

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
УКРАИНЫ
ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ



Гоков А.М., Жидко Е.А.

ОСНОВЫ ЭЛЕКТРОТЕХНИКИ И ЭЛЕКТРОНИКИ. ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ ЭЛЕКТРОНИКИ



УЧЕБНОЕ ПОСОБИЕ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ХАРКОВСКИЙ НАЦИОНАЛЬНЫЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ

**Гоков А.М.
Жидко Е.А.**

**ОСНОВЫ ЭЛЕКТРОТЕХНИКИ И ЭЛЕКТРОНИКИ
ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ ЭЛЕКТРОНИКИ**

**Учебное пособие
Часть 2**

Харьков, Изд. ХНЭУ, 2006

ББК31.2я7

Г59

УДК621.3(075)

Рецензенты: докт. техн. наук, профессор кафедры экспериментальной физики Харьковского национального университета им. В. Н. Каразина *Пойда В. П.*; докт. физ.-мат. наук, профессор кафедры полупроводниковой и вакуумной электроники Харьковского национального университета им. В. Н. Каразина *Аркуша Ю. В.*; канд. физ.-мат. наук, доцент кафедры микроэлектроники, электронных приборов и устройств Харьковского национального университета радиоэлектроники *Панченко А. Ю.*

Рекомендовано к изданию решением ученого совета Харьковского национального экономического университета.

Протокол №5 от 20.02.2006 г.

Гоков А. М.

Г59 Основы электротехники и электроники. Элементы общей теории электроники. Учебное пособие. Ч. 2 / А. М. Гоков, Е. А. Жидко. — Харьков: Изд. ХНЭУ, 2006. — 276 с. (Русск. яз.)

ISBN 966-676-142-4

Приведены основные сведения из общей теории электроники, фундаментальные идеи, термины и определения. Излагаются принципы построения, модели, характеристики и параметры простейших структур полупроводниковой электроники — диодов — и более сложных — транзисторов. Изучаются основы микроэлектроники и оптоэлектроники, модели, принципы построения и функционирования реальных устройств. Рассматриваются полупроводниковые изделия силовой электроники, их практические реализации и применение. Приведен широкий круг характерных практических примеров.

Рекомендовано для студентов, обучающихся по профилю компьютеризованных технологий и систем издательско-полиграфических производств, а также для изучающих одноименную дисциплину по другим профилям подготовки.

Наведено основні відомості із загальної теорії електроніки, фундаментальні ідеї, терміни й визначення. Викладаються принципи побудови, моделі, характеристики й параметри найпростіших структур напівпровідникової електроніки — діодів — і більш складних — транзисторів. Вивчаються основи мікроелектроніки й оптоелектроніки, моделі, принципи побудови й функціонування реальних пристроїв. Розглядаються напівпровідникові вироби сигової електроніки, їхні практичні реалізації й застосування. Наведено широке коло характерних практичних прикладів.

Рекомендовано для студентів, що навчаються за профілем комп'ютеризованих технологій і систем видавничо-поліграфічних виробництв, а також для тих, хто вивчає однойменну дисципліну за іншими профілями підготовки.

ББК 31.2я7

ISBN 966-676-142-4

© Харьковский национальный
экономический университет, 2006

© Гоков А. М.
Жидко Е. А.
2006

Введение

Характер современного экономического развития многих стран наглядно показывает, что прогресс, стратегии промышленно-технологического обновления и инновационного развития неразрывно связаны с успехами в электронике. С каждым годом все более отчетливо прослеживается следующая доминанта: большинству специалистов при решении многих задач приходится иметь дело с различными электротехническими и электронными устройствами, построенными с применением последних достижений науки и новых технологий.

При подготовке специалистов со специфической сферой профессиональных интересов неэлектротехнические вузы, учитывая диалектику современных процессов, включают в программу обязательной подготовки самостоятельную дисциплину «Основы электротехники и электроники». Этот курс, следуя по направлению научно-технического прогресса, отражая тенденции развития теоретического арсенала и элементной базы, объединяет и систематизирует наиболее важные содержательные аспекты и принципы в области электротехники и электроники.

Основой учебного пособия стал курс лекций «Основы электротехники и электроники», который читается в Харьковском национальном экономическом университете студентам, обучающимся по специальностям 7.092704 «Комп'ютерізовані технології та системи видавничо-поліграфічних виробництв» и 7.092702 «Технологія мультимедійних видань». Он состоит из двух частей: «Элементы общей теории электротехники и электроники» и «Реализация изделий электротехники и электроники». В этом учебном пособии приведен основной теоретический материал по второй части курса «Элементы общей теории электроники». Учебная дисциплина, кроме лекций состоит из практических занятий и лабораторных работ, материал которых излагается в отдельном издании. Учебный материал по первой части курса «Элементы общей теории электротехники» изложен авторами в опубликованных ранее двух пособиях (отдельно – теоретический курс и практикум).

Содержание учебного пособия основывается на концепции возможно более тесного сближения теоретического материала с практикой. Это, в первую очередь, определило принцип обора материала и степень

детальности освещения. Он отличается разнообразием содержания, обилием понятий и методов.

Основная цель авторов – четко, строго и логично изложить материал учебной дисциплины в соответствии с современными стандартами высшего образования в Украине. Пособие написано в самостоятельной форме и таким образом, чтобы у студента не возникало необходимости обращаться к дополнительным литературным источникам, широкий список которых рекомендован для самоподготовки, расширения и углубления знаний. Помогут при усвоении материала и развитии навыков самостоятельной работы приведенные в каждом разделе простые и наглядные примеры и модели, а также широкий круг контрольных вопросов.

В первом разделе последовательно излагаются принципы построения, модели, характеристики и параметры простейших структур полупроводниковой электроники – идеализированных диодов. Этот материал является основой для изучения работы более сложных приборов, поэтому подробно рассматриваются особенности построения идеальных диодов, современные представления об их функционировании, проводится расчет параметров и характеристик, позволяющих, исходя из требований к конкретному устройству, грамотно выбирать нужный полупроводниковый прибор. Подобного рода структуры служат составной частью более сложных полупроводниковых приборов. В несколько сокращенном виде излагаются наиболее общие функции диодов, связанные с преобразованием сигналов в электронных устройствах, и вопросы, касающиеся применения приборов в устройствах силовой электроники.

Второй раздел логически продолжает материал первого и посвящен изучению моделей транзисторов, с помощью которых можно отразить принципы построения и механизмы функционирования прибора, изучаются физические явления, параметры и характеристики наиболее распространенных типов транзисторов. Цель этого раздела состоит в том, чтобы помочь студенту изучить основные компоненты электрических схем, освоить работу с их математическими моделями, понять физические процессы, происходящие в них, научиться представлять внешний вид реального изделия и расшифровывать графический значок, применяемый для обозначения элементов в схеме.

В третьем разделе мы переходим к изучению основ микроэлектроники – области науки и техники, связанной с изготовлением и использо-

ванием электронных цепей малых размеров. Важнейшим компонентом в этом случае становится интегральная схема. Микроминиатюрное исполнение элементов в интегральной схеме меняет как характеристики и параметры самих элементов, так и отдельные принципы построения, в целом, самих схем. Поэтому в этом разделе рассмотрены особенности элементов интегральных схем и некоторые схемотехнические приемы, используемые при реализации функциональных блоков изделия (усилителей, формирователей и преобразователей электрических сигналов, элементов цифровых электронных схем и т. д.), а также отдельные аспекты производства интегральных схем. Изучается важнейший элемент энергонезависимой, быстро перепрограммируемой памяти. В конце раздела рассматриваются приборы преобразования неэлектрических величин в электрический сигнал. При этом упор делается на полупроводниковые датчики, в частности датчики магнитного потока.

Оптоэлектроника – направление в науке и технике, соединившее в себе возможности оптики и электроники. Являясь синтезом двух направлений, оптоэлектроника превосходит каждое из них по возможностям. Поэтому в четвертом разделе мы рассматриваем класс явлений, широко используемых в настоящее время в электронной технике, – явления излучения, поглощения, усиления оптического излучения веществом. Изучаются физические принципы, лежащие в основе работы широко распространенных изделий – оптоэлектронных приборов, а также рассматриваются свойства и параметры оптического излучения. Основное внимание уделено приборам, предназначенным для создания оптического излучения – излучающим диодам, полупроводниковым лазерам, а также наиболее распространенным типам фотоприемников.

С целью глубокого понимания физических принципов работы рассматриваются важные для оптоэлектронных устройств вопросы, посвященные поглощению света веществом и внутреннему фотоэффекту. Кратко, но достаточно информативно рассмотрены имеющие большое значение вопросы, касающиеся создания оптоэлектронных устройств, работающих в системах передачи и отображения информации.

В современной жизни в ряде областей, связанных, к примеру, с применением в промышленных устройствах электродвигателей переменного напряжения, бесшумных и энергосберегающих электродвигателей бытовой техники, исполнительных устройств промышленной авто-

матики, источников бесперебойного питания и прочее, приходится «работать» с токами в единицы, сотни и даже тысячи ампер. В ряде случаев необходимы также устройства, работающие с напряжениями до 6000 В. Область электроники, работающая с такими диапазонами напряжений и токов, получила название энергетической или силовой электроники. Современный специалист должен знать основные положения силовой электроники. Поэтому четвертый раздел посвящен изучению полупроводниковых изделий силовой электроники, начиная с изучения классических тиристоров, работающих на основе положительной обратной связи. Эти приборы в настоящее время широко применяются. Особую важность представляет материал по биполярным транзисторам с изолированным затвором. Такие транзисторы почти идеальны в качестве высоковольтных ключей в устройствах, где существенным является низкое падение напряжения в открытом состоянии, таких как, двигатели постоянного и переменного тока, источники питания для соленоидов и т. д.

По электронике написано много хороших книг. Однако их структура и объем, недостаточное отражение современных тенденций в развитии электроники, по мнению авторов, обычно недостаточно или мало соответствует требованиям, предъявляемым к обучению студентов в неэлектротехнических вузах. Авторы при написании учебного пособия постарались устранить некоторые недостатки, широко используя современные представления и методы анализа функционирования элементов различных электронных устройств, анализ работы наиболее важных, типовых современных электронных приборов различного назначения.

Данная книга может быть полезна студентам, обучающимся по другим направлениям и специальностям, а также для изучающих одноименную дисциплину по другим профилям подготовки.

Авторы благодарны за оказанную помощь в улучшении содержания учебного пособия сотрудникам кафедры физики и электроники Харьковского национального экономического университета (зав. кафедрой кандидат физико-математических наук, доцент Е.А. Бондаренко).

Авторы выражают глубокую признательность рецензентам – доктору технических наук, профессору кафедры экспериментальной физики Харьковского национального университета им. В.Н. Каразина В.П.Пойде; доктору физико-математических наук, профессору кафедры полупроводниковой и вакуумной электроники Харьковского национального универ-

ситета им. В.Н. Каразина Ю.В. Аркуше; кандидату физико-математических наук, доценту кафедры микроэлектроники, электронных приборов и устройств Харьковского национального университета радиоэлектроники А.Ю. Панченко за внимательное прочтение рукописи и высказанные полезные замечания.

1. Основные структуры полупроводниковой электроники и диоды на их основе

В электрических цепях и устройствах электроники очень широко используют разнообразные полупроводниковые приборы. С точки зрения понимания их принципа действия, функциональных возможностей они намного сложнее по сравнению с изучаемыми ранее компонентами цепи. Поэтому их изучение мы начинаем с напоминания основных положений физики полупроводников и современных представлений об эффектах в полупроводниках при различных воздействиях на них. При этом используются две модели. Модель ковалентной связи проста и наглядна, дает возможность на уровне интуиции описать физические процессы и получить первоначальные представления о работе полупроводниковых приборов. Модель энергетических зон позволяет количественно изучать явления переноса носителей заряда, оценивать электрофизические параметры и получать требуемые характеристики приборов. Затем мы приступаем к изучению принципов построения, моделей, характеристик и параметров простейших структур полупроводниковой электроники – идеализированных диодов. Этот материал является основой для изучения работы более сложных приборов, поэтому подробно рассматриваются особенности построения идеальных диодов, современные представления об их функционировании, проводится расчет параметров и характеристик, позволяющих, исходя из требований к конкретному устройству, грамотно выбирать нужный полупроводниковый прибор. В сокращенном виде излагаются наиболее общие функции диодов, связанные с преобразованием сигналов в электронных устройствах, и вопросы, касающиеся применения приборов в устройствах силовой электроники.

1.1. Основные понятия и материалы полупроводниковой электроники

1.1.1. Основные материалы полупроводниковой электроники и понятия их характеризующие

Вероятнее всего, исходя из опыта, приобретенного в школе и в повседневной жизни, вы уверенно различаете по тому, как вещества проводят ток, *проводники и диэлектрики*: проводники хорошо проводят ток, а диэлектрики, наоборот, практически его не проводят. Вы пользуетесь диэлектриками, как говорят еще изоляторами, для того, чтобы избавиться себя от воздействия напряжения, а проводами (проводниками в изоляции) чтобы подводить напряжение к различным бытовым приборам.

Как известно из физики, названные материалы имеют существенно разное удельное электрическое сопротивление.

Удельное электрическое сопротивление, если вспомнить физику, есть скалярная физическая величина, характеризующая данный материал (вещество), численно равная сопротивлению бруска материала с постоянным сечением S и длиной L .

Обычно используют брусок кубической формы со стороной один сантиметр, и материал характеризуют удельным объемным электрическим сопротивлением ρ – сопротивлением одного кубического сантиметра материала. Единица измерения удельного объемного сопротивления – Ом×см.

Диэлектрические материалы характеризуются очень большим удельным объемным сопротивлением $\rho > 10^{10}$ Ом×см (брусок со стороной 1 см имеет электрическое сопротивление $> 10^{10}$ Ом). Понятно, что через диэлектрик, исходя из закона Ома, протекает малый ток, если он даже находится под воздействием сравнительно большого напряжения (десятки и сотни тысяч вольт). В качестве диэлектрика в электротехнике используют специальные виды фарфора ($\rho > 10^{14}$), стекла ($\rho > 10^{12}$), пластмассы, фторопласт-4 ($\rho > 10^{16}$) и прочее. Диэлектрические материалы используют для отделения одной токопроводящей части электрической цепи от другой, межслойной и защитной изоляции.

Проводящие материалы характеризуются очень малым удельным объемным сопротивлением $\rho < 10^{-4}$ Ом×см (брусочек со стороной 1 см имеет электрическое сопротивление $< 10^{-4}$ Ом). Они, обладая высокой электропроводностью, используются в качестве токопроводящих элементов. В качестве проводников в электротехнике используют алюминий ($\rho = 2,6 \times 10^{-6}$ Ом×см), золото ($\rho = 2,2 \times 10^{-6}$ Ом×см), серебро ($\rho = 1,6 \times 10^{-6}$ Ом×см), вольфрам ($\rho = 2,6 \times 10^{-5}$ Ом×см) и т.п. Хорошо проводящие материалы используют в качестве токопроводящих цепей аппаратуры (проводов, металлических дорожек, внешних выводов приборов, перемычек, межсоединений и т.д.).

Существует в настоящее время еще одна весьма обширная группа материалов (около 2000 видов), которая занимает как бы промежуточное положение между проводниками и диэлектриками.

К *полупроводниковым материалам* относят вещества, занимающие по удельному электрическому сопротивлению промежуточное положение между проводящими материалами и диэлектрическими ($\rho > 10^{-4}$ – $\rho < 10^{10}$) Омсм. Конкретное значение ρ зависит от вида материала полупроводника и от количества посторонних веществ, находящихся в полупроводнике, то есть от концентрации примеси, и других факторов.

Несмотря на формальное «среднее» положение полупроводниковые материалы, как увидим в дальнейшем, обладают рядом уникальных особенностей, использование которых позволяет создавать разнообразные полупроводниковые приборы и микроэлектронные изделия.

1. Во-первых, наиболее существенной особенностью полупроводников является их способность изменять свойства электропроводности (в первую очередь величину ρ) в очень широких пределах под воздействием внешних факторов: температуры; электрического и магнитного поля; освещения.

2. Во-вторых, в полупроводниковых материалах можно относительно просто (за счет введения примесей, инжекции и прочее) менять количество носителей заряда и характер их движения.

Несмотря на то, что к полупроводникам можно отнести примерно 2000 веществ, в полупроводниковой электронике в настоящее время наибольшее распространение получили, в первую очередь, кристаллические кремний, арсенид галлия, германий. Заметим, что в современных приборах германий уже практически не используется.

Доминирующее положение в полупроводниковой электронике, и, как увидим в дальнейшем, в микроэлектронике занимают кремний и арсенид галлия. Это обусловлено следующим.

1. *Кремний* (Si от лат. silicius), самый распространенный после кислорода элемент на Земле (26%), обладает высокой технологичностью процессов изготовления из него изделий.

2. Кремний и приборы на его основе, как увидим, лучше, чем германий, сохраняет свои свойства при изменении температуры. Низкая температурная чувствительность позволяет использовать его в изделиях, работающих при относительно высоких температурах ($T < 120^{\circ}\text{C}$).

3. Кремний позволяет относительно просто получать на своей поверхности пленки диэлектрика (SiO_2 – двуокись кремния, Si_3N_4 – нитрид кремния), которые надежно изолируют кремний от окружающей среды, а, с другой стороны, схемные элементы друг от друга.

4. *Арсенид галлия* (GaAs), по сравнению с кремнием, позволяет получить высокое быстродействие.

Однако, говоря об арсениде галлия, следует иметь в виду следующее.

Распространенность в земной коре As - $5 \cdot 10^{-4}\%$, Ga - $1 \cdot 10^{-4}\%$. Поэтому добыча редких и рассеянных элементов сопряжена с большими затратами энергии.

Технология изготовления GaAs сложнее (исходные элементы его токсичны для человеческого организма, он плохо поддается легированию, по диэлектрическим и маскирующим свойствам пленки диэлектрика его уступают SiO_2 и Si_3N_4).

1.1.2. Собственные полупроводники и понятия их характеризующие

Для понимания физических процессов, происходящих в полупроводниковых приборах (а именно на них мы в дальнейшем сосредоточим внимание), прежде всего, необходимо знать, как устроен полупроводник, какими электрофизическими параметрами его характеризуют, представлять за счет чего (каких носителей заряда) через вещество протекают токи и прочее. Поэтому на этих вопросах мы сейчас и остановимся.

Напомним, что, как известно из курса физики, всегда следует различать *собственные* (от лат.- intrinsic – собственный; говорят еще типа i) и *примесные* (р – типа или n – типа) кристаллические полупроводники.

Собственным называют бездефектный (с идеальной кристаллической решеткой) тщательно очищенный от посторонних случайных примесей (как говорят химически «чистый» - количество примесей не превышает 10^{-6} %) кристаллический полупроводник.

Собственный полупроводник, в идеале, представляет собой кристаллическое тело, в котором атомы размещены в пространстве в строго определенной последовательности и на постоянном расстоянии друг от друга, как говорят, в узлах кристаллической решетки. Упрощенная картина расположения атомов кремния (фрагмент кристаллической решетки, представленный на плоскости), показана на рис. 1.1.

Диаметр атома – 10^{-10} м. Объемная плотность атомов в кристалле, т.е. их количество в одном кубическом сантиметре вещества, имеет порядок $N = 10^{22}$ 1/см³.

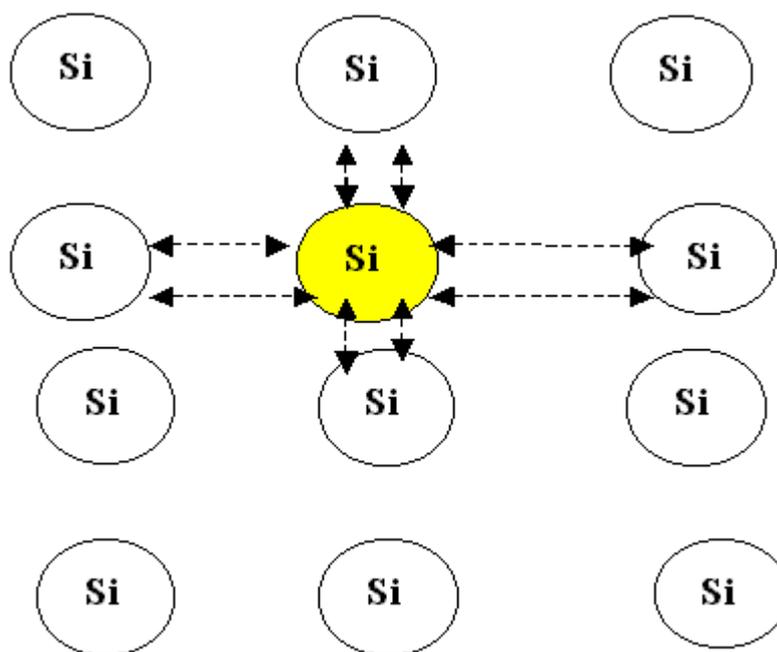


Рис. 1.1. Упрощенное плоскостное расположение атомов кристаллической решетки кремния с образованными ковалентными СВЯЗЯМИ

Вокруг ядер атомов, составляющих остов кристаллической решетки, с точки зрения планетарной модели Э. Резерфорда, по своим орбитам движутся микрочастицы – электроны. Электроны можно представить, с точки зрения корпускулярных свойств, в виде счетного количества неких «твердых шариков», которые взаимодействуют с ядром и друг с другом. Количество орбит определяется положением полупроводника в таблице периодической системы Д. Менделеева.

Если воспользоваться таблицей элементов Д. Менделеева (IV колонка таблицы), то можно констатировать, что для кремния и германия имеет место следующее распределение электронов по оболочкам (орбитам) или, как еще говорят, следующая электронная конфигурация (таблица 1.1.).

Таблица 1.1

Распределение электронов по орбитам атомов

| Вещество | Количество электронов на оболочке, имеющей условное обозначение в виде латинской буквы | | | |
|---------------|--|---|----|---|
| | K | L | M | N |
| Кремний (Si) | 2 | 8 | 4 | |
| Германий (Ge) | 2 | 8 | 18 | 4 |

Считается, что электроны на оболочках K, L (на внутренних орбитах) для кремния и на K, L, M для германия «сильно» связаны с ядром. Электроны каждого атома, расположенные на внешних орбитах, их еще называют валентными, «слабо» связаны с ядром.

Ближайшие атомы в кристаллах расположены таким образом, что их внешние электронные оболочки перекрываются друг с другом, из – за чего валентные электроны на них становятся общими. Говорят, что у ближайших соседних атомов валентные электроны на некоторое время обобществляются. Для «центрального» атома (выделен на рис. 1.1), имеющего 4 «соседа», на внешней орбите, в общем случае, побывает 8 электронов: 4 своего центрального «хозяина» и 4 «соседских» атома. «Соседские» электроны проводят лишь часть времени возле «центрального» атома. Другую часть времени они пребывают у своего искон-

ного «хозяина». В результате такого слияния возникают значительные межатомные силы, которые называют ковалентными связями.

В первом, весьма грубом, приближении можно считать, что при температуре абсолютного нуля ($T=0^0$ К) в полупроводнике, имеющем идеальную кристаллическую решетку, все электроны движутся по своим орбитам, а валентные электроны к тому же задействованы на обеспечение ковалентной связи. «Свободных» электронов нет, а, следовательно, если нет носителей заряда, ток через такой полупроводник протекать не может.

В случае, когда температура окружающей среды не равна нулю, полупроводник подвергается воздействию тепловой энергии в виде потока тепловых фононов. Из-за энергии тепловых фононов первоначальная энергия валентных электронов в ковалентных связях может возрасти настолько, что некоторые электроны могут «порвать, разорвать» ковалентные связи и стать, уйдя в межатомное пространство, «свободными». Не «привязанный» к своему и ближайшим соседним атомам «свободный» электрон уже может хаотически перемещаться по кристаллу, перенося заряд. Ток такие заряды тоже пока создавать не могут, поскольку для создания тока необходимо упорядоченное движение зарядов.

«Освобождение» какого – либо валентного электрона приводит к тому, что в системе ковалентных связей возникает «пустое место». Отсутствие электрона в решетке называют «дыркой». Причем, дырку можно рассматривать как некоторую фиктивную частицу, аналогичную электрону, но с зарядом противоположного знака.

Итак, исходя из модели ковалентных связей, можно констатировать, что в полупроводнике под воздействием внешней тепловой энергии (температуры) идет процесс образования «свободных» электронов и дырок в результате которого в кристалле имеется определенное количество «свободных» носителей заряда. Этот процесс называют термогенерацией электронно - дырочных пар (подвижных носителей заряда).

Наряду с этим существует и противоположный, конкурирующий процесс, когда, отдав полученную дополнительную энергию, электрон может опять занять место в системе ковалентных связей. Этот процесс называется *рекомбинацией* (воссоединением) электронов и дырок.

При заданной температуре окружающей среды имеет место термодинамическое равновесие между тепловой генерацией и рекомбинацией носителей заряда, в результате чего устанавливается некоторая, вполне определенная концентрация свободных электронов и дырок.

Понятно, что в состоянии термодинамического равновесия для собственного полупроводника концентрации (количество частиц в одном кубическом сантиметре) электронов и дырок равны. Это символически можно выразить в виде следующей формулы

$$n_i = p_i. \quad (1.1)$$

Следует привыкнуть к используемой в формулах подобного рода системе обозначений. Символами n и p обозначают концентрацию электронов и дырок соответственно. Нижний индекс i отражает тип полупроводника, в котором определена концентрация (здесь типа i , от лат.- intrinsic – собственный).

Произведение концентраций в таком полупроводнике называют *законом действующих масс для собственного полупроводника*:

$$n_i \cdot p_i = n_i^2. \quad (1.2)$$

Использование представлений полупроводника, основанных на ковалентных связях, позволяет наглядно представить, что происходит в полупроводнике. Однако такие ценные качества модели ковалентных связей, как простота, наглядность, понятность не могут перевесить ее основного недостатка: с помощью этой теории фактически нельзя ничего точно подсчитать. К сожалению, используемый в этом случае упрощенный подход к описанию внутреннего строения вещества, вообще не позволяет объяснить ряд интереснейших свойств, нашедших практическое применение в полупроводниковой электронике. Поэтому «ковалентные связи» используют на этапе, предшествующем строгому изучению полупроводников.

Все дело в том, что электроны полупроводника, как известно из физики, проявляют не только корпускулярные свойства. Эти микрочастицы органически сочетают в себе также квантовые свойства и, с этой

точки зрения, они являются волнами вещества (волновыми процессами или точнее, пакетами волн) и подчиняются волновым законам квантовой механики. В отличие от классических карпускулярных частиц, образно представляемых в виде «шариков», квантовые частицы, как волны, принципиально нельзя пронумеровать и пересчитать, поскольку они принадлежат всему кристаллу и по этому признаку неотличимы друг от друга. Сами волновые процессы носят вероятностный характер, и движение частиц описывается волновыми уравнениями. С позиций квантовой физики не может быть и речи ни о какой траектории движения частиц в классическом смысле, поскольку они как бы «размазаны» по всей области пространства полупроводника.

Сталкиваясь с гибридами «карпускулярно – волнового дуализма», получивших странное название «волно – частиц» или «частиц - волн», обладающими одновременно свойствами как частицы, так и волны, с представлениями того, что непосредственно не воспринимается нашими ощущениями, мы невольно ощущает некий психологический тормоз в восприятии изучаемого и нас все время мучает вопрос «как же так может быть?». С другой стороны, мы понимаем, что это вынужденный способ описания микромира и без него нам не обойтись.

Ученые, понимая сложность квантовой физики, разработали ряд наглядных средств, позволяющих «пробиться» через непредставимость наглядного образа «волны – частицы».

Поскольку энергия выступает мерилем любого движения, и изменение энергии может быть связано с любыми формами перехода из одного состояния в другое, удобно опираться при описании «волны – частицы» на упрощенные энергетические модели движения микрочастиц. По аналогии с тем, как в повседневной жизни, пытаясь выразить смысл, трудный для человеческого восприятия, пользуются языком символов и абстрактных образов, было введено понятие энергетической диаграммы полупроводника.

Энергетическая диаграмма, являясь самой по себе бессмысленным изображением в виде линий, является абстрактным образом, пытающимся передать то, что не может представить наше воображение о «частицах – волнах», о процессах их движения.

В самом простейшем виде энергетическая диаграмма одномерного приближения полупроводника представляет некий специфический гра-

фик, отображающий набор энергетических уровней (рис. 1.2). На диаграмме, по аналогии с потенциальной энергией тела, поднятого на различные высоты относительно Земли, каждому значению энергии микрочастицы W_K , отражающему процесс ее движения, ставится в соответствие на оси ординат уровень на энергетической диаграмме. При этом считается, что чем выше уровень, тем большей энергией W обладает электрон, отражаемый данным уровнем. Ось абсцисс специфическим образом характеризует пространственную координату. Если полупроводник не подвержен внешнему воздействию, то энергетический уровень микрочастицы имеет вид горизонтальной линии. Это отражает тот факт, что энергия электрона, «размазанного» по всему кристаллу полупроводника, во всех точках полупроводника одна и та же и не зависит от координаты электрона.

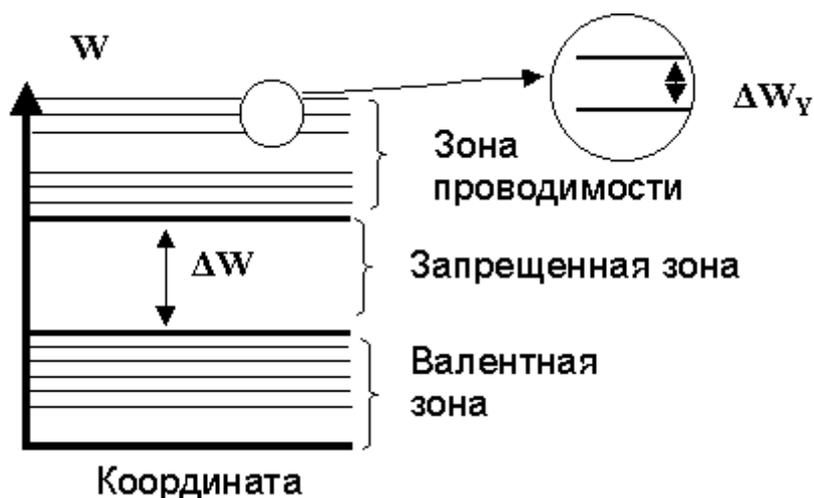


Рис. 1.2. Энергетическая диаграмма, показывающая верхние зоны полупроводника: ΔW – ширина запрещенной зоны; ΔW_γ – разность энергий между уровнями. ΔW порядка 1 эВ, ΔW_γ – 10^{-15} эВ

В общем случае, все микрочастицы – электроны, локализованные в области полупроводника, характеризуются набором уровней. При этом энергетические уровни, как известно из физики, могут быть лишь «разрешенными». Это обусловлено тем, что в полупроводнике с упорядоченным расположением атомов все микрочастицы, как говорят физики, энергетически квантованы. Иначе говоря – энергия микрочастицы может

быть только такой, которой ей «разрешено», «дозволено» быть кристаллической решеткой. Квантованность энергии микрочастицы является следствием ее волновых свойств и имеет физический смысл того, что некоторые волны в принципе не могут существовать в периодической кристаллической решетке.

Следует иметь в виду, что при изображении энергетических диаграмм широко используются различные упрощения и общепринятые соглашения по этому поводу. Основные из них следующие.

1. Энергетические диаграммы условно разделяются на зоны (рис. 1.2). Дело все в том, что количество энергетических уровней, соответствующих микрочастицам, очень велико, а энергетический зазор между уровнями ΔW_{γ} очень мал. Поэтому для полупроводника характерно то, что уровни группируются в наборы уровней (зоны), разделенные полосами, где должны находиться запрещенные энергетические состояния. В общем случае энергетическая диаграмма по своему начертанию напоминает «зевру»: полосы разрешенных уровней чередуются с участками запрещенных уровней.

Объединенные наборы близко расположенных друг от друга «разрешенных» энергетических уровней получили название разрешенных зон энергетической диаграммы. Фрагмент энергетической диаграммы, расположенный в определенном энергетическом диапазоне, на котором в принципе не могут располагаться энергетические уровни, получил название запрещенной зоны.

2. На энергетической диаграмме, как правило, изображают не все возможные зоны. Делается это по причине того, что не все микрочастицы, в классическом понимании, фактически участвуют в формировании потоков, обуславливающих протекание тока. Оказывается достаточным, для пояснения физических процессов в полупроводнике, рассматривать лишь «поведение» валентных электронов, размещенных на внешней оболочке атома. Это позволяет исключить из рассмотрения и, соответственно, не изображать на энергетической диаграмме «глубокие» уровни, лежащие ниже так называемой «валентной» зоны.

3. На энергетической диаграмме в разрешенных зонах изображают всего лишь несколько энергетических уровней. Из-за того, что количество уровней, соответствующих возможным энергетическим состояниям, даже для валентных электронов составляет величину порядка 10^{22} , то

на диаграмме либо вообще не показывают уровни разрешенных зон, либо чисто символически изображают отдельные из них, позволяющие пояснять физические процессы. Наибольший уровень валентной зоны называют «*потолком валентной зоны*», а наименьший уровень зоны проводимости – «*дном зоны проводимости*».

Для пояснения сказанного рассмотрим энергетическую диаграмму собственного полупроводника в одномерном приближении, фрагмент которой, не содержащий «глубоких» зон, показан на рис. 1.3.

Собственный полупроводник, как видно из рисунка, представлен тремя верхними зонами уровней энергии, названными валентной, запрещенной и зоной проводимости.

Стационарные энергетические состояния «связанных» валентных электронов представлены на диаграмме энергетическими уровнями валентной зоны. Энергетические состояния «свободных» электронов, обладающих энергиями, достаточными для «отрыва» от своего атома, характеризуются уровнями зоны проводимости. Зону проводимости и валентную зону разделяет запрещенная зона.

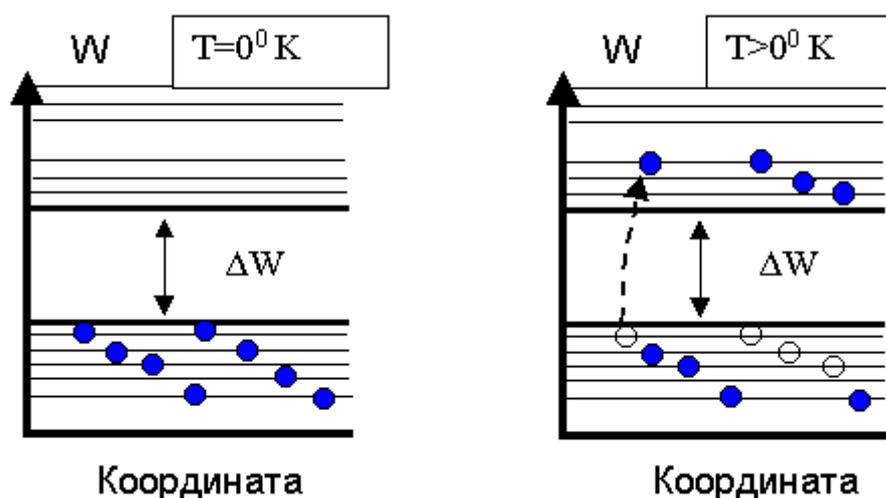


Рис. 1.3. Энергетическая диаграмма собственного полупроводника в одномерном приближении при температуре абсолютного нуля и ($T>0^0 \text{ K}$)

При температуре абсолютного нуля ($T=0^0 \text{ K}$), в первом приближении, можно считать, что все уровни валентной зоны, как говорят, полностью «заселены» электронами (по два электрона на каждом уровне), а

уровни зоны проводимости «пусты». Несколько виртуальных (воображаемых, мысленно представляемых) квазичастиц, отображающих на энергетической диаграмме «расселение» по энергетическим уровням электронов, на рисунке условно показаны зачерненными кружочками. Изображение виртуальных квазичастиц на энергетической диаграмме означает, что все электроны в кристалле полупроводника при $T=0^0$ К находятся в валентной зоне и их волновые пакеты характеризуются минимальной энергией. С точки зрения классической физики это означает, что все электроны движутся по орбитам своих атомов, не имея возможности «оторваться» от них, и участвуют в системе ковалентных связей.

Естественно, что в этом случае, из-за «незаселенности» зоны проводимости, никакое дополнительное движение квазичастиц, представляющих электроны, через кристалл невозможно. В валентной зоне также нет «свободных» уровней и, следовательно, также невозможны какие либо движения в кристалле с уровня на уровень и переносы энергии волновыми пакетами. Напомним, что «занятие» одного и того же уровня виртуальными квазичастицами, представляющими электроны на энергетической диаграмме, в количестве, превышающим число 2, недопустимо из-за принципа запрета В. Паули.

При температуре окружающей среды, превышающей температуру абсолютного нуля ($T>0^0$ К), отдельные виртуальные квазичастицы, представляющие на энергетической диаграмме валентные электроны, от тепловых фононов могут получать дополнительную порцию энергии, «переходить» в зону проводимости и «заселять» там один из свободных уровней (рис. 1.3). Образовавшиеся в результате такого перехода «незаселенные» энергетические уровни в валентной зоне окажутся на какое то время свободными.

Переходя в зону проводимости, квазичастицы – электроны (в дальнейшем для краткости, электроны), как бы становятся «свободными». В зоне проводимости «свободные электроны» могут изменять свою энергию, «перепрыгивать» с уровня на уровень – свободных уровней там достаточно – и участвовать в любых движениях.

Наличие «незаселенных» уровней в валентной зоне позволяет также перемещаться виртуальным квазичастицам, представляющим на диаграмме электроны, по уровням валентной зоны. При «уходе» какого либо электрона с энергетического уровня валентной зоны на уровень

зоны проводимости на «освободившееся место» сразу же «устремляется» электрон с другого уровня валентной зоны, расположенного на энергетической диаграмме ниже «освободившегося» уровня. В свою очередь, на это, вновь «освободившееся» место сразу же «приходит» электрон с еще более низкого уровня. Затем этот процесс повторяется для нового «свободного» уровня. Несложно сообразить, что при подобных перемещениях электронов по уровням валентной зоны «пустой» энергетический уровень движется вглубь валентной зоны, напоминая передачу эстафеты. Такое движение электрона вглубь валентной зоны удобнее трактовать как перемещение в валентной зоне нового фиктивного «свободного» носителя заряда – дырки, которой присущ положительный заряд, по абсолютной величине равный заряду электрона. При этом, ввиду изменения знака заряда у дырок по отношению к электронам, дырка, переходя на более низкий уровень валентной зоны, обретает большую энергию. Несколько дырок, «расселенных» по энергетическим уровням на рис. 1.3 условно показаны кружочками.

Тем самым, при температуре окружающей среды, превышающей температуру абсолютного нуля ($T > 0^0 \text{ K}$), в валентной зоне и в зоне проводимости в результате энергетического «возбуждения» появляются «свободные» носители заряда (электроны и дырки), способные, при определенных условиях создавать электрический ток.

Опять же следует понимать, что наряду с процессами генерации «свободных» носителей заряда существует и противоположный процесс: «свободный» электрон возвращается на незаполненный уровень в валентной зоне и происходит рекомбинация (воссоединение электрона и дырки) и пара электрон – дырка исчезает. При заданной температуре окружающей среды имеет место термодинамическое равновесие между процессами тепловой генерации и рекомбинации, в результате чего в зоне проводимости устанавливается вполне определенная концентрация свободных носителей электронов, а в валентной зоне – дырок проводимости.

Время, по истечении которого электрон, как квазичастица на энергетической диаграмме, возвращается из зоны проводимости в валентную зону, называется «временем жизни».

В условиях термодинамического равновесия свободные электроны и дырки возникают и исчезают парами, поэтому полупроводник остается в целом электронейтральным.

Следует обратить внимание на то, что электропроводность полупроводника появляется под действием внешнего фактора, способного сообщить квазичастицам – электронам валентной зоны энергию, достаточную для «переброса» их в зону проводимости. Такими факторами могут быть нагревание полупроводника, облучение его светом или ионизирующим излучением, введение специальной примеси и прочее.

Еще раз обратим внимание на то, что энергетическая диаграмма является не более, чем условным графическим изображением, которое облегчает процесс изучения генерации пар электрон - дырка и движения виртуальных квазичастиц, как носителей заряда в полупроводнике. Энергетические уровни, «заселение» их виртуальными квазичастицами и прочее, особым образом представляют некий «энергетический спектр» полупроводника и своеобразно трактуют движение электронов, отражая тот факт, что при одной совокупности уровней есть движение «свободных» квазичастиц (соответственно, есть энергия, энергетический уровень) и электропроводность возможна, при других – нет и невозможна. Уровни - это не «каналы», в которых происходит движение электронов.

Из сказанного следует, что собственный полупроводник характеризуют следующими электрофизическими параметрами.

1. *Шириной запрещенной зоны ΔW .*

Полезно запомнить, чтобы лучше представлять используемые в электронике материалы, что при комнатной температуре ширина запрещенной зоны имеет следующие, приведенные в таблице 1.2, значения.

Таблица 1.2

Ширина запрещенной зоны основных полупроводников

| Полупроводник | Кремний | Арсенид галия | Германий |
|----------------------------|---------|---------------|----------|
| Ширина ЗЗ, ΔW , Эв | 1,12 | 1,43 | 0,70 |

2. *Равновесная концентрация электронов (дырок) в зоне проводимости (в валентной зоне).*

Чтобы получить выражения, определяющие концентрацию электронов в зоне проводимости необходимо проделать следующее. Сначала найти функцию $N(W)$, которая описывает распределение уровней в соответствующей зоне. Затем следует вычислить величину $N(W)dW$, представляющую собой число разрешенных энергетических уровней, приходящихся на единицу объема и лежащих в пределах от W до $W + dW$. Также надо найти функцию, которая в условиях термодинамического равновесия описывала бы вероятность того, что некоторая частица занимает определенный энергетический уровень. Наконец, зная количество уровней и вероятность нахождения на них электронов, интегрируя произведение этих величин от нижней до верхней границы зоны проводимости, можно найти искомую концентрацию носителей заряда.

Опуская громоздкие выкладки, сразу запишем конечное выражение для определения равновесной концентрации электронов и дырок в собственном полупроводнике:

$$n_i = N_E \ell \frac{(\Delta W - W_F)}{kT}, \quad p_i = N_V \ell \frac{(\Delta W - W_F)}{kT}, \quad (1.3)$$

где N_E (N_V) – эффективная плотность состояний в зоне проводимости (в валентной зоне), значение которых примерно равны $2 \cdot 10^{-19} \text{ 1/см}^{-3}$; $k = 1,38 \cdot 10^{-23} \text{ Дж/К}$ – постоянная Больцмана.

Входящая в (1.3) величина W_F – уровень (энергия) Ферми. W_F упрощенно можно трактовать или как уровень, вероятность заполнения которого равна 0,5, или как среднюю энергию, приходящуюся на один носитель заряда. В собственном полупроводнике уровень Ферми расположен практически в середине запрещенной зоны.

Выражение (1.3), важнейшее для физики полупроводников, свидетельствует о том, что собственная концентрация электронов экспоненциально (т. е. очень, очень резко) зависит от ширины запрещенной зоны и температуры.

При постоянной температуре из-за разницы в величине запрещенной зоны у арсенида галлия и германия ($1,43 - 0,7 = 0,73 \text{ эВ}$) величины собственных концентраций носителей заряда будут отличаться примерно в миллион раз.

То же самое можно говорить о влиянии температуры на полупроводник. При этом следует учитывать, что влияние температуры на соб-

ственную концентрацию будет тем слабее, чем больше величина запрещенной зоны.

При расчетах собственной концентрации в зависимости от температуры следует учитывать, что N_E и $N_V \sim T^3$, а ширина запрещенной зоны линейно зависит от температуры. Поэтому практические расчеты удобно проводить по следующей формуле:

$$n_i(T) = BT^{3/2} \ell^{-\frac{\Delta W}{2kT}}, \quad (1.4)$$

где для кремния $B = 3,9 \cdot 10^{16} \text{ K}^{-3/2} \text{ см}^{-3}$.

Полезно иметь в виду, чтобы лучше представлять наиболее используемые в электронике материалы, что при комнатной температуре концентрация собственных носителей заряда следующая (см. табл. 1.3).

В настоящее время, наряду с арсенидом галлия, широко используются и другие соединения группы $A^{III} B^V$ (латинские номера – группы таблицы Д. Менделеева). Надо иметь в виду, что при увеличении атомных номеров элементов A^{III} и B^V , входящих в соединение, ширина запрещенной зоны будет меньше. Так, GaP (номера 31,15), полупроводниковые соединения $A^{III} B^V$, в состав которых входит алюминий (13), а также соединения, состоящие из компонентов с низкими значениями атомного номера, имеют ширину запрещенной зоны ~ 2 эВ, в отличие от GaAs (31, 33), InP (49, 15), которые имеют ширину запрещенной зоны $\sim 1,4$ эВ.

Таблица 1.3

Концентрация собственных носителей заряда

| Полупроводник | Кремний | Арсенид галлия | Германий |
|---------------------------------------|----------------------|------------------|---------------------|
| Концентрация, n_i , $1/\text{см}^3$ | $1,45 \cdot 10^{10}$ | $1,3 \cdot 10^7$ | $2,4 \cdot 10^{13}$ |

1.1.3. Примесные полупроводники и понятия их характеризующие

Для целевого изменения свойств «чистые», собственные полупроводники легируют, т. е. в них, путем использования специальных технологий, вводят атомы примесей. Легирование позволяет менять тип про-

водимости полупроводника (с собственного на примесный: «n» – электронный или «р» – дырочный) и его удельное объемное сопротивление.

Обратимся опять к таблице Д.И. Менделеева, фрагмент которой представлен в таблице 1.4.

Таблица 1.4

Фрагмент периодической таблицы элементов Д.И. Менделеева

| II | III | IV | V | VI |
|----|-----|----|----|----|
| Be | B | C | N | O |
| Mg | Al | Si | P | S |
| Zn | Ga | Ge | As | Se |
| Cd | In | Sn | Sb | Te |

Вещества, которые расположены в колонке левее кремния и германия (элементы 3 группы Периодической системы элементов) – бор (B), алюминий (Al), галлий (Ga), индий (In) имеют такое распределение электронов на оболочках, при котором у них всегда на внешней орбите находится 3 валентных электрона. Вещества, которые расположены в колонке правее кремния и германия (элементы 5 группы) - фосфор (P), мышьяк (As), сурьма (Sb) имеют такое распределение электронов на оболочках, при котором у них всегда на внешней орбите находится 5 валентных электронов. Введение в «чистый» полупроводник элементов 5 группы, например, фосфора, приводит к тому, что атом примеси замещает атом полупроводника (рис. 1.4).

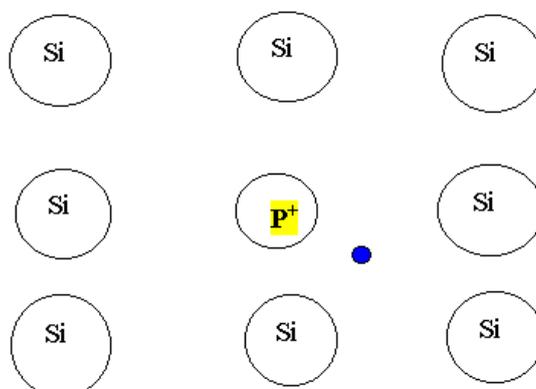


Рис. 1.4. Упрощенное плоскостное расположение атомов кристаллической решетки кремния с легированием донорной примесью

При этом четыре валентных электрона образуют с соседними четырьмя атомами кремния ковалентные связи, а пятый, не найдя себе «места», покидает атом примеси и становится «свободным». При этом примесь ионизируется (атом примеси становится положительно заряженным ионом), а в полупроводнике, даже при малых температурах, преобладает электронная проводимость. Легирующую примесь, которая «отдает» электроны во внутреннее пространство полупроводника, называют донорной.

Введение в «чистый» полупроводник элементов 3 группы, например, индия, приводит к тому, что атом примеси замещает атом полупроводника (рис. 1.5). Он, в дополнение к своим трем валентным электронам, захватывая один из валентных электронов кремния, образует ковалентные связи с ближайшими атомами кремния и становится отрицательно заряженным ионом. В покинутой связи остается положительно заряженная дырка проводимости, свободно перемещающаяся внутри кристалла. Это также меняет соотношение между концентрациями дырок и электронов по сравнению с собственным полупроводником. Дырок становится больше.

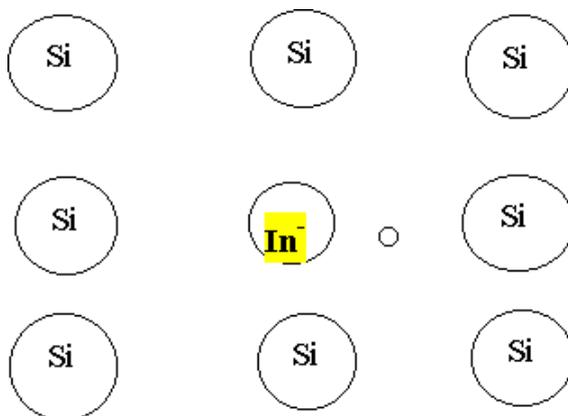


Рис. 1.5. Упрощенное плоскостное расположение атомов кристаллической решетки кремния с легированием акцепторной примесью

Для создания свободных подвижных носителей заряда в AsGa также используется легирование примесными атомами. Однако, следует иметь в виду при этом, что акцепторные свойства проявляют элементы II группы таблицы Менделеева, а донорные – VI группы (см. табл. 1.4). Наиболее распространенным акцептором является бериллий (Be); могут использоваться также магний (Mg), цинк (Zn) и кадмий (Cd). В качестве

донорных примесей могут быть использованы сера (S), селен (Se), теллур (Te). Проявляет донорные свойства и кремний, поэтому его тоже широко используют в качестве донорной примеси.

В квантовой физике доказывается, что различного рода дефекты кристаллической решетки, а также появление в узлах «нестандартного» атома примеси приводит к существенному изменению энергетического спектра системы микрочастиц в кристалле. На энергетической диаграмме это проявляется в появлении новых уровней энергии, расположенных в запрещенной зоне. Эти уровни являются следствием существенного изменения свойств полупроводников и механизмов электропроводности.

Энергетические уровни донорной примеси расположены очень близко к «дну» зоны проводимости (рис. 1.6). Количество таких уровней очень велико, однако, как и ранее, в целях упрощения энергетической диаграммы на рисунке показаны лишь отдельные уровни, позволяющие мысленно представить физическую картину происходящих процессов. Типичное значение разницы энергий между наиболее высоко расположенным донорным уровнем и дном зоны проводимости W_D имеет порядок 10^{-2} эВ.

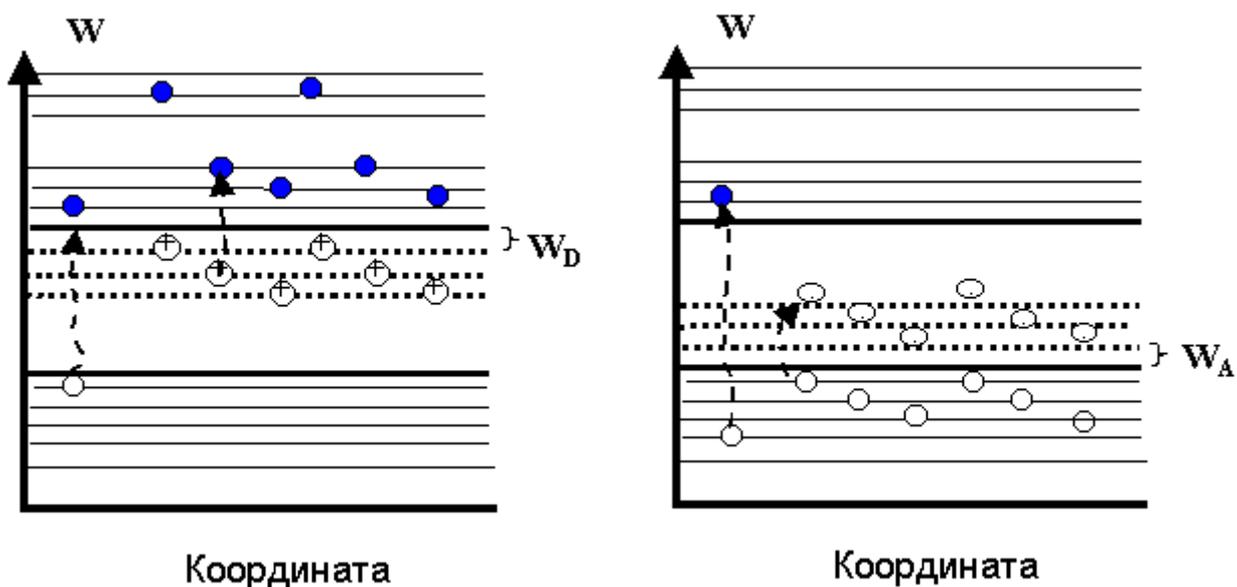


Рис. 1.6. Энергетическая диаграмма примесного полупроводника при легировании донорной (слева) и акцепторной (справа) примесью ($T > 0^\circ \text{ K}$)

При почти нулевой температуре электроны донорных примесей могут легко получить от тепловых фононов такую энергию (энергию ионизации $\sim 10^{-2}$ эВ), чтобы перейти в зону проводимости. В тоже время, чтобы возбудить собственный электрон кремния, находящийся в зоне валентности, и перевести его в зону проводимости, необходимо передать ему порцию энергии, по меньшей мере, 1,2 эВ. Не трудно сообразить, что при комнатной температуре все виртуальные квазичастицы, представленные на энергетической диаграмме на донорных уровнях, «переходят» на уровни зоны проводимости, в то время как из валентной зоны – лишь отдельные. На рис. 1.6 это условно показано таким образом: в зоне проводимости появляется 6 виртуальных квазичастиц, представляющих «свободные» электроны донорных уровней и лишь одна квазичастица, отображающая электрон с уровня валентной зоны, которая показана в левой части диаграммы. Уровень Ферми полупроводника с донорной примесью, который при заданной температуре занят электронами с вероятностью 0,5, располагается между самым верхним уровнем примеси и «дном» зоны проводимости.

Итак, в полупроводнике, легированном донорной примесью в зону проводимости попадают электроны с уровней примеси и основными носителями заряда являются электроны.

Если в запрещенной зоне вблизи верхнего края валентной зоны («потолка» валентной зоны) появляются уровни акцепторного типа (рис. 1.6), то при малейшем возбуждении (энергия ионизации $W_A \sim 10^{-2}$ эВ) на них переходят виртуальные квазичастицы, представлявшие на энергетической диаграмме электроны из валентной зоны. Только лишь отдельные квазичастицы могут переходить на уровни зоны проводимости. При комнатной температуре на все акцепторные уровни «переходят» квазичастицы - электроны с валентной зоны. В полупроводнике, легированном акцепторной примесью основными носителями заряда являются дырки. Уровень Ферми полупроводника с акцепторной примесью, который при заданной температуре занят электронами с вероятностью 0,5, располагается между самым нижним уровнем примеси и «потолком» валентной зоны.

Опираясь на сказанное, отметим еще ряд важных для электроники положений.

В полупроводнике, легированном донорной примесью, который чаще называют *примесным полупроводником n - типа*, концентрация «свободных» электронов значительно больше концентрации дырок $n_n \gg p_n$. Электронов больше и их называют основными носителями заряда, дырки – неосновными.

В полупроводнике, легированном акцепторной примесью, который чаще называют *примесным полупроводником p - типа*, наоборот, концентрация «свободных» дырок значительно больше концентрации электронов $p_p \gg n_p$. Дырок больше и их называют основными носителями заряда, электроны – неосновными.

Напомним еще раз об используемой в формулах подобного рода системе обозначений. Символами n и p обозначают концентрацию электронов и дырок соответственно. Нижний индекс p или n отражает тип полупроводника, в котором определена концентрация.

Вместе с тем, в полупроводниках n – типа и p - типа всегда выполняется условие электронейтральности. С учетом того, что в общем случае в полупроводнике могут быть примеси обоих типов и что при комнатной температуре все доноры примеси с концентрацией N_D (количество атомов доноров в одном кубическом сантиметре) и акцепторы с концентрацией N_A (количество атомов акцепторов в одном кубическом сантиметре) полностью ионизированы, можно записать:

$$p_p + n_n + N_D^+ + N_A^- = 0. \quad (1.5)$$

Это условие устанавливает, что сумма всех зарядов отрицательных (электронов, отрицательных ионов акцепторов) и положительных (дырок и положительных ионов доноров) равна нулю.

Закон действующих масс для примесных полупроводников будет звучать в такой формулировке: произведение концентрации основных носителей на концентрацию неосновных носителей есть величина постоянная и равная квадрату концентрации собственных носителей.

Для акцепторного полупроводника (p - типа):

$$p_p n_p = n_i^2. \quad (1.6)$$

Для донорного полупроводника (n – типа):

$$n_n p_n = n_i^2. \quad (1.7)$$

Часто, в практически важных случаях, можно считать, что концентрация собственных носителей существенно меньше концентрации ионизированной примеси. Это позволяет записать, из (1.7), выражение для определения концентрации неосновных носителей для полупроводника n – типа:

$$n_n \cong N_D, p_n = \frac{n_i^2}{N_D}. \quad (1.8)$$

Из (1.6) вытекает формула для определения концентрации неосновных носителей для полупроводника p – типа:

$$p_p \cong N_A, n_p = \frac{n_i^2}{N_A}. \quad (1.9)$$

Следует запомнить, что чем больше в полупроводник вводится примеси (чем больше концентрация примеси, и, соответственно, концентрация основных носителей заряда), тем меньше у него концентрация неосновных носителей заряда.

Кроме того, формулы (1.6, 1.7) свидетельствуют о том, что концентрация неосновных носителей, как и собственная концентрация, о чем говорилось ранее, очень сильно зависит от температуры полупроводника. Концентрация же основных носителей при комнатных температурах определяется главным образом концентрацией введенной в полупроводник примеси.

Уже отмечалось, что в случае, когда на полупроводник не действуют иные факторы, кроме температуры, носители заряда, имеющиеся в полупроводнике движутся хаотически, беспорядочно по кристаллу и не создают тока как, по определению, упорядоченного движения заряда.

На полупроводник могут воздействовать и внешние факторы. В первую очередь, к бруску полупроводника может быть приложено электрическое напряжение (разность потенциалов между двумя противоположными стенками). Следствием этого будет выход полупроводника из состояния термодинамического равновесия. Движение электронов в кристалле будет складываться из беспорядочного теплового движения и упорядоченного движения под действием внешнего электрического поля. В результате будет происходить перемещение всей совокупности «свободных» электронов с некоторой средней скоростью. Направленное движение совокупности свободных электронов во внешнем электрическом поле носит название дрейфа. Скорость направленного движения в этом случае называется *дрейфовой*.

Плотность дрейфового тока (заряда электронов $q = 1,6 \cdot 10^{-19}$ Кл с концентрацией N , проходящих за одну секунду через поперечное сечение площадью 1 см^2) определяется формулой

$$\vec{J} = qN\vec{V}_{др} \quad (1.10)$$

Экспериментально установлено, что дрейфовая скорость, при слабых электрических полях ($E < 3 \cdot 10^3$ В/см), прямо пропорциональна напряженности. А затем, в сильных полях, дрейфовая скорость перестает расти и стремится к насыщению. Дрейфовая скорость насыщения составит $\sim 10^7$ см/с.

Мы, главным образом, будем рассматривать полупроводники со слабыми полями. В таких ситуациях подвижность от величины напряженности электрического поля не зависит, а *дрейфовая скорость* будет равна

$$\vec{V}_{др} = \mu\vec{E}. \quad (1.11)$$

Коэффициент пропорциональности μ между скоростью дрейфа и напряженностью электрического поля называется *подвижностью*.

Подвижность характеризует скорость дрейфа, приобретаемую свободными носителями в электрическом поле единичной напряженно-

сти (1 В/см). Подвижность имеет разные значения для электронов и дырок. Величины подвижностей приведены в таблице 1.5.

Таблица 1.5

Подвижности носителей заряда полупроводников

| Полупроводник | Крем- ний | Арсенид галия | Герма- ний |
|--|--------------|------------------|---------------|
| Подвижность электронов, см ² /(В с) | 1500 | 8500 | 3900 |
| Подвижность дырок, см ² /(В с) | 450 | 400 | 1900 |

Для области слабых полей, зная подвижность носителей заряда, из (1.10), можно записать выражения для *плотности электрического тока* электронов и дырок:

$$\vec{J}_{n,\partial p} = qn\vec{V}_{\partial p} = qn\mu_n\vec{E}, \quad \vec{J}_{p,\partial p} = qp\vec{V}_{\partial p} = qp\mu_p\vec{E}. \quad (1.12)$$

Суммарная плотность электронов и дырок:

$$\vec{J}_{n,p,\partial p} = q(n\mu_n + p\mu_p)\vec{E} = \sigma\vec{E} = \frac{1}{\rho}\vec{E}. \quad (1.13)$$

Последнее выражение представляет собой *дифференциальную форму закона Ома*. Величина $\sigma = q(n\mu_n + p\mu_p)$ называется *удельной проводимостью*, а обратная ей величина ρ - *удельным сопротивлением* при величинах концентраций электронов n и дырок p .

Как увидим в дальнейшем, типичной для полупроводника является ситуация, когда концентрация носителей в одном месте бруска полупроводника отличается от концентрации в другом месте. Другими словами, имеет место неравномерное по пространству распределение носителей заряда в объеме полупроводника или, как говорят, градиент (разность) концентрации носителей заряда по пространству. Как известно из физики, в таких случаях возникает, связанная с тепловым движением носителей заряда, диффузия. Соответственно, возникает направленное, упо-

рядоченное диффузионное движение носителей заряда, связанное с неравномерным распределением носителей по пространству, и, как следствие, так называемый *диффузионный ток*.

Теоретической основой диффузии является *закон Фика*, в соответствии с которым плотность потока свободных носителей пропорциональна градиенту концентрации, взятому с обратным знаком, поскольку диффузионный поток направлен в сторону меньшей концентрации носителей. Для одномерного случая (по координате X), умножая плотность потока на заряд, можно получить *плотность диффузионного тока* для электронов и дырок:

$$\vec{J}_{n, \text{диф}} = qD_n \frac{dn}{dx}, \quad \vec{J}_{p, \text{диф}} = -qD_p \frac{dp}{dx}. \quad (1.14)$$

Здесь D_n, D_p - коэффициенты диффузии электронов и дырок, соответственно.

Параметры дрейфового и диффузионного движения связаны между собой соотношением Эйнштейна, которое имеет следующий физический смысл: независимо от причины, вызвавшей направленное движение, свободные носители встречают на своем пути одни и те же неоднородности, от взаимодействия с которыми зависит их *скорость* и, соответственно, *ток*:

$$D_n = \left(\frac{kT}{q}\right) \mu_n = \varphi_T \mu_n, \quad (1.13)$$

$$D_p = \left(\frac{kT}{q}\right) \mu_p = \varphi_T \mu_p. \quad (1.14)$$

Коэффициент пропорциональности:

$$\varphi_T = \left(\frac{kT}{q}\right) \quad (1.15)$$

имеет размерность потенциала (вольт) и называется *тепловым потенциалом*. Следует запомнить, что при комнатной температуре ($T = 300 \text{ K}$) $\varphi_T = 0,026 \text{ В} = 26 \text{ мВ}$.

1.2. Структуры полупроводниковой электроники

В полупроводниковой электронике однородные материалы для создания каких либо изделий применяются редко. В частности, как увидим в дальнейшем, в микроэлектронике однородные полупроводниковые слои находят использование для создания полупроводниковых резисторов. Большинство же элементов представляют собой систему разнородных тел, именуемую структурой.

Структура полупроводникового изделия – система граничащих и находящихся в контакте твердых тел, различающихся по типу электропроводности, по количеству введенной примеси или по величине объемного удельного электрического сопротивления, обеспечивающая выполнение изделием его функции.

Применяется много вариантов контактов разнородных материалов. В простейших случаях различают структуры:

Металл 1 – металл 2;

Полупроводник – диэлектрик;

Металл – диэлектрик – полупроводник;

Полупроводник 1 – переходный слой – полупроводник 2;

Полупроводник 1 – переходный слой – металл.

В структуре полупроводникового изделия при контакте двух разнородных тел имеется *переходный слой (электрический переход)*, который играет важную роль. В зависимости от этого слоя, в свою очередь, структуры «полупроводник 1 – переходный слой – полупроводник 2» могут быть:

а) *гомotypными*, когда для образования структуры и $p - n$ перехода в качестве полупроводников 1 и 2 используются одни и те же полупроводники (с одной шириной запрещенной зоны), но с разным типом примеси;

б) *гетерotypными*, когда для образования структуры и гетероперехода применяются разные полупроводниковые материалы, с различной шириной запрещенной зоны (Ge – Si, Ge – GaAs и т. п.).

Кроме того, гомотипные структуры, в зависимости от концентрации примеси в p - n областях могут быть симметричными и несимметричными. Если концентрация легирующих примесей и, соответственно, концентрация основных носителей заряда в контактирующих полупроводниках примерно одинакова, то такая структура называется *симметричной*; в противном случае – *несимметричной*. Подгруппы полупроводников, которые могут использоваться в несимметричных переходах, их условное обозначение и основные параметры (для Si) представлены в таблице 1.6. Несимметричные переходы обозначают $p^{++} - n$, $p - n^+$, $p^+ - n$ и т. д.

Таблица 1.6

Обозначения полупроводников в зависимости от концентрации примеси

| Название подгруппы используемого полупроводника | Условное обозначение | Концентрация введенной примеси, $1/\text{см}^3$ | Удельное сопротивление, Ом/см |
|---|----------------------|---|-------------------------------|
| Собственный (беспримесный) | p_i, n_i | $\sim 0 (< 10^{-6})$ | $\sim 2 \cdot 10^5$ |
| Слабелегированный | p, n | $\sim 10^{13} - 10^{14}$ | $\sim 2 \cdot 10^5$ |
| Умеренно легированный | p^+, n^+ | $\sim 10^{15} - 10^{16}$ | 100...10 |
| Сильно легированный | p^{++}, n^{++} | $\sim 10^{17} - 10^{18}$ | 0,1...0,001 |
| Вырожденный | p^{+++}, n^{+++} | $\sim 10^{19} - 10^{22}$ | $10^{-2} \dots 10^{-4}$ |

Неотъемлемым элементом любого, без исключения, полупроводникового прибора являются структуры полупроводник – переходный слой – металл. В зависимости от свойств, образующегося электрического перехода между металлом и полупроводником различают:

- а) *омические переходы* металл - полупроводник;
- б) *выпрямляющие переходы* металл – полупроводник.

Омические переходы металл - полупроводник имеют малое электрическое сопротивление и, в заданном рабочем диапазоне токов, его величина практически не зависит от направления тока. Такие переходы

используют для создания токоподводящих и токоотводящих электродов (выводов).

Выпрямляющие переходы металл – полупроводник имеют резко несимметричную вольт-амперную характеристику. Такие переходы используют для создания переходов и диодов Шоттки на их основе.

Следует иметь в виду, что, как правило, омические переходы металл - полупроводник не показывают, считая их наличие само собой разумеющимся.

1.3. Электронно-дырочная структура и ее свойства

1.3.1. Образование и равновесное состояние электронно-дырочного перехода

Предположим, что мы располагаем образцом монокристалла полупроводника, к примеру, кремния, полученного («выращенного») таким образом, что относительная доля примесных атомов в нем ничтожно мала, а его кристаллическая решетка идеальна и не содержит дефектов. Можно считать, что данный кристалл является собственным полупроводником. Из этого монокристалла вырезается брусок, в который, скажем пока так, специальными технологическими методами внедряются донорная (с концентрацией N_D) и акцепторная (с концентрацией N_A) примеси. Условное изображение такого бруска на плоскости (проекция на плоскость, одномерная модель) в тот момент, когда было закончено введение примеси, показано на рис. 1.7а. При этом необходимое равновесное распределение зарядов еще не сформировалось. Брусок мысленно разобьем на пространственные p и n области, которые разделяет воображаемая плоскость (на рисунке линия), называемая металлургической границей. Будем полагать также, что рассматривается случай термодинамического равновесия, при котором на брусок воздействует тепловая энергия (пусть комнатной, все время неизменной, температуры), а другие внешние воздействия, к примеру, напряжение, свет, радиация отсутствуют.

С момента окончания введения примесей в состоянии термодинамического равновесия в $p - n$ структуре начнут происходить следующие процессы.

А. Поскольку концентрация электронов, как основных носителей заряда в n области, значительно больше концентрации электронов, как неосновных носителей заряда, в p области ($n_n \gg n_p$), то в соответствии с законом Фика, возникнет явление диффузии электронов из n области в p область и, как следствие, появится упорядоченное движение носителей заряда: диффузионный электрический ток, обусловленный движением электронов (рис. 1.7.б). Этот ток будем характеризовать вектором плотности электрического тока $\vec{J}_{n, \text{диф}}$. Следует обратить внимание на то, что направление движения электронов и направление вектора плотности электрического тока противоположны, поскольку за направление протекания тока принято движение положительных носителей заряда.

Аналогично, из-за разности концентраций между основными носителями и неосновными будет наблюдаться перемещение дырок из p области в n область (рис. 1.7.в). Соответственно, появится упорядоченное движение дырок и диффузионный электрический ток. Этот ток будем характеризовать вектором плотности электрического тока $\vec{J}_{p, \text{диф}}$.

Б. Через некоторое время в n области окажутся избыточные дырки, а в p области – избыточные электроны. Большая часть избыточных зарядов будет находиться вблизи металлургической границы. Как уже отмечалось, по истечении «временем жизни» будет происходить рекомбинация (воссоединение электрона и дырки) пришедших зарядов и пары электрон – дырка будут интенсивно исчезать. Вблизи металлургической границы полупроводник перестанет быть электронейтральным. С одной стороны, это обусловлено тем, что вследствие рекомбинации электронов будут, как говорят, «обнажаться» некомпенсированные заряды акцепторных ионов. С другой стороны, из-за ухода электронов в n области «обнажаться» некомпенсированные положительные заряды донорных ионов. Об аналогичных процессах можно говорить и для дырок. Следует учитывать, что ионы неподвижны, поэтому по обе стороны от металлургической границы образуется слой с пониженной концентрацией подвижных носителей заряда. Этот слой называют обедненным (подвижными носителями заряда). Так как этот слой (из-за отсутствия подвижных носителей заряда) обладает большим сопротивлением, то его еще называют запирающим. Используют также название слой «объемного заряда».

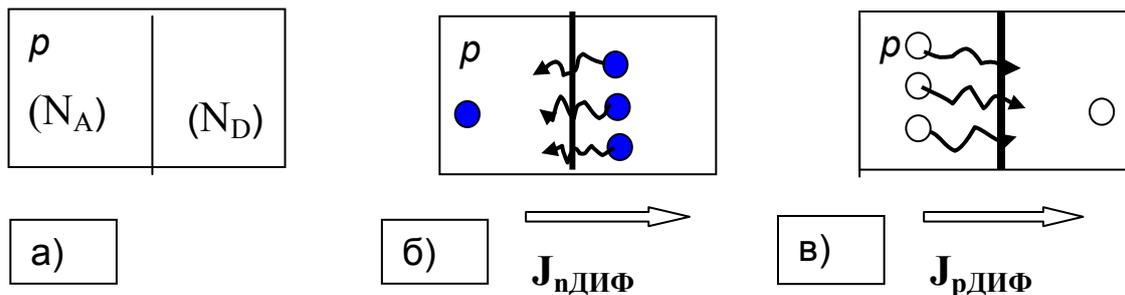


Рис. 1.7. Условное изображение потоков носителей заряда и протекающих при этом токов в начальной стадии образования перехода

В результате этих процессов, из-за дисбаланса зарядов вблизи металлургической границы, появляется две разноименно заряженных области объемного заряда ионов примесей (рис. 1.8). Как известно из физики, любая система неподвижных заряженных тел с отличным от нуля суммарным зарядом, расположенным в конечной области пространства, может рассматриваться как локальное электростатическое электрическое поле. Поэтому вблизи металлургической границы появится внутреннее электрическое поле \vec{E} (рис. 1.8). Это поле, как известно, обуславливает появление напряжения или, как еще говорят, *контактной разности потенциалов* между p и n областями.

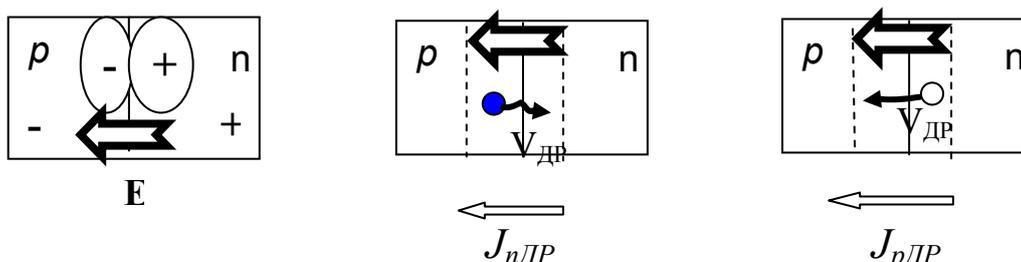


Рис. 1.8. Условное изображение потоков носителей заряда и протекающих при этом токов в второй стадии образования перехода

В. Неосновные носители заряда, находящиеся в обедненном слое, под воздействием внутреннего поля начинают дрейфовое перемещение со скоростью $\vec{V}_{др}$. Если вблизи металлургической границы оказываются электрон p области или дырка n области, то создаваемое системой неподвижных относительно наблюдателя разноименно заряженных ионов

примеси электростатическое поле \vec{E} , создает дрейфовое движение носителей заряда и, соответственно, дрейфовый ток электронов $\vec{J}_{n,\partial p}$ и дырок $\vec{J}_{p,\partial p}$ (рис. 1.8).

