частоты переменного напряжения отрицательное дифференциальное сопротивление пропадает при фазовом сдвиге между током и напряжением, достигающем 270° . Таким образом, лавинно-пролетные диоды обладают отрицатель-

ным дифференциальным сопротивлением только для СВЧ-колебаний.

Всякий прибор с отрицательным дифференциальным сопротивлением может быть использован для генерации и усиления электромагнитных колебаний. Лавинно-пролетные диоды используют для генерации СВЧ-колебаний большой мощности. При этом не обязательно подавать на лавинно-пролетный диод переменное напряжение необходимой частоты, что мы мысленно проделали при рассмотрении принципа действия прибора. Лавинно-пролетный диод вместе с резонансной камерой, в которую его обычно помещают, способен выделить из импульсов, возникающих во время подачи постоянного смещения, и усилить колебания определенной частоты.

Лавинно-пролетные диоды впервые были созданы в СССР на основе открытого А.С.Тагером с сотрудниками эффекта генерации и усиления СВЧ-колебаний при лавинном пробое.

Кроме рассмотренного *лавинно-пролетного режима* работы, который в английской литературе называют IMPATT-режимом (сокращение от impact ionization avalanche transit time), лавинно-пролетные диоды могут работать и в *режиме с захваченной плазмой* или TRAPATT-режиме (trapped plasma avalanche triggered transit).

Принцип действия при этом режиме работы связан с тем, что скорость перераспределения электрического поля в структуре диода может значительно превышать скорость дрейфа носителей заряда. На рис. 5.15 показано распределение напряженности электрического поля в слаболегированной n -области p^+-n-n^+ структуры лавинно-пролетного диода в различные моменты времени после включения диода на обратное напряжение, превышающее пробивное напряжение.

В первый момент (t_1) напряженность электрического поля максимальна около металлургической границы. Именно здесь из-за ударной ионизации начинается образование электронно-дырочной плазмы. Это приводит к перераспределению электрического поля в n -области.

В следующий момент времени (t_2) ударная ионизация будет происходить в соседнем слое n -области. Скорость

дрейфа носителей заряда ограничена даже в сильных электрических полях скоростью насыщения

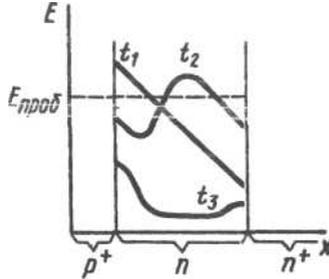


Рисунок 5.15. Распределение напряженности электрического поля в слаболегированной n -области p - n -перехода в различные моменты времени при работе лавинно-пролетного диода в режиме с захваченной плазмой

Скорость дрейфа электронов плазмы может оказаться значительно меньше скорости насыщения, если напряженность электрического поля в слое с плазмой успеет уменьшиться. Перераспределение напряженности электрического поля может произойти быстрее, если источник питания диода обеспечит большую плотность тока через диод (с учетом плотности тока смещения), и если концентрация примесей в слаболегированной области будет достаточно мала.

В результате фронт волны ионизации быстро пересекает всю n -область, которая заполняется высокопроводящей электронно-дырочной плазмой. Напряженность электрического поля в это время (t_3 на рис.5.14) и напряжение на диоде становятся малыми, что приводит к относительно медленному рассасыванию носителей плазмы из p - n перехода. Задержка экстракции носителей из p - n перехода обусловила название «режим с захваченной плазмой».

Так как скорость направленного движения носителей заряда в лавинно-пролетных диодах в режиме с захваченной плазмой значительно ниже скорости насыщения, то частота генерируемых колебаний обычно не превышает 10 ГГц, в то время как при лавинно-пролетном режиме эта частота может составлять несколько сотен гигагерц. Другие отличия в свойствах и параметрах при различных режимах работы вызваны тем, что при

лавинно-пролетном режиме уменьшение скорости дрейфа ниже скорости насыщения нежелательно, а при режиме с захваченной плазмой — наоборот. Поэтому большая амплитуда колебаний может быть получена в режиме с захваченной плазмой — до нескольких сотен киловатт при импульсной работе (при непрерывной работе — до нескольких ватт). А так как напряжение на диоде оказывается малым при больших токах и, наоборот, большим при малых токах, то КПД достигает 40% для диодов из арсенида галлия и кремния.

Лавинно-пролетные диоды имеют высокий уровень шума, присущий ударной ионизации, так как небольшие беспорядочные отклонения тока (дробовой шум) через электрический переход усиливаются при ударной ионизации в число раз, равное коэффициенту лавинного размножения M . Поэтому лавинно-пролетные диоды применяют только для генерации СВЧ-колебаний и не используют для усиления слабых сигналов. Поэтому же лавинный пробой используют для создания шумовых диодов.

5.7. Туннельные и обращенные диоды

Туннельный диод — это полупроводниковый диод, на прямой ветви ВАХ которого имеется участок отрицательного дифференциального сопротивления N-типа. Отличительной особенностью p-n перехода туннельного диода является то, что он изготовлен на основе сильнолегированных (вырожденных) полупроводников.

Физической основой работы туннельного диода является эффект туннелирования носителей через очень узкую область пространственного заряда.

При высокой степени легирования ($>10^{19}$ см⁻³) полупроводники становятся вырожденными, в них уровень Ферми заходит в зоны разрешенных энергий (рис. 5.16). При этом все уровни между E_F и E_C в n-полупроводнике — заняты, а уровни между E_V и E_F в p-полупроводнике — свободны.

На рисунке 5.17 приведены зонная диаграмма ЭДП туннельного диода в состоянии термодинамического равновесия.

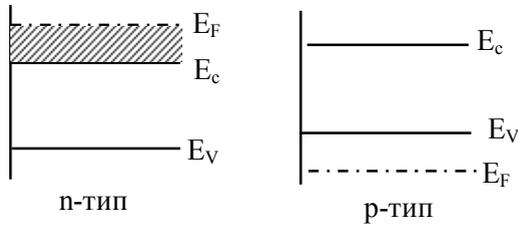


Рисунок 5. 16. Зонные диаграммы вырожденных полупроводников n- и p-типов

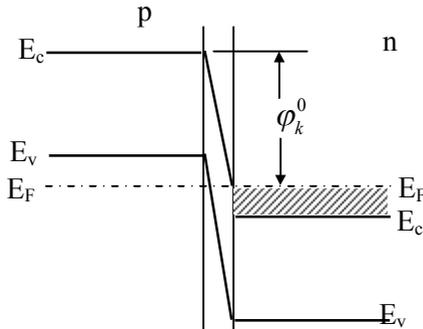


Рисунок 5.17. Зонная диаграмма p-n-перехода туннельного диода

В силу высокой степени легирования высота потенциального барьера φ_k^0 велика (больше ширины запрещенной зоны). Поэтому переход основных носителей через этот барьер невозможен, а количество неосновных носителей ничтожно мало. Обмен носителями между p- и n-областями за счет туннелирования равновероятен, так что ток через диод при $U = 0$ также равен нулю. При приложении обратного смещения (рис.5.18) возникает ситуация аналогичная туннельному пробую.

Из-за наличия сильного электрического поля и малого зазора между зонами электроны, находящиеся в интервале энергий ΔW , туннелируют из p- в n-область, создавая достаточно большой ток, т.е. при обратном смещении через туннельный диод протекает большой ток.

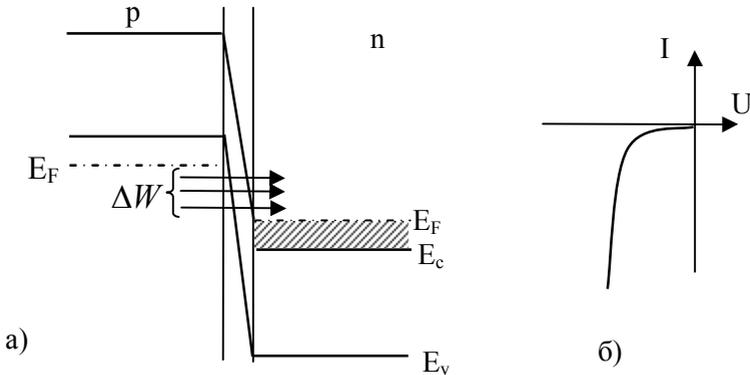


Рисунок 5.18. Зонная диаграмма ЭДП туннельного диода (а) и обратная ветвь ВАХ (б) при обратном смещении перехода.

Вид прямой ветви ВАХ туннельного диода поясняет рис. 5.19, а - д.

В диоде без внешнего напряжения существует туннелирование электронов из n-области в p-область и обратно. Встречные потоки электронов равны, поэтому суммарный ток через диод равен нулю (рис. 5.19,а).

При небольшом прямом напряжении на туннельном диоде происходит уменьшение высоты потенциального барьера p-n перехода или смещение энергетической диаграммы n-области относительно энергетической диаграммы p-области. Свободные энергетические уровни p-области (занятые дырками), расположенные непосредственно над уровнем Ферми, оказываются на одной высоте по энергетической диаграмме или при одних и тех же значениях с энергетическими уровнями n-области, занятыми электронами (рис. 5.19,б). Поэтому будет происходить преимущественное туннелирование электронов из n-области в p-область.

При прямом напряжении на диоде, когда свободные энергетические уровни валентной и примесной зон p-области окажутся на одной высоте с занятыми электронами энергетическими уровнями зоны проводимости и примесной зоны n-области, туннельный ток через диод будет максимальным (рис. 5.19,в).

При дальнейшем увеличении прямого напряжения на диоде туннельный ток через диод будет уменьшаться, так как из-за смещения энергетических диаграмм будет уменьшаться количество электронов, способных туннелировать из n-области в p-область (рис. 5.19,г).

Туннельный ток через диод окажется равным нулю при некотором еще большем прямом напряжении, когда из-за относительного смещения энергетических диаграмм n- и p-областей для свободных электронов n-области не будет свободных энергетических уровней в p-области (рис.5.19,д).

Однако при этом через диод будет проходить прямой ток, обусловленный переходом носителей заряда через понизившийся потенциальный барьер p-n перехода, т.е. ток, связанный с инжекцией.

Если в запрещенной зоне p- и n-областей имеются локализованные уровни, то форма прямой ветви вольтамперной характеристики будет иметь более сложный вид. В этом случае при $U \geq U_{\min}$ возможно туннелирование электрона из зоны проводимости на локальный уровень (рис.5.20, переход 1), а оттуда электрон, потеряв энергию за счет взаимодействия с колебаниями атомов кристаллической решетки, перейдет в валентную зону p-области. Возможен и другой путь перехода: вначале электрон из зоны проводимости n-области переходит на локализованный уровень (рис.5.20, переход 2), а затем туннелирует в валентную зону p-области. При достаточно большой концентрации локализованных уровней различных типов переход из одной разрешенной зоны в другую может осуществляться по «лестнице» (рис.5.20, путь 3).

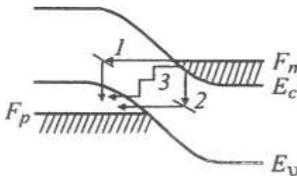


Рисунок 5.20. Различные пути туннелирования через локализованные состояния в запрещенной зоне

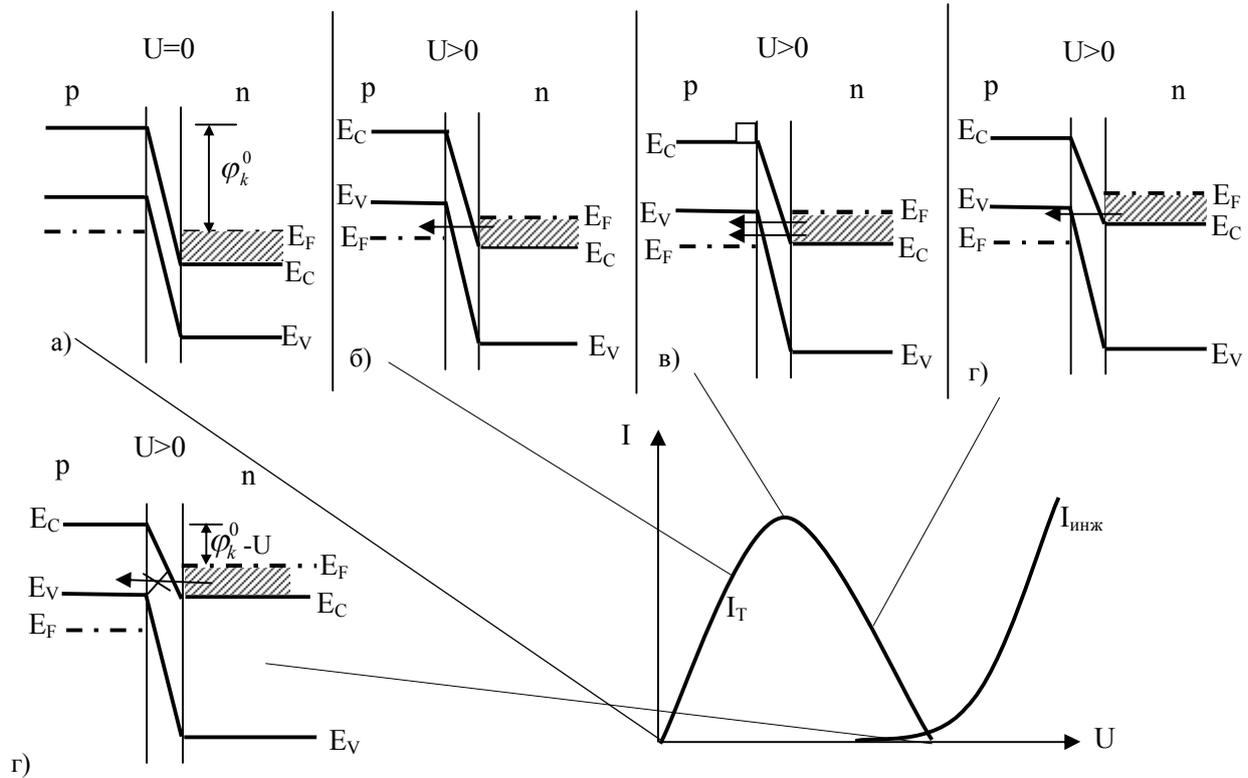


Рисунок 5.19. Зонная диаграмма ЭДП туннельного диода при различных значениях прямых смещений

В соответствии с вышеизложенным наличие мелких локализованных состояний вблизи E_c и E_v приводит к появлению избыточного тока при $U \geq U_{\min}$.

Таким образом, туннельный диод обладает отрицательным дифференциальным сопротивлением в некотором диапазоне прямых напряжений. Это и является самым интересным свойством туннельного диода, так как всякий прибор с отрицательным дифференциальным сопротивлением может быть использован для генерации и усиления электромагнитных колебаний, а также в переключающих схемах.

Параметры

Промышленность выпускает туннельные диоды из арсенида галлия и из германия. Их вольтамперные характеристики показаны на рис. 5.21. Из рисунка видно, что чем больше ширина запрещенной зоны исходного полупроводника, тем при больших напряжениях наблюдается отрицательное дифференциальное сопротивление.

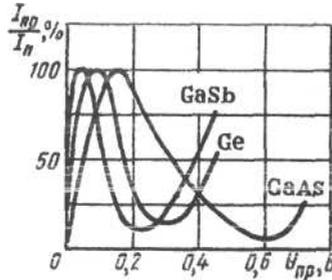


Рисунок 5.21. ВАХ туннельных диодов из различных полупроводниковых материалов

Туннельные диоды характеризуются специфическими параметрами:

1) Пиковый ток I_{Π} — прямой ток в точке максимума ВАХ, при котором значение $\frac{dI}{dU}$ равно нулю. Этот ток различен для туннельных диодов разного назначения. Значение его может составлять от десятых долей миллиампера до сотен миллиампер.

2) Ток впадины I_B — прямой ток в точке минимума ВАХ, при котором значение $\frac{dI}{dU}$ равно нулю.

3) Отношение токов туннельного диода $\frac{I_{\Pi}}{I_B}$ — отношение пикового тока к току впадины. Для туннельных диодов из арсенида галлия $\frac{I_{\Pi}}{I_B} \geq 10$, для германиевых туннельных диодов

$$\frac{I_{\Pi}}{I_B} = 3 \div 6.$$

4) Напряжение пика U_{Π} — прямое напряжение, соответствующее пиковому току. Для туннельных диодов из арсенида галлия $U_{\Pi} = 100 \div 150$ мВ, для германиевых $U_{\Pi} = 40 \div 60$ мВ.

5) Напряжение впадины U_B — прямое напряжение, соответствующее току впадины. У туннельных диодов из арсенида галлия $U_B = 400 \div 500$ мВ, у германиевых $U_B = 250 \div 350$ мВ.

6) Напряжение раствора U_{pp} — прямое напряжение, большее напряжения впадины, при котором ток равен пиковому.

7) Удельная емкость туннельного диода $\frac{C_{\Delta}}{I_n}$ — отношение

емкости туннельного диода к пиковому току.

8) Предельная резистивная частота f_R — расчетная частота, на которой активная составляющая полного сопротивления последовательной цепи, состоящей из р-п перехода и сопротивления потерь, обращается в нуль.

9) Резонансная частота туннельного диода f_0 — расчетная частота, на которой общее реактивное сопротивление р-п перехода и индуктивности корпуса туннельного диода обращается в нуль.

Температурные зависимости параметров

Изменение температуры туннельного диода может по-разному влиять на туннельную составляющую тока и на составляющую, связанную с инжекцией.

На температурную зависимость туннельной составляющей тока могут влиять следующие физические факторы:

1. С повышением температуры уменьшается ширина запрещенной зоны арсенида галлия и германия — основных исходных полупроводниковых материалов для туннельных диодов. Уменьшение ширины запрещенной зоны приводит к уменьшению толщины потенциального барьера, сквозь который туннелируют электроны, при этом вероятность туннелирования растет. Туннельная составляющая тока и, в частности, пиковый ток увеличиваются.

2. При увеличении температуры изменяется распределение электронов по энергетическим уровням - количество электронов под уровнем Ферми в зоне проводимости n-области уменьшается, так как часть свободных электронов переходит на более высокие энергетические уровни, а уровень Ферми смещается вниз. Поэтому уменьшается число электронов, которые могут туннелировать из n-области в p-область. Туннельная составляющая прямого тока уменьшается.

Так как эти факторы действуют, так сказать, в разные стороны, то суммарное их влияние, во-первых, должно быть малым, а во-вторых, оно может привести как к увеличению, так и к уменьшению пикового тока туннельного диода с увеличением температуры.

Инжекционная составляющая тока туннельного диода растет с увеличением температуры по двум причинам, имеющим место и в выпрямительных диодах: уменьшение высоты потенциального барьера и перераспределение носителей заряда по энергетическим уровням. Поэтому у туннельного диода ток впадины растет с увеличением температуры.

Частотные свойства туннельных диодов

Механизм действия туннельных диодов связан с туннелированием электронов сквозь потенциальный барьер. Время, необходимое для завершения этого процесса, составляет 10^{-13} – 10^{-14} с. Эффекта накопления неосновных носителей в базе туннельных диодов практически нет, так как они используются при малых напряжениях соответствующих падающему участку ВАХ (с

отрицательным дифференциальным сопротивлением). Поэтому туннельные диоды способны работать на частотах до сотен гигагерц, что соответствует миллиметровому диапазону радиоволн. Верхний предел частотного диапазона работы туннельных диодов ограничен лишь паразитными реактивностями - собственной емкостью, основу которой составляет барьерная емкость р-п перехода, и индуктивностью выводов и корпуса.

Для анализа и расчета параметров, характеризующих частотные свойства туннельных диодов, воспользуемся эквивалентной схемой туннельного диода для малого переменного сигнала при наличии постоянного напряжения, которое смещает рабочую точку на падающий участок ВАХ. Эквивалентная схема туннельного диода отличается от эквивалентной схемы обычного диода только тем, что здесь вместо активного сопротивления перехода введено отрицательное дифференциальное сопротивление r_- и учитывается индуктивность выводов L (рис.5.22).

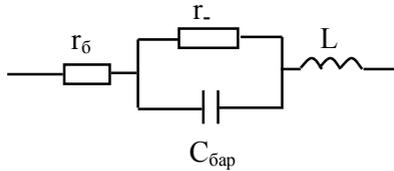


Рисунок 5.22. Эквивалентная схема туннельного диода

Полное сопротивление туннельного диода при малом синусоидальном напряжении

$$\begin{aligned} \dot{z} = r_{\delta} + j\omega L + \frac{1}{g + j\omega C_{\text{бар}}} = r_{\delta} + \frac{g}{g^2 + \omega^2 C_{\text{бар}}^2} \\ + j \left(\omega L - \frac{\omega C_{\text{бар}}}{g^2 + \omega^2 C_{\text{бар}}^2} \right), \end{aligned} \quad (5.13)$$

где $g = 1/r_-$ - отрицательная дифференциальная проводимость туннельного диода.

Отрицательное дифференциальное сопротивление у туннельного диода будет существовать, если вещественная часть полного сопротивления будет меньше нуля, т. е. при

$$r_{\bar{o}} < \left| \frac{g}{g^2 + \omega_2 C_{\bar{o}ap}^2} \right| \quad (5.14)$$

При большой частоте второе слагаемое в вещественной части (5.13) окажется меньше (по абсолютному значению) первого слагаемого. Это значит, что туннельный диод при больших частотах не будет обладать отрицательным дифференциальным сопротивлением. Определим *предельную резистивную частоту*, на которой активная составляющая полного сопротивления туннельного диода обращается в нуль, т. е. неравенство (5.14) обращается в равенство $r_{\bar{o}} = \frac{g}{g^2 + \omega_R^2 C_{\bar{o}ap}^2}$.

$$\text{Отсюда } \omega_R = \frac{1}{|r_{-}| C_{\bar{o}ap}} \sqrt{\frac{|r_{-}|}{r_{\bar{o}}}} - 1. \quad (5.15)$$

Из соотношения (5.15) видно, что предельная резистивная частота туннельного диода зависит от соотношения сопротивления базы диода и абсолютного значения отрицательного дифференциального сопротивления. Для нахождения экстремальных значений предельной резистивной частоты надо продифференцировать (5.15) по $|r_{-}|$ и приравнять полученное выражение нулю. Тогда получим условие

$$|r_{-}| = 2r_{\bar{o}}, \quad (5.16)$$

при выполнении которого предельная резистивная частота будет максимальной. Связь ее с параметрами эквивалентной схемы можно получить, подставив (5.16) в (5.15):

$$f_R = \frac{1}{2\pi |r_{-}| C_{\bar{o}ap}} \quad (5.18)$$

Усредненное отрицательное дифференциальное сопротивление можно представить в виде

$$|r_-| = \frac{\Delta U}{\Delta I} = \frac{U_\epsilon - U_\Pi}{I_\Pi - I_\epsilon} \approx \frac{const}{I_\Pi}.$$

Если полученное соотношение подставить в (5.18), то получим

$$f_R \approx I_\Pi / C_{\delta ap}.$$

Поэтому отношение емкости туннельного диода к пиковому току является одним из основных параметров туннельного диода, который характеризует его частотные свойства.

В туннельном диоде, имеющем паразитную индуктивность и емкость, при некоторой частоте могут возникнуть условия для резонанса, что нарушит нормальную работу схемы с туннельным диодом. Поэтому резонансная частота туннельного диода f_0 является еще одним параметром, характеризующим частотные свойства туннельного диода.

Резонансную частоту туннельного диода можно определить из соотношения (5.13) при условии равенства нулю мнимой части полного сопротивления диода:

$$\omega_0 L = \frac{\omega_0 C_{\delta ap}}{g^2 + \omega_0^2 C_{\delta ap}^2},$$

$$\text{тогда } \omega_0 = \frac{1}{|r_-| C_{\delta ap}} \sqrt{\frac{|r_-|^2 C_{\delta ap}}{L} - 1}. \quad (5.19)$$

Туннельные диоды должны быть сконструированы так, чтобы резонансная частота была выше предельной истинной частоты ($\omega_0 > \omega_R$), тогда возможные нежелательные резонансные явления возникают только на частотах, на которых туннельный диод уже не будет обладать отрицательным дифференциальным сопротивлением. Неравенство ($\omega_0 > \omega_R$) с учетом соотношений (5.19) и (5.15) преобразуем следующим образом: $L < r_\delta |r_-| C_{\delta ap}$.

Следовательно, индуктивность эквивалентной схемы туннельного диода, которая определяется в основном индуктивностью внутренних и внешних выводов диода, должна быть по возможности малой. Поэтому внутренние и внешние

выводы у туннельных диодов делают не из тонкой проволоки, а из плющенко или в виде мембран, пластин и т.п.

ВАХ туннельного диода

Исходя из качественного рассмотрения характера физических процессов в ТД, можно записать, что

$$I_{ТД} = I_T + I_{узб} + I_{инж}.$$

Выражения для тока $I_{инж} = I_s \exp \frac{qU}{kT}$, а для

$$I_{узб} = I_{\min} \exp \frac{A}{kT} (U - U_{\min})$$
 - это выражение записано, исходя

из общей теории туннелирования через локальные центры.

A – некоторая константа. Действительно, если $U = U_{\min}$, то

$I_{узб} = I_{\min}$, а далее с ростом U $I_{узб}$ (как это установлено экспериментом) растет по экспоненте.

Определим выражение для оценки I_T :

Ток I_T можно считать равным $I_T = I_{n \rightarrow p} - I_{p \rightarrow n}$, причем величины этих токов определяются количеством туннелировавших частиц:

$$dN_{n \rightarrow p} = C \cdot W \cdot q_n(E) \cdot f_n(E) \cdot q_p(E) \cdot (1 - f_p(E)) dE,$$

которое определяется количеством электронов в интервале $\Delta E_n - q_n(E) f_n(E)$, числом свободных состояний в p-области $\Delta E_p - q_p(E) (1 - f_p(E))$ и прозрачностью барьера W

$$W = \exp \left[- \frac{h(2m_n^*)^{1/2}}{3q\pi E} \Delta E^{3/2} \right],$$

где E - напряженность поля в p-n переходе

ΔE - высота потенциального барьера, равная $\Delta E = \Delta E_n + \Delta E_p$

$q_n(E), q_p(E)$ - плотность состояний в зоне ΔE_n и ΔE_p ;

$f_{n,p}(E)$ - функция Ферми для электронов и дырок;

C – коэффициент пропорциональности.

$dN_{p \rightarrow n} = C \cdot W \cdot q_p(E) \cdot f_p(E) \cdot q_n(E)(1 - f_n(E))dE$. Домножая dN на заряд и интегрируя в интервале энергий, соответствующих перекрытию зон, получим I_T , полагая $W_n = W_p = W$,

$$I_T = qCS \int_{E_c}^{E_v} W [f_n(E) - f_p(E)] \cdot q_n(E) \cdot q_p(E) dE,$$

где C – константа, S – площадь p-n перехода.

Интеграл для I_T берется при следующих приближениях:

- 1) Прозрачность барьера W не зависит от E .

2) Степень вырождения полупроводника такова, что допускает линейную аппроксимацию функции Ферми-Дирака ($\Delta E_n, \Delta E_p \approx kT$). Разлагая $f_{n,p}(E)$ в ряд Тейлора, можно записать, используя два члена разложения.

$$f_{n,p} \approx \frac{1}{2} + \frac{(E_{F_{n,p}} - E)}{4kT}.$$

3) Плотность состояний в зонах зависит от энергии, как $(E - E_c)^{1/2}$ и $(E_v - E)^{1/2}$. С учетом этих приближений, а также беря за начало отсчета E_c , т.е. полагая $E_c = 0$, запишем:

$$I_T = CS \frac{q^2 W U}{4kT} \int_0^{E_v} E^{1/2} (E_v - E)^{1/2} dE$$

и это дает после интегрирования (с учетом того, что

$$\Delta E_n + \Delta E_p = qU_{\min}) \quad I_T = D \frac{qU}{kT} [(\Delta E_n + \Delta E_p) - qU]^2,$$

где $D = CS \frac{qW}{4}$ - константа. Это уравнение хорошо описывает ВАХ

ТД при $U < U_{\min}$.

Напряжение минимума ВАХ – полное отсутствие перекрытия зон ($\Delta E_n, \Delta E_p$).

$$U_{\min} = \frac{1}{q} (\Delta E_n + \Delta E_p)$$

Напряжение U_{\max} определяется по условию $\frac{dI}{dU} = 0$ и равно

$$U_{\max} = \frac{\Delta E_n + \Delta E_p}{3q}$$

$$I_{TD} = D \frac{qU}{kT} [(\Delta E_n + \Delta E_p) - qU]^2 + I_{\min} \exp \frac{A}{kT} (U - U_{\min}) + I_s \exp \frac{qU}{kT}$$

ВАХ ТД во всем диапазоне напряжений.

Применение туннельных диодов

По назначению туннельные диоды делятся на следующие основные группы: усилительные, генераторные, переключающие.

Ниже приведены схемы, поясняющие работу туннельных диодов.

Туннельный диод в режиме усиления. Если положение рабочей точки выбрать на падающем участке вольтамперной характеристики, туннельный диод будет обладать отрицательным дифференциальным сопротивлением - R_d . Это свойство прибора широко используется для создания усилителей с относительно низким уровнем собственных шумов, успешно работающих в различных частотных диапазонах. Особенно перспективным является применение таких усилителей в диапазоне СВЧ.

На рис.5.23,а приведена упрощенная эквивалентная схема усилителя на туннельном диоде.

Допустим, что ключ S1 замкнут (сопротивление R_d из схемы исключено), тогда максимальная мощность, отдаваемая в нагрузку при условии $R_H = R_G$, будет равна

$$P_H = \frac{U_{ex}^2}{4R_H} \quad (5.20)$$

Разомкнем ключ. Мощность, выделяющаяся на R_H составит

$$P_{H \max} = \left[\frac{U_{ex}}{R_H + R_G + R_d} \right]^2 R_H \quad (5.21)$$

Из выражения (5.21) видно, что мощность в нагрузке при включении отрицательного сопротивления возрастает. При $|-R_d| = R_H = R_G$

$$P_{H \max} = \frac{U^2}{R_H} = 4P_H \quad (5.22)$$

На рис. 5.23,б приведены графические построения, при помощи которых можно проанализировать работу «последовательного» (R_H и R_d включены последовательно) усилителя на туннельном диоде.

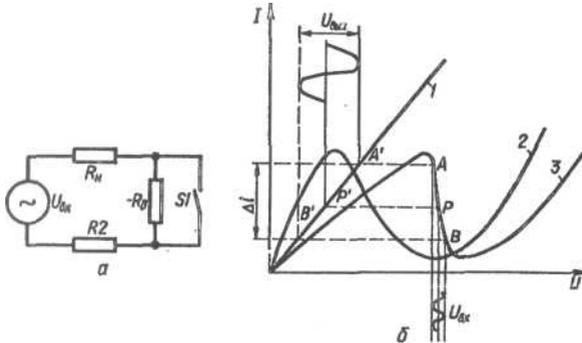


Рисунок 5.23. Упрощенная эквивалентная схема (а) ($R_{г}$ - внутреннее сопротивление генератора напряжения $U_{ВХ}$; $R_{н}$ - сопротивление нагрузки; $R_{д}$ - отрицательное дифференциальное сопротивление туннельного диода) и графическое пояснение работы (б) усилителя на туннельном диоде (1 - вольтамперная характеристика нагрузки; 2 - вольтамперная характеристика туннельного диода; 3 - суммарная вольтамперная характеристика схемы).