Г. В обедненном слое $p - n$ структуры все потоки носителей заряда взаимосвязаны. Вектор напряженности электрического поля направлен таким образом, что он препятствует движению основных носителей заряда, поэтому диффузионные токи и, соответственно, объемные заряды ионов примеси, ограничиваясь электрическим полем, не могут все время увеличиваться. Напряженность внутреннего электрического поля, в свою очередь, также нарастает до определенной величины, которая зависит от степени диффузии и рекомбинации. Поэтому, в конце концов, при отсутствии внешнего напряжения и постоянной температуре в приконтактной области устанавливается режим динамического равновесия, при котором напряженность внутреннего электрического поля нарастает до тех пор, пока вызванное им дрейфовое движение неосновных носителей заряда не уравновесит встречное диффузионное движение, обусловленное разностью концентраций (градиентом концентрации) электронов и дырок. Другими словами, в режиме динамического равновесия количество основных носителей заряда, проходящих через плоскость контакта, будет равно количеству неосновных носителей заряда, проходящих через ту же плоскость. Сказанное, для единичной площадки плоскости между p и n полупроводниками, можно выразить следующей формулой равенства токов:

$$\vec{J}_{n,диф} + \vec{J}_{p,диф} = \vec{J}_{n,\partial p} + \vec{J}_{p,\partial p} \Leftrightarrow \vec{J}_{диф} = \vec{J}_{\partial p}. \quad (1.18)$$

Таким образом, при соединении двух полупроводников с различным типом электропроводности, образовавшуюся структуру можно представить состоящей из электронейтральных $p - n$ областей и переходного слоя объемного заряда между ними (шириной $L_{об}$), находящегося в состоянии динамического равновесия (рис. 1.9).

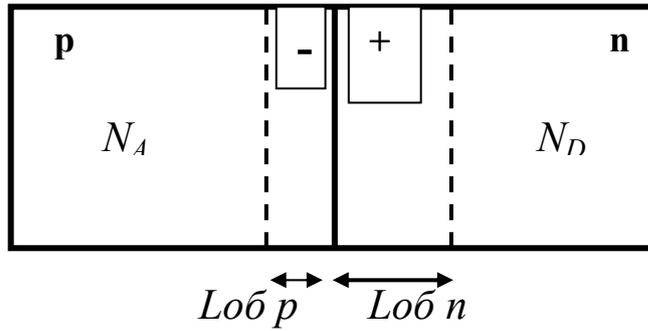


Рис. 1.9. Электронно – дырочная структура в состоянии равновесия

Следует понимать, что, хотя при рассуждениях внешние металлические выводы нигде не использовались и нигде не изображались, ток через внешние выводы не протекал бы и в том случае, если бы они и были. Токи протекают только внутри обедненного слоя.

Если считать, что в p области расположена часть обедненного слоя толщиной $L_{об p}$, которая содержит отрицательный заряд ионов акцепторов, а в n области – толщиной $L_{об n}$ и положительные ионы доноров, то полная толщина обедненного слоя $L_{об} = L_{об p} + L_{об n}$. Обедненный слой в целом нейтрален: отрицательный заряд в p области $Q_{об}^- = qN_A L_{об p} S$ равен положительному заряду в $Q_{об}^+ = qN_D L_{об n} S$, где S – площадь перехода. Из равенства зарядов следует

$$L_{об n} / L_{об p} = N_A / N_D. \quad (1.19)$$

Для несимметричного перехода при $N_A \gg N_D$ получаем $L_{об n} \gg L_{об p}$ и $L_{об} = L_{об n}$. Следовательно, обедненный слой сосредоточен в области электронно - дырочной структуры с меньшей концентрацией введенной примеси, именуемой базой.

Рассмотрим теперь, что будет представлять собой энергетическая диаграмма структуры, состоящей из двух полупроводников с различным типом электропроводности.

В момент образования структуры (как бы при соприкосновении полупроводников) с различным характером пространственного легирования монокристалл, из-за асимметрии энергий в p и n полупроводниках, о чем свидетельствует различное положение уровней Ферми, представляет собой неустойчивую систему микрочастиц, которая стремится дос-

тичь состояния энергетического равновесия. Достичь такого состояния внутри замкнутой системы микрочастиц нельзя иным способом, чем создать внутреннее электрическое поле, которое бы, меняя структуру волновых полей в монокристалле, выравнивало бы уровни Ферми, сдвигая исходные диаграммы по вертикали друг относительно друга.

Как следует из рис. 1.6, уровни энергии различных полупроводников, существенно различаются степенью «заселенности» квазичастицами: дырками и электронами зон проводимости и валентности. При этом уровень Ферми p полупроводника (располагается между уровнями акцепторной примеси и потолком валентной зоны) находится ниже уровня Ферми n полупроводника (располагается между уровнями донорной примеси и дном зоны проводимости). В начальный момент «стыковки» энергетических диаграмм полупроводников концы линий одинаковых уровней соединяются. Соединение в одну линию уровней с различной «заселенностью виртуальными квазичастицами» дает возможность основным носителям заряда одного полупроводника «заселять» свободные уровни другого полупроводника. Если судить по рисунку 1.6 и мысленно представить, что линии уровней соединены, у электронов n полупроводника появляется возможность «переходить» на свободные уровни зоны проводимости p полупроводника. И, наоборот, у дырок p полупроводника имеется возможность переходить на свободные уровни валентной зоны n полупроводника.

Подобное «перемещение» микрочастиц по уровням энергетической диаграммы свидетельствует о том, что в системе имеют место процессы переноса энергии. Это «перемещение» может быть истолковано также как протекание тока. По аналогии с тем, что мы ранее рассмотрели, этому току можно придать смысл диффузионного тока, состоящего из электронной и дырочной компоненты.

Из-за того, что электромагнитные волны в монокристалле переносят энергию в направлениях, определяемых движением микрочастиц, меняется структура поля: энергетические уровни полупроводника p типа поднимаются вверх, а n типа – опускаются вниз (представьте с помощью энергетических диаграмм этот процесс пока мысленно). Линии уровней при этом приобретают вид кусочно - ломаных линий (рис. 1.10). Подобное изменение положения испытывают и уровни Ферми (показаны на рис. 1.10 штриховыми толстыми линиями).

Происходящее «искривление» уровней, получающееся в местах «стыковки» исходных энергетических диаграмм, свидетельствует о том, что в монокристалле, вблизи металлургической границы, начинает формироваться внутреннее электрическое поле. Очевидно, что чем интенсивнее идут процессы обмена энергией и чем большей величины протекает ток, тем больше внутреннее электрическое поле, что находит свое образное выражение в более сильном «изгибе» уровней.

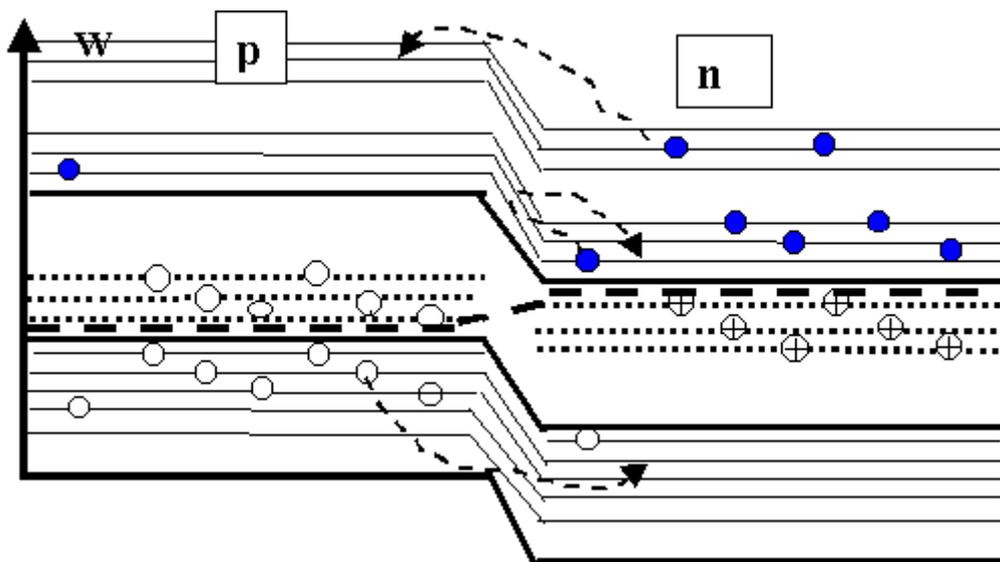


Рис. 1.10. Энергетическая диаграмма электронно - дырочной структуры в стадии установления равновесия

Появление «изгиба» уровней между зонами на пути движущихся на свободные уровни квазичастиц свидетельствует о появлении «препятствия»: для электронов и дырок, создающих диффузионный ток, в виде «горки», некоего «барьера», который надо преодолеть. Напомним, что для дырок на энергетических диаграммах используется несколько иной способ отсчета энергий: дырка, переходя на более низкий уровень валентной зоны, обретает большую энергию. Чтобы преодолеть эти «горку», «барьер», электрону, «перемещающемуся» по уровню уже требуется иметь определенный запас энергии, чтобы «взобраться на горку». Если учитывать несколько иной способ отсчета энергии, то можно сказать, то же самое, и о дырке. Так как не все электроны и дырки обладают необходимым для преодоления барьера запасом энергии, то часть из них будет «возвращаться назад». Можно утверждать, что появление

внутреннего электрического поля («изгиба» уровней) приводит к уменьшению количества квазичастиц, переносящих энергию, и к уменьшению диффузионного тока.

Рассмотренная упрощенная картина процессов имеет и «волновой» физический смысл. Электромагнитные волны квазичастиц, взаимодействуя с внутренним электрическим полем, либо теряют часть своей энергии, либо как бы «отражаются» от барьера. Хорошо известна подобная аналогия, касающаяся бегущих и стоячих волн. Поэтому электромагнитные волны одних квазичастиц, имеющих структуру типа «бегущей» волны и достаточную энергию, превосходящую, как говорят, высоту потенциального барьера, «проходят» сквозь барьер, перенося энергию. Другие же волны квазичастиц «отражаются» от барьера, создавая «стоячие» волны, не переносящие энергию.

Если посмотреть на перемещение квазичастиц, то несложно заметить следующее. Электроны, располагающиеся в зоне проводимости полупроводника типа p (на рис. 1.10 показан слева в виде одного «неосновного» электрона, попавшего при «перебрасывании» в зону проводимости за счет процессов термогенерации в собственном полупроводнике), не имеют потенциального барьера. Даже, наоборот, им не надо «взбираться на горку», они могут «скатываться» с нее. Такие квазичастицы начнут «перемещаться» из полупроводника p - типа в полупроводник n – типа. Аналогично можно говорить о том, что «неосновные» дырки полупроводника n – типа начнут «всплывать», перемещаясь из полупроводника n – типа в полупроводник p – типа. Чисто формально подобное перемещение квазичастиц и перенос ими энергии можно толковать как протекание тока. Если воспользоваться аналогией, то этому току можно придать физический смысл дрейфового тока, состоящего из дырочной и электронной компоненты.

В конце концов, отмеченные выше процессы обязательно, самопроизвольно будут эволюционизировать (развиваться) таким образом, что придут в состояние динамического равновесия. В монокристалле полупроводника установится только для него характерный режим комбинированных волн.

Покажем, для примера, как будет осуществляться регулирование потоков квазичастиц, пересекающих плоскость контакта двух полупроводников.

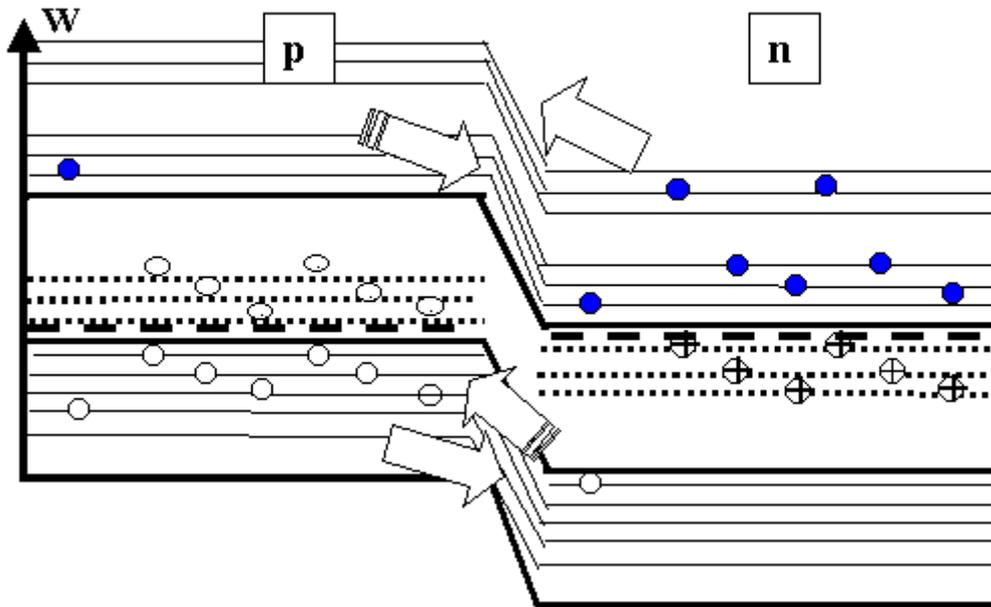


Рис. 1.11. Энергетическая диаграмма электронно - дырочной структуры в стадии окончания процесса установления состояния равновесия

Предположим вначале, что потоки квазичастиц, создающих диффузионный ток, превышают встречные потоки квазичастиц, создающих дрейфовый ток. В этом случае, как уже отмечалось, будет расти напряженность внутреннего электрического поля (увеличиваться изгиб уровней). Это обязательно приведет к тому, что будет расти потенциальный барьер и уменьшаться количество квазичастиц, его преодолевающих (уменьшаться диффузионный ток). С другой стороны, увеличение изгиба уровней можно толковать как увеличение потока квазичастиц, создающих дрейфовый ток. В результате, тенденция уменьшения диффузионного тока и роста дрейфового тока будет сохраняться до тех пор, пока, грубо говоря, потоки через границу не уравниваются и не наступит равновесное состояние.

Если же по какой либо причине поток квазичастиц, создающий дрейфовый ток, превысит поток, создающий диффузионный ток, то уменьшится напряженность внутреннего электрического поля. Вследствие этого увеличится диффузионный поток и опять наступит состояние динамического равновесия.

Окончание процесса установления состояния динамического равновесия отражает энергетическая диаграмма, показанная на рис. 1.11. На ней различными стрелками показаны потоки, характеризующие все компоненты диффузионного и дрейфового тока. Положение уровней Ферми таково, что они располагаются на одной линии.

1.3.2. Параметры, характеризующие электронно-дырочную структуру (идеальный диод) в состоянии термодинамического равновесия и при отсутствии напряжения

Математическая модель электронно-дырочной структуры (идеальный диод) строится на основе следствий, вытекающих из предположений о термодинамическом равновесии.

Поскольку внутри обедненного слоя есть внутреннее электрическое поле E , то можно говорить о том, что между границами обедненных областей имеет место контактная разность потенциалов (рис. 1.12):

$$\varphi_k = \varphi_n - \varphi_p. \quad (1.20)$$

Для любой поверхности, расположенной в области объемного заряда, суммарные токи электронов и дырок равны, поэтому на основании (1.12, 1.14, 1.18) можем записать:

$$-qD_p \frac{dp}{dx} = qp\mu_p \vec{E}. \quad (1.21)$$

Так как E есть производная от разности потенциалов с другим знаком, то, решая дифференциальное уравнение:

$$qD_p \frac{dp}{dx} = qp\mu_p \frac{d\varphi}{dx}, \quad (1.22)$$

можно получить выражение, определяющее контактную разность потенциалов на $p - n$ переходе:

$$\varphi_k = \varphi_n - \varphi_p = \varphi_T \ln \frac{p_p n_n}{n_i^2}. \quad (1.23)$$

Контактная разность потенциалов есть произведение теплового потенциала $\varphi_T = \left(\frac{kT}{q}\right)$ на натуральный логарифм произведения концентраций основных и неосновных носителей заряда, деленного на квадрат собственной концентрации. Для кремния: $\varphi_K^{Si} = (0,6 - 0,8)B$.

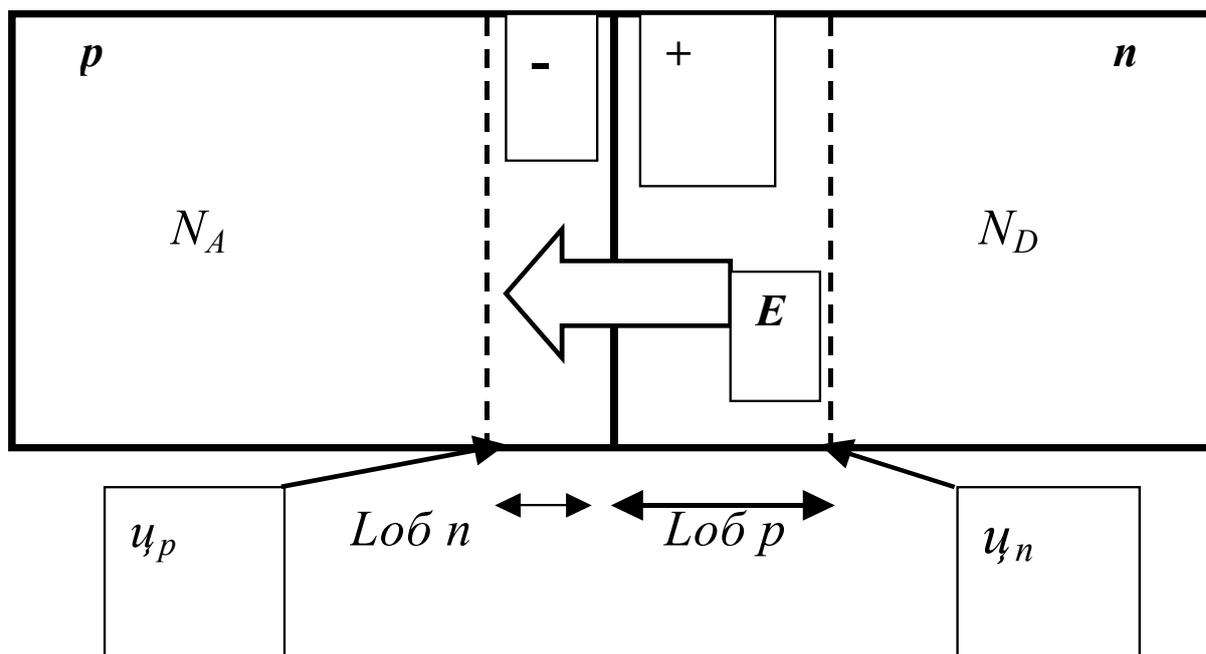


Рис. 1.12. Условное изображение электронно - дырочной структуры и параметров ее характеризующих

Помножив величину φ_K на величину заряда q и используя (1.3) получаем величину потенциального барьера, образующегося на энергетической диаграмме $p - n$ структуры:

$$\varphi_K q = \Delta W - kT \ln \frac{N_{E(V)}^2}{n_n p_p}, \quad (1.24)$$

где N_E (N_V) – эффективная плотность состояний в зоне проводимости (в валентной зоне), значение которых примерно равны $2 \times 10^{-19} \text{ 1/см}^{-3}$;

$k = 1,38 \times 10^{-23}$ Дж/К – постоянная Больцмана; ΔW – ширина запрещенной зоны.

Можно численными расчетами показать, что величина потенциального барьера зависит, главным образом, от величины запрещенной зоны полупроводника.

Решая дифференциальное уравнение Пуассона для $p - n$ структуры, можно найти *полную ширину обедненного слоя*:

$$L_{об} = L_{об p} + L_{об n} = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_K}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}. \quad (1.25)$$

Обедненный слой несимметричного перехода, как отмечалось уже, расширяется в сторону меньшей концентрации и при резко несимметричном переходе практически полностью лежит в слабо легированной области полупроводника.

1.3.3. Функционирование электронно-дырочной структуры (идеального диода) при приложении к ней внешнего прямого напряжения

При приложении к электронно-дырочной структуре внешнего постоянного напряжения возможны два способа его подачи на $p - n$ переход.

1. *Прямым включением* идеального диода называется такое его включение при котором к p полупроводнику приложен положительный потенциал источника, а n полупроводнику – отрицательный.

2. *Обратным включением* идеального диода называется такое его включение при котором к n полупроводнику приложен положительный потенциал источника, а p полупроводнику – отрицательный.

Математическая модель электронно-дырочной структуры (идеального диода) предполагает, что объемное сопротивление областей полупроводника существенно меньше, чем объемное сопротивление обедненного (запирающего) слоя. По этой причине, если воспользоваться правилом простейшего делителя напряжения для трех последовательно включенных резисторов, имитирующих 3 области полупроводника (p ,

обедненный «запирающий» слой и n), можно утверждать, что почти все внешнее напряжение, приложенное к боковым торцам бруска полупроводника, падает на обедненной области.

Всякая разность потенциалов, как известно из физики, создает электрическое поле, поэтому приложенное напряжение создаст внутри обедненного слоя дополнительное электрическое поле, которое мы будем называть внешним.

Если имеет место прямое включение диода, то вектор напряженности внешнего электрического поля имеет направление, противоположное вектору напряженности внутреннего электрического поля. В результате суммарное электрическое поле внутри обедненного слоя будет становиться меньше исходного (в состоянии равновесия) внутреннего поля (рис. 1.13).

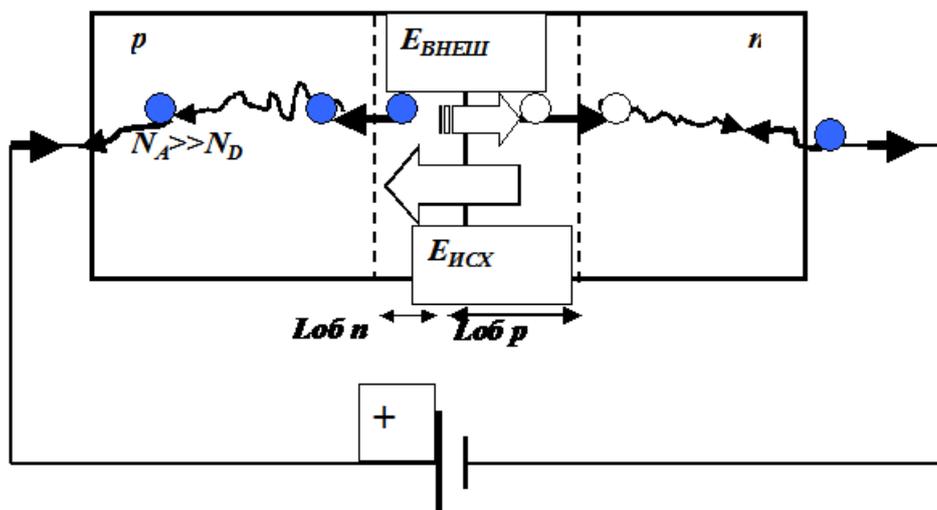


Рис. 1.13. Электронно-дырочная структура при приложении к ней внешнего положительного напряжения (прямое включение)

Если же имеет место обратное включение диода, то вектор напряженности внешнего электрического поле имеет такое же направление, как и у вектора напряженности внутреннего электрического поля. В этом случае результирующее электрическое поле внутри обедненного слоя будет становиться больше исходного (в состоянии равновесия) внутреннего поля (рис. 1.14).

При прямом включении идеального диода в нем будут происходить следующие процессы:

А. В результате суперпозиции исходного внутреннего $E_{ИСХ}$ и внешнего $E_{ВНЕШ}$ электрических полей результирующая напряженность электрического поля внутри обедненного слоя уменьшится.

Это приводит к тому, что, имевшее место в состоянии равновесия, равенство (1.18) токов нарушается. Диффузионный ток становится больше дрейфового.

Б. Следствием увеличения диффузионного тока является то, что на правой и левой границах обедненного слоя появляются дополнительные (избыточные) носители заряда (электроны слева, а дырки справа границ обедненного слоя). Происходит некий «вброс, впрыск» через границы обедненного слоя в p и n областей носителей заряда, которые были прежде основными, а стали неосновными носителями. Как говорят, на границах обедненного слоя имеет место инжекция (впрыскивание) неосновных носителей заряда (рис. 1.13).

В. Проникшие через границу обедненного слоя и ставшие теперь неосновными подвижные носители заряда, вследствие конечного времени жизни дырок и электронов, не будут сразу рекомбинировать, а продолжат свое «движение» от границы. Так как вблизи границы обедненного слоя концентрация носителей больше равновесной, а в местах подключения выводов источников концентрация их будет почти равновесной (т. е. меньше), то можно утверждать, что имеет место пространственная разность концентраций и двигаться от границы вглубь бывших прежде электронейтральных p и n областей инжектируемые носители будут вследствие диффузии. При этом по мере их движения вглубь p и n областей к выводам все большее их количество прорекомбинирует.

Г. Чтобы из-за инжекции и появления избыточных носителей заряда в областях, прилегающих к обедненному слою, p и n полупроводники по-прежнему оставались электронейтральными, вступает в действие механизм компенсации заряда, обеспечивающий протекание тока через внешние выводы. В первом приближении его работу можно пояснить следующим образом.

Чтобы p область оставалась электронейтральной из-за инжекции электронов, из p области в металлический вывод, имеющий положительный потенциал, уходит поток электронов, имеющий такой же самый заряд. Чтобы n область оставалась электронейтральной из-за инжекции дырок, из металлического вывода, имеющего отрицательный потенциал,

в n область входит поток электронов, имеющий такой же самый заряд, что и инжектированные дырки. Таким образом, через металлические внешние выводы диода протекает ток.

Из формулы (1.23), потенцируя, можно получить для правой (ПГОС) и левой (ЛГОС) границы обедненного слоя следующие выражения:

$$p_{n0}^{ПГОС} = p_{p0} \ell^{-\frac{\varphi_K}{\varphi_T}} \quad , \quad n_{p0}^{ЛГОС} = n_{n0} \ell^{-\frac{\varphi_K}{\varphi_T}} . \quad (1.26)$$

Нижний индекс 0 подчеркивает, что выражение справедливо для равновесного состояния.

При прямом включении диода, вместо φ_K , так как разность потенциалов на обедненном слое уменьшилась, следует в выражения (1.26) подставить $\varphi_K - U$.

Тогда получаем:

$$p_n^{ПГОС} (U > 0) = p_{p0} \ell^{-\frac{\varphi_K - U}{\varphi_T}} = p_{p0} \ell^{-\frac{\varphi_K}{\varphi_T}} \ell^{\frac{U}{\varphi_T}} = p_{n0}^{ПГОС} \ell^{\frac{U}{\varphi_T}} . \quad (1.27)$$

Проделав аналогичные выкладки для левой границы, можно записать выражения, получившие в литературе название *граничных условий Шокли*:

$$p_n^{ПГОС} (U > 0) = p_{n0}^{ПГОС} \ell^{\frac{U}{\varphi_T}} . \quad (1.28)$$

$$n_p^{ЛГОС} (U > 0) = n_{p0}^{ЛГОС} \ell^{\frac{U}{\varphi_T}} . \quad (1.29)$$

Граничные условия Шокли свидетельствуют о том, что по сравнению с равновесной, при положительном напряжении концентрация неосновных носителей заряда на границах обедненного слоя очень резко (экспоненциально) возрастает.

Явление резкого возрастания концентраций неосновных носителей по обе стороны обедненного слоя при приложении к электронно-дырочной структуре прямого напряжения получило название *инжекции* (впрыскивания) неосновных носителей (если быть предельно точным, то инъекции основных носителей, которые, затем, становятся неосновными).

В несимметричных электроно-дырочных структурах, типа $p^{++} - n$, имеющих $N_A \gg N_D$, в соответствии с (1.8, 1.9) $p_{n0}^{ПГОС} \gg n_{p0}^{ПГОС}$. Следовательно, как вытекает из граничных условий Шокли (1.28, 1.29), $p_n^{ПГОС}(U > 0) \gg n_p^{ПГОС}(U > 0)$, то есть, имеет место односторонняя инъекция неосновных носителей заряда (главным образом, инъекция дырок).

В электронейтральных областях, окружающих обедненный слой, имеет место диффузия избыточных носителей заряда, за счет чего через внешние выводы протекает ток. Для области n , используя (1.14) можно записать для неравновесного состояния (НС):

$$j^{ПГОС} = j_p^{ПГОС} + j_n^{ПГОС} = -qD_p \frac{dp^{НС}}{dx} + qD_n \frac{dn^{НС}}{dx}. \quad (1.30)$$

Можно показать, что (1.30) преобразуется к виду:

$$j^{ПГОС} = \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right) \left(\ell^{\frac{U}{\varphi_T}} - 1 \right), \quad (1.31)$$

где D_p, D_n – коэффициенты диффузии дырок и электронов; L_p, L_n – *диффузионная длина дырок и электронов* (расстояние проходя которое, из-за рекомбинации количество носителей уменьшается в 2,7 раза).

Помножив плотность тока на площадь S $p - n$ перехода, получаем выражение для тока, протекающего через идеальный диод:

$$I = I_0 \left(\ell^{\frac{U}{\varphi_T}} - 1 \right), \quad (1.32)$$

где $I_0 = S \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right)$ - тепловой ток.

С точки зрения энергетических диаграмм функционирование электронно-дырочной структуры при приложении к ней внешнего прямого напряжения, в первом приближении, можно пояснить следующим образом.

Если осуществить прямое включение диода, то есть к p полупроводнику приложить положительный потенциал, а к n полупроводнику – отрицательный, то состояние термодинамического равенства токов будет нарушено. Имея перед глазами рис.1.11, мысленно это можно представлять так.

У n полупроводника, подключенного к отрицательному потенциалу источника все уровни, в том числе и уровень Ферми, будут смещаться вверх на величину qU , а p полупроводника – вниз. Все это приведет к понижению энергетического барьера для основных носителей заряда, уменьшению напряженности поля в приконтактной области (уменьшению «изгиба» уровней диаграммы) и, в целом, к нарушению равенства диффузионного и дрейфового токов: диффузия (поток) электронов из n области и встречная диффузия дырок становится преобладающей по сравнению с дрейфовым движением.

Поскольку эффекты диффузии более ощутимы, чем эффекты дрейфа, то результирующая плотность тока при прямом включении электронно-дырочной структуры будет теперь в основном диффузионной и тем больше, чем больше величина приложенного напряжения. Следовательно, опять же имеет место инжекция неосновных носителей заряда и возникает ток через внешние выводы..

Надо иметь в виду, что, так как в прилегающих к обедненному слою полупроводниках концентрация «своих» неосновных носителей заряда мала, то попавшие «чужие» носители существенно изменяют равновесную концентрацию носителей. Требуется некоторое время для их «нейтрализации» или, как еще говорят, «релаксации». «Переходящие чужие»

электроны, имеющие большую энергию, чем «свои», за время $t \sim (10^{-5} - 10^{-10})$ секунд теряют на ее соударениях имеющих место при рекомбинации.

1.3.4. Функционирование электронно-дырочной структуры (идеального диода) при приложении внешнего обратного напряжения

При обратном включении идеального диода в нем будут происходить следующие процессы:

А. В результате суперпозиции исходного внутреннего $E_{ИСХ}$ и внешнего $E_{ВНЕШ}$ электрических полей результирующая напряженность электрического поля внутри обедненного слоя увеличится.

Это приводит к тому, что, имевшее место в состоянии равновесия, равенство (1.18) токов нарушается. Диффузионный ток становится меньше дрейфового.

Б. Из-за того, что в обедненном слое преобладает дрейфовое движение носителей заряда, происходит постоянное «вытягивание» результирующим электрическим полем неосновных носителей из p и n областей, граничащих с обедненным слоем (рис. 1.14). Как говорят, на границах обедненного слоя имеет место экстракция неосновных носителей заряда. «Вытянутые» неосновные носители электрическим полем «перемещаются» через обедненный слой в противоположную соседнюю область, где они становятся основными. Важно, что концентрация неосновных носителей вблизи границ обедненной области уменьшается и становится меньше, чем в пространстве вблизи контактов.

В. Так как вблизи границы обедненного слоя концентрация неосновных носителей меньше равновесной, а в местах подключения выводов источников концентрация их будет почти равновесной (т. е. больше), то можно утверждать, что имеет место пространственная разность концентраций и возникнет диффузия: неосновные носители от приконтактных областей начнут двигаться к границам обедненного слоя. При этом по мере их движения вглубь p и n областей от выводов они также рекомбинируют.

Г. Чтобы прилегающие к обедненному слою, p и n полупроводники по-прежнему оставались электронейтральными, вступает в действие

механизм компенсации заряда, обеспечивающий протекание тока через внешние выводы. Его работу, в первом приближении, можно пояснить следующим образом.

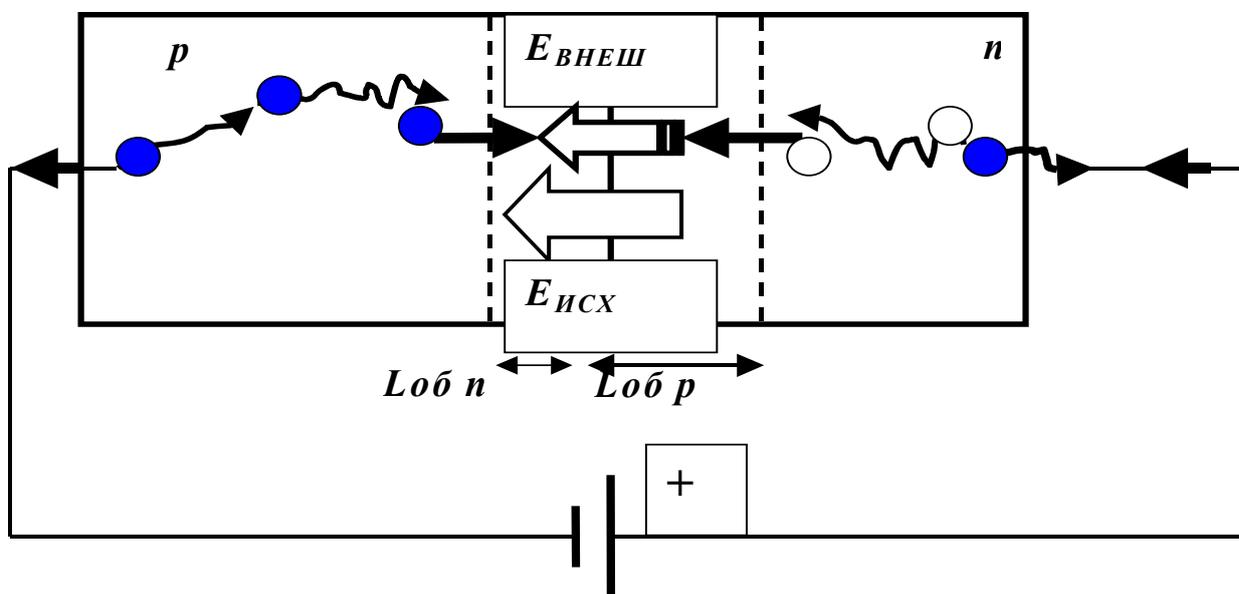


Рис. 1.14. Электронно-дырочная структура при приложении к ней внешнего отрицательного напряжения (обратное включение)

Чтобы p область оставалась электронейтральной из-за экстракции электронов, из металлического вывода, имеющего отрицательный потенциал, в p область входит поток электронов, имеющий такой же самый заряд, что и «вытянутые» электрическим полем электроны. Чтобы n область оставалась электронейтральной, из-за экстракции дырок, из n области в металлический вывод, имеющий положительный потенциал, уходит поток электронов, имеющий такой же самый заряд, что и «вытянутые» электрическим полем дырки. Таким образом, через металлические внешние выводы диода протекает ток. Однако, на этот раз речь идет о движении неосновных носителей заряда. Поскольку их количество значительно меньше чем основных, то можно предполагать, что ток при обратном включении электронно-дырочной структуры будет на несколько порядков меньше тока, протекающего через ту же структуру при прямом включении.

Для диода при обратном включении ($U < 0$) из граничных условий Шокли (1.28, 1.29) следует:

$$p_n^{ПГОС}(U < 0) = p_{n0}^{ПГОС} \ell^{\frac{-U}{\phi_T}}, \quad n_p^{ЛГОС}(U < 0) = n_{p0}^{ЛГОС} \ell^{\frac{-U}{\phi_T}}. \quad (1.33)$$

То есть граничные условия Шокли свидетельствуют о том, что концентрация неосновных носителей заряда на границах обедненного слоя при отрицательном напряжении очень резко (экспоненциально) уменьшается по сравнению с равновесной.

Явление уменьшения концентрации неосновных носителей заряда на границах обедненного слоя получило название *экстракции (вытягивания) неосновных носителей заряда*.

В несимметричных электроно-дырочных структурах, типа $p^{++} - n$, имеющих $N_A \gg N_D$, в соответствии (1.8, 1.9) $p_{n0}^{ПГОС} \gg n_{p0}^{ЛГОС}$.

Следовательно, как вытекает из граничных условий Шокли, $p_n^{ПГОС}(U < 0) \gg n_p^{ЛГОС}(U < 0)$, то есть, имеет место односторонняя экстракция неосновных носителей заряда (главным образом, экстракция дырок из области структуры в которой концентрация введенной легирующей примеси меньше).

Для диода при обратном включении ($U < 0$) из формулы (1.32) для тока следует

$$I = I_0 \left(\ell^{\frac{-U}{\phi_T}} - 1 \right) = -I_0, \quad (1.34)$$

где $I_0 = S \left(\frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \right)$ - *тепловой ток*.

Название «тепловой» ток получил по двум причинам.

Во-первых, это название подчеркивает то обстоятельство, что I_0 сильно зависит от температуры.

Если концентрацию неосновных носителей выразить, используя закон действующих масс (1.6, 1.7), через собственную концентрацию и концентрации примесей, то:

$$I_0 = qSn_i^2 \left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A} \right). \quad (1.35)$$

То, что тепловой ток связан с собственной концентрацией как раз и подчеркивает его зависимость от температуры, что уже отмечалось.

Во-вторых, это название связано с механизмом его образования: тепловой ток обусловлен тепловой генерацией неосновных носителей заряда в электронейтральных областях, прилегающих к обедненному слою. Докажем это.

Так как коэффициенты диффузии связаны с диффузионными длинами дырок L_p и электронов L_n и временами жизни носителей τ_p, τ_n следующими соотношениями

$D_p = \frac{L_p^2}{\tau_p}$, $D_n = \frac{L_n^2}{\tau_n}$, то выражение (1.35) можно переписать в виде:

$$I_0 = \left(\frac{qSL_p p_{n0}}{\tau_p} + \frac{qSL_n n_{p0}}{\tau_n} \right). \quad (1.36)$$

где SL_p, SL_n – это объемы областей, где происходит термогенерация дырок и электронов, соответственно, а $\frac{SL_p p_{n0}}{\tau_p}, \frac{SL_n n_{p0}}{\tau_n}$ – полное число неосновных носителей, генерированных в слоях толщиной L_p и L_n за единицу времени. Что и требовалось доказать.

С точки зрения энергетических диаграмм функционирование электронно-дырочной структуры при приложении к ней внешнего обратного напряжения, в первом приближении, можно пояснить следующим образом.

Если осуществить обратное включение диода, то есть к p полупроводнику приложить отрицательный потенциал, а к n полупроводнику – положительный, то состояние термодинамического равенства токов бу-

дет нарушено. Имея перед глазами рис.1.11, мысленно это можно представлять так.

У n полупроводника, подключенного к положительному потенциалу источника все уровни, в том числе и уровень Ферми, будут смещаться вниз на величину qU , а p полупроводника – вверх. Все это приведет к увеличению энергетического барьера для основных носителей заряда, повышению напряженности поля в приконтактной области (увеличению «изгиба» уровней диаграммы) и, в целом, к нарушению равенства диффузионного и дрейфового токов: диффузия электронов из n области и встречная диффузия дырок становятся существенно меньшими по сравнению с дрейфовым движением.

Поскольку эффекты дрейфа более ощутимы, чем эффекты диффузии, то результирующая плотность тока при обратном включении электронно - дырочной структуры будет теперь в основном дрейфовой. Следовательно, опять же имеет место экстракция неосновных носителей заряда и возникает ток через внешние выводы.

1.3.5. Вольт-амперная характеристика и основные свойства электронно-дырочной структуры (идеального диода)

Вольт-амперная характеристика (ВАХ) электронно-дырочной структуры описывается формулой (1.32). При прямом включении диода ($U>0$) график ВАХ (говорят график прямой ветви), имеет вид, показанный на рис.1.15.

Анализируя ВАХ, следует отметить следующее.

ВАХ диода при прямом включении существенно нелинейна. Она имеет сходство с хоккейной клюшкой. Зависимость между током и напряжением носит «пороговый» характер: при малых напряжениях ток близок к нулю, а затем, достигая некоего «порога» напряжения U^* , начинает резко расти.

График ВАХ содержит два участка. Первый участок – начальный ($0 < U < U^*$) – характеризуется током малой величины и медленным его нарастанием с ростом напряжения. Для кремния это диапазон напряжений (0 – 0,6) В, для германия (0 – 0,3) В. На втором участке (а,б) происходит резкое возрастание тока при незначительном ($U \sim U^*$) изменении напряжения (в пределах десятых долей вольта). При этом изменению

тока на порядок (в 10 раз) соответствуют изменения напряжения на $2,3 \varphi_T (< 60 \text{ мВ})$.

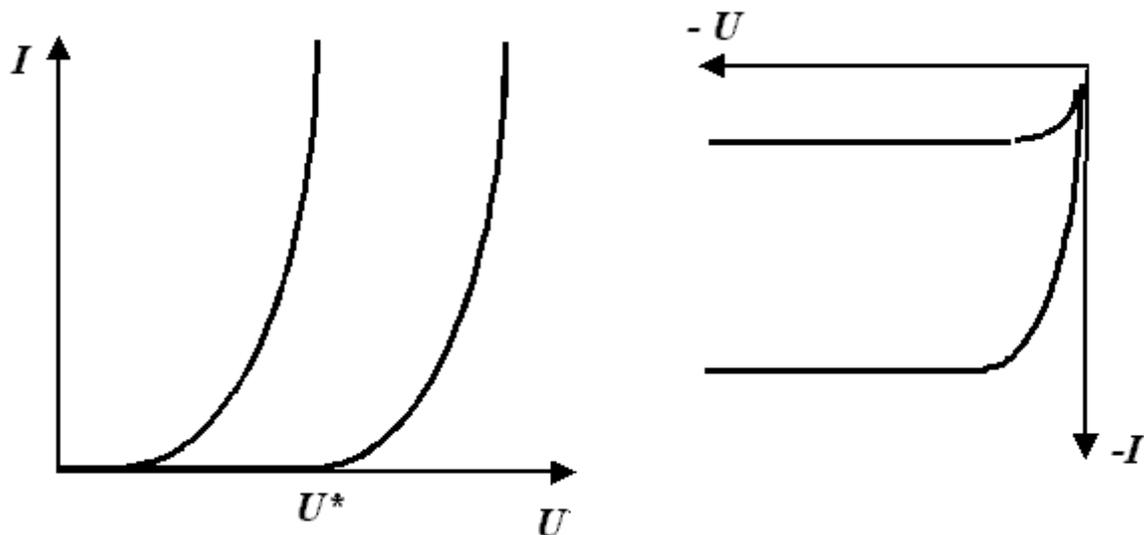


Рис. 1.15. Вольт-амперная характеристика идеального диода при прямом (слева) и обратном (справа) включениях

Если вспомнить классификацию компонентов электрической цепи, то диоду на этих участках можно придать разный смысл. На первом участке диод в большей степени напоминает разрыв цепи. На втором участке, диод напоминает источник напряжения, приближающийся по свойствам к идеальному (к источнику с ЭДС $\sim U^*$ и малым внутренним дифференциальным сопротивлением r)

$$r = \frac{dU}{dI} = \frac{1}{dI/dU} = \frac{1}{\left(\frac{1}{\varphi_T}\right) I_0 e^{\frac{U}{\varphi_T}}} = \frac{\varphi_T}{I + I_0}. \quad (1.37)$$

При токе $I = 1 \text{ мА}$ дифференциальное сопротивление равно $r = 0,026 \text{ Ом}$.

Следует иметь в виду, что ВАХ электронно-дырочной структуры с увеличением температуры смещается влево. Это обусловлено зависимостью ряда электрофизических параметров полупроводников от температуры. Доминирующей является зависимость собственной концен-

трации от температуры, поэтому падение напряжения на идеальном диоде уменьшается с ростом температуры. Температурный коэффициент напряжения получается отрицательным. Для кремниевого диода он равен $-2 \text{ мВ}^\circ\text{С}$.

Из физики известно, что если при изменении напряжения меняется заряд, то это равносильно появлению электрической емкости. При прямом включении возникает инжекция и, затем, «впрыснутый» заряд перемещается за счет диффузии. В основе этого – напряжение. Меняя напряжение, можно менять заряд и, следовательно, электронно-дырочную структуру можно характеризовать емкостными свойствами, связанными с диффузией:

$$C_{\text{диф}} = \frac{dQ}{dU} = \frac{dQ}{dU} \frac{dI}{dI} = \frac{dQ/dI}{dU/dI} = \frac{dQ/dI}{\frac{(I + I_0)}{\varphi_T}} = \tau \frac{(I + I_0)}{\varphi_T}, \quad (1.38)$$

где τ - *время жизни избыточных носителей заряда*.

Величину $C_{\text{диф}} = \tau \frac{(I + I_0)}{\varphi_T}$ называют *диффузионной емкостью*.

При обратном включении диода ($U < 0$) график ВАХ (говорят график обратной ветви), имеет вид, показанный на рис. 1.15.

Следует отметить, что ВАХ диода при обратном включении также нелинейна и содержит два участка. Первый участок – начальный – характеризуется ростом тока. На втором участке (с,d) ток не меняется.

Если опять же вспомнить классификацию компонентов электрической цепи, то диод на втором участке напоминает источник тока, приближающийся по свойствам к идеальному (к источнику тока I_0 с большим внутренним дифференциальным сопротивлением R):

$$R_{\text{диф}} = \frac{dU}{dI} = \frac{\varphi_T}{I_0}. \quad (1.39)$$

При $I_0 = 10^{-9} \text{ А}$ получаем $R_{\text{диф}} = 26\,000\,000 \text{ Ом} = 26 \text{ МОм}$.

Сравнивая величины дифференциальных сопротивлений диода при прямом и обратном включении, можно сделать вывод, что диод обладает свойством однонаправленной проводимости: сопротивление

протеканию тока при прямом включении очень мало, а при обратном – очень велико, поэтому ток через электронно-дырочную структуру может фактически протекать только в одном направлении.

В обедненном слое по обе стороны от металлургической границы существуют равные по значению и противоположные по знаку объемные заряды, обусловленные ионами примесей. В зависимости от приложенного напряжения меняется толщина обедненного слоя и, следовательно, величина зарядов. Это свидетельствует о том, что и при обратном включении электронно-дырочная структура обладает емкостными свойствами и ведет себя подобно электрическому конденсатору. Емкость такого конденсатора называется барьерной емкостью электронно-дырочной структуры:

$$C_{бар} = \frac{dQ}{dU} = S \sqrt{\frac{q \epsilon \epsilon_0 N}{2(\phi_K - U)}} \Rightarrow N = \frac{N_A N_D}{N_A + N_D}. \quad (1.40)$$

График зависимости барьерной емкости от напряжения имеет вид, показанный на рис. 1.16.

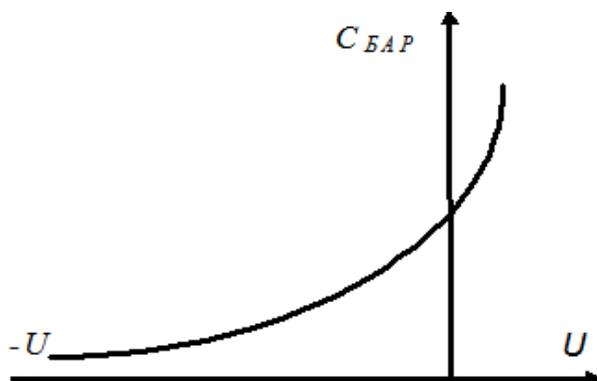


Рис. 1.16. Зависимость барьерной емкости от напряжения

1.4. Переход металл – полупроводник и его свойства

1.4.1. Основные понятия

Прежде чем изучать структуры металл – полупроводник (Me – п/п), напомним известные из физики понятия и сведения об исходных веще-

ствах. Металл и полупроводник это совершенно разные (разнородные) материалы, поэтому для понимания природы контакта Me – п/п надо сначала четко представлять условия при которых электрон может «выходить» за пределы вещества, понимать, что такое так называемая «работа выхода».

Энергетические диаграммы обоих материалов показаны на рис. 1.17. Будем считать, что поверхность анализируемого материала граничит с вакуумом. Уровень вакуума является отсчетным и определяет энергию, которую надо сообщить «покоящемуся» электрону для того, чтобы он покинул материал.

Для определенности возьмем n полупроводник, у которого уровень Ферми находится в верхней половине запрещенной зоны. У металла уровень Ферми находится внутри зоны проводимости. Известно, что электроны не могут покинуть пределы вещества из-за потенциального барьера. Последний отражает «притяжение» электронов положительными ионами кристаллической решетки. Чтобы преодолеть силу, действующую со стороны кристалла, требуются затраты работы по преодолению удерживающей силы. Соответственно, выход электронов из любого вещества определяется некоторой энергией, именуемой работой выхода, которую нужно затратить для преодоления приповерхностного барьера

Различают две работы выхода.

1. *Термодинамическая работа выхода* (W^{Me} , $W^{п/п}$) – работа по перемещению электрона средней энергии за пределы твердого тела, которую необходимо затратить, чтобы электрон «перешел» с уровня Ферми на уровень Ферми в вакууме.

2. *Внешняя работа выхода* ($W_{ВН}^{Me}$, $W_{ВН}^{п/п}$) (ее еще называют электронным сродством, сродством электрона и вакуума, истинной работой выхода) – работа, которую необходимо затратить для перемещения электрона за пределы твердого тела, чтобы электрон «покинул» зону проводимости и «перешел» на уровень Ферми в вакууме. Внешнюю работу выхода принято отсчитывать от дна зоны проводимости.

Известно, что самой низкой работой выхода обладают щелочные металлы (цезий), самой высокой иод. У металлов используемых в электронике она составляет в электрон – вольтах: Al – 4,5; Au - 4,76; Pt - 6,0; W – 4,55. Внешняя работа выхода полупроводников: Si – 4,05; GaAs – 4,07.

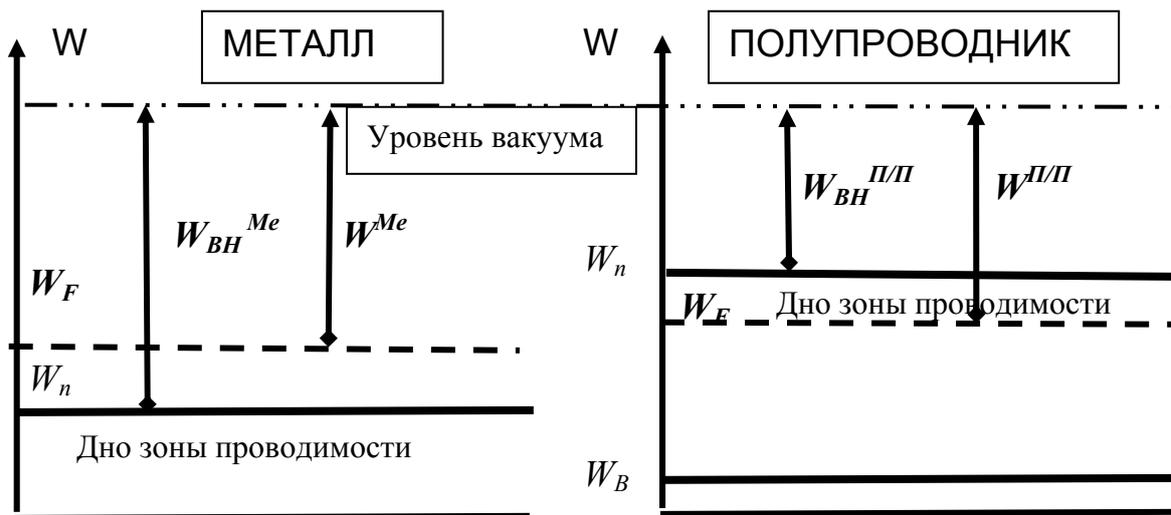


Рис. 1.17. Энергетические диаграммы металла и n полупроводника

Термодинамическая работа выхода полупроводников зависит от типа полупроводника и степени его легирования примесью. Поэтому возможны ситуации, когда термодинамическая работа выхода металла больше чем у полупроводника (обычно это справедливо для полупроводника n типа) и, наоборот, – металла меньше чем у полупроводника (обычно для p типа). Плотность тока электронов, преодолевающих потенциальный барьер и выходящих из твердого тела. Определяется формулой термоэлектронной эмиссии, называемой также формулой Ричардсона – Дешмена:

$$j_{П/П} = A_{П/П} T^2 \ell^{-\frac{W^{П/П}}{kT}}, \quad (1.41)$$

$$j_{Me} = A_{Me} T^2 \ell^{-\frac{W^{Me}}{kT}}, \quad (1.42)$$

где $A = \frac{4\pi m q (kT)^2}{h^3}$ постоянная Ричардсона.

1.4.2. Образование контакта металл – полупроводник и процессы в нем в равновесном состоянии

Пусть термодинамическая работа выхода из металла больше чем работа выхода из полупроводника. Кроме того, будем полагать, что при соприкосновении, зазор между металлом и полупроводником достаточно мал. Тогда, в идеальной структуре металл – п/п сразу же после соприкосновения начнут происходить следующие физические процессы.

А. После соприкосновения металла и полупроводника согласно (1.41, 1.42), появятся два потока электронов. Так как $W^{Me} > W^{П/П}$, то поток электронов из металла будет меньше потока электронов из полупроводника. Соответственно, поверхность металла будет заряжаться отрицательно, а поверхность полупроводника положительно (вблизи границы контакта будут «оголяться» ионы доноров). С точки зрения энергетических диаграмм (рис. 1.17) это приводит к выравниванию уровней Ферми: исходная диаграмма металла начнет «подниматься» по вертикали, а полупроводника – «опускаться».

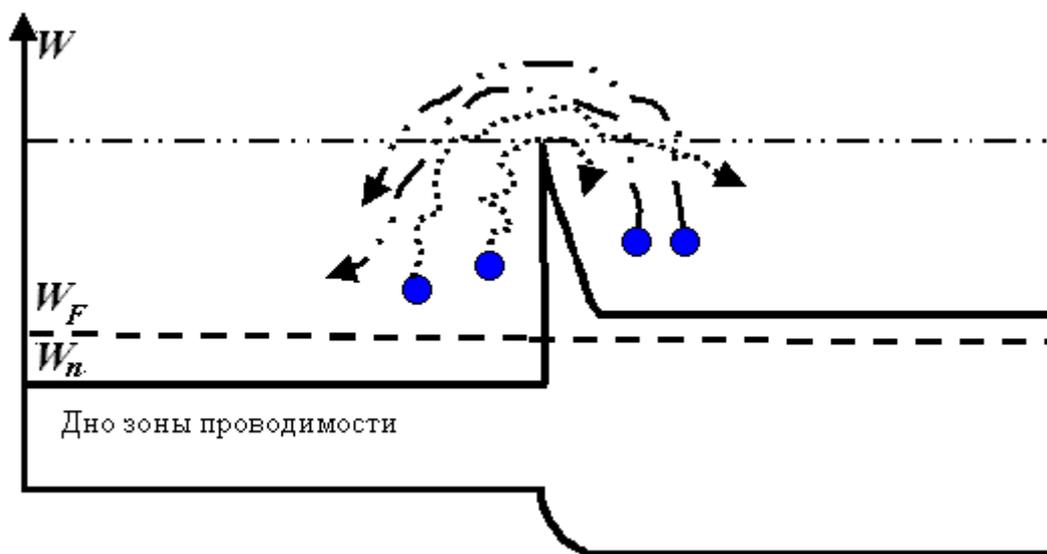


Рис. 1.18. Энергетическая диаграмма структуры металл – полупроводник в состоянии равновесия

Б. Появление разноименных зарядов эквивалентно появлению дополнительного «барьера», поэтому при неизменном потоке электронов

из металла ток из полупроводника в металл начнет уменьшаться и, в конце концов, наступит равенство токов $j^{Me} = j^{П/П}$. С точки зрения энергетических диаграмм равенство токов наступит при равных уровнях Ферми (рис. 1.18). Возникает состояние динамического равновесия.

В. При $j^{Me} = j^{П/П}$ между металлом и полупроводником возникает контактная разность потенциалов, которую в первом приближении можно оценить так: $\varphi_K = 1/q(W^{Me} - W^{П/П})$. На энергетической диаграмме также виден «энергетический барьер» $q\varphi_K$. Для «перехода» из металла в полупроводник и наоборот электроны должны затратить энергию чтобы преодолеть «горку» (рис. 1.18).

Следует иметь в виду, что в состоянии термодинамического равновесия встречные потоки равны, поэтому токи через внешние выводы структуры равны нулю.

Заметим, что если имеет место контакт металл – полупроводник р типа, то «оголяются» отрицательные ионы акцепторов и «искривление» энергетических зон идет в противоположную сторону. Но также возникает контактная разность потенциалов и энергетический барьер.

1.4.3. Функционирование структуры металл – полупроводник (идеального диода Шоттки) при приложении к ней внешнего напряжения

Если осуществить прямое включение диода Шоттки, то есть к металлу приложить положительный потенциал, а полупроводнику – отрицательный, то состояние термодинамического равенства токов будет нарушено.

У полупроводника, подключенного к отрицательному потенциалу источника все уровни, в том числе и уровень Ферми, будут смещаться вверх на величину qU , что приведет к понижению энергетического барьера для электронов, переходящих в металл. Плотность электронов, идущих из полупроводника в металл станет равной:

$$j_{П/П} = j_S \ell^{\frac{qU}{kT}}. \quad (1.43)$$

Результирующая плотность тока при прямом включении структуры Me – п/п будет:

$$j_{PP} = j_{П/П} - j_{Me} = j_S \ell^{\frac{qU}{kT}} - j_S = j_S (\ell^{\frac{qU}{kT}} - 1). \quad (1.44)$$

Уравнение, умноженное на площадь контакта, чисто внешне напоминает уравнение для электронно-дырочной структуры (1.32). Однако сходство чисто внешнее. С точки зрения внутренних физических процессов идеальный диод Шоттки и электронно-дырочная p-n структура различаются существенным образом.

В p-n переходе, как отмечалось, имеют место инжекция неосновных носителей заряда, появление избыточных носителей на границах обедненного слоя, последующие диффузии и рекомбинации носителей заряда.

В структуре Me – п/п нет инжекции неосновных носителей. В такой структуре напряжение управляет потоком основных носителей, попадающих в металл. Так как в металле значительная концентрация «своих» электронов, то попавшие «чужие» не очень изменяют равновесную концентрацию носителей и происходит быстрая их «нейтрализация» или, как еще говорят, «релаксация». «Переходящие чужие» электроны, имеющие большую энергию, чем электроны металла, быстро (за время $t \sim 10^{-16}$ с) ее теряют на соударениях. Говорят, релаксация происходит за максвелловское время релаксации $t \sim 10^{-16}$ с). Важно отметить, что в структуре Me – п/п никогда не бывает накопления заряда. Это свойство диода Шоттки широко используется.

Понятно, что контакт металл – полупроводник не обладает диффузионной емкостью.

Если осуществить обратное включение диода Шоттки, то есть к металлу приложить отрицательный потенциал, а полупроводнику – положительный, то состояние термодинамического равенства токов также будет нарушено.

У полупроводника, подключенного к положительному потенциалу источника все уровни, в том числе и уровень Ферми, будут смещаться вниз на величину $-qU$, что приведет к увеличению энергетического

барьера для электронов, переходящих в металл. Плотность электронов, идущих из полупроводника в металл станет равной:

$$j_{\Pi/\Pi} = j_S \ell^{\frac{-qU}{kT}}. \quad (1.45)$$

Результирующая плотность тока при обратном включении структуры Me – п/п будет:

$$j_{\Pi\Pi} = j_{\Pi/\Pi} - j_{Me} = j_S \ell^{\frac{-qU}{kT}} - j_S = j_S (\ell^{\frac{-qU}{kT}} - 1) \approx -j_S. \quad (1.46)$$

Тепловой ток I_S диода Шоттки значительно больше теплового тока электронно-дырочной структуры. Это обуславливает то, что при одинаковом прямом токе структура Me – п/п имеет большее прямое напряжение. На рис. 1.19 для сравнения показаны вольт – амперные характеристики идеализированных кремниевого диода и диода Шоттки структуры алюминий - кремниевый n полупроводник.

Следует помнить, контакт металл – полупроводник, также как и электронно-дырочная структура, обладает барьерной емкостью.

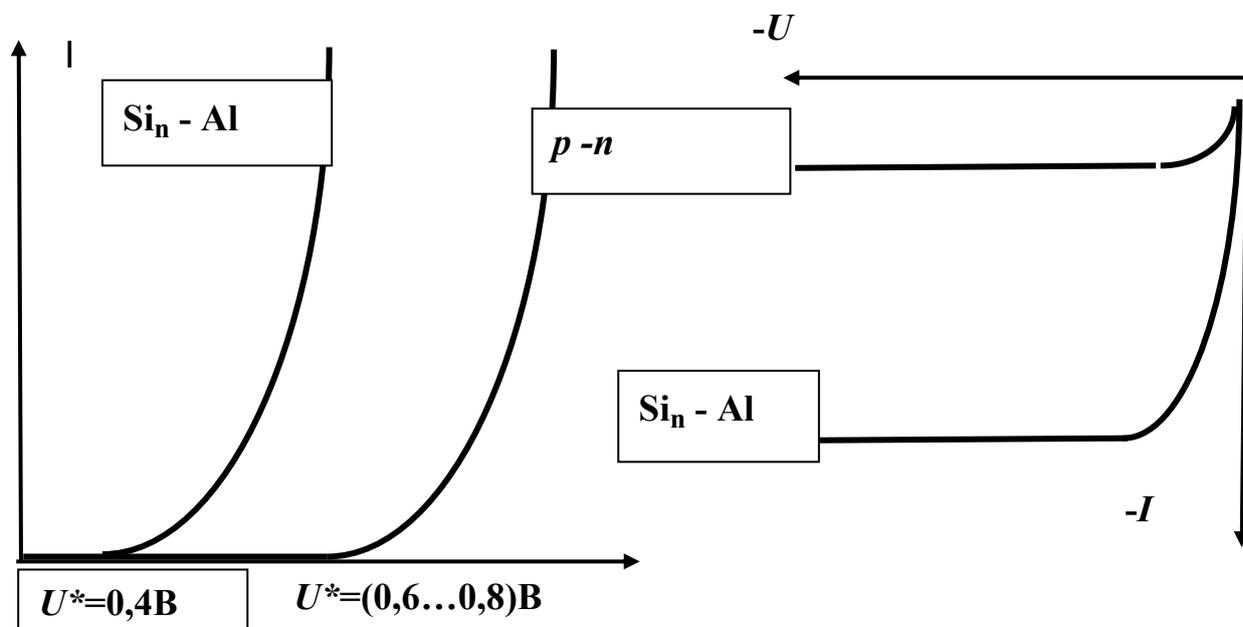


Рис. 1.19. Вольт-амперные характеристики электронно-дырочной структуры и идеального диода Шоттки

1.4.4. Омические (невыпрямляющие) контакты

Контакты такого типа используются практически во всех полупроводниковых приборах для формирования внешних выводов от полупроводниковых областей. Для омического контакта характерна близкая к линейной зависимость тока от напряжения (как для резистивного элемента). Сопротивление протекающему току омического контакта должно быть малым.

Существует два способа формирования омических контактов.

1. Перед созданием контакта проводят шлифовку поверхности полупроводника, в результате которой на его поверхности образуются «дефекты» кристаллической решетки, являющиеся центрами рекомбинации. Они способствуют подавлению инжекции носителей и формированию линейной ВАХ.

2. Формирование на поверхности полупроводника тонкого слоя полупроводника с высокой концентрацией примеси.

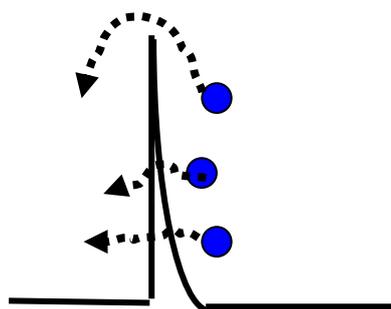


Рис. 1.20. Упрощенное изображение процесса «туннелирования» электронов через барьер

Алюминий, являясь акцепторной примесью, дает хорошие омические контакты на p – кремнии. Однако сложно создать хорошие омические контакты к слаболегированным n полупроводникам, поскольку при этом возникает барьер. Выходом из положения является создание на поверхности n – полупроводника сильно легированного слоя n^{+++} , на который наносится металл. В этом случае работает туннельный механизм протекания тока (рис. 1.20). При высоких степенях легирования (используется вырожденный полупроводник) толщина обедненного слоя получается чрезвычайно малой (несколько нанометров $\sim 10^{-9}$ м). Поэтому

электроны, вместо того, чтобы преодолевать горку «туннелируют» через барьер.

1.5. Полупроводниковые диоды общего применения

1.5.1. Общие сведения

Полупроводниковым диодом называют полупроводниковый прибор с двумя внешними выводами (двухполюсник), создаваемый на базе сочетания различных выпрямляющих и омических структур, в основу принципа действия которого положено то или иное свойство выпрямляющего электронно-дырочного или металл– полупроводник перехода.

Структуры типичных диодов на основе электронно-дырочного выпрямляющего перехода и выпрямляющего перехода металл – полупроводник (диода Шоттки) и их условное обозначение на схемах показаны на рис. 1.21.

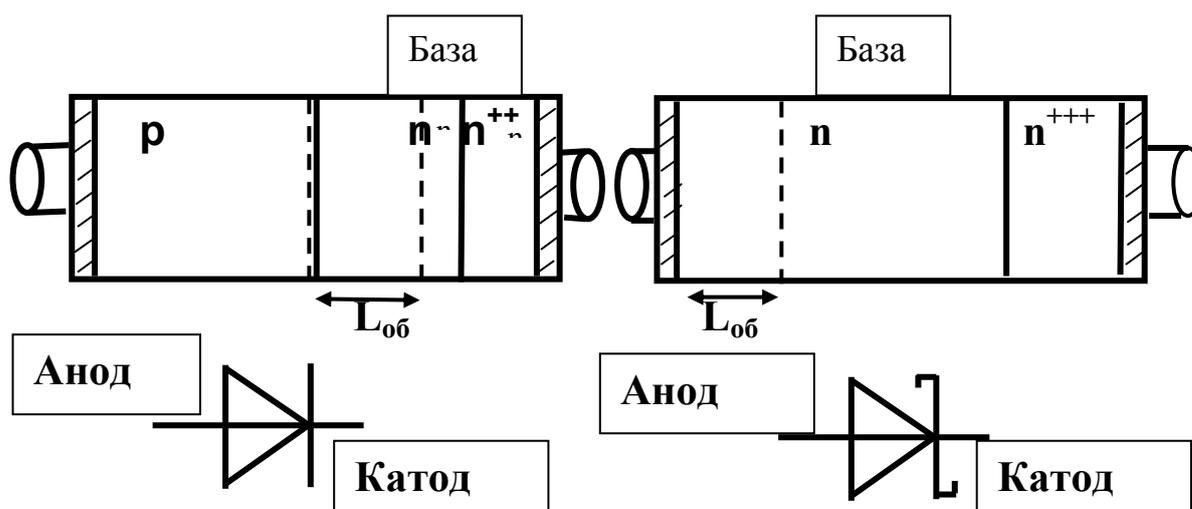


Рис. 1.21. Структуры типичных диодов и их условное изображение на схемах

На краях диода расположены омические контакты металл – полупроводник, к которым припаяны внешние выводы. Эти выводы именуют анод, катод. Левый контакт диода Шоттки представляет собой выпрямляющий контакт металл – полупроводник. Направление стрелки, которую легко выделить из «треугольника» условного обозначения, показывает направление тока, протекающего через диод при прямом включе-

нии. Принято считать, исходя из свойств однонаправленной проводимости, что в диодах ток протекает только от анода к катоду. Области n полупроводника (с меньшей концентрацией легирующей примеси и, соответственно, с самым большим удельным объемным сопротивлением) принято называть базой.

По типу используемого материала диоды могут быть кремниевыми, из арсенида галлия и германиевыми. Большинство современных диодов изготавливаются на основе кремния.

Класс диодов может быть разделен на следующие основные подклассы.

Выпрямительные диоды.

Стабилитроны.

Варикапы.

Импульсные диоды.

Все выпускаемые промышленностью диоды характеризуют большим числом допустимых и предельных электрических параметров, которые приводятся в справочниках или размещаются на страницах сайтов фирм производителей. *Допустимое значение параметра* – это значение, при котором ожидается удовлетворительная работа прибора. *Предельное* – это значение параметра, за пределами которого прибор может быть поврежден или выведен из строя.

1.5.2. Отличия вольт-амперных характеристик реальных диодов от идеальных

Рассматривая идеализированный диод, мы учли лишь самое главное. В реальных диодах имеют место и другие физические эффекты. Эти дополнительные эффекты находят свое выражение в отличиях ВАХ реального и идеального диода. Вариант ВАХ реального (сплошной линией) и идеального (штриховой линией) диодов при прямом и обратном включениях диодов показаны на рис. 1.22.

Отличия ВАХ при прямом включении проявляются в следующем.

1. При малых напряжениях возникает дополнительный ток рекомбинации.

Если прямое напряжение мало, то высота потенциального барьера еще относительно высока. Поэтому часть основных носителей заряда не

может преодолеть этот барьер, пересечь обедненный слой и создать ток инжекции.

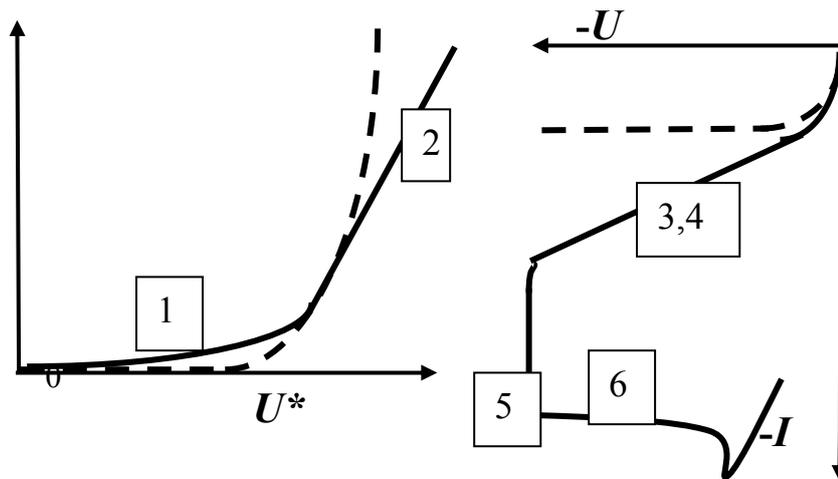


Рис. 1.22. Вольт – амперные характеристики идеального и реального диодов с нумерацией причин отличий

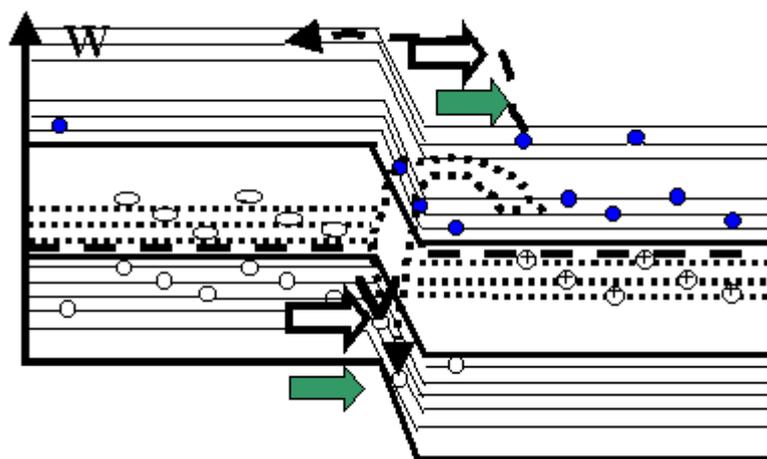


Рис. 1.23. Энергетическая диаграмма, поясняющая возникновение в обедненном слое дополнительного процесса рекомбинации

Вместе с тем, при изготовлении реальных диодов, из-за неидеальности технологии, вблизи металлургической границы возникают различного рода дефекты, получившие название центры рекомбинации Холла, Шокли, Рида (ХШР - центры). Поэтому часть носителей заряда, не преодолев весь обедненный слой, может в нем прорекомбинировать на ХШР – центрах (рис. 1.23). Вследствие рекомбинации возникает допол-

нительный ток рекомбинации (показан зачерненными стрелками), совпадающий по направлению с током инжекции (показан незачерненными стрелками) и ток реального диода больше чем идеального.

Можно показать, что отношение тока рекомбинации к току инжекции определяется выражением:

$$\frac{j_{PEK}}{j_{ИНЖ}} \approx \frac{n_i \ell \frac{U}{2\varphi_T}}{p_n L_p}. \quad (1.47)$$

Из него следует, что ток рекомбинации будет наблюдаться у диодов, выполненных из материала с большой шириной запрещенной зоны (у кремниевых и арсенид галлиевых).

2. При выводе ВАХ предполагалось, что сопротивления электронной нейтральных областей, примыкающих к обедненному слою, равно нулю. В реальных, несимметричных диодах сопротивление базы составляет десятки и даже сотни Ом. Поэтому внешнее напряжение распределяется между обедненной областью и базой. В этом случае

$$I = I_0 \left(\ell \frac{(U - I r_B)}{\varphi_T} - 1 \right) \text{ и ВАХ идет правее идеальной, все больше отклоня}$$

ясь при нарастании тока.

Чтобы снизить сопротивление базы, используют специальные методы, в частности, применяют *эпитаксиальные структуры*.

Отличия ВАХ при обратном включении проявляются в следующем.

3. У реальных диодов имеют место «утечки» тока.

На поверхности диода, вследствие «загрязнений» при изготовлении и так называемых поверхностных состояний, образуются микроскопические проводящие каналы, по которым протекает дополнительный ток утечки. Ток утечки увеличивается пропорционально напряжению.

4. В наличии у реальных диодов дополнительного тока термогенерации в обедненном слое.

В обедненном слое на ХШР-центрах возможно появление (генерация) дополнительных носителей заряда. Так как внутри обедненного слоя имеется электрическое поле, то возникшие дополнительные носи-

тели сразу же «уносятся» из обедненной области, не успевая рекомбинировать (рис. 1.24).

Для несимметричного диода тепловой ток, как следует из (1.36) равен:

$$I_0 = \left(\frac{qSL_p p_{n0}}{\tau_p} \right). \quad (1.48)$$

Напомним, что SL_p - это объем n области, где происходит термогенерация дырок, а $\frac{SL_p p_{n0}}{\tau_p}$ - полное число неосновных носителей, генерированных в слое толщиной L_p за единицу времени.

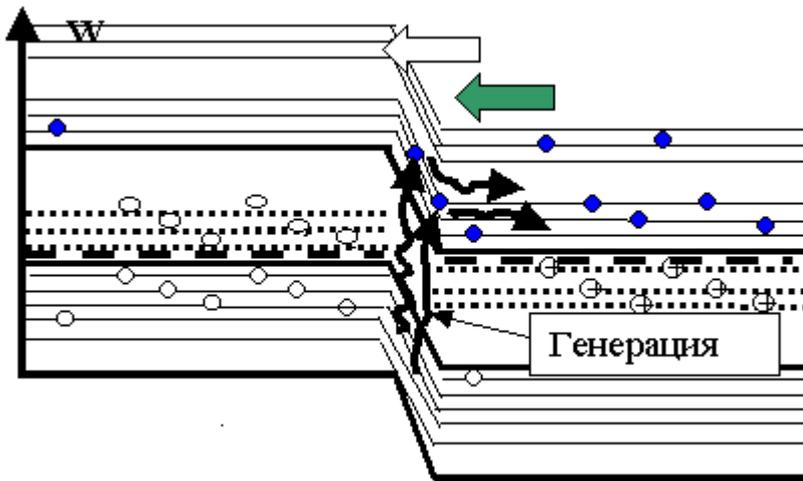


Рис. 1.24. Энергетическая диаграмма, поясняющая возникновение в обедненном слое дополнительного процесса генерации носителей заряда

По аналогии, учитывая, что генерация идет в толще обедненного слоя, можно записать:

$$I_{TG} = \left(\frac{qSL_{об} n_i}{\tau_p} \right) \Rightarrow \frac{I_{TG}}{I_0} \approx \frac{L_{об} n_i}{L_p p_n} \approx \frac{L_{об} n_i}{n_i^2} \approx \frac{L_{об}}{n_i} \ell \frac{\Delta W}{2kT}. \quad (1.49)$$

Выражение (1.49) свидетельствуют о том, что доля тока термогенерации в полном обратном токе тем выше, чем больше ширина запрещенной зоны и ниже температура. В кремниевых и арсенид галлиевых диодах собственная концентрация носителей заряда мала, поэтому в них преобладает ток термогенерации над тепловым током. Для кремниевого диода $I_0 = 10^{-14}$ А, $I_{ТГ} = 10^{-9}$ А.

Кроме того, следует иметь в виду, что с увеличением обратного напряжения ширина обедненного слоя увеличивается и, следовательно, увеличивается ток термогенерации.

5. При относительно больших обратных напряжениях в реальных диодах может наблюдаться электрический пробой. Как известно из физики при этом возможны две разновидности электрического пробоя.

А. Туннельный пробой, который наблюдается в $p - n$ структурах с сильно легированными и вырожденными $p - n$ областями (рис. 1.25).

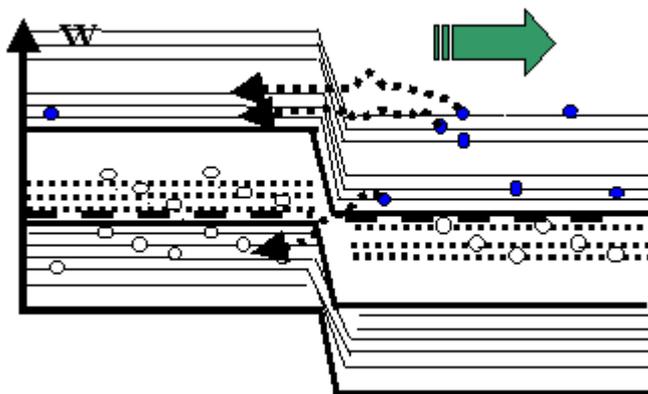


Рис. 1.25. Энергетическая диаграмма, поясняющая возникновение туннельного пробоя

В этом случае толщина обедненного слоя очень мала ($\sim 10^{-8}$ м). Электроны из зоны проводимости n области могут «туннелировать» в зону проводимости p полупроводника и, при сильном «искривлении» зон, даже в валентную зону p полупроводника. Обратный ток резко растет (рис. 1.22).

Напряжение, при котором возникает *туннельный* (или как еще говорят *Зенеровский*) пробой, можно оценить по формуле:

$$U_Z = \frac{\varepsilon\varepsilon_0 E_{KP}^2}{2q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right), \quad (1.50)$$

где $E_{KP} \sim 1,4 \cdot 10^6$ В/см.

Обычно напряжение туннельного пробоя не превышает 5 В.

Туннельный пробой отличается тем, что с ростом температуры напряжение туннельного пробоя уменьшается, из – за того, что ширина запрещенной зоны при этом уменьшается и вероятность туннелирования увеличивается. Температурный коэффициент напряжения оказывается отрицательным.

Б. Лавинный пробой.

Этот вид пробоя характерен для диодов с высокоомной базой. Он связан с образованием «лавины» из носителей заряда под действием сильного электрического поля.

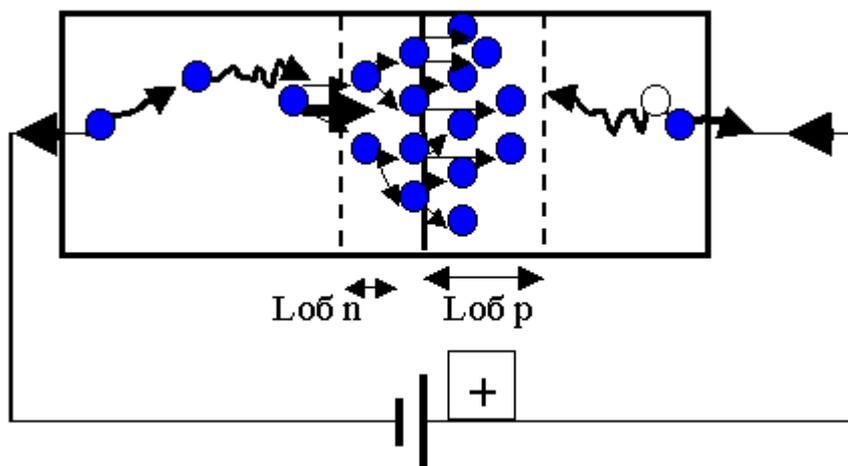


Рис. 1.26. **Модель диода, поясняющая механизм возникновения лавинного электрического пробоя**

Электроны, «разогнавшись» в электрическом поле обедненного слоя, могут набрать энергию, достаточную для ударной ионизации вещества («сбивания» электронов с внешних орбит атома). В результате разрыва валентной связи ионизирующим носителем появляется, к примеру, два новых электрона и две новых дырки (рис. 1.26). Эти носители также набирают энергию достаточную для ионизации носителей в обедненном слое. Количество вновь образовавшихся носителей будет уже 4. И далее, по аналогии, количество носителей будет расти, как говорят

«размножаться»: 8,16,32, 64 и т. д. С ростом напряженности электрического поля процесс размножения электронно дырочных пар приобретает характер «лавины», а ток резко увеличивается.

Напряжение, при котором возникает лавинный пробой, определяется формулой:

$$U_{ЛПР} \cong 60 \left(\frac{\Delta W}{1.1} \right)^{1.5} \left(\frac{N_D}{10^{16}} \right)^{-0.75} . \quad (1.51)$$

С ростом температуры напряжение лавинного пробоя увеличивается. Это связано, грубо говоря, с тем, что при повышении температуры увеличиваются тепловые колебания решетки кристалла и электронам «легче» сорваться с внешних орбит атома. Относительный температурный коэффициент напряжения положителен и для кремния составляет 0,1 %/°С.

6. У реальных диодов возможен тепловой пробой.

Тепловой пробой вызывается тем, что при протекании обратного тока в диоде выделяется тепло и его температура повышается. С ростом температуры повышается концентрация неосновных носителей заряда и увеличивается обратный ток. В свою очередь, увеличение обратного тока приводит к увеличению мощности, рассеиваемой на диоде, а это способствует дальнейшему росту температуры. Если количество теплоты, выделяемой на диоде, превышает количество отводимой теплоты, то развивается процесс непрерывного нарастания температуры, а значит, и тока. В результате процесс нарастания обратного тока приобретает лавинообразный характер и наступает тепловой пробой, при котором диод «сгорает». Напряжение теплового пробоя зависит от условий теплоотвода и снижается при повышении температуры окружающей среды.

1.5.3 Выпрямительные диоды

Выпрямительным диодом называют полупроводниковый прибор, посредством которого, используя свойство однонаправленной проводимости тока (вентильное свойство), в схемах выпрямителей преобразуют входной разнополярный сигнал в сигнал одной полярности.

В большей части выпрямительные диоды используются в устройствах электротехники, работающих с гармоническим (косинусоидальным или синусоидальным) напряжением частоты 50 Гц или 400 Гц. Такие диоды называют низкочастотными. Верхняя граница рабочих частот – так называемая предельная частота среднечастотных выпрямительных диодов – как правило, не превышает 20 кГц. Такие диоды используют в импульсных источниках питания. Обычно это диоды Шоттки. Высокочастотные диоды работают до частот 100 МГц.

1.5.4 Стабилитроны

Стабилитроном называют полупроводниковый диод, работающий при обратном включении в режиме электрического пробоя, у которого напряжение на нем слабо зависит от протекающего через него тока.

Стабилитроны предназначены для стабилизации постоянного напряжения (т. е. для того, чтобы напряжение было одной и той же величины) и используются в устройствах, которые называются стабилизаторами напряжения. По своим свойствам стабилитрон приближается к идеальному источнику напряжения (при электрическом пробое его дифференциальное сопротивление очень мало и стараются его сделать так, чтобы он был термостабильным).

Вольт-амперная характеристика стабилитрона и его условное графическое обозначение на схемах, показаны на рис. 1.27.

На участке с, d при изменении тока от минимального I_{min} до максимального значения I_{max} напряжение на стабилитроне остается практически неизменным. Как известно, диапазон изменения напряжения при варьировании тока определяются величиной внутреннего дифференциального сопротивления стабилитрона.

Величины напряжений стабилизации для конкретного типа стабилитрона приводятся в справочниках. Обычно в стабилитронах с напряжением стабилизации (3,5 – 6,3) В используется туннельный пробой, а в стабилитронах с напряжением (6,4 – 140) В – лавинный электрический пробой.

Чтобы сделать лавинный стабилитрон термостабильным (т.е. таким, чтобы его напряжение стабилизации практически не зависело от температуры окружающей среды) используют следующий прием.

Температурный коэффициент напряжения у стабилитронов, основанных на лавинном пробое, положительный. У электронно-дырочных структур при прямом включении отрицательный. Если соединить стабилитрон и несколько диодов, как показано на рис 1.28, чтобы температурные коэффициенты были примерно равны по модулю, но противоположны по знаку, то температура будет мало влиять на напряжение стабилизации.

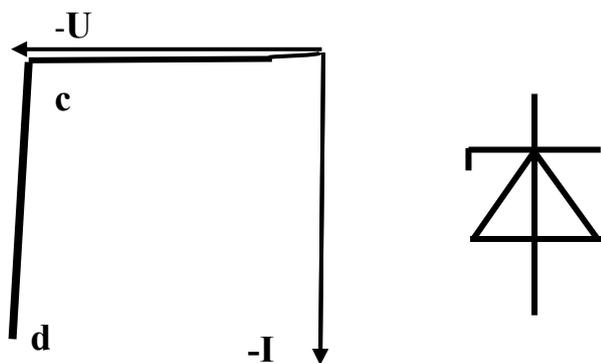


Рис. 1.27. Вольт-амперная характеристика стабилитрона и его условное обозначение на схемах

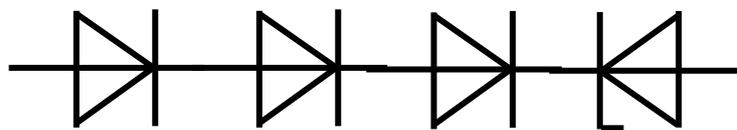


Рис. 1.28. Устройство термокомпенсированного стабилитрона

Такого вида конструкции обычно выпускают в одном корпусе, а подобные стабилитроны называют термокомпенсированными или аттестуемыми прецизионными. Их обычно используют в качестве источников эталонного напряжения.

1.6. Полупроводниковые диоды специального назначения

1.6.1. Варикапы

Варикап [от англ varicap - vari(able) – переменная cap(acity) - емкость] это специальный полупроводниковый диод, работающий при обратном включении, принцип действия которого основан на использовании барьерной емкости электрического перехода от напряжения.

Варикапы предназначены для использования в качестве электрически управляемой емкости. Они применяются в устройствах управления для изменения резонансной частоты параллельного или последовательного электрического контура.

Для описания варикапов используют *вольт-фарадные характеристики*. Вид такой характеристики показан на рис. 1.16. Чтобы получить более широкий диапазон изменения емкости от напряжения используют специальные $p-n$ структуры.

1.6.2. Импульсные диоды и высокочастотные диоды

Импульсные диоды предназначены для работы в быстродействующих импульсных схемах.

При работе этих диодов обычно происходят процессы практически мгновенного переключения, когда полярность приложенного напряжения меняется скачком. Такой режим работы диода принято называть динамическим режимом или режимом большого сигнала. На практике это проявляется в том, что к диоду сначала прикладывается прямоугольный импульс напряжения положительной полярности и очень малой длительности, и, сразу же затем, короткий импульс отрицательной полярности. Если приложено положительное напряжение, то часто говорят, что диод «включают». При переходе с уровня положительного на уровень отрицательного напряжения, диод «выключают».

Для того, чтобы изучить процессы, наблюдаемые при выключении диода, рассмотрим диод с «длинной» базой, у которого p область легирована значительно сильнее, чем n область.

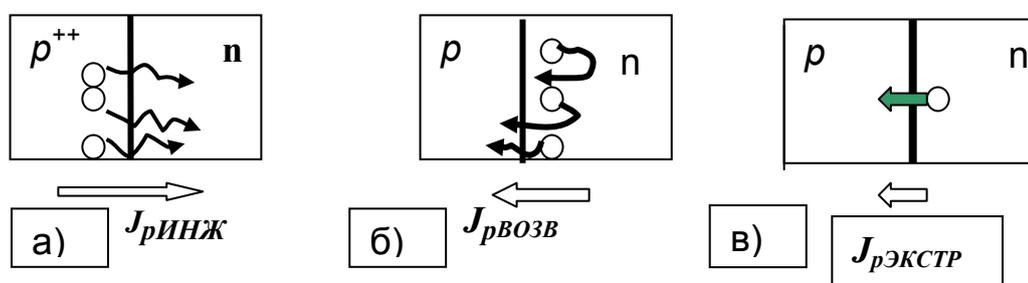


Рис. 1.29. Модели диода, поясняющие процессы накопления и рассасывания носителей в базе

Пусть сначала диод включен и в нем, как отмечалось ранее (рис. 1.29а), происходят процессы односторонней инжекции неосновных носителей заряда (дырок из р в n область). В результате через диод протекает ток I_{m1} (рис. 1.30). В базе диода имеется заряд, образованный инжекцией дырок. В момент времени t_1 напряжение генератора импульсов меняется от положительного значения U_{m1} до отрицательного значения U_{m2} (рис. 1.30). Казалось бы, что из-за этого мгновенно должен исчезнуть прямой ток и через диод должен начать протекать малый тепловой ток I_0 , обусловленный экстракцией неосновных носителей заряда. На самом деле так не происходит. Инжектированные дырки, и соответственно, заряд, накопленный в базе в результате инжекции, мгновенно исчезнуть не могут. Должно пройти определенное время $t_{PAC} = t_2 - t_1$, чтобы дырки «повернули обратно», «покинули базу, ушли назад», а накопленный заряд, как говорят, «рассосался» (рис. 1.29б). При этом ток через диод меняется на противоположный и остается практически неизменным пока ранее инжектированные туда дырки «уходят» из базы и рассасывается заряд (рис. 1.30). Кроме того, уменьшение концентрации ранее инжектированных дырок в базу происходит также за счет рекомбинации их с основными носителями заряда (электронами n области) Протекающий при этом ток I_{m2} , связанный с движением инжектированных неосновных носителей в базе и накопленных там, намного превышает тепловой обратный ток I_0 . Соответственно, сопротивление диода при обратном включении оказывается малым и он, с этой точки зрения, по своим свойствам похож на диод при прямом включении. После того как накопленный заряд рассосался (рис. 1.29в), после времени t_3 , диод переходит в состояние, характеризуемое прямым напряжением, малым обратным током I_0 и большим дифференциальным сопротивлением. Этап восстановления обратного дифференциального сопротивления диода совершается за время $t_{ВОССТ}$ называемое временем восстановления. Длительность этапа рассасывания t_{PAC} прямо пропорциональна времени жизни дырок в базе и зависит от соотношения прямого I_{m1} и обратного I_{m2} тока. С ростом I_{m1} заряд, накопленный в базе растет, что при неизменном токе I_{m2} приводит к увеличению t_{PAC} . Длительность этапа восстановления зависит от времени жизни дырок в базе и от величины

барьерной емкости. Чем барьерная емкость больше, тем медленнее спадает ток в течение этого этапа. Барьерную емкость снижают за счет уменьшения площади перехода.

Длительность процессов накопления, рассасывания носителей заряда и восстановления обратного сопротивления диода пропорциональна времени жизни неосновных носителей заряда в базе диода. Поэтому при изготовлении быстродействующих импульсных диодов снижают время жизни вводя в базу «ловушки рекомбинации» легированием полупроводника специальными примесями. Обычно для этого используют золото. Следует иметь в виду, что при большой концентрации золота растут сопротивления базы и обратный ток. Поэтому все быстродействующие диоды характеризуются повышенными обратными токами.

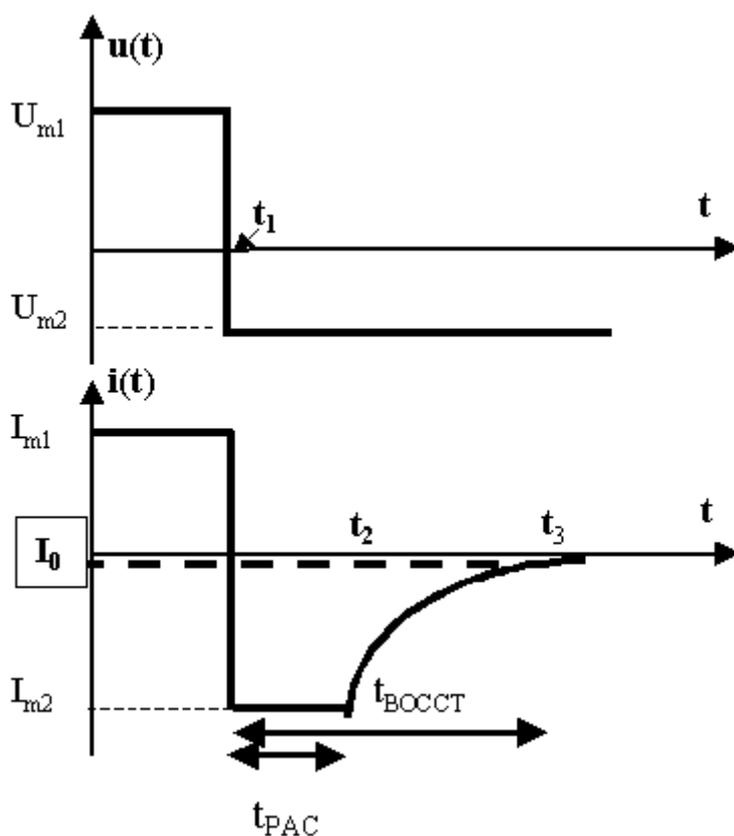


Рис.1.30. Временные зависимости напряжения и тока, протекающего через диод, при мгновенном переключении его с прямого на обратное включение

Нечто подобное имеет место и при выпрямлении диодом гармонических напряжений высокой частоты. На таких частотах, дырки, инжектированные в базу за положительный полупериод напряжения, не ус-

певают полностью рассосаться и рекомбинировать в базе. Во время отрицательного полупериода часть накопленных в n базе дырок уходит в p область, создавая «всплеск» обратного тока.

В заключение отметим следующее. В диодах Шоттки накопление носителей в базе отсутствует. Поэтому его импульсные и высокочастотные характеристики определяются только барьерной емкостью и временем пролета электронов через высокоомный слой кремния (примерно 10^{-11} с). В этой связи диоды Шоттки по своим характеристикам оказываются лучшими в сравнении с диодами других типов при работе с короткими импульсами и при выпрямлении высокочастотных сигналов.

Контрольные вопросы

1. Охарактеризуйте понятия: проводящие, диэлектрические и полупроводниковые материалы.
2. Какие основные полупроводниковые материалы Вы знаете? Перечислите их основные качества и особенности.
3. Какие полупроводники называют собственными, а какие примесными? В чем их отличие и основные свойства?
4. Определите понятия: электрон и дырка, электронно-дырочная пара.
5. Определите понятия: генерация и рекомбинация носителей зарядов в полупроводнике.
6. Запишите и поясните закон действующих масс для собственного полупроводника.
7. Определите понятие энергетической диаграммы полупроводника.
8. Охарактеризуйте понятия: «потолок валентной зоны», «дно зоны проводимости».
9. Нарисуйте и поясните энергетические диаграммы собственного полупроводника в одномерном приближении при температуре абсолютного нуля и ($T > 0^{\circ}$ К).
10. Какой процесс называется «временем жизни» электрона?
11. Какими электрофизическими параметрами характеризуют собственный полупроводник?
12. Как зависит собственная концентрация электронов в полупроводнике от ширины запрещенной зоны и температуры?

13. Охарактеризуйте понятия донорной и акцепторной примеси в полупроводнике.
14. Нарисуйте и поясните энергетические диаграммы примесного полупроводника при легировании донорной и акцепторной примесью ($T > 0^{\circ}$).
15. Какой полупроводник называют примесным полупроводником n – типа?
16. Какой полупроводник называют примесным полупроводником p – типа?
17. Охарактеризуйте понятия: основные и неосновные носители зарядов.
18. Сформулируйте закон действующих масс для примесных полупроводников.
19. Определите понятия: дрейф электронов в полупроводнике, дрейфовая скорость электронов.
20. Запишите и поясните выражение для плотности дрейфового тока.
21. Охарактеризуйте понятие подвижности носителей заряда в полупроводнике.
22. Запишите и поясните дифференциальную форму закона Ома.
23. Определите понятия диффузии зарядов и диффузного тока в полупроводниках.
24. Сформулируйте закон Фика.
25. Как связаны параметры дрейфового и диффузионного движения?
26. Перечислите основные структуры полупроводниковой электроники.
27. Определите понятия гомотипной и гетеротипной, симметричной и несимметричной структур.
28. Охарактеризуйте понятие контактной разности потенциалов между p и n областями.
29. Запишите и проанализируйте выражение, определяющее контактную разность потенциалов на $p - n$ переходе.
30. Охарактеризуйте прямое включение идеального диода.
31. Охарактеризуйте обратное включение идеального диода.
32. Запишите и поясните граничные условия Шокли.
33. Поясните физический смысл понятия инжекции (впрыскивания) неосновных носителей.
34. Охарактеризуйте понятие: диффузионная длина дырок и электронов.

35. Поясните физический смысл понятия экстракции (вытягивания) неосновных носителей.
36. Охарактеризуйте понятие: тепловой ток при обратном включении диода.
37. Нарисуйте и поясните ВАХ диода при прямом включении.
38. Определите понятие диффузной емкости.
39. Определите понятие барьерной емкости электронно-дырочной структуры. Нарисуйте и поясните зависимость барьерной емкости от напряжения.
40. Охарактеризуйте понятия: термодинамическая работа выхода и внешняя работа выхода.
41. Нарисуйте и поясните энергетическую диаграмму структуры металл – полупроводник в состоянии равновесия.
42. Нарисуйте и поясните вольт-амперные характеристики электронно-дырочной структуры и идеального диода Шоттки.
43. Какие существуют способы формирования омических контактов?
44. Назовите основные подклассы диодов.
45. Охарактеризуйте понятия: допустимое и предельное значение параметров диода.
46. Перечислите основные отличия вольт-амперных характеристик реальных диодов от идеальных.
47. Что означают понятия: туннельный, тепловой и лавинный пробой?
48. Какое устройство называют выпрямительным диодом? Для чего он служит?
49. Какое устройство называют стабилитроном? Для чего он служит?
50. Какое устройство называют варикапом? Для чего он служит?
51. Назовите назначение и принцип работы импульсных диодов.

2. Биполярные и полевые транзисторы

В предыдущем разделе мы начали изучение нового класса широко распространенных изделий – полупроводниковых приборов. Вспомнив основные положения физики полупроводников, мы изучили те простейшие структуры, которые лежат в основе функционирования многих приборов: соединение двух полупроводников с различным типом электропроводности и металл – полупроводник. Были рассмотрены различного рода диоды на основе этих простейших структур, проанализированы их параметры и характеристики.

Вместе с тем подобного рода структуры служат составной частью более сложных полупроводниковых приборов. Структура на основе трех полупроводниковых областей с двумя взаимодействующими $p - n$ переходами получила название биполярного транзистора. Термин «биполярный» связан с тем, что в работе транзистора используются носители зарядов двух типов: дырки и электроны. Биполярные транзисторы, являясь миниатюрными изделиями, получили широкое распространение в первую очередь благодаря тому, что они позволяют усиливать электрический сигнал по мощности. Хорошо они работают и в цифровых схемах, осуществляя функции ключа, который то замыкается, то размыкается под воздействием уже не механического воздействия, а электрического управляющего сигнала. Биполярные транзисторы позволяют хорошо согласовывать источники сигналов, обладающие малым внутренним сопротивлением, с высокоомной цепью нагрузки. За создание биполярного транзистора в 1956 году Шокли, Бардин и Браттейн были удостоены Нобелевской премии по физике.

Многие годы биполярный транзистор продолжал совершенствоваться, создавались все новые и новые полупроводниковые приборы. Вместе с тем все острее была необходимость в полупроводниковых приборах, которые бы могли, осуществляя те же функции, работать на основных носителях. Это давало бы ряд преимуществ, из которых главным была более высокая температурная стабильность.

Следующим шагом стало создание униполярных или, как их сейчас чаще называют, полевых транзисторов. В англоязычной литературе для обозначения таких транзисторов используют сокращение – FET – Field Effect Transistor, т. е. прибор, управляемый полем. В полевых транзи-

сторах ток основных носителей заряда протекает по так называемому каналу, образованному внутри кристалла полупроводника. Этим током можно управлять, прикладывая напряжение к управляющему электроду. В настоящее время благодаря ряду преимуществ полевые транзисторы заметно потеснили биполярные транзисторы, особенно в области цифровой и энергетической (силовой) электроники. Как правило, на полевых транзисторах создают компьютеры, микропроцессорные системы, устройства обработки электрических сигналов.

В этом разделе рассматриваются модели транзисторов, с помощью которых можно отразить принципы построения и механизмы функционирования прибора, изучаются физические явления, основные параметры и характеристики наиболее распространенных типов транзисторов. Описывается работа транзисторов для усиления электрических сигналов по мощности и особенности их функционирования в электронных цифровых схемах.

2.1. Общие сведения о биполярных транзисторах

2.1.1. Назначение и структура (устройство) биполярного транзистора

Биполярным транзистором (БТ) называют трехполюсный полупроводниковый прибор, в основу принципа действия которого положено взаимодействие чередующихся по типу проводимости электронно – дырочных структур, и предназначенный для осуществления следующих функций:

1. Усиления электрических сигналов по мощности.
2. Коммутации (переключения, подключения) участков электрической цепи.
3. Регулирования величины протекающего через два его вывода тока.

На основе БТ создаются усилители, генераторы гармонических и импульсных сигналов, логические схемы управления, регулируемые стабилизаторы напряжения, тока и многие другие электронные изделия.

БТ, если не показывать омические контакты, могут выполняться в виде следующих электронно-дырочных структур, простейшие одномерные модели которых показаны на рис.2.1.

БТ включает в себя три несимметричные, с точки концентрации введенной примеси, полупроводниковые области с чередующимися типами проводимости, либо $p^{++} - n - p^+$, либо $n^{++} - p - n^+$, которые называют соответственно эмиттером, базой и коллектором. Эти области разделены между собой обедненными слоями. Слой между эмиттером и базой называют эмиттерным $p - n$ переходом (ЭП), а между базой и коллектором - коллекторным $p - n$ переходом (КП). К внешним областям эмиттера (Э), базы (Б) и коллектора (К) присоединены металлические электроды (выводы), на которые подаются напряжения. В упрощенной модели $p - n$ переходы считаются плоскими, а носители движутся только в одном направлении – вдоль горизонтальной оси, перпендикулярной к переходам, которую мы будем именовать осью X.

Важную роль в транзисторе играет центральная область полупроводниковой структуры, которую называют базой (base – база, основание). Штриховой линией на модели показаны обедненные слои $p - n$ переходов. В этой связи следует различать физическую толщину базы, как расстояние между обедненными слоями, и технологическую толщину базы, как расстояние между металлургическими границами.

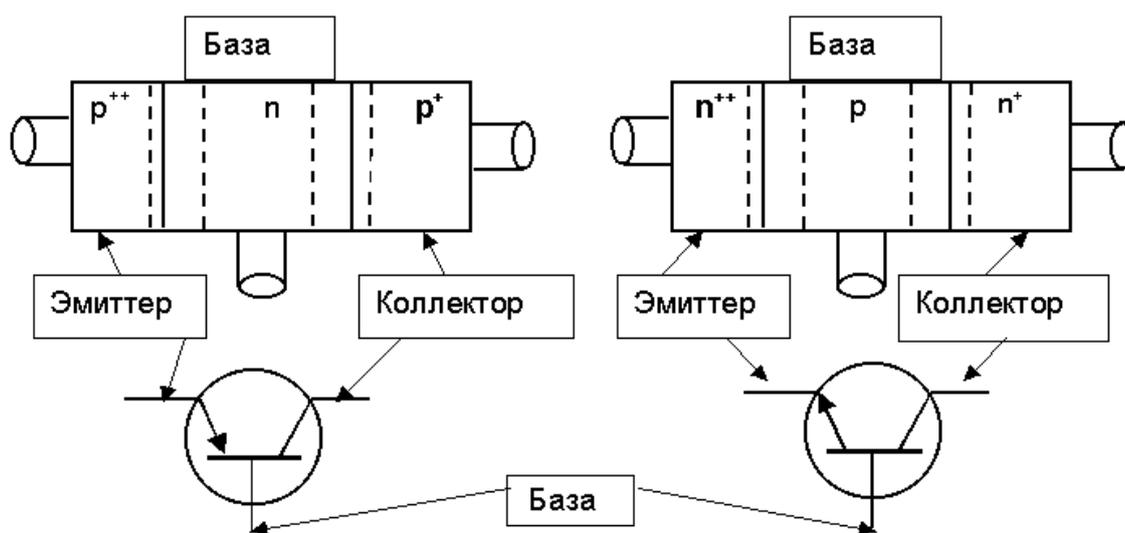


Рис. 2.1. Структуры $p^{++} - n - p^+$ и $n^{++} - p - n^+$ биполярных транзисторов и их условные графические изображения на схемах

Упрощенные модели структур $p^{++} - n - p^+$ либо $n^{++} - p - n^+$ транзистора, показанные на рис. 2.1, не отражают и тот факт, что толщина базы очень мала. Расстояние между обедненными слоями (физическая толщина базы) много меньше диффузионной длины неосновных носителей в базе, благодаря чему обеспечивается взаимодействие между потоками носителей ЭП и КП. Другими словами – при движении носителей заряда в базе вероятность их рекомбинации будет мала. Область коллектора, как мы увидим в дальнейшем, по соображениям лучшего отвода тепла, имеет самые большие геометрические размеры.

Таким образом, следует помнить, что, хотя это, как правило, не отражено на рисунках, поясняющих устройство транзистора, БТ это прибор существенно несимметричный, как с точки концентрации введенной в его области примеси, так и с точки зрения геометрических размеров его областей.

Условные графические изображения $p^{++} - n - p^+$ и $n^{++} - p - n^+$ показаны на рис. 2.1. Отметим пока тот факт, что стрелка на условном изображении транзистора направлена от p области к n ЭП. Это следует запомнить.

В настоящее время большинство выпускаемых промышленностью транзисторов имеет структуру $n^{++} - p - n^+$. Однако мы, преимущественно, будем рассматривать $p^{++} - n - p^+$ транзистор, что связано с тем, что в этом случае направление интересующих нас потоков носителей заряда и направления токов транзистора совпадают. Это проще для понимания основных процессов и не более. Однако все выводы теории в равной степени справедливы и для $n^{++} - p - n^+$ транзистора. Вы легко сможете распространить все сказанное для $p^{++} - n - p^+$ на $n^{++} - p - n^+$ транзистор.

2.1.2. Схемы включения биполярного транзистора

БТ имеет 3 вывода, поэтому это трехполюсный компонент электрической цепи. Однако, как уже отмечалось, его можно легко преобразовать в четырехполюсной компонент электрической цепи (компонент цепи класса SISO: single inputs – один вход – single outputs –один выход), сделав из одного вывода два. Тогда, в зависимости от того, какой вывод мы как бы делим на два и делаем общим, можно говорить о следующих схемах включения БТ.

Если базовый вывод транзистора является общим для входной (эмиттерной) и выходной (коллекторной) цепи, то говорят, что транзистор включен по схеме с общей базой (или, кратко, СОБ). Аналогично рассуждая, можно говорить о том, что транзистор может быть включен по схеме с общим эмиттером (СОЭ) и схеме с общим коллектором (СОК) (рис. 2.2).

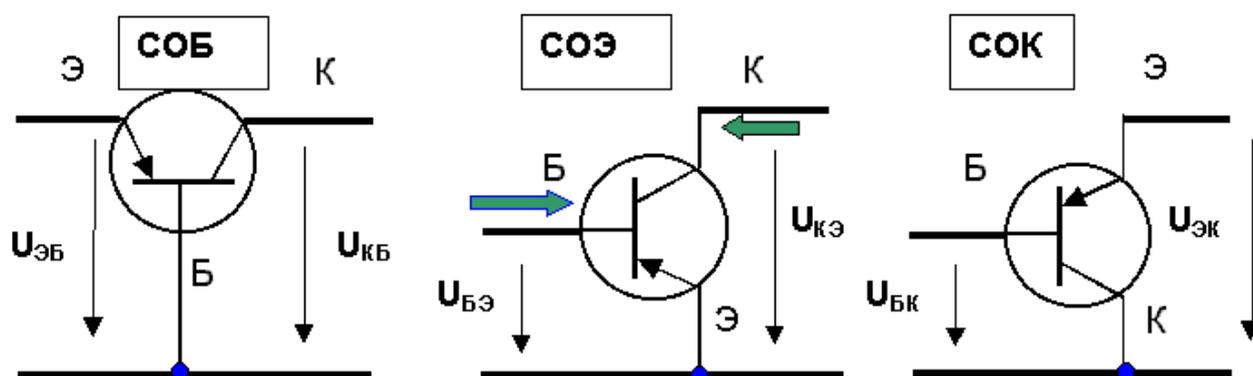


Рис. 2.2. Схемы включения $p^{++} - n - p^{+}$ биполярного транзистора

Следует привыкнуть к обозначениям напряжений в таких схемах: нижние индексы показывают, между какими выводами определяется напряжение. Напряжение отсчитывается относительно общего вывода, поэтому для СОБ $U_{ЭБ}$ – напряжение между эмиттером и базой, $U_{КБ}$ – напряжение между коллектором и базой. Поскольку напряжение это есть разность потенциалов, то можно также говорить, что $U_{ЭБ}$ – это разность потенциалов между эмиттером и базой и т. д. Аналогично определяются напряжения для СОЭ и СОК (рис. 2.2).

Надо помнить, что согласно теории четырехполюсника токам придают положительное значение в том случае, если они втекают внутрь четырехполюсника (см. рис. 2.2, зачерненные стрелки на СОЭ).

2.1.3. Электрические режимы работы биполярного транзистора

Электрическим режимом работы БТ называют совокупность условий, определяющих состояние и функционирование транзистора.

Электрический режим работы БТ однозначно определяется полярностью (знаком) и величиной потенциалов на входных и выходных выводах транзистора (напряжениями на входных и выходных электродах транзистора относительно общего вывода), включенного по СОБ.

Для $p^{++} - n - p^+$ БТ, в зависимости от $U_{ЭБ}$ и $U_{КБ}$, возможны рабочие электрические режимы названия которых приведены в табл. 2.1.

Таблица 2.1

Электрические режимы $p^{++} - n - p^+$ БТ

| № | $U_{ЭБ}$ | $U_{КБ}$ | Название электрического режима |
|----|---|--|--|
| 1 | $\begin{cases} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{cases}$ | $\begin{cases} U_{КБ} \leq 0 \\ U_{КБ} < U_{КБПР} \end{cases}$ | Активный (режим усиления электрического сигнала по мощности) |
| 2А | $\begin{cases} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{cases}$ | $\begin{cases} U_{КБ} > 0 \\ U_{КБ} \sim 0.4B < U_{КБ}^{жс} \end{cases}$ | Режим насыщения, который по своим свойствам близок к активному |
| 2Б | $\begin{cases} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{cases}$ | $\begin{cases} U_{КБ} > 0 \\ U_{КБ} \sim U_{КБ}^{жс} \end{cases}$ | Режим глубокого насыщения (неосновными носителями базы) |
| 3 | $\begin{cases} U_{ЭБ} < 0 \\ U_{ЭБ} < U_{ЭБПР} \end{cases}$ | $\begin{cases} U_{КБ} < 0 \\ U_{КБ} < U_{КБПР} \end{cases}$ | Режим отсечки (токов транзистора) |

Символами $U_{ЭБ}^{жс}$ и $U_{КБ}^{жс}$ в таблице 2.1 обозначаются пороговые напряжения эмиттерного и коллекторного переходов, соответственно. Символами $|U_{ЭБПР}|$ и $|U_{КБПР}|$ обозначены напряжения при котором происходит электрический пробой $p - n$ структуры. Так как эти напряжения отрицательны, то значения надо брать по модулю.

Для на $n^{++} - p - n^+$ БТ, в зависимости от $U_{ЭБ}$ и $U_{КБ}$, возможны рабочие электрические режимы названия которых приведены в табл. 2.2.

Если БТ включен по СОЭ, то для определения режима необходимо предварительно вычислить величины $U_{ЭБ}$ и $U_{КБ}$, а, затем, по таблицам

2.1 или 2.2 определить режим работы. Для нахождения $U_{ЭБ}$ и $U_{КБ}$ следует использовать формулы:

$$U_{ЭБ} = -U_{БЭ} , \quad U_{КБ} = U_{КЭ} - U_{БЭ} . \quad (2.1)$$

Таблица 2.2

Электрические режимы на $n^{++} - p - n^{+}$ БТ

| № | $U_{ЭБ}$ | $U_{КБ}$ | Название электрического режима |
|----|---|--|--|
| 1 | $\left\{ \begin{array}{l} U_{ЭБ} < 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}$ | $\left\{ \begin{array}{l} U_{КБ} \geq 0 \\ U_{КБ} < U_{КБПР} \end{array} \right\}$ | Активный (режим усиления электрического сигнала по мощности) |
| 2А | $\left\{ \begin{array}{l} U_{ЭБ} < 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}$ | $\left\{ \begin{array}{l} U_{КБ} < 0 \\ U_{КБ} \sim 0.4B < U_{КБ}^{жс} \end{array} \right\}$ | Режим насыщения, который по своим свойствам близок к активному |
| 2Б | $\left\{ \begin{array}{l} U_{ЭБ} < 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}$ | $\left\{ \begin{array}{l} U_{КБ} < 0 \\ U_{КБ} > U_{КБ}^{жс} \end{array} \right\}$ | Режим глубокого насыщения (неосновными носителями базы) |
| 3 | $\left\{ \begin{array}{l} U_{ЭБ} > 0 \\ U_{ЭБ} < U_{ЭБПР} \end{array} \right\}$ | $\left\{ \begin{array}{l} U_{КБ} > 0 \\ U_{КБ} < U_{КБПР} \end{array} \right\}$ | Режим отсечки (токов транзистора) |

В том случае, когда эмиттер БТ соединен с точкой нулевого потенциала $\varphi_{Э} = 0$, то напряжения $U_{БЭ}$ и $U_{КЭ}$ есть не что иное, как потенциалы. В этом случае:

$$U_{ЭБ} = -U_{Б0} = -\varphi_{Б} , \quad (2.2)$$

$$U_{КБ} = U_{КЭ} - U_{БЭ} = U_{К0} - U_{Б0} = \varphi_{К} - \varphi_{Б} . \quad (2.3)$$

Следует иметь в виду, что когда транзистор находится в активном режиме, он может быть в *нормальном* или в *инверсном включении*.

Инверсное включение (включение наоборот) это включение БТ, когда эмиттер и коллектор меняют местами. В активном инверсном режиме работы $p^{++} - n - p^+$ БТ:

$$\left\{ \begin{array}{l} U_{ЭБ} \leq 0 \\ |U_{ЭБ}| < |U_{ЭБПР}| \end{array} \right\}, \left\{ \begin{array}{l} U_{КБ} > 0 \\ U_{КБ} \sim U_{КБ}^{жс} \end{array} \right\}. \quad (2.4)$$

2.1.4. Что значит объяснить принцип действия транзистора

Для того, чтобы в дальнейшем вы понимали о чем пойдет речь, следует запомнить, что для объяснения принципа действия идеализированного БТ в одном из режимов, надо пояснить следующее.

1. Какие физические явления и процессы происходят в области Э и на ЭП? Какой величины и какого направления протекает ток через вывод Э?

2. Какие физические явления происходят в К (Б) и на КП? Какой величины и какого направления протекает ток через вывод К (Б)?

3. Как соотносятся между собой токи, протекающие через выводы БТ (Э, К, Б) и какой величины при этом сопротивления p и n областей, эмиттерного и коллекторного переходов.

Кроме того, при объяснении работы БТ, как вытекает из предыдущего раздела, необходимо помнить следующее.

А. Сопротивление областей полупроводника, прилегающих к обедненным слоям ЭП и КП, мало, поэтому внутри этих областей не бывает электрических полей.

Б. Обедненные слои ЭП и КП, как раз наоборот, обеднены подвижными носителями заряда, их сопротивление велико, в них есть внутреннее электрическое поле и внешнее электрическое поле внутри них взаимодействует с внутренним.

2.2. Принцип действия биполярного транзистора в различных электрических режимах их характеристики его работы

2.2.1. Функционирование биполярного транзистора в активном режиме

Активным режимом (АР) работы $p^{++} - n - p^+$ БТ называется такое его состояние, при котором на эмиттер относительно базы подан положительный потенциал порядка величины порогового напряжения, а на коллектор относительно базы – отрицательный или равный нулю и при этом напряжение U_{KB} меньше напряжения электрического пробоя. Математически это можно записать так:

$$\left\{ \begin{array}{l} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}, \left\{ \begin{array}{l} U_{KB} \leq 0 \\ |U_{KB}| < |U_{KBПР}| \end{array} \right\}. \quad (2.5)$$

При этом в БТ будут происходить следующие явления и процессы.

1. Поскольку несимметричный $p^{++} - n$ ЭП оказывается при прямом включении, то в нем, как следует из предыдущего раздела (п.1.3.3), будет наблюдаться односторонняя (преимущественная) инжекция дырок из эмиттера в базу (рис. 2.3).

На правой границе ЭП (ПрЭП), согласно граничных условий Шокли (1.28) и (1.29), концентрация дырок по сравнению с равновесной будет резко возрастать:

$$p_n^{ПрЭП}(U_{ЭБ} > 0) = p_{n0}^{ПрЭП} \ell \frac{U_{ЭБ}}{\varphi_T} \quad (2.6)$$

и возникнет ток дырочная компонента тока I_p .

Будет иметь место также и инжекция электронов из Б в Э, но, так как концентрация акцепторной примеси намного больше донорной, то количество неосновных носителей в области p будет совсем мало:

$$n_p^{ЛГЭП}(U_{ЭП} > 0) = n_{p0}^{ЛГЭП} \ell \frac{U_{ЭП}}{\varphi_T}. \quad (2.7)$$

Понятно, что поток электронов будет существенно меньшим и ток I_n будет мал.

Так как дырки из эмиттера уходят и в него входит немного электронов, то нарушается электронейтральность Э. Поэтому, чтобы Э по-

прежнему оставался электронейтральным, вступает в действие механизм компенсации заряда, обеспечивающий протекание тока через внешний вывод Э. В первом приближении его работу можно пояснить следующим образом.

Чтобы p^{++} область оставалась электронейтральной из-за инжекции дырок, из p^{++} области в металлический вывод, имеющий положительный потенциал, уходит поток электронов, имеющий такой же самый заряд. Покидают Э через металлический вывод также электроны, инжектированные из Б в Э. В результате электронейтральность восстанавливается и через металлический вывод эмиттера начнет протекать ток, который направлен внутрь транзистора (показан на рис. 2.3 зачерненной стрелкой).

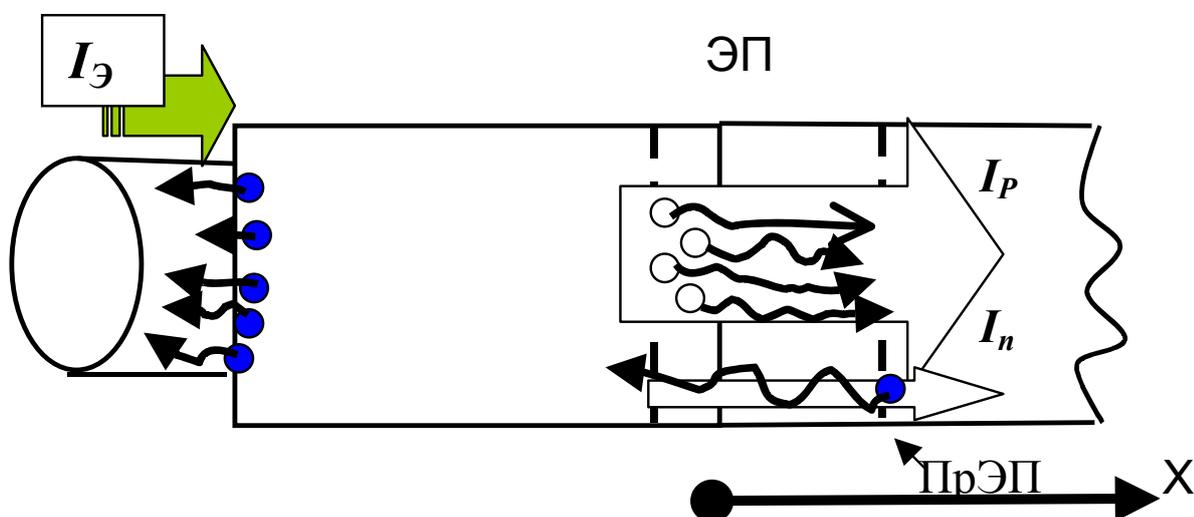


Рис. 2.3. Фрагмент эмиттера и эмиттерного перехода структуры $p^{++} - n - p^{+}$ биполярного транзистора и протекающие при этом токи

Ток эмиттера относительно большой, так как речь идет об инжекции фактически большого потока основных носителей заряда. Понятно, что величина тока эмиттера определяется величиной приложенного напряжения $U_{ЭБ}$. Эта зависимость экспоненциальная.

2. Концентрация дырок у правой границы ЭП, из-за инжекции, велика. У левой границы КП она мала (примем пока на веру). Следовательно, вдоль пространства базы имеет место разность концентраций неосновных носителей - дырок - и в Б будет наблюдаться явление диф-

фузии дырок и, соответственно, будут перемещения их от Э к К (рис. 2.4).

Помимо основного процесса диффузии в базе будет также наблюдаться процесс рекомбинации движущихся неосновных носителей (воссоединение дырок и электронов). Так как база очень тонкая (намного меньше диффузионной длины), то вероятность рекомбинации будет малой. Но, тем не менее, не все дырки «дойдут до КП, отдельные из них все же прорекомбинируют (рис. 2.4).

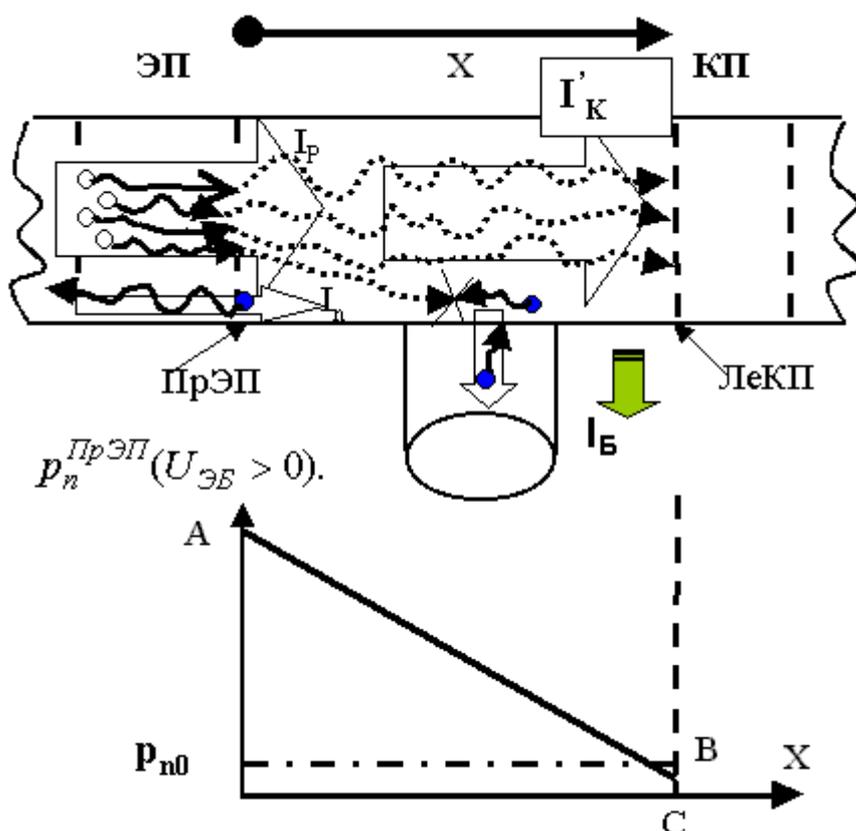


Рис. 2.4. Фрагмент базы и коллекторного перехода структуры $p^{++} - n - p^+$ биполярного транзистора и протекающие при этом токи

Чтобы обеспечивалась электронейтральность базы, для участия в рекомбинации через металлический вывод базы внутрь транзистора входят электроны. Ток базы направлен из транзистора и, так как рекомбинация мала, намного меньше тока эмиттера и коллектора.

Для того, чтобы можно было оценивать, в первом приближении, процессы, происходящие в базе, под фрагментом изображения структуры транзистора построен график зависимости концентрации неосновных

носителей (дырок) в базе транзистора от координаты X . Горизонтальной штрих – пунктирной линией на нем отмечено распределение неосновных носителей в базе в равновесном состоянии (при отсутствии напряжений на транзисторе неосновные носители распределены в базе равномерно по пространству). Ордината графика при $X = 0$ (точка А) отражает величину концентрации неосновных носителей на правой границе эмиттерного перехода при приложении напряжения между эмиттером и базой. Она определяется (2.6). Ордината графика при $X=X_B$ (точка В) отражает величину концентрации неосновных носителей на левой границе коллекторного перехода при приложении отрицательного напряжения между коллектором и базой:

$$p_n^{LeKP} (U_{KB} > 0) = p_{n0}^{LeKP} \ell \frac{-U_{KB}}{\varphi_T} . \quad (2.8)$$

Практически эта ордината равна нулю (точка С). Наклон линии графика, как это следует из (1.14), пропорционален току коллектора со штрихом I'_K . Площадь под линией распределения (трапеция OABC) пропорциональна вероятности рекомбинации, а, следовательно, и току базы.

Следует иметь в виду, что чем больше напряжение $U_{ЭБ}$, тем, согласно (2.6), выше ордината графика (точка А), тем больше площадь под кривой, тем выше вероятность рекомбинации и, тем самым, больше ток базы. Чем больше напряжение $U_{ЭБ}$, и выше ордината графика (точка А), тем больше наклон линии и тем больше, согласно (1.14), ток I'_K (ток коллектора).

3. Так как несимметричный $n - p^+$ КП смещен в обратном направлении, то в нем, как следует из предыдущего раздела (п.1.3.4), имеет место односторонняя экстракция неосновных носителей заряда (главным образом, вытягивание дырок из базы в коллектор).

Дырки, достигшие коллекторного перехода за счет диффузии (перемещения) неосновных носителей по базе втягиваются электрическим полем КП (как говорят кратко, экстрагируют) и перебрасываются в коллектор (рис. 2.5).

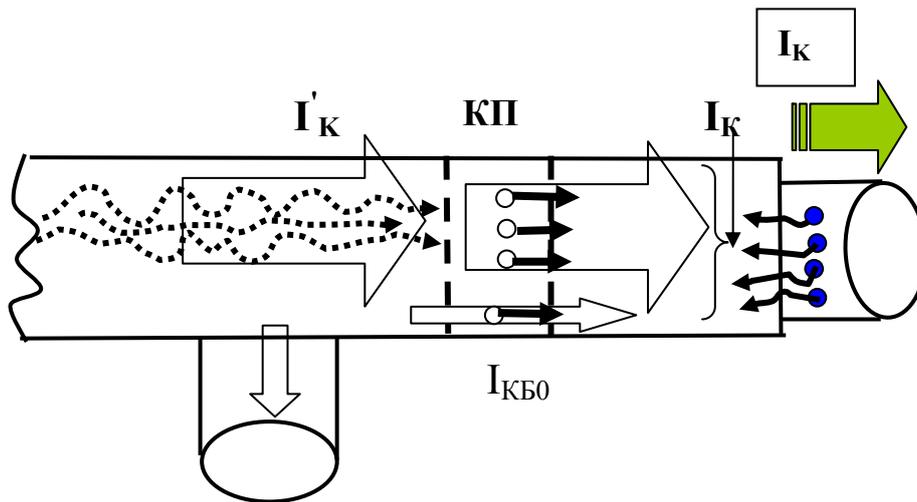


Рис. 2.5. Фрагмент базы и коллектора структуры $p^{++} - n - p^{+}$ биполярного транзистора и протекающие при этом токи

Следует иметь в виду, что помимо дырок, движущихся за счет диффузии, в базе имеются дырки, которые возникают вследствие процессов термогенерации неосновных носителей. Эти дырки электрическое поле КП также экстрагирует и перебрасывает в коллектор. Кроме того, уже отмечалось, что и в самом коллекторном переходе, как в обедненном слое кремниевого перехода, могут возникать носители за счет процессов термогенерации. Поэтому наряду с током коллектора со штрихом I'_K , обусловленным диффузионным движением носителей по базе, через коллекторный переход будет протекать еще и обратный ток КП $I_{КБО}$ подобно тому, как это имело место при обратном включении электронно – дырочной структуры. Обратный ток коллекторного перехода $I_{КБО}$ будет состоять их двух компонент: тока, возникающего за счет экстракции неосновных носителей, термогенерированных в базе; тока, обусловленного термогенерацией неосновных носителей в самом КП. Тем самым, коллектор БТ собирает (коллектирует) дырки, инжектированные в базу ЭП и дошедшие без рекомбинации до КП, а также дырки, обусловленные обратным током КП.

Так как дырки в большом количестве поступают в коллектор, то нарушается его электронейтральность. Поэтому, чтобы К по-прежнему оставался электронейтральным, вступает в действие механизм компенсации заряда, обеспечивающий протекание тока через внешний вывод К. В первом приближении его работу можно пояснить следующим образом.

Чтобы p^+ область оставалась электронейтральной из-за экстракции дырок, в p^+ область из металлического вывода, имеющего отрицательный потенциал, входит поток электронов, имеющий такой же самый заряд, и компенсирует заряд дырок. В результате через металлический вывод коллектора начнет протекать ток, который направлен из транзистора (показан на рис. 2.5 зачерненной стрелкой) и относительно велик. Если процессы рекомбинации малы, то ток коллектора будет равен примерно току эмиттера.

2.2.2. Токи биполярного транзистора в активном режиме

Для того, чтобы можно было оценивать токи в БТ вводят ряд параметров.

Для выявления доли тока инжекции дырок, как «полезного» тока для БТ, в полном токе эмиттера, вводят коэффициент инжекции:

$$\gamma = \frac{I_p}{I_{\Sigma}} = \frac{I_p}{I_p + I_n} = \frac{1}{1 + I_n/I_p}. \quad (2.9)$$

Из (1.32) следует:

$$I_n = S \left(\frac{qD_n n_{p0}}{L_n} \right) \ell^{\frac{U_{ЭБ}}{\phi_T}}, I_p = S \left(\frac{qD_p p_{n0}}{L_p} \right) \ell^{\frac{U_{ЭБ}}{\phi_T}}, \quad (2.10)$$

поэтому $I_n/I_p \approx n_{p0}/p_{n0} \approx n_n/p_p \cong N_A/N_D$. Следовательно, для несимметричного $p^{++} - n$ ЭП получаем $\gamma \approx 0,99 - 0,999$. Коэффициент I_n/I_p называют «дефект инжекции эмиттера». Он определяет «бесполезную» долю тока в общем токе эмиттера.

Можно также записать:

$$I_p = I_{\Sigma} \gamma, I_n = I_{\Sigma} (1 - \gamma). \quad (2.11)$$

Для оценки степени преобладания в базе процесса диффузии над рекомбинацией вводят коэффициент переноса, который показывает, какая часть дырок, инжектированных из эмиттера в базу, достигает КП:

$$\eta = \frac{I'_K}{I_P}. \quad (2.12)$$

Тогда:

$$I_{PEK} = I_P(1 - \eta), I'_K = \eta I_P. \quad (2.13)$$

Ток коллектора можно записать в виде:

$$I_K = I'_K + I_{KB0}, \quad (2.14)$$

где I_{KB0} - обратный ток коллекторного перехода, состоящий в общем случае из теплового тока КП, тока термогенерации КП и тока утечки КП.

Из рис. 2.3, 2.4 и 2.5 следует $I_B = I_{PEK} + I_n - I_{KB0}$.

Тогда, используя (2.11), (2.13) и выполняя последовательно преобразования, получаем:

$$\begin{aligned} I_B &= (1 - \eta)I_P + (1 - \gamma)I_{\mathcal{E}} - I_{KB0} \Leftrightarrow I_B = I_P - \eta I_P + I_{\mathcal{E}} - \mathcal{I}_{\mathcal{E}} - I_{KB0} \Leftrightarrow \\ I_B &= I_P - I'_K + I_{\mathcal{E}} - I_P - I_{KB0} \Leftrightarrow I_B = -I'_K + I_{\mathcal{E}} - I_{KB0} \Leftrightarrow \\ I_B &= I_{\mathcal{E}} \left(1 - \frac{I'_K}{I_{\mathcal{E}}}\right) - I_{KB0} \Leftrightarrow I_B = I_{\mathcal{E}}(1 - \alpha_0) - I_{KB0}, \end{aligned} \quad (2.15)$$

где $\alpha_0 = \frac{I'_K}{I_{\mathcal{E}}} = \gamma\eta$ - статический (интегральный) коэффициент передачи эмиттерного тока к коллектору.

Окончательно из (2.15) получаем:

$$\begin{aligned} I_B &= I_{\mathcal{E}} - I_{\mathcal{E}}\alpha_0 - I_{KB0} \Leftrightarrow I_{\mathcal{E}} - I_B = I_{\mathcal{E}}\alpha_0 - I_{KB0} \Leftrightarrow \\ I_K &= \alpha_0 I_{\mathcal{E}} + I_{KB0}. \end{aligned} \quad (2.16)$$

Уравнение (2.16) называют *основным уравнением биполярного транзистора, работающего в активном режиме, в схеме с общей базой*.

Основное уравнение, теоретически, является линейным: в нем ток коллектора пропорционален току эмиттера. Кроме того, оно свидетельствует о том, что ток коллектора, теоретически, не зависит от напряжения $U_{КБ}$.

Так как $I_{Э} = I_{К} + I_{Б}$, то из (2.16) можно получить:

$$I_{К} = \alpha_0(I_{К} + I_{Б}) + I_{КБ0} \Leftrightarrow I_{К} - \alpha_0 I_{К} = \alpha_0 I_{Б} + I_{КБ0} \Leftrightarrow$$

$$I_{К} = \frac{\alpha_0}{1 - \alpha_0} I_{Б} + \frac{1}{1 - \alpha_0} I_{КБ0} \Leftrightarrow$$

$$I_{К} = \beta I_{Б} + (\beta + 1) I_{КБ0}, \quad (2.17)$$

где $\beta = \frac{\alpha_0}{1 - \alpha_0}$ - статический (интегральный) коэффициент передачи базового тока в СОЭ. Обычно β равен 40 – 200.

Уравнение (2.16) называют *основным уравнением биполярного транзистора, работающего в активном режиме, в схеме с общим эмиттером*.

Основное уравнение БТ в СОЭ, также теоретически, является линейным: в нем ток коллектора пропорционален току базы. Кроме того, оно свидетельствует о том, что ток коллектора, теоретически, не зависит от напряжения $U_{КЭ}$.

Второе слагаемое, входящее в уравнение (2.17), называют *сквозным током БТ*

$$I_{КЭ0} = (\beta + 1) I_{КБ0}. \quad (2.18)$$

Уравнение (2.17) свидетельствует, что $I_{КЭ0}$ это ток коллектора БТ, у которого «оторвана база» ($I_{Б} = 0$).

Очень важен и тот факт, что сквозной ток примерно в β (40 – 200) раз больше обратного тока коллекторного перехода $I_{КБ0}$.

Объяснение возрастания сквозного тока в СОЭ по сравнению с обратным током КП в СОБ (как следует из рис. 2.2, в обоих случаях это

выходные токи БТ при токах на входе равных нулю) состоит в следующем.

При «обрыве базы» электроны для рекомбинации также уже не могут «вводиться» в базу через базовый вывод. По этой причине восстановление электронейтральности базы не может осуществляться прежним образом и, казалось бы, потенциал базы должен приобретать положительный знак и расти. Но на самом деле так не происходит. В транзисторе начинает работать положительная обратная связь, за счет которой избыточный положительный заряд «выводится» из базы.

Дело в том, что напряжение $U_{кэ}$, распределяясь между КП и ЭП, вызывает появление дополнительного прямого напряжения на ЭП. В результате этого в ЭП появляется дополнительная инжекция и увеличивается «первичный» ток. Если исходить из идеализированной теории работы БТ, то обратный ток коллектора теперь может протекать только через вывод эмиттера и сквозь базу (а не через базовый вывод, как было раньше). Его поэтому и называли «сквозным». В результате ток, протекающий через КП, оказывается увеличенным в $\beta + 1$ раз.

Другими словами, восстановление электронейтральности при «обрыве» базы не может осуществляться иным способом, как за счет появления дополнительной инжекции через ЭП. Для ликвидации заряда одного прорекомбинировавшего в базе электрона требуется, чтобы через базу прошло, унося дополнительный положительный заряд, и было «переброшено» в коллектор в $\beta + 1$ раз большее число дырок, нежели это было раньше.

2.2.3. Признаки работы биполярного транзистора в активном режиме

Характерными признаками, по которым можно определить, что БТ работает в активном режиме, являются следующие.

1. Напряжение на ЭП имеет величину примерно равную пороговому напряжению электронно – дырочной структуры. Для кремниевого БТ $U_{эб} \sim U_{эб}^{жс} \approx (0,6 - 0,8)В$, для германиевого БТ $U_{эб} \approx (0,3 - 0,4)В$.

Напряжение на коллекторе относительно базы меньше пробивного, указанного для данного типа транзистора в справочнике. При работе БТ обычно оно составляет:

а) для маломощных БТ $U_{КБ} = (5 - 30)B$;

б) для мощных БТ $U_{КБ} = (5 - 100)B$;

в) для высоковольтных БТ $U_{КБ} = (100 - 1000)B$.

Напряжение на коллекторе относительно эмиттера примерно равно напряжению на коллекторе относительно базы $U_{КЭ} \approx U_{КБ}$.

2. Ток коллектора практически равен току эмиттера $I_K \cong I_{Э}$. Важно, что ток коллектора фактически не зависит от напряжения $U_{КБ}$, а определяется напряжением $U_{ЭБ}$. В связи с (2.15) и (2.16) говорят, что БТ управляется током.

Ток базы $I_B \approx I_K / \beta$, в первом приближении, примерно в 100 раз меньше тока коллектора и эмиттера.

3. Со стороны входа сопротивление БТ определяется *дифференциальным сопротивлением эмиттерного перехода*. Согласно (1.37) можно записать:

$$r_{Э} = \frac{dU_{ЭБ}}{dI_{Э}} = \frac{\varphi_T}{I_{Э}} \quad (2.19)$$

Входное сопротивление БТ, как было определено ранее, относительно мало.

Со стороны выхода сопротивление БТ определяется *дифференциальным сопротивлением коллекторного перехода*:

$$r_K = \frac{dU_{КБ}}{dI_K} = \frac{1}{\frac{dI_K}{dU_{КБ}}} \approx \frac{1}{0} \approx \infty. \quad (2.20)$$

Выходное сопротивление БТ в СОБ очень велико. Можно считать, что БТ в активном режиме «трансформирует» сопротивления. Этот факт лег в основу названия БТ. Слово «транзистор» состав-

лено из двух английских слов: TRANSfer – преобразовывать и reSISTOR – сопротивление.

4. Входная (потребляемая от источника) и выходная (отдаваемая в нагрузку) мощности равны:

$$\Delta P_{BX} = \Delta I_{\mathcal{E}} \Delta U_{\mathcal{EB}} = \Delta I_{\mathcal{E}}^2 R_{BX} = I_{\mathcal{E}}^2 r_{\mathcal{E}}, \quad (2.21)$$

$$\Delta P_{BIX} = \Delta I_K \Delta U_{KB} = \Delta I_K^2 R_H. \quad (2.22)$$

Следовательно, коэффициент усиления по мощности:

$$K_P = \frac{\Delta P_{BIX}}{\Delta P_{BX}} = \frac{R_H}{r_{\mathcal{E}}}. \quad (2.23)$$

БТ в активном режиме при $R_H > r_{\mathcal{E}}$ способен усиливать электрический сигнал по мощности. В других режимах этого не происходит, поэтому активный режим работы БТ является основным во всех схемах усилителей.

Если вспомнить классификацию зависимых источников, то можно утверждать, что БТ в активном режиме подобен источнику тока, управляемому током (повторителю тока).

5. В активном режиме на КП ток и напряжение велики, поэтому на нем выделяется существенная мощность, которая преобразовывается в тепло. Для того, чтобы транзистор не вышел из строя необходимо принимать меры по отводу этого тепла.

Следует подчеркнуть, что все время анализировался БТ, работающий при нормальном включении. Уже отмечалось, что возможен инверсный режим включения, когда коллектор инжектирует электроны, а эмиттер их собирает. Так как при этом существенно уменьшаются и коэффициент инжекции и коэффициент переноса, то ток коллектора будет составлять примерно 50 – 60% от тока эмиттера. Коэффициент передачи по току в инверсном режиме составляет 20 – 40 % от того же коэффициента при нормальном включении БТ.

Необходимо также обратить внимание, что в транзисторе всегда выполняются первый и второй закон Густава Роберта Кирхгофа.

Так как алгебраическая сумма токов, протекающих через замкнутую (разделяющую произвольную цепь на два фрагмента, внутренний и внешний) поверхность (показана на рис. 2.6 штрих пунктирной линией), будет равна нулю в любой момент времени, то для любой схемы включения БТ можем записать:

$$I_{\text{Э}} - I_{\text{К}} - I_{\text{Б}} = 0 \Rightarrow I_{\text{Э}} = I_{\text{К}} + I_{\text{Б}}. \quad (2.24)$$

Также, поскольку алгебраическая сумма ЭДС источников, включенных в ветвях любого контура электрической цепи (контур показан на рис. 2.6 штрих пунктирной линией), равна алгебраической сумме напряжений на всех остальных элементах того же контура в любой момент времени, то для любой схемы включения БТ при обходе контура против часовой стрелки можно записать:

$$U_{\text{ЭБ}} - U_{\text{КБ}} + U_{\text{КЭ}} = 0 \Rightarrow U_{\text{КЭ}} = -U_{\text{ЭБ}} + U_{\text{КБ}} \quad (2.25)$$

$$U_{\text{БЭ}} + U_{\text{КБ}} - U_{\text{КЭ}} = 0 \Rightarrow U_{\text{КЭ}} = U_{\text{БЭ}} + U_{\text{КБ}} \quad (2.26)$$

$$U_{\text{БК}} + U_{\text{ЭБ}} - U_{\text{ЭК}} = 0 \Rightarrow U_{\text{ЭК}} = U_{\text{БК}} + U_{\text{ЭБ}} \quad (2.27)$$

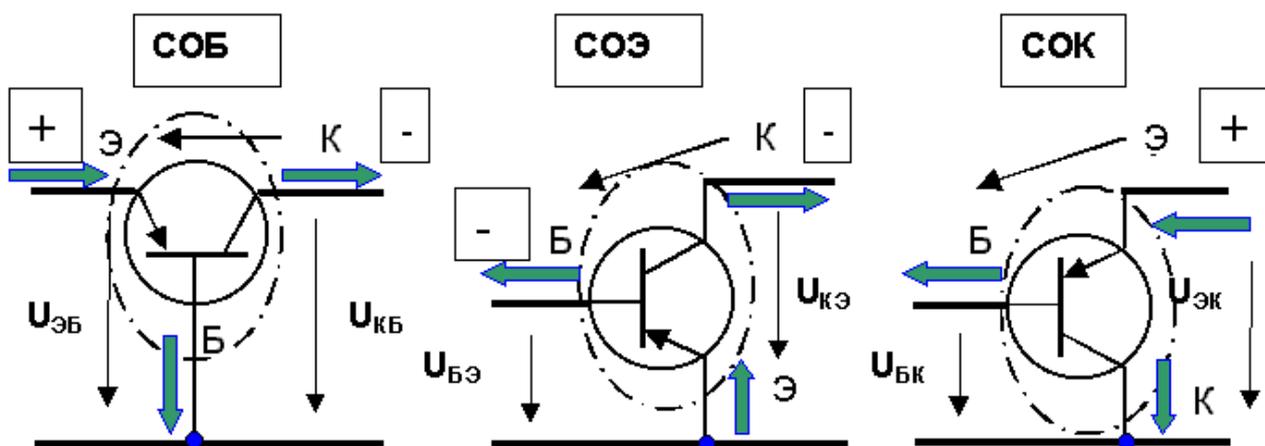


Рис. 2.6. Выполнение первого и второго закона Кирхгофа для схем включения $p^{++} - n - p^{+}$ биполярного транзистора

2.2.4. Функционирование биполярного транзистора в режиме насыщения

Режимом насыщения (НР) $p^{++} - n - p^{+}$ БТ называется такое его состояние, при котором на эмиттер относительно базы и на коллектор относительно базы *подан положительный потенциал*. Математически это можно записать так:

$$\left\{ \begin{array}{l} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}, \left\{ \begin{array}{l} U_{КБ} > 0 \\ U_{КБ} \leq U_{КБ}^{жс} \end{array} \right\}. \quad (2.28)$$

При этом в БТ будут происходить следующие явления и процессы.