При выборе рабочей точки P на падающем участке суммарной вольтамперной характеристики схемы небольшие изменения входного напряжения $U_{ВХ}$ приводят к значительным изменениям тока в цепи и напряжения на выходе (на сопротивлении $R_{н}$).

Большинство практических схем усилителей на туннельных диодах - резонансного типа. Одна из возможных схем такого усилителя приведена на рис. 5.24. Колебательный контур здесь образован катушкой индуктивности L_1 и емкостью $C_{д}$ диода. Батарея G_1 совместно с резисторами $R_{б}$ и R_1 задает рабочую точку по постоянному току. Дроссель L_6 блокирует источник питания по переменному току.

Современные усилители на туннельных диодах работают до частот порядка десятков гигагерц ($\lambda \approx 1 \text{ см}$).

Туннельный диод в схеме генератора. Генераторы гармонических колебаний на туннельных диодах отличаются простотой схемы, малыми габаритами и массой, высокой экономичностью.

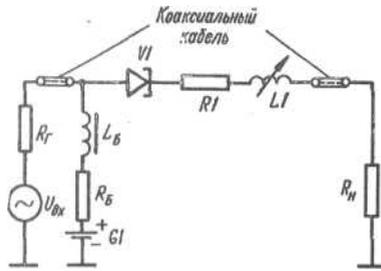


Рисунок 5.24. Принципиальная схема усилителя на туннельном диоде

Наиболее полно преимущества туннельного диода удается использовать в генераторах диапазона СВЧ, особенно на частотах выше 1 ГГц, где из-за невозможности использования транзисторов приходилось использовать сложные, громоздкие и неэкономичные генераторы на клистронах, лампах бегущей и обратной волны и др. Современные туннельные диоды позволяют генерировать электрические колебания с частотами до 100 ГГц.

Применение туннельного диода в схемах генераторов объясняется тем, что с помощью отрицательного сопротивления туннельного диода можно компенсировать потери в колебательном контуре и получить в нем незатухающие колебания.

Простейшая схема генератора на туннельном диоде приведена на рис.5.25.

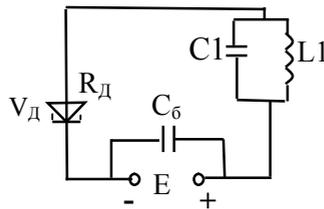


Рисунок 5.25. Схема генератора на туннельном диоде

Туннельный диод как переключатель. Переключающей схемой называется такая схема, электрическое состояние которой резко изменяется под воздействием относительно ма-

лого электрического сигнала. В основе большинства переключающих схем на туннельных диодах лежит цепь, представляющая собой последовательное соединение туннельного диода с другими элементами. Пример такой цепи показан на рис. 5.26, а. В отсутствие переменного входного сигнала под действием внешнего напряжения E в цепи диода течет постоянный ток I_0 , а на диоде действует постоянное напряжение $U_1 = E - I_0 R1$. При подаче переменного сигнала в зависимости от его полярности ток в цепи будет либо уменьшаться на величину i , либо увеличиваться на ту же величину. При токе $(I_0 - i)$ на диоде будет действовать напряжение U_2 (рис. 5.26), а при токе $(I_0 + i)$ напряжение будет равно U_3 , затем по мере уменьшения тока напряжение на диоде упадет до U_4 и потом скачком изменится до U_1 . В результате при отрицательном напряжении сигнала напряжение на диоде (и на выходе) будет равно U_2 , а при положительном значении - U_3 , причем $U_3 \gg U_2$.

Эта резкая разница между величинами напряжений на выходе и позволяет рассматривать туннельный диод как прибор с двумя устойчивыми состояниями, т.е. электронный ключ.

Переход из одного устойчивого состояния в другое совершается очень быстро – за время порядка $10^{-9} \div 10^{-8}$ с, в связи с чем туннельные диоды широко используются в быстродействующих вычислительных машинах в схемах триггеров, заминающих ячеек, логических элементов и т.д.

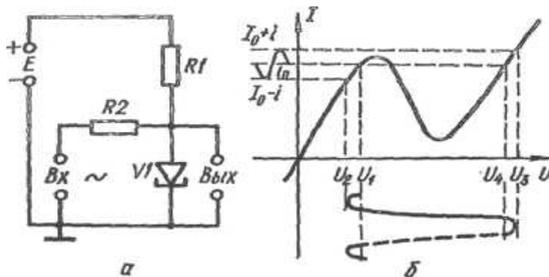


Рисунок 5.26. Применение туннельного диода в режиме ключа: а – схема, б – графическое пояснение работы

Обращенные диоды

Обращенным называют диод на основе полупроводника с критической концентрацией примесей, в котором проводимость при обратном напряжении вследствие туннельного эффекта значительно больше, чем при прямом напряжении.

При концентрациях примесей в р- 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.5.27, а. Уровень Ферми при такой средней концентрации примесей может быть расположен на потолке валентной зоны р-области и на дне зоны проводимости n-области диода, т.е. потолок валентной зоны р-области и дно зоны проводимости n-области при нулевом смещении на диоде находятся на одной высоте по энергетической диаграмме.

Обратная ветвь ВАХ обращенного диода аналогична обратной ветви ВАХ туннельного диода, так как при обратных напряжениях происходит туннелирование электронов из валентной зоны р-области в зону проводимости n-области. Поэтому обратные токи в обращенных диодах оказываются большими при ничтожно малых обратных напряжениях (десятки милливольт).

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 р-n перехода.

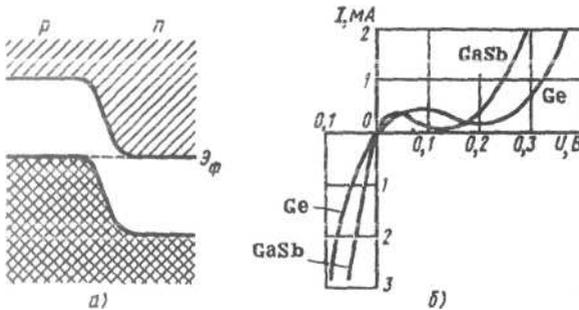


Рисунок 5.27. Энергетическая диаграмма (а) и ВАХ (б) об

Но заметная инжекция может наблюдаться только при прямых напряжениях в несколько десятых долей вольта. При меньших напряжениях прямые токи в обращенных диодах оказываются меньше обратных (рис. 5.27, б).

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 основных 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 р-п переходу областях обращенные диоды оказываются мало чувствительными к воздействиям проникающей радиации.

5.8. Фотодиоды

Фотодиод — это фотогальванический приемник излучения без внутреннего усиления, фоточувствительный элемент которого содержит структуру полупроводникового диода.

Принцип действия фотодиода. Действие фотодиода основано на фотогальваническом эффекте — появлении тока при включении фотодиода в электрическую цепь в результате разделения фотоносителей электрическим полем ОПЗ р-п перехода.

Поглощение квантов оптического излучения в р-п переходе сопровождается генерацией электронно-дырочных пар. Если излучение поглощается в квазинейтральных п- или р-областях на расстоянии, меньшем диффузионной длины, то электроны и дырки достигают р-п-перехода. Потенциальный барьер перехода способствует перемещению электронов из р-области в п-область и дырок в обратном направлении. Если излучение поглощается в ОПЗ, то электроны переносятся в п-область, а дырки в р-область. В рабочем состоянии к фо-

тодиоду приложено напряжение в обратном направлении (рис.5.28).

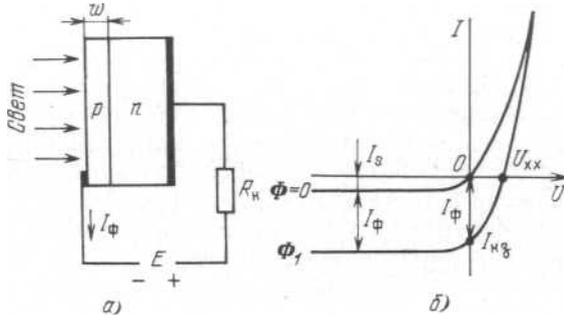


Рисунок 5.28. Схема включения фотодиода (а) и его ВАХ (б)

Воздействие света вызывает протекание помимо темного тока, равного току насыщения для узкозонных полупроводников (InSb, Ge), фототока I_ϕ , который по направлению совпадает с током насыщения. Так как из n- и p-областей в ОПЗ уходят неосновные носители заряда, то они увеличивают обратный ток перехода, ибо именно он образуется за счет неосновных носителей.

При прямом направлении напряжения электрическое поле в ОПЗ не изменяется по направлению, поэтому направление фототока также остается постоянным и противоположным темновому току. Вольтамперная характеристика фотодиода описывается уравнением

$$I = I_s(e^{\frac{qU}{kT}} - 1) - I_\phi \quad (5.23)$$

Пусть излучение поглощается в p-области таким образом, что произведение $\alpha\omega > 1$, т. е. интенсивность света быстро спадает у поверхности, а $L_n > \omega$, поэтому почти все фотоносители достигают p-n перехода. При обратном направлении напряжения перенос фотоносителей тождествен переносу неосновных носителей, образующихся в результате термогенерации в p-области, поэтому плотность фототока

$$j_{\phi} = qg\omega, \quad (5.24)$$

где объемная скорость генерации фотоносителей $g = \frac{\beta\Phi}{\omega}$.

$$\text{Следовательно, } I_{\phi} = q\beta\Phi A, \quad (5.25)$$

где Φ – световой поток,

β - квантовый выход,

A – площадь р-п-перехода,

ω - толщина полупроводника

Это выражение получено в предположении, что все фотоносители достигают р—п перехода и дают вклад в обратный ток.

Для узкозонных полупроводников ВАХ фотодиода при собственном поглощении в приповерхностной области и выполнении условия $L_n > \omega$ описывается уравнением (5.23), где I_{ϕ} определяется формулой (5.25). Для широкозонных полупроводников (Si, GaAs, GaP) преобладает обратный темновой ток генерации носителей в ОПЗ.

$$I_{RG} = qn_i\delta \frac{A}{2\tau} \quad (5.26)$$

поэтому уравнение ВАХ при обратном направлении напряжения имеет вид

$$I_{обр} = -qA(n_i\delta / 2\tau + \beta\Phi). \quad (5.27)$$

Характеристики и параметры. На рис. 5.29,б. показаны ВАХ фотодиода на основе узкозонного полупроводника. При отсутствии освещения обратный ток равен току насыщения. При освещении он возрастает и тоже практически не зависит от напряжения. Темновая ВАХ проходит через начало координат. При освещении ВАХ сдвигается по оси обратных токов, даже при отсутствии внешнего напряжения на фотодиоде через него протекает обратный ток вследствие проявления фотогоальванического эффекта. Рабочий режим фотодиода соответствует принципиальной электрической схеме, показанной на рис. 5.28,а, с внешним источником напряжения, приложенного к

фотодиоду в запирающем направлении. Этот режим называют *фотодиодным*, рабочая область ВАХ фотодиода расположена в третьем квадранте рис. 5.28,б.

Зависимость фототока от потока излучения называют световой или энергетической характеристикой фотодиода. В отличие от фоторезисторов, фототок фотодиодов линейно зависит от потока излучения при обычных уровнях освещенности

$$I_{\phi} = S_i \Phi \quad (5.28)$$

(S_i — интегральная фоточувствительность), вследствие того, что при $L_n > \omega$ почти все неосновные фотоносители, образовавшиеся в р-области, дают вклад в фототок.

Спектральная характеристика фотодиода (рис.5.29) показывает зависимость чувствительности от длины волны. Граница спектральной характеристики со стороны больших длин волн определяется шириной запрещенной зоны, а коротковолновая граница зависит от толщины облучаемой области и скорости поверхностной рекомбинации. Положение максимума спектральной характеристики зависит от скорости возрастания коэффициента поглощения с уменьшением длины волны.

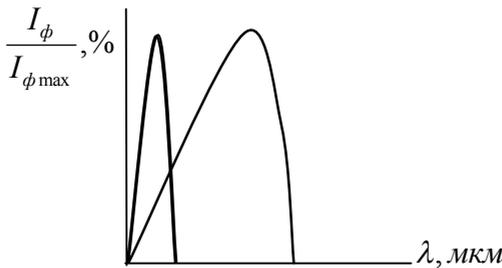


Рисунок 5. 29. Спектральные характеристики фотодиодов на основе германия и кремния

При более быстром возрастании α максимум лежит ближе к длинноволновой границе и слабо зависит от толщины облучаемой области.

Преимуществом фотодиодов по сравнению с фоторезисторами является их быстроедействие. Если фотодиод освещать

модулированным светом, то фототок появляется только после того, как фотоносители достигают р-п перехода. Так как генерация носителей происходит в слое толщиной $x_0 = \frac{1}{\alpha} \ll \omega$, то время диффузии электронов через р-

$$\text{область} \quad \tau_n = \frac{\omega^2}{2D_n} \quad (5.29)$$

Вследствие уменьшения градиента концентрации фотоносителей по мере их продвижения к р-п переходу, скорость диффузии фотоносителей уменьшается во время воздействия светового импульса. Поэтому и нарастание, и спад фототока происходят постепенно и тем сильнее, чем толще р-область, так как при этом увеличивается вклад от разброса скоростей фотоносителей. С увеличением частоты следования импульсов фронт нарастания фототока накладывается на фронт спада фототока от предыдущего импульса. При некоторой частоте фотодиод не успевает различать отдельные импульсы света — фототок через него становится практически постоянным. При одинаковых толщинах освещаемых областей предельная частота больше у фотодиода с освещаемой р-областью, так как $D_n > D_p$. Быстродействие фотодиодов в сильной степени определяется возможностями технологии. Получить толщину освещаемой области менее $30 \div 50$ мкм методом сплавления затруднительно, тогда как с помощью диффузии она получается не толще $3 \div 5$ мкм, поэтому различие в постоянной времени и быстродействии таких фотодиодов составляет почти 100 раз.

Для повышения быстродействия фотодиодов применяют диффузионные р-і-п структуры с тонкой р-областью, тогда излучение поглощается в і-слое. Высокое быстродействие имеют фотодиоды с переходом Шоттки. Излучение проходит через полупрозрачную металлическую пленку и поглощается в ОПЗ полупроводника. Инерционность фотодиода Шоттки определяется перезарядкой барьерной емкости и временем пролета ОПЗ: $\tau = \delta / v_{\max}$, где для германия и кремния максимальная скорость движения носителей в электрическом поле

$v_{\max} \approx 5 \cdot 10^6 \text{ см/с}$ и $\tau \approx 10^{-10} \text{ с}$. Использование узкозонных материалов (InAs, InSb) позволяет сдвинуть длинноволновую границу спектральной характеристики в область $4,5 \div 5,5 \text{ мкм}$. Для работы этих фотодиодов необходимо охлаждение. На рис. 5.30 показаны конструкции фотодиодов типа ФД-3 и ФД-2.

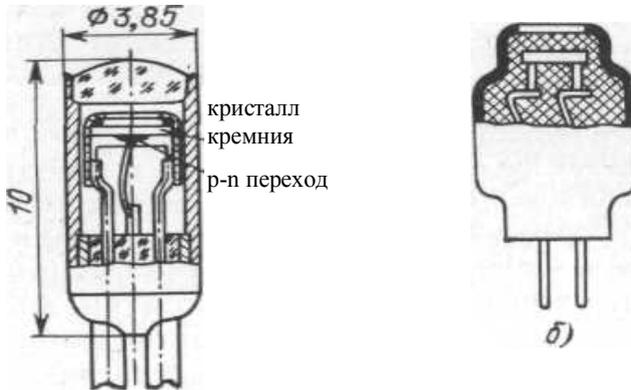


Рисунок 5.30. Фотодиоды типа ФД-3 (а) и ФД-2 (б)

Широкое применение фотодиодов обусловлено чувствительностью их в видимой, УФ и ИК областях спектра, высоким быстродействием, низкими шумами, долговечностью. Основное применение фотодиоды находят в вычислительной технике в устройствах ввода и вывода информации, фотометрии, автоматике. Перспективно применение фотодиодов в оптоэлектронике.

5.9. Светодиоды

Светоизлучающим диодом (СИД) называют полупроводниковый диод, в котором осуществляется непосредственное преобразование электрической энергии в энергию светового излучения (видимого или инфракрасного) за счет рекомбинации электронов и дырок. В обычных диодах процесс рекомбинации заканчивается выделением энергии в виде квантов тепловых колебаний решетки (фононов) и является безызлучательным. Такую рекомбинацию называют *фононной*. В струк-

турах СИД преобладает излучательная или *фотонная* рекомбинация.

В большинстве СИД излучательной рекомбинации электронов и дырок предшествует инжекция неосновных носителей заряда через p-n переход при приложении напряжения в прямом направлении. Характер рекомбинации инжектированных носителей заряда зависит от применяемого в СИД полупроводникового материала, наиболее важным физическим свойством которого является структура энергетических зон. В соответствии с этой структурой все полупроводниковые материалы делят на два класса: материалы с прямыми и непрямыми межзонными переходами электронов (рис. 5.31). При переходах электронов из зоны проводимости в валентную зону должны выполняться законы сохранения энергии и импульса частицы. В силу закона сохранения энергии для прямого перехода (рис.5.31,а) $\Gamma_1 - \Gamma_8 = p\nu$, для непрямого перехода (рис.5.31,б) $X_1 - \Gamma_8 = h\nu$, где греческими буквами обозначена энергия электронов в зонах при волновом числе $k = 0(\Gamma)$ и $k \neq 0(X)$; $h\nu$ — энергия излученного кванта.

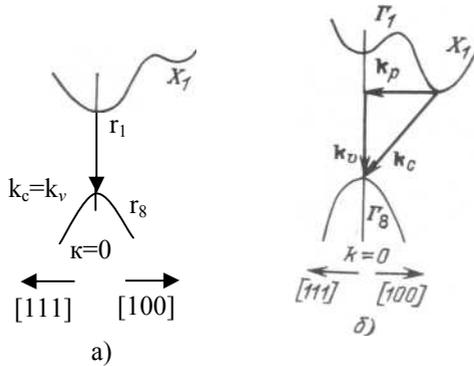


Рисунок 5.31. Прямые (а) и непрямые (б) переходы электронов

Закон сохранения количества движения (импульса) в случае прямых переходов выполняется автоматически, и эти переходы совершаются при одном и том же значении волнового числа $k = 0$, поэтому $k_c = k_v$ - квазиимпульс электрона в зоне проводимости равен квазиимпульсу электрона в валентной зо-

не. Непрямые переходы (из минимума X_1 в максимум Γ_8 , рис.5.31,6) становятся возможными только благодаря участию в них третьей частицы - фонона, которому передается часть количества движения k_p так, что векторная сумма $k_c = k_v + k_p$ символизирует выполнение закона сохранения импульса. Понятно, что интенсивность таких переходов намного меньше, чем прямых, так как требует одновременной встречи трех частиц: электрона, дырки и фонона — явления менее вероятного, чем столкновение только электрона и дырки. В табл. 5.2 некоторые характерные соединения типа $A^{III}B^V$ сопоставлены по энергии минимумов зоны проводимости с прямыми Γ_1 и непрямыми переходами X_1 относительно максимума валентной зоны Γ_8 и постоянным кристаллической решетки a (с).

Таблица 5.2. Структура зон соединений типа $A^{III}B^V$

$A^{III}B^V$	Типы переходов	$\Gamma_1-\Gamma_8$, эВ	$X_1-\Gamma_8$, эВ	a , нм
InSb	Прямые	0,17		0,6479
In As	»	0,36	-	0,6059
QaSb	»	0,69	-	0,6096
InP	»	1,35	2,20	0,5869
GaAs	»	1,43	1,86	0,5653
AlAs	Непрямые	—	2,16	0,5661
GaP	»	2,75	2,26	0,5451
AlP	»	3,6	2,45	0,5451
GaN	Прямые	3,5	-	$a = 0,3186,$ $c = 0,5176$

Специфическими параметрами светоизлучающих диодов являются *световые или фотометрические* величины, количественно характеризующие степень воздействия света на человеческий глаз. За единицу силы света в фотометрии принята кандела (кд). Одна кандела дает полный световой поток 4π

люмена (лм). Эффективность действия света на человеческий глаз определяется *функцией спектральной чувствительности* (функцией видности) или функцией спектральной световой эффективности, зависящей от длины волны (рис.5.32). Чувствительность глаза максимальна при длине волны 555 нм. При этом 1 Вт мощности излучения эквивалентен световому потоку 680 лм. Важной фотометрической величиной является яркость, которая определяется как сила света, отнесенная к видимой площади излучателя ($\text{кд}/\text{м}^2$). Интересным следствием, вытекающим из особенностей фотометрических единиц, является то, что два СИД с длиной волны излучения $\lambda = 650 \text{ нм}$ (красный свет) и $\lambda = 555 \text{ нм}$ (зеленый свет), обладающие одинаковой мощностью излучения, например 1 Вт, и равными поверхностями свечения, например 1 мм^2 , имеют разные яркости. Яркость СИД зеленого цвета свечения в 10 раз больше, чем яркость красного, их световые потоки составят 680 и 68 кд соответственно. Это обусловлено тем, что функция видности резко спадает в сторону больших и

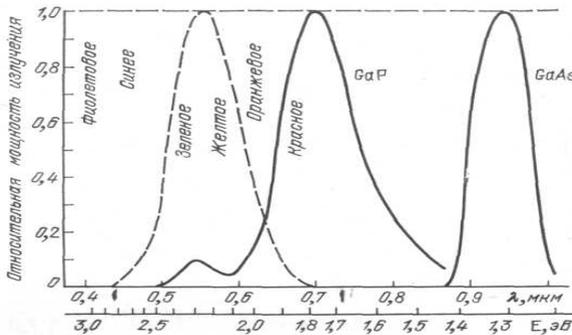


Рисунок 5.32. Функция спектральной чувствительности глаза (штриховая линия) и спектры излучения СИД из арсенида и фосфида галлия

Граничной длиной волны видимого красного света является 720 нм, что соответствует энергии непрямых переходов 1,72 эВ, поэтому для изготовления СИД требуются так называемые широкозонные материалы, у которых ширина запре-

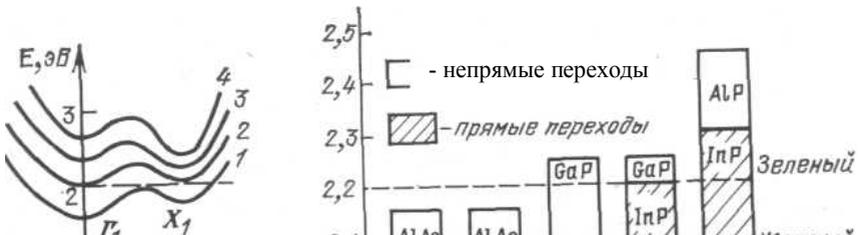
щенной зоны превышает 1,72 эВ. Большинство бинарных соединений типа $A^{III}B^V$ с прямыми межзонными переходами, как видно из табл.5.2, имеет $\Gamma_1-\Gamma_8 < 1,72$ эВ и не излучают в видимой области спектра. Соединения с непрямыми межзонными переходами, как уже отмечалось, обладают малой эффективностью излучательной рекомбинации. Распространенный метод получения эффективных СИД с большой энергией прямых межзонных переходов состоит в использовании тройных твердых растворов - из бинарного соединения с прямыми переходами с малой энергией минимума Γ_1 и соединения с непрямыми переходами с большой энергией минимума X_1 .

Структура энергетических зон таких твердых растворов изменяется в зависимости от состава, что дает возможность управлять энергией переходов и тем самым спектром излучения. На рис. 5.33 показана качественная диаграмма изменения границы зоны проводимости твердого раствора арсенида-фосфида галлия $GaAs_{1-x}P_x = (1-x)GaAs + xGaP$ в зависимости от x-мольной доли фосфида галлия в твердом растворе. Соответствующие числовые данные приведены в табл.5.3.

Таблица 5.3. Энергия межзонных переходов в $GaAs_{1-x}P_x$

Номер кривой	x	$h\nu, \text{эВ}$		Типы переходов	$\lambda, \text{нм}$	Излучаемый свет
		$\Gamma_1-\Gamma_8$	$X_1-\Gamma_8$			
1	0	1,43	1,86	Прямые	910	ИК
2	0,4	1,92	1,97	»	650	Красный
3	0,85	2,55	2,17	Непрямые	580	Желтый
4	1,0	2,78	2,26	»	560	Зеленый

Структура зон с прямыми переходами сохраняется в диапазоне составов $0 \leq x \leq 0,42$, а при $x > 0,42$ преобладают не прямые переходы. Для других типов тройных соединений также существуют области изменения энергии прямых и не прямых переходов. Диаграмма с указанием границ таких переходов показана на рис. 5.34 для соединений $InAlAs$, $GaAlAs$, $GaAsP$, $GaInP$, $InAlP$.





Эффективная фотонная рекомбинация в полупроводнике

Рисунок 5.33. Изменение зоны проводимости в $\text{GaAs}_{1-x}\text{P}_x$ в зависимости от мольной доли GaP

Рисунок 5.34. Диаграмма границ прямых и непрямых переходов в твердых растворах соединений типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$

Эффективная фотонная рекомбинация в полупроводниковых соединениях с непрямими переходами обеспечивается введением специальных примесных центров. В фосфиде галлия, который наиболее широко применяется для изготовления СИД зеленого, оранжевого и красного цветов свечения, центры, повышающие эффективность люминесценции, образуются примесными атомами азота для зеленого и цинка с кислородом для красного излучений. Оранжевое или близкое к желтому свечение получают при одновременном присутствии примесных центров обоих типов, так как суммарное воздействие зеленого и красного цветов воспринимается глазом как желтый цвет. На рис. 5.35 показана зонная диаграмма с примесными уровнями азота и донорно-акцепторной пары Zn—O. Атом азота замещает фосфор в GaP. Поскольку это элемент V группы, то азот является изоэлектронной примесью. Он создает короткодействующий притягивающий потенциал некулоновского типа. После захвата электрона центр становится отрицательно заряженным, и кулоновский потенциал даль-

него порядка позволяет ему захватить свободную дырку с образованием связанного экситона — комплекса из электрона и дырки. Излучательная рекомбинация связанного на уровне азота экситона сопровождается зеленой люминесценцией с энергией основной полосы излучения 2,23 эВ.

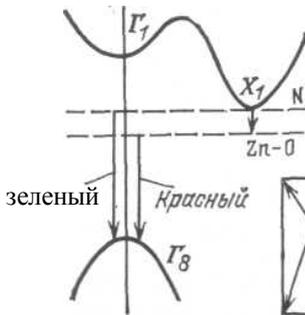


Рисунок 5.35. Зонная диаграмма фосфида галлия с примесями азота и донорно-

акцепторных пар
цинк-кислород

Красная люминесценция в фосфиде галлия возникает вследствие присутствия донорно-акцепторных пар, создаваемых мелким акцептором - цинком Zn ($\Delta E_a \approx 60$ мэВ) и глубоким донором - кислородом O ($\Delta E_d \approx 0,8$ эВ), которые замещают атомы Ga и P в соседних узлах соответственно. Донорно-акцепторная пара $Zn - O$ подобна изоэлектронной ловушке с энергией связи 0,3 эВ для электронов.

Энергетическая эффективность СИД характеризуется *квантовым выходом* η или квантовой эффективностью - отношением числа излученных фотонов к числу электронов, прошедших через р-п переход. Различают внутренний η_i и внешний η_e квантовые выходы. Внутренний квантовый выход определяет эффективность излучения непосредственно в р-п переходе, а внешний - на поверхности кристалла полупроводника. Внутренний квантовый выход зеленой люминесценции в фосфиде галлия составляет около 1%, что гораздо ниже, чем квантовый выход излучения в полупроводнике с прямыми

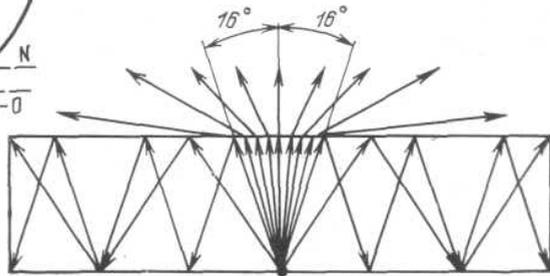


Рисунок 5.36. Полное внутренне отражение лучей от источника S в кристалле

межзонными переходами, где он достигает 50%. Это обусловлено тем, что имеется возможность термической диссоциации экситонов до рекомбинации, а также сильным влиянием конкурирующего процесса безызлучательной рекомбинации. Внутренний квантовый выход красной люминесценции с энергией фотонов 1,79 эВ значительно выше, чем зеленой, и составляет 10—20%, так как влияние безызлучательной рекомбинации слабее.

В структурах СИД примеси определяют не только тип электропроводности и удельное сопротивление материала, но и преобладающие процессы излучательной рекомбинации. Высокая квантовая эффективность обеспечивается, если излучательные переходы электронов преобладают над безызлучательными процессами рекомбинации. Увеличение концентрации примеси повышает концентрацию основных носителей заряда и вероятность рекомбинации и снижает электрическое сопротивление. Чрезмерное повышение концентрации примеси приводит к появлению дефектов, которые служат центрами безызлучательной рекомбинации. С уменьшением содержания примеси падает скорость рекомбинации инжектированных носителей заряда, возрастают сопротивление структуры и падение напряжения на ней, что вызывает разогрев и уменьшение эффективности электролюминесценции. Оптимальная концентрация примеси в данном полупроводниковом соединении определяется экспериментально с точки зрения обеспечения максимального выхода электролюминесценции и находится в пределах 10^{17} — 10^{18} см⁻³ для доноров и 10^{17} — 10^{19} см⁻³ для акцепторов.

Наряду с процессами генерации света одной из главных проблем является выведение света из полупроводникового кристалла. Показатель преломления n соединений типа $A^{III}B^V$ обычно близок к 3,5, а критический угол для лучей, падающих из области p - n перехода на границу кристалл - воздух $\varphi = \arcsin(1/n)$, составляет 16°. Основная часть излучения, т. е. все лучи, падающие под углами больше критического, испытывают полное внутреннее отражение (см. рис.5.37). Вследствие этого полный оптический путь света внутри кристалла боль-

ше размеров кристалла, что ведет к сильному поглощению света. Из-за поглощения и отражения света для излучения с энергией фотонов, близкой к ширине запрещенной зоны, внешний квантовый выход на $1 \div 2$ порядка меньше внутреннего. Обычно он составляет $0,1 \div 7\%$.

Поглощение проявляется главным образом в полупроводниках с прямыми переходами, в которых энергия излучения кванта равна ширине запрещенной зоны, и значительно слабее в тех случаях, когда излучательная рекомбинация идет через глубокие центры. В полупроводниках с непрямыми переходами энергия излучаемого фотона меньше ширины запрещенной зоны, коэффициент поглощения мал и поглощение света играет меньшую роль. Снижение поглощения достигается уменьшением толщины слоя, через который проходит излучение, уменьшением концентрации носителей в материале, выведением света через «оптическое окно», образуемое тонким слоем материала с большой шириной запрещенной зоны, наращиваемого на поверхность кристалла.