1. Поскольку несимметричный $p^{++} - n$ ЭП оказывается при прямом включении, то в нем, как следует из (п.1.3.3), будет наблюдаться односторонняя (преимущественная) инжекция дырок из эмиттера в базу (рис. 2,7)

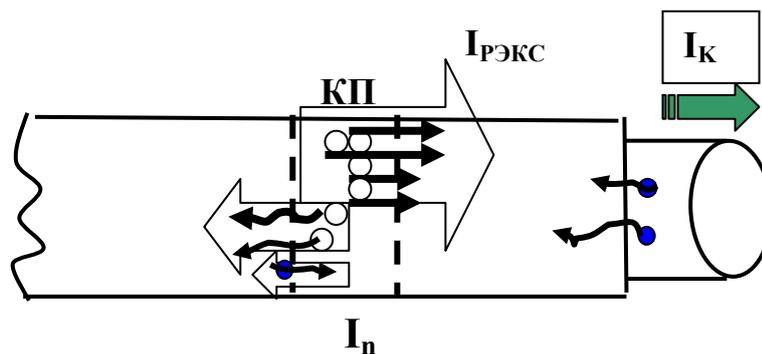


Рис. 2.7. Фрагмент базы и коллектора структуры $p^{++} - n - p^{+}$ биполярного транзистора в режиме насыщения и протекающие при этом токи

Так как дырки уходят из эмиттера, а в него входит немного электронов, то вступает в действие механизм компенсации заряда. Чтобы p^{++} область оставалась электронейтральной из-за инжекции дырок, из p^{++} области в металлический вывод, имеющий положительный потенциал, уходит поток электронов, имеющий такой же самый заряд. Покидают Э через металлический вывод также электроны, инжектированные из Б в Э. В результате электронейтральность восстанавливается и через металлический вывод эмиттера начнет протекать ток, который направлен внутрь транзистора (показан на рис. 2.3 зачерненной стрелкой). Ток

эмиттера относительно большой, его величина определяется значением приложенного напряжения $U_{ЭБ}$. Зависимость тока от напряжения экспоненциальная.

2. Чтобы упредить возможность неверного истолкования того, что происходит в КП и в К, нужно понимать, что в режиме насыщения КП присуща совокупность двух, происходящих одновременно, процессов: КП и инжектирует неосновные носители в базу и «собирает» подходящие к его левой границе неосновные носители, которые «дошли» до нее в результате диффузии (рис. 2.7).

Действительно, с одной стороны, поскольку несимметричный $n - p^+$ КП оказывается при прямом включении, то в нем, как следует из п. 1.3.3, будет наблюдаться односторонняя (преимущественная) инжекция дырок из коллектора в базу (рис. 2.7). С другой стороны, так как в обедненной области $n - p^+$ КП имеется электрическое поле, то в нем также имеет место односторонняя экстракция неосновных носителей заряда (главным образом, вытягивание дырок из базы в коллектор). Дырки, достигшие коллекторного перехода за счет диффузии, втягиваются электрическим полем КП и перебрасываются в коллектор (рис. 2.7). Какой процесс, экстракция или инжекция, преобладает в КП, а, соответственно, и какая компонента тока является преобладающей, зависит от того, как соотносятся между собой напряжения $U_{ЭБ}$ и $U_{КБ}$ и каково при этом распределение неосновных носителей в базе.

Уместно рассмотреть два случая.

А) Пусть соотношение напряжений таково

$$\left\{ \begin{array}{l} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}, \left\{ \begin{array}{l} U_{КБ} > 0 \\ U_{КБ} \sim 0.4B < U_{КБ}^{жс} \end{array} \right\}. \quad (2.28)$$

В этом случае напряжение на коллекторном переходе меньше порогового и можно полагать, что инжекция на КП практически отсутствует (ее уровень небольшой). Распределение носителей в базе подобно распределению в активном режиме, за исключением того, что, за счет небольшой инжекции, величина концентрации неосновных носителей заряда на левой границе КП (точка В) будет несколько больше равновесной (рис. 2.8).

Напомним, что на графике зависимости концентрации неосновных носителей (дырок) в базе транзистора от координаты X горизонтальной штрих – пунктирной линией отмечено распределение неосновных носителей в базе в равновесном состоянии (при отсутствии напряжений на транзисторе). Ордината графика при $X = 0$ (точка А) отражает величину концентрации неосновных носителей на правой границе ЭП при приложении напряжения между эмиттером и базой. Она определяется (2.6). Ордината графика при $X = X_B$ (точка В) отражает величину концентрации неосновных носителей на левой границе КП при приложении напряжения между К и Б. Эта ордината определяется формулой:

$$p_n^{LeKP} (U_{KB} > 0) = p_{n0}^{LeKP} \ell \frac{U_{KB}}{\varphi_T} . \quad (6.30)$$

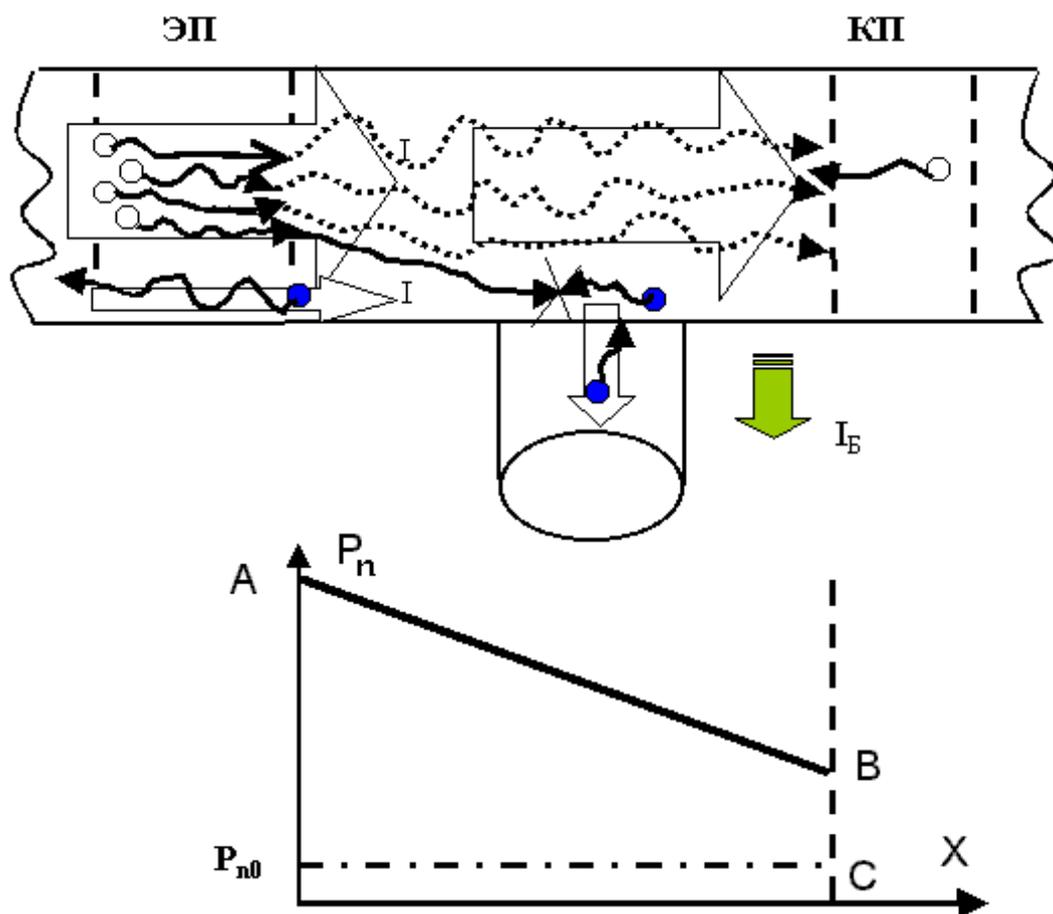


Рис. 2.8. Фрагмент базы и коллекторного перехода структуры $p^{++} - n - p^+$ биполярного транзистора в режиме насыщения, по свойствам близкого к активному режиму, и протекающие при этом токи

Наклон линии графика, как это следует из (1.14), пропорционален току коллектора со штрихом I'_K . Площадь под линией распределения (трапеция $OABC$) пропорциональна вероятности рекомбинации, а, следовательно, и току базы.

Сравнивая рис. 2.4 и рис. 2.8, можно утверждать следующее. Так как площадь под кривой в этом режиме, осталась практически такой же, как и в активном режиме, то примерно такой же величины, как и в активном режиме, остался ток базы. Наклон линии также остался почти прежним и, следовательно, почти не изменился ток I'_K (ток коллектора).

Процессы, происходящие в коллекторе и в базе по перемещению носителей, восстановлению их электронейтральности, как нетрудно догадаться, будут такими же, как на рис. 2.5.

То есть мы видим что, хотя БТ работает в режиме насыщения, в нем происходят, в первом приближении, те же процессы, что и в активном режиме. Примерно соблюдаются и соотношения токов. Поэтому такой режим получил название режима насыщения по свойствам близкого к активному режиму.

Работа БТ в таком режиме обычно происходит в цифровых электронных схемах. При этом, чтобы не наступала инжекция на коллекторном переходе, параллельно коллекторному переходу подключают фиксирующий диод Шоттки (рис. 2.9).

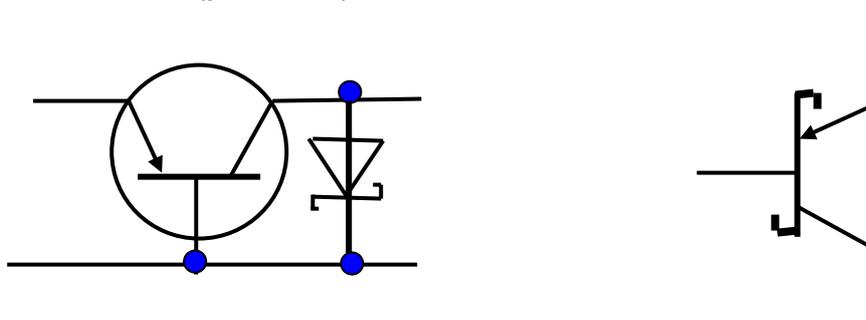


Рис. 2.9. Биполярный транзистор с фиксирующим диодом Шоттки и условное обозначение транзистора Шоттки

Для фиксирующего диода Шоттки пороговое напряжение обычно составляет $U_{AK} \sim U_{MeП}^{жс} \approx (0,3...0,4)В$. Для того, чтобы перевести КП в режим инжекции требуется $U_{KB} \sim U_{KB}^{жс} \approx (0,6...0,8)В$. Поэтому в схеме

БТ с фиксирующим диодом Шоттки, как только напряжение на коллекторе относительно базы достигнет величины (0,3...4) В, диод Шоттки «откроется», т.е. перейдет, как отмечалось ранее, в состояние «стабилизации» напряжения: на диоде напряжение не будет превышать (0,3 – 4) В, а на коллекторном переходе будет практически отсутствовать инжекция.

Режим насыщения БТ, по свойствам близкий к активному характеризуется тем, что в базе транзистора практически не накапливаются неосновные носители заряда, поэтому время перехода (переключения) БТ с режима насыщения в активный будет мало.

Б) Пусть соотношение напряжений будет теперь таким:

$$\left\{ \begin{array}{l} U_{ЭБ} > 0 \\ U_{ЭБ} \sim U_{ЭБ}^{жс} \end{array} \right\}, \left\{ \begin{array}{l} U_{КБ} > 0 \\ U_{КБ} \sim U_{КБ}^{жс} \end{array} \right\}. \quad (2.31)$$

В этом случае напряжение на коллекторном переходе равно пороговому и можно полагать, что на КП имеет место инжекция по интенсивности примерно одного порядка, что и инжекции на ЭП. Распределение носителей в базе для этого случая показано на графике рисунка 2.10.

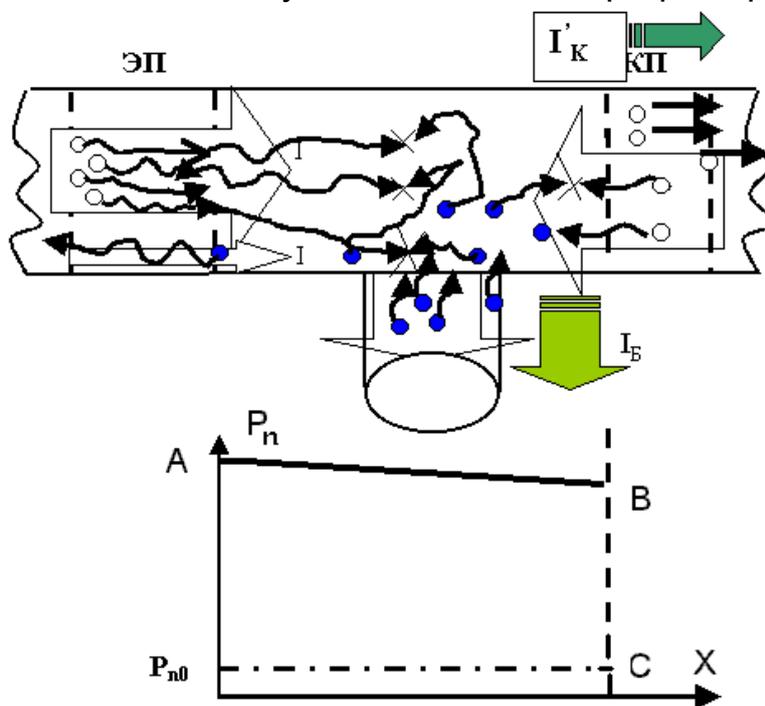


Рис. 2.10. Фрагмент базы и коллекторного перехода структуры $p^{++} - n - p^+$ биполярного транзистора в режиме глубокого насыщения, и протекающие при этом токи

Вспоминая, какую информацию несут наклон линии и площадь под линией графика, можем утверждать следующее.

Наклон линии графика, характеризующей распределение неосновных носителей в базе БТ, по сравнению с активным режимом (рис. 2.4) существенно уменьшился. Следовательно, существенно снизился уровень диффузии неосновных носителей в базе и – уменьшился ток I'_K . Соответственно, ток коллектора станет существенно меньшим и уменьшится коэффициент передачи по току БТ в СОБ. Если линия будет проходить горизонтально, что свидетельствует о равенстве процессов инжекции и экстракции на КП (инъекции переходов уравнивают друг друга), то ток коллектора будет равен нулю.

Площадь под линией (рис. 2.10), характеризующей распределение неосновных носителей в базе, по сравнению с активным режимом (рис.6.4), существенно увеличилась. Это свидетельствует о том, что значительно возрос ток базы. Соответственно, коэффициент передачи тока базы в СОЭ $\beta = I_K/I_B$ уменьшится.

2.2.5. Признаки работы биполярного транзистора в режиме насыщения

Характерными признаками, по которым можно определить, что БТ работает в режиме насыщения, являются следующие.

1. Напряжение на ЭП имеет величину примерно равную пороговому напряжению электронно – дырочной структуры. Для кремниевого БТ $U_{ЭБ} \sim U_{ЭБ}^{жс} \approx (0,6...0,8)V$, для германиевого БТ $U_{ЭБ} \approx (0,3...0,4)V$.

Напряжение на коллекторе относительно базы мало $U_{KB} \leq U_{KB}^{жс}$ или примерно равно пороговому напряжению.

Напряжение на коллекторе относительно эмиттера мало $U_{КЭ} \approx (0,1...1)V$.

2. Ток коллектора меньше тока эмиттера $I_K < I_Э$ в активном режиме. Важно, что ток коллектора в режиме глубокого насыщения существенно меньше и может быть даже равным нулю.

Ток базы I_B , существенно больше, чем он был в активном режиме.

3. Со стороны входа сопротивление БТ определяется *дифференциальным сопротивлением эмиттерного перехода*. Согласно (1.37) можно записать:

$$r_{\text{Э}} = \frac{dU_{\text{ЭБ}}}{dI_{\text{Э}}} = \frac{\varphi_T}{I_{\text{Э}}}. \quad (2.32)$$

Сопротивление эмиттерного перехода, как было определено ранее, относительно мало.

Со стороны выхода сопротивление БТ определяется *дифференциальным сопротивлением коллекторного перехода*

$$r_{\text{К}} = \frac{dU_{\text{КБ}}}{dI_{\text{К}}} = \frac{\varphi_T}{I_{\text{К}}}. \quad (2.33)$$

Сопротивление коллекторного перехода, как было определено ранее, относительно также относительно мало (обратите внимание на то, что БТ перестал трансформировать сопротивления, как это предписывается ему названием).

Понятно, что сопротивление базы также уменьшено, поскольку она «насыщена» подвижными носителями заряда.

Таким образом, нетрудно сообразить, что сопротивление между коллектором и эмиттером, как последовательное соединение трех сопротивлений малой величины, также будет малым. Это позволяет, если вспомнить классификацию компонентов электрической цепи, уподобить БТ, работающий в режиме насыщения, коротко замкнутой перемычке (точнее электронному ключу, у которого накоротко замкнуты выводы коллектора и эмиттера).

Из такого заключения вытекают два важные для практики обстоятельства.

А. В режиме насыщения ток, протекающий по цепи коллектора определяется уже не самим транзистором, а той внешней электрической цепью, в которую он включен.

Б. «Быстрота размыкания ключа» определяется тем обстоятельством, как быстро будут «удалены» неосновные носители из базы в про-

цессе переключения «ключа». Чем больше у БТ «глубина насыщения» (чем больше база насыщена дырками), тем дольше по времени будут происходить процессы удаления дырок из базы и тем медленнее будет «размыкаться» ключ. При использовании фиксирующего диода Шоттки быстродействие ключа намного повышается.

4. В режиме насыщения БТ не способен усиливать электрический сигнал по мощности. С точки зрения электротехники БТ в этом режиме представляет ключ, который потребляет очень малую долю энергии и диссипирует ее (превращает в тепло).

5. В режиме насыщения напряжения на ЭП и КП малы, поэтому на БТ выделяется очень малая мощность, которая преобразовывается в тепло.

2.2.6. Функционирование биполярного транзистора в режиме отсечки

Режимом отсечки $p^{++} - n - p^{+}$ БТ называется такое его состояние, при котором на эмиттер относительно базы и на коллектор относительно базы поданы отрицательные потенциалы. При этом величины напряжений $U_{ЭБ}$ и $U_{КБ}$ должны быть меньше значений напряжений электрического пробоя. Математически это можно записать так:

$$\left\{ \begin{array}{l} U_{ЭБ} < 0 \\ |U_{ЭБ}| < |U_{ЭБПР}| \end{array} \right\}, \left\{ \begin{array}{l} U_{КБ} < 0 \\ |U_{КБ}| < |U_{КБПР}| \end{array} \right\}. \quad (2.33)$$

При этом в БТ будут происходить следующие явления и процессы.

1. Поскольку несимметричные $p^{++} - n$ ЭП и $n - p^{+}$ КП смещены в обратном направлении, то в них, как следует из предыдущего раздела (п.1.3.4), будет иметь место односторонняя экстракция неосновных носителей заряда (вытягивание дырок из базы в эмиттер и в коллектор). Неосновных носителей в базе очень мало, поэтому потоки «вытягиваемых» носителей будут крайне малы (рис. 2.11).

После восстановления электронейтральности областей коллектора и эмиттера, подобно тому, как это было рассмотрено ранее, через ме-

таллические выводы Э и К будут протекать токи очень малой величины (обратные токи электронно – дырочного перехода).

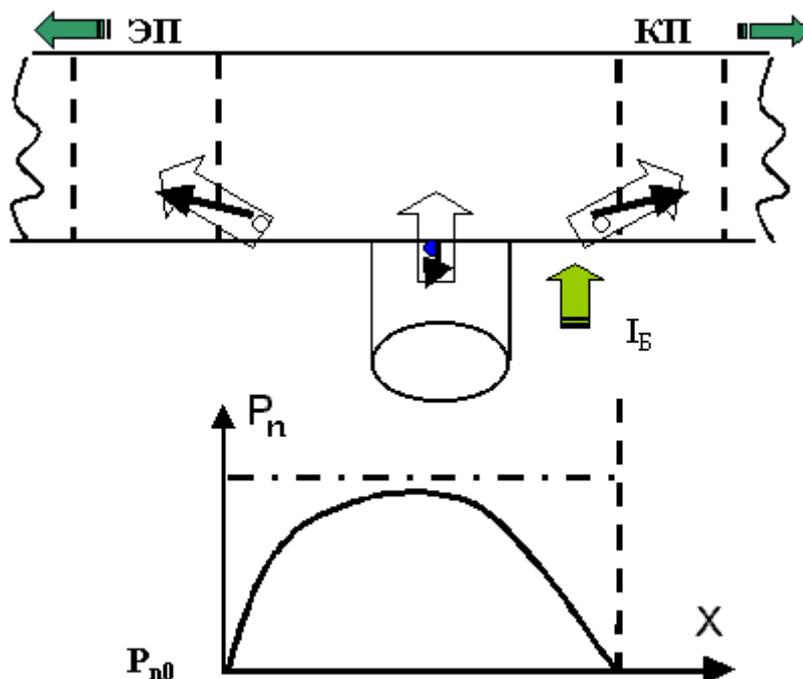


Рис. 2.11. Фрагмент базы и коллекторного перехода структуры $p^{++} - n - p^+$ биполярного транзистора в режиме отсечки, и протекающие при этом токи

Для того, чтобы можно было представить процессы, происходящие в базе БТ режиме отсечки, на рис. 2.11 под фрагментом изображения структуры транзистора показан график зависимости концентрации неосновных носителей (дырок) в базе транзистора от координаты X .

2.2.7. Признаки работы биполярного транзистора в режиме отсечки

Характерными признаками, по которым можно определить, что БТ работает в режиме отсечки, являются следующие.

1. Напряжение на ЭП и КП велики и практически имеют величину примерно равную ЭДС источников питания, подключенным к металлическим выводам Э и К.
2. Токи коллектора, базы и эмиттера, протекающие через металлические выводы крайне малы и практически равны нулю.

Исключение составляет случай «оторванной» базы, когда напряжение приложено непосредственно между Э и К. В этом случае через металлические выводы К и Э протекает, согласно (2.17), «сквозной» ток:

$$I_{КЭ} = (\beta + 1)I_{КБ0}, \quad (2.35)$$

который, при больших «бетта», может достигать относительно большой величины.

3. Дифференциальные сопротивления ЭП и КП велики:

$$r_{Э} = \frac{dU_{ЭБ}}{dI_{Э}} = 1 / dI_{Э} / dU_{ЭБ} \approx \frac{1}{0} \approx \infty, \quad (2.36)$$

$$r_{К} = \frac{dU_{КБ}}{dI_{К}} = 1 / dI_{К} / dU_{КБ} \approx \frac{1}{0} \approx \infty, \quad (2.37)$$

а сопротивление базы увеличено по сравнению с активным режимом.

Таким образом, нетрудно сообразить, что сопротивление между коллектором и эмиттером, как последовательное соединение трех сопротивлений очень большой величины, также будет очень большим. Это позволяет, если вспомнить классификацию компонентов электрической цепи, уподобить БТ, работающий в режиме отсечки, разрыву цепи (точнее, электронному ключу у которого разомкнуты выводы коллектора и эмиттера).

4. В режиме отсечки БТ не способен усиливать электрический сигнал по мощности. С точки зрения электротехники БТ в этом режиме представляет ключ, который в идеале не потребляет энергию.

5. В режиме отсечки токи, протекающие через ЭП и КП, крайне малы, поэтому на БТ выделяется очень малая мощность, которая преобразовывается в тепло.

2.2.8. Статические характеристики биполярного транзистора

Статическими характеристиками транзистора называют математические зависимости между токами и напряжениями на различных выводах транзистора. Они важны для того, чтобы можно было правильно-

но выбирать условия работы транзистора. Обычно их представляют в виде графиков.

Поскольку транзистор прибор с тремя выводами, и если подойти чисто формально, то в принципе можно получить 24 вида статических характеристик. На практике обычно используют три вида: *входные; выходные; проходные статические характеристики*.

Следует также учитывать, что статические характеристики, вообще говоря, это функции двух независимых переменных. Их графики представляют объемное изображение поверхности в трехмерном пространстве. Это не всегда удобно при практической работе, особенно для случаев, когда надо определять какие либо параметры транзистора. По этой причине пользуются обычно не самими характеристиками, а *семейством характеристик*. В этом случае одну из двух независимых переменных фиксируют (считают, что она на характеристиках не меняется; эту переменную называют *параметром семейства*) и анализируют уже функциональные зависимости, получающиеся уже как функции от одной независимой переменной.

Если взять несколько различных фиксированных переменных, то получится набор графиков одной переменной, расположенных в единой системе координат. Фиксация одной переменной с точки зрения математики означает, что трехмерную поверхность мы «сечем» плоскостью, соответствующей зафиксированной переменной. А график в этом случае – кривая, получающаяся, как результат пересечения поверхности и плоскости.

Сделав несколько подобных «сечений», получив на плоскости для них графики функций одной переменной, расположив графики в единой системе координат, мы, в итоге, получим набор «родственных» графиков (все они «дети» поверхности в трехмерном пространстве, рожденные «сечением» поверхности «фиксирующей» плоскостью). Эти графики получили название *семейств характеристик*.

Графики семейств входных и выходных характеристик БТ, включенного по СОЭ, показаны на рис. 2.12.

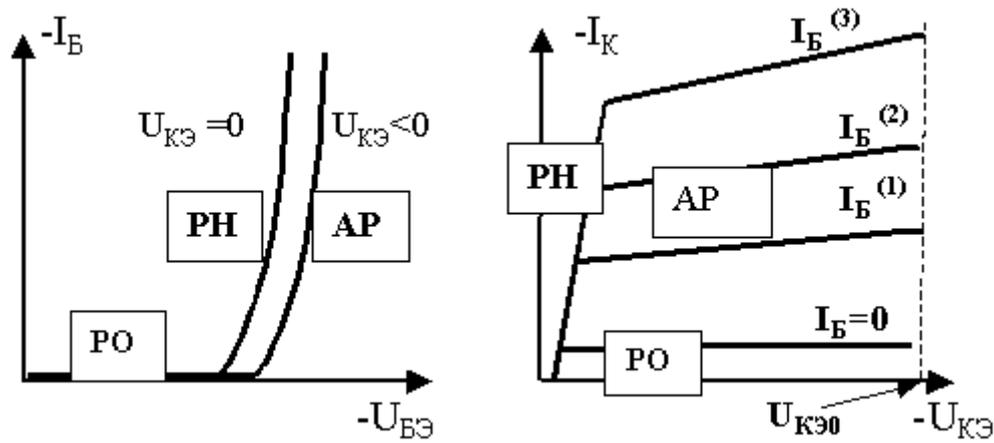


Рис. 2.12. Семейство входных и выходных статических характеристик биполярного транзистора структуры $p^{++} - n - p^{+}$

Семейство входных характеристик БТ, включенного по СОЭ, это система графиков, которая отражает зависимость тока базы от входного напряжения между базой и эмиттером, полученная при условии, что параметр – напряжение между коллектором и эмиттером – постоянно.

$$\{I_B = F(U_{БЭ})\} \text{ при } U_{КЭ}^n = const, n = 1 - 2. \quad (2.38)$$

Было принято соглашение статические характеристики изображать в первом квадранте системы координат, поэтому Вас сейчас, и в дальнейшем, не должно смущать то обстоятельство, что ток базы и напряжение база-эмиттер имеют знак «минус». Аналогичная ситуация и с выходными характеристиками.

Входные характеристики по виду напоминают вольтамперную характеристику электронно-дырочной структуры при прямом включении. Как и ранее, характеристики начинаются при пороговом напряжении и существенно нелинейны. Однако, при этом надо иметь в виду, что ток базы это есть ток рекомбинации, поэтому его величина мала и на несколько порядков меньше прямого тока электронно-дырочной структуры. Расположение характеристик зависит от режима работы: при заданном напряжении $U_{БЭ}$ входной ток в активном режиме несколько меньше.

Графики характеристик идут, после порогового напряжения, вертикально вверх. Это свидетельствует о том, что со стороны «входа» БТ обладает относительно малым входным дифференциальным сопротив-

лением и «стабилизирует» напряжение. С этой точки зрения, вид характеристик показывает, что со стороны входа БТ *должен управляться источником тока*.

Чтобы не было искажений сигнала необходимо, как известно, иметь линейную зависимость между током базы и напряжением база-эмиттер. Предположение о линейности участка входных характеристик можно допустить, если диапазон изменения входного напряжения ($U_{БЭ}$) не превышает величины $2\varphi_T$. По этой причине БТ, как свидетельствует семейство входных характеристик, можно использовать для усиления электрического сигнала без искажений в активном режиме только в том случае, *если амплитуда сигнала на входе не превышает несколько десятков милливольт*.

Для оценки *коэффициента нелинейных искажений*, возникающих при усилении сигнала, можно, в первом приближении, пользоваться формулой:

$$K_{\Gamma} = \frac{U_{mBX}^2}{\varphi_T^2}. \quad (2.39)$$

Семейство выходных характеристик БТ, включенного по СОЭ, это система графиков, которая отражает зависимость тока коллектора от выходного напряжения между коллектором и эмиттером, полученная при условии, что параметр – ток базы – постоянен

$$\{I_K = F(U_{КЭ})\} \quad \text{при} \quad I_B^n = const, n = 1 - 4. \quad (2.40)$$

Как видно из рис. 2.12 линии графика, в большей своей части идут почти горизонтально. Это соответствует активному режиму. Пологий участок графиков, в идеале, объясняется выражением (2.17). Однако на практике коэффициент усиления тока β зависит как от напряжения $U_{КЭ}$, так и от тока базы. Поэтому графики имеют наклон относительно оси абсцисс и расположены неэквидистантно (при одинаковом изменении тока базы характеристики располагаются на разном расстоянии друг от друга).

Почти «горизонтальный» вид характеристик свидетельствует о том, что со стороны «выхода» БТ можно уподобить источнику тока с большим внутренним дифференциальным сопротивлением (если быть точным, то

зависимому источнику тока, управляемому током). Наклон реальных характеристик отражает тот факт, что величина выходного дифференциального сопротивления равна не бесконечности, а конечной, но большой величине.

«Почти вертикальные, крутые» участки графиков семейства выходных характеристик (рис. 2.12), при малых напряжениях $U_{кэ}$, соответствуют *режиму насыщения*. Причем, чем больше ток базы, тем больше «глубина» насыщения неосновными носителями базы. Напряжение, соответствующее точке на линии режима насыщения при заданных токах базы и коллектора, получило название *напряжение насыщения*. Этот параметр приводят в справочных данных транзистора. Он важен для транзисторов, работающих в цифровых устройствах, и его стараются сделать минимальным.

При токе базы равном нулю (как отмечалось ранее, это соответствует случаю оторванной базы) протекает довольно большой величины ток коллектора, который мы именовали *сквозным током* (2.18).

При напряжениях, больших $U_{кэ0}$, наблюдается резкое увеличение тока, обусловленное *электрическим пробоем коллекторного перехода* (на графиках не показано). Следует иметь в виду, что $U_{кэ0}$ связано с напряжением пробоя электронно-дырочной структуры при обратном включении (напряжением пробоя КП в БТ, включенном по СОБ) следующим соотношением:

$$U_{кэ0} = \frac{U_{КБПР}}{\sqrt[n]{\beta + 1}} \approx \frac{U_{КБПР}}{(2...3)}. \quad (2.40)$$

Понижение напряжения пробоя в СОЭ объясняют проявлением внутренней положительной обратной связи. Пары носителей заряда, образующиеся в коллекторном переходе при возникновении лавинного пробоя, разделяются электрическим полем КП: электроны переносятся в базу, а дырки в – коллектор. Так как ток базы поддерживается постоянным, то электроны накапливаются в базе и понижают ее потенциал, т.е. увеличивают прямое напряжение на ЭП. Вследствие этого возрастает инжекция неосновных носителей в базу из эмиттера. Большая часть этих дополнительных носителей «проходит» к КП. Там они в свою очередь вызывают лавинный пробой, в результате чего опять увеличивает-

ся число электронов, попавших в базу. Так возникает положительная обратная связь, приводящая к резкому увеличению коллекторного тока.

Немаловажное значение для понимания работы транзистора представляет умение находить распределение напряжений в выходной цепи БТ для случая, если между выводом коллектора и источником постоянной ЭДС E включен резистор R (рис. 2.13).

Сложность этой цепи состоит в том, что в ней ток коллектора и напряжение между коллектором и эмиттером на транзисторе связаны между собой нелинейной функциональной зависимостью $i_K = F(u_{KЭ})$. Поэтому напряжения и токи можно определить только из решения нелинейного уравнения. Наша задача – научиться в цепи с нелинейным элементом определять напряжения и токи.

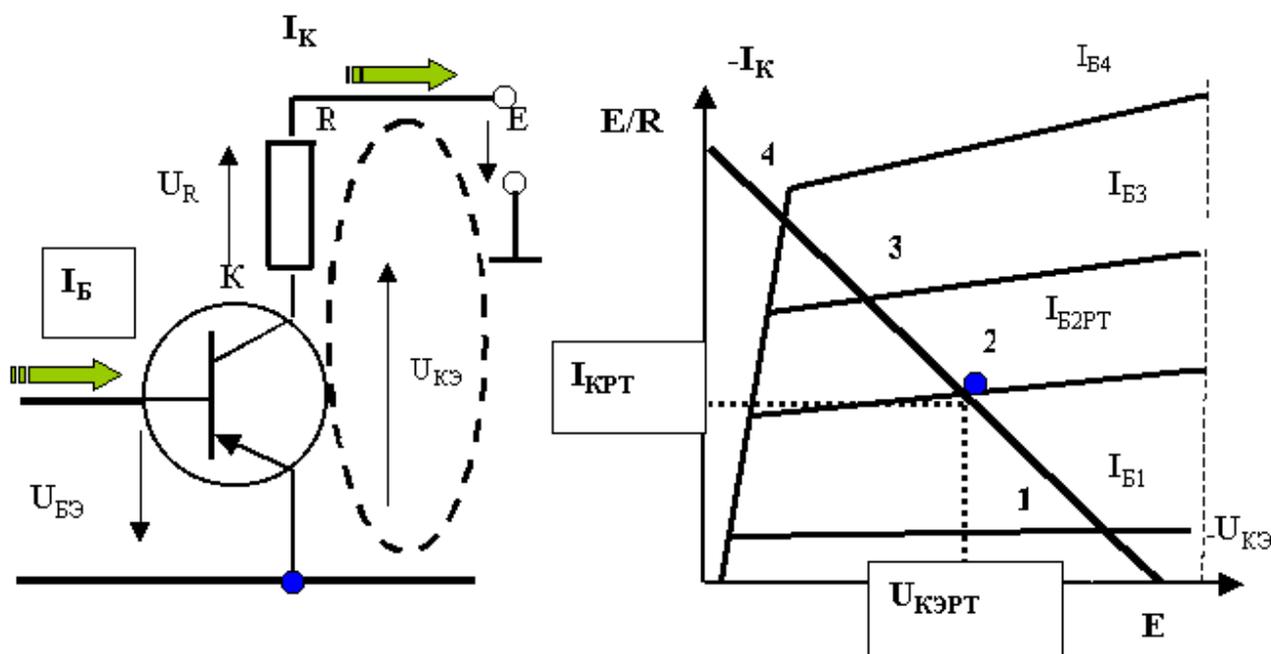


Рис. 2.13. Распределение напряжений в выходной цепи биполярного транзистора структуры $p^{++} - n - p^+$, если между выводом коллектора и источником постоянной ЭДС E включен резистор R

Согласно второму закону Кирхгофа для схемы можем записать (если обходить контур по часовой стрелке):

$$U_R + U_{KЭ} = E. \quad (2.42)$$

Используя закон Ома для резистора, и, решая уравнение относительно тока, получаем систему уравнений:

$$\begin{cases} I_K = \frac{1}{R}(E - U_{КЭ}), \\ I_K = F(U_{КЭ}). \end{cases} \quad (2.43)$$

Следовательно, ток коллектора, который будет протекать через транзистор, и напряжение $U_{КЭ}$ находятся из решения системы уравнений. Первое уравнение системы представляет уравнение прямой линии:

$$I_K = \frac{1}{R}(E - U_{КЭ}) \Rightarrow y = -k(x + b). \quad (2.44)$$

Эту линию называют *нагрузочной прямой* (линией нагрузки) и ее можно построить по двум точкам (рис. 2.13).

а) При $U_{КЭ} = 0$ получаем $I_K = E/R$. В этой точке линия пересекает ось ординат.

б) При $I_K = 0$ получаем $U_{КЭ} = E$. В этой точке линия пересекает ось абсцисс.

Второе уравнение системы это нелинейная зависимость, которую обычно представляют в виде некой кривой.

Понятно, что в изучаемой цепи могут быть только такие напряжения и токи, которые бы одновременно удовлетворяли двум уравнениям: и линии и одной из нелинейных характеристик семейства. То есть, если исходить из графиков, решение это есть точка, которая получается в результате пересечения нагрузочной прямой и одной из выходных характеристик БТ (назовем ее условно ζ характеристикой) (рис. 2.13). В свою очередь, ζ характеристика задается током базы БТ. Поэтому приходим к тому, что электрический режим БТ будет определяться системой трех значений: тока базы, тока коллектора, напряжения $U_{КЭ}$.

Точка с координатами $\{I_{КРТ}, U_{КЭРТ}, I_{БРТ}\}$ на семействе выходных характеристик БТ, одновременно располагающаяся и на нагрузочной прямой, называется *рабочей точкой*.

На рис. 2.13 рабочая точка это точка 2 с координатами $\{I_{KPT}, U_{KЭPT}, I_{Б2PT}\}$. Задавая другие величины токов базы можно перейти в новую рабочую точку, например, 1,3,4. Несложно догадаться о том, что если менять величину ЭДС E и сопротивления резистора R , тем самым меняя наклон и положение нагрузочной прямой, то можно получить новые рабочие точки в выходной цепи транзистора.

Часто на транзистор подается сигнал, представляющий сумму постоянного напряжения и косинусоидального сигнала малой амплитуды (сигнал малой амплитуды необходим для того, чтобы характеристики можно было считать линейными и чтобы, при этом, не было искажений). В этом случае постоянные напряжения, которые задают рабочую точку, называют «*напряжениями смещения*». С этой точки зрения рабочая точка определяет на выходных характеристиках точку, в окрестности которой происходят изменения косинусоидальных токов и напряжений малой амплитуды. Соотношения, характеризующие гармонические сигналы малой амплитуды в рабочей точке, называют «*дифференциальными (малосигнальными) параметрами*». Понятно, что в каждой новой рабочей точке будут свои дифференциальные параметры.

Понятия «рабочая точка», «напряжение смещения», «дифференциальный параметр» являются важнейшими при описании режимов работы всех разновидностей транзисторов. По аналогии, их используют при описании работы входной цепи транзистора. Эти понятия используются при объяснении работы практически всех полупроводниковых приборов, поэтому их следует запомнить и применять в своей работе.

Посмотрим, как будет меняться величина сопротивления БТ между выводами коллектора и эмиттера и режим работы БТ в зависимости от положения рабочей точки (РТ). С этой целью проследим, как будет меняться работа транзистора при изменении тока базы.

А) Если ток базы равен нулю $I_B = 0$, то нагрузочная прямая и выходная характеристика семейства, соответствующая $I_{Б1} = 0$, пересекаются в точке 1. Напряжение $U_{кэ} \sim E$. Ток коллектора равен сквозному току и относительно мал. Нетрудно подсчитать, что сопротивление БТ между выводами К и Э очень велико. Из рис. 2.12 следует, что рабочая точка 1 располагается в области режима отсечки, и мы еще раз убеждаемся в том, что в режиме отсечки БТ между выводами К и Э можно упо-

добить разомкнутому ключу (разрыву цепи, участку цепи, имеющему очень большое сопротивление).

Б) Если последовательно увеличивать ток базы, то линия нагрузки и соответствующая выходная характеристика будут пересекаться в точках 2, 3. Это область активного режима. Меняя ток базы и перемещаясь по нагрузочной прямой, как нетрудно подсчитать, мы меняем сопротивление между выводами К и Э БТ. Если исходить из рассмотренной классификации компонентов электрической цепи, то можно говорить о том, что БТ в активном режиме подобен переменному резистору, у которого, строго говоря, величина сопротивления меняется не за счет перемещения ползунка или поворота ручки, а за счет изменения тока базы. Теперь БТ это некий переменный «электронный» резистор, у которого отсутствуют всякие механические перемещения, вращения, контакты.

В) Если еще увеличивать ток базы, то точка пересечения линии нагрузки и характеристика семейства попадут в точку 4 и, соответственно, в режим насыщения. Для этой области характерны относительно высокие токи базы, большие токи коллектора и очень малые значения напряжения $U_{КЭ} \sim 0$. Соответственно, можно полагать $R_{КЭ} \sim 0$. Мы еще раз убеждаемся в том, что в режиме насыщения БТ между выводами К и Э можно уподобить замкнутому ключу (КЗ перемычке, участку цепи, имеющему нулевое сопротивление).

Интересен и тот факт, что в режиме насыщения, как бы мы не увеличивали ток базы, ток коллектора, достигнув своего максимального значения $I_{Кmax} = E/R$ не будет больше зависеть от I_B .

Г) Наконец, если обобщить сказанное, то можно отметить, что если, меняя ток базы (например, подавая короткие по длительности прямоугольные импульсы тока базы) переводить БТ из режима насыщения в режим отсечки, то можно получить *быстродействующий «электронный» ключ*.

Такое функционирование транзистора используется в цифровых электронных схемах для создания логического «нуля» и логической «единицы». Для таких схем очень важно уменьшать время переключения из «нуля» в «единицу», поскольку это время определяет быстродействие схем и, соответственно, скорость проведения вычислений.

2.2.9. Температурные и частотные свойства биполярного

транзистора

В процессе эксплуатации полупроводниковые изделия могут подвергаться воздействию перепада температур от -60°C до 100°C . Варьирование температуры окружающей среды обуславливает изменения внутренней энергии материалов, из которых изготовлены полупроводниковые изделия, и вызывают изменения их характеристик и параметров. Поэтому важно знать, как температура влияет на БТ.

Не приводя математических выкладок, отметим, что при увеличении температуры, в первую очередь за счет изменения концентрации неосновных носителей в базе БТ, статические характеристики будут изменяться примерно таким образом, как это показано на рис. 2.14.

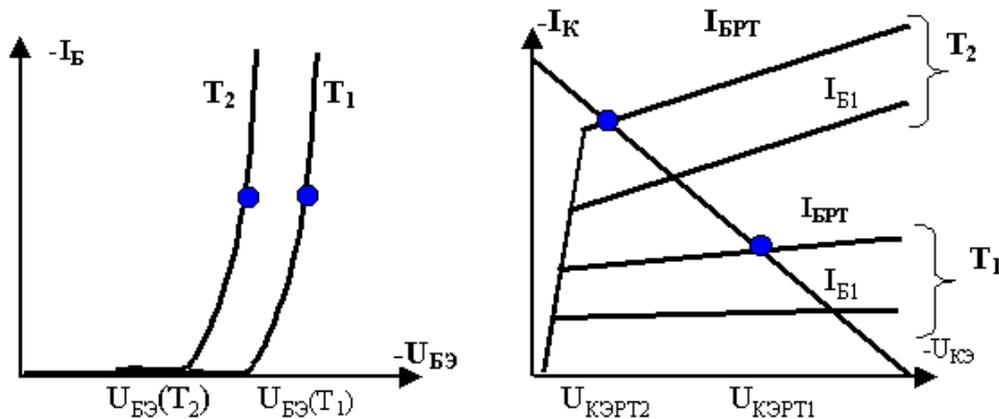


Рис. 2.14. Изменение семейства входных и выходных статических характеристик биполярного транзистора структуры $p^{++} - n - p^{+}$ при варьировании температуры

Напряжение $U_{BЭ}$ с увеличением температуры уменьшается (по модулю). Это изменение можно считать линейным:

$$U_{BЭРТ}(T_2) = U_{BЭРТ}(T_1) + TKU(T_2 - T_1). \quad (2.45)$$

Для инженерных расчетов можно принять $TKU = -2,1\text{мВ/град}$.

Известно (2.17), что ток коллектора в СОЭ определяется формулой

$$I_K = \beta I_B + (\beta + 1)I_{KB0} \quad (2.46)$$

Поэтому для оценки влияния температуры на выходные характеристики надо знать как зависит от температуры коэффициент усиления по току β и обратный ток коллекторного перехода I_{KB0} .

В первом приближении зависимость β от температуры можно оценивать следующим образом:

$$\beta(T_2) = \beta(T_1) + (1,2...1,5)(T_2 - T_1). \quad (2.47)$$

Для определения того, как зависит от температуры I_{KB0} , пользуются понятием *температуры удвоения*:

$$I_{KB0}(T_2) = I_{KB0}(T_1) \cdot 2^{\frac{T_2 - T_1}{10}}. \quad (2.49)$$

Нетрудно рассчитать, что при изменении температуры на 100 градусов, обратный ток коллекторного перехода I_{KB0} меняется примерно в $2^{10} \sim 1000$ раз.

Очень важно понимать следующее. Если не предпринимать никаких мер, то транзистор, работающий в активном режиме и способный усиливать электрический сигнал по мощности, при изменении температуры окружающей среды будет переходить либо в режим отсечки, либо в режим насыщения, в которых он не усиливает сигнал, а играет роль ключа. Происходящее можно, опираясь на рис. 2.14, пояснить следующими цепочками рассуждений:

$$T \uparrow \Rightarrow \left\{ \begin{array}{l} \beta \uparrow \\ I_{KB0} \uparrow \end{array} \right\} \Rightarrow I_{KPT} \uparrow \Rightarrow p.m. \rightarrow \text{вверх} \rightarrow PH, \quad (2.49)$$

$$T \downarrow \Rightarrow \left\{ \begin{array}{l} \beta \downarrow \\ I_{KB0} \downarrow \end{array} \right\} \Rightarrow I_{KPT} \downarrow \Rightarrow p.m. \rightarrow \text{вниз} \rightarrow PO. \quad (2.50)$$

Надо иметь в виду, что БТ, работающий в СОБ, обладает несколько лучшей температурной устойчивостью.

При работе БТ с быстроизменяющимися высокочастотными гармоническими или импульсными сигналами его усилительные свойства ухудшаются. В этом отношении БТ подобен фильтру низких частот: с увеличением частоты его коэффициенты усиления по току α, β уменьшаются. Графики зависимости α, β от частоты показаны на рис. 2.15.

Как следует из рис. 2.15, БТ, с точки зрения его частотных свойств, характеризуется тремя параметрами:

- а) предельной частотой передачи базового тока для СОЭ f_β ;
- б) предельной частотой передачи эмиттерного тока для СОБ f_α ;
- в) граничной частотой $f_{ГР}$.

Рисунок 2.15 также свидетельствует, что в активном режиме СОБ имеет значительно лучшие частотные свойства по сравнению с СОЭ. Так как СОЭ лучше усиливает по мощности, то целесообразно в диапазоне частот $(0 - f_{ГР})$ использовать включение БТ по СОЭ, а в диапазоне $(f_{ГР} - f_\alpha)$ — по СОБ.

Ухудшение усилительных свойств БТ с ростом частоты связано с двумя факторами.

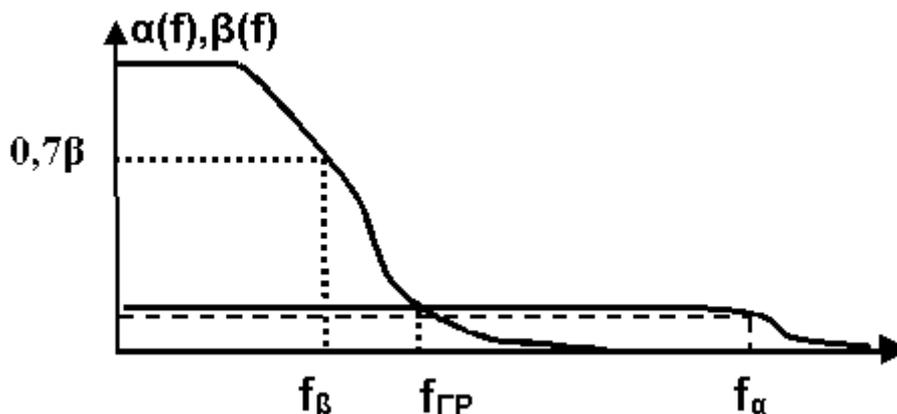


Рис. 2.15. Графики зависимости коэффициентов усиления по току α, β биполярного транзистора, включенного по СОБ и СОЭ, от частоты

1. На высоких частотах барьерная емкость коллекторного перехода «шунтирует» КП и тем самым ухудшает усилительные свойства.

2. При больших частотах изменения сигнала за счет инерционности, носители не успевают перемещаться по базе, что приводит к ухудшению усилительных свойств.

Для расширения частотного диапазона работы необходимо увеличивать скорость перемещения неосновных носителей по базе (для этого используют так называемые дрейфовые транзисторы), уменьшать толщину базы и емкость коллекторного перехода (другими словами, уменьшать размеры областей транзистора).

2.3. Полевые транзисторы

2.3.1. Общие сведения о полевых транзисторах

Биполярным транзисторам присущи недостатки, основными из которых являются.

А. У БТ работа определяется, главным образом, неосновными носителями заряда. Отсюда сильная зависимость их работы от температуры окружающей среды.

Б. В базе БТ носители перемещаются за счет диффузии. Скорость такого перемещения относительно медленная (значительно меньше, чем скорость дрейфа), поэтому быстродействие БТ не всегда отвечает требованиям практики.

В. Малое входное сопротивление БТ не позволяет их согласовывать с источниками сигналов с большим внутренним сопротивлением (вспомните условия передачи мощности от источника в нагрузку).

Чтобы свести к минимуму эти недостатки был предложен новый тип транзистора, который получил название полевого. От БТ полевой транзистор принципиально отличается следующим.

1. Используемым типом носителей. У полевого транзистора осуществляется перенос основных носителей и за счет электрического поля (дрейфа). В связи с тем, что главную роль играют основные носители, полевые транзисторы часто называют также *униполярными*.

2. Способом управления. БТ является прибором, управляемым током. У полевого транзистора управление сигналом осуществляется входным напряжением (электрическим полем).

Полевой транзистор обладает высокой устойчивостью к температурным изменениям, что, как увидим в дальнейшем, очень важно для интегральных схем, в которых транзисторы расположены очень близко друг относительно друга, и мощных приборов энергетической (силовой) электроники. Кроме того, у полевых транзисторов удастся достигнуть очень высокого входного сопротивления (сотен мегом). Это дает возможность подключать их к «высокоомным» источникам.

Полевым (униполярным) транзистором (ПТ) называют полупроводниковый прибор с тремя (иногда с четырьмя) выводами, у которого входная цепь изолирована от выходной диэлектриком или обедненным слоем при обратном включении, в основу принципа действия которого положено управляющее действие электрического поля на основные носители, и предназначенный для осуществления следующих функций:

1. Усиления электрических сигналов по мощности.
2. Коммутации (переключения, подключения) участков электрической цепи.
3. Регулирования величины протекающего через два его вывода тока.

Структуры ПТ весьма разнообразны, но обязательно в их составе имеются:

А) Слой (брусочек) полупроводника p или n типа по которому движутся основные носители заряда, называемый каналом.

Канал p или n типа (рис.2.16) может существовать изначально с момента изготовления транзистора и именоваться встроенным. Канал, отсутствующий в равновесном состоянии, но образующийся за счет внешнего напряжения, называется *индуцированным каналом*. Сечение (толщина) канала в процессе работы ПТ может изменяться, за счет чего будет меняться величина тока, протекающего по каналу.

Б) Металлические выводы (электроды, омические контакты) с помощью которых обеспечивается включение канала в электрическую цепь.

В) Исток – это электрод и область канала, из которой исходят (истекают) основные носители заряда.

Г) Сток – это электрод и область канала, в которую основные носители входят (стекают).

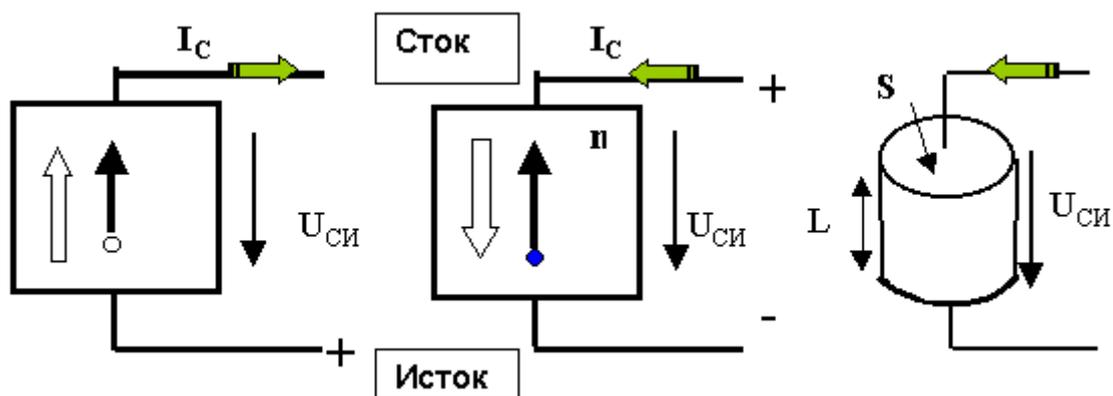


Рис. 2.16. Разновидности каналов полевого транзистора и демонстрация того, как движутся носители в канале и протекает ток стока

Основные носители движутся по каналу вследствие воздействия на них продольного электрического поля E , создаваемого в канале при приложении напряжения между стоком и истоком.

Следует запомнить, что, как бы по определению, у полевого транзистора основные носители всегда должны двигаться от истока к стоку (рис. 2.16). Этот факт однозначно определяет полярность напряжения между стоком и истоком.

Д) Затвор – это электрод, применяемый для управления величиной тока в канале.

Е) Монокристаллическая подложка – полупроводниковая или диэлектрическая область, на которой размещаются элементы полевого транзистора.

Идея управления током канала реализуется на основе следующих простых соображений.

Ток стока, протекающий через брусок полупроводника равен (рис. 2.16):

$$I_C = \frac{U_{СИ}}{R_{СИ}} = \frac{U_{СИ}}{\rho L/S} = \frac{U_{СИ} S}{\rho L} = KS, \quad (2.51)$$

где $K = \frac{U_{СИ}}{\rho L}$ - постоянная.

Тогда, как следует из (2.51), при постоянном напряжении между стоком и истоком, варьируя сечение канала S , можно менять протекающий по каналу ток стока.

На практике сечение канала S можно менять, например, следующим образом. Известно, что при изменении обратного напряжения на электронно-дырочной структуре меняется толщина обедненного слоя:

$$L_{об} = L_{обp} + L_{обn} = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_K + U)}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}. \quad (2.52)$$

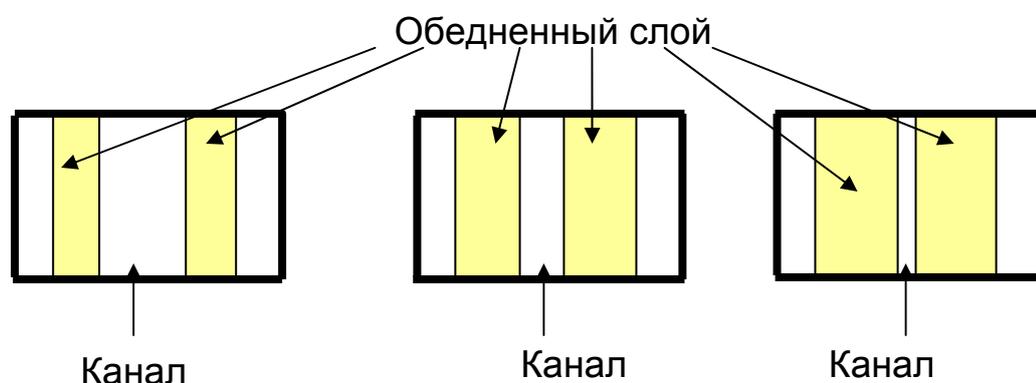


Рис. 2.17. Возможный способ регулирования сечения канала

Поэтому, если с помощью напряжения на затворе менять толщину обедненного слоя, то тем самым будет изменяться толщина проводящей части канала (рис. 2.17) и, соответственно, ток канала (ток стока).

2.3.2. Устройство и принцип действия полевых транзисторов с управляющим переходом

Идею изменения сечения канала ПТ за счет изменения обедненной области на практике в настоящее время реализуют двумя способами: путем использования обедненного слоя электронно-дырочной структуры; путем использования обедненной области структуры металл – полупроводник (затвора Шоттки). Более широкое распространение получили ПТ с затвором, на основе электронно – дырочной структуры. Они проще и сточки зрения понимания работы. На их изучении мы и сосредоточим свое внимание. ПТ с затвором Шоттки сейчас являются основным функ-

циональным элементом при создании сверхскоростных цифровых компонентов, а также схем, работающих на сверхвысоких частотах. Эти ПТ изготавливаются на основе арсенид галлия (у которого очень высокая подвижность носителей заряда и, соответственно, большая скорость дрейфа носителей в канале) и принцип их действия, в идейном смысле, не отличается от ПТ с затвором на основе $p-n$ -перехода. Он легко может быть понят при усвоении ПТ с затвором на основе электронно дырочной структуры.

Представим себе, что в нашем распоряжении имеется предмет, напоминающий по форме карандаш, однако, в качестве грифеля у него используется полупроводник n типа, в качестве деревянного цилиндра – полупроводник p типа. Вместо краски на его внешнюю поверхность напылен слой металла. С двух сторон к « n – грифелю», а также к напыленному металлу подсоединены металлические выводы, к которым приложены электрические потенциалы. Если мысленно рассечь такой «карандаш» и, как принято в инженерной графике, изобразить чертеж такой структуры на плоскости (проекцию), то получим упрощенное представление ПТ с затвором на основе электронно - дырочной структуры (часто говорят с затвором на основе $p-n$ -перехода), показанное на рис. 2.18.

К затвору (З) относительно истока (И) приложено напряжение $U_{ЗИ}$ таким образом, чтобы $p-n$ переход оказался при обратном включении. На сток (С) относительно И подано напряжение $U_{СИ}$ такой полярности, чтобы основные носители под действием силы электрического поля ($\vec{F} = q\vec{E}$) двигались от истока к стоку.

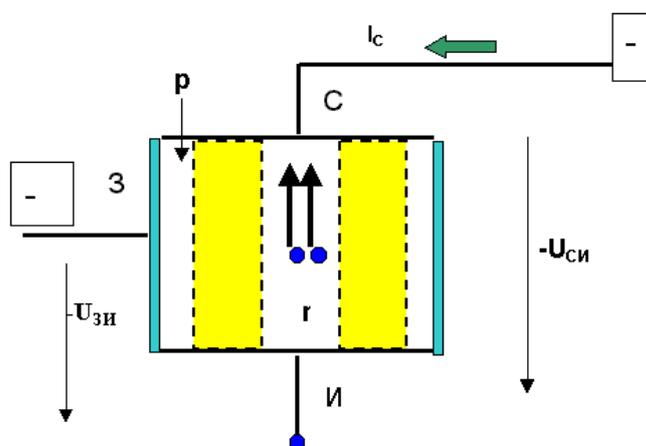


Рис. 2.18. Упрощенная структура полевого транзистора с управляющим затвором в виде $p-n$ перехода

Следует учитывать, что ток затвора представляет обратный ток, протекающий через закрытый $p-n$ -переход. Этот ток пренебрежимо мал и на практике считается, что он равен нулю. Соответственно, нетрудно догадаться, что входное сопротивление ПТ очень велико и стремится к бесконечности.

Принцип действия ПТ с затвором в виде $p-n$ -перехода заключается в том, что при изменении напряжения на затворе изменяется толщина обедненного слоя, а значит, и толщина проводящей части канала. Меняя напряжение $U_{ЗИ}$ можно изменять толщину канала и, соответственно ток стока.

Эта модель ПТ позволяет понять и достаточно наглядно представить вид статической передаточной характеристики ПТ, как зависимости тока стока от напряжения $U_{ЗИ}$ при неизменном напряжении $U_{СИ}$ (рис. 2.19).

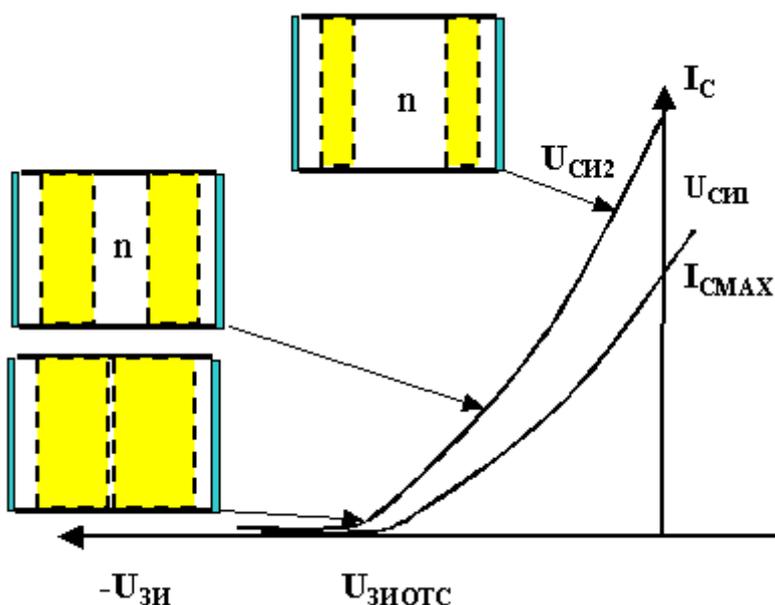


Рис. 2.19. График семейства стоко-затворных характеристик полевого транзистора с управляющим затвором в виде $p-n$ перехода

Шокли показал, что:

$$I_C = I_{СМАХ} \left(1 - \frac{U_{ЗИ}}{U_{ЗИ0ТС}} \right)^2. \quad (2.53)$$

$I_{C_{MAX}}$ - максимальный ток стока. $U_{ЗИОТС}$ - напряжение отсечки, т. е. напряжение между затвором и истоком при котором считается, что ток стока равен нулю (обычно считают при $I_C = 10$ мкА).

Управляющее действие затвора характеризуют *крутизной стоко-затворной характеристики*:

$$S = \frac{dI_C}{dU_{ЗИ}} \text{ при } U_{СИ} = const. \quad (2.54)$$

Дифференцируя (2.53) находим:

$$S = \frac{dI_C}{dU_{ЗИ}} = \frac{2I_{C_{MAX}}}{U_{ЗИОТС}} \left(1 - \frac{U_{ЗИ}}{U_{ЗИОТС}} \right). \quad (2.55)$$

Нетрудно догадаться, что максимальной крутизной ПТ с затвором в виде $p-n$ перехода $S_{MAX} = 2I_{C_{MAX}}/U_{ЗИОТС}$ обладает при напряжениях $U_{ЗИ} \sim 0$. То есть для получения хороших усилительных свойств рабочую точку полевого транзистора следует выбирать в области, где $U_{ЗИ} \sim 0$.

Степень «нелинейности» характеристики, по сравнению с БТ меньше (у БТ экспоненциальная зависимость, а у ПТ квадратичная), поэтому, при том же заданном коэффициенте нелинейных искажений, на вход полевого транзистора можно подавать большую амплитуду входного сигнала.

Прежде чем анализировать вид выходной, или как говорят, стоковой характеристики, сделаем несколько предварительных замечаний.

На первый взгляд, кажется, что сечение канала при перемещении вдоль канала от истока к стоку (как говорят, конфигурация канала) при изменении напряжения $U_{СИ}$ должно оставаться неизменным. На самом деле это не так.

Распределение потенциала внутри канала зависит от двух координат X и Y , от «поперечной» и «продольной». Если полагать потенциал истока равным нулю, то можно заметить, что значение «продольного» потенциала канала $\varphi_K(y)$ линейно меняется в пространстве, вдоль координаты Y от точки к точке канала.

Ширина обедненного слоя определяется «поперечной» разностью потенциалов, то есть, при конкретном Y , разностью потенциалов между точкой канала и точкой затвора: $U_{зк}(Y) = \varphi_K(Y) - \varphi_3 = \varphi_K(Y)$. Вдоль пространственной координаты Y будет меняться разность потенциалов затвор – точка канала $U_{зк}(Y)$, напряжение, прикладываемое к обедненному слою, внешнее поле обедненного слоя, ширина обедненного слоя. Следовательно, *первый очевидный факт – толщина канала должна быть переменной: канал должен быть самым узким в районе стока и самым широким в районе истока*. Второй факт состоит в том, что, меняя величину напряжения $U_{си}$, мы обязательно будем менять степень перекрытия канала: чем больше будет напряжение $U_{си}$ тем сильнее будет перекрытие канала в области стока (рис. 2.21).

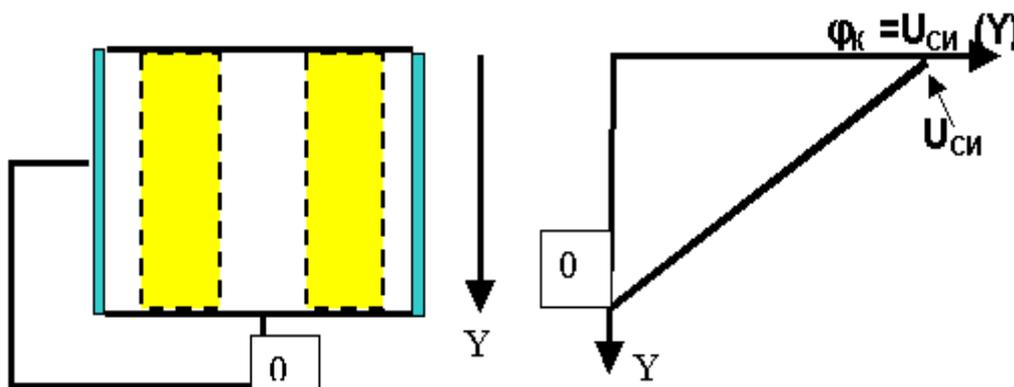


Рис. 2.20. Распределение потенциала φ_K вдоль координаты Y канала ПТ

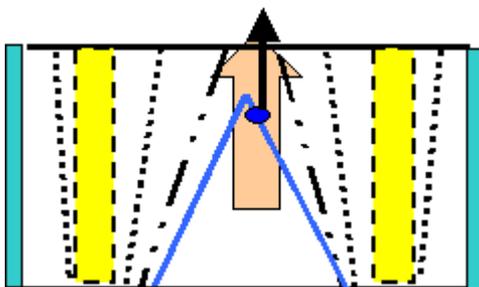


Рис. 2.21. Изменение ширины обедненной области и конфигурации канала при изменении напряжения $U_{си}$

Семейство стоковых характеристик ПТ с затвором в виде $p-n$ перехода, как система зависимостей $I_C = F(U_{СИ})$ при постоянных напряжениях $U_{ЗИ}$, показано на рис. 2.22.

Внешне эти зависимости напоминают выходные характеристики БТ. На характеристиках имеется участок резкого изменения тока, и участок, на котором изменение тока мало. По аналогии с БТ на графике семейства $I_C = F(U_{СИ})$ можно выделить три основные рабочие области.

1. Область отсечки выходного тока (ООВТ). На ПТ подано большое отрицательное напряжение $U_{ЗИ}$, толщина канала очень мала и в цепи стока протекает очень малый ток.
2. Область открытого состояния (ООС).
3. Активная область (АО), в которой наблюдается «пологая» часть характеристик.

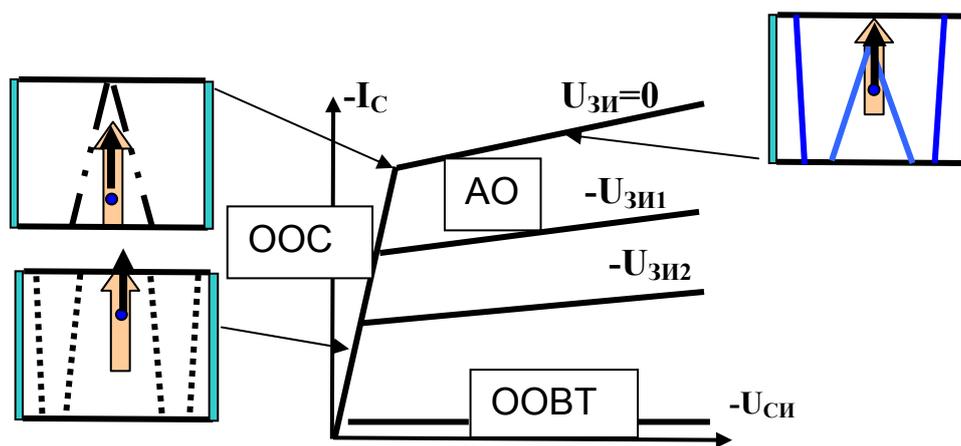


Рис. 2.22. Графики семейства стоковых характеристик полевого транзистора с управляющим затвором в виде $p-n$ перехода

Рассмотрим одну из характеристик семейства при $U_{ЗИ} = 0$.

Пока канал ПТ не перекрыт, вид характеристики определяется, в первом приближении, законом Ома: ток стока пропорционален напряжению $U_{СИ}$. На характеристиках это «вертикальный» участок (крутая часть характеристики).

В точке излома характеристик наступает перекрытие канала. Перекрытие канала называют «условным», так как при этом не происходит

«отсечки» тока (ток стока не становится равным нулю, а, по-прежнему, протекает через канал).

Далее, при увеличении напряжения $U_{СИ}$. включается в работу «механизм самоограничения тока стока», результатом работы которого является неизменность тока стока. Другими словами, в активной области у ПТ имеет место «отсечка приращений тока стока» при увеличении напряжения $U_{СИ}$.

Существуют несколько объяснений механизмов протекания тока стока при условном перекрытии канала и самоограничения тока стока.

а) Ток протекает и ограничивается за счет эффекта экстракции.

При работе в активной области появление некоторого «проводящего канала» в перекрытой стоковой области, по которому протекает ток, обусловлено экстракцией (вытягиванием) носителей из канала в область стока за счет электрического поля. При наличии напряжения $U_{СИ}$. сомкнутая часть канала работает подобно коллекторному переходу БТ. При увеличении $U_{СИ}$. растет и длина перекрытой части канала, потому продольная напряженность поля E остается неизменной, а соответственно, не меняется ток стока.

б) Ток протекает за счет эффекта экстракции, а ограничивается за счет «эффекта горячих электронов».

Поскольку условно перекрытая часть канала имеет значительно большее сопротивление по сравнению с остальной частью канала, то основная часть напряжения $U_{СИ}$. (согласно свойству простейшего делителя напряжения) падает на перекрытой части канала. В области перекрытия создается очень высокая напряженность электрического поля. Доказано, что подвижность носителей μ в очень сильных электрических полях (так называемых «горячих» носителей) не зависит от величины напряженности электрического поля. Вследствие этого дрейфовая скорость носителей заряда $V_{др} \sim \mu E$ становится постоянной и, естественно, перестает увеличиваться ток стока: $j_C = qnV_{др}$.

В реальных ПТ в активной области ток стока с увеличением напряжения сток – исток несколько увеличивается, что свидетельствует о том, что со стороны выхода ПТ имеет большое, но все же не бесконечное, выходное дифференциальное сопротивление.

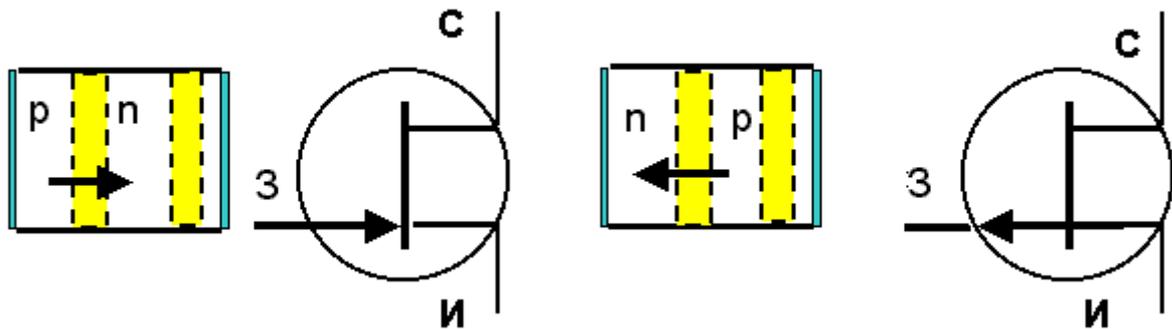


Рис. 2.23. Условное графическое изображение полевых транзисторов с управляющим затвором в виде $p-n$ перехода на схемах: с n каналом (слева) и p каналом

На рис. 2.23 показано условное графическое изображение на схемах полевых транзисторов с управляющим затвором в виде $p-n$ перехода. Линия посередине обозначает канал. Стрелка у затвора всегда направлена от p области ПТ к n области. Это позволяет по условному графическому изображению определять тип канала, структуру ПТ и, соответственно, какой полярности напряжения надо прикладывать между выводами транзистора. Напряжение $U_{ЗИ}$ должно быть таким, чтобы $p-n$ переход оказывался при обратном включении. Полярность $U_{СИ}$ должна обеспечивать движение основных носителей в канале от истока к стоку.

2.3.3. Устройство и принцип действия полевых транзисторов с изолированным затвором

Идею изменения сечения канала ПТ с изолированным затвором рассмотрим сначала на основе транзистора с индуцированным каналом. Упрощенная структура полевого транзистора с изолированным затвором и индуцированным каналом показана на рис. 2.24.

Такой ПТ создается на слаболегированной кремниевой подложке (П) p – типа с концентрацией акцепторов порядка 10^{15} см^{-3} . У поверхности подложки сформированы сильнолегированные истоковая и стоковая n^{++} области с концентрацией доноров порядка 10^{19} см^{-3} . Расстояние между сильнолегированными областями стока и истока, называемое *длиной канала* может составлять от десятых долей до нескольких микрометров. На поверхности полупроводника создан тонкий слой диэлектри-

ка толщиной 0,05 – 0,1 мкм, в качестве которого обычно используют диоксид кремния SiO_2 . На слой диэлектрика нанесен металлический слой (обычно алюминий), к которому подсоединен вывод затвора (З). Аналогично, имеются металлические слои и выводы от стока (С), истока (И), подложки (П). Если двигаться по структуре слева направо, то получим Металл – Диэлектрик – Полупроводник. Поэтому часто ПТ такого типа называют МДП-транзисторами. Если именуют диоксид кремния SiO_2 оксидом, то ПТ называют *МОП транзистором*.

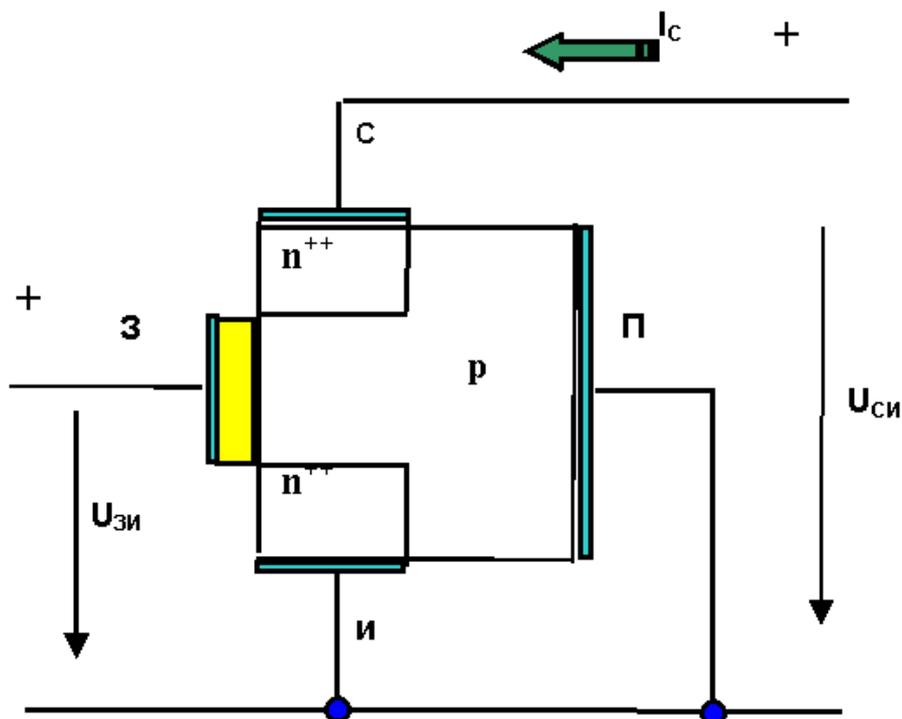


Рис. 2.24. Упрощенная структура полевого транзистора с изолированным затвором и индуцированным n каналом

Следует учитывать, что ток затвора представляет ток, протекающий через качественный диэлектрик. Этот ток пренебрежимо мал и на практике считается, что он равен нулю. Соответственно, нетрудно догадаться, что входное сопротивление ПТ с изолированным затвором очень велико и стремится к бесконечности. Полевой транзистор не потребляет мощность от источника, подключенного к затвору.

Предположим, что ПТ мы включили по схеме с общим истоком (СОИ) и его подложку соединили с И. Пусть потенциалы И и П равны нулю, а на сток, относительно истока, подано напряжение полярности, показанной на рис. 2.24. При $U_{зи} = 0$ ток стока будет равен нулю, так как p

– n^{++} переход у стока оказывается включенным в обратном направлении. Инжекция на истоковом $p - n^{++}$ переходе происходить не может, так как все напряжение $U_{СИ}$ приложено к обедненному слою стокового перехода. Напряженность продольного электрического поля под затвором всюду равна нулю. При $U_{ЗИ} < 0$ ситуация будет аналогичной.

Так как подложка соединена с истоком и потенциал подложки равен потенциалу истока, то следует иметь в виду, что напряжение $U_{ЗИ} = U_{ЗП}$ распределяется между диэлектриком и слаболегированным, а, следовательно, обладающим большим сопротивлением, полупроводником. Это напряжение создает в диэлектрике и в подложке поперечное электрическое поле. Легко усмотреть при этом аналогию с конденсатором, у которого одна обкладка – металл затвора, другая – металл подложки.

При $U_{ЗИ} > 0$ возникнет электрическое поле, вектор напряженности которого ориентирован так, как показано на рис. 2.25. Электрическое поле сосредоточено главным образом в диэлектрике, но часть его, проникая на небольшую глубину, имеется в подзатворной области p полупроводника. Под действием этого поперечного электрического поля ($\vec{F} = q\vec{E}$) основные носители подложки (дырки) будут «выталкиваться» вглубь подложки, а неосновные (электроны) – «подтягиваться» к затвору и накапливаться там. Концентрация «подтянувшихся» электронов в подзатворной области будет определяться напряженностью поперечного электрического поля, или точнее, напряжением $U_{ЗИ} = U_{ЗП}$. При малых напряжениях $U_{ЗИ}$ концентрация электронов мала и в подзатворной области преобладают дырки, как основные носители p полупроводника. По-прежнему отсутствует канал в ПТ и ток стока пренебрежимо мал.

С увеличением $U_{ЗИ}$ в подзатворной области ПТ будет расти концентрация электронов и уменьшаться концентрация дырок. При некотором напряжении $U_{ЗИ}$, называемом *пороговым напряжением*, концентрация электронов в подзатворной области превысит концентрацию дырок, количество которых все время уменьшалось. Произойдет инверсия типа проводимости p полупроводника: в подзатворной области он станет полупроводником n типа. Слово «инверсия» – изменение обычного порядка. Следовательно, в подзатворной области образуется канал n – типа (условно показан на рис. 2.25 штриховой линией). По мере дальнейшего

увеличения напряжения $U_{ЗИ}$ толщина канала будет увеличиваться. А раз есть канал, то возникнет ток стока.

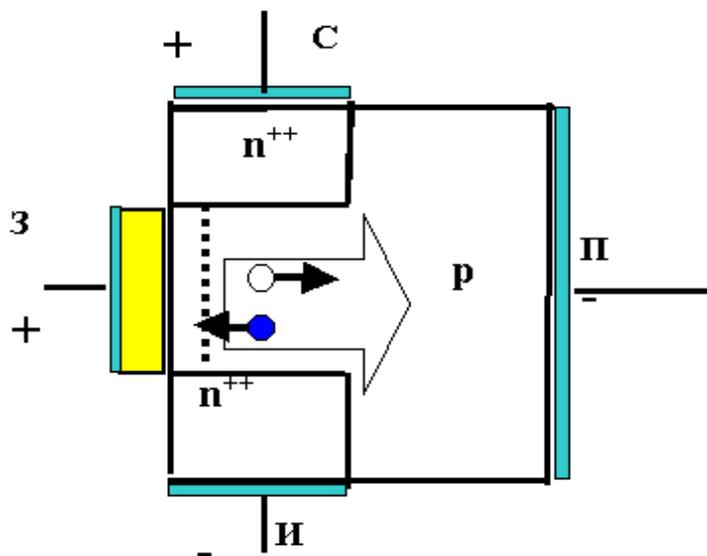


Рис. 2.25. Поперечное электрическое поле и «наведение» канала в полевом транзисторе с изолированным затвором

Итак, при подаче на затвор относительно истока напряжения, превышающего пороговое значение, в ПТ с изолированным затвором возникает (индуцируется, наводится) канал. Такие транзисторы называют транзисторами с индуцированным каналом.

Из рассуждений несложно представить вид передаточной стокозатворной характеристики ПТ с изолированным затвором и индуцированным n каналом (рис. 2.26).

В большинстве случаев подложка объединяется внутри корпуса ПТ с истоком. Однако в ряде случаев такое соединение не производят, поэтому разность напряжений между истоком и подложкой может воздействовать на канал и изменять пороговое напряжение (рис. 2.26). Вывод от подложки может использоваться в качестве второго управляющего электрода. Но при этом, следует иметь в виду, что напряжение подложка-исток влияет гораздо слабее, чем затвор исток.

В ряде случаев подзатворную область заранее легируют донорами, которые создают в ней проводящий канал. Следовательно, у таких с изолированным затвором ПТ, проводящий канал существует (является как бы «встроенным») при нулевом напряжении $U_{ЗИ}$ ПТ подобного типа называют *транзисторами со встроенным каналом*.

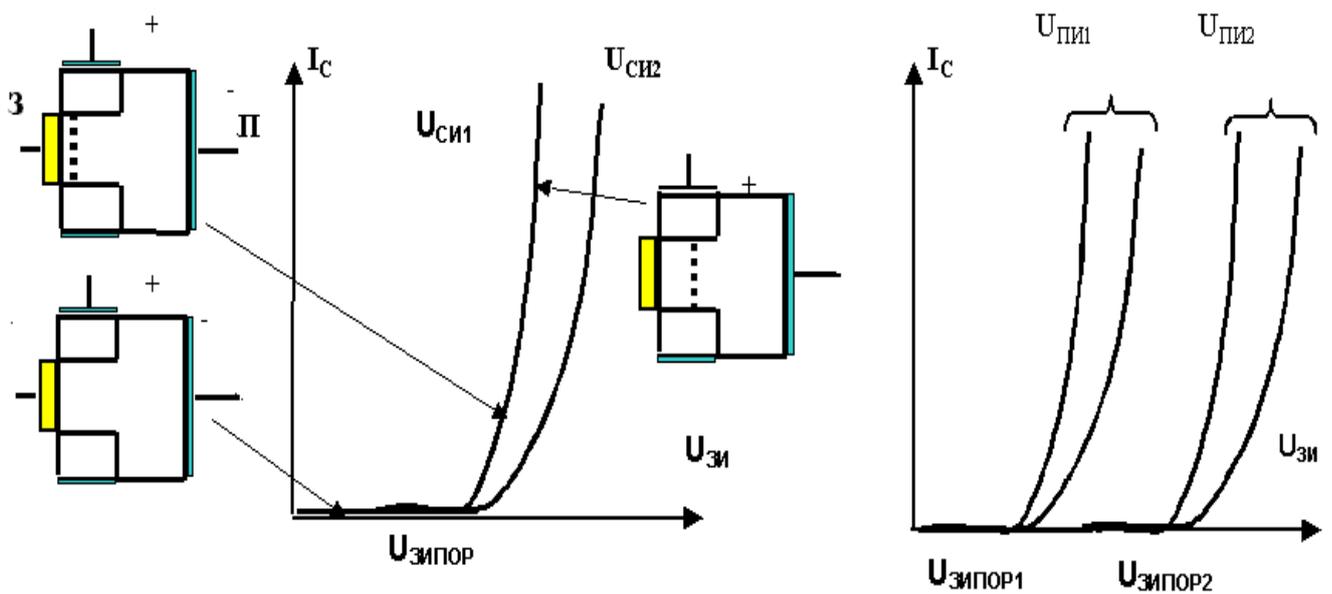


Рис. 2.26. График семейства стоко-затворных характеристик полевого транзистора с изолированным затвором

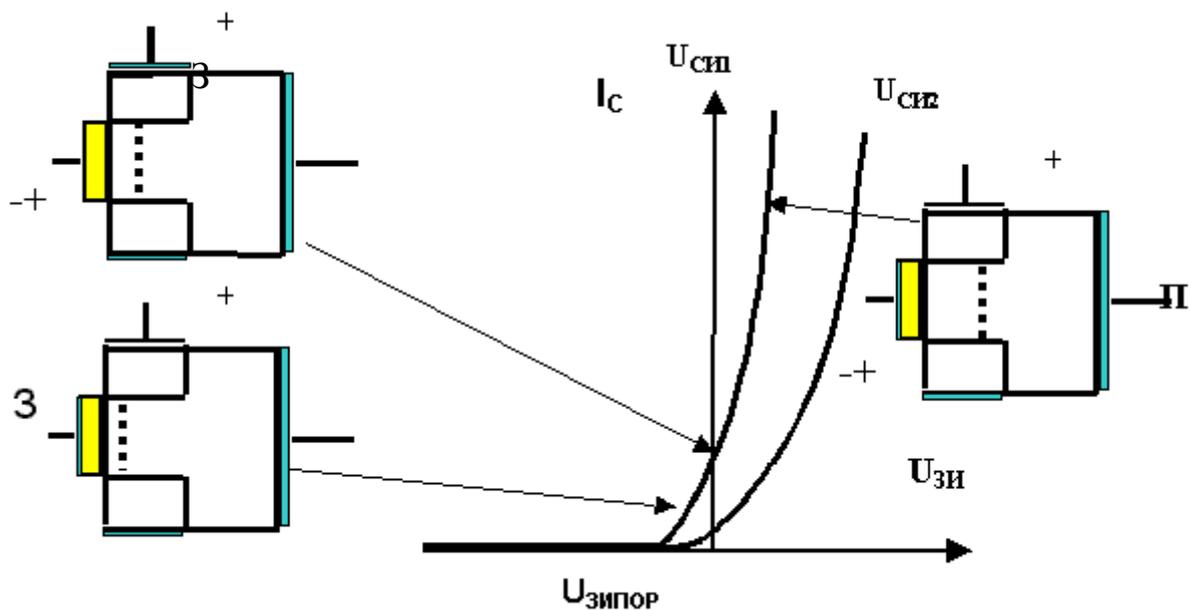


Рис. 2.27. График семейства стоко-затворных характеристик полевого транзистора с изолированным затвором и встроенным каналом

В транзисторах со встроенным каналом при подаче $U_{зи} >$ поперечное электрическое поле будет увеличивать толщину канала за счет

«подтягивания» электронов из подложки. Говорят, что происходит «обогащение» канала носителями. При приложении отрицательного напряжения $U_{ЗИ}$ поперечное электрическое поле «выталкивает» в подложку электроны и тем самым уменьшает толщину канала. Часто говорят, что происходит «обеднение» канала.

Стокзатворные характеристики ПТ с изолированным затвором можно представлять, в первом приближении, следующей зависимостью:

$$I_C = b(U_{ЗИ} - U_{ЗИПОР})^2. \quad (2.56)$$

Стокзатворные характеристики квадратичные, поэтому степень их нелинейности меньше, чем у БТ.

Кривизна стокзатворной характеристики определяется следующей формулой:

$$S = \frac{dI_C}{dU_{ЗИ}} = 2b(U_{ЗИ} - U_{ЗИПОР}). \quad (2.57)$$

На рис. 2.28 показаны условное графическое изображение на схемах полевых транзисторов с изолированным затвором с встроенным и индуцированным каналами.

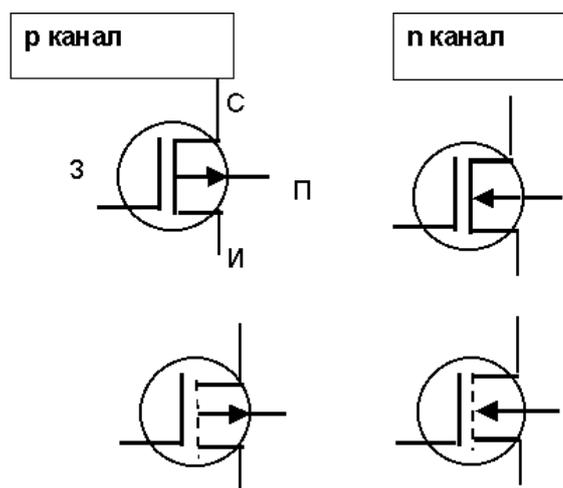


Рис. 2.28. Условное графическое изображение на схемах полевых транзисторов с изолированным затвором: с встроенным (вверху) и индуцированным (внизу) каналом

Линия посередине между стоком и истоком по -прежнему обозначает канал. Если канал изображен сплошной линией, то это встроенный канал. Штриховая линия свидетельствует о том, что это ПТ с индуцированным каналом. Затвор отделен от канала промежутком, подчеркивающим, что затвор изолирован от канала диэлектриком. Стрелка, идущая от подложки к каналу или наоборот от канала к подложке, всегда направлена от p области к области n . Это позволяет по условному графическому изображению определять тип канала и, соответственно, какой полярности напряжения надо прикладывать между выводами транзистора. Поскольку основные носители в канале всегда должны двигаться от истока к стоку, ставим знаки потенциалов так, чтобы обеспечить такое перемещение (рис. 2.16). Для ПТ с n каналом $U_{СИ}$ должно быть положительным, а для ПТ с p каналом – отрицательным. Напряжение $U_{ЗИ}$ должно быть таким, чтобы в ПТ с индуцированным каналом обеспечивалось наведение канала. Обычно знак потенциала стока и затвора относительно подложки совпадают. Полярность $U_{ЗИ}$ у ПТ со встроенным каналом, в зависимости от того что требуется увеличивать или уменьшать сечение канала, может быть и положительной и отрицательной и равной нулю.

Семейство стоковых характеристик ПТ с изолированным затвором, как система зависимостей $I_C = F(U_{СИ})$ при постоянных напряжениях $U_{ЗИ}$, подобно графикам ПТ с затвором в виде $p - n$ перехода. Различаются напряжения $U_{ЗИ}$, как показано на рис. 2.29.

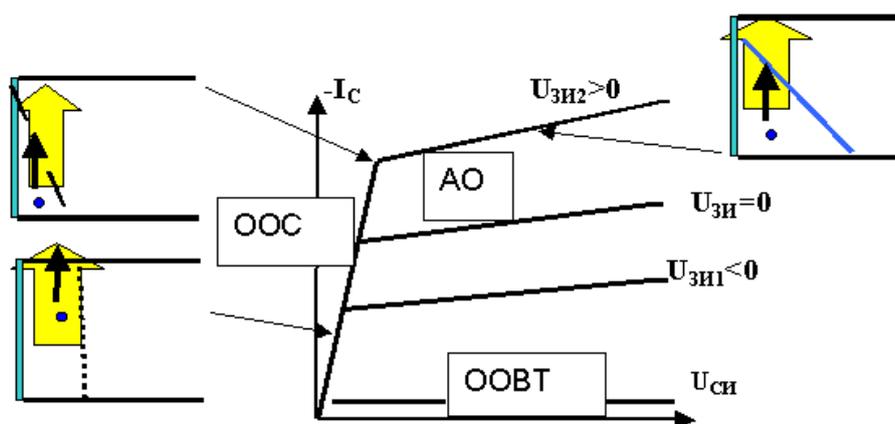


Рис. 2.29. График семейства стоковых характеристик полевого транзистора с изолированным затвором

Как отражено на фрагментах, представляющих подзатворную область, в ПТ с изолированным затвором при наличии напряжения $U_{си}$ толщина вдоль канала также является переменной: канал самый узкий в районе стока и самый широкий в районе истока. Наступает и перекрытие канала, которое мы называли «условным», так как при этом не происходит «отсечки» тока (ток стока не становится равным нулю, а протекает через канал). В ПТ работает и «механизм самоограничения тока стока», результатом работы которого является неизменность тока стока.

В заключение отметим, что полевые транзисторы могут работать в усилительных схемах, где они будут усиливать сигнал по мощности, и в цифровых схемах - в качестве ключа. Следует обратить внимание на то, что ПТ с индуцированным каналом в исходном состоянии представляет собой разомкнутый ключ. Это позволяет создавать схемы, не потребляющие энергии от источника питания, пока они не работают.

Контрольные вопросы

1. Какое устройство называют биполярным транзистором? Для осуществления каких функций он предназначен?
2. Охарактеризуйте структуры $p^{++} - n - p^+$ и $n^{++} - p - n^+$ биполярных транзисторов и приведите их условные графические изображения на схемах.
3. Приведите и охарактеризуйте схемы включения $p^{++} - n - p^+$ биполярного транзистора.
4. Что называют электрическим режимом работы БТ? Как он определяется?
5. Назовите и поясните электрические режимы $p^{++} - n - p^+$ БТ.
6. Назовите и поясните электрические режимы $n^{++} - p - n^+$ БП.
7. Поясните понятия: нормальное и инверсное включение транзистора.
8. Какое состояние называют активным режимом работы $p^{++} - n - p^+$ БТ?
9. Поясните понятия: коэффициент инжекции и «дефект инжекции эмиттера».
10. Охарактеризуйте понятие: статический (интегральный) коэффициент передачи эмиттерного тока к коллектору.

11. Запишите и поясните основное уравнение биполярного транзистора, работающего в активном режиме, в схеме с общей базой.
12. Запишите и поясните основное уравнение биполярного транзистора, работающего в активном режиме, в схеме с общим эмиттером.
13. Определите понятие сквозного тока БТ.
14. Определите понятие: «обрыв» базы БТ.
15. Перечислите основные признаки работы биполярного транзистора в активном режиме.
16. Как определены дифференциальные сопротивления эмиттерного и коллекторного переходов?
17. Как определены входная и выходная мощности, коэффициент усиления по мощности?
18. Какой режим работы называют режимом насыщения $p^{++} - n - p^{+}$ БТ?
19. С какой целью в схеме с БТ параллельно коллекторному переходу подключают фиксирующий диод Шоттки?
20. Охарактеризуйте режим насыщения БТ по свойствам близкий к активному.
21. Перечислите характерные признаки работы биполярного транзистора в режиме насыщения.
22. Какой режим работы называют режимом отсечки $p^{++} - n - p^{+}$ БТ?
23. Перечислите характерные признаки работы биполярного транзистора в режиме отсечки.
24. Какие зависимости называют статическими характеристиками транзистора? Охарактеризуйте 3 основных вида семейств характеристик.
25. Нарисуйте и поясните семейство входных и выходных статических характеристик биполярного транзистора структуры $p^{++} - n - p^{+}$.
26. Охарактеризуйте семейство входных характеристик БТ, включенного по СОЭ.
27. Охарактеризуйте семейство выходных характеристик БТ, включенного по СОЭ.
28. Охарактеризуйте понятие нагрузочной прямой.
29. Как определена на семействе выходных характеристик БТ рабочая точка?
30. Поясните понятие: «напряжение смещения»?
31. Какие соотношения называют «дифференциальными (малосигнальными) параметрами»?

32. Как на основе БТ можно получить быстродействующий «электронный» ключ?
33. Нарисуйте и поясните изменение семейства входных и выходных статических характеристик биполярного транзистора структуры $p^{++} - n - p^+$ при варьировании температуры.
34. Охарактеризуйте понятие температуры удвоения.
35. Какое устройство называют полевым (униполярным) транзистором (ПТ)? Назовите основные отличия полевого транзистора от БТ.
36. Для осуществления каких функций предназначены ПТ?
37. Назовите и охарактеризуйте разновидности каналов полевого транзистора.
38. Кратко охарактеризуйте устройство и принцип действия полевых транзисторов с управляющим переходом.
39. Нарисуйте и поясните упрощенную структуру полевого транзистора с управляющим затвором в виде $p - n$ перехода.
40. Что характеризует крутизна стоко-затворной характеристики ПТ?
41. Нарисуйте и поясните графики семейства стоковых характеристик полевого транзистора с управляющим затвором в виде $p - n$ перехода.
42. Кратко охарактеризуйте устройство и принцип действия полевых транзисторов с изолированным затвором.
43. Нарисуйте и поясните упрощенную структуру полевого транзистора с изолированным затвором и индуцированным n каналом.
44. ПТ какого типа называют МДП (МОП)-транзисторами?
45. Нарисуйте и поясните график семейства стоко-затворных характеристик полевого транзистора с изолированным затвором.
46. Нарисуйте и поясните график семейства стоко-затворных характеристик полевого транзистора с изолированным затвором и встроенным каналом.
47. Нарисуйте условное графическое изображение на схемах полевых транзисторов с изолированным затвором: с встроенным и индуцированным каналом.
48. Нарисуйте и поясните график семейства стоковых характеристик полевого транзистора с изолированным затвором.

3. Основы микроэлектроники

Мы уже изучили широко распространенные классы полупроводниковых изделий – биполярные и полевые транзисторы. При этом транзисторы мы рассматривали как отдельные компоненты электрической цепи. Это позволило нам объяснить их механизмы функционирования, изучить основные параметры и характеристики, описать работу транзисторов для усиления электрических сигналов по мощности и особенности их работы в электронных цифровых схемах

Теперь мы приступаем к изучению полупроводниковых изделий, у которых в едином, небольшом по размерам, кристалле создана целая (единая) сложная электрическая схема, состоящая из большого количества самых различных элементов: транзисторов, резисторов, конденсаторов. То есть, мы переходим к микроэлектронике – области науки и техники, связанной с изготовлением и использованием электронных цепей весьма малых размеров. Важнейшим компонентом электроники в этом случае становится интегральная схема.

Интегральной схемой называют самостоятельное изделие в виде некоторой полупроводниковой микросистемы, образованной неразъемными элементами с высокой плотностью их размещения. Все эти элементы, а также межэлементные соединения, расположены на общей полупроводниковой пластине, что с точки зрения технолога означает, что в различные точки пространства кристалла внедрены, по определенной логике, примеси как различной концентрации, так и различного типа проводимости.

Мы рассмотрим отдельные аспекты производства интегральных схем. Термин «производство», применительно к интегральным схемам, означает совокупность всех операций, начиная с замысла и кончая отправкой готовых изделий к заказчику. Здесь очень много вопросов, на которые требуется дать ответ. К примеру, надо рассчитать электрофизические параметры отдельных элементов цепи (геометрические размеры, дозировку примеси в каждой точке кристалла, тепловое и электрическое взаимодействие между отдельными элементами и т. д.), расположение однородных областей в объеме полупроводника, чтобы, на их основе, можно было создать систему межсоединений, при которой реализуется нужное законченное изделие. При этом электрофизические па-

раметры, оптимальные для одних элементов, могут быть малоприспособлены для других. Выполнить все это «вручную» не представляется возможным, поэтому все в интегральных схемах, начиная от процессов проектирования и заканчивая проверкой правильности функционирования этой сложной электрической цепи, установленной в корпус, выполняются с использованием ЭВМ. Производственные процессы изготовления автоматизированы и, практически, выполняются без участия человека.

Микроминиатюрное исполнение элементов в интегральной схеме меняет как характеристики и параметры самих элементов, так и отдельные принципы построения, в целом, самих схем. Поэтому в дальнейшем мы рассмотрим особенности элементов интегральных схем и некоторые схемотехнические приемы, используемые при реализации функциональных блоков (узлов) изделия (усилителей, формирователей и преобразователей электрических сигналов, элементов цифровых электронных схем и т. д.). Будет изучен важнейший элемент энергонезависимой, быстро перепрограммируемой памяти.

В конце раздела рассматриваются приборы преобразования неэлектрических величин в электрический сигнал. При этом упор делается на полупроводниковые датчики, в частности датчики магнитного потока.

3.1. Общие сведения о микроэлектронных изделиях

3.1.1. Тенденции развития элементной базы электроники и их воплощение в поколениях электронной аппаратуры

Люди, с момента появления первых электронных изделий, всегда стремились сделать их так, чтобы они мало потребляли энергии от источников питания и, соответственно, характеризовались высоким коэффициентом полезного действия (в мире все время ужесточаются требования к экономии электроэнергии), имели небольшие массогабаритные показатели, высокую надежность, большой срок службы, низкую стоимость и большое число одновременно выполняемых функций.

При создании реальной аппаратуры удовлетворить всем этим желаниям было весьма непросто. История воплощения в жизнь этих требований знает ряд характерных этапов, на которых доминировало отвечающее своему времени «поколение» изделий электроники.

В электронных изделиях так называемого «*первого поколения*» использовались электровакуумные лампы, как правило, мощные, крупногабаритные компоненты элементной базы, объемный метод их размещения. Соответственно, аппаратура была тяжелой, малонадежной и дорогой, потребляла много электрической энергии. Автоматизировать производство и конструирование такой аппаратуры было невозможно.

В изделиях «*второго поколения*» применялись отдельные (как говорят, дискретные) транзисторы и миниатюрные навесные компоненты (резисторы, конденсаторы, индуктивности и прочее). Для монтажа их в изделии использовался модульный метод и размещение их на односторонних «печатных» платах. За счет этого был, конечно, достигнут определенный прогресс, в первую очередь, в массогабаритных показателях, надежности и КПД. Однако при этом форма компонентов была самой разнообразной и плохо согласовывалась с поверхностью монтажной плоскости. Автоматизации такое производство поддавалось плохо.

В изделиях *третьего поколения* использовалась широкая гамма микроминиатюрных ЧИП (CHIP) компонент в SMD – исполнении (см. раздел 2), интегральные схемы (группа «планарных» полупроводниковых приборов, расположенных на поверхности единого кристалла, исполненных в одном корпусе), многослойные печатные платы для электрического соединения компонентов. Для монтажа изделий использовался принцип планарной компоновки (при нем практически отказались от цилиндрической формы), который позволил автоматизировать производство. Использование такой элементной базы позволило существенно уменьшить массогабаритные характеристики, повысить надежность, улучшить частотные характеристики, увеличить число функций, выполняемых изделием. Производство было в сильной степени автоматизировано.

Современные электронные изделия «*четвертого поколения*» являются результатом конструктивных и технологических успехов в изготовлении универсальных биполярно–полевых монокристаллических полупроводниковых структур. В таких изделиях применяют микросхемы очень высокой степени интеграции (размеры структурных элементов порядка долей нанометра). Элементы (транзисторы) внутри интегральной схемы, как правило, имеют «вертикальную» структуру и размещаются в объеме кристалла. Для получения хо-

роших тепловых характеристик применяются вертикальные многоячеечные структуры приборов с множеством параллельных каналов. Везде, где требуется высокая плотность монтажа компонент, применяют безкорпусные микросхемы, упакованные в общий корпус, внутренний объем которого, для защиты от воздействия внешней среды, заполнен инертным газом.

Использование «сверхбольших» интегральных схем, безкорпусных микросборок позволило при одинаковой функциональной сложности увеличить плотность упаковки в 5 – 6 раз, повысить надежность, увеличить число функций, выполняемых изделием, снизить стоимость аппаратуры. Все производство является автоматизированным.

Современное исполнение электронных изделий «породило» следующие *основные проблемы*.

А) Повышение плотности упаковки структурных элементов, как несложно догадаться, приводит к увеличению удельной мощности, приходящейся на один элемент, к ухудшению их тепловых режимов.

Б) Близкое расположение элементов друг к другу порождает «резистивные» и «емкостные», как говорят «паразитные», связи между ними.

В) Требуется, чтобы электронное изделие было универсальным, а не узкоспециализированным, поскольку только в таком случае изделие будет иметь приемлемую стоимость.

3.1.2. Понятие об интегральной микросхеме

Интегральной микросхемой (ИМС) называют конструктивно законченное изделие, выполняющее определенную полезную функцию, имеющее высокую плотность упаковки электрически соединенных элементов (в ряде случаев и кристаллов), которое с точки зрения требований к испытаниям, приемке и эксплуатации рассматривается как единое целое.

Термин «интегральная» отражает:

А) Конструктивную интеграцию – объединение значительного числа транзисторов, диодов, конденсаторов и соединительных проводников в единую конструкцию.

В одной ИМС размером 2мм x 2 мм сейчас размещают примерно 1 миллион транзисторов.

Б) Технологическую интеграцию – выполнение в едином технологическом цикле, как всех элементов кристалла (чипа), так и формирование на пластине слитка большого числа чипов.

За один цикл изготовления, после «разламывания» пластины слитка на «осколки» - чипы, получают примерно 20 000 микросхем. Понятно, что стоимость даже очень сложной схемы не может превышать 10 –100 долларов.

В) Схемотехническую интеграцию – выполнение микроэлектронным изделием сложной, качественно новой функции, которую раньше выполняли по раздельности отдельные устройства.

Полупроводниковые ИМС по сравнению с обычными устройствами, в которых используются дискретные элементы, имеют ряд существенных преимуществ:

А. Производство ИМС рентабельно (выгодно), так как здесь в едином технологическом цикле изготавливается большое число идентичных изделий.

Б. Все кристаллы – чипы, полученные из единой пластины, имеют одинаковые характеристики, чего невозможно добиться, применяя дискретные элементы.

В. ИМС отличается высокой надежностью, так как отпадает необходимость малонадежной пайки, поскольку все внутренние межсоединения образуются в процессе изготовления ИМС.

Г. Малые габаритные размеры, хорошие эксплуатационные характеристики и низкая стоимость ИМС позволяют создавать сложные в функциональном отношении, универсальные и относительно недорогие изделия.

Потеря функциональной универсальности, как известно, приводит к сокращению объемов производства. Чтобы этого не допустить идут на усложнение ИМС, введение в них функциональной избыточности, что повышает универсальность микросхем. В частно-

сти, именно это обстоятельство привело к идее создания микропроцессора.

3.1.3. Принципы изготовления интегральных микросхем

В настоящее время наибольшее распространение получили кремниевые полупроводниковые ИМС, в которых элементы устройства формируются в приповерхностном объеме полупроводниковой пластины. Процессы их изготовления мы и будем рассматривать.

Процесс изготовления ИМС заключается в физико-химической обработке кремния и состоит в последовательном выполнении ряда технологических операций, при которых формируются требуемые структуры.

Прежде всего, выращивают слиток из монокристаллического кремния цилиндрической формы диаметром 5 – 15 см. Этот слиток разрезают в нужном направлении с помощью алмазной пилы на круглые пластины. Потом следует механическая «доводка», обеспечивающая заданную толщину пластины, шероховатость (высота неровностей 0,03 – 0,05 мкм), а также требуемую плоскостность ее полированных поверхностей (отклонение от параллельности поверхностей не более 1 мкм по всему диаметру). При этом исключительно важна и чистота поверхности, поэтому пластину постоянно очищают, удаляя посторонние вещества различными методами.

Всю пластину разбивают на фрагменты (на прямоугольники, подобно тому, как это делается на поверхности вафли), которые в конце изготовления будут служить основой чипа и соответствующей микросхемы. Затем выполняют цикл операций по формированию на пластине структурных элементов каждого чипа и их соединений. Часто говорят, что выполняют технологический цикл операций первого этапа – формирования электрической схемы чипа.

Для реализации структурных полупроводниковых элементов электрической схемы с требуемым типом электропроводности и удельным объемным сопротивлением, в определенные места пластины вводят соответствующие примеси. Геометрия и расположение (конфигурация) легированных областей задается обычно

«масками», которые формируются на поверхности пластины с помощью литографии (литография – процесс формирования отверстий в масках). «Маски» позволяют «рисовать геометрию» элементов (вводить примесь в определенную область с размером до 1 мкм). Рисунок будущей маски, которую разрабатывает конструктор микросхемы, задается фотошаблоном. Маска, в первом приближении, представляет собой тонкую непрозрачную (с точки зрения проникновения примесных атомов) пленку с окнами (через них атомы примеси могут проникать), соответствующими конфигурации расположения элементов. В качестве маски могут использоваться нанесенные на пластину пленки диоксида SiO_2 или нитрида Si_3N_4 кремния с вытравленными в них специальным образом окнами. Окна позволяют ввести примесь в требуемое место чипа.

Введение примесей нужной концентрации в пластину через окна маски осуществляют с помощью технологической операции называемой диффузией примеси. При высокой температуре примесные атомы поступают через окно на поверхность пластины и распространяются вглубь ее. В результате одной операции формируется целая система областей с соответствующей геометрией и дозировкой примеси и, в целом, определенный фрагмент будущей структуры ИМС.

Для создания нескольких слоев с различным типом электропроводности, диффузия проводится многократно с использованием соответствующих комбинаций необходимых примесей и специально разработанных комплектов масок. Задумка конструктора микросхемы последовательно, используя различное число технологических операций, «переводится» в пространственное расположение различных многослойных областей с различным типом электропроводности. Тем самым, на пластинах образуется единая матрица полупроводниковых структур, соответствующая реализации электрической схемы ИМС. Благодаря планарности пластины, удается одновременно, как говорят, групповым методом, сформировать на одной пластине примерно 7 тысяч чипов и, соответственно, миллионы идентичных элементов (транзисторов) каждого чипа. При этом следует иметь в виду, что в технологическом цикле обычно участвует до 10 пластин.

Для введения примесей через маску диоксида SiO_2 или нитрида Si_3N_4 кремния может использоваться технологическая операция, получившая название – ионное легирование. Получение соответствующих ионов примеси, их ускорение, фокусировку ионного пучка и «бомбардировку» ионами примеси поверхностного слоя пластины осуществляют в специальных установках. Так как ионный пучок направлен перпендикулярно поверхности, то горизонтальные размеры легированной области точно соответствуют отверстию в маске. Это является преимуществом по сравнению с легированием путем диффузии, так как позволяет получить области меньших размеров (слои с субмикронными горизонтальными размерами и толщиной менее 0,1 мкм). Принципиально возможно локальное легирование без применения маски с помощью сканирования остросфокусированного ионного пучка по поверхности чипа, включаемого и выключаемого по заданной программе.

Завершает процесс изготовления микросхем (чипов) на пластине формирование тонкопленочных проводников, соединяющих элементы, и контактов, необходимых в дальнейшем для подсоединения к ним внешних выводов. В ходе процесса металлизации в нужных для соединения элементов местах осаждаются тонкие металлические пленки из алюминия (удельное поверхностное сопротивление $2,6 \cdot 10^{-6}$ Ом·см и, соответственно, при толщине слоя 0,5 – 1 мкм сопротивление слоя равно 0,025 – 0,05 Ом). Так как алюминий может взаимодействовать с кремнием, в результате чего могут возникать разрывы проводящих участков, используют двухслойные проводники. В качестве проводников первого слоя используют силициды (соединения с кремнием) тугоплавких металлов: вольфрама, молибдена и др. В качестве второго слоя по-прежнему выступает алюминий. После создания рисунка соединений производится вжигание контактов при температуре примерно 550 градусов.

После того, как выполнены описанные ранее технологические операции и чип становится работоспособным, выполняется второй этап создания микросхем: сборочно-контрольный. Он начинается с операций контроля функционирования микросхем на пластине. С помощью механических зондов – тонких игл – к каждому чипу подключается испытательный прибор, который производит проверку

микросхем и выявляет дефектные чипы. После контроля пластины «разламывают» на кусочки (отсюда и название chip – осколок), соответствующие отдельным микросхемам. Дефектные чипы отбраковывают, а годные помещают в корпус микросхемы. Корпус предохраняет чип от механических, химических и других воздействий внешней среды.

В общих чертах процесс установки чипа в корпус и присоединения к нему выводов сводится к следующему. Чип закрепляют с помощью специального клея к металлическому или керамическому основанию, а к его контактными площадкам и выводам корпуса микросхемы с помощью ультразвуковой сварки или термокомпрессии (сжатия при высокой температуре) присоединяют алюминиевые проволочки. Металлические основания обычно выполнены из ковар. Ковар – это сплав железо – никель – кобальт, который хорошо согласуется с кремнием по температурному коэффициенту расширения. Затем проводят герметизацию корпуса. В конце концов, проводят выходной контроль и испытания готовых микросхем с помощью автоматизированных систем, работающих по заданной программе.

3.1.4. Особенности элементов электрических схем интегрального исполнения

Интегральная схемотехника основывается на полупроводниковых структурах. Специфика этих структур оказалась настолько существенной, что привела к значительному изменению ряда схемных решений, существовавших в дискретной электронике.

Характерными особенностями выполнения структурных элементов современных ИМС являются следующие.

1. По ряду причин, связанных с электрофизическими параметрами полупроводниковых материалов (в первую очередь, из-за того, что они занимают меньше места на кристалле), основным элементом электрических схем, реализуемых в интегральном исполнении, является транзистор.

Широко используют биполярные и полевые транзисторы с изолированным затвором и индуцированным каналом. В цифровых схемах

широко используют транзистор Шоттки (биполярный транзистор с диодом Шоттки). Также применяют специфические транзисторы – многоэмиттерные, многоколлекторные и «супербетта».

Конструкции биполярных транзисторов интегрального исполнения различаются, прежде всего, способами их изоляции друг от друга.

В первых микросхемах, из-за простоты технологии, применяли, «горизонтальные» эпитаксиально – планарные транзисторы с изоляцией обратно смещенным $p - n$ переходом. Вы, наверное, помните, что электронно-дырочный переход при обратном включении обладает большим сопротивлением. Однако хорошая изоляция $p - n$ переходом требует глубокой диффузии соответствующих примесей и относительно длительной термообработки. При этом область изоляции в значительной степени «размывается», вследствие боковой диффузии, в результате чего отдельные транзисторы занимают относительно большую площадь. Кроме того, большая площадь изолирующего перехода вносит пропорциональную ей «паразитную» барьерную емкость, что сказывается на частотных характеристиках микросхемы (уменьшается граничная частота и увеличивается время переключения импульсных схем). Изоляция этого типа обладает также низкой тепловой стойкостью (обратный ток $p - n$ перехода резко увеличивается при повышении температуры).

Все это обусловило необходимость широкого применения изоляции транзисторов друг от друга диэлектриком. По конструкции исполнения боковых изолирующих областей различают изоляцию диоксидом кремния SiO_2 (ширина изолирующей области ~ 7 мкм), изоляцию нитридом кремния Si_3N_4 в виде V – образной канавки с последующим заполнением ее поликристаллическим кремнием (ширина ~ 5 мкм), изоляцию в виде аналогичной U – образной канавки (ширина области ~ 3 мкм). При изоляции диэлектриком увеличивается напряжение пробоя изолирующего перехода до 50 вольт. Однако, изоляция диэлектриком стоит дороже.

Основным методом изоляции элементов современных биполярных микросхем является способ комбинированной изоляции, сочетающей изоляцию диэлектриком и $p - n$ переходом, включенным в обратном направлении.

В качестве биполярных транзисторов ИМС широко применяются *эпитаксиально - планарный $n^{++} - p - n$ и изопланарный $n^{++} - p - n$ тран-*

зисторы. Изопланарный транзистор, по сравнению с первым, имеет несколько лучшие импульсные и частотные характеристики. Кроме того, при одинаковой площади эмиттерного перехода общая площадь изопланарного транзистора меньше почти на порядок.

В последнее время в качестве эмиттерных контактов при изготовлении БТ применяют пленки поликристаллического кремния. Это позволяет получить следующие преимущества: а) повысить коэффициент усиления по току β ; б) значительно уменьшить сопротивление базы и, тем самым, повысить быстродействие транзистора. Усиление коэффициента усиления по току обусловлено образованием тонкой пленки окиси на поверхности раздела монокристаллического и поликристаллического кремния, которая обладает свойствами барьера, препятствующего инжекции дырок из базы в эмиттер (увеличивается коэффициент инжекции \mathcal{U}).

При изготовлении транзисторов ИМС, работающих до частоты ~ 1 ГГц, и сверхбыстродействующих сверхбольших ИМС применяют технологии, основанные на принципе самосовмещения. Это позволяет: а) существенно уменьшить емкость перехода коллектор – база; б) значительно уменьшить сопротивление базы. Самосовмещение и повышению их быстродействия ведет к увеличению плотности «упаковки» элементов.

Для создания аналоговых и цифровых схем кроме $n^{++} - p - n$ транзисторов необходимы также $p^{++} - n - p$ транзисторы. Однако, как оказалось, планарные $p^{++} - n - p$ транзисторы в ИМС имеют электрические параметры намного хуже транзисторов $n^{++} - p - n$. Это связано со следующими причинами.

А. Для обеспечения максимального коэффициента передачи по току α необходимо иметь коэффициент инжекции γ близкий к единице. А это возможно только тогда, когда эмиттер будет легирован намного сильнее, чем база. Поскольку базовые и эмиттерные области транзистора формируются диффузией то, вследствие того, что предельная растворимость фосфора больше, чем растворимость бора, получить $n^{++} - p$ области значительно проще, чем $p^{++} - n$.

Б. В $n^{++} - p - n$ транзисторах неосновными носителями базы являются электроны. Их подвижность приблизительно в 2,5 раза превышает подвижность дырок. Следовательно, время пролета базы в $n^{++} - p - n$

транзисторе меньше, чем у транзисторов $p^{++} - n - p$. Граничная частота получается выше.

То есть, «горизонтальные» $p^{++} - n - p$ транзисторы имеют существенно меньший коэффициент передачи по току α и граничную частоту. Наиболее эффективный путь улучшения параметров $p^{++} - n - p$ транзисторов это использование «вертикальной» структуры. Однако для ее формирования требуются дополнительные технологические операции. Вертикальный транзистор характеризуется более высоким коэффициентом передачи и граничной частотой. Типичные значения параметров биполярных транзисторов в ИМС приведены в таблице 3.1.

Таблица 3.1

Типичные значения параметров биполярных транзисторов

| Тип транзистора | Коэффициент усиления по току β | Предельная частота усиления по току |
|---|--------------------------------------|-------------------------------------|
| Эпитаксиально - планарный $n^{++} - p - n$ | 50 – 200 | ~500 МГц |
| Изопланарный $n^{++} - p - n$ транзистор | 50 – 150 | ~ (1 – 2) ГГц |
| Поликремниевый эмиттер самосовмещенный $n^{++} - p - n$ | 200 – 300 | ~ (2 – 4) ГГц |
| Горизонтальный $p^{++} - n - p$ транзистор | 5 – 20 | ~1 МГц |
| Вертикальный $p^{++} - n - p$ транзистор | 30 – 50 | ~(100 – 500) МГц |

Таблица 3.2

Сравнение технологий изготовления биполярных и полевых транзисторов

| Биполярная технология | Технология МОП |
|---|---|
| Используется 130 технологических операций | Используется 40 технологических операций |
| Применяется 10 процессов с температурой около 1000 ⁰ С | Применяется 2 процесса с температурой около 1000 ⁰ С |
| Используется 8 фотошаблонов | Используется 5 фотошаблонов |

Большинство цифровых ИМС создано на основе полевых транзисторов с изолированным затвором. Обычно их называют *МОП – транзисторами* (металл – оксид – полупроводник).

Следует сразу отметить, что, с точки зрения технологии, МОП транзистор существенно проще биполярного. Это легко заметить из таблицы 3.2, в которой проводится их общее сопоставление.

Важно также то, что МОП транзистор занимает на кристалле малую площадь и, обладая высоким входным сопротивлением, потребляет малую мощность от источника сигнала. Как мы увидим в дальнейшем, имеются такие схемные решения на МОП транзисторах, при которых устройство потребляет малую мощность от источников электрической энергии (источников питания) только во время работы.

Технология МОП позволяет создавать три базовые структуры логических элементов – инверторов (элементов логики «НЕ»), которые, в области цифровой обработки информации, получили название цифровых ИМС с определенной логикой: с каналом p – типа (логика p – МОП); с каналом n – типа (логика n – МОП); с использованием *комплементарных* (т. е. взаимно дополняющих друг друга) транзисторов с каналом n – и p – типа (логика КМОП).

Благодаря использованию слаболегированных подложек, с высоким удельным сопротивлением, поликристаллических кремниевых затворов, вместо металлических, многослойной металлизации, принципов самосовмещения удалось преодолеть трудности, существовавшие ранее при изготовлении МОП транзисторов с каналом n – типа. Поэтому такие ПТ оказались весьма привлекательными для изготовления цифровых ИМС сверх высокой степени интеграции, для которых характерны очень большая плотность «упаковки» транзисторов, простота изготовления, высокое быстродействие. Примерами, где используется n – логика, служат высокопроизводительные процессоры и микроконтроллеры фирм Motorola, Intel.

Комплементарная МОП-пара (КМОП) образована двумя МОП – транзисторами, размещенными на общей подложке. Один транзистор имеет канал n – типа, другой – канал p – типа. КМОП-структуры являются предпочтительными, в первую очередь, благодаря малому потреблению энергии от источников питания, широкому диапазону напряжений питания. Быстродействие КМОП – устройств не ниже быстродействия n

– МОП ИС, а в ряде случаев и несколько выше. Достоинством таких схем является устойчивость к перегрузкам, а также возможность их выполнения на одном кристалле (чипе) вместе с аналоговыми биполярными цепями.

Недостаток КМОП – структуры обусловлен необходимостью применять транзисторы с каналом p – типа, которые занимают на кристалле «большую» площадь и имеют меньшее быстродействие. Однако, тщательное конструирование и правильный выбор технологических процессов позволяет КМОП ИМС успешно конкурировать с изделиями, выполненными по технологии n –МОП.

В большинстве микросхем используются полевые транзисторы с «горизонтальным» каналом (параллельным поверхности пластины), однако применяются транзисторы и с «вертикальным» каналом, образующимся на стенках вытравленных канавок. Такие транзисторы занимают меньшую площадь на кристалле.

Вертикальные структуры и ИМС на их основе принадлежат к направлению микроэлектроники, получившей название «трехмерной» интеграции, которое в настоящее время интенсивно развивается. Важным направлением трехмерной электроники являются многомерные структуры, содержащие не один, а несколько слоев, в которых транзисторы размещаются один над другим в несколько «этажей».

При одинаковой конструкции n – канальные транзисторы имеют более высокую крутизну и, вследствие большей подвижности электронов по сравнению с дырками, более высокую граничную частоту, чем p – канальные. Поэтому логика p – МОП в последнее время применяется сравнительно редко.

1. Диоды специально не изготавливают, а используют транзисторы в диодном включении (у таких транзисторов коллектор соединен с базой).

В качестве стабилитронов используют переход эмиттер база n – p – n транзистора, напряжение пробоя которого составляет 6 – 8 вольт.

2. В полупроводниковых ИМС стараются избегать использовать резисторы. Это обусловлено следующим. Точное значение величины резистора получить очень трудно (точность $\sim 30\%$), «подогнать» ее до требуемого значения нельзя, величина сопротивле-

ния сильно зависит от температуры, они занимают много места. Если все-таки избежать применения резисторов не удастся, то в качестве резистора используют изолированные от остальных частей полупроводниковые области кристалла. Низкоомные резисторы изготавливают также из поликремния. Кроме того, используя поликремний, можно получить резисторы относительно больших номиналов с малыми геометрическими размерами и пониженными паразитными емкостями. Диапазон значений сопротивлений, которые можно создать в ИМС, составляет 100 Ом – 50 кОм.

3. В качестве конденсаторов используют барьерные емкости коллекторного и эмиттерного переходов транзистора. Емкости величиной более 200 пФ получить очень трудно. При этом такие конденсаторы могут работать при подаче на них напряжения одной полярности.

4. Индуктивные элементы практически не применяют, так как изготовить их очень трудно. Не удастся изготовить в интегральном исполнении трансформаторы тока и напряжения, переменные резисторы и другие настроечные элементы.

3.2. Особенности и примеры схемных решений типовых узлов микроэлектронных изделий

3.2.1. Необходимость нового подхода к созданию схем

Технологические особенности элементов ИМС потребовали во многих случаях по - новому подходить к реализации схем типовых, казалось бы, ставших уже «классическими» в дискретной схемотехнике, узлов электронных изделий.

При разработке электрической схемы надо было: во-первых, учесть, что технологически в полупроводниковых ИМС получать транзисторы и диоды значительно проще и намного дешевле, чем пассивные элементы (резисторы и конденсаторы); во-вторых, помнить, что серьезным препятствием на пути развития микроэлектроники является ограниченный набор элементов, в частности, отсутствие индуктивностей, конденсаторов большой емкости и резисторов номиналов больших 50 кОм. Наконец, в-третьих, при изготовлении ИМС сложно было достигнуть вы-

сокой точности воспроизведения заданных параметров элементов, их температурной стабильности. В дискретной схемотехнике, во многом, все было наоборот. Например, транзисторы стоили намного дороже, чем резисторы и конденсаторы. Параметры элементов перед установкой в схему измерялись и были известны довольно точно.

Создание электрических схем типовых узлов в ИМС пошло по пути использования несколько иных принципов, положенных в основу создания базовых устройств, нежели это было в дискретной схемотехнике.

Интегральная технология дает возможность, вследствие того, что используются групповые способы изготовления, получить на одной подложке, на одном чипе совокупность элементов у которых характеристики и параметры взаимно согласованы (если ввели несколько меньше донорной примеси в эмиттер транзистора, то не довели ее всем тысячам и миллионам транзисторов, расположенным на подложке). Однотипные элементы в этом случае имеют одинаковые параметры во всем диапазоне внешне допустимых воздействий. По этой причине стал применяться принцип создания схем, при котором достигается взаимная компенсация нестабильности параметров отдельных одинаковых элементов.

Кроме того, в основу создания схем закладывается принцип минимизации не количества элементов схемы, а минимизации площади кристалла и простоты технологической ее реализации. В ИМС может быть, с точки зрения «классической дискретной» схемотехники, избыточное количество простых в изготовлении и «дешевых» транзисторов, лишь бы при этом была минимальной площадь кристалла и не надо было реализовывать «технологически невыгодных» резисторов и конденсаторов большой величины.

Наконец, когда решается задача объединения различных узлов (схем) в одно устройство, то обязательно используется принцип, при котором реализуется согласование между собой сопрягаемых узлов, чтобы, в конечном счете, построить устройство, являющееся оптимальным (наилучшим) с точки зрения решения общей поставленной задачи. То есть, в схеме могут быть устройства, казалось бы «бессмысленные» с точки зрения исполнения одного каскада, но которые «работают» на «решение общей задачи», причем на ее оптимальное решение.

Особенности схемотехнического исполнения типовых устройств ИМС мы и рассмотрим в дальнейшем.

3.2.2. Особенности схем установки рабочей точки во входной цепи усилительного каскада интегрального исполнения на биполярном транзисторе в СОЭ

Мы уже рассматривали схемы установки рабочей точки во входной цепи биполярного транзистора в СОЭ с помощью «гасящего» резистора и с помощью делителя напряжения на двух резисторах. Эти схемные решения требуют, как известно, резисторов больших номиналов, что в ИМС неприемлемо. Кроме того, чтобы источник сигнала не влиял на «установленную рабочую точку» требовалось использовать на входе усилителя (а также на его выходе) «развязывающий» конденсатор большой емкости, который бы благодаря своему заградительно-шунтирующему свойству, пропускал бы на вход транзистора переменный сигнал и «отключал бы от входа» источник сигнала по постоянному току. Понятно, что развязывающие конденсаторы емкостью в сотни микрофарад также использовать в ИМС невозможно.

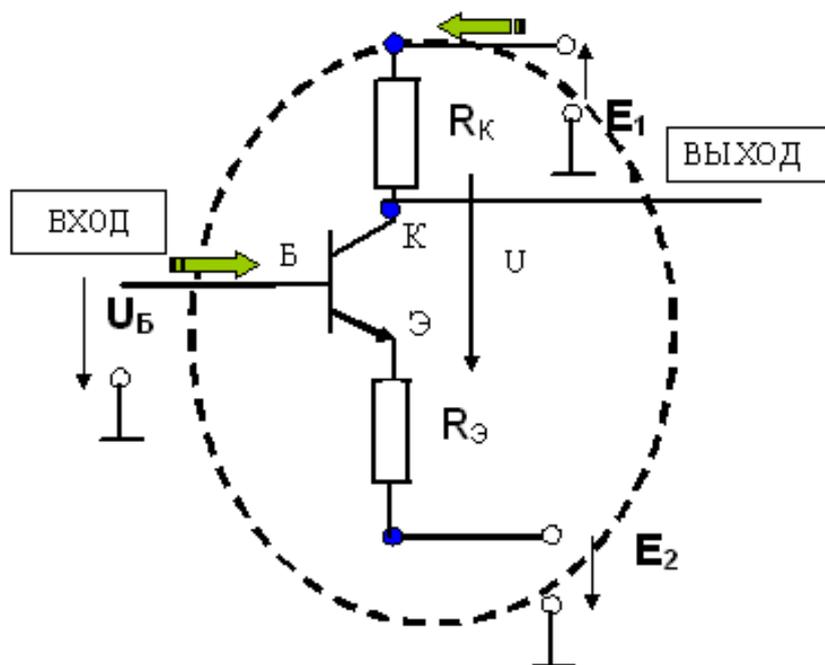


Рис. 3.1. Схема каскада усилителя на биполярном транзисторе с двуполярным питанием

По названным причинам в интегральной схемотехнике используют так называемые непосредственные (без конденсаторов) связи между каскадами. Обычно при этом (рис. 3.1) используют два источника с ЭДС противоположной полярности E_1 , E_2 (говорят – «расщепленные» источники). Как следует из второго закона Кирхгофа $U_B = E_1 - E_2$, и, следовательно, потенциал базы можно сделать равным нулю, взяв величины ЭДС E_1 и E_2 одинаковыми. Тогда, преимущество такой схемы состоит в том, что отпадает необходимость в конденсаторе большой емкости на входе и в «гасящем резисторе» или делителе напряжения. Достаточно на вход такой схемы подключать источник переменного сигнала малой амплитуды, у которого отсутствует на выходе постоянное напряжение. Когда источник напряжения непосредственно включается на вход усилителя, потенциал базы по прежнему будет равен нулю.

Повысить температурную стабильность работы, а также улучшить ряд характеристик усилителя с двухполярным питанием и непосредственными связями позволяет включение в цепь эмиттера небольшого резистора $R_Э$, приемлемого для ИМС по величине сопротивления. Это резистор, как вскоре узнаем, создает «обратную связь по току» за счет которой и достигается положительный эффект.

3.2.3. Использование «эталонов» тока и напряжения – характерная черта каскадов интегрального исполнения

«Эталоны» это вспомогательные устройства интегральной схемотехники, которые определенным образом замещая идеальные источники тока и напряжения, позволяют обеспечивать надежную и стабильную работу интегральных устройств в широких диапазонах изменения температур, питающих напряжений. С их помощью можно реализовать (наилучшим образом имитировать) резисторы больших номиналов сопротивлений, обеспечить «согласование» каскадов по напряжению (к примеру, соединение нескольких каскадов усилителей с непосредственными связями) и прочее.

Наиболее часто в ИМС используют идеальные источники тока, стараясь получить внутри схемы ток определенного значения, который

не должен меняться при изменении падения напряжения на элементах (сопротивления нагрузки), ЭДС источников питания, температуры окружающей среды. Их часто называют источниками стабильного тока (ИСТ) или генераторами стабильного тока (ГСТ).

Ранее мы изучили независимые и управляемые источники тока. Было отмечено, что ИСТ может быть охарактеризован двумя видами его внутренних сопротивлений: сопротивлением по постоянному току (статическим сопротивлением) R_{0cm} и сопротивлением по переменному току (дифференциальным) $R_{диф}$:

$$R_{0cm} = \frac{U_0}{I_0}, R_{диф} = \frac{dU}{dI}. \quad (3.1)$$

В идеале ГСТ выдает стабильный ток и на его величину не оказывают влияние дестабилизирующие факторы (изменения температуры, напряжений питания, сопротивлений нагрузки и прочее), если его внутреннее дифференциальное сопротивление стремится к бесконечности.

ИМС «питаются» от источников электрической энергии, обладающих малым внутренним сопротивлением (другими словами от идеальных источников ЭДС).

Чтобы получить ГСТ, надо «увеличить внутреннее сопротивление» источника напряжения. Такое решение годится только в том случае, если использовать резистор, находящийся вне ИМС. Очевидно, что последовательное подключение к идеальному источнику ЭДС резистора с большим сопротивлением, располагающегося внутри ИМС неприемлемо с точки зрения технологии его изготовления. Само собой напрашивается решение: последовательно с источником ЭДС включить резистор очень большого сопротивления и тем самым создать источник стабильного тока на все микросхемы, а внутри ИМС «общий» стабильный ток «распределять» с помощью специальных устройств, отвечающих требованиям интегральной технологии.

Такая идея будет успешно работать, если, к тому же, обеспечить, чтобы потребитель тока имел малое внутреннее сопротивле-

ние. В качестве желаемой нагрузки, в простейшем случае, может выступить обладающий малым сопротивлением полупроводниковый диод (рис. 3.2) или интегральный биполярный транзистор в диодном включении (у которого соединены накоротко коллектор и база).

Для тока, втекающего в ИМС, будет выполняться соотношение:

$$I_{И} = \frac{E - U_{Д}}{R_{ВН}} \cong \frac{E}{R_{ВН}}. \quad (3.2)$$

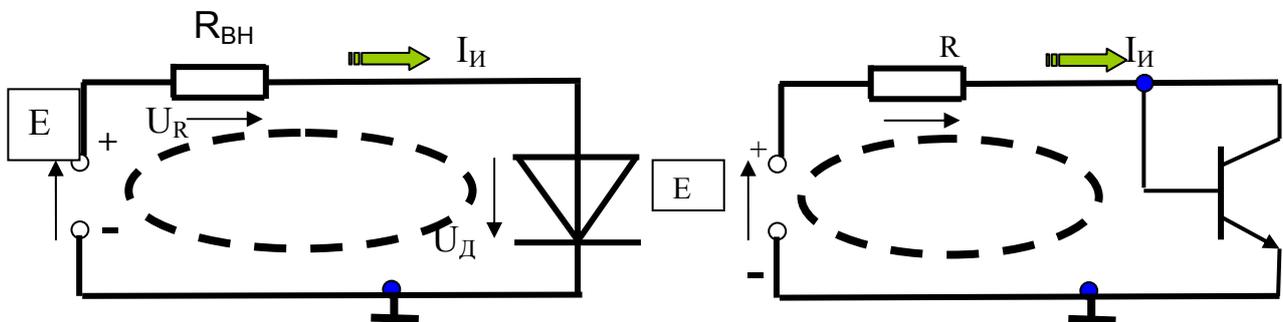


Рис. 3.2. Схемы простейших источников стабильного тока

Так как E и $R_{ВН}$ стабильны, то ток, как частное от деления, также будет весьма стабильным.

Чтобы «распределить» втекающий в ИМС ток, между ее многочисленными внутренними потребителями, используют два варианта устройств: «повторители» тока; «отражатели» тока.

Типичной схемой источника тока, которая «повторяет» ток, полученный от общего источника питания, является ГСТ на полевом транзисторе, схема которого показана на рис. 3.3.

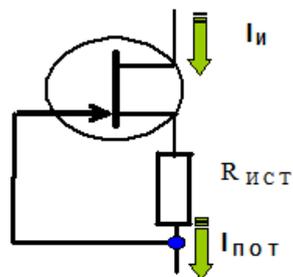


Рис. 3.3. Схема источника тока на полевом транзисторе

Можно показать, что ГСТ на полевом транзисторе, который является двухполюсным компонентом электрической цепи, характеризуется следующими параметрами»:

сопротивлением по постоянному току:

$$R_{0cmГСТ} = \frac{U_{СИ}}{I_{И}} + R_{ИСТ}, \quad (3.3)$$

дифференциальным сопротивлением ГСТ:

$$R_{дифГСТ} = \frac{dU}{dI} \approx (1 + SR_{ИСТ}) \frac{U_A}{I_{И}}, \quad (3.4)$$

где $U_A \sim (100 - 300)В$, и

крутизной стокзатворной характеристики:

$$S = \frac{2\sqrt{I_{И}I_{Снач}}}{U_{3Иотс}}. \quad (3.5)$$

Резистор автоматического смещения $R_{ИСТ}$, помимо того, что он задает напряжение рабочей точки в цепи затвор – исток, «делает» за счет «обратной связи по току» стоковые характеристики транзистора более горизонтальными, а, следовательно, увеличивает внутреннее дифференциальное сопротивление ГСТ.

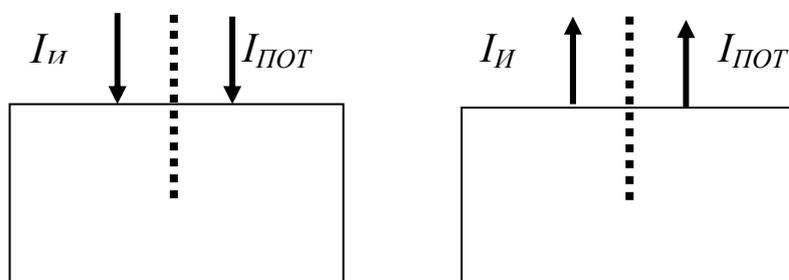


Рис. 3.4. Понятие отражателя тока

Отражатели тока (а их еще называют токовыми зеркалами) получили свое название по той причине, что ток, который они отдают потребителю $I_{\text{ПОТ}}$, является «отражением» поступающего общего тока $I_{\text{И}}$. Ток правой части схемы (рис. 3.4) как бы является зеркальным отражением левой части схемы (относительно штриховой линии).

Схема широко используемого в настоящее время отражателя тока (с эмиттерной стабилизацией и термокомпенсацией) показана на рис. 3.5 А.

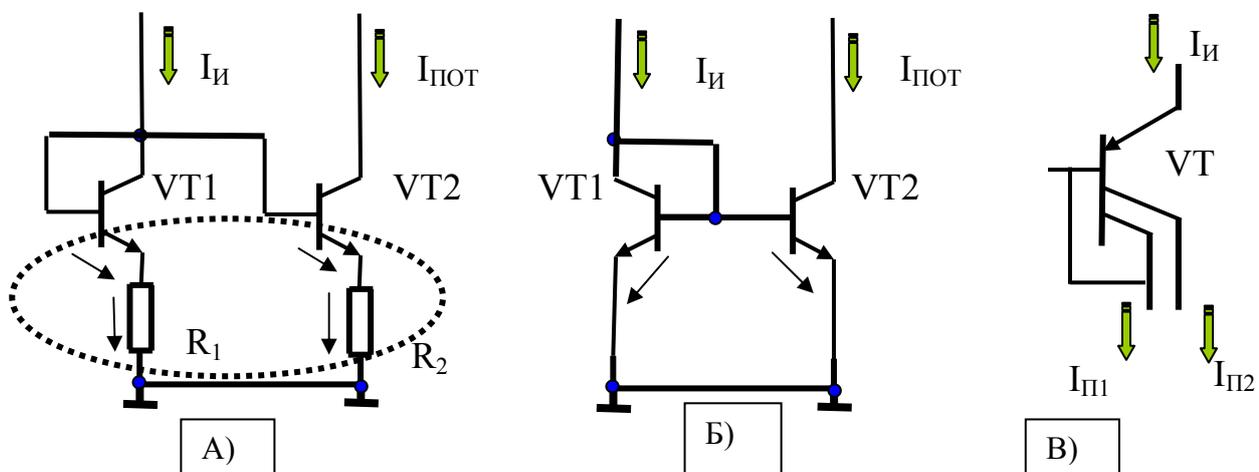


Рис. 3.5. Схемы отражателей тока на биполярных транзисторах

На схеме рис. 3.5 А транзистор VT1, в диодном включении, создает малое сопротивление для «общего» источника тока, подобно тому как показано на рис. 3.2, обеспечивая тем самым стабильность общего тока. Транзистор VT2 создает «частный» ток для потребителя, расположенного внутри микросхемы.

Если предположить, что транзисторы VT1 и VT2 идентичны (следовательно, у них почти тождественно равное пороговое напряжение $U_{\text{БЭ}}$), имеют большой и одинаковый коэффициент усиления по току β , то можно определить, как связаны между собой общий и частный ток. Действительно, согласно второго закона Кирхгофа, (при обходе контура против часовой стрелки) имеем:

$$U_{БЭ1} + U_{R1} - U_{БЭ2} - U_{R2} = 0, \Rightarrow U_{R1} = U_{R2} \Rightarrow \quad (3.6)$$

$$I_{ПОТ} = I_{И} \frac{R_1}{R_2}. \quad (3.7)$$

При условии, что $R_1 = R_2$ получаем, что ток, отдаваемый потребителю, расположенному внутри ИМС, является, по существу, зеркальным отражением общего тока источника.

Используют источники тока, в которых сопротивления резисторов не равны. В них происходит передача тока потребителю определенной доли общего тока.

ГСТ с эмиттерной стабилизацией характеризуется следующими показателями.

А. Сопротивление по постоянному току ГСТ равно:

$$R_{0cmГСТ} = R_{КЭ} + R_2 = \frac{U_{КЭ}}{I_{И}} + R_2. \quad (3.8)$$

Б. Внутреннее дифференциальное сопротивление ГСТ равно:

$$R_{дифГСТ} = F \frac{dU_{КЭ}}{dI_{ПОТ}} \approx F \frac{U_{ЭРЛИ}}{I_{ПОТ}}, \quad (3.9)$$

где: $U_{ЭРЛИ} \sim 100 - 300$ В, $F = 40 - 50$ – коэффициент, определяющий глубину отрицательной обратной связи (растет с увеличением сопротивлений R_1 и R_2).

В. Относительное изменение тока в процентах при изменении напряжения на один вольт:

$$\delta_U = \frac{100}{I_{ПОТ}} \frac{dI_{ПОТ}}{dU_{КЭ}} = \frac{100}{I_{ПОТ} R_{дифГСТ}}. \quad (3.10)$$

Г. Относительное изменение тока в процентах при изменении температуры на один градус Цельсия:

$$\delta_T = \frac{100TKI_{II}}{I_{ПОТ}} = \frac{-1}{R_2 I_{ПОТ}} \frac{dU_{БЭ}}{dT} \frac{100}{1 + R_{II}/R_2}. \quad (3.11)$$

Использование резисторов в ИМС не всегда приемлемо, поэтому в 1965 году Вайдлером была предложена схема простого отражателя без резисторов (рис. 3.5 Б). Показатели такого ГСТ несколько хуже, чем ранее рассмотренного.

Для одновременной стабилизации тока для нескольких различных «потребителей» ИМС используют многоплечевые ГСТ (рис. 3.5 В) в которых используют многоколлекторные биполярные транзисторы (ГСТ с «расщепленным» коллектором). За счет выбора размеров коллекторов в таких ГСТ можно менять «долю» тока, отдаваемого каждому потребителю. Поскольку такие ГСТ работают в инверсном активном режиме, при котором величина β невелика, то показатели такого ГСТ несколько хуже ранее рассмотренных.

В современных ИМС применяют весьма сложные отражатели тока, которые обеспечивают более высокие показатели, чем рассмотренные ГСТ.

3.2.4. Использование повторителей напряжения в ИМС

Наряду с ГСТ в ИМС широко используют повторители напряжений. Название их вытекает из того обстоятельства, что их коэффициент усиления (передачи) по напряжению близок к единице, поэтому на выходе они «повторяют» напряжение на входе. Причем речь идет не только о повторении амплитуды гармонического сигнала, но и фазы. Начальные фазы входного и выходного напряжения равны, поэтому в повторителях напряжения синфазны.

Повторитель напряжения, выполненный на биполярном транзисторе, называют *эмиттерным повторителем*. В таком устройстве транзистор включен по схеме с общим коллектором. В *истоковых повторителях* полевой транзистор включен по схеме с общим стоком.

На первый взгляд может показаться, что повторители это устройства, которые не очень важны. Ведь два проводника также «повторяют» напряжение. На самом деле это не так. Повторители напряжения на транзисторах обладают рядом уникальных свойств, благодаря которым их широко используют как для «согласования» каскадов, так и для реа-

лизации оптимальных возможностей всего устройства, выполненного на ИМС. Схема простейшего эмиттерного повторителя на биполярном транзисторе показана на рис. 3.6.

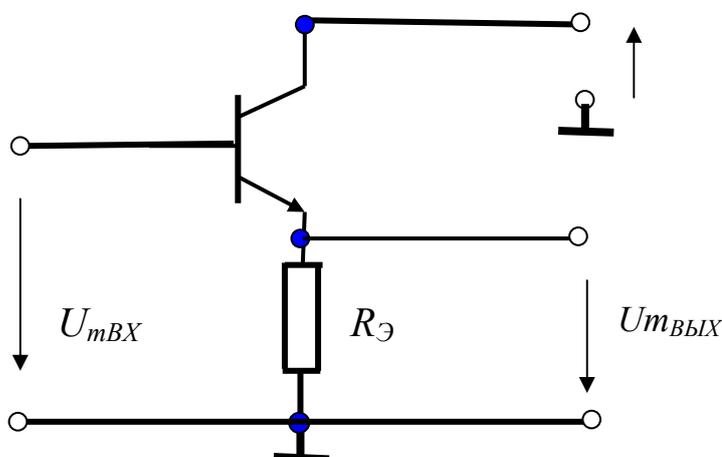


Рис. 3.6. **Схема простейшего эмиттерного повторителя на биполярном транзисторе**

Приведем без вывода отдельные показатели эмиттерного повторителя (ЭП), чтобы понять, в чем состоят его достоинства:

А) хотя ЭП не усиливает сигнал по напряжению, он усиливает сигнал по току, а, следовательно, и по мощности:

$$K_I = \frac{I_{mBBI}}{I_{mBX}} \approx \beta + 1. \quad (3.12)$$

Б) Входное сопротивление эмиттерного повторителя:

$$R_{BX} = \frac{U_{mBX}}{I_{mBX}} \approx \beta R_{\text{Э}} \quad (3.13)$$

велико и, примерно, в β раз больше сопротивления нагрузки. С точки зрения теории электрических цепей, повторитель является конвертором (преобразователем) сопротивления, т. е. он как бы преобразует малое сопротивление на выходе ЭП в большое сопротивление на входе. Это

широко используют для согласования сопротивлений каскадов с целью передачи максимальной мощности в нагрузку. С другой стороны, при большом входном сопротивлении ЭП не потребляет энергии от источника сигнала. По этой причине, как правило, сигналы от маломощных источников, обладающих большим внутренним сопротивлением (датчиков), всегда поступают к последующим устройствам посредством ЭП, который стоит у таких устройств на входе.

В) Выходное сопротивление эмиттерного повторителя:

$$R_{ВЫХ} = \frac{U_{mВЫВ}}{I_{mВЫВ}} \approx r_{\text{Э}} + \frac{R_{ИС}}{\beta} \quad (3.14)$$

мало (при сопротивлении источника сигнала равном нулю примерно равно сопротивлению эмиттерного перехода при прямом включении электронно дырочной структуры). С точки зрения теории электрических цепей, со стороны выхода ЭП представляет источник напряжения с малым внутренним сопротивлением, т. е., по сути, идеальный источник напряжения. Очень важно, что к выходу эмиттерного повторителя можно подключать нагрузки, имеющие малое сопротивление, и это почти не будет сказываться на его работе. По этой причине ЭП используется для обеспечения работы многих исполнительных механизмов характеризующихся малым сопротивлением (соленоидов, шаговых двигателей, устройств с прецизионно управляемым вращением и перемещением, а также кабелей, звуковоспроизводящих динамиков и пр.). Другие схемы включения транзисторов таким свойством не обладают.

Г) ЭП обладает улучшенными частотными свойствами: у него очень высокая граничная частота и, следовательно, ЭП способен равномерно передавать сигнал в весьма широкой полосе частот.