В большинстве СИД эффективность излучательных процессов для р- и п-слоев одинакова, тем не менее излучение в основном генерируется в р-области. Это связано с меньшей инжекцией дырок в п-область вследствие их более низкой подвижности, а также с тем, что уровень Ферми в полупроводнике п-типа лежит немного выше границы зоны проводимости, а в полупроводнике р-типа - выше валентной зоны. Кроме того, излучение, возникающее в п-области, сильно поглощается в полупроводнике р-типа, поскольку энергия излучаемых квантов в материале п-типа больше, чем в материале р-типа. Излучение, возникающее в р-области, проходит через п-область с минимальными потерями на поглощение.

Структуры СИД обычно изготавливают на основе эпитаксиальных слоев, выращенных на монокристаллических подложках из арсенида или фосфида галлия. Выбор подложки зависит от параметра решетки эпитаксиального слоя и коэффициента термического расширения. Слитки фосфида галлия, используемые при изготовлении подложек, выращивают в печи высокого давления с жидким уплотнителем методом

Чохральского. Слитки арсенида галлия растят либо методом Чохральского, либо горизонтальной плавкой по методу Бриджмена при атмосферном давлении. Эпитаксиальные слои наращивают на подложки из жидкой и газовой фазы. Для твердых растворов $\text{GaAs}_{1-x}\text{P}_x$ (арсенида - фосфида галлия) успешно применяется газофазная эпитаксия (ГФЭ), в то время как для GaP и $\text{Ga}_{1-x}\text{Al}_x\text{As}$ (арсенида галлия - алюминия) лучшие результаты получены при жидкофазной эпитаксии (ЖФЭ). Излучающий p-n переход образуют либо эпитаксиальным наращиванием p- и n-областей, либо диффузией цинка в эпитаксиальный слой n-типа. Диффузию цинка обычно используют для изготовления СИД из арсенида - фосфора галлия, так как она позволяет применить планарный процесс. Лучшие красные и желтые приборы из фосфида галлия изготовлены методом двойной эпитаксии. В табл. 5.4 приведены наиболее употребляемые материалы и методы изготовления СИД.

Эффективность однослойных СИД, у которых p-n переход совпадает с границей раздела эпитаксиальный слой - подложка, обычно существенно ниже, чем у СИД, изготовленных с выращиванием буферного слоя того же типа электропроводности, что и подложка. Это обусловлено тем, что число безызлучательных центров рекомбинации и дислокаций в подложке очень велико.

Таблица 5.4. Материалы для изготовления СИД

Материал	Цвет свечения	Типы переходов	Технология		V %
			p-слой	p-слой	
GaP : Zn—0	Красный	Непрямые	ЖФЭ	ЖФЭ	15
GaP: N	Зеленый	»	ЖФЭ	ЖФЭ	0,7
GaP:N	»	»	ГФЭ	Диффузия	0,1
GaP : NN	Желтый	»	ГФЭ	»	0,1
GaAs _{0,6} P _{0,4}	Красный	Прямые	ГФЭ	»	0,5
GaAs _{0,35} P _{0,65} : N	Оранжевый	Непрямые	ГФЭ	»	0,5
GaAs _{0,15} P _{0,85} : N	Желтый	»	ГФЭ	»	0,2
Ga _{0,7} Al _{0,3} As	Красный	Прямые	ЖФЭ	ЖФЭ	1.3

$Ga_{0,71}In_{0,3}P$	Оранже- вый	»	ГФЭ	Диффузия	0,1
----------------------	----------------	---	-----	----------	-----

Эффективность однослойных СИД, у которых р-п переход совпадает с границей раздела эпитаксиальный слой - подложка, обычно существенно ниже, чем у СИД, изготовленных с выращиванием буферного слоя того же типа электропроводности, что и подложка. Это обусловлено тем, что число безызлучательных центров рекомбинации и дислокаций в подложке очень велико.

Световой поток или яркость являются важнейшими эксплуатационными параметрами СИД. Они пропорциональны плотности тока через р-п переход (рис.5.37). Интенсивность красного излучения в СИД на основе фосфида галлия и его твердых растворов выше интенсивности зеленого излучения при плотности токов до 10 A/cm^2 . Однако вблизи этого значения зависимость яркости свечения от плотности тока выходит на участок насыщения вследствие того, что концентрация донорно-акцепторных пар $Zn - O$ не может быть выше 10^{17} см^{-3} без ущерба для структуры кристаллической решетки.

Концентрация азота в фосфиде галлия может достигать без всяких вредных последствий 10^{19} см^{-3} , поэтому зеленая люминесценция насыщается при более высоких плотностях токов.

Светоизлучающие диоды являются источниками некогерентного излучения, спектральные кривые излучения имеют широкие пики (рис.5.33), причем у фосфид-галлиевых СИД часто присутствует два основных максимума: красного и зеленого свечений.

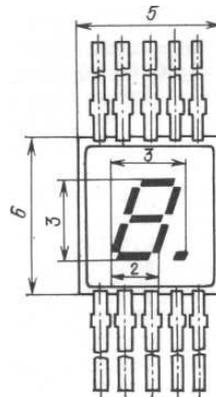
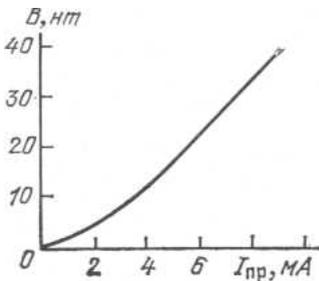


Рисунок 5.37. Яркостная характеристика СИД

Рисунок 5.38. Внешний вид цифрового полупроводникового индикатора

Широкое применение получили не отдельные СИД, а матрицы СИД, позволяющие воспроизводить букву или цифру, - цифро-знаковые индикаторы (ЦЗИ). Внешний вид ЦЗИ из семи излучающих сегментов показана на рис.5.38. ЦЗИ изготавливаются в виде монолитных матриц или гибридных ИС а керамической подложке.

Полупроводниковые ЦЗИ могут собираться в знаковое табло, линейную шкалу или экран и применяются в устройствах отображения аналоговой и цифровой информации. Широкое применение полупроводниковые ЦЗИ нашли в бытовой технике (микрокалькуляторы, электронные наручные часы).

5.10. Оптопары

Оптопарой называется оптоэлектронный прибор, содержащий излучатели и фотоприемники, оптически и конструктивно связанные друг с другом.

Принцип действия оптопары основан на двойном преобразовании энергии. В излучателях энергия электрического сигнала преобразуется в оптическое излучение, а в фотоприемниках, наоборот, оптический сигнал вызывает электрический ток или напряжение. Таким образом, оптопара представляет собой прибор с электрическими входными и выходными сигналами, т. е. связь с внешней схемой электрическая. Внутри оптопары связь входа с выходом осуществляется с помощью оптических сигналов. В электрической схеме такой прибор выполняет функцию выходного элемента - фотоприемника с одновременной электрической изоляцией (гальванической развязкой) входа и выхода. Однако сводить назначение оптопары только к обеспечению электрической изоляции было бы неверно. Введение с помощью оптопары оптического управления позволяет

получить электронные устройства с исключительно своеобразными параметрами и характеристиками.

С точки зрения применения именно фотоприемник является определяющим элементом оптопары, а излучатель выбирается «под фотоприемник». С точки зрения конструктивно-технологической оба элемента - приемник и излучатель - являются «равноправными». Более того, излучателем в большинстве случаев определяются эффективность преобразования энергии и срок службы оптопары. Тем не менее еще раз следует подчеркнуть, что схемотехническое «лицо» оптопары определяет именно фотоприемник.

Принципиальные достоинства оптопар, обусловленные использованием фотонов в качестве носителей информации, заключаются в обеспечении высокой электрической изоляции входа и выхода, однонаправленности потока информации, отсутствии обратной связи с выхода на вход и широкой полосе пропускания.

Кроме того, важными достоинствами оптопар являются: возможность бесконтактного (оптического) управления электронными объектами и обусловленные этим разнообразие и гибкость конструкторских решений управления; невосприимчивость оптических каналов связи к воздействию электромагнитных полей, что в случае оптопар с протяженным оптическим каналом обуславливает высокую помехозащищенность, а также исключает взаимные наводки;

возможность создания функциональных микроэлектронных устройств с фотоприемниками, характеристики которых под действием оптического излучения изменяются по заданному (сколь угодно сложному) закону;

расширение возможностей управления выходным сигналом оптопары путем воздействия (в том числе и неэлектрического) на оптический канал и, как следствие этого, создание разнообразных датчиков и приборов для передачи информации.

Современным оптопарам присущи и определенные недостатки:

низкий КПД, обусловленный необходимостью двойного преобразования энергии (электричество - излучение - электричество), и значительная потребляемая мощность;

сильная температурная зависимость параметров;
 высокий уровень собственных шумов;
 конструктивно-технологическое несовершенство, связанное
 в основном с использованием гибридной технологии.

Перечисленные недостатки оптопар по мере совершенствования материалов, технологии, схмотехники постепенно устраняются. Широкое применение оптопар определяется прежде всего неповторимостью достоинств этих приборов.

Рассмотрим отдельные элементы оптопар подробнее (рис.5.39).

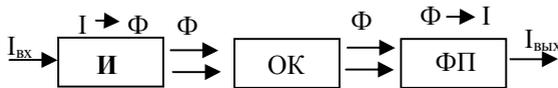


Рисунок 5.39. Структура оптопары:

И – излучатель; ОК – оптический канал; ФП - фотоприемник

Излучатель оптопары. Основным излучателем современных оптопар является инжекционный излучающий диод. В будущем - для создания сверхбыстродействующих оптопар ($t_{пер} \div 10^{-9} - 10^{-10}$ с) и мощных оптопар с протяженным оптическим каналом - применение полупроводниковых лазеров, характеризующихся высоким быстродействием и высокой направленностью излучения, может оказаться целесообразным и экономически оправданным.

Следует подчеркнуть отличия в конструкции оптического окна излучателя оптопары по сравнению с обычным излучающим диодом. Излучатель излучающего диода изготавливают с кольцевой излучающей областью вокруг расположенной в центре контактной площадки. В результате видимая область излучения как бы увеличивается на площадь контактной площадки. Для оптопары излучающая область должна быть минимальной по площади, что уменьшает краевые потери излучения. Минимальная площадь излучающей области ограничивается допустимой плотностью тока через излучатель. Контактная площадка в излучателе оптопары смещается из центра излучающей области. Это создает минимальное затенение и также

уменьшает потери излучения при передаче к фотоприемнику. Малый размер излучающей области позволяет обеспечить, кроме того, стабильность условий оптической связи, сделать их практически независимыми от точности совмещения с приемным окном фотоприемника.

Оптический канал. Качество оптопары в значительной степени зависит от эффективности передачи энергии от излучателя к приемнику, т. е. от свойств оптического канала. Для уменьшения потерь на отражение необходимо создать иммерсионную систему, в которой показатели преломления n оптической среды и материалов излучателя и фотоприемника были бы одинаковыми. Такая цель принципиально достижима, так как у основных используемых полупроводниковых материалов значения n близки друг к другу.

В оптопарах используют следующие конструктивные виды оптических каналов:

связь через воздух, которая характеризуется простотой и высокой электрической изоляцией;

связь через воздух с использованием оптической фокусировки с помощью линз, что обеспечивает лучшую передачу излучения по сравнению с прямой связью через воздух;

связь с использованием иммерсионной среды, что обеспечивает наилучшие параметры оптического канала;

связь с использованием отрезка световода (обычно жесткого моноволокна), которая удобна при создании оптоизоляторов с допустимым напряжением изоляции свыше $(20 \div 50)$ кВ.

Таким образом, при выборе типа оптического канала необходимо в зависимости от применения оптопары удовлетворить следующие требования:

1) обеспечить заданный уровень электрической изоляции между входом и выходом оптопары (между излучателем и фотоприемником);

2) согласовать материал оптического канала с излучателем и фотоприемником спектрально, т. е. обеспечить высокую прозрачность для излучения в рабочем диапазоне длин волн;

3) снизить до минимума потери на отражение на границах излучатель — оптический канал и оптический канал — фотоприемник.

При выборе оптического канала требования к изоляции оказываются решающими, если расстояние между излучателем и фотоприемником мало. Если же расстояние достаточно велико, например, при использовании световодов, изолирующие свойства становятся менее важными. Зато большое значение приобретает спектр пропускания оптического канала, который является составной частью спектральной характеристики системы И—ОК—ФП (рис. 5.40).

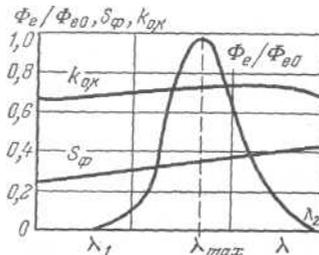


Рис. 5.40. Спектральные характеристики элементов оптопары

6. БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

6.1. Общие сведения об биполярных транзисторах (БТ)

БТ — это полупроводниковый прибор, содержащий два взаимодействующих ЭДП, три или более выводов. Усилительные генераторные и переключательные свойства БТ обусловлены явлениями инжекции и экстракции неосновных носителей заряда. Взаимодействие переходов осуществляется при условии, что толщина базы W должна быть много меньше длины диффузионного смещения носителей, инжектированных в базу L_0 , т.е. $W \ll L_0$. Это соотношение является основным требованием к базе БТ и условием взаимодействия переходов. В противном случае носители, инжектированные в базу, прорекомбинируют в слое базы и не дойдут до коллекторного перехода.

БТ имеет выводы: эмиттер, базу и коллектор. ЭДП, прилегающий к выводу от эмиттера называют эмиттерным переходом (Э), а прилегающий к коллектору - коллекторным (К) (рис. 6.1).

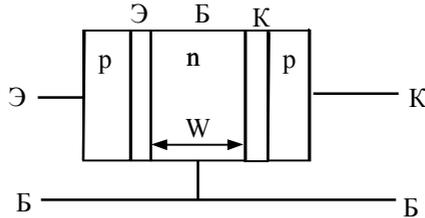


Рисунок 6.1. Структура и выводы БТ

Между эмиттерным и коллекторным переходами находится слой базы (Б). Основное назначение эмиттерного перехода - инжекция неосновных носителей в базу. Основное назначение коллекторного перехода – экстракция неосновных носителей из базы.

Рассматриваемый вид транзистора называют биполярным потому, что в его работе принимают участие носители двух знаков – электроны и дырки.

БТ – это управляемый ток активный элемент электроники. Управление выходным током в БТ осуществляется за счет изменения входного тока. БТ является активным элементом, так как имеет возможность производить усиление по мощности.

Различают p-n-p и n-p-n транзисторы. Их условно-графические изображения представлены на рис. 6.2.

Направление стрелки в эмиттерном выводе показывает направление тока эмиттера.

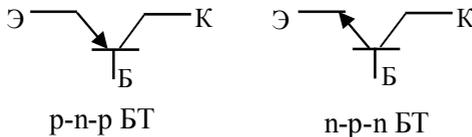


Рисунок 6.2. Условно-графическое обозначение БТ

БТ могут включать в схемах с общей базой (ОБ), общим эмиттером (ОЭ) и общим коллектором (ОК) (рис.6.3.).

Общим считается электрод, являющийся общим для входной и выходной цепей.

Отличительной особенностью схем с ОЭ и ОБ является малое входное сопротивление, тогда как схема с ОК имеет высокое входное сопротивление.

БТ может быть включен нормально и инверсно. Нормальное включение – это когда эмиттерный переход включен в прямом, а коллекторный в обратном направлении.

Инверсное включение – когда коллекторный переход включен в прямом, а эмиттерный в обратном направлении.

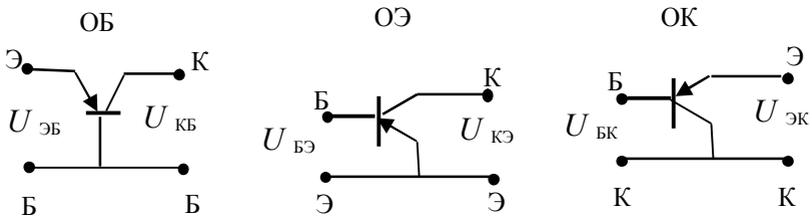


Рисунок 6.3. Схемы включения БТ

Отличие свойств эмиттерного и коллекторного переходов, а следовательно и нормального и инверсного включения, хорошо видно из рис. 6.4, где показана структура реального БТ.

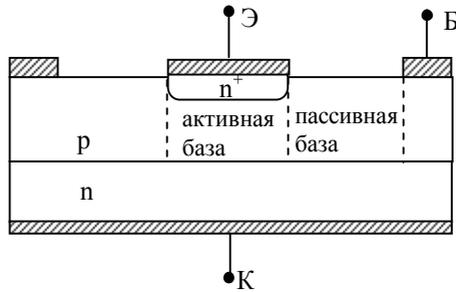


Рисунок 6.4. Структура реального БТ

Важно, что концентрация примеси в эмиттере N_E много больше, чем в базе N_B , а площадь эмиттера много меньше площади коллекторного перехода. В зависимости от смещения переходов различают три режима работы БТ: активный, насыщения и отсечки.

Активный режим – эмиттерный переход смещен в прямом, коллекторный в обратном направлениях.

Режим насыщения – оба перехода смещены в прямом направлении.

Режим отсечки – оба перехода смещены в обратном направлении.

Из трех указанных режимов управляемым режимом является активный режим работы БТ.

По характеру движения носителей зарядов от эмиттера к коллектору БТ подразделяются на дрейфовые и бездрейфовые.

Бездрейфовый БТ – это транзистор, у которого носители от эмиттера к коллектору перемещаются за счет диффузии.

Дрейфовый БТ – это транзистор, у которого носители заряда перемещаются от эмиттера к коллектору и за счет диффузии и за счет внутреннего электрического поля в базе транзистора. Внутреннее электрическое поле в базе возникает из-за неоднородного легирования базы. База бездрейфового БТ легирована однородно.

6.2. Поток носителей зарядов в БТ

На рис. 6.5. представлена структура $p^+ - n - p$ транзистора в активном режиме. При прямом смещении эмиттерного перехода из эмиттера в базу инжектируются дырки, создавая ток $I_{ЭР}$, а из базы в эмиттер – электроны, создавая ток $I_{Эн}$. Частично носители рекомбинируют в ОПЗ эмиттерного перехода, создавая ток $I_{Эрек}$. Таким образом, ток эмиттера состоит из трех компонент: $I_{ЭР}$, $I_{Эн}$, $I_{Эрек}$.

$$I_{Э} = I_{ЭР} + I_{Эн} + I_{Эрек}.$$

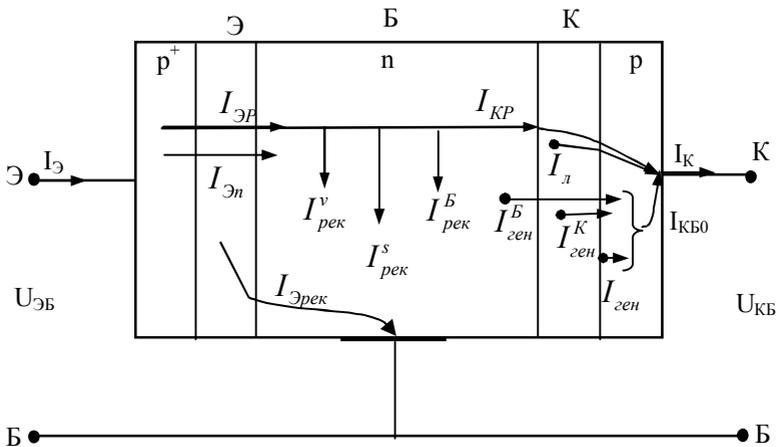


Рисунок 6.5. Потоки носителей зарядов в БТ

С учетом того, что поток дырок из эмиттера в базу много больше, чем поток электронов из базы в эмиттер, ток рекомбинации $I_{Эрек}$ мал. В дальнейшей работе БТ главную роль играют дырки, инжектированные из эмиттера в базу. Двигаясь по базе дырки частично рекомбинируют, создавая ток $I_{рек}^v$ за счет рекомбинации носителей в объеме, $I_{рек}^Б$ – за счет рекомбинации на базовом контакте, $I_{рек}^s$ – за счет рекомбинации на поверхности. Не про-рекомбинировавшие дырки дойдут до коллекторного ЭДП (ток $I_{КР}$) и экстрагируются полем коллекторного перехода, создавая

ток коллектора. Кроме того, в цепи коллектора будут протекать токи, обусловленные тепловой генерацией носителей в базе ($I_{ген}^B$), в коллекторном переходе ($I_{ген}^K$) и слое коллекторного полупроводника ($I_{ген}$). При наличии процессов ударной ионизации в ОПЗ коллекторного перехода в цепи коллектора появится ток лавинного размножения $I_{Л}$. Таким образом, в цепи коллектора протекают следующие компоненты тока: $I_{КР}$, $I_{ген}^B$, $I_{ген}^K$, $I_{ген}$, $I_{Л}$.

$$I_K = I_{КР} + I_{ген}^B + I_{ген}^K + I_{ген} + I_{Л}.$$

Обычно токи тепловой генерации обозначают как ток $I_{КБ0}$:

$I_{КБ0} = I_{ген}^B + I_{ген}^K + I_{ген}$ и называют обратным током коллекторного перехода. Из пяти компонент тока I_K только одна компонента является зависящей от тока эмиттера ($I_{КР}$), остальные компоненты тока I_K являются неуправляемыми, т.е. не зависящими от тока эмиттера. Исходя из схемы потоков носителей зарядов в БТ, можно записать два уравнения:

$$I_{Э} = I_K + I_B \quad (6.1)$$

$$I_K = \alpha I_{Э} + I_{КБ0} \quad (6.2)$$

Первое уравнение говорит о том, что ток эмиттера распадается на две компоненты – ток базы I_B и ток коллектора I_K . Причем ток базы – это ток, связанный с рекомбинацией носителей в базе и чем он меньше, тем лучше БТ.

Второе уравнение говорит о том, что ток эмиттера передается в цепь коллектора с некоторым коэффициентом α и кроме того в цепи коллектора протекает неуправляемая компонента тока, связанная с тепловой генерацией ($I_{КБ0}$).

6.3. Внутренние и внешние параметры БТ

Для характеристики свойств БТ используются внутренние и внешние физические параметры. К числу внутренних парамет-

ров относятся параметры, характеризующие свойства эмиттера, базы и коллектора и измерить которые нельзя.

Эффективность эмиттера – внутренний параметр, характеризующий инжекционную способность эмиттерного перехода.

$$\left. \begin{array}{l} \text{Определяется как } \gamma = \frac{I_{\text{Эп}}}{I_{\text{Э}}} = 1 - \frac{\sigma_n W}{\sigma_p L_n} \text{ - для р-п-р БТ} \\ \text{Определяется как } \gamma = \frac{I_{\text{Эн}}}{I_{\text{Э}}} = 1 - \frac{\sigma_n W}{\sigma_p L_n} \text{ - для р-п-р БТ} \end{array} \right\} (6.3)$$

Можно так же записать выражения для γ в общем виде

$$\gamma = 1 - \frac{\sigma_B W}{\sigma_{\text{Э}} L_{\text{Э}}}, \quad (6.4)$$

где $\sigma_B, \sigma_{\text{Э}}$ - электропроводность базы и эмиттера соответственно.

$L_{\text{Э}}$ - длина диффузионного смещения носителей, инжектированных из базы в эмиттер.

Эффективность эмиттера γ всегда меньше единицы и чем ближе она к единице, тем лучше свойства эмиттерного перехода.

Коэффициент переноса носителей через базу α_n - характеризует свойства базы и определяется для р-п-р БТ следующим выражением:

$$\alpha_n = \frac{I_{\text{КР}}}{I_{\text{Эп}}} = 1 - \frac{W^2}{2L_b^2}, \quad (6.5)$$

где L_b - длина диффузионного смещения носителей, инжектированных из эмиттера в базу. α_n всегда меньше единицы и зависит от толщины базы и параметров материала базы (L_b).

Эффективность коллектора α^* - характеризует способность коллекторного перехода экстрагировать неосновные носители из базы. Для р-п-р БТ α^* определяется как

$$\alpha^* = \frac{I_K}{I_{\text{КР}}} \approx 1 \quad (6.6)$$

Обычно эффективность коллектора полагают равной единице, поскольку из-за разницы в площадях эмиттерного и коллекторного переходов ($S_K \gg S_\ominus$) все носители, инжектированные эмиттером, собираются коллекторным переходом.

К числу внешних физических параметров БТ относятся следующие:

Статический коэффициент передачи тока эмиттера α - определяется следующим образом :

$$\text{Для нормального включения} - \alpha_N \approx \frac{I_K}{I_\ominus} / U_{KB} = const \text{ или}$$

$$\text{более точно} \quad \alpha_N = \frac{I_K - I_{KB0}}{I_\ominus} / U_{KB} = const. \quad (6.7)$$

$$\text{Для инверсного включения} - \alpha_I = \frac{I_\ominus}{I_K} / U_{ЭБ} = const \text{ или}$$

$$\text{более точно} \quad \alpha_I = \frac{I_\ominus - I_{ЭБ0}}{I_K} / U_{ЭБ} = const \quad (6.8)$$

$$\text{В БТ выполняется соотношение} \quad \alpha_N I_{KB0} = \alpha_I I_{ЭБ0} \quad (6.9)$$

В дальнейшем индекс, указывающий нормальное включение, будем упускать, обозначая нормальный коэффициент передачи тока эмиттера через α ;

Дифференциальный коэффициент передачи тока эмиттера $\tilde{\alpha}$. По определению

$$\tilde{\alpha} = \frac{dI_K}{dI_\ominus} / U_{KB} = const .$$

С учетом выражения $I_K = \alpha I_\ominus + I_{KB0}$

$$\tilde{\alpha} = \frac{dI_K}{dI_\ominus} = \alpha + I_\ominus \frac{d\alpha}{dI_\ominus} \quad (6.10)$$

При условии, что коэффициент передачи тока эмиттера α не зависит от тока эмиттера ($\frac{d\alpha}{dI_\ominus} = 0$), дифференциальный коэффициент

коэффициент $\tilde{\alpha}$ равен статическому α . В дальнейшем будем счи-

тать, что эти коэффициенты равны. Значение α всегда меньше единицы, если в коллекторном переходе не происходит лавинного размножения носителей заряда. При наличии лавинного размножения $\alpha_M = \alpha \cdot M > 1$.

Указание о том, что α измеряется при постоянном напряжении на коллекторном переходе связано с эффектом Эрли, который будет рассмотрен в разделе 6.5.

В БТ выполняется легко доказываемо соотношение

$$\alpha = \gamma \alpha_n \alpha^*, \quad (6.11)$$

из которого следует ряд формул:

$$\alpha = \gamma \left(1 - \frac{W^2}{2L_6^2}\right), \quad \alpha = 1 - \frac{\sigma_B}{\sigma_3} \frac{W}{L_3} - \frac{W^2}{2L_6^2} \quad (6.12)$$

Для БТ, включенного в схему с ОЭ основным параметром является **коэффициент передачи тока базы** β , который определяется

следующим образом:

$$\beta \approx \frac{I_K}{I_B} / U_{КЭ} = const = \frac{I_K - I_{КЭ0}}{I_B} / U_{КЭ} = const, \quad (6.13)$$

где $I_{КЭ0}$ - обратный ток коллекторного перехода при включении БТ в схему с ОЭ. Причем $I_{КЭ0} = (1 + \beta)I_{КБ0}$.

Значения β всегда больше единицы, что связано с принципом функционирования БТ в схеме с ОЭ. Так для р-п-р БТ при подаче сигнала управления (тока базы) в базу поступают электроны, вызывая нарушение электронейтральности базы. Для устранения этого из эмиттера в базу начинают поступать дырки. С учетом того, что время пролета носителей через базу t_B много меньше времени жизни τ_B , часть дырок уходит в коллектор без рекомбинации. Таким образом, на один электрон, поступивший в базу, в коллекторный переход попадает значительно большее число дырок, что и обеспечивает значения $\beta > 1$. БТ, имеющие $\beta > 400$, называют транзисторами супер-бета.

В БТ выполняется соотношения:

$$\alpha = \frac{\beta}{1 + \beta}; \quad \beta = \frac{\alpha}{1 - \alpha}; \quad \frac{1}{1 - \alpha} = 1 + \beta \quad (6.14)$$

6.4. Статические параметры

Статическими параметрами БТ называют значения токов или напряжений, измеряемых на постоянном напряжении при определенных условиях, а также соотношения между этими величинами.

Режим отсечки. Статическими параметрами режима отсечки являются начальные и обратные токи переходов.

Начальными токами переходов являются токи в одном из переходов при его обратном смещении при условии, что другой переход закорочен, т.е. условиями измерения этих токов является короткое замыкание одного из переходов. К числу начальных токов относятся токи $I_{КБК}$, $I_{ЭБК}$, $I_{КЭК}$.

$I_{КБК}$ - ток в цепи коллектора БТ, включенного по схеме с ОБ, при его обратном смещении при коротком замыкании цепи эмиттер – база.

$I_{ЭБК}$ - ток в цепи эмиттера транзистора, включенного в схему с ОБ, при обратном смещении эмиттерного перехода и коротком замыкании коллекторного перехода.

$I_{КЭК}$ - ток в цепи коллекторного перехода, включенного в схему с ОЭ при обратном смещении коллекторного перехода и коротком замыкании цепи эмиттер – база ($U_{БЭ} = 0$).

Обратные токи переходов – это токи в одном из переходов, когда он обратно смещен, при условии, что ток другого перехода равен нулю, т.е. обратные токи измеряются при условии холостого хода одного из переходов.

$I_{КБ0}$ - ток в цепи коллекторного перехода БТ, включенного в схему с ОБ, при включении его в обратном направлении при условии, что ток эмиттера равен нулю.

$I_{ЭБ0}$ - ток в цепи эмиттерного перехода, когда он обратно смещен при условии, что ток коллектора равен нулю.

$I_{KЭ0}$ - ток в цепи обратно смещенного перехода БТ, включенного в схему с ОЭ, при условии, что ток базы равен нулю.

Необходимо обратить внимание, что условия измерения токов влияют на величину измеряемых токов. Так ток $I_{КБК}$ не равен току $I_{КБ0}$. Дело в том, что при коротком замыкании эмиттерного перехода (условия измерения тока $I_{КБК}$) одна из составляющих ток эмиттера ($I_{ЭР}$) не обращается в нуль. В результате $I_{КБК}$ всегда больше $I_{КБ0}$. Аналогичные рассуждения применимы и к токам $I_{ЭБК}$ и $I_{ЭБ0}$.

Значения токов также зависят от схемы включения БТ. Так ток $I_{КБ0}$ и $I_{КЭ0}$ не равны между собой. Легко показать, что они связаны соотношением (6.15).

$$I_{КЭ0} = \frac{I_{КБ0}}{1 - \alpha} \quad (6.15)$$

К числу статических параметров режима отсечки относятся также напряжения $U_{КБ0}$ и $U_{КЭ0}$. Напряжения пробоя коллекторного перехода в схеме с ОБ ($U_{КБ0}$) и с ОЭ ($U_{КЭ0}$) при условии равенства нулю тока эмиттера и тока базы соответственно.

Режим активной работы. Статическими параметрами активного режима являются статические коэффициенты передачи тока эмиттера α и тока базы β .

$$\alpha = \frac{I_K - I_{КБ0}}{I_{Э}} / U_{КБ} = const \quad (6.16)$$

$$\beta = \frac{I_K - I_{КЭ0}}{I_B} / U_{КБ} = const$$

Режим насыщения. К числу статических параметров в режиме насыщения относятся $U_{КБнас}$ как падение напряжения между выводами база – коллектор в режиме насыщения для схемы с ОБ и $U_{КЭнас}$ - в схеме с ОЭ. $I_{Кнас}$ - ток коллектора в режиме

насыщения, $R_{нас.}$ сопротивление насыщенного транзистора, γ_H - степень насыщения транзистора, определяемая как

$$\gamma_H = \frac{\beta \cdot I_B}{I_{Кнас.}} \quad (6.17)$$

Параметры режима насыщения подробно рассмотрены в 6.9.

6.5. Явления в БТ при больших токах

Стремление получить максимальные выходные параметры от БТ сопровождается работой транзистора при больших токах. При этом возникает ряд явлений, которые надо учитывать при проектировании схем. К числу эффектов, возникающих при больших токах, относятся следующие:

1. В базе бездрейфового БТ возникает внутреннее электрическое поле;
2. Увеличивается толщина базы;
3. Возникает эффект квазинасыщения БТ;
4. Происходит оттеснение тока эмиттера на периферию.

Причина возникновения внутреннего электрического поля в базе аналогична той, что рассмотрены в разделе 4.10. Возникающее поле является ускоряющим для инжектированных в базу носителей.

Поскольку ширина ОПЗ перехода зависит от количества носителей вблизи перехода, то при больших токах из-за уменьшения ширины ОПЗ коллекторного и эмиттерного переходов происходит увеличение ширины базы и, в соответствии с (6.12) уменьшение коэффициента передачи ток эмиттера.

При больших токах коллектора напряжение $U_{KB} = U_{KB}^{p-n} + U_K$ (U_{KB}^{p-n} - падение напряжения на ОПЗ коллекторного перехода; U_K - падение напряжения на слое коллекторного полупроводника; $U_K = I_K R_K$, где R_K - сопротивление коллекторного полупроводника) перераспределяется между коллекторным переходом (U_{KB}^{p-n}) и слоем полупроводника так, что при определенном токе коллектора падение напряжения на ОПЗ коллекторного перехода фактически уменьшится до нуля, т.е. БТ будет находиться на границе перехода между активным

режимом и режимом насыщения, т.е. возникает режим квазинасыщения.

Напряжение $U_{ЭБ}$ падает на ОПЗ эмиттерного перехода и слое базы $U_{ЭБ} = U_{ЭБ}^{p-n} + U_B = U_{ЭБ}^{p-n} + I_Э R_B$. Поскольку пути протекания тока в БТ, вытекающего из различных точек эмиттера, различны, то разным будет падение напряжения на сопротивлении базы (рис. 6.6).

Из-за увеличенного падения напряжения на слое базы истинное напряжение на эмиттерном переходе, определяющее величину тока эмиттера, оказывается наименьшим в центральной части эмиттерного перехода, т.е. происходит оттеснение тока из центральной части эмиттера на периферию.

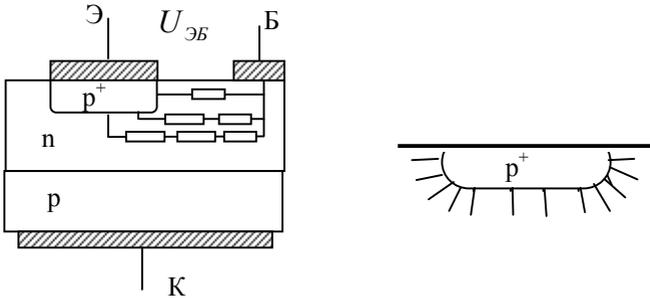


Рисунок 6.6. Рисунок, поясняющий оттеснение тока эмиттера на периферию при больших токах

Необходимо отметить, что основной параметр БТ – коэффициент передачи тока эмиттера (базы) также зависит от величины тока эмиттера (рис. 6.7). При выборе режима работы БТ, необходимо обеспечить его работу при токах, соответствующих максимальным значениям α .

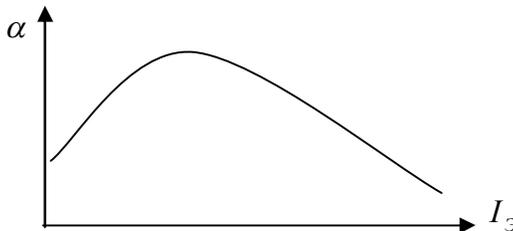


Рисунок 6.7. Зависимость коэффициента передачи тока эмиттера от тока эмиттера

6.6. Модуляция толщины базы коллекторным напряжением (эффект Эрли)

Эффект Эрли – это изменение толщины базы БТ при изменении напряжения на коллекторном переходе. Физически эффект Эрли прост и понятен – при изменении обратного напряжения изменяется ширина ОПЗ коллекторного перехода, следовательно, модулируется толщина базы. Необходимо иметь в виду, что эффект Эрли является очень важным эффектом и объяснение очень многих закономерностей в БТ базируется на этом эффекте. Можно выделить шесть основных следствий эффекта Эрли:

1) Основной параметр БТ – коэффициент передачи тока эмиттера определяется в соответствии с формулой

$$\alpha = \gamma \left(1 - \frac{W^2}{2L_B^2}\right), \text{ т.е. при изменении толщины базы изменяется}$$

основной параметр БТ, таким образом $\alpha = f(U_{KB})$. Именно

$$\text{поэтому по определению } \alpha = \frac{I_K - I_{KB0}}{I_{\mathcal{O}}} / U_{KB} = const \text{ задается}$$

при фиксированном значении напряжения на коллекторном переходе.

2) Обратный ток эмиттера зависит от толщины базы:

$$I_{\mathcal{O}B0} \approx \frac{S_{\mathcal{O}} D_p P_{n0}}{W}. \text{ Так как } I_{\mathcal{O}} = I_{\mathcal{O}B0} (\exp \frac{qU_{\mathcal{O}B}}{KT} - 1), \text{ то, следо-}$$

вательно, при фиксированном напряжении на эмиттерном переходе ток эмиттера зависит от напряжения на коллекторном переходе.

3) Одним из следствий эффекта Эрли является явление прокола базы БТ. Прокол базы – это смыкание ОПЗ эмиттерного и коллекторного переходов при увеличении напряжения на коллекторном переходе. Напряжение прокола базы $U_{\text{прок}}$ определяется исходя из того, что прокол происходит тогда, когда

$d_K = W$. Тогда для резкого коллекторного перехода

$$W = d_K = \sqrt{\frac{2\varepsilon\varepsilon_0}{qN_B} U_{\text{прок}}}; \quad U_{\text{прок}} = \frac{W^2 q N_B}{2\varepsilon\varepsilon_0}; \quad (6.18)$$

Или, домножая числитель и знаменатель на подвижность носителей в базе μ_B и заменяя $q\mu_B N_B = \frac{1}{\rho_B}$, получаем формулу

(6.19)

$$U_{\text{прок}} = \frac{W^2}{2\varepsilon\varepsilon_0 \rho_B \mu_B} \quad (6.19)$$

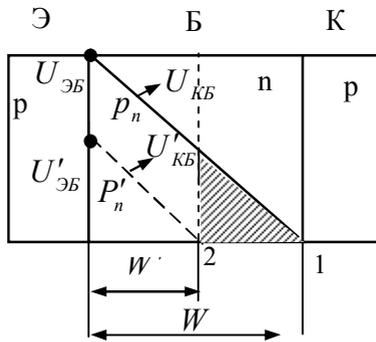


Рисунок 6.8. Рисунок, поясняющий появление отрицательной обратной связи по напряжению

4) Возникает отрицательная обратная связь по напряжению. Действительно, при увеличении $U_{КБ}$ уменьшается толщина базы W . Если задать постоянным значение тока эмиттера, то новое распределение инжектированных носителей p'_n должно быть параллельно распределению p_n , соответствующее исходному значению напряжения $U_{КБ}$. Из рисунка 13.8. видно, что при постоянном значении тока эмиттера увеличение напряжения на коллекторном переходе от $U_{КБ}$ до $U'_{КБ}$ приводит к

уменьшению напряжения на эмиттерном переходе от $U_{ЭБ}$ до $U'_{ЭБ}$.

Количественно величина отрицательной обратной связи характеризуется коэффициентом отрицательной обратной связи по напряжению μ .

$$\mu = \frac{dU_{ЭБ}}{dU_{КБ}} / I_{Э} = const \quad (6.20)$$

Численное значение μ лежит в пределах $10^{-4} - 10^{-3}$, что говорит о наличии неглубокой отрицательной обратной связи по напряжению.

5) При изменении толщины базы меняется время пролета носителей через базу: $t_B = \frac{W^2}{2D_B}$. Как будет показано далее (раздел 13.11.) одна из характеристических частот БТ – частота f_α

определяется временем пролета носителей через базу, и кроме того f_α задает значение двух других частот $f_{ГР}$ и f_{\max} . Таким образом можно говорить о том, что коллекторное напряжение влияет на частотные параметры БТ.

6) В коллекторном переходе БТ возникает диффузионная емкость. Это происходит из-за того, что при изменении напряжения $U_{КБ}$ граница ОПЗ коллекторного перехода из сечения 1 переходит в сечение 2 (рис. 6.8) очень быстро ($\sim 10^{-13}$ с), а имеющийся вблизи коллекторного перехода заряд создает диффузионную емкость коллекторного перехода C_{KD} . Её величина определяется формулой 6.21.

$$C_{KD} = \frac{\tau_B}{R_{КБ}}, \quad (6.21)$$

где τ_B - время жизни неосновных носителей заряда в базе

$R_{КБ}$ - дифференциальное сопротивление коллекторного перехода.

6.7. Пробой транзистора

Одной из причин, наряду с проколом базы, ограничивающих напряжение на коллекторном переходе, является пробой транзистора. Говоря о пробое БТ, обычно имеют в виду пробой коллекторного перехода. Из трех известных механизмов применительно к БТ говорят обычно о лавинном механизме пробоя, поскольку туннельный пробой коллекторного перехода маловероятен, а тепловой может иметь место, но как вторичный эффект. Пробой характеризуется напряжением пробоя, которое в зависимости от схемы включения, измеряется при токе $I_{\mathcal{C}} = 0$ в схеме с ОБ ($U_{КБ0}$) или $I_B = 0$ в схеме с ОЭ ($U_{КЭ0}$). При измерении напряжения пробоя в схеме с ОБ пробой происходит как в отдельном р-п переходе: появляющиеся в ОПЗ коллекторного перехода электроны выталкиваются полем перехода в базу и уходят из нее через вывод базы. Напряжение пробоя для схемы с ОБ обозначается $U_{КБ0}$ и измеряется при $I_{\mathcal{C}} = 0$.

Пробой в схеме с ОЭ имеет некоторую особенность (рис.6.9). Появляющиеся в результате ударной ионизации носители выталкиваются из ОПЗ коллекторного перехода: дырки увеличивают ток коллектора, а электроны выталкиваются область базы. Поскольку вывод базы разорван, то они не могут покинуть базу и, создавая отрицательный заряд, уменьшают барьер на эмиттерном переходе, что вызывает рост тока эмиттера и соответственно коллектора.

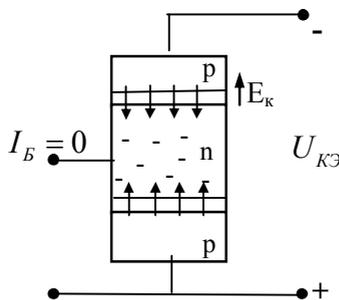


Рисунок 6.9. Пробой БТ в схеме с ОЭ при $I_B = 0$

Пробой по указанному механизму вызывает резкий и фактически неограниченный рост тока в цепи коллектора и может при-

вести к разрушению БТ, если не принять специальных мер. Поэтому включение БТ при условии, что $I_B = 0$, является недопустимым. Обычно БТ включают так, чтобы в цепи базы стоял резистор. Тогда процессы лавинообразного роста тока в цепи коллектора становятся невозможными. Напряжение пробоя БТ в этом случае обозначается как $U_{КЭР}$. Величины напряжений пробоя в схеме с ОБ и ОЭ связаны между собой соотношением 6.22.

$$U_{КЭ0} = U_{КБ0} \sqrt[6]{1 - \alpha}, \quad (6.22)$$

где ν - коэффициент, зависящий от материала ($\nu = 2 \div 6$).

6.8. Статические характеристики

Под статическими характеристиками понимаются взаимосвязи между токами и напряжениями в БТ, измеренные на постоянном напряжении при отсутствии нагрузки в выходной цепи. Поскольку в БТ имеются входные и выходные токи и входные и выходные напряжения, то можно установить значительное количество соотношений между ними. В практике полупроводниковой электроники используются четыре семейства характеристик: входные, выходные, характеристики прямой передачи и характеристики обратной связи.

Статические характеристики для схемы с ОБ.

Входные характеристики - $I_{\mathcal{E}} = f(U_{\mathcal{ЭБ}}) / U_{КБ} = const$.

Внешний вид входных характеристик (рис.6.10) соответствует прямым ветвям ВАХ электронно-дырочного перехода. Аналитически входная характеристика приближенно задается выражением

$$I_{\mathcal{E}} = I_{\mathcal{ЭБ0}} \left(\exp \frac{qU_{\mathcal{ЭБ}}}{KT} - 1 \right).$$

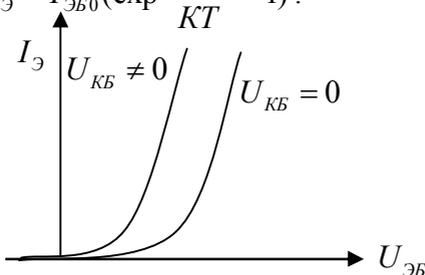


Рисунок 6.10. Входные характеристики для БТ в схеме с ОБ

Смещение характеристик при изменении напряжения U_{KB} обусловлено эффектом Эрли: с ростом U_{KB} уменьшается толщина базы, что при постоянном напряжении $U_{ЭБ}$ вызывает рост тока эмиттера.

Выходные характеристики - $I_K = f(U_{KB})/U_{Э} = const.$

Аналитический вид характеристик имеет вид - $I_K = \alpha I_{Э} + I_{KB0}$, а графические зависимости представлены на рис. 6.11. Внешний вид кривых соответствует обратной ветви ВАХ ЭДП. На участке активной работы выходные характеристики имеют небольшой угол наклона, связанный с зависимостью $\alpha = f(U_{KB})$.

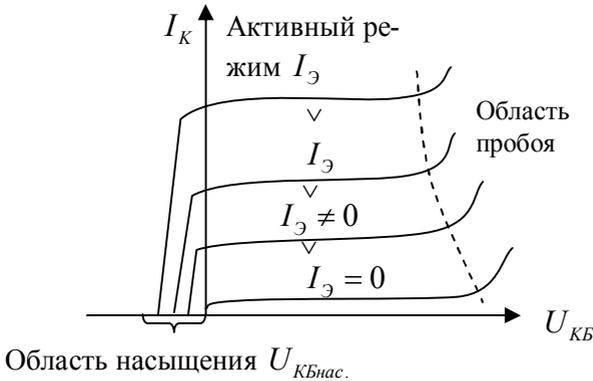
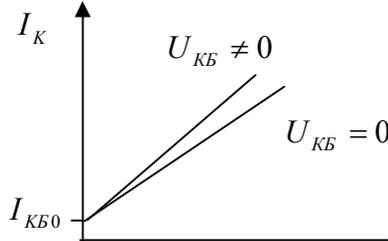


Рисунок 6.11. Семейство выходных характеристик для БТ в схеме с ОБ

Наличие тока при $U_{KB} = 0$ говорит о том, что ток в цепи коллектора связан не с наличием поля в цепи коллектора, а с наличием градиента концентрации неосновных носителей заряда в базе. Для уменьшения тока до нуля БТ надо перевести в режим насыщения.

Характеристики передачи тока - $I_K = f(I_{Э})/U_{КБ} = 0$.

В соответствии с выражением $I_K = \alpha I_{Э} + I_{КБ0}$ характеристики передачи должны представлять прямые линии с углом наклона, определяемым коэффициентом α (рис.6.12).



При $I_{Э} = 0$ на оси абсцисс отсекается ток $I_{КБ0}$. Смещение характеристик вправо вводит непрерывность прямой передачи тока в БТ в нормальном переходе $U_{КБ}$ связано с эффектом Эрли.

Характеристики обратной связи представлены на рис.6.13 и особого пояснения для схемы с ОБ не требуют.

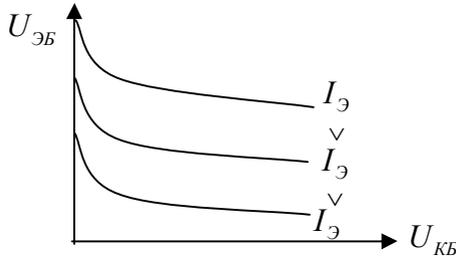


Рисунок 6.13. Статические характеристики обратной связи БТ в схеме с ОБ

Статические характеристики БТ для схемы с ОЭ

Входные характеристики - $I_B = f(U_{БЭ})/U_{КЭ} = const$. Поскольку вид входных характеристик задается выражением $I_B = (1 - \alpha)I_{Э} - I_{КБ0}$, то они представляют собой прямые ветви ВАХ эмиттерного перехода (рис.6.14). Сдвиг характеристик при изменении напряжения $U_{КЭ}$ обусловлен Эффектом Эрли: при увеличении $U_{КЭ}$ уменьшается толщина базы и, следовательно,

снижается количество рекомбинирующих частиц, создающих ток базы.

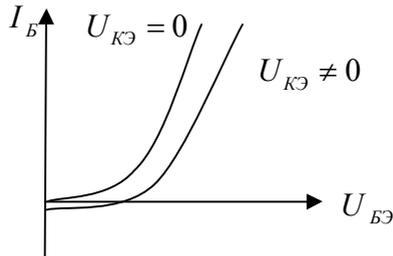


Рисунок 6.14. Входные статические характеристики БТ в схеме с ОЭ

Выходные характеристики - представлены на рис. 6.15. Зависимость тока коллектора в схеме с ОЭ от тока управления определяется уравнением (6.23).

$$I_K = \beta I_B + I_{КЭ0} \quad (6.23)$$

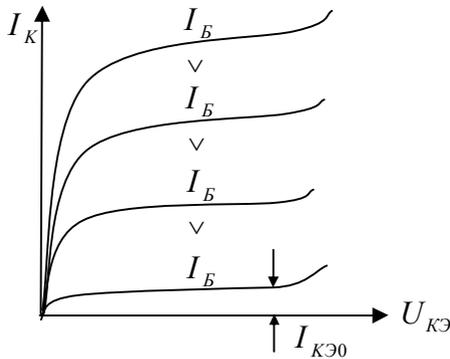


Рисунок 6.15. Выходные характеристики БТ, включенного по схеме с ОЭ

На участке активной работы рост тока I_K при увеличении напряжения $U_{КЭ}$ обусловлен эффектом Эрли. Особенностью семейства характеристик является то, что они располагаются только в первом квадранте.

Характеристики передачи тока - $I_K = f(I_B)/U_{КЭ} = const$

представлены на рис. 6.16. В принципе эти зависимости должны иметь вид прямых линий с углом наклона пропорциональным β , т.к. выходной ток I_K задается выражением 6.23.

На практике эти зависимости сильно отклоняются от прямо-

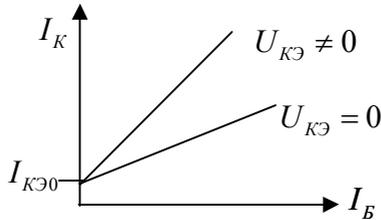


Рисунок 6.16. Характеристики передачи тока для БТ в схеме с ОЭ

линейных из-за сильной зависимости β от режима работы.

Характеристики обратной связи –

$U_{БЭ} = f(U_{КЭ})/I_B = const$ представлены на рис. 6.17. Увеличение $U_{БЭ}$ с ростом $U_{КЭ}$ обусловлено тем, что эти зависимости снимаются при постоянном токе базы.

Ток базы пропорционален площади треугольника, занимаемого

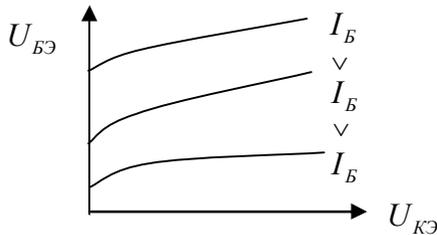


Рисунок 6.17. Характеристики обратной связи БТ в схеме с ОЭ

распределением инжектированных в базу носителей. С увеличением напряжения $U_{КЭ}$ толщина базы уменьшается и для того, чтобы $I_B = const$, напряжение $U_{БЭ}$ должно увеличиваться.

6.9. Динамический режим работы

Для построения нагрузочной прямой используется уравнение

$$E_K = U_{KЭ} + I_K R_H \quad (6.24)$$

При отсутствии сигнала на входе ($I_B = 0$) БТ находится в закрытом состоянии. Рабочая точка соответствует точке A и реализуется режим отсечки. При этом через коллекторный переход протекает ток $I_{KЭ0}$. Поскольку он мал, то все напряжение от источника питания E_K падает между эмиттером и коллектором БТ. При подаче большого сигнала БТ открывается и его рабочая точка смещается в точку B , что соответствует режиму насыщения. За счет протекания тока $I_{Kнас.}$ практически все напряжение от источника питания E_K падает на нагрузке R_H , а напряжение между коллектором и эмиттером падает до $U_{KЭнас.}$, которое составляет десятые доли вольта. Совокупность точек пересечения статических характеристик и нагрузочной прямой называют динамической характеристикой БТ. Переход БТ в режим насыщения происходит при определенном токе базы. Дальнейшее увеличение тока базы не приводит к смещению рабочей точки, но увеличивает степень насыщения γ_H .

$$\begin{aligned} \gamma_H &= \frac{\beta I_B}{I_{Kнас.}}; \quad I_{Kнас.} = \frac{E_K - U_{KЭнас.}}{R_H} \approx \frac{E_K}{R_H} \\ \gamma_H &= \frac{\beta I_B R_H}{E_K} \end{aligned} \quad (6.25)$$

Считается, что транзистор переходит в режим насыщения при токе базы, обеспечивающем $\gamma_H \geq 1$. Таким образом в режиме отсечки ($I_B = 0$) $\gamma_H = 0$, а режиму насыщения соответствует $\gamma_H \geq 1$. Динамический режим работы БТ ограничен снизу областью отсечки, слева – насыщения, справа – пробоем, сверху – предельным значением тока через коллектор.

6.10. Усилительные свойства

Усилительные свойства БТ задаются коэффициентами усиления по току, напряжению и мощности, которые в общем виде

определяются отношением выходного параметра к входному. Запишем эти коэффициенты для схем с ОБ, ОЭ и ОК.

Схема с ОБ.

Для схемы с ОБ (рис.6.20) имеем:

$$\left. \begin{aligned} k_i &= \frac{I_{\text{вых}}}{I_{\text{вх}}} = \frac{I_K}{I_{\text{Э}}} \approx \alpha < 1 \\ k_u &= \frac{U_{\text{вых}}}{U_{\text{вх}}} = \frac{I_k R_H // R_{\text{вых}}}{I_{\text{Э}} R_{\text{вх}}} \approx \alpha \frac{R_H // R_{\text{вых}}}{R_{\text{вх}}} \\ k_p &= k_i \cdot k_u = \alpha^2 \frac{R_H // R_{\text{вых}}}{R_{\text{вх}}} \end{aligned} \right\} \quad (6.26)$$

где $R_H // R_{\text{вых}}$ - сопротивление параллельно включенных резисторов нагрузки R_H и выходного сопротивления БТ;

$R_{\text{вх}}$ - входное сопротивление БТ.

Из выражений 6.26 видно, что в схеме с ОБ нельзя получить усиление по току, но можно при соответствующем подборе нагрузочного резистора получить усиление по напряжению и мощности.

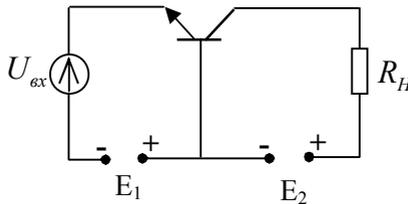


Рисунок 6.20. Схема включения БТ с ОБ

Схема с ОЭ

Для схемы с ОЭ (рис. 6.21) имеем:

$$\left. \begin{aligned} k_i &= \frac{I_{\text{вых}}}{I_{\text{вх}}} = \frac{I_k}{I_B} \approx \beta > 1 \\ k_u &= \frac{U_{\text{вых}}}{U_{\text{вх}}} = \frac{I_k R_H // R'_{\text{вых}}}{I_B R'_{\text{вх}}} = \beta \frac{R_H // R'_{\text{вых}}}{R'_{\text{вх}}} \end{aligned} \right\} \quad (6.27)$$

$$k_p = k_i \cdot k_u = \beta^2 \frac{R_H // R'_{\text{вых}}}{R'_{\text{вх}}}$$

где, $R_H // R'_{\text{вых}}$ - сопротивление параллельно включенных резисторов R_H и $R'_{\text{вых}}$;

$R'_{\text{вых}}$ и $R'_{\text{вх}}$ - выходное и входное сопротивления транзистора для схемы с ОЭ.

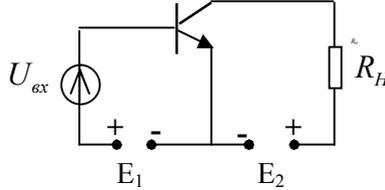


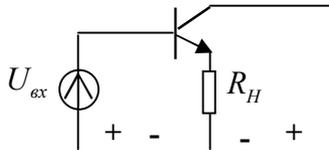
Рисунок 6.21. Схема включения БТ с ОЭ

Из 6.27. видно, что схема с ОЭ обеспечивает получение значительных коэффициентов усиления по току, напряжению и мощности. Причем усиление достигается за счет собственно транзистора и путем правильного подбора сопротивления нагрузки. Усилительные свойства БТ в схеме с ОЭ значительно лучше, чем в схеме с ОБ.

Схема с ОК

Для схемы с ОК (рис. 6.22) имеем

$$\left. \begin{aligned} k_i &= \frac{I_{\text{вых}}}{I_{\text{вх}}} = \frac{I_{\text{Э}}}{I_{\text{Б}}} = \frac{I_{\text{К}} + I_{\text{Б}}}{I_{\text{Б}}} = 1 + \beta \\ k_u &= \frac{U_{\text{вых}}}{U_{\text{вх}}} = \frac{U_{\text{вх}} - U_{\text{БЭ}}}{U_{\text{вх}}} < 1 \\ k_p &\approx k_i \cdot k_u \approx 1 + \beta \end{aligned} \right\} \quad (6.28)$$



Из 6.28. видно, что в схеме с ОК нельзя получить усиление по напряжению. Но имея высокое входное сопротивление схема с ОК широко используется как согласующий каскад.

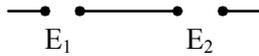


Рисунок 6.22. Схема включения БТ с ОК

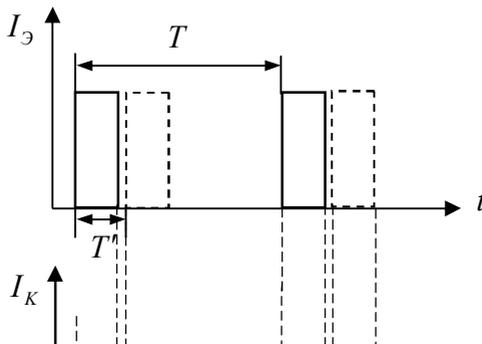
6.11. Частотные параметры

Пусть БТ включен в схему с ОБ, током управления является ток эмиттера. Сигналы во входной цепи следуют с периодом T (рис. 6.23). Ток в цепи коллектора появится не сразу, а с некоторой задержкой, обусловленной наличием времени пролета носителей от эмиттера к коллектору. Далее, импульс тока в цепи коллектора будет отличаться от прямоугольного, что связано с разбросом носителей по скоростям и наличием барьерных емкостей у переходов. Кроме того, амплитуда тока коллектора меньше, чем у тока эмиттера, поскольку часть носителей рекомбинирует в базе. Более того, когда импульс тока эмиттера закончится, в цепи коллектора еще протекает ток. Он исчезнет, когда носители уйдут из базы в коллектор. Таким образом можно говорить о том, что для переменного сигнала БТ представляет из себя инерционный элемент, вносящий в схему амплитудные и фазовые искажения.

При увеличении частоты следования импульсов тока управления (период $T' < T$) сигналы в цепи коллектора начинают накладываться друг на друга. Это позволяет утверждать, что БТ может работать не на любых частотах.

Для характеристики частотных свойств БТ вводится четыре характеристических частоты. Рассмотрим их:

Предельная частота коэффициента передачи тока эмиттера - f_α .



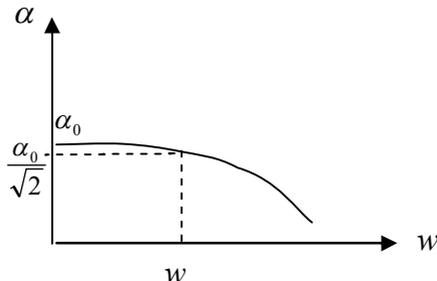
Коэффициент передачи тока эмиттера α является комплексным параметром: $\dot{\alpha} = \frac{\alpha_0}{1 + jw/w_\alpha}$. Модуль $|\dot{\alpha}|$ задается выражением

$$|\dot{\alpha}| = \frac{\alpha_0}{\sqrt{1 + \left(\frac{w}{w_\alpha}\right)^2}} \quad (6.29)$$

Тогда частота, на которой модуль коэффициента передачи тока эмиттера уменьшается в $\sqrt{2}$ раз по сравнению с низкочастотным значением α_0 , называется предельной частотой f_α . Частота f_α определяется в схеме с ОБ в основном временем пролета носителей через базу t_B .

$$f_\alpha = \frac{1}{2 \cdot \pi \cdot t_B} \quad (6.30)$$

Причиной уменьшения α при увеличении частоты (рис. 6.24) является уменьшение эффективности эмиттера из-за шунтирующего действия барьерной емкости эмиттерного перехода.



Предельная частота коэффициента передачи тока базы f_B .

На рис. 6.25 представлена зависимость модуля коэффициента передачи тока базы $|\dot{\beta}|$ от частоты:

$$|\dot{\beta}| = \frac{\beta_0}{\sqrt{1 + \left(\frac{w}{w_\beta}\right)^2}} \quad (6.31)$$

Частота, на которой модуль $|\dot{\beta}|$ уменьшается в $\sqrt{2}$ раз по сравнению с низкочастотным значением β_0 называется предельной частотой коэффициента передачи тока базы f_B .

$$\text{В БТ } f_\beta = \frac{1}{2\pi\tau_B}, \quad (6.32)$$

где - τ_B - время жизни неосновных носителей в базе.