Д) ЭП обладает малым коэффициентом нелинейных искажений и за счет своих свойств ослабляет уровень всякого рода внутренних помех (шумов, различного рода наводок и т. п.). Свойства ЭП делают усиление сигнала по мощности мало зависящим от параметров используемых в нем транзисторов, колебаний напряжения источников питания и прочее.

Многие свойства ЭП зависят от β . В интегральной схемотехнике реализовать большие величины коэффициента усиления по току β трудно, поэтому в ИМС широко используют составные транзисторы.

Если соединить два транзистора, как показано на рис. 3.7 А (коллекторы обоих транзисторов соединены вместе, а эмиттерный ток транзистора Т1 является базовым током транзистора Т2), то полученная схема будет работать как один $n - p - n$ транзистор (его называют *составным транзистором Дарлингтона* или *парой Дарлингтона*), у которого коэффициент усиления по току β будет равен:

$$\beta_{DAR} = \beta_1 \beta_2. \quad (3.15)$$

Пара Дарлингтона имеет очень высокий коэффициент усиления по току. Такой способ имитации одного $n - p - n$ транзистора двумя $n - p - n$ транзисторами весьма полезен в ряде случаев, особенно для схем, работающих с большими токами (для стабилизаторов напряжения, выходных усилительных каскадов, работающих на малоомные нагрузки, и для случаев, когда необходимо обеспечить высокое входное сопротивление).

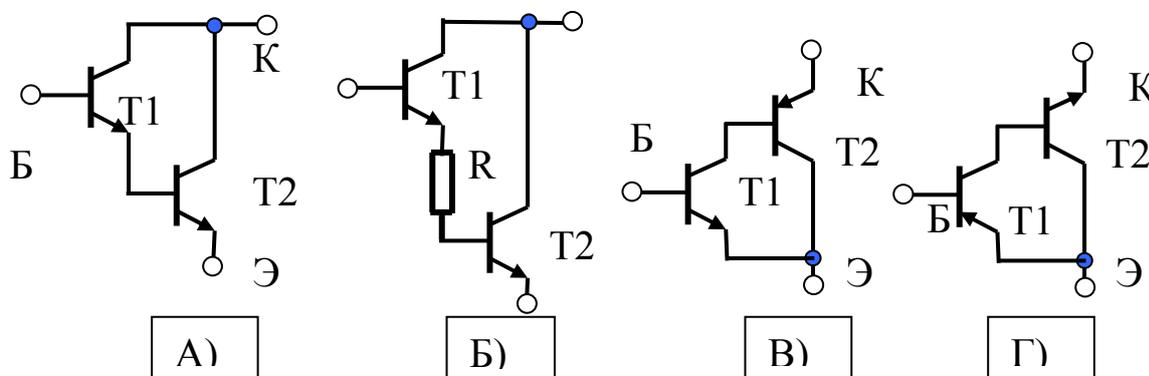


Рис. 3.7. Схемы составных транзисторов Дарлингтона и Шиклаи

В транзисторе Дарлингтона величина напряжения между базой и эмиттером равна примерно сумме пороговых напряжений каждого входящего в его состав транзистора (~1,4 В). Обычно между базой Т2 и эмиттером транзистора Т1 включают резистор (рис. 3.7 Б), который «отводит» часть тока эмиттера первого транзистора от базы второго, и тем самым, предотвращает переход его в режим насыщения. Обычно величина резистора составляет несколько сотен Ом в мощных транзисторах и несколько кОм в маломощных. Промышленность выпускает транзисторы Дарлингтона в виде законченных модулей, в едином корпусе с тремя выводами.

Подобно тому, как мы только что рассмотрели, для образования составного транзистора используют также схему Шиклаи (Sziklai). На рис 3.7 В показан составной транзистор Шиклаи, который ведет себя как транзистор $n - p - n$, обладающий большим коэффициентом β . В транзисторе Шиклаи величина напряжения между базой и эмиттером равна примерно пороговому напряжению одного транзистора (~0,7 В). Транзистор Шиклаи, который ведет себя как транзистор $n - p - n$, часто называют *комплементарным (дополняющим) транзистором Дарлингтона*. На рис. 3.7 Г показан составной транзистор Шиклаи, который ведет себя как транзистор $p - n - p$, обладающий большим коэффициентом β . Так как значение β для горизонтального $p - n - p$ транзистора мало, то использование его в комбинации с транзистором $n - p - n$, соединенном по схеме Шиклаи, позволяет в ИМС повысить коэффициент передачи тока $p - n - p$ транзистора.

Коэффициент передачи тока базы составного транзистора Шиклаи равен:

$$\beta_{SZ} = (1 + \beta_{npn})\beta_{pnp}. \quad (3.16)$$

Следует иметь в виду, что частотные свойства составных транзисторов несколько хуже, чем одиночных.

Использование в эмиттерных повторителях пар Дарлингтона и Шиклаи позволяет значительно снизить выходное сопротивление повторителя, особенно для случаев работы с высокоомными источниками сигналов, и, тем самым, привести его почти к идеальному источнику напряжения.

Выходное сопротивление ЭП на составном транзисторе Дарлингтона равно:

$$R_{ВЫХ}^{ДАР} \approx r_{\text{Э}} + \frac{r_{\text{Э}}}{\beta_2} + \frac{R_{ИС}}{\beta_1\beta_2} \approx r_{\text{Э}}. \quad (3.17)$$

Выходное сопротивление ЭП на составном транзисторе Шиклаи равно:

$$R_{ВЫХ}^{ШИК} \approx \frac{r_{\mathcal{E}}}{\beta_2} + \frac{R_{ИС}}{\beta_1\beta_2} \approx \frac{r_{\mathcal{E}}}{\beta_2}. \quad (3.18)$$

Как следует из формул (3.17 и 3.18), использование составных транзисторов в ЭП позволяет избавиться от влияния внутреннего сопротивления источника сигнала (включенного на входе) на выходное сопротивление ЭП. Важным является и то обстоятельство, что при прочих равных условиях ЭП на составном транзисторе Шиклаи имеет меньшее выходное сопротивление, чем ЭП с парой Дарлингтона.

Из формулы (3.13) несложно усмотреть, что ЭП с составными транзисторами имеют также очень высокое входное сопротивление.

3.2.5. Дифференциальный усилитель – важнейшее звено интегральных схем

Чтобы понять, почему в ИМС широко используются усилители, получившие название дифференциальных, рассмотрим несколько ситуаций, с которыми вы встретитесь на практике.

Для управления каким-либо объектом необходимо знать его состояние. Информацию о состоянии обычно получают с помощью самых разнообразных датчиков (температуры, давления, перемещения, деформации, вибрации, скорости и т. п.). Как правило, несмотря на большое функциональное многообразие и различное их конструктивное исполнение, датчики информацию об указанных неэлектрических величинах выдают в виде электрического сигнала (изменяющегося во времени напряжения). Чтобы получить такой сигнал датчики включают по схеме, которую часто называют мостовой (рис. 3.8). Три «плеча» моста образуют стабильные резисторы высокой точности R_1 , R_3 , R_4 , а в четвертое «плечо» включен резистор, сопротивление которого, под воздействием какого-нибудь фактора, меняется. К примеру, при измерении температуры меняется сопротивление терморезистора; при измерении величины давления или усилия – тензометрическое сопротивление.

В начале мост находится в состоянии баланса $R_2 R_3 = R_1 R_4$ и напряжение на выходе преобразователя равно нулю. Если затем

воздействует какой-либо фактор, то сопротивление R_2 меняется, мост «разбалансируется» и на выходе появляется электрический сигнал (напряжение) пропорциональный величине воздействующего фактора.

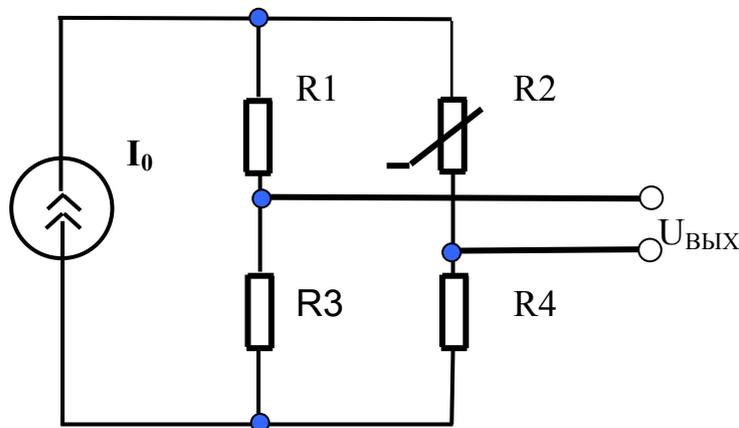


Рис. 3.8. Мостовая схема включения датчика

Данная электрическая схема обладает рядом особенностей:

А) Электрический сигнал на выходе может быть очень малым (единицы милливольт). Поэтому выходной сигнал надо усиливать.

Б) Многие неэлектрические величины (температура, давление) могут меняться очень медленно. Частота изменения сигнала на выходе может составлять единицы, и даже доли, герц.

В) По условиям работы датчики и мостовые преобразователи могут располагаться на большом расстоянии от усилителя и провода, по которым передается сигнал к усилителю, могут быть очень длинными. На длинные провода, как на «антенны», могут наводиться всякого рода помехи. Ситуация усугубляется еще тем, что по условиям работы нельзя обычно ни один провод соединять с точкой нулевого потенциала (заземлять) и за счет этого снижать уровень помех. Так что, чаще всего, на проводах появляются индуктивные помехи от сети, от работающих рядом электротехнических устройств, от передатчиков мобильных телефонов, радио и телевизионных станций. Амплитуда наводимых напряжений может достигать величины сотен вольт.

При этом важно иметь в виду, что поскольку провода от датчиков, как правило, расположены рядом, то на них будут наводить-

ся одинаковые помехи. Такие одинаковые по амплитуде и фазе «наводки» называют синфазными.

Итак, из анализа следует, что во многих практически важных случаях необходимо усиливать медленно или быстро меняющийся сигнал амплитудой всего несколько милливольт на который «налагается» синфазный сигнал помех с амплитудой в сотни вольт.

Простые усилительные каскады, которые мы рассматривали ранее, с такой задачей справиться не могут, поэтому потребовался усилитель принципиально иного типа.

Дифференциальным (т.е. разностным) усилителем (ДУ) называют усилитель (часто говорят усилитель постоянного тока – УПТ) с двумя «незаземленными» входами, осуществляющий усиление разности напряжений, поданных на его входы (дифференциальные сигналы), и, в идеале, абсолютно не реагирующий на одинаковые по величине потенциалы (напряжения), которые одновременно приложены к двум его входам (синфазные сигналы).

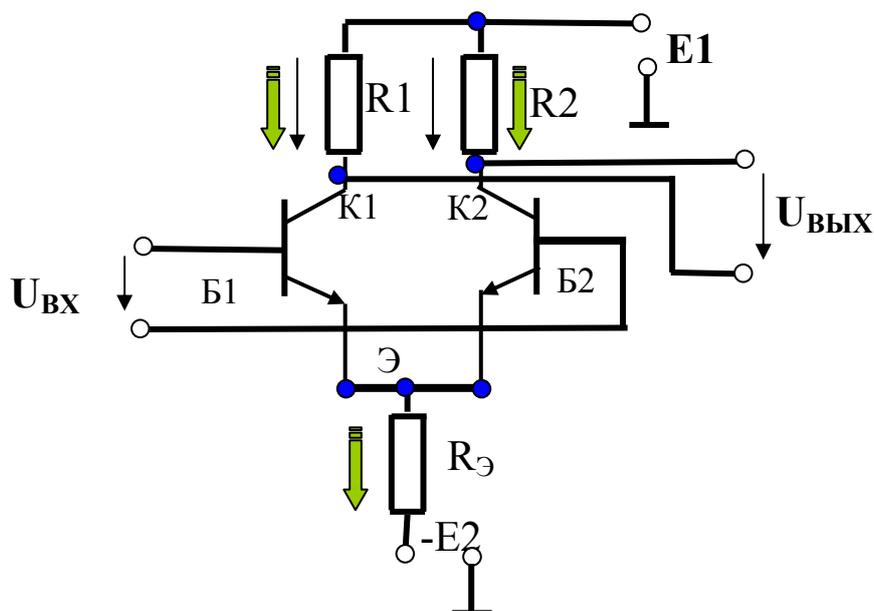


Рис. 3.9. Схема простейшего дифференциального усилителя на биполярных транзисторах

Схема простейшего ДУ показана на рис. 3.9. ДУ выполняют по интегральной технологии, поэтому уместно полагать, что транзисторы и коллекторные резисторы идентичны (одинаковы), а схема

симметрична относительно прямой, проведенной между транзисторами. Каждая «половинка» ДУ подобна схеме усилителя с расщепленным источником (рис. 3.1).

Если на базы транзисторов подать одинаковые потенциалы $\varphi_{Б1} = \varphi_{Б2}$, то в обеих половинах ДУ будут протекать абсолютно одинаковые токи коллекторов. Потенциалы рабочих точек коллекторов также будут одинаковы, поэтому напряжение на выходе, как разность потенциалов $U_{ВЫХ} = \varphi_{К1} - \varphi_{К2}$, будет в идеале равно нулю. Аналогичная ситуация будет во всех случаях, когда на базы будут поданы одинаковые потенциалы, даже если они будут одинаково меняться во времени. Следовательно, все синфазные (т.е. одинаковые по амплитуде и начальной фазе) сигналы ДУ не будут усиливать, а будут их подавлять.

Если провести строгий анализ, то можно показать, что, в первом приближении, коэффициент усиления синфазных помех определяется формулой:

$$K_U^{СИН} = \frac{R_{1,2}}{2R_{\mathcal{E}}}. \quad (3.19)$$

Чтобы синфазные помехи подавлялись, необходимо, чтобы величина сопротивления $R_{\mathcal{E}}$ была очень большой. В интегральной схемотехнике вместо резистора, трудно поддающемуся изготовлению и занимающего большую площадь кристалла, используют ГСТ (отражатель тока), который «имитирует» резистор с большим дифференциальным сопротивлением и малым сопротивлением по постоянному току. Малое сопротивление по постоянному току позволяет, обеспечивая заданное положение рабочей точки, не использовать источники ЭДС E_1, E_2 с большой величиной напряжения.

Если теперь к базам транзисторов приложить одинаковые по величине, но противоположные по знаку потенциалы, то ток коллектора одного транзистора увеличится, а другого – уменьшится. Соответственно, потенциалы коллекторов у одного транзистора увеличатся, а у второго уменьшатся. Напряжение на выходе ДУ, как разность потенциалов коллекторов $U_{ВЫХ} = \varphi_{К1} - \varphi_{К2}$, уже не бу-

дет равна нулю и на выходе появится сигнал. Понятно, что ДУ будет усиливать любые дифференциальные сигналы, лишь бы их амплитуды на базах были одинаковы по амплитуде и противоположны по знаку.

Можно показать, что коэффициент усиления дифференциальных сигналов определяется формулой:

$$K_U^{ДИФ} \approx \frac{-R_{1,2}}{r_{\Sigma}} = \frac{-R_{1,2}I_0}{2\varphi_T}. \quad (3.20)$$

Чтобы получить большой коэффициент усиления дифференциальных сигналов и «не греть» при этом кристалл микросхемы (т. е. работать при очень малом токе I_0 , порядка единиц микроампер), используют для имитации резисторов $R_{1,2}$ в цепи коллекторов транзисторов токовое зеркало. При этом удается получить очень большую величину дифференциального сопротивления $R_{1,2}$ и, в тоже время, малое сопротивление их по постоянному току. Опять же, это позволяет использовать источники питания с малой величиной ЭДС.

Таким образом, по сравнению с другими усилительными каскадами ДУ, являясь весьма технологичным для исполнения в ИМС, позволяет реализовать одновременное усиление очень малых дифференциальных сигналов и подавление достаточно больших синфазных помех. К сказанному следует добавить, что использование ДУ также позволяет:

а) усиливать сигналы в полосе частот от нуля герц, до сотен мегагерц;

б) усиливать сигналы с очень малыми нелинейными искажениями. Коэффициент нелинейных искажений ДУ можно оценить по формуле:

$$K_{НИ}^{ДУ} = \frac{U_{mBX}^4}{4\varphi_T^4}. \quad (3.21)$$

При амплитуде на входе порядка 5 милливольт усилитель на биполярном транзисторе дает коэффициент нелинейных искажений 4%, а ДУ – 0,04%, т. е. почти в 100 раз меньше.

ДУ обладает рядом и других достоинств, которые касаются его работы при изменении температуры, напряжений питания. Названные свойства и достоинства ДУ привели к тому, что он стал практически незаменимым элементом устройств, выполненных в виде ИМС.

3.2.6. Элемент флэш-памяти

Типовыми узлами современных ИМС становятся энергонезависимые («работают» при выключенных источниках питания) устройства для хранения информации. В настоящее время особый интерес представляет «память», которую непосредственно «встраивают» в большие микросхемы (микроконтроллеры), используемые для управления различными объектами. При этом пользователю предоставляется возможность программировать эту память, т. е. стирать, что в ней было записано ранее, и записывать в нее свои программы. Причем, перепрограммирование управляющих устройств становится возможным даже во время их работы. Такие операции осуществляются путем подачи на определенные выводы микросхемы электрических сигналов. Соответствующие устройства называют электрически стираемые и перепрограммируемые запоминающие устройства (EPROM).

Энергонезависимая и электрически перепрограммируемая память основывается на МДП (МОП) полевых транзисторах, которые могут находиться в одном из двух состояний «включено - выключено», соответствующих логическому «нулю» или логической «единице». Наиболее распространенными являются полевые транзисторы с «плавающим» затвором, у которых между обычным затвором (его называют управляющим) и подложкой в диэлектрике расположен изолированный со всех сторон второй затвор. Потенциал второго затвора изменяется в зависимости от заряда на нем; отсюда и название – «плавающий».

Электрически перепрограммируемая память требует относительно больших напряжений, поэтому в нее обычно встраивается генератор, который вырабатывает напряжения для осуществления

стирания и программирования. Это в определенной степени ограничивает объем памяти, размещаемой на кристалле, и требует некоторого времени для осуществления необходимых операций.

Значительным шагом вперед явилось использование таких конструкций транзисторов, при которых значительно сократилось время стирания и программирования. Это позволило создавать «память» большой емкости с высокими скоростями стирания и программирования. Синонимы «вспышка», «мгновение», «проблеск», «очень короткий промежуток времени» объединило слово FLASH (флэш). Энергонезависимую перепрограммируемую «память» с высокой скоростью стирания и программирования стали называть флэш-памятью.

Основу флэш-памяти составляет накопитель информации, занимающий почти всю площадь кристалла. Вдоль и поперек кристалла проложены образующие некую прямоугольную сетку пересекающиеся шины, соединенные с запоминающими элементами. Каждый запоминающий элемент, являясь элементарной ячейкой памяти, служит для записи и хранения одной двоичной единицы информации – бита, принимающего одно из двух возможных значений: либо логический «ноль» либо логическую единицу. Все ячейки образуют матрицу, размер которой, а, следовательно, и информационная емкость памяти, зависит от произведения количества шин по горизонтали и по вертикали. Поэтому объем памяти может выражаться в таких единицах как байт (1 байт = 8 бит), килобайт (1 килобайт $\sim 2^{10}$ байт), мегабайт (1 мегабайт $\sim 2^{20}$ байт). Код любой ячейки состоит из двух координат X и Y. По этой причине «выбор» интересующей ячейки осуществляется тем, что на одну вертикальную и одну горизонтальную шину подается сигнал «выбора» (обычно импульс положительного напряжения с определенной амплитудой).

На рис. 3.10 показана структура и эквивалентная схема одной ячейки флэш-памяти.

Элементарная ячейка памяти представляет МОП транзистор с индуцированным каналом, имеющий два затвора. Оба затвора выполнены из сильнолегированного поликремния и расположены в толще диэлектрика (слое диоксида). Один затвор, названный

управляющим, как и в обычном полевом транзисторе, имеет внешний вывод. Этот затвор соединен с выводом, предназначенным для осуществления программирования. Второй затвор предназначен для того, чтобы на его поверхности располагались заряды, с помощью которых идентифицируют наличие или отсутствие логической единицы. Он не имеет внешнего вывода, его потенциал меняется в зависимости от наличия на нем зарядов, поэтому этот затвор называют «плавающим». Слой диоксида, отделяющий сток от плавающего затвора, имеет очень малую толщину – порядка 0,01мкм. Для осуществления управления ячейкой служит второй транзистор, сток которого соединен с металлической шиной столбца Y, а затвор совмещен с шиной строки X. Весь кристалл содержит еще один внешний вывод CS (на рисунке не показан) с помощью которого можно «выбрать» кристалл, т.е. включить его в «работу». Включение осуществляется подачей на вход CS напряжения 5 В.

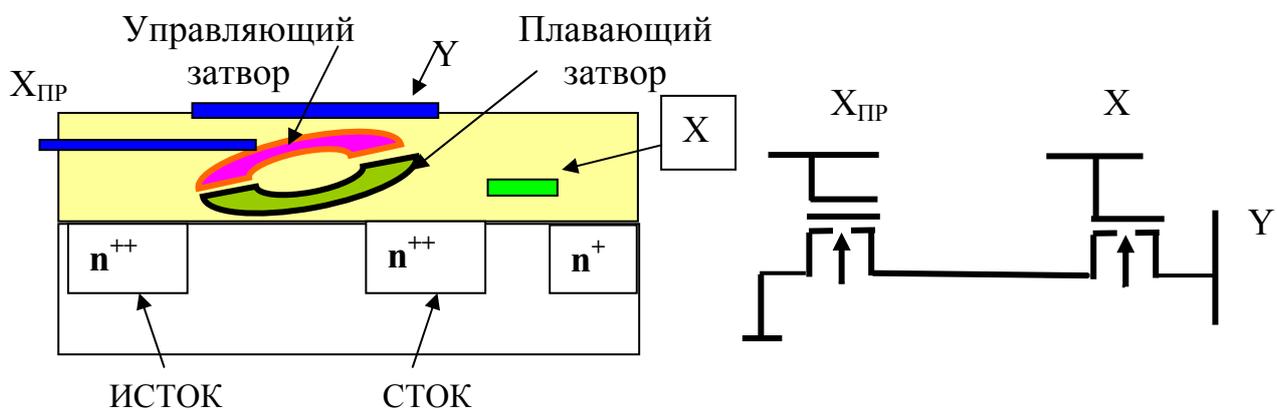


Рис. 3.10. Структура и эквивалентная схема одной ячейки флэш-памяти

Элементарная ячейка может функционировать в следующих режимах:

- а) хранения ранее записанной информации;
- б) записи в нее новой информации – «нового» бита – (программирования);
- в) стирания в ней ранее записанной информации;

г) считывания из нее ранее записанной информации– «старого» бита.

Без подачи напряжений на выводы, а также в случае, если $U_{CS} = 0$, ячейка флэш-памяти хранит ранее записанную в нее информацию (логическую единицу, если имеется заряд на плавающем затворе или логический нуль – если заряда нет).

Программирование и стирание информации в ячейке основывается на эффекте «туннелирования» «горячих» электронов из стока в диэлектрик или наоборот. «Горячие» электроны генерируются в сильном электрическом поле на границе перехода полупроводник диэлектрик, а их «туннелирование» обусловлено очень малой толщиной слоя диоксида кремния в пространстве между плавающим затвором и стоком.

Предположим, что мы хотим «записать» в ячейку логическую единицу, то есть нанести на поверхность плавающего затвора заряд электронов. Для выполнения этого необходимо предварительно «включить» кристалл, создав $U_{CS} = 5$, и «выбрать» нужную ячейку, подав на определенные шины X и Y соответствующие потенциалы. На шину X подают напряжение порядка 3 вольт и тем самым индуцируют канал во втором полевом транзисторе. К шине Y прикладывают нулевой потенциал, который по открытому каналу второго транзистора «попадает» на сток ячейки памяти. Потенциал стока становится равным нулю. На управляющий затвор ячейки подают положительный потенциал порядка 20 вольт (импульс напряжения с такой амплитудой).

Так как между управляющим затвором и стоком имеется разность потенциалов (напряжение затвор сток), то в диэлектрике очень малой толщины между плавающим затвором и стоком возникнет очень сильное электрическое поле, вектор напряженности которого направлен вертикально вниз. Это поле, как отмечалось, вызовет «впрыскивание» из стока в окисел «горячих» электронов, которые, затем, под действием того же поля, начнут «дрейфовать» к плавающему затвору, накапливаться на нем и фиксироваться на его поверхности. После окончания импульса программирующего напряжения заряд на затворе не исчезнет, поскольку двуокись кремния является хорошим изолятором (по истечении 10 лет на за-

творе сохранится еще примерно 80 % первоначальных зарядов). Таким образом, на плавающем затворе появился заряд электронов, что соответствует записи в ячейку логической единицы. Ток, протекающий в процессе программирования через диэлектрик весьма мал (порядка единиц пикоампер $\sim 10^{-12}$ А), поэтому время, необходимое для программирования одной ячейки памяти, весьма мало. Если не надо «вводить» заряд (т. е. требуется «записать» в элементарную ячейку логический ноль), то на шину Y подают потенциал, примерно равный напряжению на управляющем затворе. Этот потенциал через канал второго транзистора поступит на сток ячейки, в результате чего не будет напряжения между управляющим затвором и стоком, электрическое поле не возникнет и не появится «горячих» электронов.

После окончания процесса программирования отключают сигналы «выбора» кристалла и ячейки.

Аналогично проводится процесс стирания ранее записанной информации, однако, в этом случае схема подачи напряжений будет несколько иной. Предварительно «включают» кристалл, создавая $U_{CS} = 5$, и «выбирают» нужную ячейку. На шину X подают напряжение порядка 3 вольт и тем самым индуцируют канал во втором полевом транзисторе. К шине Y, в этом случае, прикладывают высокий потенциал, который по открытому каналу второго транзистора «попадает» на сток ячейки памяти. Потенциал стока становится равным порядка 20 вольт. На управляющий затвор ячейки подают нулевой потенциал (импульс напряжения с такой амплитудой).

Так как между управляющим затвором и стоком имеется разность потенциалов (напряжение затвор сток), то в диэлектрике очень малой толщины между плавающим затвором и стоком опять же возникнет очень сильное электрическое поле. Только теперь вектор напряженности будет направлен вертикально вверх. Электроны, расположенные на поверхности плавающего затвора, под действием электрического поля, начнут «дрейфовать», но теперь уже от плавающего затвора к стоку, совершая как бы обратный переход. В конце концов, на плавающем затворе не останется элек-

тронов. «Стирание» информации в ячейке памяти соответствует установлению в ней логического нуля.

Таким образом, программирование и стирание идет примерно с одинаковой скоростью. Эти процессы отличаются только знаком электрического поля в диэлектрике и направлением движения электронов. Но самое важное, что флэш-память оправдывает свое название позволяя «с быстротой молнии» выполнять операции программирования и стирания.

Для считывания информации по заданному адресу ячейки «включают» кристалл, создавая $U_{CS} = 5$, и «выбирают» нужную ячейку. На управляющий затвор ячейки подают потенциал порядка 5 вольт (импульс напряжения с такой амплитудой). На нужные шины X и Y подают напряжение порядка 3 вольт. При этом на «невывбранных» ячейках памяти канал у второго транзистора не наводится. В выбранной же ячейке канал может существовать (быть индуцированным) или отсутствовать в зависимости от хранимой там информации.

Если на плавающем затворе «хранится» заряд электронов, то напряжение затвор исток оказывается меньшим порогового и канал в первом транзисторе ячейки не индуцируется. Из-за этого по шине Y выбранного элемента памяти ток протекать не будет и ее потенциал останется прежним.

Если ли же на плавающем затворе отсутствует заряд электронов, то напряжение затвор исток оказывается большим порогового и в первом транзисторе ячейки индуцируется канал. Из-за этого по шине Y выбранного элемента памяти начнет протекать ток, а ее потенциал станет близким к нулю. Информацию о токе шины или ее потенциале подают на усилители «считывания» и, затем, необходимым образом обрабатывают.

Главными критериями, по которым оценивают флэш-память, являются скорость стирания и программирования и гарантированное число циклов стирания и записи. Стараясь увеличить количество ячеек памяти на кристалле, уменьшить время стирания-записи, а также добиваясь роста числа циклов, изготовители флэш-памяти постоянно предлагают новые решения и совершенствуют старые конструкции. Поэтому в новых изделиях описанные

ранее процессы в «деталях» могут отличаться. К примеру, возможно, что «горячие» электроны могут инжектироваться со стороны источника. Может быть, что электроны будут «уходить» с плавающего затвора «вверх» к управляющему затвору.

Также могут использоваться различные конструкции затворов и т. п. Однако, при этом, идея работы элементарной ячейки памяти будет почти прежней: как можно быстрее записать и считать информацию, используя различные комбинации электрических сигналов, подаваемых на внешние выводы.

3.3. Датчики

3.3.1. Общие сведения о датчиках

Эффекты внешнего мира весьма разнообразны в своих проявлениях. Свет, температура, давление, влажность, газовый состав, уровень радиации, перемещение и т. д. – таких физических явлений можно насчитать несколько десятков. Для их оценки и последующей обработки на ЭВМ нужны элементы, преобразующие внешнее воздействие в электрический сигнал. Каковы бы не были мощность компьютера и его возможности в плане обработки информации, он не способен «воспринять» информацию об изучаемом явлении. В этом ему должна «помочь» система датчиков, которая бы преобразовывала интересующие внешние воздействия в электрические сигналы.

Задачи, которые решают с помощью датчиков, весьма многообразны. Функциональное назначение датчиков перемещений – преобразовать механическое перемещение в пропорциональный электрический сигнал. Для определения величины давления используются датчики давлений. У термодатчиков изменения температуры приводят к изменению сопротивления материала.

Назначение *тензодатчиков* – преобразовать деформации, вызванные изменением внешней нагрузки, в вариацию его внутреннего сопротивления. Как правило, сам по себе датчик является пассивным элементом электрической цепи, у которого под действием внешнего возмущения изменяется лишь какой то его показатель, но он сам не генерирует электрический сигнал. Однако, если такой датчик подключить к элек-

трическому источнику, то соответствующие проявления внешней информации будут уже отражаться в выходном сигнале как ток или напряжение. Не потребляя непосредственно энергии от источника внешних воздействий, датчик «генерирует» электрический сигнал за счет работы источника питания. Понятно, что датчики такого типа (преобразователи сигнала) не могут функционировать при отсутствии источников питания.

Существуют и датчики преобразователи энергии. Как увидим в дальнейшем, солнечный элемент преобразовывает энергию света в электрическую энергию. В пьезоэлектрических датчиках давления при воздействии возникает электрический заряд.

Длительное время в технике использовались датчики, основанные на механическом и магнитоэлектрическом принципе. Они имели многообразные формы и сложные конструкции и для их эксплуатации нужны были обширные специальные знания. Развитие полупроводниковой техники и технологии привели к тому, что в настоящее время наиболее широко применяются полупроводниковые датчики. По сравнению с прежними, эти датчики миниатюрны, просты в обращении и на их основе легко создавать интегральные устройства.

Микроэлектроника внесла значительный вклад в уменьшение размеров чувствительных элементов датчиков, снижение их веса, в сокращение потребления энергии, а также в расширение их возможностей. Благодаря миниатюризации датчиков стало возможным их массовое производство, что, в свою очередь, сыграло большую роль в снижении их стоимости. Важным моментом является и то, что миниатюризация датчиков способствует повышению их надежности.

Миниатюризация датчиков позволила повысить пространственное разрешение, что в ряде случаев имеет очень важное значение. Например, для магнитных датчиков, в качестве которых раньше использовались катушки индуктивности, миниатюризация которых затруднена, использование полупроводниковых элементов для регистрации магнитного поля позволило использовать элементы с размерами $2\text{мкм} \times 2\text{мкм}$. А это, в свою очередь, позволило внедрить в электрические двигатели электронное управление. В частности, полупроводниковые датчики сыграли важную роль в решении задач стабилизации числа оборотов двигателя, регулировки

скорости вращения вала, устранения влияния изменений нагрузки на двигатель, позиционирования вала при остановке.

3.3.2. Полупроводниковые датчики температуры и тензодатчики

Простейшим датчиком температуры является полупроводниковый брусок с омическими контактами на концах, именуемый терморезистором. Как известно, концентрация собственных носителей в полупроводнике существенным образом зависит от температуры. Соответственно меняется электропроводность датчика.

Обратный ток полупроводникового диода также сильно зависит от температуры. Используя эту зависимость обратного тока и сопротивления диода от температуры, их можно применять в качестве датчиков температуры. Преимуществом диодных датчиков является высокая чувствительность и малый потребляемый ток. Кремниевые диоды можно применять для измерения температур до 170 °С.

Обратный ток коллекторного перехода транзистора также зависит от температуры. Поэтому и транзистор можно использовать в качестве датчика. При этом обычно используют схему с общим эмиттером и «оторванной» базой. При деформации пластины полупроводника происходит изменение расстояния между атомами. Вследствие этого меняется структура энергетических зон и концентрация носителей в полупроводнике. *Простейшим тензодатчиком* является полупроводниковая пластинка с омическими контактами на концах, один конец которой закреплен. При деформации пластинки меняется ее сопротивление. Для уменьшения влияния температуры на датчик используют примесные полупроводники. *Тензорезисторы* широко используются при изменении деформаций, давления, силы, смещения, ускорения.

Большая *тензочувствительность*, по сравнению с *p-n* диодами, обнаружена у диодов Шоттки. Диод Шоттки широко используют в качестве тензодатчика.

3.3.3. Полупроводниковые датчики магнитного поля

Рассмотрим пластинку полупроводника *p* типа, через которую протекает ток, созданный идеальным источником тока. Пластинка помеще-

на в магнитное поле таким образом, что поверхность пластинки перпендикулярна магнитному полю, а вектор индукции имеет направление, как показано на рис. 3.11. Из физики известно, что на заряд, движущийся в магнитном поле со скоростью V , действует магнитная сила Лоренца, величина которой определяется формулой:

$$\vec{F} = q[\vec{V} \times \vec{B}], \quad (3.22)$$

где B – индукция магнитного поля.

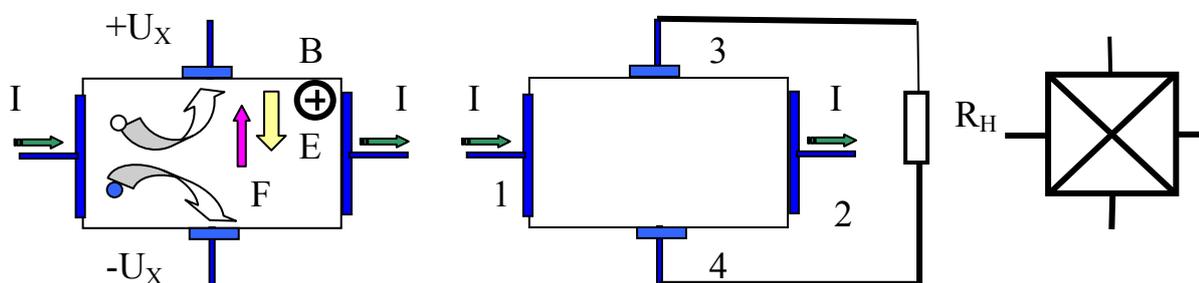


Рис. 3.11. Элемент Холла и его условное обозначение на схемах

Сила Лоренца отклоняет дырки к верхней грани полупроводника, вследствие чего их концентрация там увеличивается, а у нижней грани уменьшается. Аналогично происходит отклонение неосновных носителей заряда – электронов, только к нижней грани. В результате пространственного разделения зарядов возникает электрическое поле E , направленное от верхней грани к нижней. Это поле препятствует дальнейшему разделению зарядов, и, как только создаваемая им электрическая сила станет равной силе Лоренца $qE = qVB$, дальнейшее разделение зарядов прекратится. Поток дырок при таком равновесии движется через пластину не отклоняясь. Соответственно, между верхней и нижней гранью полупроводникового образца возникнет ЭДС, величина которой равна:

$$U_x = Ed = VBd \quad (3.23)$$

Явление возникновения поперечной ЭДС в полупроводниковой пластинке, помещенной в магнитное поле, получило название эффекта Холла, а U_x – напряжения Холла. Элемент, использующий это явление, получил название датчика Холла. Его условное графическое изображение на схемах показано на рис. 3.11.

Для создания тока через полупроводниковую пластинку датчик Холла подключают к источнику тока. Если трудно обеспечить этот режим, то датчик подключают (к выводам 1, 2) к источнику ЭДС. Выводы 3,4 часто используют как «выход» датчика, с которого снимается U_x . К выходу подключают нагрузку.

Работу датчика Холла можно характеризовать коэффициентом использования датчика, который определяется отношением мощности, выделенной во внешней нагрузке, к мощности, потребляемой входной цепью от источника. Можно показать, что этот коэффициент пропорционален квадрату подвижности основных носителей заряда. По этой причине для изготовления датчиков Холла используют полупроводники с высокой подвижностью, например, арсенид галлия. GaAs датчики Холла имеют высокую линейность характеристики, широкий температурный диапазон работы, они приспособлены для поверхностного монтажа. Размеры их могут быть порядка 1 мкм.

Датчики являются приборами, которые служат для обработки информации и поэтому неразрывно связаны со схемами обработки сигналов. Поэтому имеются также интегральные устройства, объединяющие в одном кристалле датчик и схему обработки сигнала. В конструкции ИМС Холла на одном кремниевом кристалле объединены элементы Холла, чувствительные к магнитному полю, а также усилительная и переключающая схемы. В таких устройствах низкая чувствительность к магнитному полю, обусловленная малой подвижностью электронов в кремнии, компенсируется включением усилительной схемы. Выход таких схем может быть логическим.

В настоящее время холловские ИС находят широкое применение в качестве бесконтактных измерителей тока переключателей, обеспечивающих возможность автоматической регулировки электрических двигателей. Используют их также для измерения перемещений и в качестве бесконтактных потенциометров.

В качестве датчиков магнитного потока используют многослойные структуры магнитных и немагнитных материалов (*GMR - магниторезисторы*), которые чувствительны только к направлению магнитного поля. Их применяют как датчики вращений, бесконтактных и поворотных переключателей.

Контрольные вопросы

1. Какие основные проблемы имеют место в современном исполнении электронных изделий?
2. Какое изделие называют интегральной микросхемой (ИМС)?
3. Перечислите наиболее очевидные существенные преимущества полупроводниковых ИМС по сравнению с обычными устройствами, в которых используются дискретные элементы.
4. Охарактеризуйте кратко основные принципы изготовления интегральных микросхем.
5. Перечислите основные характерные особенности выполнения структурных элементов современных ИМС.
6. Что является основным методом изоляции элементов современных биполярных микросхем?
7. Что представляет собой и для чего предназначена комплементарная МОП-пара (КМОП)?
8. Охарактеризуйте понятие микроэлектроники– «трехмерная» интеграция.
9. Каким образом в полупроводниковых ИМС реализуют диоды, резисторы, емкости?
10. В чем заключаются особенности схем установки рабочей точки во входной цепи усилительного каскада интегрального исполнения на биполярном транзисторе в СОЭ?
11. Охарактеризуйте понятие – «эталон» интегральной схемотехники.
12. Нарисуйте и поясните схемы простейших источников стабильного тока.
13. Нарисуйте и поясните схемы отражателей тока на биполярных транзисторах.
14. Охарактеризуйте особенности применения ГСТ в современных ИМС.

15. Охарактеризуйте особенности использования повторителей напряжения в ИМС.
16. Нарисуйте и поясните схему простейшего эмиттерного повторителя на биполярном транзисторе.
17. Нарисуйте и поясните схемы и назначение составного транзистора Дарлингтона.
18. Нарисуйте и поясните схемы и назначение составного транзистора Шиклаи.
19. Что называют комплементарным (дополняющим) транзистором Дарлингтона? В чем его особенность?
20. Какую схему называют дифференциальным (т.е. разностным) усилителем? Для чего он предназначен?
21. Нарисуйте и поясните схему простейшего дифференциального усилителя на биполярных транзисторах.
22. Назовите основные характерные положительные характеристики дифференциального усилителя.
23. На каких элементах основывается энергонезависимая и электрически перепрограммируемая память?
24. Охарактеризуйте понятие – флэш-память. Дайте краткую характеристику.
25. Что представляет элементарная ячейка памяти? В каких режимах она может функционировать?
26. На каком физическом эффекте основывается программирование и стирание информации в ячейке флэш-памяти?
27. Назовите главные критерии, по которым оценивают флэш-память.
28. Назовите некоторые задачи, которые решают с помощью датчиков.
29. Какие типы датчиков (по функциональному назначению) Вы знаете?
30. Охарактеризуйте простейший датчик температуры.
31. В чем заключается основное преимущество диодных термодатчиков?
32. Охарактеризуйте понятия – тензодатчик, тензорезистор.
33. Что представляет собой, на каком явлении основан и для чего служит датчик Холла?
34. Какие структуры используют в качестве датчиков магнитного потока?

4. Оптоэлектронные полупроводниковые приборы и компоненты средств отображения информации

Мы приступаем к изучению нового класса явлений, широко используемых в настоящее время в электронной технике, – явлений излучения, поглощения, усиления оптического излучения (света) веществом – и, соответственно, широко распространенных изделий – оптоэлектронных полупроводниковых приборов.

Оптоэлектроника – направление в науке и технике, соединившее в себе возможности оптики и электроники. При этом, являясь синтезом двух направлений, оптоэлектроника превосходит каждое из них по возможностям.

В 1922 году девятнадцатилетний нижегородский радиолюбитель Олег Владимирович Лосев впервые публично сообщил о своем открытии: превращении электрической энергии в световую. Во многих странах это явление называют свечением Лосева («Losev Light»). Спустя много лет он отказался от эвакуации из блокадного Ленинграда, чтобы завершить исследование фоточувствительных сплавов кремния. К сожалению, этому человеку, одному из пионеров полупроводниковой электроники, выполнить полностью задуманное не удалось: он умер зимой 1942 года в блокадном Ленинграде.

После блестящих работ О.В. Лосева прошло уже много десятилетий и фундаментальные исследования в области физики твердого тела привели оптоэлектронику к современным излучающим приборам и фотоприемникам.

Мы изучим физические принципы, лежащие в основе работы оптоэлектронных приборов, а также рассмотрим основные свойства и параметры оптического излучения. Основное внимание мы уделим приборам, предназначенным для создания оптического излучения – излучающим диодам и полупроводниковым лазерам.

Далее мы затронем важные для оптоэлектронных устройств вопросы, посвященные поглощению света веществом и внутреннему фотоэффекту. Мы изучим наиболее распространенные типы фотоприемников. Наконец, мы кратко затронем имеющие большое значение вопросы, касающиеся создания оптоэлектронных устройств, работающих в системах передачи и отображения информации.

4.1. Общие сведения об оптическом излучении

4.1.1. Понятия, энергетические и световые параметры оптического излучения

Из курса физики вы получили представление о двойкой природе оптического излучения. С одной стороны, электромагнитное излучение оптического диапазона это волновой процесс с частотой колебаний ν или длиной волны λ , а с другой – поток элементарных частиц, называемых фотонами, с энергией W_{opt} (обычно измеряемой в эВ - электронвольтах):

$$W_{opt} = \hbar \nu = c \cdot \hbar / \lambda \Rightarrow W_{opt} = 4.1 \cdot 10^{-15} \nu = 1236 / \lambda, \quad (4.1)$$

где h – постоянная Планка;

c – скорость света.

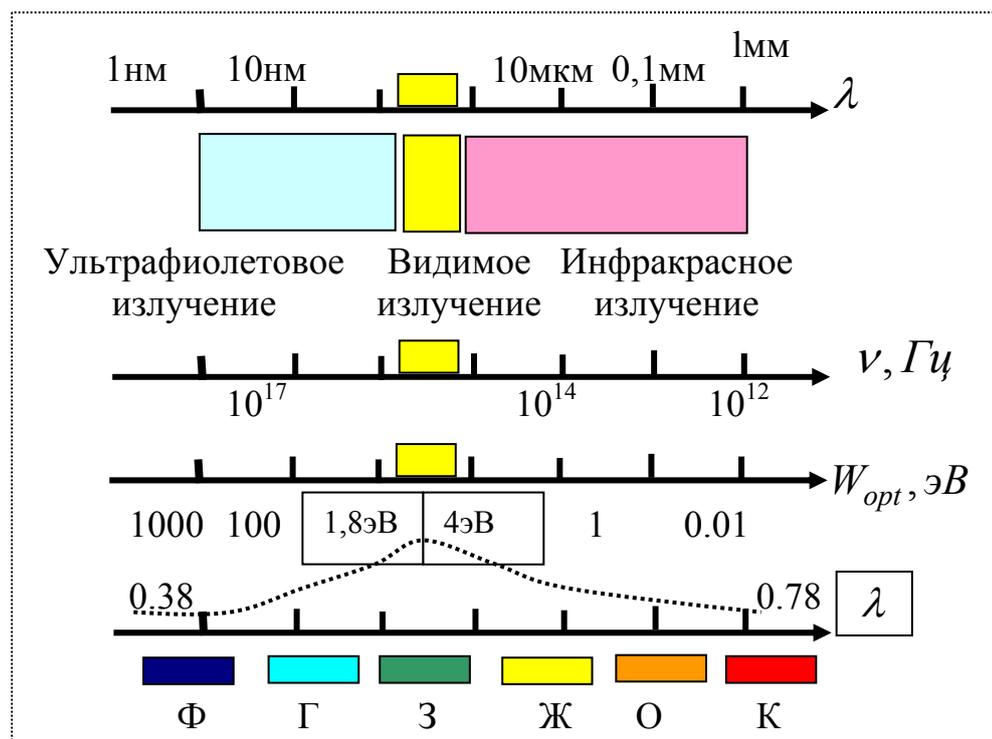


Рис. 4.1. Представление оптического диапазона в виде ряда шкал электромагнитных волн

Принято считать, что оптическое излучение это электромагнитные волны с длиной волны λ от 10 нм до 1 мм. По физическим свойствам оптический диапазон волн неоднороден. Поэтому оптический диапазон делят на поддиапазоны, в которых физические свойства в определенной степени одинаковы (рис. 4.1). Различают ультрафиолетовое излучение ($\lambda = 0,01 - 0,38$ мкм), видимое излучение ($\lambda = 0,38 - 0,78$ мкм), инфракрасное излучение ($\lambda = 0,78$ мкм – 1мм). В свою очередь, ультрафиолетовый и инфракрасный оптические поддиапазоны делят на три субподдиапазона (УФ - А, УФ - В, УФ - С, ИК - А, ИК - В, ИК - С).

Оптическое излучение характеризуют фотометрическими параметрами. При этом различают фотометрические параметры энергетические и световые.

Энергетические параметры характеризуют излучение безотносительно к его действию на какой-либо приемник излучения и отражают, в конечном счете, переносимую излучением энергию.

Энергию электромагнитных волн, приходящуюся на данную площадку облученной поверхности в единицу времени, называют *поток излучения*, отнесенным к этой площадке. Поток излучения измеряют в *ваттах*.

Падая на поверхность тела, поток излучения создает *энергетическую освещенность* поверхности, принимающей излучение. С помощью световых параметров оценивают излучение, в случае, если приемником излучения служит средний нормальный человеческий глаз. Как известно, световое ощущение на сетчатке глаза создает только сравнительно узкий диапазон электромагнитных волн ($\lambda = 0,01 - 0,38$ мкм). Кроме того, чувствительность глаза к свету с разными длинами волн неодинакова (она, как следует из штриховой кривой рис. 4.1, имеет максимум при длине волны $\lambda = 0,555$ мкм и быстро снижается при удалении от этого максимума).

Во многих случаях пользуются понятием *точечного источника света*. Это означает, что действие источника на какой-либо приемник света будет зависеть только от расстояния приемника до центра источника, и не будет зависеть от направления радиуса, проведенного к приемнику от центра источника. Полный световой поток характеризует излучение, которое распространяется от источника по всем направлениям.

Выделенная из всего потока излучения мощность, создающая световые ощущения в среднем по светочувствительности человеческого глазу, называется *световым потоком*. Световой поток измеряется в люменах.

Падая на поверхность тела, световой поток создает *освещенность* этой поверхности. При нормальном падении лучей от точечного источника света освещенность точки поверхности равна отношению светового потока, падающего на элемент поверхности, к площади этого элемента. Единица измерения освещенности – *люкс*.

На практике важно получать световой поток, распространяющийся в определенном направлении. В этой связи необходима величина, которая бы характеризовала световой поток, распространяющийся в данном направлении. Обычно полагают, что в вершине воображаемого конуса находится точечный источник света и считают, что часть его светового потока распространяется только в некотором телесном угле. Напомним, что под телесным углом, который выражается встерадианах (пространственных радианах), подразумевается некоторая область пространства, ограниченная конической поверхностью. За 1 ср принимают такой телесный угол, который опирается на площадь сферы в один ее квадратный радиус при условии, что вершина угла совпадает с центром сферы.

Силой света источника видимого излучения в данном направлении называется отношение созданного источником светового потока к достаточно малому ориентированному в заданном направлении телесному углу, в котором этот поток распространяется. Сила света измеряется в канделах (от лат. *candela* – свеча).

Точечный источник света является идеализацией. Все реальные светящиеся тела являются протяженными (мы различаем их форму и величину). Протяженные источники характеризуются величиной, называемой яркостью светящейся поверхности в определенном направлении (подразумевается в направлении к глазу). При этом важно отметить, что когда мы рассматриваем какой-либо источник света, для нас важное значение имеет не сама величина излучающей поверхности, а ее видимая величина, т. е. проекция излучающей поверхности на плоскость, перпендикулярную к направлению наблюдения.

Яркость источника в точке поверхности равна световому потоку, испускаемому с единицы видимой величины поверхности источника

внутри единичного телесного угла. Если направление, под которым наблюдается светящаяся поверхность, ей перпендикулярно, то яркость определяется отношением силы света, излучаемого с достаточно малого элемента этой поверхности в данном направлении, к площади этого элемента. *Единицей яркости* в СИ является кандела на квадратный метр. Яркость, с одной стороны, является объективной характеристикой светящейся поверхности, а с другой – субъективной характеристикой уровня зрительного ощущения при наблюдении этой поверхности.

Вы столкнетесь с тем, что в паспортных данных на многие оптоэлектронные приборы приводятся частично и энергетические и световые параметры. Вам, в зависимости от предназначения прибора, потребуется умение правильно выбирать необходимые параметры для оценки прибора, а также переводить имеющиеся световые параметры в энергетические и наоборот. Из физики вам также известно, что оптическое излучение, являясь электромагнитной волной, у которой векторы напряженности электрического и магнитного полей перпендикулярны друг к другу, может быть поляризованным и неполяризованным.

Естественный свет является неполяризованным. Он представляет собой оптическое излучение с быстро и беспорядочно изменяющимися направлениями напряженности электромагнитного поля, причем все направления векторов напряженности, перпендикулярные оптическим лучам, равноправны. Раскаленные тела, светящиеся газы испускают свет, близкий к естественному, но все же обычно в небольшой степени поляризованный. Близок к естественному свету прямой солнечный свет.

Полностью поляризованное оптическое излучение в общем случае, это – электромагнитная волна, у которой две взаимно перпендикулярные компоненты E_x и E_y электрического вектора напряженности E совершают колебания с постоянной во времени разностью фаз. Состояние поляризации оптического излучения изображается с помощью эллипса поляризации. Различают при этом линейную, эллиптическую и круговую поляризации. Известно также, что оптическое излучение может взаимодействовать друг с другом, к примеру, интерферировать. Для этого необходимо обеспечить когерентность или, другими словами, согласованность двух систем волн. *Когерентное оптическое излучение* это, в первом приближении, электромагнитные волны, имеющие одну часто-

ту (период) колебаний и обладающие неизменной разностью фаз в пространстве и на протяжении времени, достаточном для наблюдения.

Абсолютно когерентный источник излучения – это монохроматический точечный источник излучения. Однако, в действительности не существует ни монохроматических, ни точечных источников. Отклонение реального источника оптического излучения от монохроматичности характеризуется степенью временной когерентности; а отклонение размеров источника от точечных – степенью пространственной когерентности.

Мы столкнемся с тем, что атомы вещества излучают электромагнитные волны короткими сериями – волновыми пакетами (импульсами). Длительность излучаемого волнового пакета и спектр излучения связаны между собой. В идеале, монохроматическое излучение можно получить в таком случае для волны бесконечной длины. Однако, если волновой пакет, как это делают в интерферометрах, разделить на два волновых пакета и обеспечить, чтобы время запаздывания одного пакета относительно другого было много меньше длительности каждого из волновых пакетов, то на выходе интерферометра можно получить четкую, интерференционную картину. В этом смысле можно говорить, что между рассматриваемыми колебаниями существует временная когерентность.

Рассмотрим теперь достаточно протяженный – не точечный – источник оптического излучения, состоящий из большого числа точечных источников, например, из атомов излучающего вещества. Если размеры источника малы, то все атомы дадут одну и ту же интерференционную картину. Общая картина имеет такой же вид, как и интерференционная картина для отдельного атома, но гораздо ярче. В этом смысле можно говорить о *пространственной когерентности*. При увеличении диаметра источника излучения расстояния между интерференционными полосами, создаваемыми различными атомами, из-за разности хода волн (т. е. из-за рассогласования фаз колебаний в пространстве), неодинаковы и картина более «размыта», нечеткая. В таком случае говорят, что имеет место *частичная пространственная когерентность* оптического излучения. Надо иметь в виду, что следствием высокой пространственной когерентности является возможность получения высокой *направленности излучения*. Некогерентное излучение имеет низкую направленность.

4.1.2. Общие сведения о механизме генерации оптического излучения в полупроводниках

Важно сразу подчеркнуть тот факт, что далее речь пойдет о так называемой электролюминесценции, т. е. о создании оптического излучения нетеплового происхождения. В отличие от свечения накаливаемых тел для люминесценции не требуется нагревания тела, а необходимо подведение возбуждающей энергии от электрических источников. В оптоэлектронных полупроводниковых приборах используется инжекционная электролюминесценция кристаллических примесных полупроводников. Излучающие приборы преобразуют электрическую энергию в энергию оптического излучения с определенной длиной волны (точнее – в узком диапазоне длин волн).

Инжекционная электролюминесценция, как ясно из самого названия, включает в себя два основных этапа. На первом из них под воздействием возбуждающей энергии происходит генерация носителей заряда. Самый распространенный способ осуществления этого состоит в приложении к электронно-дырочной структуре (к $p - n$ -переходу) прямого напряжения; в этом случае, как известно, основные носители инжектируются в зону проводимости. На втором этапе генерированные носители заряда рекомбинируют: «свободный» электрон зоны проводимости, спустя некоторое время, называемое временем жизни, спонтанно вступает во взаимодействие с дыркой в валентной зоне и из зоны проводимости переходит в валентную зону. Этот процесс сопровождается исчезновением пары электрон зоны проводимости – дырка валентной зоны и называется рекомбинационным излучением. Выделяющаяся при излучательной рекомбинации энергия превращается в оптическое излучение и частично в теплоту.

Теперь настало время отметить, что в приведенных на рис. 1.3, 1.6 и т. д. энергетических диаграммах полупроводников, не обращалось внимание на то обстоятельство, что существующие полупроводники могут различаться с точки зрения того, как у них потолок валентной зоны соотносится с дном зоны проводимости.

Как свидетельствуют энергетические диаграммы, показанные на рис. 4.2, существует два рода полупроводников.

Во-первых, могут быть полупроводники, у которых потолок валентной зоны, в котором находятся дырки, совпадает с дном зоны проводимости, где расположены свободные электроны. Это, с точки зрения квантовой механики, соответствует нулевому значению волнового числа ($\hbar k$ или k) в волновой функции электрона и, соответственно, импульса (минимум зоны проводимости и максимум валентной зоны находятся в точке нулевого импульса). В них, при межзонных переходах электронов, значение импульса электрона не меняется. Для того, чтобы электрон и дырка могли прорекомбинировать в таких полупроводниках с излучением фотона, должно выполняться условие сохранения энергии: в процессе перехода разность энергий электрона в зоне проводимости и в валентной зоне должна равняться энергии фотона $W_1 = \hbar \nu$. Такие полупроводники называют *прямозонными*.

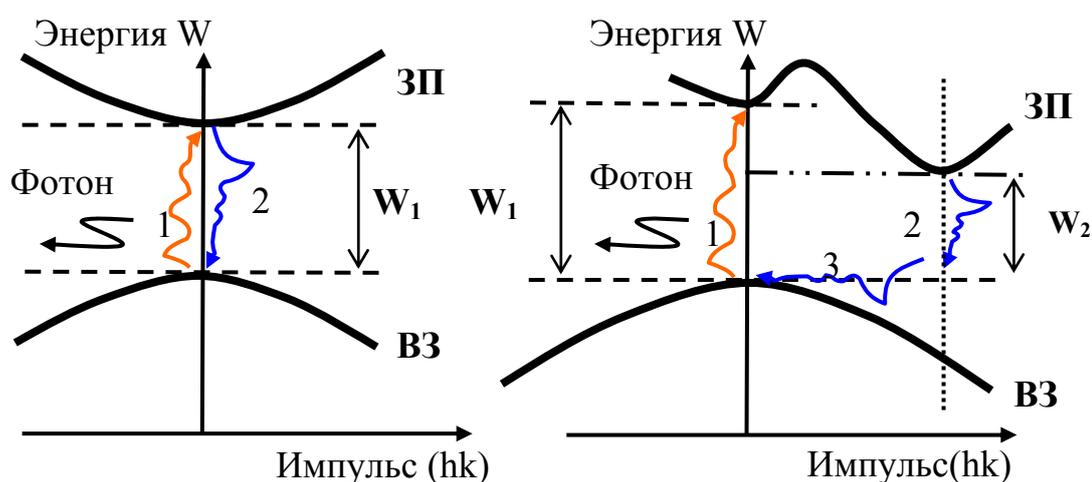


Рис. 4.2. Энергетические диаграммы прямозонных и непрямозонных полупроводников

Во-вторых, существуют непрямозонные полупроводники, в которых минимум зоны проводимости и максимум валентной зоны расположены при разных значениях импульса. В таких полупроводниках, для того, чтобы электрон и дырка могли прорекомбинировать с излучением фотона, должно выполняться условие сохранения и энергии и импульса. В этом случае имеет место двухступенчатый непрямой переход. Сначала электрон из зоны проводимости переходит на виртуальный энергетический уровень и при сохранении импульса происходит излучение оптического кванта – фотона. Затем излучается тепловой квант энергии – фо-

нон, уносящий избыточный импульс, и электрон перемещается к потолку валентной зоны. В процессе перехода разность энергий электрона в минимуме зоны проводимости и в максимуме валентной зоны и энергии импульса должна равняться энергии фотона $W_2 \pm W_1 = \hbar \nu$.

Время существования рекомбинационного излучения в прямозонных полупроводниках составляет несколько наносекунд. Двухступенчатое (совершаемое в две стадии) рекомбинационное излучение непрямозонных полупроводников существенно длительнее – от нескольких десятков, до нескольких сотен микросекунд.

Когда электроны инжектируют в p область электронно-дырочной структуры, создавая там избыточные электроны плотностью Δn , то вероятность того, что избыточный электрон зоны проводимости прорекомбинирует с одной из многих $(p + \Delta p)$ дырок валентной зоны, пропорциональна произведению $\Delta n(p + \Delta p)$. В случае односторонней инжекции ($p \gg \Delta p$), скорость излучательной рекомбинации r определяется как:

$$r = R(\Delta n)p. \quad (4.2)$$

Здесь коэффициент пропорциональности R называется *коэффициентом излучательной рекомбинации*.

Таблица 4.1

Полупроводники различного типа зонной структуры и их параметры

| Материал | Ширина запрещенной зоны W | Тип зонной структуры | Коэффициент рекомбинации R |
|----------|-----------------------------|----------------------|------------------------------|
| Si | 1.1 | Непрямозонный | $1,7 \cdot 10^{-15}$ |
| Ge | 0.67 | Непрямозонный | $5,25 \cdot 10^{-14}$ |
| GaP | 2.25 | Непрямозонный | $5,37 \cdot 10^{-14}$ |
| GaAs | 1.4 | Прямозонный | $7,21 \cdot 10^{-10}$ |
| InAs | 0.31 | Прямозонный | $8,5 \cdot 10^{-11}$ |
| InSb | 0.18 | Прямозонный | $4,6 \cdot 10^{-11}$ |

При рассмотрении процесса излучательной рекомбинации необходимо обратить внимание на то, что вероятность рекомбинации зависит от концентрации, как избыточных электронов, так и дырок. Так, например, при наличии в полупроводнике только электронов (или только дырок) вероятность рекомбинации крайне мала. Это позволяет накапливать заряды в приборах с зарядовой связью. С другой стороны, например, в полупроводниковых лазерах, вероятность излучательной рекомбинации увеличивают за счет повышения концентрации как электронов, так и дырок в узкой области структуры.

В таблице 4.1 представлены отдельные прямозонные и непрямозонные полупроводники и их коэффициенты излучательной рекомбинации. Таблица свидетельствует, что возможность рекомбинации электрона и дырки в двухступенчатом процессе существенно ниже, чем в прямом рекомбинационном процессе. Эта разница четко отражается большими значениями R (более чем на три порядка) для прямозонных, нежели для непрямозонных полупроводников.

Каждая рекомбинация носителя заряда при переходе из зоны проводимости в валентную зону полупроводника сопровождается излучением фотона, длина волны которого приближенно определяется соотношением:

$$\lambda = \frac{1.24}{W}, \quad (4.3)$$

где λ – длина волны в микрометрах; W – ширина запрещенной зоны в электрон-вольтах.

Как следует из (4.3), для работы в диапазоне видимого человеческого глаза излучения ($\lambda = 0,38 - 0,78$ мкм) необходимы полупроводники с шириной запрещенной зоны 1,5 – 3,0 эВ. Это требование сразу исключает использование германия и кремния, технология которых хорошо отработана, и обуславливает переход к применению иных полупроводниковых материалов, к примеру, бинарным соединениям типа $A^{III}B^V$.

Следует также иметь в виду, что оптическое излучение, получаемое в результате рекомбинации носителей заряда при спонтанном переходе из зоны проводимости в валентную зону полупроводника, представляет собой суперпозицию (сумму) электромагнитных волн, излучае-

мых большим числом переходящих электронов, которые до этого «находились» на разрешенных уровнях зоны проводимости. Если каждый электрон «переходит», и, соответственно, излучает независимо от остальных, а также направления поляризации различны для всех излучающих атомов, то имеет место некогерентное оптическое излучение. Оно является хаотическим многочастотным, не имеет строгой направленности и характеризуется интенсивностью (мощностью).

Если предположить, что инжектированные в зону проводимости электроны «расселены» по уровням так, что наибольший вклад в создание оптического излучения вносят носители, размещенные на уровнях в полосе (диапазоне энергий) шириной $3kT$, то полная ширина спектра (на уровне 0,5) некогерентного оптического излучения определяется формулой:

$$\Delta\lambda = \frac{3kT}{\hbar c} \lambda^2, \quad (4.4)$$

где \hbar - постоянная Планка; k - постоянная Больцмана; T – температура.

Большинство бинарных соединений $A^{\text{III}}B^{\text{V}}$ с прямыми переходами типа GaAs, InP, GaSb и т. п. имеют значение ширины запрещенной зоны много меньше 1,5 эВ и не пригодны для получения видимого излучения ($1.8 \leq W_{33} \leq 3.2 \text{ эВ}$). Один из способов получения материалов с прямыми переходами, при относительно большой ширине запрещенной зоны, заключается в образовании трехкомпонентных твердых растворов (тройных сплавов) прямозонных бинарных соединений с малой шириной запрещенной зоны и непрямозонных бинарных соединений с большой шириной запрещенной зоны. При таком способе энергия прямого перехода между зонами для материала с меньшей шириной запрещенной зоны монотонно увеличивается в растворе по мере возрастания содержания второго соединения.

К примеру, часто для увеличения эффективности излучательных процессов в непрямозонных полупроводниках используют трехкомпонентные твердые растворы фосфида и арсенида галлия $GaAs_{1-x}P_x$. Величина x – молярная доля (в относительных единицах). Рассмотрим,

как меняется энергетическая диаграмма $GaAs_{1-x}P_x$ при изменении доли фосфора в решетке арсенида галлия (рис. 4.3)

При $x = 0$ энергетическая диаграмма соответствует прямозонному чистому арсениду галлия с шириной запрещенной зоны 1,42 эВ. При $x = 1$ диаграмма соответствует непрямозонному фосфиду галлия с шириной запрещенной зоны 2,32 эВ. При возрастании x от 0 до 0,45 ширина запрещенной зоны трехкомпонентного раствора увеличивается от 1,42 до 1,98. При этом, как следует из рисунка, имеет место энергетическая диаграмма, соответствующая прямозонному полупроводнику. Дальнейшее увеличение содержания фосфора приводит к диаграмме, соответствующей непрямозонному полупроводнику, что, соответственно, значительно уменьшает вероятность излучательной рекомбинации.

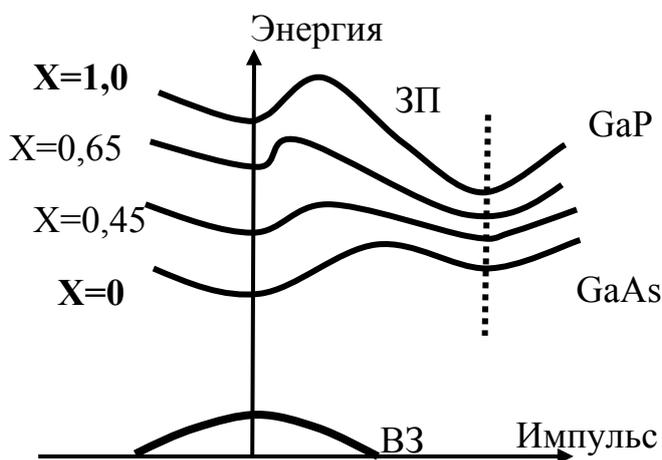


Рис. 4.3. Энергетические диаграммы твердых растворов

Итак, как следует из рис. 4.3, трехкомпонентный твердый раствор $GaAs_{1-x}P_x$ позволяет, меняя молярную долю, регулировать ширину запрещенной зоны и сохранять при этом преимущественно прямые переходы из зоны проводимости в валентную зону (до энергий, соответствующих красному цвету).

Для увеличения эффективности излучательных процессов в непрямозонных полупроводниках используют внедрение различных донорных и акцепторных примесей, которые создают рекомбинационные центры – ловушки. Излучательные переходы между донорными и акцепторными уровнями ловушек позволяет получить генерацию света на

различных длинах волн. При внедрении цинка, кадмия и кислорода реализуется красное излучение, а кадмия, серы и азота – зеленое.

4.1.3. Общие сведения о поглощении полупроводником оптического излучения

Наряду с оптическими излучателями существуют полупроводниковые фотоэлектрические приемники оптического излучения (коротко, фотоприемники), которые преобразуют энергию оптического излучения в электрическую энергию. Под действием оптического излучения происходит изменение электрофизических параметров полупроводника, обусловленное увеличением энергии электронов вещества, переходом их на более высокие уровни энергии, образованием дополнительных свободных носителей заряда.

Процесс образования дополнительных носителей заряда (фотоносителей) внутри полупроводника под действием оптического излучения называется внутренним фотоэлектрическим эффектом (внутренним фотоэффектом).

Принято различать две разновидности фотоэффекта.

1. *Эффект фотопроводимости* состоит только в создании фотоносителей. Его результат – изменение концентрации носителей в полупроводнике и, как следствие, увеличение электрической проводимости полупроводника. Этот эффект используют в фоторезисторах.

2. *Фотогальванический эффект* возникает в структурах с внутренним электрическим полем ($p - n$ – переходом, с переходом металл – полупроводник), когда электрическое поле перехода, ко всему ранее сказанному, еще и «разделяет» в пространстве дырки и электроны, возникшие под воздействием оптического излучения. Этот эффект используют в фотодиодах, фототранзисторах и ряде других приборов.

Оптическое излучение, направленное на полупроводник, как известно из физики, при взаимодействии с кристаллом полупроводника частично отражается от границы раздела воздух – полупроводник и частично поглощается.

Чтобы оценивать доли отраженной, поглощенной и прошедшей через кристалл энергии излучения пользуются соответствующими коэффициентами.

Коэффициент отражения R – отношение отраженной мощности от поверхности кристалла к мощности падающего оптического излучения. Для большинства материалов, используемых для изготовления фотоприемников, показатель преломления света на границе воздух – полупроводник $\sim 3,5$ и соответствующий этому значению коэффициент френелевского отражения при нормальном падении света $R = 0,3$. Один из способов уменьшения коэффициента отражения заключается в «искусственном» изменении условий отражения электромагнитной волны от поверхности кристалла. Потери на отражение могут быть почти полностью устранены при нанесении на поверхность полупроводника соответствующих «просветляющих» покрытий.

Коэффициент поглощения α численно равен значению обратного расстояния от поверхности полупроводника (глубины), при котором первоначальная мощность падающего излучения уменьшается в e раз, где e – основание натурального логарифма.

В случае, если координата x направлена в глубь полупроводника, нормально к его поверхности, можно записать:

$$P(x) = P_{\text{ПАД}}(0)e^{-\alpha x}, \quad (4.5)$$

где $P_{\text{ПАД}}(0)$ – мощность оптического излучения, падающего на поверхность полупроводника.

Глубина поглощения $\chi = 1/\alpha$ – толщина слоя полупроводника, после прохождения которого поток излучения уменьшается в $e = 2,7$ раза.

Зависимость коэффициента поглощения от длины волны излучения (частоты излучения) называется спектром поглощения полупроводника или *спектральной характеристикой поглощения*.

Число фотонов, падающих на единицу поверхности полупроводника в 1 с при монохроматическом оптическом излучении, называется *плотностью фотонов N_0* .

Световой поток, взаимодействующий с кристаллом, учитывая коэффициент отражения R , определяется формулой:

$$\Phi(x) = (1 - R)N_0\hbar \exp(-\alpha x). \quad (4.6)$$

После того, как оптическое излучение миновало границу воздух – полупроводник начинается его поглощение веществом и образование фотоносителей в полупроводнике (образование дополнительных дырок и электронов при поглощении полупроводников фотонов). При этом различают несколько механизмов поглощения

При собственном поглощении энергия оптического излучения затрачивается на разрыв валентной связи в атоме и перевод электрона из валентной зоны полупроводника в зону проводимости. Для перевода электрона в зону проводимости необходимо, чтобы энергия фотона превышала ширину запрещенной зоны:

$$W_{FOT} = \hbar \nu \geq W_{33}. \quad (4.7)$$

Значит, собственный фотоэффект в полупроводниках возможен при воздействии на полупроводник излучения с длиной волны, меньшей некоторого граничного значения:

$$\lambda \leq \lambda_{ГР} = \frac{\hbar c}{W_{33}} \approx \frac{1.23}{W_{33}}. \quad (4.8)$$

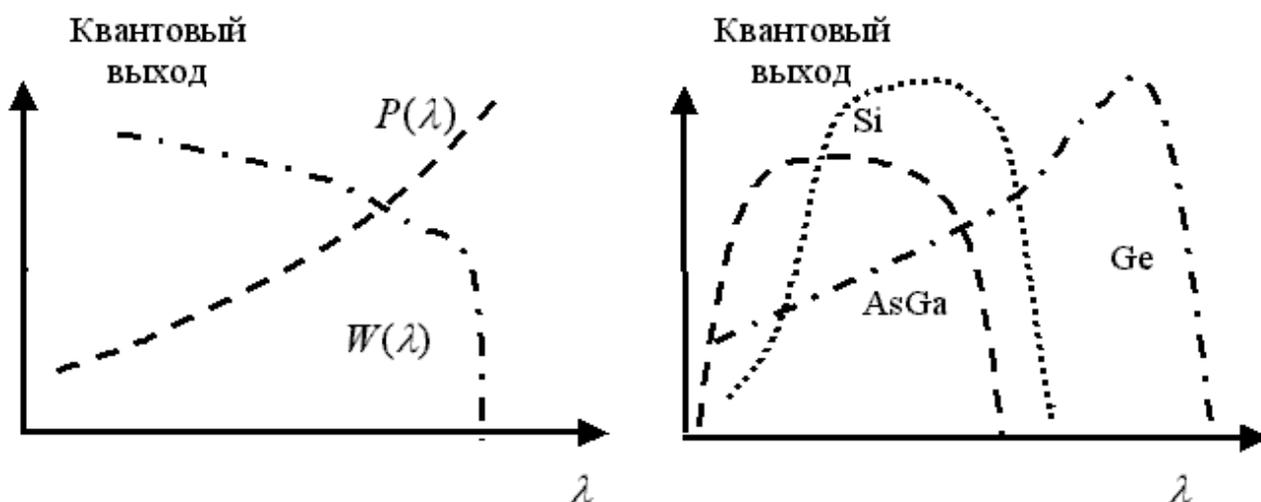


Рис. 4.4. Спектральные характеристики полупроводников

Как известно, четко выраженную длинноволновую границу называют «красной границей фотоэффекта». Красная граница соответствует минимальной энергии фотона, необходимой для перевода электрона из валентной зоны в зону проводимости. При длине волны $\lambda > \lambda_{ГР}$ собственный фотоэффект в полупроводнике невозможен; при $\lambda \leq \lambda_{ГР}$ фотоэффект в полупроводнике возможен.

Зависимость квантового выхода (количества генерированных падающим оптическим излучением фотоносителей заряда) от длины волны имеет вид, показанный на рис. 4.4.

При примесном поглощении энергия фотона излучения затрачивается на ионизацию атомов примеси. При комнатной температуре примесные атомы в основном ионизированы, поэтому энергия фотонов не поглощается из-за отсутствия неионизированных атомов примеси. По этой причине, если хотят использовать примесное поглощение, то полупроводники охлаждают до низких температур $\sim 77^{\circ}$ К.

4.2. Некогерентные и когерентные оптические излучатели

4.2.1 Устройство оптического излучателя, его характеристики и параметры

Оптический излучатель это общее название группы полупроводниковых приборов, преобразующих электрическую энергию в электромагнитное излучение оптического диапазона, создающих оптическое излучение при пропускании тока через $p - n$ - переход (через структуру полупроводников, подобных электронно-дырочному переходу), включенный при прямом включении. Такой прибор формально имеет устройство, подобное полупроводниковому диоду, несимметричному, с точки зрения концентрации введенной примеси, с той лишь разницей, что в нем имеется специальная область полупроводника, предназначенная для вывода излучения. По способу вывода излучения оптические излучатели подразделяются на диоды с поверхностным и торцевым излучением. В диодах с поверхностным излучателем «свет» излучается в направлении, перпендикулярном поверхности перехода. Излучающей (активной) является область только по одну сторону $p - n$ - перехода, как правило, p - область. В диодах с торцевыми излучателями вывод излучения, выхо-

дящего из активного слоя, осуществляется с торца электронно-дырочной структуры. Полупроводники излучающих структур, как уже отмечалось, должны иметь соответствующую ширину запрещенной зоны.

Как следует из принципа инжекционной электролюминесценции (4.2), желательно, чтобы именно в излучающей p – области было максимальное количество инжектированных электронов. Также известно, что для получения односторонней инжекции следует в n - область ввести значительное количество донорной примеси. Полезной компонентой тока, обеспечивающей излучательную рекомбинацию в p – области, является ток, создаваемый инжектированными электронами. Как следует из п. 2.2.2 эффективность инжекции определяется тем, насколько ток отличается от полного тока и характеризуется коэффициентом инжекции. Для выявления доли тока инжекции электронов, как «полезного» тока, в полном токе, вводят коэффициент инжекции:

$$\gamma = \frac{I_n}{I} = \frac{I_n}{I_p + I_n + I_{рек} + I_{упов}}, \quad (4.9)$$

где – I_p – дырочная составляющая тока, обусловленная инжекцией дырок из p – области в n - область;

$I_{рек}$ – ток безизлучательной рекомбинации в области p – n - перехода;

$I_{упов}$ – ток утечки по поверхности p – n - перехода.

Если ток инжекции обозначить через I_n , то мощность оптического излучения выражается в виде:

$$P = \hbar \nu \frac{I_n}{q} \eta_{ВН} \eta_{ОПТ} = \hbar \nu \frac{\gamma I}{q} \eta_{ВН} \eta_{ОПТ}, \quad (4.10)$$

где $\frac{I_n}{q}$ – число электронов, инжектированных в активный слой в единицу времени;

$\hbar \nu$ – энергия фотона;

η_{BH} – внутренний квантовый выход (отношение числа возникших фотонов к числу электронов, инжектированных в p – область);

η_{OPT} – коэффициент вывода оптического излучения (коэффициент потерь при выводе оптического излучения из диода).

Тогда, как следует из (4.10), качество излучающего диода можно характеризовать интегральным показателем излучательной способности – внешним квантовым выходом η (отношением числа фотонов, выходящих из диода $P/\hbar\nu$, к полному числу носителей заряда, протекающих

через него $\frac{I}{q}$):

$$\eta = \gamma\eta_{BH}\eta_{OPT}. \quad (4.11)$$

Произведение $\gamma\eta_{BH}$ определяет эффективность инжекционной электролюминисценции. При этом, однако, следует иметь в виду, что даже при большом значении $\gamma\eta_{BH}$ внешний квантовый выход может оказаться малым вследствие плохого вывода оптического излучения диода во внешнюю среду (малого η_{OPT}).

К уменьшению η_{BH} приводят безизлучательная рекомбинация на дефектах структуры и поглощение возникших фотонов в самом полупроводнике (самопоглощение).

Уменьшение выводимого излучения во внешнюю среду (малое значение η_{OPT}) происходит из-за френелевского отражения от границы раздела сред, полного внутреннего отражения, а также из-за «обратного и торцевого» излучения. Как известно из физики, излучение, падающее на поверхность раздела двух сред под углом, меньшим критического $\theta < \theta_{KP}$, частично отражается от границы. Это френелевские потери. Если на поверхность полупроводника нанести диэлектрическую пленку с соответствующим значением толщины и показателем преломления, то она будет оказывать просветляющее действие и выход оптического излучения увеличится.

При падении излучения из оптически более плотной среды (полупроводник) на границу раздела, где вторая среда оптически менее плот-

ная (воздух), под углом, большим критического $\theta > \theta_{KP}$ для оптического излучения выполняется условие полного внутреннего отражения. Эта часть излучения, отразившись от границы кристалла внутрь, теряется, в конце концов, за счет самопоглощения. Генерируемое в активной области полупроводника оптическое излучение является спонтанным, некогерентным и характеризуется тем, что «лучи» направлены равновероятно во все стороны. Поэтому часть излучения направляется не туда, куда это нужно (в «окно» $p-n$ области, предназначенное для его вывода), а в совсем нежелательные области диода: в n – область, где быстро поглощается; в «торец» $p-n$ - перехода.

Обычно внешняя квантовая эффективность излучающих диодов с поверхностным излучением составляет около 3%, а с торцевым – 1%.

В настоящее время наилучшие параметры излучения имеют диоды, изготовленные на основе так называемых гетероструктур. «Гетеро» – другой, поэтому гетероструктура – это электронно-дырочная структура, образованная соединением двух различных по химическому составу (по ширине запрещенной зоны) материалов.

Необходимость сильного легирования n - области («эмиттера») полупроводника с большой шириной запрещенной зоны приводит к образованию большого числа дефектов в полупроводнике и, соответственно, к снижению η_{BH} .

При применении гетеропереходов значение $\gamma \cong 1$ достигается использованием в качестве инжектирующей области $p-n$ перехода полупроводника с шириной запрещенной зоны, большей чем та область, куда происходит инжекция. При этом односторонняя инжекция уже практически не связана со степенью легирования «эмиттерной» и « p – базовой» областей, как это имеет место в обычном (гомогенном) $p-n$ переходе. Это позволяет обеспечить малое количество дефектов и практически полное отсутствие самопоглощения при выводе излучения.

Другой отличительной особенностью гетероструктур является разница в оптических свойствах «базы» и «эмиттера». Спектральная характеристика «базы», с малой шириной запрещенной зоны, оказывается сдвинутой в область более длинных волн по отношению к спектральной характеристике широкозонного «эмиттера». Скачок коэффициента преломления на границе активного слоя получается большим, чем 0,2, что

это позволяет выводить оптическое излучение через n – область практически без поглощения.

Все преимущества гетероструктур могут быть достигнуты при высоком качестве гетероперехода. Для этого необходимо иметь хорошее совпадение параметров структуры по обе стороны металлургической границы; различие постоянных кристаллических решеток не должно превышать 0,01%; близкими должны быть коэффициенты температурного расширения. Для использования в гетеропереходах весьма перспективными материалами являются полупроводниковые соединения, получаемые на основе бинарных (тройных и четверных) структур (твердых растворов).

Принципиально важной особенностью твердых растворов является возможность плавно и в широких пределах изменять параметры материала путем изменения его состава. Следует иметь в виду и тот факт, что для ряда таких соединений существует слабое рассогласование постоянных решеток, что дает возможность создавать «идеальные» гетеропереходы, у которых на границе раздела двух материалов отсутствуют поверхностные состояния.

Для изготовления излучающих диодов широкое распространение получили структуры на основе прямозонных или близких к этому полупроводников $Al_xGa_{1-x}As / Al_yGa_{1-y}As$, где x и y – молярные доли. Возможны, например, значения $x = 0,33$, $y = 0,66$; $x = 0,35$, $y = 0$ и прочие.

Рассмотрим энергетические диаграммы излучающей гетероструктуры $GaAs / AlGaAs$, показанные на рис. 4.5.

Полупроводник $GaAs$ p - типа имеет узкую запрещенную зону, а полупроводник $AlGaAs$ n – типа – широкую. При создании контакта этих двух материалов уровни Ферми выравниваются, как это было и в обычном гомогенном $p - n$ - переходе. Однако при этом, в отличие от энергетической диаграммы гомогенного $p - n$ - перехода, на диаграмме гетероперехода появляются разрывы (скачки энергии) в зоне проводимости и в валентной зоне, а также на металлургической границе одного из разрывов образуются «пик» и «яма».

В зоне проводимости величина разрыва обусловлена разностью работ выхода электронов из p и n – полупроводника в вакуум, а в валентной зоне, кроме того, еще и добавкой, вызванной неравенством ширины запрещенной зоны. Поэтому потенциальные барьеры для электро-

нов и дырок в гетеропереходе будут различными: потенциальный барьер для электронов в зоне проводимости существенно меньше, чем для дырок в валентной зоне

Если приложить прямое напряжение к $p - n$ - гетероструктуре (рис. 4.6), то потенциальный барьер для электронов уменьшается и электроны из n - полупроводника легко «инжектируют» в p - полупроводник. Потенциальный барьер для дырок, находящихся в p - области также уменьшится, но все же останется достаточно большим для того, чтобы инжекции дырок из p - в n - область практически не было.

Итак, при приложении прямого напряжения к $p - n$ - гетероструктуре имеет место односторонняя инжекция – только электронов из широкозонного слоя («эмиттера») в узкозонный слой («базу»).

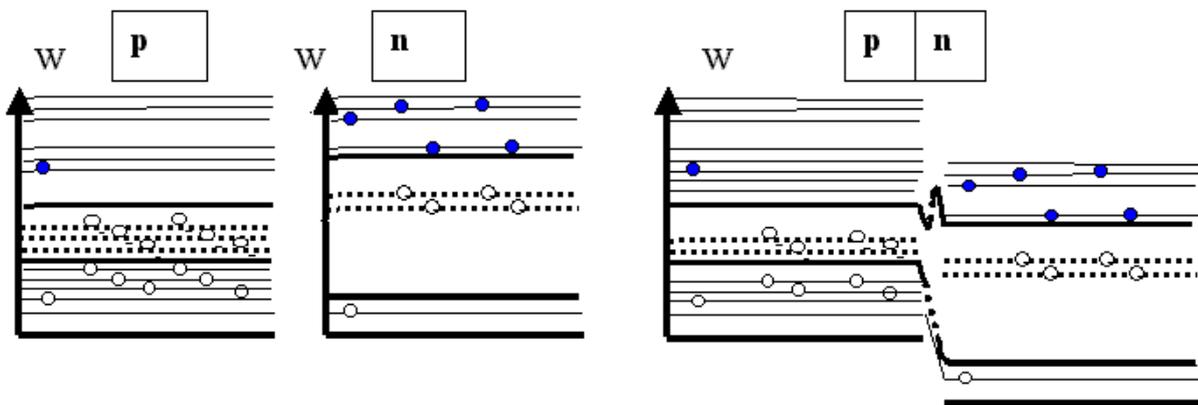


Рис. 4.5. Полупроводники с различной шириной запрещенной зоны и энергетическая диаграмма гетероструктуры в равновесном состоянии

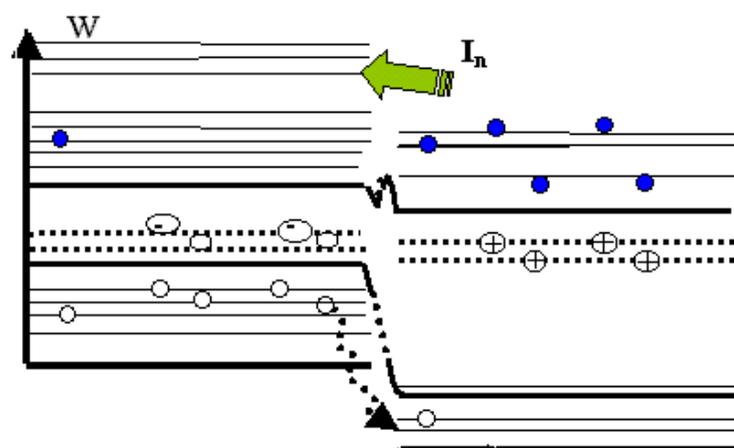


Рис. 4.6. Гетероструктура при прямом включении

Структура, содержащая широкозонный эмиттер и узкозонную базу, называется одинарной гетероструктурой. Наряду с одинарной в излучающих диодах используют двойную гетероструктуру, в которой посредине находится дополнительный широкозонный p – слой, того же, что и база типа проводимости. Двойная гетероструктура позволяет значительно повысить быстродействие излучающего диода. Для лучших образцов на одинарной гетероструктуре внешний квантовый выход составляет 3 – 4%, а время переключения 40 – 80 нс; двойные гетероструктуры имеют примерно такое же значение квантового выхода, а время переключения 20 – 30 нс.

4.2.2 Некогерентные излучатели – светодиоды и полупроводниковые средства отображения информации

Значительную часть информации (около 80 %) человек получает благодаря органам зрения. Однако, что касается современной электронной техники, то в ней информация закодирована в характеристиках и параметрах электрических сигналов. К примеру, в основе принципа действия любой вычислительной машины лежит работа миллиардов простейших электронных элементов – ячеек логики, которые выполняют логические операции над закодированными в виде двоичных чисел (электрических потенциалов) «высказываниями». Чтобы ее визуализировать, т. е. представить в форме, удобной для восприятия органами зрения (а также для фотографирования, печати и т. п.), требуются специальные приборы. Такие приборы должны иметь возможность преобразовывать электрический сигнал в видимое изображение, адекватно отображающее характеристики сигнала.

Средства визуального отображения информации могут быть построены на различных физических принципах. Уникальные возможности для преобразования электрических сигналов в видимое изображение предоставляют полупроводниковые излучающие приборы. Важными их достоинствами является совместимость уровней управляющих их напряжений с напряжениями логических схем и токами микроэлектронной техники, возможность интегрального исполнения.

Полупроводниковые излучающие приборы визуального отображения информации сейчас применяются для осуществления следующих функций.

1. *Сигнализации* – сообщения человеку о факте перехода контролируемой величины из одной области значений в другую. Простейшие примеры этого – сигнализация об отсутствии или наличии питающего напряжения, включение нижнего или верхнего регистра клавиатуры, сигнализация о поступившем сообщении, об опасности и т. п. Исполнение такого прибора – в виде единичного элемента отображения (светодиода), предназначенного для представления информации в виде точки или другой простейшей геометрической фигуры. СД может быть в виде отдельного кристалла, без какого либо корпуса, либо с полимерной герметизацией.

2. *Освещения* – использования прибора в качестве источника света для освещения, в случае их функционирования в условиях плохой освещенности, постоянных надписей, надписей на кнопках (к примеру, в мобильных телефонах), меток на экране, шкале, подсвета жидкокристаллических индикаторов и т.п. Реализацию функции освещения Вы многократно наблюдали на панели автомобиля, мобильного телефона. Исполнение такого прибора – в виде единичного элемента отображения (мощного яркого светодиода), который создает большую освещенность.

3. *Индикации* – представления в безстрелочных измерительных приборах результатов контроля или измерения непрерывно меняющейся информации в аналоговой форме, а также визуального отображения тенденций изменения наблюдаемой величины. Исполнение такого прибора – в виде шкального индикатора, предназначенного для отображения информации в виде линии или набора прямоугольников, заполняющих пространство индикатора по мере роста наблюдаемой величины. Исполнение такого прибора – в виде выполненного на одном кристалле набора излучающих элементов.

4. *Отображения цифровой, буквенной, буквенно-цифровой, графической информации* – воспроизведения в форме удобной для наблюдения букв, цифр, мнемонических символов и простейших графиков, а также различных изображений. В отличие от приборов сигнализации и индикации, имеющих возможность наблюдать бинарную информацию (типа включен – выключен) или же использовать для этого положение

светящейся точки внутри полоски, приборы отображения позволяют воспроизводить значительно больший объем информации. Исполнение такого прибора может быть: а) *сегментным*, в таком виде, когда единичные излучающие элементы представлены и сгруппированы в виде сегментов цифры или буквы; б) *матричным*, когда конфигурация единичных индикаторов представляет собой матрицу элементов, позволяющую синтезировать цифру, букву или какое-либо другое изображение. Могут выпускаться в виде отдельного кристалла с выводами, без какого либо корпуса, либо в корпусе с полимерной герметизацией.

Полупроводниковые излучатели, работающие в диапазоне видимого для глаза оптического излучения, получили название *светоизлучающего диода (СИД)* или, светодиода (СД). В зарубежной литературе используют английскую аббревиатуру LED –light-emitting diode.

Упрощенная конструкция единичного СИД показана на рис. 4.7.

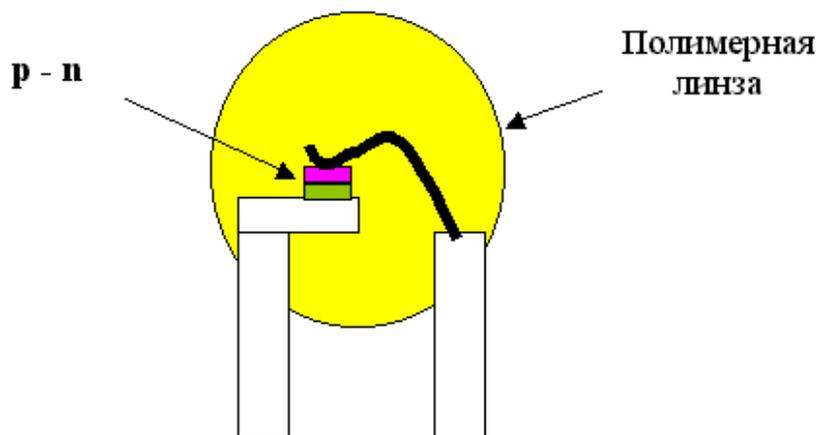


Рис. 4.7. Упрощенная конструкция светодиода

На одном из медных рамочных держателей размещается $p-n$ - гетероструктура, полученная, к примеру, на основе трехкомпонентного твердого раствора $Al_{0.35}Ga_{0.65}As/GaAs$. Излучение такой структуры имеет красный цвет, что соответствует длине волны 700 нм. Эффективный вывод излучения осуществляется с помощью отражательной пластины. К $p-n$ структуре подведены выводы от анода и катода. Конструкция залита полимерным компаундом с диспергирующим светорассеивающим наполнителем. Наполнитель позволяет расширить размеры светящегося пятна (расширить диаграмму направленности излучения). Одновременно он резко понижает интенсивность отраженного диодом

внешнего света, и, тем самым, снижает эффект отсвечивания для «невключенного» диода.

Подобные СИД, с компаундом или без него, применяются в большинстве случаев как миниатюрные единичные излучатели и в более сложных средствах отображения информации. Простая конструкция с $p-n$ переходом обеспечивает низкую стоимость индикаторных средств.

Для описания излучающих диодов используют набор характеристик и параметров, которые описывают СИД как элемент системы «оператор – устройство отображения информации» (с точки зрения качества отображения информации и надежности ее восприятия) и как элемент электрической цепи.

Одной из важнейших характеристик СИД является *цвет свечения*. Существующие стандарты четко определяют назначение каждого цвета, к примеру, в устройствах сигнализации: красный - опасность; желтый – внимание; зеленый – все в норме и т. д. За последние годы требования к цвету свечения существенно возросли, когда появились большие табло, элементы экрана, «бегущая строка» и т. д. Поэтому важно знать спектральную характеристику СИД, определяющую цвет свечения индикатора (доминирующую длину волны, координаты цветности и т.д.). Спектральную характеристику следует учитывать и при выборе фильтра, без которого применение средств отображения информации обычно малоэффективно. Для получения нужного цвета используют материалы с различной шириной запрещенной зоны для изготовления $p-n$ перехода.

Среди *эргономических параметров* следует отметить угол обзора свечения СИД. Под углом обзора понимают максимальный угол между нормалью к центру информационного поля и направлением от этого центра к глазу оператора, при котором обеспечивается безошибочное считывание отображаемой информации при заданном значении силы света, внешней освещенности и расстояния наблюдения. Для единичных индикаторов угол обзора совпадает с углом оптического излучения $p-n$ перехода.

Следующим эргономическим параметром является *контраст прибора отображения информации*.

СИД обычно закрепляется на внешней панели прибора и при этом он должен быть хорошо виден при окружающем освещении. Внешняя освещенность может меняться в значительных пределах, что вызывает

определенные трудности при восприятии информации. Особенно сложные условия освещенности в приборах подвижных (кабинах автомобилей, самолетов и т. д.) и переносимых средств. Естественная освещенность у них меняется в зависимости от времени года, времени суток, от порядка размещения на приборной доске и т. д. В случае достаточно высокой окружающей освещенности для улучшения восприятия визуальной информации следует усиливать контраст между излучением и фоном. Контраст усиливают часто путем использования избирательных оптических фильтров.

Основным светотехническим параметром СИД является *сила света*, определяемая как световой поток, приходящийся на единицу телесного угла в направлении, перпендикулярном плоскости излучающего кристалла. Характеризуют свечение светодиода в избранном направлении также яркостью. Основные светотехнические параметры зависят от вида (постоянный или импульсный) и величины тока, протекающего через диод в прямом направлении, а также от температуры окружающей среды. При импульсном воздействии светодиод «светит» ярче, поэтому во многих случаях на него подают импульсное напряжение, а средства отображения информации работают в так называемом «мультиплексном» режиме. При повышении температуры сила света уменьшается на 50 – 70%. При понижении температуры до минус 60⁰ С сила света увеличивается в 1,5 – 3,5 раза.

В последние годы стали выпускать СИД повышенной яркости, предназначенные не только для цветных знаковых и текстовых табло, но и для цветных телевизионных дисплеев больших размеров (для воспроизведения изображений на площадях, стадионах и т. д.). В них используются материалы группы $A^{III} B^V$ (AlInGaP, InGaN и др.). Их называют СИД «солнечной серии» и «революционной» конструкции (SunPower Series LEDs, Super Flux LEDs). Такие СИД пригодны для применения в любое время суток, независимо от времени года и погоды, а также в помещениях – независимо от освещения.

4.2.3 Некогерентные излучатели – диоды инфракрасного диапазона

Некогерентные излучатели в настоящее время находят широкое применение также в качестве передатчика оптического излучения сис-

тем управления, оптических систем связи и передачи данных различной протяженности.

Хорошо известным примером использования диода для получения оптического излучения инфракрасного диапазона (ИК диапазона) может служить передатчик (излучатель) сигналов устройств дистанционного управления телевизорами, видеомагнитофонами, кондиционерами и т.д. В качестве источника оптического излучения ИК диапазона диод может использоваться в устройствах считывания информации с различных этикеток, упаковок товаров, лекарств, имеющих, к примеру, так называемый штриховой код международной классификации.

Идея использования излучателей в качестве передатчиков оптических систем связи восходит к 1880 г., когда Александр Белл осуществил, с помощью фототона, беспроводную (с помощью оптического луча) передачу звукового сигнала на расстояние 213 м. В настоящее время, используя так называемое «оптическое волокно», построены многочисленные волоконно-оптические системы, которые используются для передачи различной информации в магистральных линиях связи (на далекие расстояния), в абонентских и межучрежденческих линиях, в линиях передачи данных различных компьютерных сетей, для внутренней связи в пределах здания, корабля, самолета, автомобиля и, даже, в линиях межсоединений между сверхбольшими интегральными схемами.

Можно указать основные преимущества оптических систем связи.

1. Нечувствительность к воздействию паразитных электромагнитных полей (ЭМ помех). Известно, что из-за широкого применения различных радиоэлектронных устройств (телевизионных передатчиков, мобильных телефонов, передатчиков систем связи различных служб и т.п.) и электроэнергетических установок уровень электромагнитных помех постоянно растет.

2. Невозможно «перехватить» информацию.

3. Очень большая скорость передачи данных.

4. Сравнительно невысокая стоимость.

При применении излучающего диода в качестве передатчика системы связи, управления или передачи данных к числу важнейших требований к прибору следует отнести:

а) высокую интенсивность излучения (чем больше мощность оптического излучения, тем реже надо устанавливать ретрансляторы - усилители сигнала);

б) высокое быстродействие (чем оно больше, тем больше информации за одно и то же время можно передать).

В принципе, источники оптического излучения могут работать как в видимом, так и в инфракрасном диапазоне длин волн. Однако, из-за необходимости обеспечить минимум затухания в «волокну» в настоящее время используют диапазон $\lambda = 0,9 - 1,7$ мкм. Наиболее дешевы системы связи на $\lambda = 1,3$ мкм. Для линий связи длиной более 100 км выгодно использовать $\lambda = 1,55$ мкм (уровень затухания – 0,2 дБ/км; снижение амплитуды сигналов в 2 раза происходит на расстоянии 15 км, что позволяет использовать ретрансляторы на расстоянии 100 км).

Теоретически возможно получить затухание $10^{-2} - 10^{-3}$ на длинах волн $\lambda = 3 - 5$ мкм. Это позволяет осуществить связь, к примеру, между Америкой и Японией вообще без применения ретрансляторов.

Для получения инфракрасного оптического излучения используют диодные $p - n$ переходы на основе арсенида галлия, легированные кремнием и цинком. По способу вывода излучения они подразделяются на диоды с поверхностным и торцевым излучением. Спонтанное оптическое излучение диода инфракрасного диапазона является изотропным, поэтому принимаются специальные меры по уменьшению потерь при передаче излучения диода в оптическое волокно. Для вывода излучения часто в полупроводнике вырезается отверстие, куда вводится оптическое волокно. Также используются специальные линзы, повышающие эффективность связи излучения диода с оптическим волокном. Обычно делают диаметр источника излучения меньше диаметра сердцевины оптического волокна.

Изменение интенсивности излучения в ИК диодах осуществляется путем вариации тока инжекции. Для повышения быстродействия диода (увеличения скорости передаваемой информации) стараются уменьшать толщину активного слоя $p-n$ перехода. Увеличение тока инжекции также способствует повышению быстродействия.

4.2.4 Когерентные излучатели – полупроводниковые лазеры

В излучающих диодах инжекция электронов в активную область приводит к спонтанной рекомбинации и вызывает некогерентное оптическое излучение. В дальнейшем, при прохождении излучения через полупроводник, если частота излучения $\nu \geq W / \hbar$, кванты света поглощаются, переводя электроны из зоны проводимости в валентную зону. Соответственно, при прохождении излучения через полупроводник его интенсивность уменьшается.

Предположим, что в дополнение к сказанному в полупроводнике действует еще и внешнее, как говорят, «возбуждающее» оптическое излучение. Пока не важно, как оно возникает. Оно стимулирует переходы «возбужденных» им электронов (назовем их «активными») из зоны проводимости в валентную зону, а также, сопровождается практически мгновенными обратными переходами «активных» электронов с более высокого уровня на более низкий. Естественно, что обратный процесс сопровождается испусканием фотонов.

Испускание фотонов, вызванных переходами «активных» электронов с более высокого уровня на более низкий, может происходить по-разному. Возможен спонтанный обратный переход «активных» электронов, при котором момент испускания и направление вектора поляризации каждого генерируемого при этом фотона случайны, а результирующий поток излучения определяется лишь среднестатистическими параметрами. Имеется вероятность возникновения и так называемого «вынужденного» обратного перехода электрона. При этом все «активные» электроны «сваливаются» на нижние уровни почти одновременно, взаимосвязано и так, что испускаемые фотоны почти неотличимы, как между собой, так и от вызвавших их «возбуждающих» фотонов. При этом, как говорят, возникает «вынужденное» когерентное излучение – электромагнитное излучение оптического диапазона, которое совпадает по направлению излучения, частоте, фазе и поляризации с вынуждающим излучением.

Слово «лазер» (LASER) – аббревиатура английского выражения Light Amplification by Stimulated Emission of Radiation, которое переводится как «усиление света за счет вынужденного излучения». Первоначально это слово указывало на усиление света, но сегодня это название

большой группы приборов, позволяющих генерировать монохроматическое когерентное излучение оптического диапазона определенной длины волны.

Определим условия, при котором возможно появление лазерного излучения на GaAs $p - n$ -переходе.

Во-первых, необходимо «выделить» из всего оптического излучения «вынужденное» излучение, как-то его «сохранить» и «сконцентрировать» в пространстве (повысить «плотность» потока «вынужденного» излучения). Целесообразно для решения этой задачи, как известно, применить некую «резонансную» систему, которая бы вынужденные фотоны делала преобладающими над «спонтанными». В полупроводниковых лазерах для этой цели обычно применяют оптический резонатор Фабри-Перо, образованный плоскопараллельными отполированными гранями кристалла (зеркальными сколами граней кристалла). По сути это система двух зеркал обеспечивающая многократное прохождение волны «вынужденного» оптического излучения через «активное» вещество $p - n$ – перехода и «подающая» часть излучения с выхода на вход. Причем между зеркалами могут существовать только длины волн, определяющие «вынужденные» фотоны.

Во-вторых, лазерное оптическое излучение будет возможно лишь в том случае, если количество лазерных переходов будет существенно большим, чем число спонтанных переходов и переходов, связанных с поглощением вынужденного излучения. Это можно осуществить только за счет внешних воздействий или, как говорят, за счет «накачки». В полупроводниковых лазерах накачка осуществляется за счет создания «инверсии» населенности и инжекции носителей заряда.

Уже отмечалось, что испускание фотона происходит только тогда, когда в зоне проводимости имеется электрон, а в валентной зоне одновременно «соответствующая» дырка. Если рассмотреть функции, отражающие вероятности распределения электронов по уровням зоны проводимости и валентной зоны, зафиксировать ситуации одновременного наличия электрона в зоне проводимости и дырки в валентной зоне, превышения скорости излучения над скоростью поглощения, то можно получить, что условием «накачки» и для лазерного излучения в $p - n$ – структуре является:

$$W_{Fn} - W_{Fp} > W_{3П} - W_{B3}. \quad (4.12)$$

То есть, разность уровней Ферми n и p полупроводников $W_{Fn} - W_{Fp}$, участвующих в контакте, должна быть больше разностей уровней «дна» зоны проводимости и «потолка» валентной зоны $W_{3П} - W_{B3}$.

Выполнить условие (4.12) можно, если осуществить так называемую «инверсию населенности», т. е. сделать так, чтобы уровни вблизи «дна» зоны проводимости были «заселены» электронами сильнее, чем уровни вблизи «потолка» валентной зоны. Это можно обеспечить в том случае, если концентрации введенных примесей таковы, что материал хотя бы одной из областей (p или n) является вырожденным (у таких полупроводников уровень Ферми располагается выше «дна» зоны проводимости и ниже «потолка» валентной зоны).

На рис. 4.8 показана энергетическая диаграмма полупроводникового лазера на основе GaAs. Чтобы создать режим инверсии населенности в полупроводниковом лазере используется электронно-дырочная структура $p^{+++} - n^{+++}$ с двумя вырожденными полупроводниками (с высокими уровнями легирования).

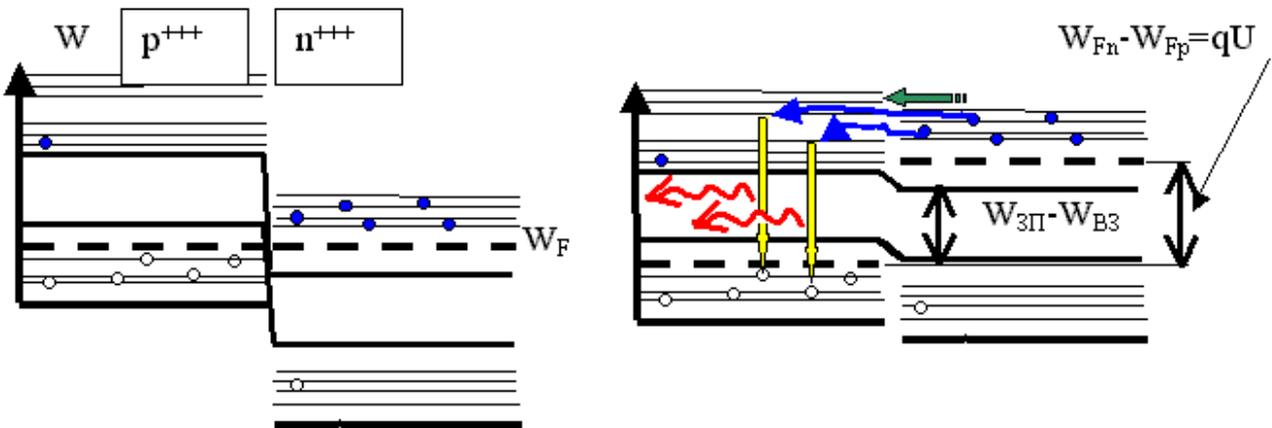


Рис. 4.8. Энергетические диаграммы, поясняющие работу полупроводникового лазера в равновесном состоянии и при приложении прямого напряжения

При подаче на $p^{+++} - n^{+++}$ - переход небольшого прямого напряжения (уровни энергетической диаграммы n^{+++} полупроводника относительно равновесного состояния слегка перемещаются вверх) с ростом тока инжекции, поскольку условие (4.12) не выполняется, происходит спонтанная рекомбинация. Лазер работает как излучающий диод и мощность излучения пропорциональна силе тока.

При протекающем через $p^{+++} - n^{+++}$ - переход токе, который называется пороговым, разность между уровнями Ферми в p^{+++} и n^{+++} полупроводников начинает превышать разницу уровней «дна» зоны проводимости и «потолка» валентной зоны $W_{зП} - W_{вЗ}$. Выполнение условия (4.12) приведет к тому, что количество лазерных переходов будет существенно большим, чем число спонтанных переходов и переходов, связанных с поглощением вынужденного излучения, и процесс спонтанной рекомбинации перейдет в новое качество – генерации оптического излучения. В результате согласованного «вынужденного» перехода электронов из зоны проводимости в валентную зону во всей излучающей части p - области и благодаря использованию оптического резонатора образуется и усиливается «вынужденное» когерентное и монохроматическое лазерное излучение. Длина волны оптического излучения определяется свойствами материалов лазерной структуры. Мощность излучения резко увеличивается. Изменяя приложенное к полупроводниковому лазеру напряжение, а точнее силу тока, можно изменять выходную мощность лазера. Также можно «модулировать» оптический сигнал, создавая «логический ноль» отсутствием излучения и «единицу» – генерацией излучения определенной мощности. Частота следования импульсов модуляции может достигать несколько ГГц.

Итак, важно помнить, что полупроводниковый лазер это «пороговый» прибор. При токах меньших порогового значения излучение является спонтанным и некогерентным; как только $I > I_{пор}$ излучение становится монохроматическим и когерентным (рис. 4.9). Обычно пороговый ток составляет от единиц до сотен миллиампер.

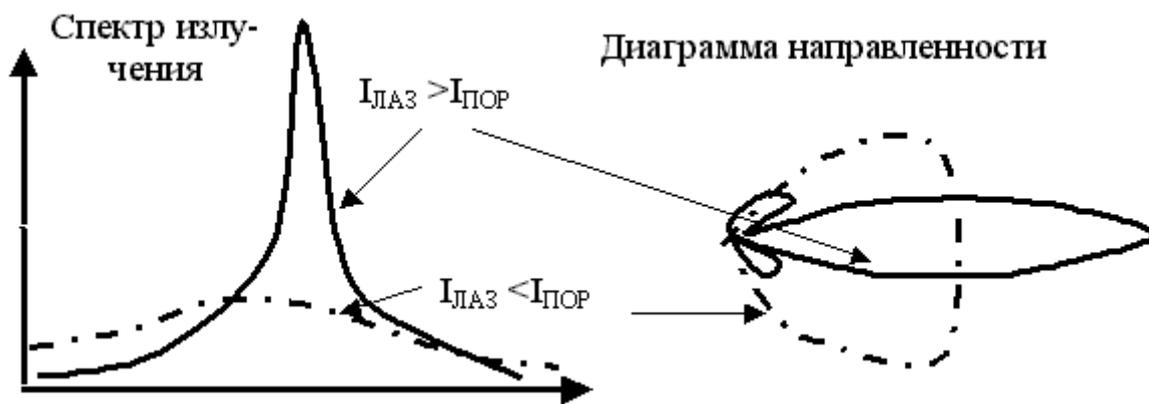


Рис. 4.9. Спектр и диаграмма направленности излучения лазера при токах меньших и больших «порога»

В настоящее время полупроводниковый лазер на основе GaAs, о котором речь шла выше, уступил место лазеру на гетеропереходе $Al_xGa_{1-x}As/GaAs$. В лазере на гетеропереходе инжектированные носители сосредотачиваются в полупроводнике с узкой шириной запрещенной зоны. При этом концентрация инжектированных носителей может на несколько порядков превышать свое равновесное значение. По этой причине состояние инверсии населенности наступает при более низких значениях тока, чем в лазере на p - переходе. В гетеролазере нет необходимости легировать полупроводник до вырождения, так как условие инверсии населенности энергетических уровней выполняется за счет разницы в ширине запрещенных зон. Снижение уровня легирования способствует уменьшению потерь на безизлучательную рекомбинацию и повышению внутренней квантовой эффективности.

В гетероструктуре оптические свойства слоев эмиттера и базы отличаются, так как ширина запрещенной зоны эмиттера значительно шире запрещенной зоны базы. А так как показатель преломления зависит от ширины запрещенной зоны, то в гетеролазере «легче» выводить оптическое излучение.

В конечном итоге, гетеролазеры при относительно небольших токах имеют достаточный КПД, что в свою очередь, позволяет осуществить непрерывный режим генерации при комнатной температуре.

Помимо лазеров с простым гетеропереходом (SH) существуют лазеры с двойным (DH) гетеропереходом. Лазер DH имеет более высокий КПД и более низкий пороговый ток.

Подчеркнем, что лазерное излучение является монохроматичным (одномодовым). Лазер генерирует когерентные оптические колебания на частоте максимального усиления и минимальных потерь излучения в резонаторе. Важно, что лазер преобразует электрическую энергию низкого качества в когерентное излучение – предельно высококачественную энергию.

Кроме резкого сужения спектральной характеристики происходит также некая «фокусировка» оптического излучения, сужение «диаграммы направленности излучения» – угла расходимости оптического излучения в направлении, перпендикулярном поверхности излучения (рис. 4.9).

Лазерное излучение характеризуется пространственно-временными и энергетическими параметрами.

Основными пространственно-временными параметрами являются: а) частота (длина волны) лазерного излучения; б) расходимость лазерного излучения – телесный угол, характеризующий угловое распределение энергии лазерного излучения.

К энергетическим параметрам лазера относятся: а) мощность лазерного излучения; б) энергией лазерного излучения, приходящейся на единицу площади сечения пучка лазерного излучения; в) КПД.

Полупроводниковые лазеры, по сравнению с другими типами лазеров, имеют следующие достоинства.

А. Малые массогабаритные показатели.

Б. Высокий КПД (20 – 50%).

В. Простота осуществления «накачки»: инжекция не требует высоких питающих напряжений и мощностей.

Г. Высокое быстродействие.

Д. Технологическая и эксплуатационная совместимость с элементами волоконной оптики и интегральных схем.

Лазерные диоды применяются в дальномерах (лазерных локаторах для измерения расстояний), лидарах (приборах контроля за состоянием атмосферы), оптических средствах связи, в устройствах считывания информации с оптических дисков.

В устройствах лазерной печати (лазерных принтерах) лазерный луч, включаясь и выключаясь в соответствии с выводимой из компьютера информацией, перемещается по поверхности фоточувствительного

барабана (говорят, сканирует по барабану). При таком «проецировании» изображения с участков поверхности барабана, на которые попал «включенный» оптический луч, уходит заряд и, в целом, на него «записывается» состоящая из зарядов «картина». Если на поверхность барабана нанести специальный порошок, то он «прилипнет» только в тех местах, где «побывал» лазерный луч. Если затем барабан прижать к бумаге и нагреть, то на бумаге получится требуемая копия изображения. Скорость такой печати высока и разрешение составляет примерно 10 – 30 линий на 1 мм. Использование полупроводниковых лазеров позволяет создавать миниатюрные устройства, дающие быструю и качественную печать.

4.3. Приемники оптического излучения

4.3.1. Фоторезисторы

Фоторезистор – это фотоэлектрический прибор с двумя выводами, электрическое сопротивление которого меняется под воздействием оптического излучения.

Простейший пример фоторезистора – полупроводниковая пластина с двумя омическими контактами на концах, помещенная в корпус с «окошком» для проникновения оптического излучения.

При отсутствии оптического излучения в полупроводниковой пластине, при определенной температуре, имеются «свободные» носители заряда с равновесной концентрацией p_0 , n_0 . При приложении к пластине напряжения величина «темновой» удельной проводимости будет равна:

$$\sigma_0 = q(n_0\mu_n + p_0\mu_p). \quad (4.13)$$

При воздействии на полупроводниковую пластину оптического излучения в ней происходит генерация электронно-дырочных пар и концентрации носителей увеличиваются на величины Δp и Δn , пропорциональные интенсивности воздействующего оптического излучения. Величина квантового выхода при генерации электронно-дырочных пар практически равна нулю при $W_{FOT} = \hbar\nu < W_{33}$ и единице при

$W_{FOT} = \hbar \nu \geq W_{33}$ (квант не может произвести более одной электронно – дырочной пары).

Неравновесные носители Δp и Δn создают дополнительную, как говорят, фотопроводимость. Полная удельная проводимость фоторезистора:

$$\sigma = \sigma_0 + \Delta\sigma = q[(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p]. \quad (4.14)$$

Следует иметь в виду, что фотопроводимость может быть, в общем случае, двух типов. При собственной фотопроводимости, как уже отмечалось, поглощение кванта света с $W_{FOT} = \hbar \nu \geq W_{33}$ приводит к переходу электрона из валентной зоны в зону проводимости и образования пары электрон – дырка. При примесной фотопроводимости поглощение кванта наступает при $W_{FOT} = \hbar \nu \geq W_{ИОНПР}$, где $W_{ИОНПР}$ - энергия ионизации примеси. При ионизации донорной или акцепторной примеси образуются «свободные» носители только одного знака – электроны или дырки.

Схема включения фоторезистора в электрическую цепь, в простейшем случае, имеет вид, показанный на рис. 4.10.

При отсутствии оптического излучения фотопроводимость фоторезистора мала. Соответственно, его электрическое сопротивление имеет величину порядка 1 – 50 МОм. В соответствии со свойством простейшего делителя напряжения выходное напряжение имеет величину близкую к нулю.

При воздействии на фоторезистор импульса оптического излучения с интенсивностью $\Phi(t)$ происходит генерация носителей заряда, снижающая сопротивление фоторезистора. Это приводит к тому, что на выходе цепи появляется выходное напряжение. При этом имеет место «фронт» - процесс нарастания напряжения и «спад» – процесс уменьшения напряжения.

Следует иметь в виду, что зависимость фотопроводимости от интенсивности оптического излучения сначала увеличивается, достигает максимума, а затем, уменьшается. Это связано с тем, что при увеличении интенсивности оптического излучения наряду с процессами генерации электронно-дырочных пар все большую роль играют процессы ре-

комбинации, что приводит к уменьшению концентрации свободных носителей и, соответственно, фотопроводимости.

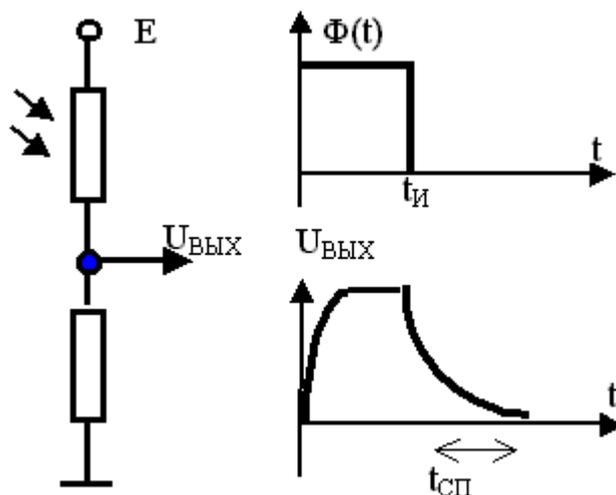


Рис. 4.10. Схема включения фоторезистора и временная зависимость выходного напряжения цепи при воздействии оптического излучения

В видимой части оптического диапазона в качестве материала полупроводниковой пластины фоторезистора используют сульфид кадмия CdS и селенид кадмия CdSe. В инфракрасной области спектра на длинах волн 1 – 5 мкм используют сульфид PbS, селенид PbSe и теллурид PbTe свинца, а на длинах волн 5 – 40 мкм – тройные кристаллические соединения $Hg_{1-x}Cd_xTe$ с узкой шириной запрещенной зоны и примесные полупроводники, к примеру германий. Фоторезисторы на основе примесных и узкозонных полупроводников при эксплуатации охлаждают до температуры 77К.

4.3.2. Фотодиоды

Фотодиод – управляемый оптическим излучением прибор с двумя выводами, фоточувствительный элемент которого имеет структуру полупроводникового диода, работа которого основана на использовании фотогальванического эффекта при обратном включении $p - n$ - перехода.

Упрощенная структура фотодиода изображена на рис. 4.11

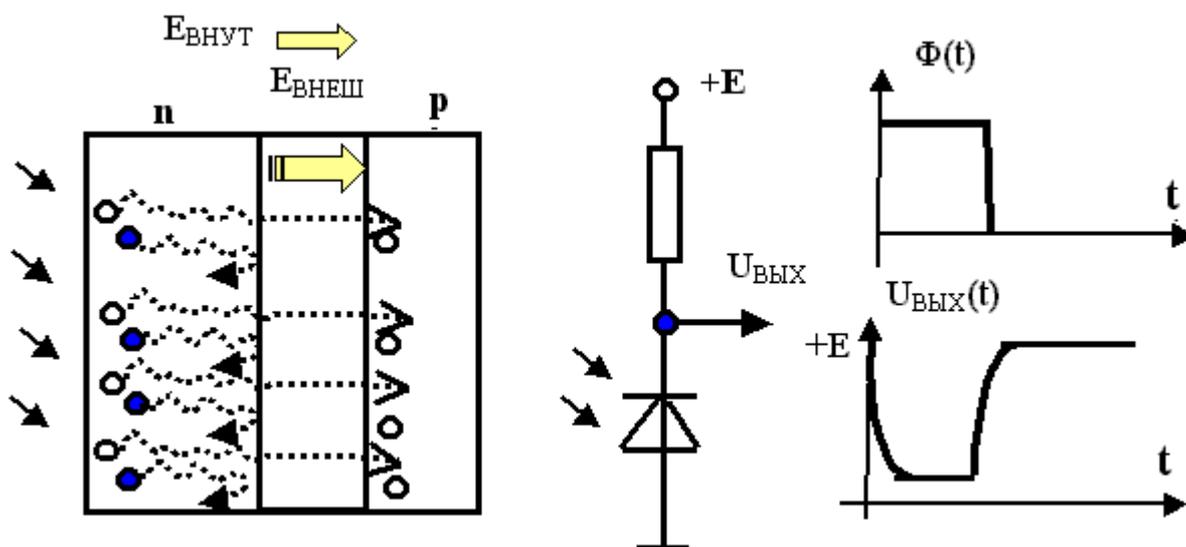


Рис. 4.11. Структура фотодиода, схема его включения и временная зависимость выходного напряжения цепи при воздействии оптического излучения

В исходном состоянии при отсутствии потока оптического излучения, воздействующего на фотодиод, через фотоприемник, как и через обычный диод, протекает обратный ток (его в этом случае называют темновым током). Его величина несколько микроампер. Падение напряжение на резисторе мало, и выходное напряжение равно напряжению источника питания $+E$.

Рассмотрим теперь, что произойдет с фотодиодом, при воздействии на него монохроматического потока оптического облучения с энергией фотонов, превышающей ширину запрещенной зоны полупроводника. Для определенности будем считать, что оптическое излучение воздействует на n – область в направлении, перпендикулярном плоскости p – n -перехода.

1. В результате поглощения фотонов в n – области до глубины примерно χ_0 от поверхности полупроводника возникают электронно-дырочные пары (генерируются фотоносители).

2. Фотоносители диффундируют (рис. 4.11) в глубь n – области. Ширина n – области такова, что во время диффузионного движения фотоносители не успевают рекомбинировать в n – области и доходят до границы обедненного слоя.

3. Суммарное электрическое поле, действующее в обедненном слое, «разделяет» электроны и дырки. Дырки под воздействием элек-

трического поля перемещаются в p – область. Воздействие силы, созданной электрическим полем на электроны имеет противоположный характер, поэтому электроны как бы «отражаются» и скапливаются в n – области у левой границы p – n - перехода.

4. Разделение созданных оптическим излучением фотоносителей приводит к тому, что в p – области возникает дополнительный положительный заряд, обусловленный переходом дырок, а в n – области – отрицательный. Эти заряды создают электрическое поле, вектор напряженности которого направлен в противоположную сторону внутреннего электрического поля.

Теперь возможны две крайние ситуации, которые получили название режимов работы фотодиода.

В фотогальваническом режиме (его еще называют вентильным режимом) внешнее напряжение не подается ($+E = 0$), а «работает» только оптическое излучение. В этой ситуации образованные оптическим излучением и «разделенные» дырки «заряжают» p – область положительно (относительно n – области), а электроны n – область отрицательно (по отношению к p – области). Возникшая таким образом разность потенциалов называется фото – ЭДС. Фото – ЭДС можно измерить, подключив вольтметр с большим внутренним сопротивлением между анодом и катодом фотодиода. Так как сопротивление «нагрузки» при этом стремится к бесконечности (что, фактически соответствует отсутствию нагрузки), то говорят, что фото – ЭДС фиксируется при «холостом ходе» и равна (с противоположным знаком) напряжению холостого хода $U_{\text{ХХ}}$.

Если фотодиод в фотогальваническом режиме это источник ЭДС определенной мощности, то, в общем случае, от него можно «питать» электрической энергией электронные устройства. Если такому источнику подключить короткозамкнутую перемычку, то через нее будет протекать ток «короткого замыкания»

В фотодиодном режиме «работает» оптическое излучение и на диод подается обратное внешнее напряжение.

Без оптического излучения протекание обратного тока обусловлено экстракцией неосновных носителей заряда. Из рис. 4.11. несложно заметить, что движение дырок, генерированных излучением, есть не что иное, как движение дополнительных неосновных носителей заряда (дополнительная экстракция дырок из n – области). В этой связи становится

очевидным, что генерируемые оптическим излучением фотоносители увеличивают обратный ток $p - n$ – перехода ($I_{\text{ОБЩ}} = I_{\text{ОБР}} + I_{\Phi}$). Далее можно записать формулу для фототока:

$$I_{\Phi} = \left(\frac{qSL_{\text{nobl}}\Delta p_n}{\tau_p} \right), \quad (4.15)$$

где L_{nobl} – толщина n – области.

Возвращаясь теперь к электрической цепи с фотодиодом (рис. 4.11), можем отметить следующее. Оптический сигнал в виде то включающегося, то выключающегося потока оптического излучения приведет к тому, что в моменты воздействия на фотодиод оптического излучения будет возникать дополнительный фототок. Падение напряжения на резисторе от фототока увеличится и, соответственно, уменьшится напряжение на выходе цепи. Таким образом, выходное напряжение практически повторяет закон изменения оптического излучения. Фототок, правда, появляется не сразу; он возникает после того, как генерированные оптическим излучением носители заряда дойдут до $p - n$ перехода, перебросятся через обедненный слой и пройдут p – область. По этой причине отрицательный импульс выходного напряжения имеет «фронт». Соответственно, носители, генерированные в конце импульса оптического излучения, должны за какое – то время «покинуть» фотодиод. Вследствие этого у отрицательного импульса появляется «спад».

Следует иметь в виду, что фотодиод «реагирует» на оптическое излучение только определенного спектрального диапазона (диапазона длин волн). С этой точки зрения он представляет полосовой оптический фильтр. Снижение чувствительности в области коротких волн связано с тем, что при уменьшении длины волны энергия оптического излучения поглощается во все более тонком приповерхностном слое, где как раз очень много дефектов кристалла и велика рекомбинация. Спад чувствительности в области длинных волн соответствует длинноволновой границе спектральной чувствительности материала.

При изготовлении фотодиодов используют кремний и германий. Кремниевые фотодиоды предпочтительны в видимой области оптического излучения. Германий является единственным имеющимся в при-

роде полупроводником, ширина запрещенной зоны которого позволяет создавать фотоприемники на спектральный диапазон 0,8 – 1,6 мкм. Однако, поскольку германий непрямозонный полупроводник с малой шириной запрещенной зоны, он имеет повышенный уровень «шума» (случайных флуктуаций тока) при комнатной и повышенной температурах. Весьма перспективными материалами, заменяющими германий, являются твердые растворы. Существуют три системы соединений пригодные для этой цели: AlGaSb, GaInAsP, HgCdTe.

4.3.3. Специальные фотодиоды

При передаче информации по волоконно-оптическим линиям связи важнейшую роль играет быстродействие фотоприемника (чем оно больше, тем с большей скоростью можно передавать информацию). Быстродействие фотодиода, в общем случае, зависит от трех факторов: от длительности процесса диффузии (передвижения) генерированных фотоносителей к $p - n$ - переходу; от времени прохождения $p - n$ - перехода; от величины барьерной емкости $p - n$ - перехода. Как показывают расчеты, быстродействие определяется, в основном, первым фактором. Напрашивается простое решение проблемы – уменьшать толщину n – области. Однако так поступать нельзя, так как при этом снижается чувствительность диода к оптическому излучению.

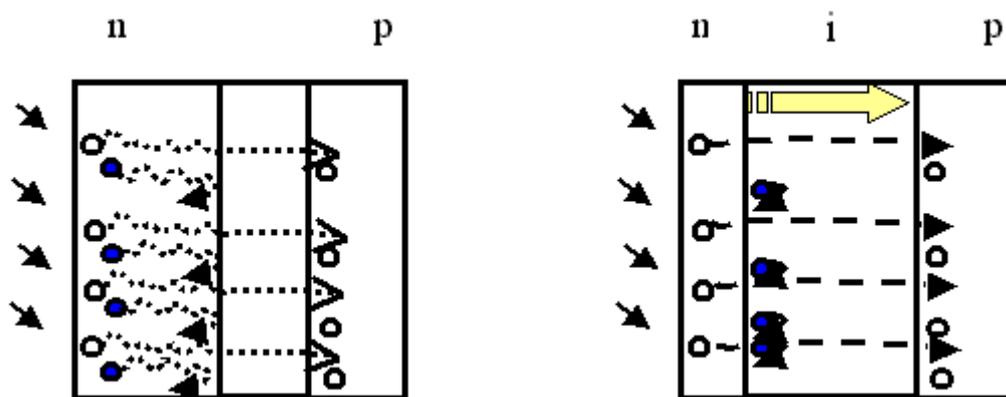


Рис. 4.12. Структура фотодиода и $p - i - n$ фотодиода

Для повышения быстродействия фотодиода, без снижения его чувствительности используют так называемые $p - i - n$ диоды (рис. 4.12.)

В $p - i - n$ фотодиоде между двумя легированными примесями областями с противоположными типами проводимости встроена достаточно широкая область полупроводника с собственным типом проводимости. При этом поглощение оптического излучения происходит теперь как в n , так и в i области. За счет этого, с одной стороны, расширяется область, в которой происходит поглощение оптического излучения, а, с другой стороны, уменьшается барьерная емкость $p - n$ – перехода. Область i имеет удельное сопротивление в $10^6 - 10^7$ раз больше, чем сопротивление легированных областей n и p типов. При обратном включении диода обедненный слой, простираясь от p - к n - области, включает в себя и слой собственной проводимости. В этом случае практически все внешнее электрическое поле «сосредотачивается» в i - области. Величина напряженности электрического поля здесь очень велика.

Дырки и электроны, появившиеся в i области за счет поглощения оптического излучения, фактически генерируются в области сильного электрического поля, поэтому они быстро разделяются электрическим полем и электроны за счет дрейфового перемещения очень быстро «пролетают» через диод.

Таким образом, повышение быстродействия обусловлено тем, что относительно медленный процесс диффузии через n – область, характерный для обычной структуры, заменен в $p - i - n$ структуре относительно быстрым процессом дрейфа фотоносителей в сильном электрическом поле через i – область.

Следует отметить, что малые рабочие напряжения $p - i - n$ фотодиода обеспечивают ему хорошую электрическую совместимость с интегральными схемами.

Обычный фотодиод и $p - i - n$ фотодиод не усиливают фототок. Их квантовый выход не может быть больше единицы. По этой причине они плохо работают, когда приходится иметь дело со слабыми потоками оптического излучения, сравнимых с «шумами» фотоприемника.

Одним из путей создания быстродействующих фотоприемников с высокой чувствительностью является использование явления лавинного пробоя $p - n$ - перехода при обратном его включении. В *лавинном фотодиоде* генерируемые оптическим излучением фотоносители размножаются за счет ударной ионизации атомов в сильном электрическом по-

ле, поэтому лавинный фотодиод отличается усилением исходного фототока в M раз, где M – коэффициент лавинного размножения.

Лавинный фотодиод имеет структуру, сходную с $p - i - n$ структурой. Но к нему прикладывается обратное напряжение такой величины, создается напряженность электрического поля такой силы, что электронно-дырочные пары, рожденные оптическим излучением, не только разделяются электрическим полем, но и ускоряются до такой степени, что их энергия начинает превышать энергию ионизации атомов кристалла. Сталкиваясь затем с нейтральными атомами, сильно ускоренные фотоносители вызывают увеличивающееся в геометрической прогрессии рождение новых электронов и дырок. При таком явлении, именуемом лавинным эффектом, коэффициент усиления фототока может достигать значений порядка 1000. Благодаря сильному электрическому полю времени пролетов носителей через диод ничтожно малы, поэтому лавинные диоды имеют высокое быстродействие.

Следует отметить, что широкое применение лавинных диодов связано с трудностями. Прежде всего, эти трудности определяются тем, что коэффициент лавинного размножения зависит от величины обратного напряжения. Для надежной работы лавинного фотодиода необходима высокая стабильность питающего напряжения (относительные отклонения не должны превышать 0,001%). Кроме того, так как типичными значениями напряжений для фотодиода являются 50 – 100 вольт, то при его работе имеют место большие потери энергии и существует необходимость обеспечения температурной стабилизации.

Фотогальванический эффект, возникающий в $p - n$ - переходе можно использовать для непосредственного преобразования энергии оптического излучения в электрическую. Обычно преобразовывают световую энергию падающего солнечного излучения или свет комнатных светильников в электрическую энергию. Фотоприемник, реализующий этот принцип, называют обычно солнечным преобразователем.

Единичные преобразователи обычно соединяют в батарею, чтобы можно было получить источники питания для калькуляторов или других устройств, работающие от света. Солнечные батареи и преобразователи применяются для питания аппаратуры космических аппаратов, в наземных маломощных энергетических установках различных областей

хозяйства. Солнечные батареи из большого числа элементов устанавливаются в местах, долгое время освещаемых солнцем.

Солнечный преобразователь, по сути, представляет диод с $p - n$ переходом большой площади, способный преобразовывать энергию падающего солнечного излучения (плюс искусственных источников света) в электрическую энергию с КРД порядка 20%.

Спектр излучения Солнца включает в себя практически весь оптический диапазон, но свыше 90 % энергии солнечного света все же приходится на видимое излучение. В качестве материалов для солнечных батарей могут использоваться полупроводники с малой шириной запрещенной зоны, поглощающие большую часть спектра, но они создают большой обратный ток перехода и не позволяют получить высокие значения напряжений холостого хода и приемлемые величины КПД. По этой причине чаще используются кремний, материалы группы $A^{III} B^V$, имеющие низкую стоимость и высокий коэффициент поглощения в области солнечного спектра. Для снижения потерь лучистой энергии используют просветляющие покрытия, а также нанесение с тыльной стороны отражающего покрытия (отражает лучистую энергию в глубь кристалла).

4.3.4. Фототранзисторы

Фототранзистор это полупроводниковый прибор, с двумя взаимодействующими между собой переходами, который, кроме преобразования оптического излучения в электрический ток (фототок), обеспечивает увеличение (усиление) фототока.

В электрическую цепь фототранзистор включается обычно по схеме с общим эмиттером (СОЭ). Входным сигналом для фототранзистора является промодулированное по интенсивности оптическое излучение, выходным ток коллектора или напряжение коллектор – эмиттер.

Предположим, что база транзистора не имеет внешнего вывода и на нее воздействует оптическое излучение. Тогда в фототранзисторе будут происходить следующие процессы (рис. 4.13).

1. Если оптический сигнал отсутствует, то, как известно из теории, через транзистор с «оторванной» базой протекает «сквозной ток»

$$I_{КЭ0} = (\beta + 1)I_{КБ0}.$$

Если через фотоприемное окно оптическое излучение попадает на базу транзистора, то в ней происходит генерация фотоносителей (пар электрон – дырка).

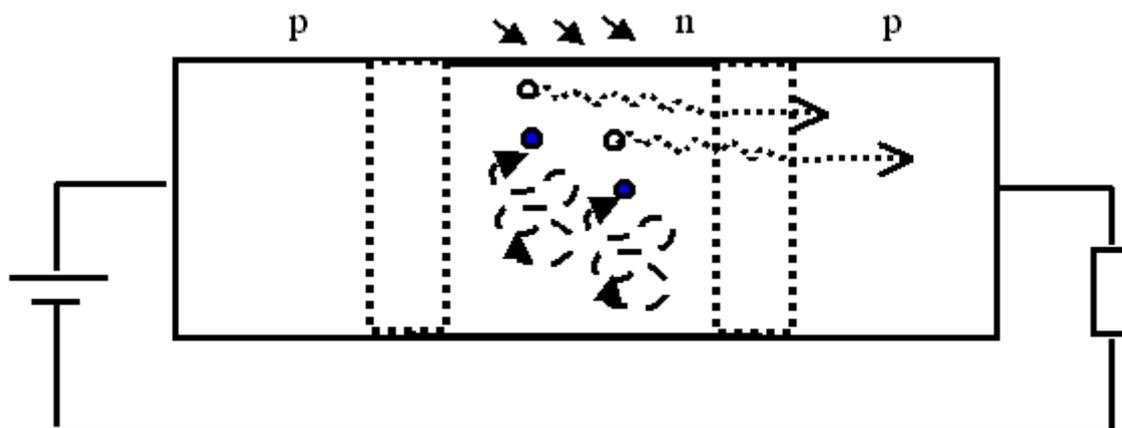


Рис. 4.13. Структура фототранзистора

2. Неравновесные носители, возникшие в базе, начинают за счет диффузии перемещаться по базе.

3. Дырки, дошедшие до коллекторного перехода, как неосновные носители базы, электрическим полем этого перехода перемещаются в коллектор, тем самым создают дополнительный фототок коллектора. Электроны, являясь основными носителями базы, не могут преодолеть силу электрического поля коллекторного перехода, поэтому они накапливаются в базе, понижая, по мере накопления, потенциал базы.

4. Понижение потенциала базы эквивалентно увеличению напряжения на эмиттерном переходе. По этой причине фотоэлектроны базы, опосредованно, вызывают дополнительную инжекцию основных носителей из эмиттера в базу. Инжектированные носители диффундируют через базу к коллекторному переходу и втягиваются его электрическим полем в область коллектора. Ток инжектированных носителей и, соответственно, образованный ими коллекторный ток создает еще одну компоненту фототока, которая многократно превышает первую – фототок оптически генерируемых носителей.

Общий коллекторный фототок – это сумма фототока базы $I_{\Phi Б}$ и тока инжектированных эмиттером дырок, диффундирующих и прошедших коллекторный переход $I_{КР}$. Коэффициент усиления фототока M равен:

$$M = \frac{I_{KP} + I_{\Phi B}}{I_{\Phi B}} = \beta + 1, \quad (4.16)$$

где $\beta = \frac{\alpha_0}{1 - \alpha_0}$ – статический (интегральный) коэффициент передачи базового тока в СОЭ.

Усиленный в M раз фототок создает падение напряжения на резисторе нагрузки, изменяя потенциал коллектора.

Вывод базы фототранзистора иногда используют для установки рабочей точки фототранзистора и его температурной стабилизации.

Используются также составные фототранзисторы, включенные по схеме Дарлингтона. Чувствительность составного фототранзистора на 3 – 4 порядка выше. Коэффициент M его равен $M = \beta_1 \beta_2$.

4.4. Компоненты оптико-электронных схем и средств отображения информации

4.4.1. Оптопары и микроэлектронные реле

Оптоэлектронной парой, или оптопарой, называется прибор, который содержит источник оптического излучения и фотоприемник, связанные между собой конструктивно и через оптическую среду. При этом очень важным является то, что в оптопаре, как четырехполюснике, отсутствует в идеале электрическая связь между входным и выходным портом (говорят, что вход и выход гальванически развязан). Другими словами вход электрически изолирован от выхода оптопары.

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. Внутри оптопары связь входа с выходом осуществляется с помощью оптических сигналов. Введение оптического управления внутри оптопары позволяет получить электронные устройства с исключительно своеобразными характеристиками.

Принципиальные достоинства оптопар, обусловленные использованием фотонов в качестве носителей информации, заключаются в обеспечении высокой электрической изоляции входа и выхода, односторонности потока информации, отсутствии обратной связи с выхода на вход и широкой полосе пропускания.

Основным излучателем современных оптопар является излучающий диод. В качестве оптических каналов в оптопарах обычно используют: **а)** связь через воздух; **б)** связь с использованием иммерсионной линзы; **в)** связь с использованием световода (жесткого оптического волокна). Наиболее распространенными типами фотоприемников в оптопарах являются *p – i – n фотодиод, фототранзистор и фототиристор*. Спектральная характеристика таких фотоприемников охватывает весь видимый диапазон спектра и часть ближней ИК - области.

В оптопарах используют следующие сочетания излучатель – фотоприемник: излучающий диод – фотодиод, излучающий диод – фототранзистор, излучающий диод - составной фототранзистор.

На основе оптопар выпускают твердотельные микроэлектронные реле, предназначенные для коммутации токов от 1 до 50 А при напряжении от 24 до 600 В. Они используются для подключения электродвигателей, трансформаторов, нагревательных элементов так же, как и обычные электромагнитные реле. Они управляются постоянным и переменным напряжением, имеют высокую надежность и скорость срабатывания. Такие реле совместимы по входу с интегральными микросхемами. Главные преимущества перед электромагнитными реле: гальваническая развязка входа и выхода; малые размеры; отсутствиедребезга контактов; нечувствительность к внешним полям, ударам и вибрациям; более высокая надежность и долговечность (большее время жизни).

4.4.2. Линейные и матричные датчики изображений

Одной из важных областей использования современных оптоэлектронных приборов является применение их в качестве приемных фотоэлементов, которые превращают изображение в электрический сигнал. «Картинка» с помощью объектива проецируется на фотоприемник, образованный огромным числом мельчайших «фотоэлементов». Каждый такой фотоэлемент воспринимает один «элемент» изображения и «пе-

реводит» его в заряд. Чем сильнее освещается тот или иной микроскопический фотоэлемент, чем более яркая точка картинка на него попала, тем больший заряд на нем накапливается. Минимальная единица изображения (элементарный элемент изображения), цвет и яркость которого можно различить независимо от остального изображения, называется пикселом (от англ. pixel – picture element – элемент изображения).

Для «считывания» изображения в электровакуумных телевизионных передающих трубках, давно реализован способ, когда электронный луч быстро «оббегает», как говорят, сканирует, все фотоэлементы с информацией об изображении (пикселе) и определяет («считывает») имеющийся на них заряд. Пробежав всю поверхность фотоприемника, точка за точкой, строку за строкой, считывающий электронный луч создает «видеосигнал» – развернутое во времени описание картинки.

В принципе, подобную систему можно было бы создать из отдельных полупроводниковых приборов, к примеру, фототранзисторов и отказаться от большой по габаритам и непростой по своему устройству электронно-лучевой трубки. Но как при этом произвести «считывание» картинки в таком плоском многоэлементном приборе? Как поочередно подключиться к каждому элементарному фототранзистору? При современных стандартах четкости картинки в фотоприемниках может быть миллион фотоэлементов, и очень трудно создать переключающее устройство, которое бы смогло обеспечить быстрое и без внесения «помех» в видеосигнал отключение и подключение каждого элементарного фотоэлемента. И ко всему прочему, это потребовало бы обеспечить на фотоприемнике прокладку ко всем элементам многочисленных металлических проводников.

И все же задача получения и считывания картинки изображения посредством малогабаритных полупроводниковых кристаллов была решена. Это удалось сделать с помощью так называемого прибора с зарядовой связью (ПЗС). В таких приборах рожденный падающим светом заряд накапливается в специально создаваемых внутри единого полупроводника «потенциальных ямах». Также, в ПЗС удалось объединить в одной структуре чувствительный фотоэлемент и механизм сканирования. В результате преобразователь изображения стал полностью твердотельным, малогабаритным, высоконадежным. К его положительным качествам следует отнести также низкую потребляемую мощность, от-

сутствие остаточного изображения и точную адресацию при считывании всех элементов изображения.

Упрощенная структура одного «фотоэлемента ПЗС» показана на рис. 4.14.

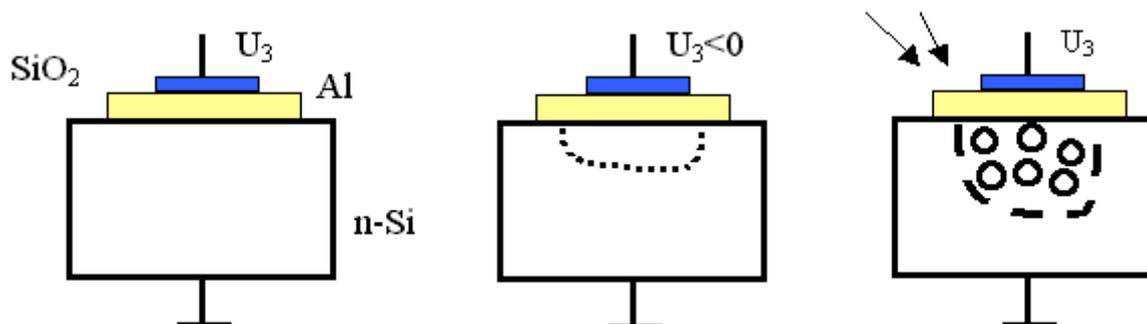


Рис. 4.14. Структура одного элемента ПЗС, формирование в нем «потенциальной ямы» и запись в него пиксела изображения

Как несложно догадаться, что один элемент ПЗС это есть не что иное, как МОП транзистор. Если на его затвор подать отрицательное напряжение относительно подложки (к металлическому электроду прикладывается потенциал одной полярности с основными носителями полупроводника), то в подзатворной области возникнет обедненный слой, в котором отсутствуют носители заряда, способные перемещаться по поверхности полупроводника. Потенциал этой области ниже остального объема подложки, поэтому говорят, что под затвором возникла «потенциальная яма». В потенциальной яме могут накапливаться неосновные носители, возникающие, к примеру, под действием света. На рисунке «потенциальная яма» показана линией, составленной из точек. Если сквозь прозрачный диэлектрик на полупроводник ячейки воздействует оптическое излучение, то вследствие внутреннего фотоэффекта в нем образуются свободные носители заряда. В рассматриваемом примере это дырки. Чем сильнее освещен данный участок кристалла, тем больше дырок возникнет в «потенциальной яме». Обычно используемый полупроводник имеет малую концентрацию собственных носителей заряда, является как бы полуизолятором, поэтому образованные фотоносители никуда не «растекаются» а, следовательно, образованный заряд может храниться в потенциальной яме относительно длительное время.

В ПЗС рассмотренный элемент «интегрирован» с другими подобными элементами за счет размещения всех фотоэлементов ПЗС на

единой полупроводниковой подложке в виде некой матрицы. При изготовлении ПЗС поверх диэлектрика прокладываются металлические электроды, соединяющие затворы с внешними выводами ПЗС. «Картинка» изображения с помощью объектива проецируется на поверхность полупроводника так, что каждый элементарный интегрированный элемент ПЗС - матрицы воспринимает один «элемент» изображения и «переводит» его в заряд. Таким образом, с помощью ПЗС удается представить двумерную оптическую информацию об изображении в виде расположенной в объеме единого полупроводникового кристалла «матрицы зарядов». Причем заряды, размещенные в потенциальных ямах, могут храниться и их величина определяется тем, насколько сильно освещается тот или микроскопический фотоэлемент ПЗС (чем более яркая точка картинки на него попала, тем больший заряд на нем накапливается).

ПЗС представляет собой, в общем случае матрицу МОП элементов, размещенных на единой подложке, и систему вывода зарядов, которая включает в себя три группы проводников и затворов. Рассмотрим принцип ее действия.

Решение задачи «считывания», полученное в виде зарядов информации об изображении (видеосигнала), осуществляется в простейшем случае с помощью последовательного вывода зарядов «строки» матрицы ПЗС. При этом перемещение зарядов «строки изображения» в ПЗС при сканировании осуществляется таким образом, что полностью разделены во времени проецирование картинки и перемещение (вывод) зарядов всех элементов (потенциальных ям).

Сначала подается отрицательное напряжение U_1 на группу затворов МОП транзисторов (показан фрагмент ПЗС, состоящий из трех, воспринимающих изображение, и шести вспомогательных элементов). Под ними, как уже отмечалось, формируются потенциальные ямы способные «принять и сохранить» фотоносители (рис. 4.15 А).

Затем проецируется изображение, в результате чего в потенциальных ямах в зависимости от интенсивности освещенности, формируется заряд. На рис. 4.15 Б показано, что освещенность первого затвора высока, третьего слабее, а второй затвор не освещен (условно отмечено количество «плюсиков»). Введенные в ямы заряды могут определенное время там храниться.