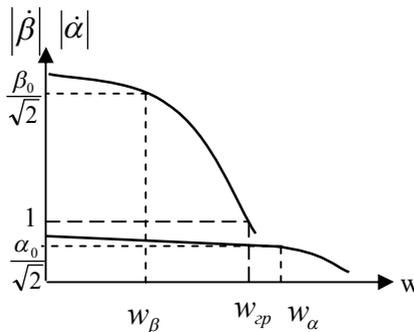


Рисунок 6.26. Зависимость модулей $|\dot{\beta}|$ и $|\dot{\alpha}|$ от частоты

Граничная частота f_{zp} - частота, на которой модуль $|\dot{\beta}|$ становится равным единице (рис. 6.26). Для БТ выполняется соотношение $\alpha f_\alpha = \beta f_\beta = f_{zp}$. Значение f_{zp} можно найти, используя соотношение

$$f_{zp} = |\dot{\beta}| \cdot f \quad (6.33)$$

Частоты f_α и f_β численно существенно отличаются, причем $f_\alpha \gg f_\beta$.

$$f_\alpha = (1 + \beta)f_\beta; \quad f_\beta = (1 - \alpha)f_\alpha \quad (6.34)$$

По значению граничной частоты БТ делятся на низкочастотные ($f_{zp} < 3\text{МГц}$), среднечастотные ($f_{zp} = 3 \div 30\text{МГц}$) и высокочастотные ($f_{zp} > 30\text{МГц}$). Транзисторы, у которых $f_{zp} > 300\text{МГц}$, являются СВЧ транзисторами.

Максимальная частота генерации f_{\max} - частота, на которой коэффициент усиления по мощности равен единице.

$$f_{\max} = \sqrt{\frac{f_{zp}}{8\pi R'_B C_{KB}}}, \quad (6.35)$$

где R'_B - омическое сопротивление базы;

C_{KB} - барьерная емкость коллекторного перехода.

Уменьшение коэффициента усиления по мощности с увеличением частоты происходит из-за того, что барьерная емкость C_{KB} начинает шунтировать нагрузку.

Выражение, стоящее в (6.35) под корнем называют коэффициентом качества БТ. Из четырех частот наиболее важное значение имеет частота f_{\max} , поскольку при более высоких частотах БТ перестает быть активным элементом.

6.12. Эквивалентная схема биполярного транзистора

Для анализа и инженерных расчетов транзисторных схем БТ удобно представлять в виде эквивалентной схемы. При этом параметры эквивалентной схемы можно рассматривать как параметры транзистора.

Существует три варианта эквивалентных схем: Т-образная, П-образная и гибридная. Наиболее широкое применение имеет Т-образная схема. Поэтому в дальнейшем речь будем вести о ней.

При представлении физической эквивалентной схемы БТ необходимо исходить из определения: БТ – это полупроводниковый прибор на основе двух взаимодействующих переходов. Поэтому суть эквивалентной схемы БТ достаточно понятна: должны быть представлены эквивалентные схемы двух р-п переходов и элементы, отражающие взаимодействие переходов. Взаимное влияние переходов заключается в том, что ток из входной цепи передается в коллекторный выход, а изменение напряжения в выходной цепи вызывает изменение входного напряжения.

Можно представить два варианта эквивалентной схемы БТ для области низких частот, когда емкостными элементами можно пренебречь (рис. 6.27 а,б)

На указанных схемах влияние эмиттерного перехода на коллекторный задано генератором тока αI_{ε} (рис.6.27, а) или генератором напряжения $r_m I_{\varepsilon}$ (рис.6.27, б), а коллекторного перехода на эмиттерный – генератором напряжения $\mu U_{КБ}$.

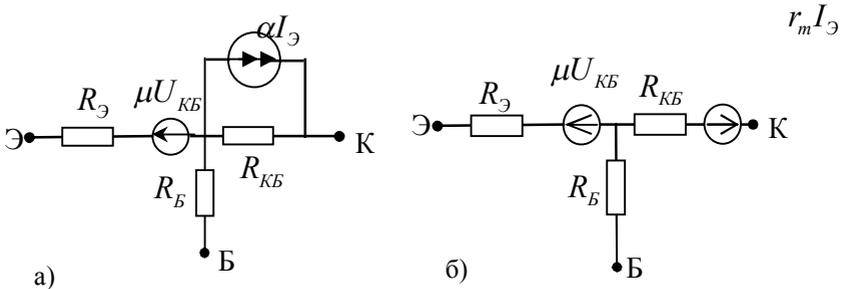


Рисунок 6.27. Низкочастотная эквивалентная схема БТ, включенного в схему с ОБ с генератором тока (а), и генератором напряжения в коллекторной цепи (б)

На рис. 6.28 а, б представлены эквивалентные схемы для случая включения БТ в схему с ОЭ.

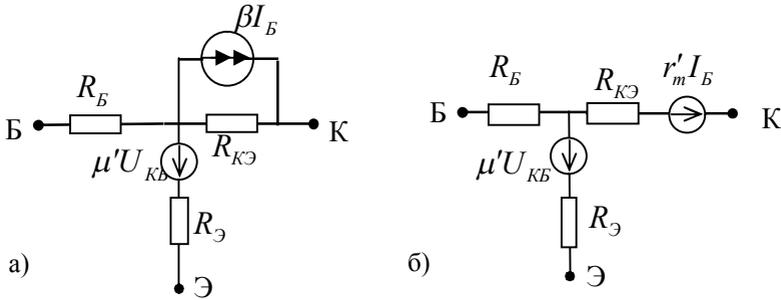


Рисунок 6.28. Низкочастотная эквивалентная схема БТ, включенного в схему с ОЭ с генератором тока (а), и генератором напряжения в коллекторной цепи (б)

На практике чаще используются эквивалентные схемы с генераторами тока, т.к. определение коэффициента пропорциональности r_m и r'_m является более сложным, чем определение α и β . Достаточно часто наличие отрицательной обратной связи по напряжению в эквивалентной схеме задается не генератором напряжения, а включением резистора R_B'' в цепь базы (рис.6.29).

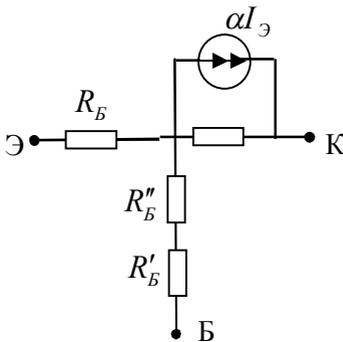


Рисунок 6.29. Эквивалентная схема БТ в области низких частот в схеме с ОБ

По определению $R_B'' = \frac{dU_{ЭБ}}{dI_K} = \frac{dU_{ЭБ}}{dU_{КБ}} \cdot \frac{dU_{КБ}}{dI_K} = \mu R_{КБ}$, т.е.

действительно отражает наличие обратной связи. Однако, учитывая, что отрицательная обратная связь по напряжению не велика, часто ее в эквивалентной схеме не учитывают.

Для области высоких частот эквивалентная схема представлена на рис. 6.30.

Параметры эквивалентной схемы приведены ниже.

$$R_{Э} = \frac{dU_{ЭБ}}{dI_{Э}} = \frac{kT}{q(I_{Э} + I_{ЭБ0})} - \text{дифференциальное сопротивление эмиттера. Имеет значение доли – единицы Ом.}$$

R_B' - омическое сопротивление базы. Зависит от конструкции прибора и лежит в диапазоне десятков Ом. Для точечного и сплавного БТ R_B' рассчитывается по следующим формулам:

$$R_B' = \rho_B \frac{1}{4\pi W} - \text{точечный БТ (} W - \text{толщина базы)}$$

$$R_B' = \frac{1}{3} \rho_B \frac{a}{bW} - \text{сплавной БТ (} a, b - \text{высота и ширина базы).}$$

$$R_B'' = \frac{dU_{ЭБ}}{dI_K} = \frac{kT}{q\gamma I_{Э}} \left(\frac{L_B}{W} \right)^2 - \text{дифференциальное сопротивление базы. Возможные значения – сотни Ом.}$$

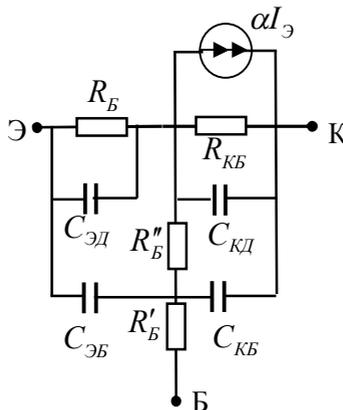


Рисунок 6.30. Эквивалентная схема БТ в области низких частот в схеме с ОБ

Часто при упрощенных расчетах полагают $R_B = R'_B + R''_B$.

$$R_{KB} = \frac{dU_{KB}}{dI_K} = \sqrt{\frac{2qN_B}{\varepsilon\varepsilon_0}} \frac{L_B^2}{W} \frac{\sqrt{U_{KB}}}{\gamma I_\varepsilon} - \text{дифференциальное сопротивление коллекторного перехода. Возможные значения – доли единицы МОм.}$$

Возможные значения – доли единицы МОм.

$$\mu = \frac{kT}{qW\sqrt{U_{KB}}} \sqrt{\frac{\varepsilon\varepsilon_0}{2qN_B}} - \text{коэффициент отрицательной обратной связи по напряжению. Ориентировочные значения}$$

Ориентировочные значения $\mu = 10^{-4} \div 10^{-3}$.

$C_{ЭД} = \frac{q}{kT}(I_\varepsilon + I_{ЭБ0})t_B$ - диффузионная емкость эмиттерного перехода;

$C_{КД} = \frac{\tau_B}{R_{KB}}$ - диффузионная емкость коллекторного перехода;

$C_{ЭБ} = \frac{\varepsilon\varepsilon_0 S_\varepsilon}{d_\varepsilon}$ - барьерная емкость эмиттерного перехода;

$C_{KB} = \frac{\varepsilon\varepsilon_0 S_K}{d_K}$ - барьерная емкость коллекторного перехода;

$\alpha = \gamma(1 - \frac{W^2}{2L_B^2})$ - коэффициент передачи тока эмиттера.

Анализ выражений для расчета параметров эквивалентной схемы показывает, что практически все они зависят от температуры, тока эмиттера и от напряжения на коллекторном переходе. А это означает, что параметры БТ также зависят от температуры, тока эмиттера и от напряжения на коллекторном пе-

реходе (T , $I_{\mathcal{E}}$, U_{KB}) и при $T = const$ рабочая точка задается током эмиттера и напряжением на коллекторе.

Для решения задач по БТ используется внутреннее уравнение транзистора (6.36).

$$R_{\mathcal{E}} = 2\mu R_{KB}(\gamma - \alpha) \quad (6.36)$$

и следующие из соотношения (6.36) формулы:

$$R_B'' = \frac{R_{\mathcal{E}}}{2(\gamma - \alpha)}; \quad R_B'' \approx R_{\mathcal{E}}(1 + \beta) \quad (6.37)$$

6.13. Системы z , y и h – параметров

Биполярный транзистор в любой схеме включения можно рассматривать как линейный четырехполюсник (рис. 6.31).

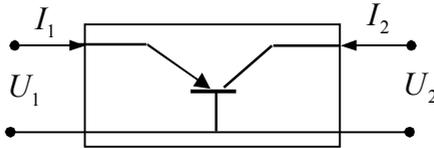


Рисунок 6.31. Биполярный транзистор как четырехполюсник

Для анализа электронных схем на БТ широко используются три системы уравнений. Их отличие заключается в том, какие из параметров четырехполюсника (U_1, U_2, I_1, I_2) используются в качестве аргументов, а какие являются функциями.

z – параметры. В системе z -параметров в качестве аргументов выступают токи, а напряжения являются функциями. Система уравнений БТ как линейного четырехполюсника представлена формулами 6.38.

$$\left. \begin{aligned} U_1 &= z_{11}I_1 + z_{12}I_2 \\ U_2 &= z_{21}I_1 + z_{22}I_2 \end{aligned} \right\} \quad (6.38)$$

Значения токов и напряжений определяется путем измерения на внешних выводах, а z -параметры находят в режиме холостого хода на входе ($I_1 = 0$) и выходе ($I_2 = 0$).

$z_{11} = \left. \frac{U_1}{I_1} \right|_{I_2=0}$ - входное сопротивление БТ при холостом ходе на

выходе;

$z_{12} = \left. \frac{U_1}{I_2} \right|_{I_1=0}$ - сопротивление обратной связи;

$z_{21} = \left. \frac{U_2}{I_1} \right|_{I_2=0}$ - сопротивление прямой передачи при холостом

ходе на выходе;

$z_{22} = \left. \frac{U_2}{I_2} \right|_{I_1=0}$ - выходное сопротивление при холостом ходе на

входе.

Эта система параметров широко используется, но имеет недостаток – трудно на практике реализовать холостой ход на высокоомном выходе. Поскольку все z-параметры имеют размерность сопротивления, то данная система параметров называется параметрами сопротивлений.

На рисунке 6.32. представлены схемы замещения транзистора в z(a), y(б), и h-параметрах (с).

y – параметры. В системе y-параметров в качестве независимых переменных выступают напряжения U_1 и U_2 , а токи I_1 и I_2 являются функциями. Уравнения БТ в y-параметрах представлены выражениями 6.39.

$$\left. \begin{aligned} I_1 &= y_{11}U_1 + y_{12}U_2 \\ I_2 &= y_{21}U_1 + y_{22}U_2 \end{aligned} \right\} \quad (6.39)$$

y - параметры определяются в режиме короткого замыкания на входе ($U_1 = 0$) и выходе ($U_2 = 0$).

$$y_{11} = \left. \frac{I_1}{U_1} \right|_{U_2=0} \quad \text{-входная проводимость при коротком замыкании}$$

на выходе;

$$y_{12} = \left. \frac{I_1}{U_2} \right|_{U_1=0} \quad \text{-проводимость обратной связи при коротком за-}$$

мыкании на входе;

$$y_{21} = \left. \frac{I_2}{U_1} \right|_{U_2=0} \quad \text{- проводимость прямой передачи при коротком}$$

замыкании на выходе;

$$y_{22} = \left. \frac{I_2}{U_2} \right|_{U_1=0} \quad \text{- выходная проводимость при коротком замыкании}$$

на входе.

Все параметры имеют размерность проводимости и поэтому система называется системой параметров проводимости. Недостатком системы у-параметров является трудность реализации короткого замыкания на низкоомном входе.

Недостатков, характерных для z и у-параметров лишена система смешанных h-параметров. В этой системе параметров в качестве аргументов выступает ток на входе и напряжение на выходе, а функциями является напряжение на входе и ток на выходе. Уравнения БТ в системе h-параметров представлены в (6.40)

$$\left. \begin{aligned} U_1 &= h_{11}I_1 + h_{12}U_2 \\ I_2 &= h_{21}I_1 + h_{22}U_2 \end{aligned} \right\} \quad (6.40)$$

Значения h-параметров определяются в режиме холостого хода на входе ($I_1 = 0$) и короткого замыкания на выходе ($U_2 = 0$).

Указанные режимы достаточно просто реализуются в низкоомной входной цепи путем постановки в цепь Э-Б большой индуктивности, а режим короткого замыкания на высокоомном выходе – путем закорачивания цепи К-Б достаточно большой емкостью.

$$h_{11} = \left. \frac{U_1}{I_1} \right|_{U_2=0} \quad \text{- входное сопротивление при коротком замыкании}$$

выходной цепи;

$$h_{12} = \left. \frac{U_1}{U_2} \right|_{I_1=0} \quad \text{- коэффициент обратной связи по напряжению при}$$

холостом ходе во входной цепи;

$$h_{21} = \left. \frac{I_2}{I_1} \right|_{U_2=0} \quad \text{- коэффициент передачи тока при коротком замы-}$$

кании выходной цепи;

$$h_{22} = \left. \frac{I_2}{U_2} \right|_{I_1=0} \quad \text{- выходная проводимость при холостом ходе на}$$

входе.

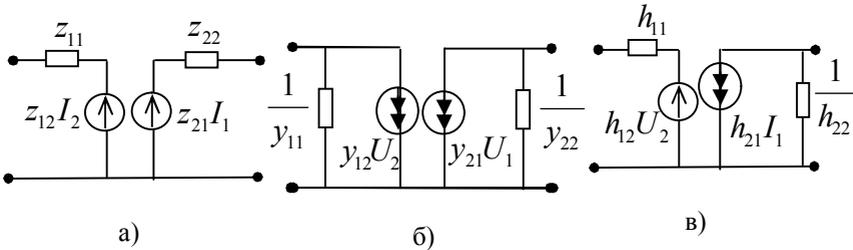


Рисунок 6.32. Схемы замещения БТ в системе z (а), y (б) и h (в)-параметров

Поскольку один и тот же объект может описываться с помощью разных схем, то между ними существует связь. В таблице 6.2 представлены формулы переходов между системами параметров.

Таблица 6.2. Формулы перехода между системами параметров

	$ z $	$ y $	$ h $

$ z $	$z_{11} \quad z_{12}$ $z_{21} \quad z_{22}$	$-\frac{y_{22}}{\Delta_y} \quad -\frac{y_{12}}{\Delta_y}$ $-\frac{y_{21}}{\Delta_y} \quad \frac{y_{11}}{\Delta_y}$	$\frac{\Delta_h}{h_{22}} \quad \frac{h_{12}}{h_{22}}$ $-\frac{h_{21}}{h_{22}} \quad \frac{1}{h_{22}}$
$ y $	$\frac{z_{22}}{\Delta_z} \quad -\frac{z_{12}}{\Delta_z}$ $-\frac{z_{21}}{\Delta_z} \quad \frac{z_{11}}{\Delta_z}$	$y_{11} \quad y_{12}$ $y_{21} \quad y_{22}$	$\frac{1}{h_{11}} \quad -\frac{h_{12}}{h_{11}}$ $\frac{h_{21}}{h_{11}} \quad \frac{\Delta_h}{h_{11}}$
$ h $	$\frac{\Delta_z}{z_{22}} \quad \frac{z_{12}}{z_{22}}$ $-\frac{z_{21}}{z_{22}} \quad \frac{1}{z_{22}}$	$-\frac{1}{y_{11}} \quad -\frac{y_{12}}{y_{11}}$ $\frac{y_{21}}{y_{11}} \quad \frac{\Delta_y}{y_{11}}$	$h_{11} \quad h_{12}$ $h_{21} \quad h_{22}$

Некоторые соотношения по малосигнальным параметрам:

1. Взаимосвязь физических параметров биполярного транзистора и h -параметров в схеме с ОБ и ОЭ:

Схема с ОБ	Схема с ОЭ
$R_{\mathcal{O}} = 2 \left[h_{11B} - \frac{h_{12B}}{h_{22B}} (1 + h_{21B}) \right]$	$R_{\mathcal{O}} = 2 \left(\frac{h_{11\mathcal{O}}}{1 + h_{21\mathcal{O}}} - \frac{h_{12\mathcal{O}}}{h_{22\mathcal{O}}} \right)$
$R_{KB} \approx \frac{1}{h_{22B}}$	$R_K \approx \frac{1}{h_{22\mathcal{O}}}$
$\alpha = -h_{21B}$	$\beta = h_{21\mathcal{O}}$
$R_B = 2 \frac{h_{12B}}{h_{22B}} - \frac{h_{11B}}{1 + h_{21B}}$	$R_B = 2 \frac{h_{12\mathcal{O}}}{h_{22\mathcal{O}}} (1 + h_{21\mathcal{O}})$
$\mu = \frac{h_{11B} h_{22B}}{1 + h_{21B}} - h_{21B}$	$\mu = \frac{h_{11\mathcal{O}}}{h_{22\mathcal{O}}} (1 + h_{21\mathcal{O}}) - h_{12\mathcal{O}}$

2. Представление h -параметров через параметры эквивалентной схемы:

<u>Схема с ОБ</u>	<u>Схема с ОЭ</u>
$h_{11B} = R_{\mathcal{O}} + R_B(1 - \alpha)$	$h_{11\mathcal{O}} = R_B + R_{\mathcal{O}}(1 + \beta)$
$h_{12B} = \frac{R_B}{R_B + R_{K\mathcal{O}}}$	$h_{12\mathcal{O}} = \frac{R_{\mathcal{O}}}{R_{K\mathcal{O}} - R_{\mathcal{O}}}$
$h_{21B} = -\alpha$	$h_{21\mathcal{O}} = \beta$
$h_{22B} = \frac{1}{R_{KB}}$	$h_{22\mathcal{O}} = \frac{1}{R_{K\mathcal{O}}}$
$R_{KB} = R_{K\mathcal{O}}(1 + h_{21\mathcal{O}})$	

3. Представление Z -параметров через параметры эквивалентной схемы:

<u>Схема с ОБ</u>	<u>Схема с ОЭ</u>	<u>Схема с ОК</u>
$z_{11B} = R_{\mathcal{O}} + R_B$	$z_{11\mathcal{O}} = R_{\mathcal{O}} + R_B$	$z_{11K} = R_B + R_{K\mathcal{O}}$
$z_{12B} = R_B$	$z_{12\mathcal{O}} = R_{\mathcal{O}}$	$z_{12K} = R_{K\mathcal{O}}$
$z_{21B} = R_B + \alpha R_{KB}$	$z_{21\mathcal{O}} = R_{\mathcal{O}} + \beta R_{K\mathcal{O}}$	$z_{21K} = R_{K\mathcal{O}} + (1 + \beta)R_{\mathcal{O}}$
$z_{22B} = R_B + R_{KB}$	$z_{22\mathcal{O}} = R_{\mathcal{O}} + R_{K\mathcal{O}}$	$z_{22K} = R_{\mathcal{O}} + R_{K\mathcal{O}}$

4. Взаимосвязь h -параметров в схеме с ОБ и ОЭ:

<u>Схема с ОБ</u>	<u>Схема с ОЭ</u>
$h_{11\mathcal{O}} = \frac{h_{11B}}{1 + h_{21B}}$	$h_{11B} = \frac{h_{11\mathcal{O}}}{1 + h_{21\mathcal{O}}}$
$h_{12\mathcal{O}} = \frac{h_{11B}h_{22B}}{1 + h_{21B}} - h_{12B}$	$h_{12B} = \frac{h_{11\mathcal{O}}h_{22\mathcal{O}}}{1 + h_{21\mathcal{O}}} - h_{12\mathcal{O}}$
$h_{21\mathcal{O}} = \frac{-h_{21B}}{1 + h_{21B}}$	$h_{21B} = \frac{h_{21\mathcal{O}}}{1 + h_{21\mathcal{O}}}$

$$h_{22\varepsilon} = \frac{h_{22Б}}{1 + h_{21Б}} \qquad h_{22\varepsilon} = \frac{h_{22\varepsilon}}{1 + h_{21\varepsilon}}$$

При установлении связи между h -параметрами и параметрами эквивалентной схемы легко доказывается, что коэффициент передачи тока в схеме с ОБ при коротком замыкании выходной цепи $h_{21Б}$ равен физическому параметру БТ - коэффициенту передачи тока эмиттера α , а коэффициент передачи тока в схеме с ОЭ - равен β .

$$h_{21Б} = -\alpha; \quad h_{21\varepsilon} = \beta \qquad (6.41)$$

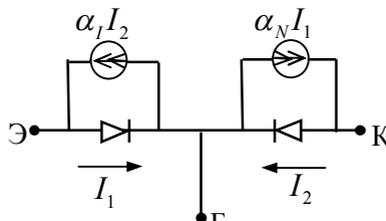
В настоящее время в справочной литературе используют как параметры БТ не значения α и β , а значения коэффициентов h -параметров - $h_{21Б}$ и $h_{21\varepsilon}$, а также частоту f_α обозначают $f_{h_{21Б}}$ и частоту f_β обозначают $f_{h_{21\varepsilon}}$.

По поводу метода определения h -параметров можно отметить, что их можно найти, используя формулы, измерить с помощью приборов (измерителей h -параметров) или рассчитать по семейству входных (h_{11} и h_{12}) и выходных (h_{21} и h_{22}) характеристик.

6.14. Модели биполярного транзистора

Для проектирования электронных схем на БТ с помощью ЭВМ необходимо представить транзистор в виде определенной модели. В любом случае модель прибора представляет совокупность схемы замещения и аналитических выражений, позволяющих определять параметры модели. Наиболее широко для этих целей используются две модели: Эберса-Молла и зарядоуправляемая модель.

Схема замещения БТ по Эберсу-Моллу представлена на рис. 6.33.



Исходные выражения модели следующие:

$$\left. \begin{aligned} I_{\text{Э}} &= I_1 - \alpha_I I_2 \\ I_K &= \alpha_N I_1 - I_2 \end{aligned} \right\} \quad (6.42)$$

Значения токов I_1 и I_2 задаются выражениями (6.43):

$$\left. \begin{aligned} I_1 &= I_{\text{ЭБК}} \left(\exp \frac{qU_{\text{ЭБ}}}{kT} - 1 \right) \\ I_2 &= I_{\text{КБК}} \left(\exp \frac{qU_{\text{КБ}}}{kT} - 1 \right) \end{aligned} \right\} \quad (6.43)$$

Заменяя значения начальных токов на обратные токи через соотношения (6.44)

$$I_{\text{ЭБК}} = \frac{I_{\text{ЭБ0}}}{1 - \alpha_I \alpha_N}; \quad I_{\text{КБК}} = \frac{I_{\text{КБ0}}}{1 - \alpha_I \alpha_N} \quad (6.44)$$

Запишем выражение 6.42 в конечном виде (6.45):

$$\left. \begin{aligned} I_{\text{Э}} &= \frac{I_{\text{ЭБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{ЭБ}}}{kT} - 1 \right) - \frac{\alpha_I I_{\text{КБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{КБ}}}{kT} - 1 \right) \\ I_K &= \frac{\alpha_N I_{\text{ЭБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{ЭБ}}}{kT} - 1 \right) - \frac{I_{\text{КБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{КБ}}}{kT} - 1 \right) \\ I_B &= \frac{(1 - \alpha_N) I_{\text{ЭБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{ЭБ}}}{kT} - 1 \right) + \frac{(1 - \alpha_I) I_{\text{КБ0}}}{1 - \alpha_I \alpha_N} \left(\exp \frac{qU_{\text{КБ}}}{kT} - 1 \right) \end{aligned} \right\} \quad (6.45)$$

Таким образом совокупность схемы замещения (рис. 6.29) и системы уравнений (6.45) представляют модель Эберса-Молла.

Основное достоинство модели Эберса-Молла – наглядность БТ, заложенная в уравнениях (6.45). Модель применима для всех режимов работы БТ.

Зарядоуправляемая модель.

Зарядоуправляемая модель относится к категории динамических. Схема замещения БТ представлена на рис. 6.34, а уравнения модели заданы выражениями (6.46).

$$\left. \begin{aligned} I_{\text{Э}} &= \frac{dQ_{\text{Э}}}{dt} + \frac{Q_{\text{Э}}}{\tau_{\text{Э}}} - \alpha_1 \frac{Q_{\text{К}}}{\tau_{\text{Э}}} \\ I_{\text{К}} &= \frac{dQ_{\text{К}}}{dt} + \frac{Q_{\text{К}}}{\tau_{\text{К}}} - \alpha_1 \frac{Q_{\text{Э}}}{\tau_{\text{К}}} \end{aligned} \right\} \quad (6.46)$$

Недостатком модели является отсутствие наглядной связи между параметрами БТ и значением зарядов.

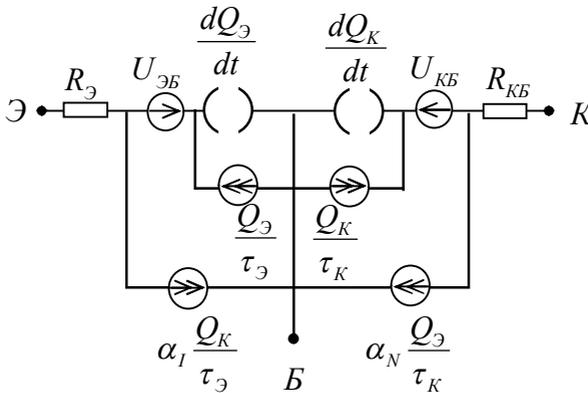


Рисунок 6.34. Схема замещения БТ

6.15. Некоторые разновидности биполярных транзисторов

В данном разделе рассмотрены некоторые разновидности БТ: составной транзистор (транзистор Дарлингтона), однопереходной транзистор, инжекционный транзистор, лавинный и гетеробиполярный транзисторы.

Составной транзистор (транзистор Дарлингтона).

Для получения мощных транзисторов необходимо обеспечить возможность подавать большие (до 1000 В и более) напряжения на коллекторный переход. Для исключения прокола

базы толщина базы должна быть достаточно большой, что задает низкое значение коэффициента передачи тока базы. Для мощных одиночных БТ значение β обычно не превышает десяти единиц. Проблема повышения β в мощных транзисторах решена путем применения составного транзистора, называемого также транзистором Дарлингтона, представленного на рис. 6.35. В этом случае коэффициент передачи тока базы β определяется произведением β_1 и β_2 транзисторов VT_1 и VT_2 :

$\beta \approx \beta_1 \cdot \beta_2$ и достигает приемлемых для мощных транзисторов значений.

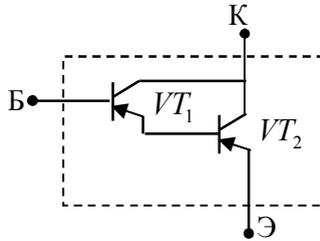
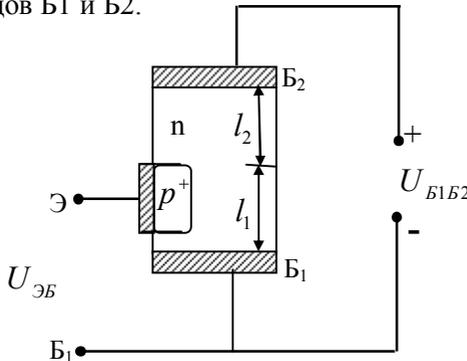


Рисунок 6. 35. Схема составного транзистора

Однопереходной транзистор (двухбазовый диод). Структура однопереходного транзистора представлена на рис. 6.36.

Транзистор представляет брусок полупроводника с омическими контактами в торцевых частях, создающими выводы электродов Б1 и Б2.



В нижней части кристалла формируется эмиттерный переход. Рисунок 6.36. Структура однопереходного транзистора

нопереходного транзистора является входная характеристика, представленная на рис. 6.37.

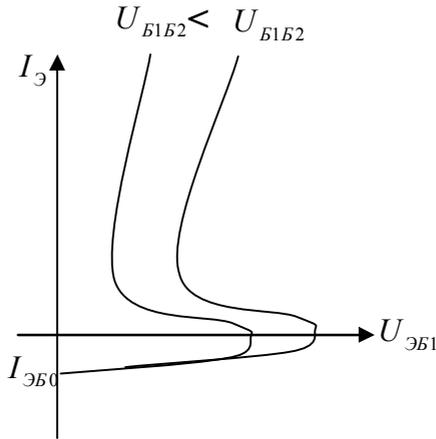


Рисунок 6.37. Входная характеристика однопереходного транзистора

Если напряжение на эмиттерный переход не подано, то за счет протекания тока между базой Б1 и базой Б2 создается падение напряжения на участке l_1 , вызывающее обратное смещение эмиттерного перехода и протекание в цепи эмиттерного тока $I_{\text{эб0}}$.

При подаче прямого смещения между эмиттером и базой Б1 происходит постепенное уменьшение обратного напряжения на эмиттерном переходе.

Когда падение напряжения на участке l_1 станет равным напряжению $U_{\text{эб1}}$, эмиттерный переход начнет смещаться в прямом направлении. При этом из эмиттера в базу на участке l_1 начнется инжекция дырок. Это приведет к уменьшению сопротивления на участке l_1 и уменьшению напряжения обратного смещения эмиттерного перехода, что равносильно увеличению прямого смещения эмиттерного перехода. Увеличение прямого смеще-

ния эмиттерного перехода вызывает увеличение уровня инжекции дырок и дальнейшее уменьшение сопротивления на участке l_1 , что приводит к уменьшению падения напряжения ΔU_{l_1} , и увеличению прямого смещения эмиттерного перехода. Таким образом в транзисторе возникает положительная обратная связь, приводящая к появлению участка отрицательного дифференциального сопротивления на ВАХ.

Эквивалентная схема однопереходного транзистора представлена на рис. 6.38, а условно-графические обозначения на рис. 6.39.

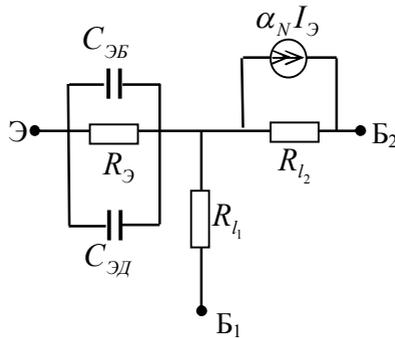


Рисунок 6.38. Эквивалентная схема однопереходного транзистора



Рисунок 6.39. Условно-графические обозначения однопереходных транзисторов

Инжекционный транзистор (биполярный транзистор с инжекционным питанием). Структура БТ с инжекционным питанием представлена на рис. 6.40. Кроме традиционных для БТ выводов в рассматриваемом приборе существует еще один вывод – инжектор.

Структуру инжекционного БТ можно представить как совокупность двух БТ:

$p_1^+ - n_1 - p_2$ – токозадающий БТ
 $n_1 - p_1 - n_2$ – переключающий БТ.

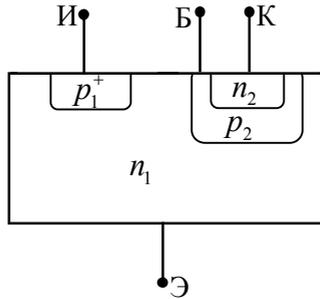


Рисунок 6.40. Структура БТ с инжекционным питанием

Принцип работы инжекционного транзистора сводится к следующему. При прямом включении инжекторного р-п перехода в слой n_1 инжектируются дырки. За счет градиента концентрации они доходят до эмиттерного перехода $n_1 - p_2$ и его полем экстрагируются в слой p_2 -базы переключающего транзистора и вызывают прямое смещение эмиттерного перехода. Инжектируемые из эмиттерного перехода $n_1 - p_2$ электроны собираются коллекторным переходом, создавая в цепи коллектора ток, величина которого зависит от заряда дырок, поступающих из инжекторного перехода, т.е. от тока инжектора. Выходные характеристики транзистора представлены на рис. 6.41, а его условно-графические обозначения – на рис. 6.42.

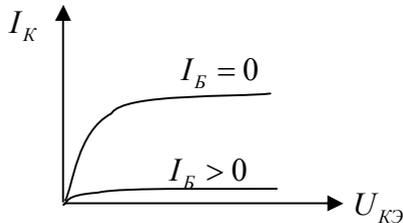


Рисунок 6.41. Выходные характеристики инжекционного транзистора

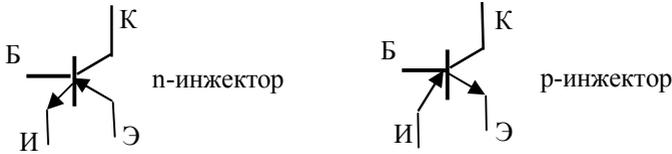


Рисунок 6.42. Условно-графическое изображение инжекционного транзистора

Лавинный транзистор

На рисунке 6.43. представлена схема включения БТ, поясняющая сущность работы лавинного БТ, а на рис. 6.44. выходные ВАХ.

Сопротивление резистора R_B подобрано так, что при малых напряжениях $U_{КЭ}$ сопротивление эмиттерного перехода $R_Э \gg R_B$ и ток течет по цепи 1, т.е. реализуется условие $I_Э = 0$.

Однако при определенном напряжении $U_{КЭ}$ падение прямого напряжения на эмиттерном переходе возрастает и сопротивление $R_Э$ становится меньше, так, что выполняется условие $R_Э \ll R_B$. При этом ток протекает по цепи 2, т.е. реализуется ситуация, что $I_B = 0$ (рис.6.43).

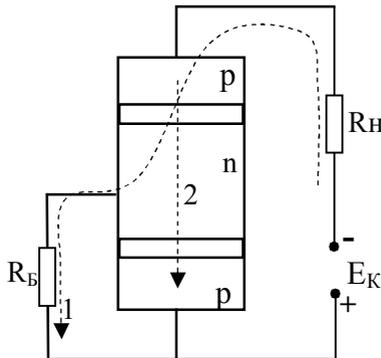


Рисунок 6.43. Схема включения лавинного транзистора

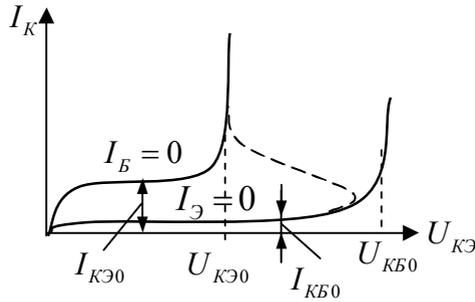


Рисунок 6.44. Выходные характеристики БТ при $I_B = 0$ и $I_B \neq 0$

Переход от режима пробоя при условии $I_B = 0$ к условию $I_B \neq 0$ происходит с наличием положительной обратной связи, так что выходная ВАХ лавинного БТ имеет вид, представленный на рис. 6.45.

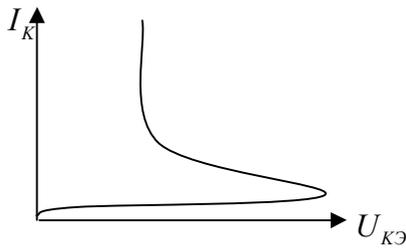


Рисунок 6.45. Выходные ВАХ лавинного транзистора

Условно-графическое обозначение лавинного БТ представлено на рис. 6.46.

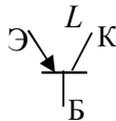


Рисунок 6.46. Условно-графическое обозначение лавинного транзистора

Основное применение лавинных транзисторов – создание мощных генераторов импульсов с крутыми фронтами.

Гетеробиполярный транзистор – это БТ, у которого эмиттерный переход выполнен в виде гетероперехода или эмиттерный и коллекторный переходы являются гетеропереходами.

Замена эмиттерного гомоперехода на гетеропереход позволяет увеличить эффективность эмиттера до единицы при возможности низкого уровня легирования слоя эмиттера и высокого уровня легирования базы.

Использование коллекторного перехода в виде гетероперехода позволяет уменьшить длительность переходных процессов из-за отсутствия инжекции неосновных носителей из коллектора в базу в режиме насыщения.

И в том и другом случаях используется основное отличие гетероперехода от гомоперехода – ток через переход осуществляется носителями одного знака.

6.16. Основные параметры биполярных транзисторов и их ориентировочные значения

1) коэффициенты передачи эмиттерного и базового тока (дифференциальные коэффициенты передачи, которые в первом приближении считают равными интегральным)

$$h_{21Э} = \left. \frac{dI_K}{dI_B} \right|_{U_{KB} = \text{const}} > 1; \quad h_{21Б} = \left. \frac{dI_K}{dI_Э} \right|_{U_{KB} = \text{const}} < 1;$$

2) дифференциальное сопротивление эмиттерного перехода

$$R_Э = \left. \frac{dU_{ЭБ}}{dI_Э} \right|_{U_{KB} = \text{const}} \quad (r_Э - \text{Ом} \div \text{десятки Ом});$$

3) обратный ток коллекторного перехода при заданном обратном напряжении

$$I_{КБ0} = I_K \Big|_{I_Э=0}; \quad U_{KB} < 0 \quad (I_{КБ0} - \text{несколько нА} \div \text{десятки мА});$$

4) объемное сопротивление базы $R'_б$ (десятки - сотни Ом);

5) коэффициент внутренней обратной связи по напряжению ($h_{12} = 10^{-3} \div 10^{-4}$);

б) выходная проводимость h_{22} или дифференциальное сопротивление коллекторного перехода

$$R_{KB} = \frac{1}{h_{22B}} = \left. \frac{dU_{KB}}{dI_K} \right|_{I_3 = const}; \quad R_{KЭ} = \frac{1}{h_{22Э}} = \left. \frac{dU_{KЭ}}{dI_K} \right|_{I_B = const}$$

(h_{22} - доли ÷ сотни микро См); $R_{KB} = R_{KЭ}(1 + h_{21Э})$.

7) максимально допустимый ток коллектора $I_{K \max}$ (сотни мА ÷ десятки А);

8) напряжение насыщения коллектор - эмиттер $U_{KЭнас}$ (десятые доли ÷ один вольт);

9) наибольшая мощность рассеяния коллектором

$P_{K \max}$ (мВт ÷ десятки Вт);

10) емкость коллекторного перехода C_K (пФ ÷ десятки пФ);

11) тепловое сопротивление между коллектором транзистора и корпусом $R_T = \frac{\Delta T}{P_{K \max}}$, где ΔT - перепад температур между коллекторным переходом и корпусом;

12) предельная частота коэффициента передачи тока $f_{h_{21}}$ или $\omega_{h_{21}}$, на которой коэффициент передачи тока h_{21} уменьшается до 0,7 своего низкочастотного значения: $\omega_{h_{21Э}} \approx \omega_\beta$;

$\omega_{h_{21Б}} \approx \omega_\alpha$ (задаются или $f_{h_{21Б}}$ или $f_{h_{21Э}}$). $f_{h_{21Э}}$ - кГц ÷ сотни МГц. Иногда вместо предельной задают граничную частоту f_{ep} или ω_{ep} , как частоту, при которой $f_{h_{21Э}} \rightarrow 1$;

13) максимальная частота генерации $f_{\max} \approx \sqrt{f_{ep} / (30R'_B C_{KB})}$ - это наибольшая частота, при которой транзистор может работать в схеме автогенератора. На этой частоте коэффициент усиления транзистора по мощности равен единице.

6.17. Маркировка транзисторов

Обозначения биполярных транзисторов состоят из шести или семи элементов. **Первый элемент** – буква, указывающая исходный материал: Г – германий, К – кремний, А – арсенид галлия. Для транзисторов специального назначения первый элемент – цифра: 1 – германий, 2 – кремний, 3 – арсенид галлия. **Второй элемент** – буква Т. **Третий элемент** – число, присваиваемое в зависимости от частоты и мощности транзистора (табл. 6.3). **Четвертый, пятый и шестой элементы** – цифра, означающая порядковый номер разработки. **Шестой (седьмой) элемент** – буква, указывающая разновидность типа из данной группы приборов. Например, ГТ108А, 2Т144А, КТ3102А и т.д.

Классификация БТ проводится

1) **по мощности**: маломощные – менее 0,3 Вт; средняя мощность – 0,3 ÷ 1,5 Вт; мощные – более 1,5 Вт;

2) **по граничной частоте**: низкочастотные – до 3МГц; средней частоты - 3 ÷ 30МГц; высокочастотные – более 30 МГц.

Таблица 6.3.

Обозначение транзистора	Мощность, рассеиваемая транзистором	Граничная частота, МГц				
		до 3	до 30	более 30	30 -300	свыше 300
Шести-значное	Малая	1	2	3		
	Средняя	4	5	6		
	Большая	7	8	9		
Семи-значное	До 1 Вт		1		2	4
	Свыше 1 Вт		7		8	9

По конструктивно-технологическим признакам БТ делятся на сплавные, диффузионно-сплавные, диффузионно-планарные, эпитаксиально-планарные и мезапланарные.

7. ТИРИСТОРЫ

Тиристор — это полупроводниковый прибор с тремя или более р-п переходами, на вольт-амперной характеристике которого имеется участок с отрицательным дифференциальным сопротивлением. Тиристор может находиться в одном из двух устойчивых состояний - закрытом или открытом. В закрытом со-

стоянии сопротивление тиристора высоко и он пропускает малый ток. В открытом состоянии сопротивление тиристора мало и через него протекает большой ток.

Типичная структура тиристора - четырехслойная, с чередующимися слоями полупроводника р-типа и n-типа: $p_1-n_1-p_2-n_2$ (рис. 7.1,а)

Схематическое устройство тиристора показано на рис. 7.1. Исходным материалом служит кремний n-типа, в кристалле которого создается структура $p-n-p-n$.

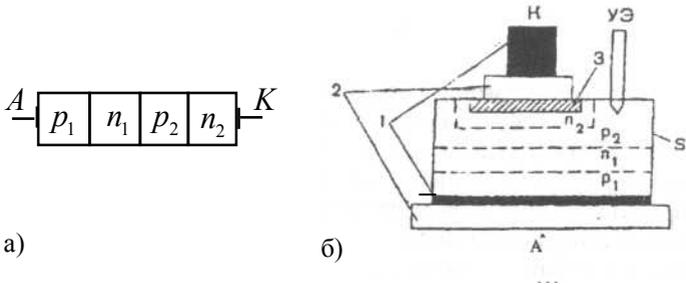


Рис. 7.1. Тиристор:

а — упрощенная структура; б — схематическое устройство;
1 — алюминий; 2 — молибден; 3 — золото-сурьма

Пластины кремния с готовой четырехслойной структурой припаивают к кристаллодержателю. Контактные площадки создают металлизацией, а соединение их с внешними выводами осуществляется через вольфрамовые прокладки. Герметизированный корпус предохраняет кристалл от воздействия окружающей среды.

В зависимости от конструктивных особенностей и свойств тиристоры делят на диодные и триодные.

В *диодных* тиристорах различают: тиристоры, запираемые в обратном направлении; проводящие в обратном направлении; симметричные.

Триодные тиристоры подразделяют: на запираемые в обратном направлении с управлением по аноду или катоду; проводящие в обратном направлении с управлением по аноду или катоду; симметричные (двунаправленные). Кроме того, в их составе различают группу выключаемых тиристоров.

Условные обозначения тиристоров приведены на рис. 7.2.

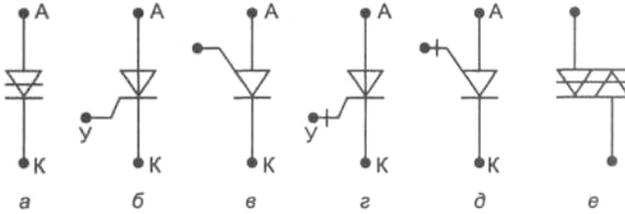


Рисунок 7.2. Условные графические обозначения тиристоров: а – диодный тиристор (динистор); б – незапираемый триодный тиристор (тринакстор) с управление по катоду; в – незапираемый триодный тиристор с управлением по аноду; г – запираемый тринакстор с управлением по катоду; д – запираемый тринакстор с управлением по аноду; е – симметричный диодный тиристор

Для понимания работы всех видов тиристоров необходимо уяснить принцип действия диодного тиристора (рис. 7.3). Для этого необходимо представлять тиристор как совокупность двух разноструктурных БТ, работающих при прямом включении тиристора (рис. 7.3), в активном режиме. При таком включении переходы 1 и 3 смещены в прямом направлении, а переход 2 – в обратном направлении.

Ток в цепи анода в этом случае определяется сопротивлением обратнсмещенного коллекторного перехода 2, он мал и равен $I_{КЕ0}$ (рис. 7.5, участок 1).

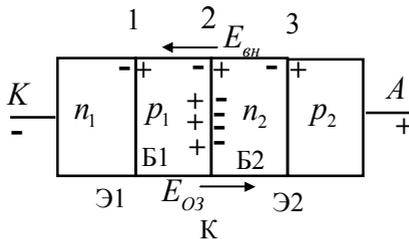


Рисунок 7.3. Рисунок, поясняющий принцип действия диодного транзистора

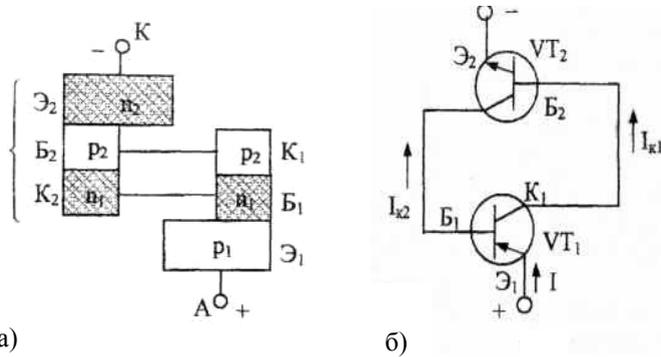


Рисунок 7.4. Представление тиристора как совокупности двух биполярных транзисторов (б,в)

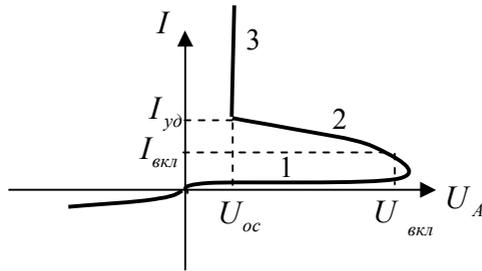


Рисунок 7.5. ВАХ диодного тиристора

По мере увеличения напряжения на аноде возрастает падение напряжения на эмиттерных переходах 1 и 3. Поскольку транзисторы работают в активном режиме, то это приводит к тому, что электроны из эмиттера Э1 инжектируются в базу Б1 и пройдя ее экстрагируются полем коллекторного перехода E_{ex} и накапливаются в базе Б2 (знаки “-“ в Б2). Одновременно инжектированные переходом Э2 дырки накапливаются в базе Б1, создавая положительный объемный заряд (знаки “+” в Б1). Возникающие заряды в базах создают электрическое поле E_{O3} , вектор которого направлен противоположно вектору электрического поля коллекторного перехода $E_{вн}$. Как только поле E_{O3} станет

больше поля $E_{вн}$, коллекторный переход из обратносмещенного перейдет в прямосмещенный и тиристор откроется. Т.е. в нем установится большой ток, поскольку все три p-n перехода будут смещены в прямом направлении и имеют малое сопротивление (участок 3 на ВАХ). Таким образом, для того чтобы тиристор открылся необходимо, чтобы коллекторный переход из обратносмещенного стал бы прямосмещенным. Это достигается за счет поля объемных зарядов E_{O3} , создаваемого электронами и дырками в базах Б1 и Б2. Если по какой-либо причине это поле станет меньше $E_{вн}$, то тиристор закроется, т.е. перейдет в состояние с высоким сопротивлением (участок 1 на ВАХ).

Переход из состояния 1 в состояние 3 происходит достаточно быстро из-за наличия положительной обратной связи. Она проявляется в том, что накапливающиеся в базе Б2 электроны увеличивают прямое смещение перехода Э2, что вызывает повышенный уровень инжекции перехода Э2 и увеличение положительного заряда в базе Б1. Это, в свою очередь, увеличивает прямое смещение перехода Э1 и уровня инжекции электронов, которые попадая в базу Б2 вновь увеличивают прямое смещение перехода Э1 и уровня его инжекции. Состояние 2 в тиристоре является неустойчивым и в работе не используется. Процессом накопления носителей зарядов в одной из баз можно управлять с помощью дополнительного электрода к базам Б1 или Б2 в триодном тиристоре, ВАХ которой представлена на рис. 7.6.

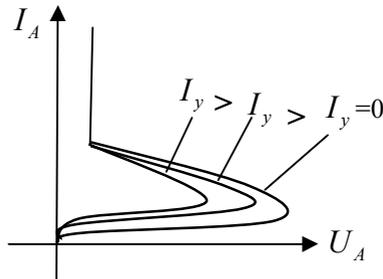


Рисунок 7.6. ВАХ триодного тиристора

Триодные тиристоры отличаются от диодных тем, что одна из баз имеет внешний вывод, который называют управляющим

электродом (УЭ). УЭ необходим для подключения цепи управления. Назначение цепи управления состоит в управлении моментом включения тиристора при напряжениях в основной цепи меньших, чем напряжение переключения.

Основные параметры тиристорov:

- ✓ напряжение переключения (постоянное - $U_{пр.к}$, импульсное (десятки – сотни вольт)
- ✓ напряжение в открытом состоянии U_{oc} - падение напряжения на тиристоре в открытом состоянии ($U_{oc} = 1 \div 3$ В);
- ✓ обратное напряжение $U_{обр}$ - напряжение, при котором тиристор может работать длительное время без нарушения его работоспособности (единицы—тысячи В);
- ✓ ток в открытом состоянии I_{oc} - максимальное значение тока открытого тиристора (сотни мА - сотни А).
- ✓ ток удержания $I_{уд}$ (десятки - сотни мА);
- ✓ обратный ток $I_{обр}$ (доли мА);
- ✓ отпирающий ток управления I_{yom} — наименьший ток управляющего электрода, необходимый для включения тиристора (десятки мА);
- ✓ скорость нарастания напряжения в закрытом состоянии - максимальная скорость нарастания напряжения в закрытом состоянии (десятки - сотни В/мкс);
- ✓ время включения $t_{вкл}$ - интервал времени, в течение которого тиристор переходит из закрытого состояния в открытое (мкс - десятки мкс);
- ✓ время выключения $t_{выкл}$ - наименьший интервал времени, в течение которого восстанавливаются запирающие свойства тиристора (десятки - сотни мкс); по прошествии этого времени на тиристор может быть снова подано прямое напряжение; тиристор остается закрытым до следующего отпирающего импульса.

8. ПОЛЕВЫЕ ТРАНЗИСТОРЫ

Полевыми транзисторами называют трехэлектродные полупроводниковые приборы, в которых управление током осуществляется изменением проводимости токопроводящего канала путем воздействия электрического поля, поперечного к направлению тока. Токопроводящий канал соединяет две сильнолегированных области. Область, из которой носители заряда уходят в канал, называется *истоком*, а область, в которую они приходят, - *стоком*. Электрическое поле, изменяющее проводимость канала, создается путем подачи управляющего напряжения на электрод, называемый *затвором*. В полевых транзисторах от истока к стоку перемещаются только основные носители заряда (либо электроны, либо дырки), поэтому их часто называют *униполярным*, в отличие от биполярных, в которых перемещаются как основные, так и неосновные носители заряда.

Существует две разновидности полевых транзисторов, различающихся физической структурой и способом управления проводимостью канала: полевые транзисторы с управляющим переходом (p-n переходы или контакт Шоттки) и транзисторы с изолированным затвором. В транзисторах *с управляющим p-n переходом* в качестве затвора используется область, тип электропроводности которой противоположен типу электропроводности канала, в результате чего между затвором и каналом образуется p-n переход. В транзисторах с управляющим переходом *металл—полупроводник* металлический затвор образует с приповерхностным слоем канала выпрямляющий контакт (барьер Шоттки). В транзисторах *с изолированным затвором* между металлическим затвором и проводящим каналом расположен тонкий слой диэлектрика так, что образуется структура *металл-диэлектрик - полупроводник* (МДП-структура). Такие транзисторы обычно называют *МДП-транзисторами*. Они подразделяются на транзисторы с индуцированным и встроенным каналом. Особенностью всех полевых транзисторов является незначительный ток в цепи затвора. В этом состоит важнейшее отли-

чие полевых транзисторов от биполярных, во входной цепи которых протекает сравнительно большой ток.

8.1. Полевые транзисторы с управляющим р-п переходом

В транзисторах с управляющим р-п переходом проводимость канала, соединяющего исток со стоком, изменяется путем изменения ширины ОПЗ р-п перехода, изолирующего затвор от канала.

Устройство транзистора с управляющим р-п переходом и каналом п-типа показано на рис. 8.1. На кремниевой подложке р-типа создается карман с электронной электропроводностью, в котором формируются две сильнолегированные области p^+ -типа, выполняющие функции истока и стока.

Над слоем п-типа формируется слой р-типа, выполняющий функции затвора. Между затвором р-типа и каналом п-типа образуется р-п переход, отделяющий затвор от канала. Концентрация примеси в затворе значительно превышает концентрацию примеси в канале, поэтому р-п переход толщиной d почти целиком расположен в канале. Толщина канала d_K зависит от отрицательного напряжения на затворе $u_{з-у}$, изменяющего d , то есть $d_K = d_0 - d$, где d_0 - металлургическая толщина канала.

Чем больше отрицательное напряжение $u_{з-у}$, тем больше ширина ОПЗ р-п перехода и, соответственно, меньше толщина канала d_K и его проводимость. При некотором напряжении $u_{з-у} = u_{омс}$, называемом *напряжением отсечки*, канал полностью перекрывается.

$$U_{омс} = \frac{qN_K d_0^2}{2\epsilon\epsilon_0} - \varphi_k^0,$$

где d_0 - начальная ширина канала,

N_K - концентрация примесей в канале.

Принцип действия данного вида ПТ очень прост: ток между истоком и стоком модулируется напряжением на затворе. При этом причиной изменения тока является изменение геометрических размеров канала. За счет изменения обратного напряжения

на управляющем р-п переходе меняется ширина его ОПЗ и, как следствие, ширина канала.

При подаче на сток положительного напряжения u_{c-u} в канале возникает ток i_c , создающий падение напряжения вдоль канала $\varphi(x)$, вследствие чего толщина d и, соответственно, d_K оказываются зависящими от координаты (рис. 8.2).

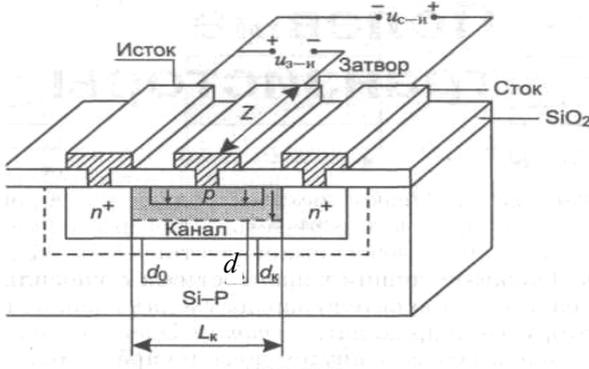


Рисунок 8.1. Устройство транзистора с управляющим р-п переходом

При небольших значениях u_{c-u} (рис. 8.2, а) канал вдоль оси x постепенно сужается.

При некотором значении $u_{c-u} = u_{нас}$ называемом *напряжением насыщения*, канал у стока полностью перекрывается (рис.8.2, б). При напряжении $u_{c-u} > u_{нас}$ участок перекрытия δ расширяется, а проводящая часть канала L'_k укорачивается (рис. 8.2, в).

В этом режиме внешнее напряжение u_{c-u} перераспределяется между каналом и областью перекрытия.

На проводящем участке канала напряжение сохраняется равным $u_{нас}$, а к перекрытому участку канала оказывается приложенным напряжением $u_{c-u} - u_{нас}$. Режим работы с частично

перекрытым каналом называется *режимом насыщения*, а с неперекрытым каналом - *линейным режимом*.

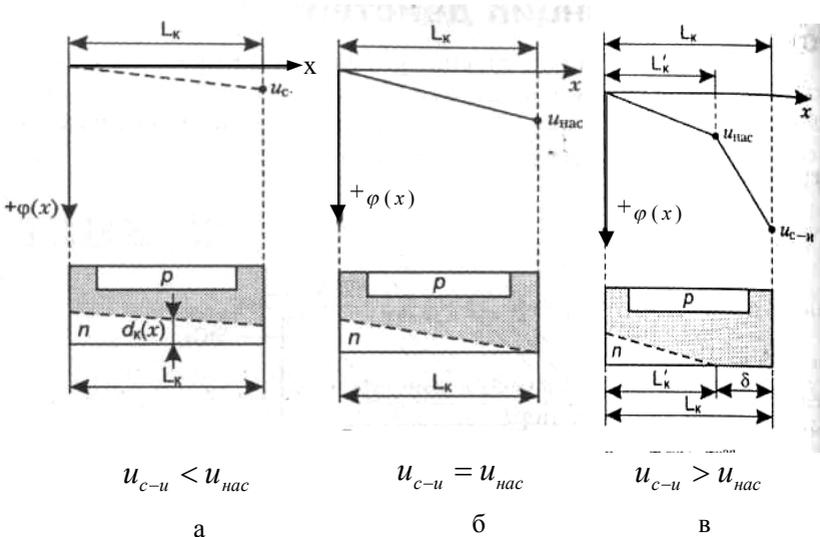


Рисунок 8.2. Зависимость параметров полевого транзистора от координаты

8.2. Дифференциальные параметры

Ток стока зависит от напряжения на затворе $u_{з-и}$ и от напряжения стока $u_{c-и}$.

Поэтому выражение для полного дифференциала тока можно представить в виде

$$di_c = \frac{\partial I_c}{\partial u_{з-и}} du_{з-и} + \frac{\partial I_c}{\partial u_{c-и}} du_{c-и} \quad (8.1)$$

В этом выражении частные производные, определяющие приращения тока при изменении соответствующих напряжений, можно рассматривать как дифференциальные параметры транзистора.

Крутизна S : $S = \partial i_c / \partial u_{3-u}$ характеризует управляющее действие затвора. Ее измеряют в миллиамперах на вольт $\left[\frac{mA}{B} \right]$

и определяют по передаточной характеристике, как это показано на рис. 8.3,б, заменяя бесконечно малые приращения конечными, т.е.

$$S = \frac{\Delta I_C}{\Delta U_{3-u}} \Big|_{U_{c-u} = const}$$

Крутизна обратно пропорциональна сопротивлению канала, которое равно $R_{K0} = \frac{\rho \cdot L_K}{z \cdot d_0}$ где z - ширина канала (см. рис.8.1)

Следовательно, для получения высокой крутизны необходимо иметь канал с малой длиной L_k и большой шириной Z .

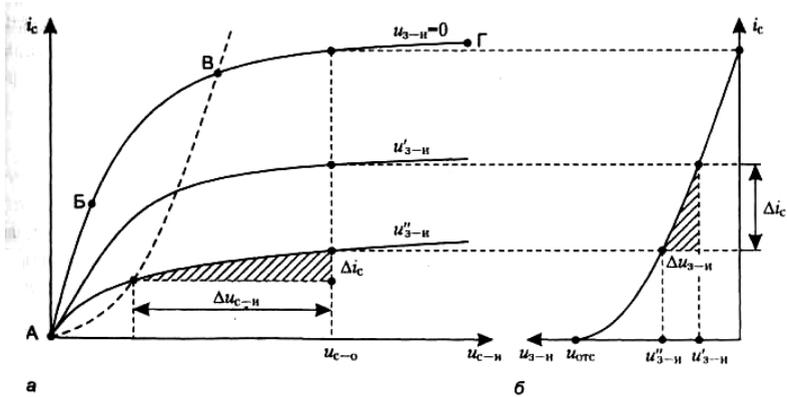


Рисунок 8.3. Выходные и передаточные характеристики

Увеличивать толщину канала d_k и уменьшать удельное сопротивление ρ нецелесообразно, т.к. это ведет к увеличению напряжения отсечки. Практически длина канала составляет несколько микрон, а его ширина – тысячи микрон.

Выходная проводимость G_i $G_i = \frac{\Delta I_C}{\Delta U_{c-н}} \Big|_{U_{3-н} = const}$ характеризует влияние напряжения стока на ток стока. Ее опреде-

ляют по выходным характеристикам (рис.8.3), заменяя бесконечно малые приращения конечными, т.е. $G_i = \frac{\Delta I_C}{\Delta U_{C-II}}$, где

ΔI_C - изменение тока (вертикальный катет характеристического треугольника), вызванное изменением напряжения ΔU_{C-II} (горизонтальный катет характеристического треугольника). Величина выходной проводимости очень мала и обусловлена изменением длины канала при изменении напряжения стока. Чем короче канал, тем больше относительное изменение его длины при одном и том же приращении ΔU_{C-II} , следовательно, тем больше проводимость G_i . Часто вместо параметра G_i применяют обратную величину $R_i = 1/G_i$. Этот параметр называется внутренним сопротивлением транзистора.

Коэффициент усиления μ характеризует сравнительное воздействие напряжений стока и затвора на ток стока. Он равен отношению приращений напряжений стока и затвора, вызывающих одинаковые по величине и противоположные по знаку приращения тока стока

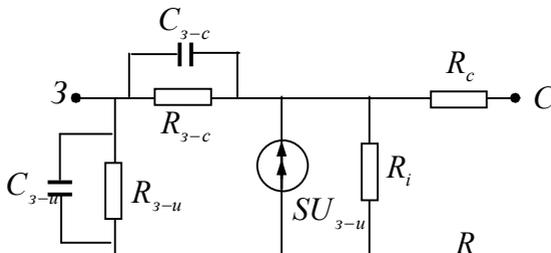
$$\mu = - \frac{dU_{C-II}}{dU_{3-II}} \Big|_{I = const} \quad (8.2)$$

Знак «минус» в этой формуле учитывает, что положительному приращению dU_{C-II} увеличивающему ток на dI_C , соответствует отрицательное приращение dU_{3-II} , уменьшающее ток на ту же самую величину dI_C , благодаря чему обеспечивается постоянство тока I_C .

Дифференциальные параметры полевого транзистора связаны между собой соотношением:

$$SR_i = \mu. \quad (8.3)$$

Эквивалентная схема транзистора представлена на рис. 8.4.



8.3. Полевые транзисторы с управляющим переходом металл - полупроводник

Полевые транзисторы с управляющим переходом металл-полупроводник (МЕП-транзисторы) находят применение в арсенид-галлиевых интегральных схемах, характеризующихся повышенным быстродействием, что обусловлено рядом преимуществ арсенида галлия по сравнению с кремнием. Арсенид галлия имеет более широкую запрещенную зону ($\Delta E_g = 1,43$ эВ), благодаря чему подложка может рассматриваться как диэлектрик ($\rho \approx 10^7 \dots 10^9 \text{ Ом} \cdot \text{см}$). Подвижность электронов в слабых электрических полях в арсениде галлия примерно в 5 раз выше, чем в кремнии ($\mu_n \approx 8 \cdot 10^3 \text{ см}^2/\text{В} \cdot \text{с}$), а скорость насыщения в сильных полях больше примерно в 2,5 раза ($V_{нас} \approx 2 \cdot 10^7 \text{ см/с}$). По некоторым параметрам арсенид галлия хуже кремния. Так, у него низкая подвижность дырок и малое время жизни неосновных носителей заряда, что затрудняет разработку биполярных транзисторов. Слишком высокая плотность поверхностных состояний не позволяет создавать на его основе МДП-транзисторы.

Простейшая структура арсенид-галлиевого МЕП-транзистора представляет из себя следующее. Транзистор создается на подложке из нелегированного арсенида галлия. У поверхности создаются области истока и стока n^+ -типа, а также тонкий слой n -типа толщиной $d_0 = 0,1 \div 0,2$ мкм с концентрацией примеси $N_d = 10^{17} \text{ см}^{-3}$. На поверхность подложки над слоем n -типа наносят металлический электрод затвора в виде сплава титан-вольфрам. Металлические электроды и из сплава золото-германий образуют омические контакты к областям истока и стока. Выводы от всех электродов формируются в окнах диэлектрика. Металлический электрод затвора образует со слоем n -типа выпрямляющий контакт-барьер Шоттки. Толщина обед-

ненногo слоя зависит от напряжения, подаваемого на затвор по указанным ранее причинам. Проводящий канал формируется между обедненным слоем и подложкой. Принцип действия МЭП-транзистора заключается в том, что при изменении напряжения на затворе изменяется толщина проводящего канала, а следовательно, его проводимость и ток стока. То есть принцип действия практически не отличается от принципа действия транзистора с управляющим р-п переходом. Для МЭП-транзистора справедливы соотношения, данные в предыдущем разделе. Существуют лишь количественные различия.

Полевые транзисторы с управляющим р-п переходом и МЭП-транзисторы относятся к категории нормально открытых: т.е. при отсутствии напряжения на затворе ток в цепи исток-сток при приложенном напряжении U_{C-II} протекает.

8.4. Полевые транзисторы с изолированным затвором

Полевые транзисторы с изолированным затвором отличаются от транзисторов с управляющим р-п переходом тем, что у них затвор, выполненный в виде металлической пленки, изолирован от канала тонким слоем диэлектрика. В качестве диэлектрика обычно используют диоксид кремния SiO_2 .

Полевые транзисторы делятся на транзисторы с индуцированным и встроенным каналами.

• Устройство и принцип действия

Устройство транзистора с изолированным затвором и встроенным затвором и каналом п-типа показано на рис. 8.5.

На подложке р-типа создают две сильнолегированные области n^+ -типа. Одна из этих областей является истоком, другая - стоком. Между истоком и стоком существует канал, который либо создается в процессе изготовления транзистора (встроенный канал), либо индуцируется при подаче на затвор положительного напряжения (индуцированный канал).

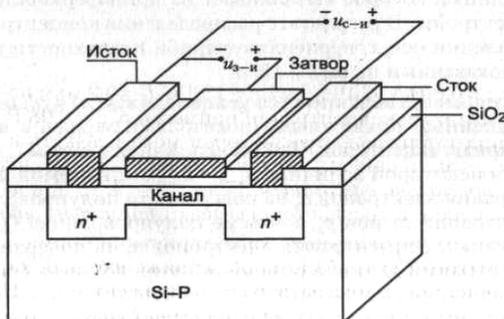


Рисунок 8.5. Устройство транзистора с изолированным затвором условно-графическое обозначение транзисторов со встроенным каналом n-типа показано на рис. 8.6, а, а транзисторов с индуцированным каналом - на рис. 8.6, б.

Помимо транзисторов с электронным каналом существуют транзисторы с дырочным каналом. Их схематические изображения показаны на рис. 8.6, в и г. В p-канальном транзисторе на сток подается отрицательный потенциал.

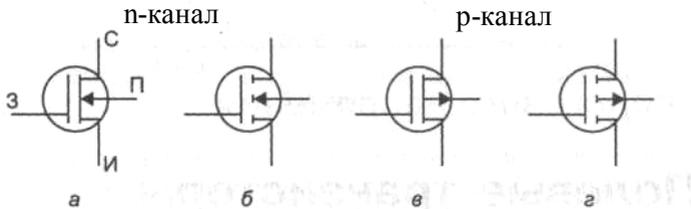


Рисунок 8.6. Схематическое обозначение транзисторов

При подаче на затвор отрицательного напряжения возникает вертикальное электрическое поле между подложкой и затвором, которое выталкивает электроны из канала, в результате чего проводимость канала уменьшается. При положительном напряжении канал обогащается электронами, и его проводимость возрастает.

В транзисторе с индуцированным каналом (рис.8.7) при подаче на сток положительного напряжения и нулевом напряжении на затворе проводящий канал между истоком и стоком отсутствует, т.к. один из p-n переходов (стоковый) смещен в обратном направлении. При подаче на затвор положительного напряжения возникает поперечное электрическое поле, направленное перпендикулярно поверхности полупроводника, которое выталкивает из приповерхностного слоя дырки и притягивает электроны.

При определенном напряжении на затворе, называемом пороговым напряжением ($U_{3-u.пор}$) произойдет явление инверсии поверхностной проводимости: на поверхности р-полупроводника под затвором образуется слой с n-типом проводимости и в цепи стока потечет ток.

Изменяя напряжение u_{3-u} можно изменять толщину канала d_K , а следовательно, его проводимость.

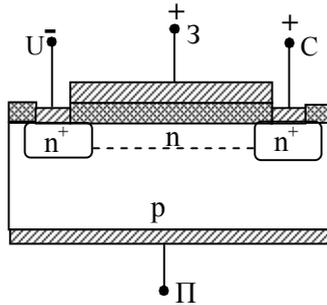


Рисунок 8.7. Структура ПТ с изолированным затвором, индуцированным каналом n-типа

Семейство выходных (а) и передаточных (б) характеристик для ПТ с индуцированным и встроенным каналом приведены на рис. 8.8 - 8.9.

На рис. 8.9, а - U'_{3-u} - напряжение на затворе, вызывающее обогащение канала носителями; U''_{3-u} - напряжение на затворе, вызывающее обеднение канала носителями.

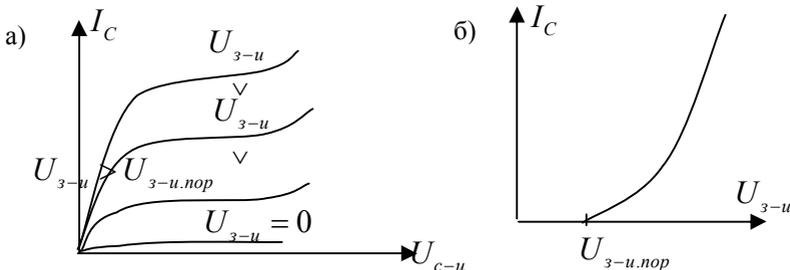


Рисунок 8.8. Семейство выходных (а) и передаточных (б) характеристик ПТ с индуцированным каналом

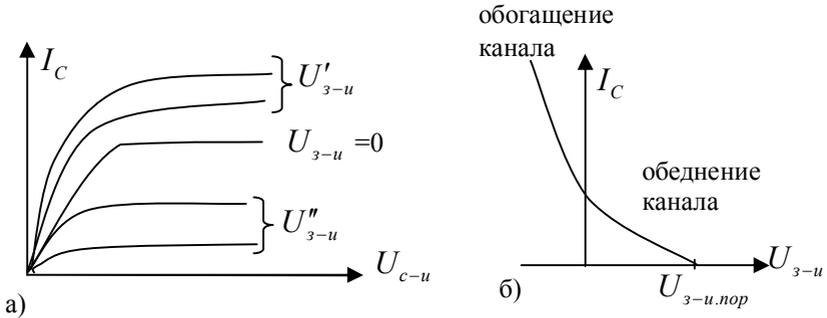


Рисунок 8.9. Семейство выходных (а) и передаточных (б) характеристик ПТ со встроенным каналом

Эквивалентная схема ПТ с изолированным затвором приведена на рис. 8.10.

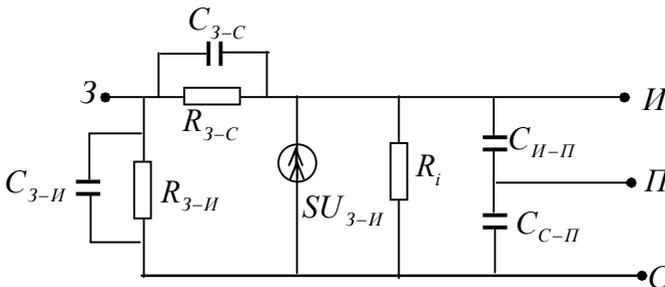


Рисунок 8.10. Эквивалентная схема ПТ с изолированным каналом

Параметры ПТ с изолированным каналом такие же, что и для транзисторов с управляющими переходами: S , R_i , μ .

8.5. Импульсный режим полевых транзисторов

В современных цифровых интегральных схемах широкое применение находят МДП-транзисторы, работающие в режиме электронного ключа (рис.8.11, а).

В этом режиме на затвор транзистора от источника сигнала $u_{Г}$, имеющего внутреннее сопротивление $R_{Г}$ подаются импульсы напряжения, отпирающие и запирающие МДП-транзистор.

Если на вход транзистора подается напряжение $u_{з-и0} < u_{пор}$, транзистор находится в запертом состоянии и на его выходе будет напряжение $u_{с-и1} \approx E_{u-n}$ (точка А на рис. 8.11,в). Если на затвор подается напряжение $u_{з-и1} > u_{пор}$, то транзистор находится в открытом состоянии и на его выходе будет невысокое напряжение $u_{си0}$, определяемое точкой пересечения нагрузочной линии с выходной характеристикой, соответствующей напряжению $u_{з-и} = u_{з-и1}$ (точка В на рис.8.11, в). Как правило, эта точка пересечения находится на восходящем участке выходной характеристики, соответствующем линейному режиму работы транзистора. Желательно, чтобы этот участок был достаточно крутым, для того чтобы напряжение $u_{си0}$ было как можно меньше. Ввиду наличия паразитных емкостей

$$C_{вх.экр.} = C_{з-и} + C_{з-с} (K_u + 1) \text{ и } C_H = C_{с-н} + C_{вх.экр.2}.$$

($C_{вх.экр.2}$ - входная емкость следующего каскада) переключение транзистора из одного состояния в другое не может протекать мгновенно.

В закрытом состоянии режим работы транзистора характеризуется точкой А (см. рис. 8.11, в). В момент времени t_0 напряжение генератора скачком изменяется от значения $U_{з-и0}$ до значения $U_{з-и1}$ (рис. 8.11, б). В течение интервала времени $t_3 = t_1 - t_0$, называемого *временем задержки включения*, происходит заряд паразитной емкости $C_{вх.экр.}$:

$$\tau_3 = \frac{C_{вх.экр.} u_{пор}}{i_{вх}} = C_{вх.экр.} R_{Г} \frac{u_{пор}}{u_{з-и1}}$$

Если внутреннее сопротивление $R_{Г}$ небольшое, то время задержки τ_3 пренебрежимо мало. Время задержки уменьшается также при снижении порогового напряжения.

В момент t_1 напряжение на затворе достигает порогового значения, транзистор открывается, появляется ток стока (рис. 8.11, з) и начинается разряд емкости C_H через открывшийся транзистор. В интервале времени $t_{и} = t_2 - t_1$ происходит фор-

мирование проводящего состояния канала, в ходе которого ток стока быстро нарастает с постоянной времени $\tau_s = C_{з.к} R_x$ и в момент времени t_2 ток достигает значения I_p , при этом рабочая точка переходит из положения А в положение A_1 .

С момента времени t_2 емкость C_H разряжается практически постоянным током I_p . При этом для тока разряда конденсатора справедливо соотношение:

$$I_c = C_H \frac{du_{c-u}}{dt}.$$

По мере разряда конденсатора C_H напряжение на нем уменьшается и в интервале времени $\tau_p = t_3 - t_2$ транзистор находится в режиме насыщения, рабочая точка переходит из положения A_1 в положение A_2 . Интервал времени можно определить

$$\tau_p = \frac{C_H (E_{u.n} - u_{нас})}{I_c}.$$

В момент t_3 транзистор переходит в линейный режим, и в интервале времени $\tau_k = t_4 - t_3$ разряд емкости происходит током стока, который равен току канала - I_c . В момент времени t_4 процесс включения завершается и на выходе схемы устанавливается напряжение u_{c-u0} , определяемое точкой В (рис.8.11,в).

Длительность временного интервала τ_k можно ориентировочно определить с помощью соотношения для элементарной RC цепочки $\tau_k \approx 2,3 R_{k0} C_H$. Здесь R_{k0} - сопротивление канала, через который происходит разряд конденсатора.

Для ориентировочных расчетов время включения можно оценить, поделив накопленный емкостью заряд C_H на постоянный ток разряда I_p .

$$\tau_{вкл} \approx \frac{E_{u.n} C_H}{I_p}$$

По окончании импульса транзистор закрывается, и емкость C_H заряжается от источника питания через резистор R_H . При

этом время выключения $\tau_{\text{выкл}}$ значительно отличается от времени $\tau_{\text{вкл}}$. Этот недостаток удачно разрешается использованием коллекторной пары, в которой заряд и разряд C_H идет через открытые разноканальные транзисторы, один из которых заменяет R_H в схеме (рис. 8.11)

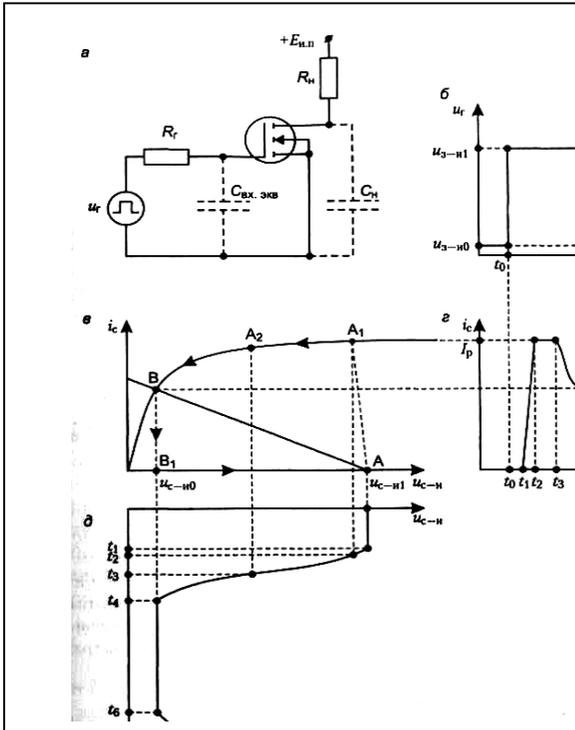


Рисунок 8.11. Переходные процессы в транзисторах

9. Сенсоры, датчики, преобразователи

9.1. Датчики температуры

9.1.1. Терморезисторы

Простейшим датчиком температуры является полупроводниковый резистор, изготовленный в виде стержня с двумя омическими контактами на концах. Зависимость концентрации носителей заряда от температуры изображена на рис.9.1.

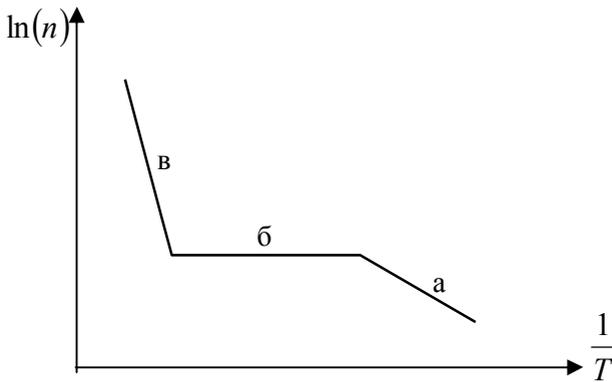


Рисунок 9.1 – Зависимость концентрации носителей заряда в полупроводнике n-типа от температуры

а – ионизация примесей; б – истощение примесей; в - изменение собственной проводимости

При температурах вблизи комнатной концентрация носителей заряда меняется в основном за счет собственных носителей. Для увеличения относительного изменения концентрации с изменением температуры необходимо использовать полупроводники с проводимостью, близкой к собственной. Электропроводность собственного полупроводника $\delta_i = qn_i(\mu_n + \mu_p)$. Так как $n_i \sim \exp(-E_g/2kT)$, то, пренебрегая зависимостью $\mu(T)$, сопротивление полупроводника: $R = R_\infty \exp(B/T)$, где $B = -E_g/2k$, а R_∞ - постоянная (формально, сопротивление при $T = \infty$). Формула для практического применения:

$R = R_0 \exp[B(T_0 - T)/(T_0 T)]$, где R_0 - номинальное сопротивление при $T_0 = 293 \text{ K}$. Кроме величин B и R_0 терморезистор характеризуется параметрами:

температурный коэффициент сопротивления α_T [%/C]

$$\alpha_T = (1/R)(\partial R / \partial T) = -(B/T^2) \cdot 100 \text{ [%/C]};$$

максимально допустимая мощность рассеяния P_{MAX} и максимально допустимая температура T_{MAX} , при которых не происходят необратимые изменения характеристик терморезистора;

коэффициент рассеяния H , численно равный мощности, отводимой от терморезистора в окружающую среду при разности температур поверхности терморезистора и окружающей среды 1 град;

постоянная времени τ , равная времени, за которое превышение температуры рабочего тела терморезистора над окружающей средой при охлаждении его в спокойном воздухе уменьшается в 2,7 раз.

Из вольтамперной характеристики терморезистора (рис. 9.2) видно, что при малых напряжениях электрическая энергия, выделяемая в терморезисторе, мала, его температура в зависимости от протекающего тока не изменяется и ВАХ линейна (выполняется закон Ома).

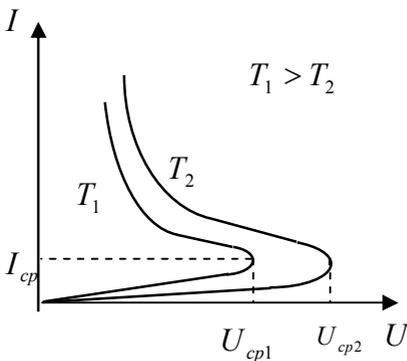


Рисунок 9.2. Вольтамперная характеристика терморезистора

При увеличении напряжения выделяемая энергия приводит к повышению температуры и уменьшению сопротивления. Это вызывает увеличение тока, что обуславливает повышение температуры терморезистора, уменьшение его сопротивления и дальнейший рост тока.

Такая положительная обратная связь по току приводит к его лавинному нарастанию и появлению участка с отрицательным сопротивлением.

ВАХ терморезистора имеет S-образный вид. Уравнения статической ВАХ резистора в параметрической форме:

$$U^2 = HR_0(T - T_1) \exp\left(B \frac{T_0 - T}{T_0 T}\right),$$

$$I^2 = \frac{H}{R_0} (T - T_1) \exp\left(B \frac{T_0 - T}{T_0 T}\right).$$

Максимальное напряжение U_{cp} и ток I_{cp} можно определить, подставив $T_{cp} = 0,5B(1 \pm \sqrt{1 - 4T_0/B})$, где T_{cp} - температура в точке (U_{cp}, I_{cp}) .

В зависимости от назначения схемы терморезисторы работают в двух основных режимах. При малых напряжениях (меньше U_{cp}) ток, протекающий через терморезистор, не вызывает его заметного разогрева и определяется по закону Ома: $I = U/R$, где R зависит от температуры окружающей среды. Этот режим используется в устройствах датчиков температуры и схемах температурной компенсации. Как прибор с ВАХ S-типа, терморезистор может использоваться в схеме переключателя (рис.9.3).

Сопротивление нагрузки R_H и напряжение источника питания E выбираются таким образом, чтобы нагрузочная линия пересекала ВАХ терморезистора в трех точках. После включения источника питания устанавливается рабочая точка и в цепи протекает небольшой ток. Напряжение на терморезисторе уменьшается при повышении температуры окружающей среды или при уменьшении коэффициента рассеяния H .

Коэффициент H уменьшается, например, при помещении терморезистора в среду с меньшей теплопроводностью, при снижении скорости потока окружающего воздуха или другой среды.

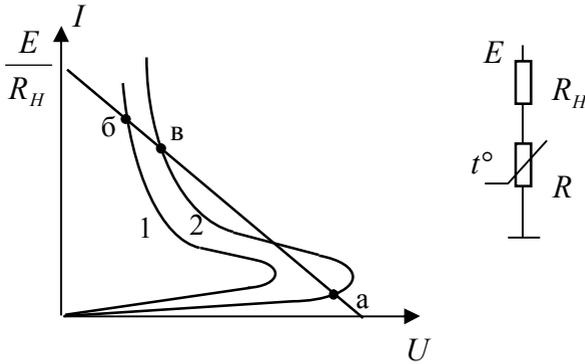


Рисунок 9.3 – Схема переключателя на основе терморезистора

При воздействии любого из указанных факторов, например, повышении температуры, ВАХ терморезистора меняется от кривой 2 к кривой 1. Рабочая точка перемещается из а в б, и ток в цепи резко увеличивается. При возвращении внешней температуры к первоначальному значению рабочая точка перемещается из б в в и высокое значение тока в цепи сохраняется. Часто в такой схеме вместо резистора нагрузки включается обмотка электромагнитного реле, которое срабатывает при увеличении тока в цепи. На этой основе могут быть построены реле, реагирующие на изменение температуры, влажности, скорости потока, уровня жидкости и т.д.

Терморезисторы изготавливаются из поликристаллических полупроводников с большим температурным коэффициентом сопротивления – окислов металлов переходного ряда периодической системы элементов или смеси окислов этих металлов (никель и марганец; никель, марганец и кобальт; титан и магний и др.). Для измерения низких температур (от азотных до гелиевых) используются терморезисторы из графита. Широко применяются терморезисторы из легированного германия и кремния. Введение в эти полупроводники примесей позволяет подобрать

оптимальный температурный коэффициент для заданного интервала рабочих температур. Монокристаллические терморезисторы более стабильны и надежны, их параметры более воспроизводимы – они используются при создании аппаратуры с более жесткими требованиями по надежности.

9.1.2. Диоды, как датчики температуры

Ток насыщения полупроводникового диода

$$I_{НАС} = A \exp(-E_g / kT),$$

где A - параметр, слабо зависит от температуры,
 E_g - ширина запрещенной зоны.

Статическое сопротивление диода можно записать в том же виде, что и для терморезистора

$$R = U / I_{НАС} = (U / A) \exp(E_g / kT) = R_{\infty} \exp(B / T),$$

где $I_{НАС}$ - ток насыщения диода.

Но в данном случае $B = E_g / k$ в два раза больше значения B для терморезистора. Используя зависимость обратного тока и сопротивления диодов от температуры, их можно использовать как датчики температуры. Преимуществом диодных датчиков является высокая чувствительность и малый потребляемый ток.

Из-за наличия токов по поверхности изготавливаемые промышленностью диоды имеют значительный разброс в значениях $I_{НАС}$. На практике для расчета $I_{НАС}(T)$ пользуются формулой:

$$I_{НАС} = I_{НАС0} \exp\left(B \frac{T - T_0}{TT_0}\right), \text{ где } T = 293K,$$

$I_{НАС0}$ - экспериментально измеряемый обратный ток диода при этой температуре.

Германиевые диоды используются для измерения температур до $90 \div 100^\circ C$, кремниевые – до $150 \div 170^\circ C$, арсенид-галлиевые – более высоких.

При выборе в качестве термочувствительного параметра обратного тока диода термочувствительность максимальна, но характеристика нелинейна. Для получения линейной зависимости

напряжения с датчика от температуры используется зависимость прямого напряжения на р-п-переходе от температуры. Лучше подходит р-п-переход с тонкой базой, так как его ВАХ более стабильна (меньше зависит от времени жизни неосновных носителей и от состояния поверхности). Для р-п перехода при $I \gg I_{НАС}$

$$U = \frac{kT}{q} \ln \left(\frac{I}{I_{НАС}} + 1 \right) \approx \frac{kT}{q} \ln \frac{IW_n}{qD_p p_n}.$$

Так как

$$p_n = n_i^2 / n_n \sim (T^3 / n_n) \exp(-E_g / kT), D_p = \mu_p kT / q, \mu_p \sim T^{-\alpha},$$

$$\text{то } U = \frac{kT}{q} \ln \frac{IW_n n_n}{CT^{4-\alpha}} + \frac{E_g}{q},$$

где в C входят все не зависящие от температуры постоянные. р-п переход работает в режиме $I = const$. Для уменьшения энергопотребления используется режим $I < CT^{4-\alpha} / (W_n n_n)$. Логарифм отрицателен, поэтому с ростом температуры напряжение на кремниевом р-п переходе уменьшается на величину порядка 2 мВ/град. Такой датчик работоспособен в интервале температур 77...400 К. Незначительная нелинейность зависимости $U(T)$ существует из-за множителя $T^{4-\alpha}$ в логарифме. Для ее уменьшения используют полупроводник с большей зависимостью подвижности от температуры (больше α). Зависимость $\mu(T)$ усиливается с уменьшением степени легирования полупроводника. В кремнии с концентрацией основных носителей $\sim 10^{10} \text{ см}^{-3}$ $\alpha = 2,5$ для μ_n и $\alpha = 2,7$ для μ_p . В Ge и GaAs α меньше.

При больших токах через р-п переход логарифм становится положительным и U растет с увеличением температуры. В реальных диодах это возрастание начинается при токах, меньших, чем это следует из аналитической зависимости напряжения от температуры, т.к. при больших токах растет часть общего напряжения, падающая на сопротивлении базы, а оно увеличивается с увеличением температуры в результате уменьшения под-

вижности. Такая зависимость характерна для всех переходов независимо от их структуры. Отсюда же следует, что существует точка на ВАХ, где I и U не зависят от температуры. На практике в качестве термочувствительного диода часто используется эмиттерный переход биполярного транзистора с коллектором, замкнутым на базу. Зависимость $U(T)$ аналогична.

9.1.3. Транзисторы, как датчики температуры

Обратный ток коллекторного p-n перехода транзистора $I_{КБО}$ так же зависит от температуры, как ток насыщения p-n перехода. При работе транзистора в качестве датчика температуры обычно используют схему включения с ОЭ и отключенной базой (рис.9.4).

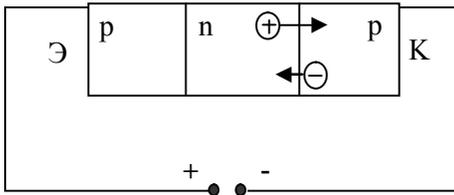


Рисунок 9.4. Включение транзистора в схеме с ОЭ

при отключенной базе

Ток через транзистор $I \approx h_{21Э} I_{КБО}$, зависимость сопротивления транзистора от температуры та же, что для диода, но R_{∞} в $h_{21Э}$ раз меньше.

Экспериментально измеряемые значения α_T у транзисторов выше, чем у диодов. Причина этого заключается в зависимости коэффициента передачи тока в схеме с ОБ и в схеме с ОЭ от тока эмиттера. Значение ТК сопротивления для транзистора в схеме с ОЭ и отключенной базой:

$$\alpha_{TT} = \frac{\alpha_{ТД}}{1 - (\partial h_{21Э} / \partial I) I_{КБО}},$$

где α_{TD} - коэффициент для р-п перехода. Для увеличения термочувствительности транзистора нужно усилить зависимость $h_{21Э}(I)$. Если $h_{21Э} = const$, термочувствительность диода α_{TD} и транзистора α_{TT} одинакова.

9.2. Полупроводниковые термоэлектрические преобразователи

Полупроводниковое термоэлектрическое устройство - это устройство, действие которого основано на использовании термоэлектрических эффектов Пельтье или Зеебека в полупроводниках, предназначенное для получения теплоты или холода с использованием электрической энергии или получения электрической энергии с использованием тепловой.

Конструкция и принцип действия

В полупроводниковых термоэлектрических устройствах используются полупроводниковые термоэлементы, состоящие из двух ветвей, обладающих разным типом электропроводности. Ветвь термоэлемента, материал которого имеет электропроводность р-типа, называют положительной ветвью, а ветвь с электропроводностью n-типа – отрицательной ветвью. Ветви полупроводникового термоэлемента соединяют последовательно контактной пластиной (рис. 9.5). Зону их электрического соединения называют спаем. При работе термоэлемента спаи имеют разную температуру: один является теплопоглощающим, а другой - тепловыделяющим.

Полупроводниковое термоэлектрическое устройство обычно имеет большое число последовательно соединенных между собой термоэлементов в едином конструктивном исполнении – полупроводниковую термобатарю. Полупроводниковое термоэлектрическое устройство, включающее одну или несколько электрически соединенных полупроводниковых термобатарей в едином конструктивном исполнении с системой теплообмена, называют полупроводниковым термоэлектрическим блоком.

Возникновение термо-ЭДС (эффект Зеебека)

При наличии разности температур спаев в цепи с термоэлементом появляются три составляющих термоэлектродвижущей силы (термо-ЭДС):

Первая составляющая обусловлена диффузией носителей заряда от нагретого спая, температура которого из-за подводимой к нему тепловой мощности от источника выше температуры тепловыделяющего спая. Диффузия носителей заряда возникает из-за градиента концентрации (у нагретого спая большее число ионизированных примесей, при дополнительной ионизации примесей увеличивается концентрация основных носителей заряда на нагретых концах ветвей термоэлемента) и вследствие выравнивания средней энергии, приходящейся на один носитель определенного знака (на нагретых концах ветвей носители заряда приобретают большие энергии).

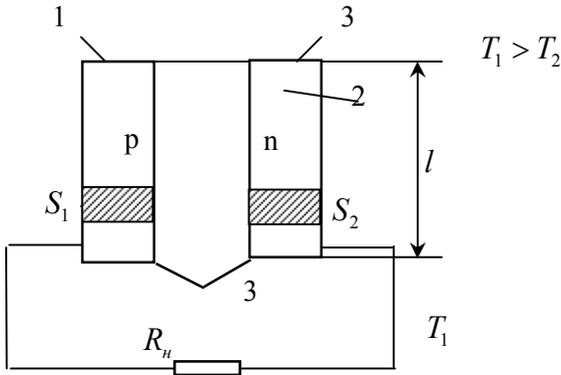


Рисунок 9.5 – Схема полупроводникового термоэлемента с сопротивлением нагрузки: 1- положительная ветвь; 2- отрицательная ветвь; 3- металлические контактные пластины.

Диффузия электронов в отрицательной ветви происходит только от нагретого конца этой ветви вдоль нее. Переходу электронов в положительную ветвь препятствует потенциальный барьер нагретого спая. Аналогично, дырки в положительной ветви могут диффундировать вдоль нее от нагретого конца.

Перемещение носителей заряда, связанное с их диффузией, нарушает электрическую нейтральность в ветвях термоэлемента

– на нагретых концах ветвей остаются нескомпенсированные ионизированные примесные атомы, а на противоположных концах образуется избыток основных носителей заряда.

Вторая составляющая – следствие температурной зависимости контактной разности потенциалов. Если оба спая термоэлемента имеют одну и ту же температуру, то контактные разности потенциалов на этих спаях равны, направлены в противоположные стороны при обходе контура с термоэлементом и не дают результирующей термо-ЭДС. Если же температура спаев термоэлемента различна, то значение контактной разности потенциалов на спаях будет различно.

Третья составляющая возникает вследствие увлечения носителей заряда квантами тепловой энергии – фононами. Если в ветвях термоэлемента есть градиент температуры, то будет существовать направленное движение фононов от нагретых концов ветвей. В результате столкновений с носителями заряда фононы увлекают за собой электроны в отрицательной ветви и дырки в положительной. Эффект может быть преобладающим при низких температурах.

Результирующая термо-ЭДС зависит от разности температур спаев термоэлемента и от электрофизических свойств полупроводниковых материалов, образующих ветви термоэлемента. В небольшом диапазоне температур термо-ЭДС ε можно считать пропорциональной разности температур спаев термоэлемента и коэффициенту термо-ЭДС $\alpha_{1,2}$: $\varepsilon = \alpha_{1,2} \Delta T$

Поглощение и выделение теплоты в спаях термоэлемента (эффект Пельтье)

При прохождении постоянного тока через термоэлемент в спаях поглощается или выделяется некоторое количество теплоты: $Q_{II} = \pm P_{1,2} I t$,

где $P_{1,2}$ - коэффициент Пельтье.

Если направление тока в цепи с термоэлементами, как показано на рис. 9.6, то на спае 1 свободные электроны и дырки, возникшие в результате тепловой генерации, движутся в разные стороны под действием суммарного электрического поля. При тепловой генерации носителей заряда в области спая 1 на пере-

брос электрона из валентной зоны в зону проводимости затрачивается тепловая энергия кристаллической решетки полупроводника. При заданном направлении тока спай 1 будет охлаждаться; спай 2 нагревается, так как к нему с разных сторон подходят электроны и дырки, которые отдают при рекомбинации некоторую энергию в виде теплоты.

Эффект Пельтье является обратным эффекту Зеебека. Есть связь между коэффициентом Пельтье и коэффициентом термо-ЭДС: $\Pi_{1,2} = \alpha_{1,2} T$

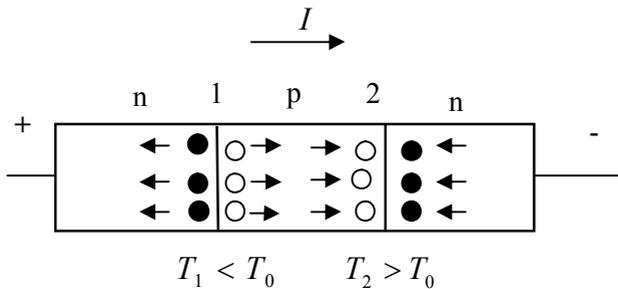


Рисунок 9.6. Охлаждение спая 1 и нагрев спая 2 полупроводниковых термоэлементов при прохождении тока

Термоэлектрические генераторы

Термоэлектрический генератор (ТЭГ) – это полупроводниковое термоэлектрическое устройство, состоящее из полупроводникового термоэлектрического блока и предназначенное для непосредственного преобразования тепловой энергии в электрическую энергию.

Коэффициент полезного действия термоэлемента.

Одним из основных параметров термоэлектрического генератора является коэффициент полезного действия, т.е. отношение полезной мощности, выделяемой в нагрузку генератора, к количеству теплоты, поступающей в единицу времени на теплопоглощающие спаи.

Рассмотрим работу отдельного термоэлемента в качестве преобразователя тепловой энергии в электрическую. Если удельное сопротивление ветвей термоэлемента ρ_1 и ρ_2 , токо-

вая высота ветви – длина пути тока в ветви термоэлемента l , токовое сечение ветвей – сечение ветвей термоэлемента, определяемое по нормали к векторным линиям тока S_1 и S_2 (рис.9.5), то полное сопротивление термоэлемента:

$$R = \rho_1 \frac{l}{S_1} + \rho_2 \frac{l}{S_2}. \text{ Полную теплопроводность термоэлемента}$$

можно выразить через удельные теплопроводности ветвей χ_1 и

$$\chi_2: K = \chi_1 \frac{S_1}{l} + \chi_2 \frac{S_2}{l}. \text{ В результате процесса теплопроводности}$$

от теплопоглощающего спая или от источника тепловой энергии к тепловыделяющему спаяу в единицу времени передается теплота: $Q_{II} = K(T_2 - T_1)$. Ток, проходящий в цепи термоэлемента:

$$I = \frac{\alpha_{1,2}(T_2 - T_1)}{R + R_H}. \text{ В нагрузке выделяется полезная}$$

$$\text{мощность: } P_H = I^2 R_H = \frac{\alpha_{1,2}^2 (T_2 - T_1)^2 R_H}{(R + R_H)^2}$$

Ток, проходящий в цепи термоэлемента и обусловленный эффектом Зеебека, вызывает выделение и поглощение теплоты из-за эффекта Пельтье на спаях и выделение джоулевой теплоты в ветвях термоэлемента. Направление тока в термоэлементе соответствует поглощению в единицу времени на теплопоглощающем спаяе теплоты: $Q_{II2} = \Pi_{1,2} I = \alpha_{1,2} I T_2$

Эффект Пельтье, заключающийся в перекачивании некоторого количества теплоты от теплопоглощающего спая к тепловыделяющему, увеличивает теплопроводность термоэлемента. Теплота, отдаваемая тепловыделяющему спаяу в результате этого эффекта: $Q_{II1} = \alpha_{1,2} I T_1$.

Из-за разных значений T_2 и T_1 значения Q_{II2} и Q_{II1} также различны.

$$Q_{II2} - Q_{II1} = \alpha_{1,2} I (T_2 - T_1) = \mathcal{E} I = I^2 (R + R_H)$$

Т.о. вся электрическая мощность, вырабатываемая термоэлементом, - это разница теплоты, отданной источником теплоты, поглощающему спаю, и теплоты, перенесенной на тепловыделяющий спай из-за эффекта Пельтье. Часть электрической мощности $I^2 R_H$ выделяется в виде полезной мощности на нагрузке, а $I^2 R$ - в виде джоулевой теплоты в ветвях термоэлемента. Приблизительно, джоулева теплота распределяется пополам между спаями, т.е. к теплопоглощающему спаю или источнику тепловой энергии возвращается мощность:

$$\frac{1}{2} Q_{Дж} = \frac{1}{2} I^2 R = \frac{\alpha_{1,2}^2 (T_2 - T_1)^2 R}{2(R + R_H)^2}.$$

Т.о. коэффициент полезного действия термоэлемента:

$$\eta = \frac{P_H}{Q_K + Q_{П2} - \frac{1}{2} Q_{Дж}} = \frac{\frac{\alpha_{1,2}^2 (T_2 - T_1)^2 R_H}{(R + R_H)^2}}{K(T_2 - T_1) + \frac{\alpha_{1,2}^2 (T_2 - T_1)}{R + R_H} - \frac{\alpha_{1,2}^2 (T_2 - T_1)^2 R}{2(R + R_H)^2}}$$

Обозначим $R_H / R = m$, тогда

$$\begin{aligned} R_H &= mR = ml(\rho_1 / S_1 + \rho_2 / S_2) && \text{и} \\ R + R_H &= (m + 1)R = (m + 1)l(\rho_1 / S_1 + \rho_2 / S_2) \\ \eta &= \frac{T_2 - T_1}{T_2} = \frac{m / (m + 1)}{1 + \frac{KR(m + 1)}{\alpha_{1,2}^2 T_2} - \frac{T_2 - T_1}{2(m + 1)T_2}} \end{aligned}$$

Первый множитель представляет собой КПД обратимой тепловой машины, второй характеризует уменьшение КПД вследствие необратимых потерь на теплопроводность и джоулеву теплоту в термоэлементе.

Для обеспечения наибольшего КПД (при данных T_2 , T_1 , $\alpha_{1,2}$, m , ρ и χ ветвей) необходимо выбрать токовые сечения ветвей так, чтобы произведение KR было минимальным. Функция KR имеет минимальное значение при оптимальном отношении токовых сечений ветвей:

$(S_1 / S_2)_{opt} = \sqrt{\rho_1 \chi_2 / (\rho_2 \chi_1)}$. При этом

$$(KR)_{MIN} = \left(\sqrt{\chi_1 \rho_1} + \sqrt{\chi_2 \rho_2} \right)^2.$$

Величину $\alpha_{1,2}^2 / KR$ обычно обозначают Z и называют эффективностью или добротностью термоэлемента:

$$Z = \frac{\alpha_{1,2}^2}{\left(\sqrt{\chi_1 \rho_1} + \sqrt{\chi_2 \rho_2} \right)^2}.$$

Т.о. КПД термоэлемента оптимальных размеров зависит от следующих факторов:

эффективности термоэлемента Z , определяемой электрофизическими параметрами полупроводниковых материалов ветвей термоэлемента;

разности температур спаев термоэлемента;

отношения сопротивления нагрузки к сопротивлению термоэлемента.

Для нахождения максимального КПД необходимо выбрать оптимальное отношение сопротивления нагрузки к сопротивлению термоэлемента: $m_{opt} = (R_H / R)_{opt} = \sqrt{1 + ZT}$, где $T = (T_2 - T_1) / 2$ - средняя температура спаев.

Максимальное значение КПД термоэлемента, зависящее лишь от температуры спаев и эффективности термоэлемента:

$$\eta_{MAX} = \frac{T_2 - T_1}{T_2} = \frac{m_{opt} - 1}{m_{opt} + T_1 / T_2}$$

При определенных температурах T_2 и T_1 и при $Z \rightarrow \infty$ максимальный КПД стремится к КПД идеальной тепловой машины (цикл Карно): $\eta_{MAX} \rightarrow (T_2 - T_1) / T_2$.

Часто от термоэлемента необходимо получить максимальную мощность даже в ущерб КПД. При таком режиме работы сопротивление нагрузки должно быть равно сопротивлению

термоэлемента, т.е. $R_H / R = m = 1$. При этом

$$\eta|_{m=1} = \frac{1}{2} \frac{T_2 - T_1}{T_2 + \frac{2}{Z} - \frac{1}{4}(T_2 - T_1)}$$

В большинстве случаев максимальный КПД и КПД в режиме максимальной отдаваемой мощности отличаются друг от друга на $2 \div 3\%$.

Конструкция, свойства и параметры полупроводниковых термопреобразователей

Коэффициент термо-ЭДС большинства термоэлементов ($\alpha_{1,2} = \alpha_1 - \alpha_2$) составляет $300 \div 500$ мкВ/К. Учитывая, что почти все термоэлектрические генераторы работают при разности температур спаев, не превышающей 600 К, от одного термоэлемента можно получить термо-ЭДС 0,3 В. При этом оптимальные условия получения электрической мощности на нагрузке достигаются при сопротивлении нагрузки, сравнимом с сопротивлением термоэлемента, т.е. около нескольких миллиОм. Для получения приемлемых напряжений и для согласования с реальными нагрузочными сопротивлениями необходимо последовательное соединение многих термоэлементов.

В большинстве термоэлектрических генераторов (ТЭГ) полупроводниковая термобатарея представляет собой плоскую конструкцию, состоящую из нескольких десятков термоэлементов. По форме термоэлементы могут быть разнообразными в зависимости от конструктивных особенностей источника тепловой энергии. Наиболее распространены термоэлементы из ветвей в виде параллелепипедов с различным соотношением токового сечения и токовой высоты.

Соединение отдельных ветвей между собой металлическими контактными пластинами должно быть сделано так, чтобы контактные сопротивления были намного меньше сопротивления ветвей термоэлемента. Иначе контактные сопротивления могут уменьшить эффективность термоэлемента из-за увеличения общего сопротивления термоэлемента при неизменном значении термо-ЭДС.

В качестве источников тепловой энергии в ТЭГ применяют атомные реакторы, источники теплоты с радиоактивными изотопами, используют химические реакции горения и тепловую энергию, излучаемую Солнцем.

Существуют три основные области применения ТЭГ:

для катодной защиты магистральных газо- и нефтепроводов от коррозии;

для питания электрической энергией космических аппаратов;

для питания электрической энергией навигационных и метеорологических установок.

Холодильники и тепловые насосы

Максимальное снижение температуры теплопоглощающих спаев термоэлемента

При прохождении тока через термоэлементы наряду с поглощением на спае 1 (рис. 9.6) в единицу времени некоторого количества теплоты $Q_{II} = -\Pi_{1,2}I$ происходит еще выделение джоулевой теплоты, пропорциональной квадрату силы тока и сопротивлению спаев 1. В первом приближении можно считать, что сопротивления спаев 1 и 2 равны, т.е. каждое из них составляет половину полного сопротивления термоэлемента R .

Тогда Джоулева теплота, выделяющаяся на первом спае:

$$Q_{Дж} = \frac{1}{2}I^2R.$$

Для теплопоглощающего спаев:

$$Q = Q_{II} + Q_{Дж} = -\Pi_{1,2}I + \frac{1}{2}I^2R$$

При токе, равном нулю, оба эффекта отсутствуют и охлаждения не происходит, а при достаточно большом токе $Q_{Дж}$ превысит Q_{II} и охлаждение первых спаев перейдет в нагрев. Поэтому существует оптимальный ток, при котором эффект охлаждения максимален (от теплопоглощающего спаев отводится

максимальное количество теплоты): $I_{omn} = P_{1,2} / R$. При этом токе $Q_{I_{omn}} = -P_{1,2} / (2R)$.

Чем меньше сопротивление термоэлемента, тем больше количество теплоты $Q_{I_{omn}}$, отводимого от теплопоглощающего спая. Однако, увеличивая токовые сечения или уменьшая токовую длину ветвей термоэлемента, нельзя достичь более низких температур на теплопоглощающем спае, так как одновременно в такой же мере возрастают теплопроводность и поток теплоты по ветвям термоэлемента от тепловыделяющего спая к теплопоглощающему. Температура теплопоглощающего спая будет понижаться до тех пор, пока количество теплоты Q_K , переносимое между спаями в результате теплопроводности по ветвям термоэлемента, не сбалансирует отводимое от спая 1 количество теплоты Q_I при токе I .

Условие теплового баланса: $-Q_1 = Q_K = K(T_2 - T_1)$ или $T_2 - T_1 = -Q_I / K$.

Т.о. требования, которым должен удовлетворять полупроводник ветвей термоэлемента, работающего в холодильнике, сводятся к достижению возможно большего значения эффективности Z .

Применение полупроводниковых холодильников

Полупроводниковые холодильники применяют в радиоэлектронике, медицине, сельском хозяйстве, метрологии и для бытовых целей (например, автомобильные холодильники).

Но экономичность полупроводниковых холодильников при охлаждении объемов более десятков литров ниже экономичности компрессионных машин. Поэтому полупроводниковые холодильники могут конкурировать с компрессионными только при охлаждении малых объемов.

Использование принципа полупроводникового теплового насоса

В полупроводниковой термобатарее при прохождении через нее постоянного тока происходит охлаждение одних спаев и нагрев других из-за поглощения и выделения теплоты в связи с эффектом Пельтье. Это явление используется в термостатах:

изменение направления тока позволяет либо охлаждать, либо нагревать спаи термоэлементов, находящиеся внутри термоста-та. Разогрев тепловыделяющих спаев в полупроводниковой тер-мобатарее происходит из-за выделения обычной джоулевой те-плоты и в результате переноса теплоты Пельтье с теплопогло-щающих спаев на тепловыделяющие. На тепловыделяющих спаях может выделяться большее количество теплоты, чем было затрачено электрической энергии. Принцип теплового насоса может оказаться экономически целесообразным для обогрева и охлаждения жилых и производственных помещений при суще-ственном снижении стоимости термоэлементов, когда выигрыш в расходе электрической энергии компенсирует стоимость по-лупроводникового термоэлектрического подогревателя.

9.3. Датчики магнитного поля. Преобразователи.

На заряд q , движущийся в магнитном поле со скоростью V , действует сила Лоренца $F=q[VB]$, где B - индукция магнитного поля. Если $V \perp B$, то $F=qVB$.

Рассмотрим пластину полупроводника p -типа, через кото-рую протекает ток, направленный перпендикулярно внешнему магнитному полю (рисунок 9.7).

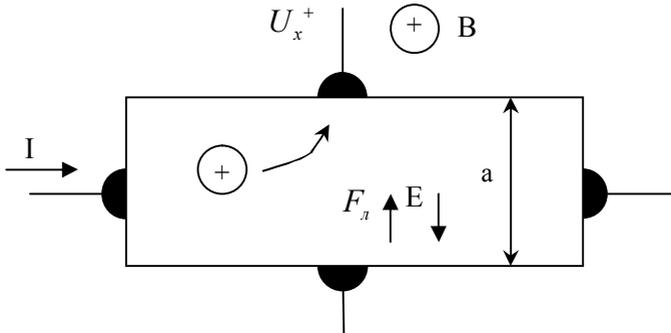


Рисунок 9.7. Возникновение ЭДС Холла

Сила Лоренца отклоняет дырки к верхней грани полупро-водника, вследствие чего их концентрация там увеличивается, а у нижней границы уменьшается. В результате пространственно-го разделения зарядов возникает электрическое поле, направ-

ленное от верхней границы к нижней. Это поле препятствует разделению зарядов, и, как только создаваемая им сила станет равной силе Лоренца, $qE=qVB$, дальнейшее разделение зарядов прекратится. При таком равновесии поток дырок движется через пластину не отклоняясь. Разность потенциалов между верхней и нижней гранью образца равна:

$$U_x = Ed = VBd .$$

Явление возникновения поперечной ЭДС в полупроводнике, помещённом в магнитное поле, называется эффектом Холла, а U_x - соответственно напряжение Холла.

Часто датчик Холла используется как элемент цепи с токовыми контактами в качестве входных и контактами, с которых снимается U_x , в качестве выходных. Примером такого режима работы является преобразователь постоянного тока в переменный на основе датчика Холла. В этом случае датчик помещается в переменное магнитное поле и при пропускании через входную цепь постоянного тока на входе возникает переменное напряжение U_x , изменяющееся с частотой магнитного поля.

Работу датчика Холла в таких схемах можно характеризовать коэффициентом использования датчика $\eta = P_n / P_{ex}$, который является отношением мощности, выделяемой во внешней нагрузке P_n , к мощности, потребляемой во входной цепи P_{ex} .

9.3.1. Магниторезисторы

В полупроводнике, помещённом в магнитное поле, при прохождении электрического тока на носители действует сила Лоренца, искривляющая их линии тока. Если в полупроводнике один тип носителей, то в нём возникает поперечное электрическое поле - поле Холла. Напряжённость этого поля такова, чтобы его сила точно скомпенсировала действие на носители силы Лоренца: линии тока должны быть параллельны граням образца.

Однако скорости движения носителей неодинаковы из-за теплового размытия, поэтому поле Холла компенсирует действие силы Лоренца только для носителей, движущихся со средней скоростью. На более медленные носители сильнее действует

поле Холла, а на более быстрые - сила Лоренца. Поэтому вклад в проводимость носителей, движущихся со скоростью, отличной от средней, в магнитном поле оказывается меньше. Уменьшение проводимости полупроводника в магнитном поле называется эффектом Гаусса или магниторезистивным (эффектом магнитосопротивления). В полупроводнике с двумя видами носителей поле Холла меньше, поэтому линии тока в нём не параллельны граням и эффект магнитосопротивления соответственно больше.

Исключить эффект Холла можно и в полупроводнике с одним видом носителей, например, используя образцы специальной формы. Если сделать образец в виде пластинки с концентрическим расположением контактов: один в центре, а другой по окружности на равных расстояниях от первого, то ток будет протекать по радиусам. При помещении в магнитное поле, пронизывающее перпендикулярно пластину, носители будут отклоняться вдоль поверхности. Линии тока удлиняются, но накопление заряда происходить не будет, и поле Холла не возникает. В такой структуре, называемой диском Корбино, наблюдается максимальный эффект магнитосопротивления.

Исключить эффект Холла можно также, если на поверхность полупроводниковой пластинки нанести узкие металлические полоски для закорачивания Холловской разности потенциалов. Они должны быть расположены перпендикулярно линиям тока и направлению индукции магнитного поля. Вместо нанесённых на поверхность полосок могут использоваться монокристаллы с пронизывающими их металлическими иглами. Такой вариант используется для создания магниторезисторов из $\text{InSb}+\text{NiSb}$. При выращивании монокристаллов из этого материала NiSb образует иглы высокой проводимости, которые при выборе соответствующей ориентации практически полностью исключают эффект Холла.

9.3.2. Магнитодиоды

В длинных диодах прямой ток определяется равновесной проводимостью базовой области. Распределение носителей в базовой области зависит от подвижности и эффективного времени жизни. В магнитном поле вследствие эффекта магнитосопротивления уменьшается подвижность носителей и, следова-

тельно, сильно уменьшается проводимость диода. Эффект магнитосопротивления за счёт изменения инжекции усиливается в десятки и сотни раз. Магниточувствительность длинных диодов, названных магнитодиодами, во много раз превышает магниточувствительность магниторезисторов так же, как фоточувствительность обычных фотодиодов и фоторезисторов.

Магнитное поле в магнитодиодах не только уменьшает подвижность, но и искривляет линии тока, так как в них концентрации электронов и дырок практически одинаковы и поле Холла отсутствует. Удлинение линий тока приводит к уменьшению глубины проникновения неравновесных носителей и дополнительно уменьшению модуляции проводимости базовой области инжектированными носителями, то есть к повышению магниточувствительности.

Искривление линий тока повышает концентрацию носителей у одной грани и понижает у другой. Поскольку эффективное время жизни носителей в тонких пластинках определяется поверхностной рекомбинацией, то перераспределение носителей приводит к изменению роли поверхностной рекомбинации и эффективного времени жизни. Роль рекомбинации на грани, к которой отклоняются носители, возрастает, рекомбинация на второй грани почти перестаёт играть роль.

Основной характеристикой магниторезисторов и магнитодиодов, а также некоторых других гальваномагнитных приборов является вольтовая магниточувствительность. Она равна отношению изменения напряжения на образце к произведению тока через него и изменения индукции магнитного поля, приведшего к данному изменению напряжения:

$$\gamma_n = \Delta U / (\Delta BI)$$

В германиевых и кремниевых магнитодиодах $\gamma_n = 30 \dots 90$ В/(А·Т).

В качестве магниточувствительных диодов могут использоваться и светодиоды, причём в зависимости от индукции магнитного поля может меняться как интенсивность, так и длина волны излучения. Например, при длине базы $W < L$ (L- длина диффузионного смещения инжектированных носителей заряда)

часть носителей рекомбинирует в базе с излучением, а остальные рекомбинируют на контакте к базе без излучения.

При помещении диода с длинной базой в магнитное поле, параллельное его продольной оси, инжектированные в базу электроны и дырки, начинают вращаться вокруг оси в перпендикулярной плоскости. При этом происходит образование сгустков плазмы, чередующихся с обеднёнными областями. Если на боковых гранях диода установить два симметричных зонда, то вращение сгустков и обеднение областей плазмы в перпендикулярной плоскости приведёт к появлению переменной ЭДС между зондами. Колебаний тока, протекающего через диод симметричной структуры, при этом не происходит, так как перераспределение плазмы в перпендикулярной плоскости не меняет суммарной концентрации и общего сопротивления. При нарушении симметрии распределения исходной концентрации носителей заряда относительно оси возникновение переменной ЭДС на поперечных зондах сопровождается возникновением колебаний тока, протекающего через диод.

9.3.3. Биполярные магнитотранзисторы

В обычном биполярном транзисторе, например сплавном р-п-р типа (рисунок 9.8), инжектированные из эмиттера носители движутся симметрично относительно оси транзистора. Основная их часть проходит путь, равный толщине базы W_0 . Магнитное поле приводит к отклонению траекторий движения инжектированных носителей от первоначального направления приблизительно на угол Холла ϕ .

Поскольку толщина базы много меньше размеров эмиттера и коллектора, то практически все носители, прошедшие базовую область, попадают в коллектор.

Однако средний путь, проходимый носителями в базовой области, несколько возрастает, что приведёт к увеличению их доли, рекомбинировавшей в базовой области. Коэффициент передачи тока уменьшится.

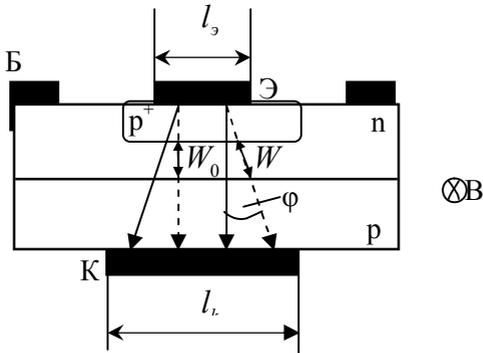


Рисунок 9.8. Траектории движения инжектированных носителей в магнитном поле в p-n-p-биполярном транзисторе с узкой базой

При оценке магниточувствительности таких магнитотранзисторов можно приближённо считать, что возросла эффективная толщина базы: $W \approx W_0 \sec \varphi \approx W_0 \times (1 + \mu^2 B^2 \sqrt{2})$.

Угол Холла мал, поэтому магниточувствительность “обычных” транзисторов мала. В настоящее время транзисторы изготавливают только из германия и кремния. Поскольку подвижность носителей в германии выше, чем в кремнии, то германиевые транзисторы обладают большей чувствительностью к магнитному полю. Возможно также создание магнитотранзисторов из антимонида индия, который имеет существенно большую подвижность носителей и, следовательно, может обеспечить большую магниточувствительность. Однако транзисторы из антимонида индия могут работать только при низких температурах.

Для увеличения магниточувствительности транзисторов необходимо, чтобы искривление траектории движения инжектированных носителей приводило не только к увеличению W , но и к отклонению части носителей от коллектора.

Наибольшее практическое распространение получили горизонтальные (латеральные) магнитотранзисторы, то есть транзи-

стор, коллекторы которых расположены в стороне от эмиттера на одной поверхности пластины. Их магниточувствительные свойства обусловлены двумя основными эффектами: изменением эффективной длины базы и отклонением инжектированных носителей от коллектора (либо к нему).

Список рекомендуемой литературы

1. Пасынков В.В., Чиркин Л.К. Полупроводниковые приборы.- СПб.: Лань, 2001.- 480 с.
2. Тугов Н.М., Глебов Б.А., Чарыков Н.А. Полупроводниковые приборы./ Под ред. – В.А.Лабунцова. – М.: Энергоатомиздат, 1990. -576 с.
3. Гаман В.И. Физика полупроводниковых приборов. Томск: изд-во НТЛ, 2000.- 426 с.
4. Кутякова М.Г., Чарыков Н.А., Юдин В.В. Полупроводниковые приборы и основы их проектирования. – М.: Радио и связь, 1983. – 352 с.
5. Викулин И.М., Стафеев В.И. Физика полупроводниковых приборов. – М.: Советское радио, 1980. – 296 с.

6. Зи С. Физика полупроводниковых приборов. – М.: Мир, 1984. – 453 с.
7. Легостаев Н.С. Электроника и микроэлектроника. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2003. – 185 с.
8. Гуртов В. Твердотельная электроника. – М.: Техносфера, 2005. – 408 с.
9. Батушев В.А. Электронные приборы. – М.: Высшая школа, 1980. – 383 с.
10. Степаненко И.П. Основы микроэлектроники. – М.: Лаборатория базовых знаний, 2004. – 488 с.
11. Воробьев Г.А., Троян П.Е. Электронные процессы в тонкопленочных структурах металл – диэлектрик – металл. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2006. – 167 с.
12. Легостаев Н.С., Троян П.Е., Четвергов К.В. Физические основы твердотельной электроники. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2006. – 130 с.