После этого, подают отрицательный потенциал на вторую группу затворов, не изменяя напряжение U_1 . При этом величина напряжения U_2 задается такой, чтобы она была больше по модулю, чем U_1 . В этом случае под второй группой затворов формируются потенциальные ямы, «глубина» которых больше. При такой конфигурации потенциалов в объеме полупроводника начнется «перетекание» заряда в «яму», глубина которой больше. Понятно, что заряды из ячейки 1 и 3 переместятся в более глубокую яму рис. 4.15 В.

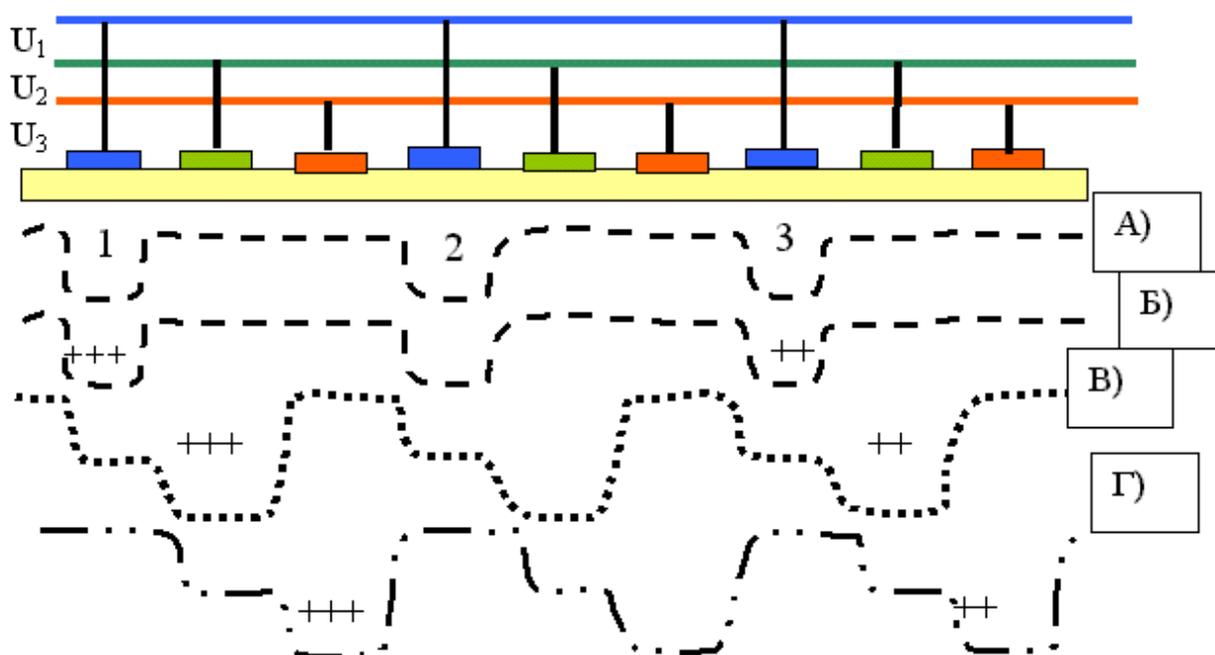


Рис. 4.15. Фрагмент простейшей ПЗС и осуществление «считывания» информации с трех элементов

Наконец, подают еще больший отрицательный потенциал на третью группу затворов и заряды, далее, «сдвигаются» вправо. При этом отключается напряжение в первой группе затворов и тем самым «ликвидируются» потенциальные ямы, способные воспринимать изображение.

Чтобы заряд с левого элемента строки матрицы ПЗС переместился вправо, осуществляют, как говорят, «трехфазное тактирование» напряжений (подается определенная комбинация импульсов напряжения отрицательной полярности на проводники U_1 , U_2 , U_3). Вывод требуемой информации о видеосигнале строки осуществляется установкой рядом с

самым правым элементом строки $p - n$ перехода, который осуществляет экстракцию последовательно поступающих в каждом такте зарядов.

ПЗС в настоящее время широко используются в качестве линейных и матричных датчиков изображений. Эти датчики используются для сканирования плоских изображений, получения неподвижных и подвижных изображений при помощи фото- и видеокамер. Регистрируемые изображения могут быть черно-белыми и цветными, и иметь разрешение, превышающее 10 тысяч элементов для линейного и 16 миллионов пикселей для прямоугольного изображения.

Контрольные вопросы

1. Какую величину называют потоком излучения и в каких единицах измеряют?
2. Какую величину называют световым потоком и в каких единицах измеряют?
3. Какую величину называют силой света источника видимого излучения в данном направлении и в каких единицах измеряют?
4. Как определена яркость источника в точке поверхности?
5. Охарактеризуйте понятия: когерентное оптическое излучение и абсолютно когерентный источник излучения.
6. Какие полупроводники называют прямозонными и непрямозонными?
7. Нарисуйте и поясните энергетические диаграммы прямозонных и непрямозонных полупроводников.
8. Какие растворы используют для увеличения эффективности излучательных процессов в непрямозонных полупроводниках?
9. Какой процесс называется внутренним фотоэлектрическим эффектом и где он используется?
10. Охарактеризуйте кратко эффект фотопроводимости и фотогальванический эффект.
11. Что называют спектральной характеристикой поглощения?
12. Нарисуйте и поясните спектральные характеристики основных полупроводников.
13. Как подразделяются по способу вывода излучения оптические излучатели?
14. Какую структуру называют одинарной гетероструктурой?

15. С какой целью в излучающих диодах используют двойную гетероструктуру?
16. Для осуществления каких функций применяются полупроводниковые излучающие приборы визуального отображения информации?
17. Охарактеризуйте понятие и назначение светоизлучающего диода.
18. Нарисуйте и поясните упрощенную конструкцию светодиода.
19. Охарактеризуйте понятие: угол обзора свечения СИД.
20. Что является основным светотехническим параметром СИД?
21. Укажите основные преимущества оптических систем связи.
22. Какие материалы используют для получения инфракрасного оптического излучения?
23. Нарисуйте энергетические диаграммы и поясните работу полупроводникового лазера в равновесном состоянии и при приложении прямого напряжения.
24. Нарисуйте и поясните спектр и диаграмму направленности излучения лазера при токах меньших и больших «порога».
25. Перечислите основные пространственно-временные параметры лазерного излучения.
26. Перечислите энергетические параметры лазерного излучения.
27. Перечислите основные достоинства полупроводниковых лазеров по сравнению с другими типами лазеров.
28. Назовите основные известные Вам применения лазеров.
29. Что такое фоторезистор? Приведите примеры.
30. Нарисуйте и поясните схему включения фоторезистора и временную зависимость выходного напряжения цепи при воздействии оптического излучения.
31. Что такое фотодиод? Поясните принцип его работы.
32. Нарисуйте и поясните структуру фотодиода, схему его включения и временную зависимость выходного напряжения цепи при воздействии оптического излучения.
33. Что такое $p-i-n$ диоды и с какой целью их используют?
34. Что такое лавинный фотодиод где и с какой целью их используют?
35. Опишите устройство, принцип работы и назначение солнечного преобразователя.
36. Что такое фототранзистор? Для чего он предназначен?
37. Что называют оптопарой? Опишите принцип действия оптопары.

5. Полупроводниковые ключи в энергетических (силовых) устройствах электроники

Мы уже изучили широко распространенные классы слаботочных полупроводниковых изделий – биполярные и полевые транзисторы. При этом транзисторы мы рассматривали как компоненты электрической цепи, работающие при токах порядка микроампер – миллиампер и напряжениях до сотни вольт.

В ряде областей, связанных, к примеру, с применением в промышленных устройствах электродвигателей переменного напряжения (380Вх3ф), бесшумных и энергосберегающих электродвигателей бытовой техники (220В), исполнительных устройств промышленной автоматики, источников бесперебойного питания и прочее, приходится «работать» с токами в единицы, сотни и даже тысячи ампер. В ряде случаев необходимы также устройства, работающие с напряжениями до 6000 В.

Область электроники, работающая с такими напряжениями и токами, получила название *энергетической или силовой электроники*.

Изучение полупроводниковых изделий силовой электроники мы начнем с изучения классических тиристоров, работающих на основе положительной обратной связи. Эти приборы в настоящее время широко применяются. Новые разновидности тиристоров позволяют достигнуть более высоких показателей и, отчасти, устранить недостатки, присущие традиционным тиристорам.

Материал по мощным полевым транзисторам не излагается, поскольку анализ их работы был уже частично проведен. Рассмотрению особенностей исполнения, построению цепей управления ДМОП, VМОП мощных полевых транзисторов мы уделим внимание по мере надобности, в дальнейшем. Будет изучен лишь мощный транзистор со статической индукцией, имеющий ряд особенностей.

Особый интерес представляет материал по биполярным транзисторам с изолированным затвором. Эти приборы имеют большое входное сопротивление, как у полевых транзисторов, и низкое падение напряжения в открытом состоянии, как у биполярных (обычно меньше 2 В). Такие транзисторы почти идеальны в качестве высоковольтных ключей в устройствах, где существенным является низкое падение напряжения в открытом состоянии, таких как двигатели постоянного и переменного

тока, источники питания для соленоидов, реле пускателей и т. д. При этом они температурно-устойчивы и могут работать в температурном диапазоне от – 40 до 150⁰С.

В конце мы кратко ознакомимся с терминологией, используемой при управлении силовыми полупроводниковыми ключами.

5.1. Общие сведения о применении полупроводниковых ключевых приборов в энергетических (силовых) устройствах электроники

Энергетической (силовой) полупроводниковой электроникой, принято называть область электроники, когда электронные устройства работают при очень больших токах (< 3000 А), напряжениях (< 6000 В) и мощностях (< 10М Вт). На практике с такими предельными электрическими режимами работы электронных устройств, с мегаватными мощностями приходится иметь дело, как правило, в сверхмощных системах промышленности. К силовой электронике, вместе с тем, относят также многие электротехнические и бытовые устройства, имеющие токи до 100 А и напряжения до 600 В. В частности, подобного рода устройства находят применение при управлении различного рода электродвигателями (бытовая и промышленная техника, электроинструмент, электрифицированный транспорт, электропривод и т. п.), в преобразователях электрической энергии в тепловую (электронагревательные установки, различного рода печи, электросварка и прочее), а также в современных вторичных источниках питания, в светотехнике, в том числе, в современных лампах дневного света.

Основная задача полупроводникового прибора в энергетической электронике состоит в том, чтобы без искрений, «подгорания», «дребезга» контактов, что характерно для механических контактных элементов, отключать или подключать за требуемое время (коммутировать) различные электрические устройства, замыкать или размыкать какую-то часть электрической схемы. Поэтому, с точки зрения функционального назначения, полупроводниковые приборы энергетической электроники являются, по своей сути, электрическими ключами. Напомним, что электрический ключ – это двухполюсник, осуществляющий под воздействием управляющего сигнала мгновенный переход от состояния разрыва цепи

к состоянию, когда его выводы замкнуты накоротко, к короткому замыканию (КЗ). При коммутации ключа его сопротивление в идеале меняется с нуля на бесконечность или наоборот.

Переход полупроводникового энергетического ключа из закрытого состояния в открытое и наоборот сопровождается изменением проводящих свойств его структуры. Открытое состояние ключа определяется наличием в структуре прибора проводящего канала с высокой концентрацией подвижных носителей заряда, а значит, низким сопротивлением протекающему току. Чтобы «разомкнуть» ключ необходимо обеспечить условия прерывания движения носителей в проводящем канале. Это достигается за счет «формирования» области «обедненной» носителями, называемой областью пространственного (объемного) заряда, в которой практически отсутствуют «свободные» электроны и дырки. Такой механизм переключения свойственен всем типам ключевых энергетических полупроводниковых приборов. Различны только физические эффекты, лежащие в основе принципов переключения.

Из-за предельных режимов его работы, ключ силовой электроники должен обладать почти идеальными свойствами. Он должен мгновенно, при нулевой мощности управления, переключать бесконечно большие токи, блокировать подачу бесконечно больших потенциалов на какие – то точки электрической цепи. Практические ключи могут лишь в той или иной степени приближаться к «идеальным».

Все более ужесточающиеся требования по экономии электроэнергии, повышению КПД электротехнических и бытовых устройств, уменьшению их размеров, веса, стоимости выдвинули задачи по созданию высокоскоростных силовых ключей с минимальными потерями в широком диапазоне мощностей. Дополнительные требования к характеристикам управляемости ключа накладывает применение практически во всех системах управления электротехническими и электробытовыми устройствами микроконтроллеров (ЭВМ).

Требования потребителей в отношении коммутируемых токов, блокируемых напряжений, частоты и времени переключения энергетического ключа весьма разнообразны и широки, поэтому не существует единого типа полупроводникового ключа универсального применения. По этой причине мы в дальнейшем рассмотрим полупроводниковые ключевые приборы, получившие наибольшее распространение в тех энергетиче-

ских (силовых) устройствах электроники, с которыми вам придется иметь дело.

5.2. Тиристоры

5.2.1. Устройство и принцип действия диодного тиристора

Тиристором называют полупроводниковый прибор с двумя устойчивыми состояниями, состоящий из трех и более электронно-дырочных структур, который имеет на вольт-амперной характеристике участок с отрицательным дифференциальным сопротивлением и который может переключаться из закрытого состояния в открытое и наоборот.

Слово тиристор (thyristor) происходит от греческого *thura* – дверь (в смысле – имеющая закрытое или открытое состояние), а также от английского *resistor* – сопротивление. В целом, оно отражает тот факт, что переход тиристора из закрытого состояния в открытое и наоборот сопровождается изменением проводящих свойств его структуры.

Все тиристоры отличаются друг от друга количеством внешних выводов. В этой связи различают:

- а) диодные тиристоры (динисторы), имеющие два вывода – анод и катод;
- б) триодные тиристоры, имеющие три вывода – анод, катод, и управляющий электрод.

Структура простейшего диодного тиристора (динистора) показана на рис. 5.1.

Диодный тиристор представляет собой структуру, состоящую из четырех чередующихся полупроводниковых областей дырочного и электронного типов проводимости с различной степенью легирования примесями. В равновесном состоянии (без приложения внешнего напряжения), согласно п. 1.3.1 раздела 1 в тиристоре образуются три области (слоя) объемного заряда, которые обозначены на рис. 5.1, как области П1, П2, П3.

Если к тиристору приложить обратное напряжение (к аноду – отрицательный потенциал, к катоду – положительный), то тиристор будет пребывать в закрытом, непроводящем состоянии.

Согласно правилу простейшего делителя напряжения, напряжение между катодом и анодом, должно, в зависимости от сопротивлений переходов, поделиться между тремя переходами. Но при этом, как оказывается, все внешнее напряжение будет приложено к обладающим большим сопротивлением переходам П1 и П3. Действительно, при таком напряжении переходы П1 и П3 окажутся при обратном включении и, следовательно, будут иметь очень большое сопротивление. Переход П2 включен при прямом включении, но напряжение на нем будет очень мало, по крайней мере, значительно меньше порогового, поэтому инжекции носителей не будет и сопротивление П2 будет также большим. Следовательно, в структуре прибора отсутствует проводящий канал и ток от катода к аноду через них протекать не может.

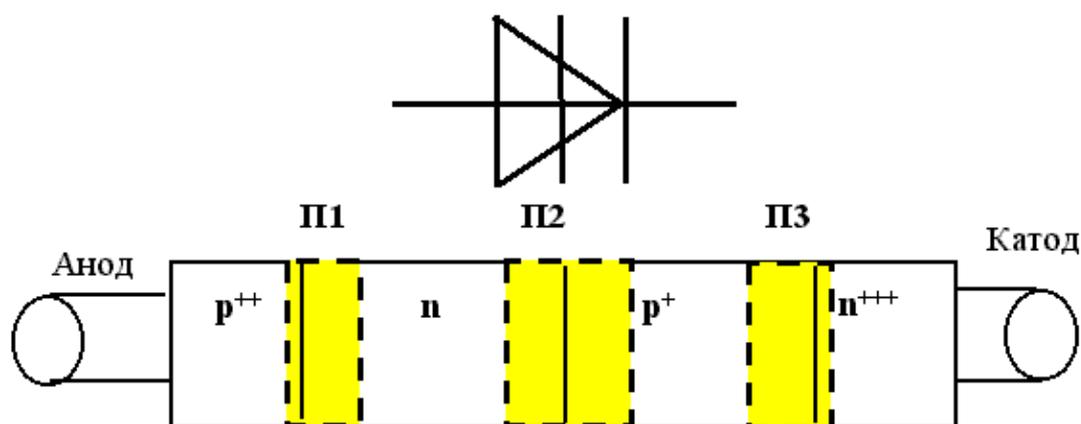


Рис. 5.1. Структура и условное обозначение на схемах диодного тиристора (динистора)

Итак, сопротивление тиристора при обратном включении (между катодом и анодом) определяется сопротивлениями обратно включенных переходов П1 и П3, а поскольку они велики, то через тиристор протекает малый обратный ток. Тиристор при обратном включении он подобен разрыву цепи.

Рассмотрим теперь, что будет происходить в структуре, если к тиристорному приложить прямое напряжение (к аноду – положительный потенциал, к катоду – отрицательный). При этом возможны два случая, когда прямое напряжение меньше и, наоборот, больше некоторой величины, называемой напряжением включения $U_{ВКЛ}$.

При прямом включении тиристора напряжение источника питания перераспределяется таким образом, что переходы П1 и П3 окажутся при прямом включении, а переход П2 – в обратном. При этом большая часть внешнего напряжения будет прикладываться к переходу П2, и лишь незначительная его часть будет падать на переходах П1 и П3. Этого напряжения, с одной стороны, не достаточно, чтобы в переходах П1 и П3 началась инжекция неосновных носителей заряда, поэтому эти переходы обладают большим сопротивлением. С другой стороны, при напряжениях между анодом и катодом $U_{AK} < U_{ВКЛ}$ на переходе П2 напряжение недостаточно, чтобы в нем начался лавинный пробой. Следовательно, в структуре прибора отсутствует проводящий канал и ток от анода к катоду протекать не может. Тиристор пребывает в закрытом состоянии. При напряжениях между анодом и катодом $U_{AK} \sim U_{ВКЛ}$ в обратном включенном переходе П2 начинается явление лавинного пробоя (п.1.5.2). Это, с одной стороны, увеличивает ток через переход П2 и, соответственно, уменьшает его сопротивление. А с другой стороны, согласно правилу простейшего делителя напряжения, это влечет за собой рост напряжений на переходах П1 и П3. В свою очередь, это усиливает процессы инжекции на П1 и П3 и способствует дальнейшему увеличению коэффициента размножения носителей в П2. В результате ток через тиристор будет быстро нарастать, а сопротивление всех его трех переходов будет уменьшаться.

Таким образом, при прямых напряжениях между анодом и катодом $U_{AK} \sim U_{ВКЛ}$ в тиристоре возникает проводящий канал с высокой концентрацией подвижных носителей заряда, а значит, низким сопротивлением протекающему току. Через тиристор в этом случае может протекать большой ток при малом падении напряжения между анодом и катодом.

Описанный процесс переключения тиристора из закрытого в проводящее состояние является весьма упрощенным. Он был необходим для понимания идей, положенных в основу функционирования прибора. Поэтому для более тонкого анализа происходящих в тиристорной структуре процессов сделаем следующее.

Мысленно проведем разрез структуры, как показано на рис. 5.2 линией, составленной из точек, и представим четырехслойную $p-n-p-n$ –структуру в виде комбинации двух $p-n-p$ и $n-p-n$ транзисторов. При этом коллектор одного транзистора будет являться базой другого и

наоборот. Будем полагать, что усилительные свойства транзисторов характеризуются коэффициентами передачи тока эмиттера в СОБ α_p и α_n или коэффициентами передачи тока базы в СОЭ β_p и β_n . Переходы П1 и П3 находятся под прямым напряжением – это эмиттерные переходы транзисторов; к переходу П2 приложено обратное напряжение – это коллекторный переход для обоих транзисторов.

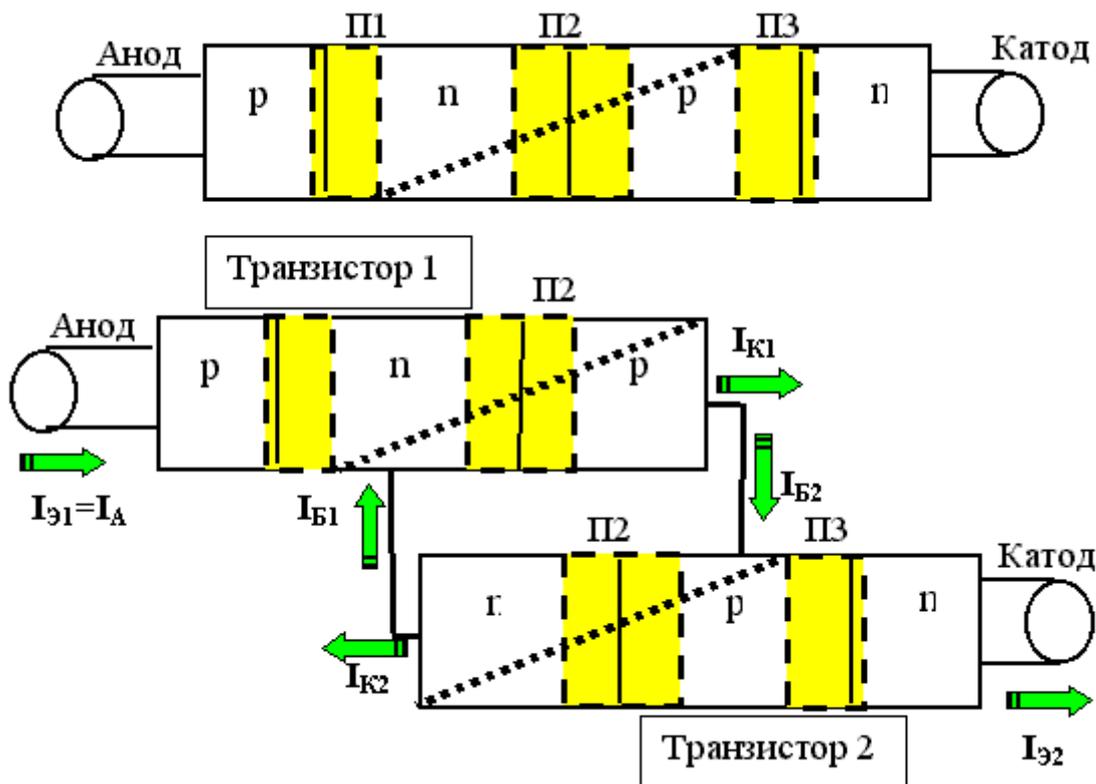


Рис. 5.2. Структура динистора и представление диодного тиристора в виде комбинации двух $p-n-p$ и $n-p-n$ транзисторов

Если напряженность электрического поля в коллекторном переходе транзистора достаточна для размножения носителей заряда, то для этого случая основное уравнение биполярного транзистора, работающего в активном режиме, в схеме с общей базой можно записать, введя поправочный коэффициент размножения носителей M , учитывающий степень размножения носителей заряда в коллекторном переходе. Тогда для первого $p-n-p$ транзистора можно записать:

$$I_{K1} = M_1 \alpha_p I_{\mathcal{E}1} + M_1 I_{KB01}. \quad (5.1)$$

Аналогично для второго $n - p - n$ транзистора:

$$I_{K2} = M_2 \alpha_n I_{\mathcal{E}2} + M_2 I_{KB02}. \quad (5.2)$$

Если предположить для простоты, что $M_1 = M_2 = M$, $I_{\mathcal{E}1} = I_A = I_K = I_{\mathcal{E}2} = I$, $I_{K1} = I_{B2} = I_{\mathcal{E}2} - I_{K2}$, то можно получить следующие соотношения:

$$M \alpha_p I + M I_{KB01} = I - I M \alpha_n - M I_{KB02} \Leftrightarrow$$

$$I = \frac{M(I_{KB01} + I_{KB02})}{1 - M(\alpha_p + \alpha_n)}. \quad (5.3)$$

Известно, что коэффициенты M , α_p и α_n , входящие в (5.3), зависят от напряжений приложенных к переходам транзистора. Вид этих зависимостей показан на графиках рис. 5.3.

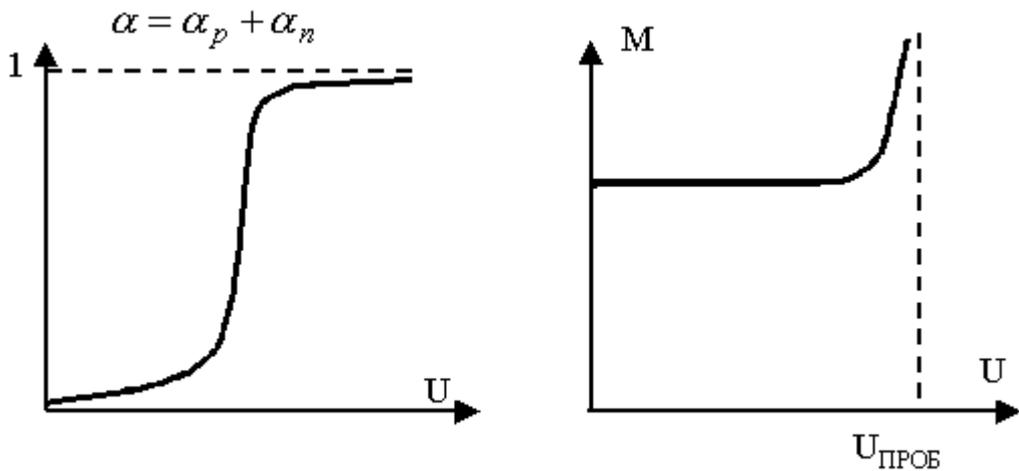


Рис. 5.3. Графики зависимости коэффициентов передачи тока эмиттера и коэффициента размножения носителей заряда

Как следует из (5.3) при малых напряжениях на аноде относительно катода напряжения на переходах транзисторов не велики, поэтому

$M=1$, а $\alpha_p + \alpha_n \approx 0$ и, следовательно, ток через тиристор равен $I = I_{KB01} + I_{KB02}$. Тиристор пребывает в закрытом состоянии.

С повышением разности потенциалов на аноде относительно катода тиристора напряжение на эмиттерных переходах увеличивается. Соответственно, усиливается инжекция основных носителей заряда и коэффициенты α_p, α_n увеличиваются ($\alpha_p + \alpha_n \Rightarrow \uparrow$). Электроны, инжектированные из n^{+++} эмиттера в p^+ область, втягиваются электрическим полем перехода П2 и попадают в n - область (n – базу первого транзистора). Там они создают неравновесный заряд, который снижает потенциал n – базы, что, в свою очередь, повышает разность потенциалов на эмиттерном переходе первого транзистора и увеличивает инжекцию дырок на нем ($\alpha_p \Rightarrow \uparrow$). Дырки, инжектированные переходом П1, через n область, посредством электрического поля П2 попадают в p – базу второго транзистора. Там они создают неравновесный заряд, который увеличивает потенциал p – базы, что, в свою очередь, повышает разность потенциалов на эмиттерном переходе П3 ($\alpha_n \Rightarrow \uparrow$). Таким образом в тиристорной структуре возникает положительная обратная связь, которая приводит к самопроизвольному лавинообразному увеличению тока через тиристор.

Как только член $M(\alpha_p + \alpha_n)$ в уравнении (5.3) приблизится к единице, произойдет мгновенный переход тиристора от состояния, когда между анодом и катодом был практически разрыв цепи, к состоянию, когда выводы тиристора замкнуты накоротко.

Можно найти, что напряжение включения тиристора равно:

$$U_{ВКЛ} = U_{ПРОБ} (1 - \alpha_p - \alpha_n)^{0.25}. \quad (5.4)$$

Из выражения (5.4) видно что напряжение включения всегда меньше напряжения пробоя перехода П2, что обусловлено наличием положительной обратной связи в тиристорной структуре.

После того, как тиристор включился и все три перехода смещены в прямом направлении, оба эквивалентных транзистора переходят в режим глубокого насыщения.

График вольт-амперной характеристики (ВАХ) динистора имеет вид, показанный на рис. 5.4. Там же показаны линия нагрузки, схема включения динистора в электрическую цепь, а также эквивалентная схема, замещающая тиристор, в виде ключа.

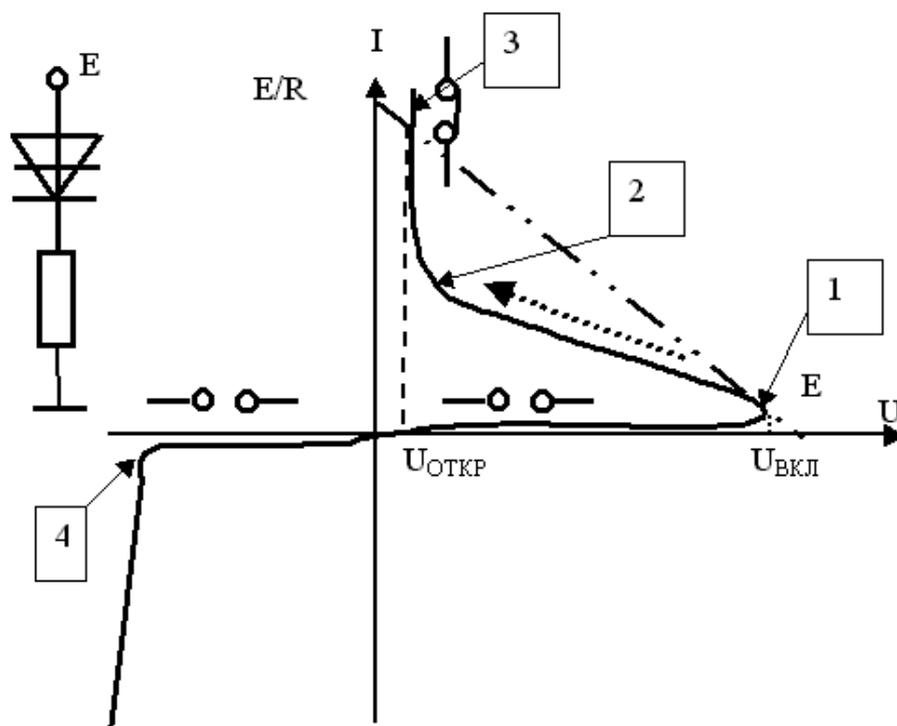


Рис. 5.4. График вольт-амперной характеристики диодного тиристора и эквивалентные схемы, характеризующие состояние динистора

В зависимости от напряжения на аноде и тока, протекающего через диодный тиристор, можно выделить несколько режимов его работы.

Режим 1 (область ВАХ 0-1). Напряжение на аноде относительно катода положительно, но меньше напряжения включения $U_{ВКЛ}$, ток через тиристор незначителен (несколько микроампер). Этот участок ВАХ соответствует режиму прямого запираания, что отражает эквивалентная схема в виде разрыва цепи между анодом и катодом.

Режим 2 (область ВАХ 1-2). Как только напряжение на аноде относительно катода превысит напряжение включения $U_{ВКЛ}$, тиристор очень быстро «проходит» участок характеристики с отрицательным дифференциальным сопротивлением от точки 1 к точке 2. Точка 1, соответствующая $dU/dI = 0$, характеризуется напряжением включения $U_{ВКЛ}$ и током включения $I_{ВКЛ}$.

Режим 3 (область ВАХ 2-3). На этом участке работа тиристора соответствует режиму прямой проводимости, что отражено эквивалентной схемой в виде короткозамкнутой перемычки. Напряжение в точке 3 называется напряжением тиристора в открытом состоянии $U_{ОТКР}$ или напряжением удержания. В этом режиме через тиристор протекает ток удержания.

Режим 4 (область ВАХ 0-4). Напряжение на аноде относительно катода отрицательно, ток через тиристор незначителен (несколько микроампер). Этот участок ВАХ соответствует режиму обратного запираения, что отражает эквивалентная схема в виде разрыва цепи между анодом и катодом.

5.2.2. Устройство и принцип действия триодного тиристора

Для того, чтобы можно было управлять процессом включения тиристора с помощью отдельного специального вывода, используют триодные тиристоры, имеющие три вывода – анод, катод, и управляющий электрод. Для этой цели к «базам» эквивалентных транзисторов подключают металлические проводники, которые и образуют управляющие электроды (рис. 5.5.).

Управление включением триодного тиристора может осуществляться с помощью управляющего электрода (УЭ) либо «по аноду», либо по «катоду». В первом случае в n – базу подается отрицательный ток управления; во втором – в p – базу положительный ток управления.

Можно показать, что если в цепи управляющего электрода протекает ток, то выражение (5.3) будет иметь вид:

а) при управлении по цепи анода:

$$I = \frac{I_{УЭ}\alpha_n + M(I_{КБ01} + I_{КБ02})}{1 - M(\alpha_p + \alpha_n)}; \quad (5.5)$$

б) при управлении по цепи катода:

$$I = \frac{I_{УЭ}\alpha_p + M(I_{КБ01} + I_{КБ02})}{1 - M(\alpha_p + \alpha_n)}. \quad (5.6)$$

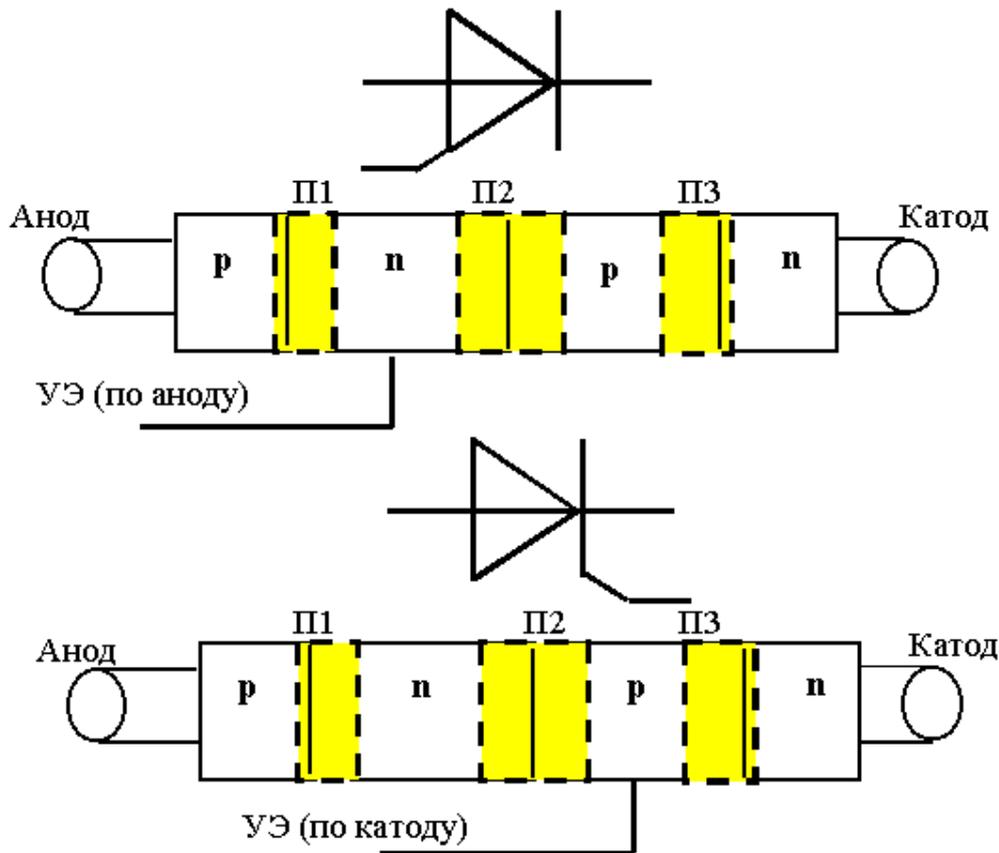


Рис. 5.5. Структуры и условное обозначение на схемах триодных тиристоров с управлением по аноду и катоду

Если в цепи управляющего электрода протекает ток $I_{УЭ}$, то это приводит к увеличению инжекции через переход (либо через П1, либо через П3, соответственно). Это вызывает увеличение $\alpha_n \Rightarrow \uparrow$ или $\alpha_p \Rightarrow \uparrow$, соответственно. Таким образом, ток анода достигает значения, при котором выполняется условие включения тиристора $M(\alpha_p + \alpha_n) \approx 1$ при меньшем анодном напряжении. ВАХ триодного тиристора будет иметь вид как на рис 5.6.

Для подачи требуемого тока в цепь управляющего электрода создается специальная маломощная электрическая цепь управления. В большинстве случаев тиристор отпирается прямоугольным импульсом тока малой длительности. Требования стабильности, взаимозаменяемости и надежности тиристорных схем приводят к необходимости использования источников питания, работающих в режиме источника тока.

В этом случае форма и величина тока управления практически не зависят от величины и характера входного сопротивления тиристора. При этом принимаются специальные меры по обеспечению устойчивости (помехозащищенности) цепи управления к воздействию «паразитных» импульсов управления, обусловленных внешними электромагнитными помехами («наводками») или внутренними процессами, например, переключениями других тиристорных схем.

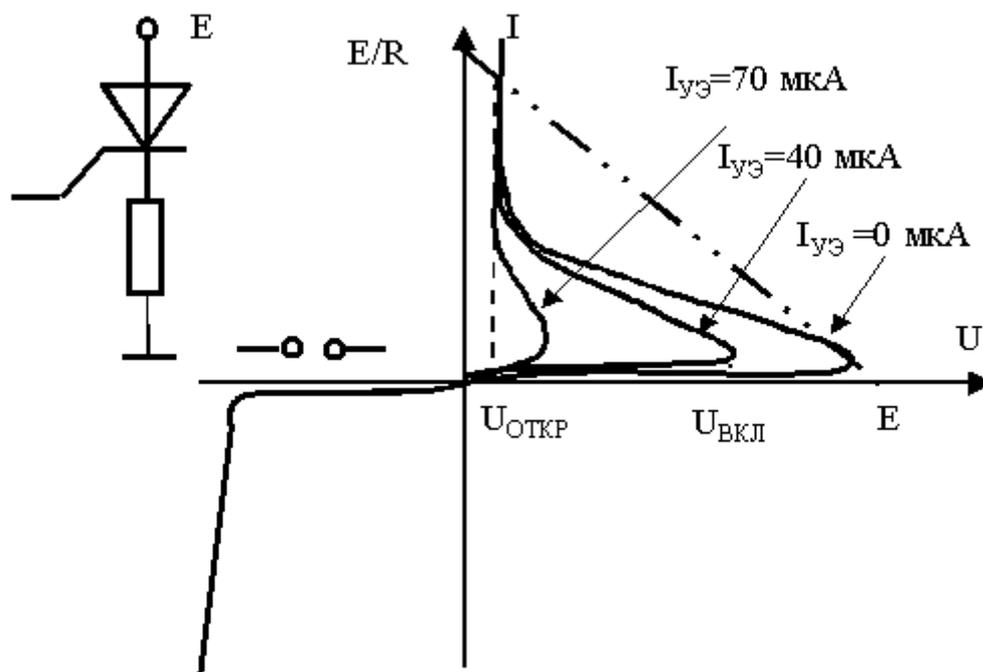


Рис. 5.6. Семейство вольт-амперных характеристик триодного тиристора при различных токах управляющего электрода

Обратим также внимание на те следствия, которые обуславливает вид ВАХ триодного тиристора.

1. Включить триодный тиристор, т. е. перевести его из режима прямого запирающего (~разрыва цепи между анодом и катодом) в режим прямой проводимости (~короткого замыкания между анодом и катодом) можно, используя маломощные импульсы тока, не связанные с «силовой» цепью.

2. Выключить же триодный тиристор, т. е. перевести его в режим запирающего (~разрыва цепи между анодом и катодом) можно, опять же, только уменьшив напряжение в «силовой» сети до $U = 0$.

Если выключение тиристора обеспечивается по выходной анодной цепи («силовой» сети), то такие тиристоры называются *незапираемыми*. Их также часто называют однооперационными тиристорами. Существуют также так называемые *запираемые тиристоры*, которые могут переключаться из закрытого состояния в открытое и наоборот при подаче на управляющий электрод сигналов соответствующей полярности. Их также называют двухоперационными (выключаемыми) тиристорами. Запираемый тиристор принципиально и конструктивно подобен обычному незапираемому тиристор. Для выключения в цепь управляющего электрода подается длительный импульс тока полярности, противоположной импульсу включения. При достаточной амплитуде и длительности импульса управления один из эквивалентных «транзисторов» выводится из режима глубокого насыщения в активный режим, в результате чего уменьшается концентрация носителей в базе и падение напряжения на коллекторном переходе увеличивается. Вслед за первым выходит из насыщения и второй транзистор. Это вызывает прекращение действия положительной обратной связи и тиристор переходит в режим прямого запираания.

5.2.3. Устройство и принцип действия симметричного тиристора (симистора)

Рассмотренные выше диодные и триодные тиристоры предназначены для работы в цепях постоянного или знакопостоянного (пульсирующего положительной полярности) тока. На практике весьма часто приходится коммутировать цепи, по которым протекает переменный ток. Часто требуется, чтобы включение тиристора проводилось путем подачи на его электроды импульсов тока любой полярности. Эти задачи позволяют решить симисторы.

Симистор (TRIAC) – переключающий прибор, работающий, как при положительном, так и при отрицательном напряжении на аноде относительно катода, который, при подаче сигнала на его управляющий электрод, включается как в прямом, так и в обратном направлении.

Симистор состоит из пяти чередующихся областей и по сути, представляет собой два встречно-параллельно соединенных триодных ти-

ристора. На рис. 5.7. изображена типичная структура симмистора и условное изображение его на схемах.

ВАХ симистора имеет вид, показанный на рис. 5.8.

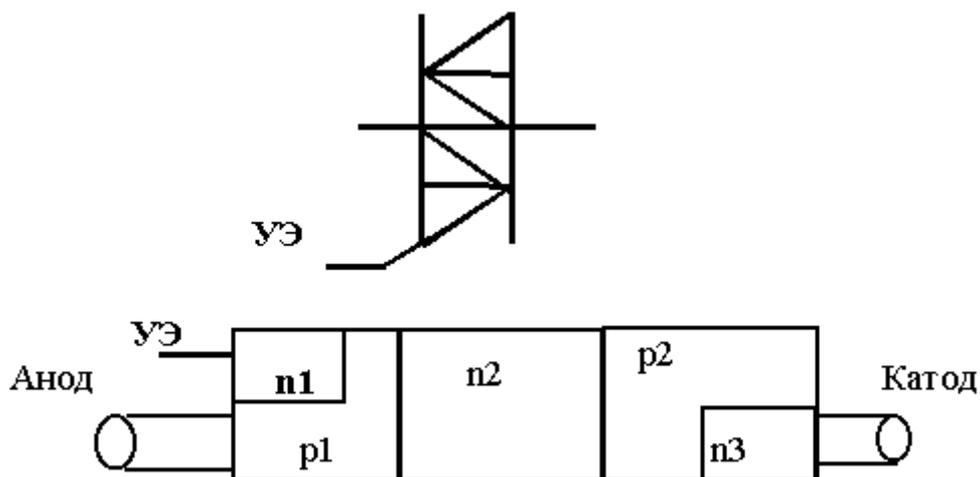


Рис. 5.7. Структура и условное обозначение на схемах симистора

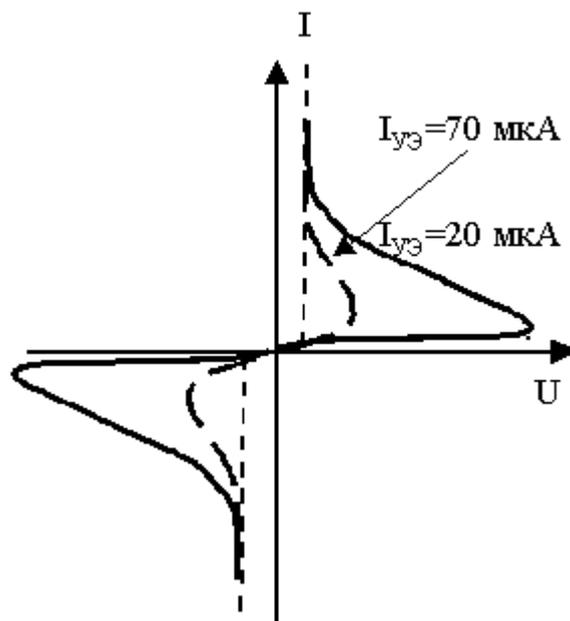


Рис. 5.8. Семейство вольт-амперных характеристик симистора при различных токах управляющего электрода

Симисторы, благодаря универсальности и простоте, широко используются в электротехнических и бытовых устройствах, для подключения их к электросети и регулирования подаваемой на устройства амплитуды напряжения силовой сети. В частности, симисторы применя-

ются для питания различного рода электрических нагревателей. При этом мощность в нагрузке, а следовательно и степень нагрева, регулируется путем изменения отношения длительности замкнутого состояния ключа (симистор открыт) к длительности его разомкнутого состояния (симистор закрыт). Пример подобного использования симистора показан на рис. 5.9.

Если мощный электронагреватель подключать непосредственно к электрической сети, то в момент включения вилки в розетку, из-за большого начального (как говорят, пускового) тока, будет возникать «искрение», «расплавление и выгорание» отдельных частей контактов. Контакты такого разъемного соединителя очень быстро выйдут из строя и потребуют замены. Из-за искрений во время коммутации возникнут различного рода помехи. Кроме того, из-за малого сопротивления «холодного» нагревателя, в момент включения уменьшится напряжение в сети и другие потребители электроэнергии «ощутят» ваши действия, к примеру, потому, что у них примерно на пол-секунды снизится яркость свечения лампочек.

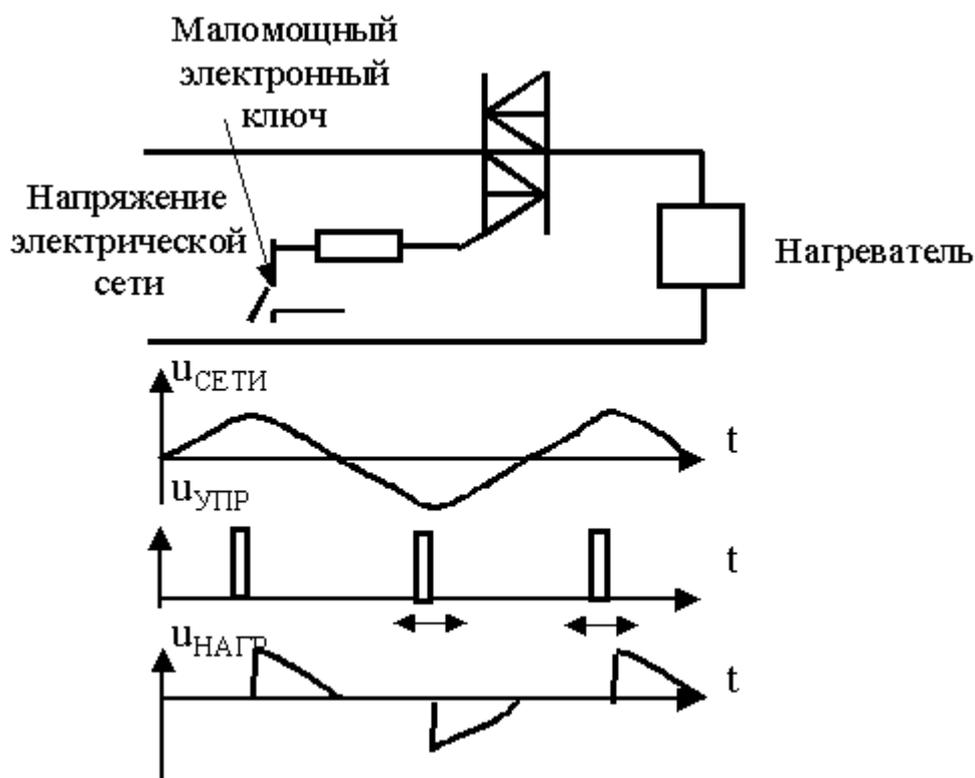


Рис. 5.9. Пример использования симистора для подключения мощного электронагревателя к электрической сети и регулирования температуры нагрева

Использование симистора позволяет подключаться к сети (коммутировать большую мощность) посредством маломощного электронного ключа, у которого, как известно, отсутствует искрение и подгорание контактов. Симистор открывается импульсами напряжения, подаваемыми на управляющий электрод в момент замыкания контактов маломощного электронного ключа. В дальнейшем мы подробно познакомимся с реализацией электронных ключей и увидим, что в них относительно просто можно менять положение (время начала) импульса.

Из физики известно, что «холодный» нагреватель имеет очень малое сопротивление и в момент включения будет иметь большой пусковой ток. Если импульсы управления сначала подавать в момент перехода напряжения через «нуль» (осуществить синхронную коммутацию или коммутацию при нулевом напряжении), то пусковой ток окажется малым. Увеличивая на небольшую величину ток в нагрузке, можно постепенно повышать температуру нагревателя (тем самым увеличивать его сопротивление) и симистор «плавно» (без резких «бросков» тока) подключит нагреватель к электрической сети. Уменьшения напряжения сети не произойдет и электрическая сеть «не почувствует» подключения мощной нагрузки. То есть, с симистором относительно просто реализовать очень важный принцип «плавного» включения нагрузки – синхронную коммутацию или коммутацию при нулевом напряжении.

Регулировка мощности, подаваемой в нагрузку, осуществляется путем «перемещения» импульса включения тиристора по оси времени. При этом меняется начальная фаза (фазовый угол) напряжения, при котором отпирается тиристор и, следовательно, среднеквадратическое (действующее) значение тока в нагрузке. Такой способ регулирования напряжения и тока в нагрузке часто называют фазовым управлением.

5.2.4. Разновидности мощных тиристорov

В настоящее время применяются различные по технологии исполнения, конструкции и особенностям работы мощные тиристоры, коммутирующие потоки электрической энергии.

Однооперационные тиристоры (SCR- Silicon Controlled Rectifier – кремниевый управляемый переключатель) в настоящее время изготов-

ливают на основе сверхвысоковольтной технологии и имеют мощностные параметры 8000 В и 3600 А для линий передач постоянного тока и 12000В и 1500А для переключателей переменного тока. Применяют также мощные однооперационные тиристоры (SCR), в которых ключ управляется оптическим излучением. Для переключения структуры тиристора из закрытого в открытое состояние используется световой сигнал, передаваемый по световолоконному кабелю. При создании мощных тиристоров большие успехи достигнуты также при создании запираемых *двухоперационных тиристоров* (GTO – Gate Turn Off Thyristor). С их помощью удастся осуществлять коммутацию токов до 3000А и напряжений до 6000 В. Конкуренцию запираемым тиристорам составляют *индукционные тиристоры* (SiTh – Static Induction Thyristor). Важным преимуществом индукционных тиристоров перед GTO является отсутствие в них внутренней положительной обратной связи. Это существенно повышает рабочую температуру ключа (до 200⁰ С). Такие тиристоры сочетают высокую перегрузочную способность по току с быстрым протеканием процессов переключения. Параметры такого ключа по предельным токам и напряжениям составляют 1500 А и 2500В. К недостаткам такого тиристора относится его нормальная открытость, т. е. способность проводить ток при отсутствии управляющего воздействия. Это требует более сложной схемы управления. Интенсивно развиваются *полевые тиристоры с изолированным затвором МСТ* (MOS Composite SiTh, т. е. комбинированная транзисторно-тиристорная структура). МСТ, в первом приближении, представляют собой монолитную полупроводниковую структуру индукционного тиристора (SiTh) с интегрированным в нее управляющим МДП полевым транзистором. В ней используется принцип катодного управления тиристором. В катодную цепь SiTh тиристора последовательно включается низковольтный, обладающий малым сопротивлением канала МДП транзистор, который по своим электрическим характеристикам способен пропускать полный ток SiTh ключа. Этот новый класс силовых полупроводниковых приборов управляется по изолированному поликремниевому затвору; он является двухоперационным ключом, выключение которого осуществляется с помощью изолированного затвора. В настоящее время существуют различные подклассы таких МСТ приборов: *p*- и *n*- канальные, с одно – или двухсторонним затвором управления, с различными способами включения.

5.3. Базовые транзисторные структуры силовых ключей

5.3.1. Общие сведения о базовых транзисторных структурах силовых ключей

В настоящее время среди мощных транзисторов полевые существенно «потеснили» биполярные транзисторы. Это связано с целым рядом преимуществ полевых транзисторов над биполярными:

А) коэффициент усиления биполярных транзисторов из-за снижения коэффициента инжекции резко уменьшается в области больших токов. Крутизна полевых транзисторов с коротким каналом остается практически неизменной в области рабочих токов. Это обеспечивает большую устойчивость полевых транзисторов к токовым перегрузкам по сравнению с биполярными.

Б) Значительные токи утечки у мощных биполярных транзисторов уменьшают область безопасной работы ключа. Кроме того, они требуют большой мощности управления. Наоборот, высокое входное сопротивление полевых транзисторов позволяет существенно снизить мощность управления.

В) Отсутствие накопленного заряда неосновных носителей обеспечивает высокую скорость переключения полевых транзисторов по сравнению с биполярными.

Г) Еще одно преимущество полевых транзисторов связано с однородностью их полупроводниковой структуры. Однородный полупроводник обладает положительным температурным коэффициентом сопротивления. По этой причине полевой транзистор обладает большей помехоустойчивостью и способностью к самовыравниванию плотностей протекающего тока. В биполярных структурах, в противоположность этому, присутствует эффект локального перегрева.

К сказанному также необходимо добавить следующее. Во всех силовых ключах на передний план выступают энергетические показатели. Большую роль, в частности, играют потери мощности в транзисторе и, прежде всего, потери в открытом его состоянии (замкнутом состоянии ключа) $P_{ПОТ} = I^2 r_{КАН}$. Значение тока I задается внешней нагрузкой и

для снижения мощности потерь $P_{ПOT}$ необходимо обеспечить минимальное значение сопротивления открытого транзистора, в частности, для полевого транзистора сопротивления канала $r_{КАН}$.

Сопротивление канала можно снизить за счет уменьшения его длины. Это возможно достигнуть, если от «горизонтального» расположения канала перейти к «вертикальному», т.е. направленному перпендикулярно поверхности, вглубь кремниевой пластины.

Снизить сопротивления канала $r_{КАН}$ можно также за счет того, что вместо одного канала использовать в монокристаллической полупроводниковой пластине большое число (более тысячи) одинаковых, коротких, плотно расположенных друг возле друга, вертикальных, параллельных каналов (одинаковых ячеек).

Наконец, уменьшить сопротивление канала можно, если в него вводить (инжектировать) дополнительные носители заряда, которые бы увеличивали проводимость канала.

Опираясь на технологические методы обработки кремниевых пластин, используя субмикронные технологии (нанотехнологии), которые успешно применялись при изготовлении низковольтных цифровых интегральных схем, разработчики силовых полупроводниковых приборов решили ряд задач по снижению сопротивления канала. Во-первых, им удалось создать сложные, но стабильно работающие структуры (элементарные ячейки) с однородными свойствами. Во-вторых, дальнейшее совершенствование технологии позволило осуществить и приблизить по стоимости к обычным транзисторам монокристаллические многоячейковые структуры, а также создать полностью управляемые гибридные модули, работающие в мегаватном диапазоне мощностей. Наконец, была создана комбинированная структура, позволившая объединить положительные качества как биполярного, так и полевого МДП - транзистора, и названная как биполярный транзистор с изолированным затвором.

5.3.2. Полевой транзистор со статической индукцией

Статическим индукционным транзистором (СИТ - SIT - Static Induction Transistor) называется полевой транзистор, структура которого имеет множество параллельных каналов с затворами малой длины (ячеек), встроенных в монокристаллический полупроводник.

Упрощенное строение элементарной структурной ячейки СИТ транзистора показано на рис. 5.10 а.

Элементарная ячейка СИТ транзистора имеет короткий вертикальный канал, расположенный между затворами близкими по форме к цилиндру, малые размеры области затвора и характеризуется малым расстоянием от истока до затвора. При отсутствии напряжения на затворе относительно истока и наличии положительного напряжения сток – исток, через канал элементарной ячейки протекает ток. Толщина области n^- определяет пробивное напряжение транзистора. Реальный СИТ транзистор представляет собой параллельное соединение большого числа отдельных ячеек. Интегральная технология позволяет довольно плотно располагать эти элементарные ячейки друг возле друга, создавая возможность протекания тока стока параллельно по многим каналам.

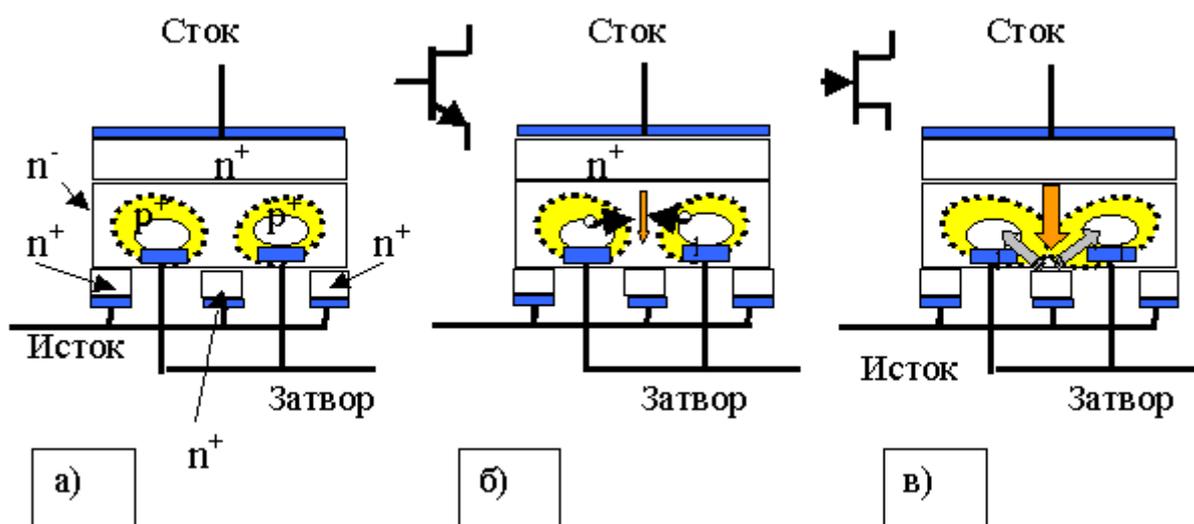


Рис. 5.10. Структура элементарной ячейки и условное обозначение на схемах статического индукционного транзистора в различных режимах его работы

СИТ транзистор, в общем случае, может функционировать в двух режимах: в биполярном, как ключ, и в полевом, как усилительный прибор для нагрузок малой величины.

Сопротивление канала элементарной ячейки при подаче напряжения сток – исток определяется омическим сопротивлением слаболегированной области канала n^- . Чтобы уменьшить сопротивление канала

элементарной ячейки, а, соответственно, сопротивление сток – исток СИТ (сопротивление ключа в замкнутом состоянии) на затвор относительно истока подают положительное напряжение. При этом из p – областей в n^- область канала возникает инжекция неосновных носителей заряда (рис. 5.10 б), что приводит к увеличению количества подвижных носителей в канале и, соответственно, к уменьшению сопротивления канала. При определенных напряжениях на затворе можно создать «насыщение» канала носителями заряда и тем самым обеспечить достаточно малое сопротивление канала.

В классическом варианте использования СИТ, как усилительного прибора, на его затвор подают отрицательное напряжение относительно истока. В этом случае вокруг p - n - переходов формируется широкий обедненный слой, простирающийся главным образом в n^- область. При определенном отрицательном напряжении на затворе и при $U_{СИ} = 0$ происходит смыкание обедненных областей (рис. 5.10 в) и возникает «отсечка» проводящего канала. Ток стока в этом случае протекать не может.

При приложении положительного напряжения сток – исток к СИТ в обедненной носителями перекрытой области результирующее электрическое поле будет определяться разностью напряженностей полей сток – исток и результирующего электрического поля p - n перехода. При заданном отрицательном напряжении затвор – исток и, одновременно, при увеличении положительного напряжения сток исток «тормозящее» поле у истока будет уменьшаться, что будет создавать условия для прохождения основных носителей от истока к стоку. При определенном напряжении сток – исток потенциальный барьер исчезнет и в канале появится ток. При дальнейшем увеличении напряжения $U_{СИ}$ ток стока будет расти, поскольку в этом случае не наступает перекрытие канала.

Если увеличивать напряжение на затворе – относительно истока, то результирующее электрическое поле p - n перехода будет сильнее, а, следовательно, большее напряжение $U_{СИ}$ придется прикладывать к СИТ для компенсации запирающего действия поля p - n перехода и обеспечения того, чтобы через канал протекал ток. Следовательно, стоковые ВАХ СИТ при увеличении величины отрицательного напряжения (по модулю) смещаются вправо. Типичное семейство стоковых характе-

ристик СИТ в полевом режиме (ограничения тока объемным зарядом) имеет вид, показанный на рис. 5.11.

Как видно из рис. 5.11 ВАХ СИТ не имеет пологой области насыщения, как у выходных характеристик рассмотренных нами ранее биполярных и полевых транзисторов. По этой причине полевой транзистор со статической индукцией имеет малое выходное дифференциальное сопротивление (является управляемым источником напряжения), что существенно повышает его энергетические показатели при работе на низкоомные нагрузки. Следует также иметь в виду, что характеристики СИТ более линейны, чем характеристики биполярных транзисторов. Это позволяет создавать усилительные устройства с малыми коэффициентами нелинейных искажений.

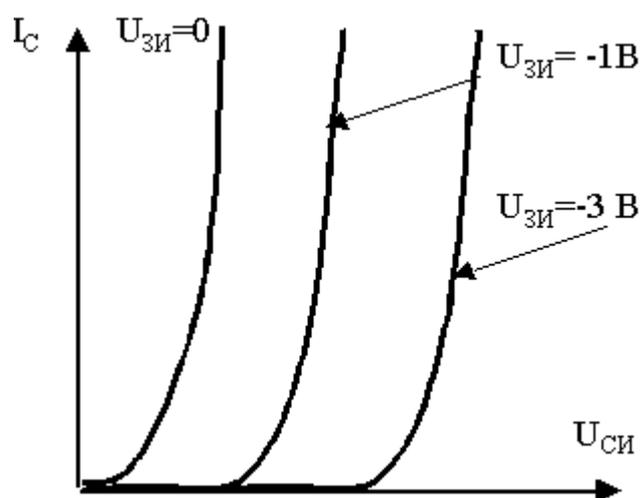


Рис. 5.11. Семейство стоковых характеристик СИТ в полевом режиме

5.3.3. Биполярный транзистор с изолированным затвором

Биполярным транзистором с изолированным затвором IGBT (Insulated Gate Bipolar Transistor) называют монокристаллический транзистор, биполярно – полевая структура которого имеет множество параллельных ячеек, в которых совмещаются функции биполярной проводимости и полевого управления.

Базовая ячейка в верхней своей части содержит полевой транзистор с изолированным затвором и индуцированным n – каналом (рис. 5.12). Сопротивление канала этого полевого транзистора является пер-

вой составляющей общего сопротивления на пути протекания тока. Структура, составленная из $p^+ - n -$ областей нижней и средней части ячейки, является несимметричным с точки зрения введенной примеси диодом, у которого область p^+ – это эмиттер (анод), а n – база (катод). Области p, n, p , располагающиеся в ячейке, в свою очередь, можно рассматривать как два соединенных $p^+ - n - p$ и $n - p - n^+$ транзистора. Схема соединения этих транзисторов такова, что они образуют (см. п. 5.2.1) комбинацию, подобную тиристор. Это «нежелательный», образуемый вследствие такой технологии изготовления тиристор. Его часто называют «паразитным» тиристором и чтобы он не «включался» в работу эмиттерный переход верхнего $n - p - n^+$ транзистора шунтируют (закорачивают) резистором.

Если к затвору IGBT не прикладывают напряжение, то канал в полевом транзисторе не наводится, а, следовательно, отсутствует путь протекания тока между эмиттером и коллектором. Это соответствует разомкнутому состоянию ключа. Внешнее напряжение, имеющееся между коллектором и эмиттером, распределяется в структуре таким образом, что большая часть его приложена к широкому обедненному слою, располагающемуся в слаболегированной n – базе. Это определяет высокие предельные значения рабочего напряжения элементарной ячейки в закрытом состоянии и IGBT в целом.

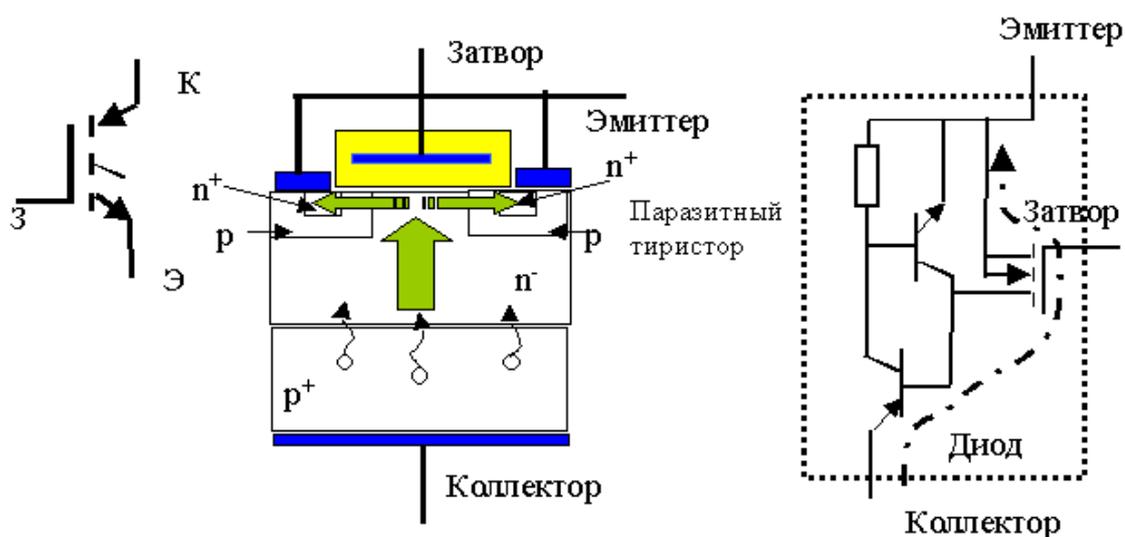


Рис. 5.12. Условное обозначение на схемах, структура элементарной ячейки с n -каналом и эквивалентная схема биполярного транзистора с изолированным затвором

При подаче на затвор IGBT относительно эмиттера положительного напряжения в полевом транзисторе индуцируется проводящий канал (в p – областях ячейки) и между выводами. Если между коллектором и эмиттером имеется внешнее положительное напряжение, то по цепи коллектор - диод ($p^+ - n -$ области нижней и средней части ячейки) – канал полевого транзистора начинает протекать ток. В этом состоянии IGBT подобен замкнутому ключу.

Так как схемотехнически структура элементарной ячейки представляет (без паразитных элементов) комбинацию полевого транзистора и биполярного $p^+ - n - p$ транзистора, то прямое напряжение на замкнутом ключе будет складываться из двух компонентов: порогового напряжения на прямосмещенном эмиттерном переходе $p^+ - n - p$ транзистора и падения напряжения на сопротивлении канала полевого транзистора:

$$U_{КЭ}^{IGBT} = U_{БЭ}^{pnp} + I_B^{pnp} r_{КАН}. \quad (5.7)$$

Часто говорят, что напряжение на замкнутом ключе состоит из двух составляющих – диодной и омической составляющей.

Очевидно, что для снижения мощности потерь IGBT в открытом состоянии (замкнутом состоянии ключа) $P_{ПOT}$ необходимо обеспечить минимальное значение сопротивления канала $r_{КАН}$.

В IGBT при подаче положительного напряжения на коллектор относительно эмиттера под воздействием этого напряжения из высоколегированного p^+ слоя в глубь n области начинается односторонняя инжекция неосновных носителей заряда - дырок (рис. 5.12). Инжектированные дырки снижают сопротивление протекающему току в самой n области, а также, в области индуцированного (наведенного) $n -$ канала. Это явление называют в литературе модуляцией проводящего канала, а саму структуру IGBT – структурой с модулируемой проводимостью. При этом оказывается возможным в значительной степени снизить сопротивление открытого канала полевого транзистора.

Поскольку омическая составляющая напряжения на ключе расположена в базовой цепи $p^+ - n - p$ транзистора, то модулирующее дейст-

вие инъекции по снижению сопротивления канала можно, в первом приближении, оценивать уменьшением его в β раз:

$$U_{КЭ}^{IGBT} = U_{БЭ}^{pnp} + I_B^{pnp} r_{КАН} \approx U_{БЭ}^{pnp} + I_K^{IGBT} \frac{r_{КАН}}{\beta^{pnp}}, \quad (5.8)$$

где β^{pnp} - коэффициент передачи базового тока $p^+ - n - p$ транзистора.

Выражение (5.8) свидетельствует о следующем.

1. Напряжение на IGBT в открытом состоянии (при замкнутом состоянии ключа) не может быть меньше, чем пороговое напряжение $U_{БЭ} = U_{pn}$.

2. Температурная зависимость напряжения складывается из двух противоположных составляющих: отрицательного температурного коэффициента диодной компоненты и положительной омической. Результирующий температурный коэффициент можно сделать малым, поэтому IGBT является температурно-устойчивым (может работать при температурах до 200°C).

3. Уменьшения прямых падений напряжений на открытом IGBT можно достигнуть двумя путями: повышением степени легирования n^- - области базы $p^+ - n - p$ транзистора и увеличением коэффициента тока базы β^{pnp} . Оба эти пути неоднозначны. Первый путь имеет ограничения, связанные со снижением предельного напряжения ключа в закрытом состоянии. Значительное увеличение β^{pnp} связано с тем, что может «включиться в работу» паразитный тиристор, образованный комбинацией соединенных $p^+ - n - p$ и $n - p - n^+$ транзисторов. Для IGBT включение паразитного тиристора крайне нежелательно, так как это может привести к выходу из строя ячейки (катастрофическому отказу). Чтобы этого не происходило эмиттерный переход $n - p - n^+$ транзистора «шунтируют» сопротивлением, а β^{pnp} выбирают в разумных пределах. Для уменьшения потерь в открытом состоянии ключа структуры IGBT оптимизируют (выбирают наиболее приемлемые размеры, концентрации примесей, формы областей и т. п.).

Отметим, что элементарная ячейка в настоящее время имеет линейные размеры порядка 1 мкм. Подобные IGBT имеют прямые падения напряжения не более 1,5 В при токах порядка 1000 А. Имеются высоковольтные структуры IGBT (HV-IGBT), рассчитанные на рабочие напряжения более 1700 В. Их структуры аналогичны рассмотренным выше, за исключением деталей касающихся уменьшения утечек и предотвращения поверхностного пробоя.

Следует иметь в виду, что в настоящее время выпускаются IGBT не только с n -каналом, но и с p – каналом. Базовые ячейки таких биполярных транзисторов с изолированным затвором не содержат внутреннего «антипараллельного, демпферного» диода, свойственного мощным структурам МДП. Поскольку наличие быстро восстанавливающегося демпферного диода становится необходимым условием применения и защиты ключей, то, как правило, используют специально разработанные и согласованные по характеристикам с IGBT отдельные кристаллы диодов, встроенные в корпус прибора.

Важной отличительной особенностью IGBT по сравнению с полевыми транзисторами является наличие инжекции неосновных носителей заряда, а, следовательно, и процессов накопления их в базовых областях $p^+ - n - p$ транзистора элементарной ячейки. Быстрое выключение IGBT (выключением напряжения на затворе) приводит к «пропаданию» канала и пути их отвода. Дальнейшее уменьшение накопленного заряда возможно только на основе рекомбинационных эффектов. Это, в свою очередь, определяет наличие дополнительной фазы в выключаемом токе IGBT, называемой хвостовой частью коллекторного тока (tail current). Для уменьшения этого явления, ухудшающего частотные свойства IGBT, используют различные технологические приемы, связанные с уменьшением времени жизни неосновных носителей заряда.

5.4. Общие сведения об управлении силовыми полупроводниковыми ключами

В устройствах силовой электроники всегда можно выделить силовую (с мощными потоками электрической энергии) и управляющую части, которые находятся в непрерывном взаимодействии.

Силовая часть, имеющая в своем составе мощные электрические ключи, обеспечивает процессы коммутации, преобразования электрической энергии. *Задача управляющей части* – управление потоком энергии, сбор и обработка информации о состоянии функционирования всей системы, диагностика ее неисправностей. В простейшем случае система управления вырабатывает импульсы управления силовыми ключами и строится на основе регулирования фазы (положения на оси времени) управляющих импульсов. Уже отмечалось, что суть, к примеру, импульсно-фазового управления состоит в изменении момента подачи отпирающих импульсов (изменении угла управления) по отношению к синусоидальной кривой напряжения сети. Функционально система управления представляет собой отдельный, самостоятельный блок.

Часть системы управления, которая формирует логику входных сигналов силовых ключей, а затем их усиливает до требуемых уровней тока и напряжения, называется *формирователем импульсов управления (ФИУ)*. В структуре ФИУ, в свою очередь, можно выделить информационно – логическую часть и усилитель управляющих импульсов.

Усилитель импульсов управления, который формирует выходные сигналы заданной мощности и формы, изготовленный в виде отдельной интегральной микросхемы, называется драйвером.

Обычно драйвер содержит входной узел, принимающий информационные сигналы, узел согласования, преобразующий информационный маломощный сигнал в сигнал управления нужного уровня, и выходной узел, осуществляющий окончательное формирование импульсов управления требуемой формы и мощности. Дополнительно на драйвер могут быть возложены функции защиты силового ключа от перегрузок или функции слежения за уровнем напряжения питания микросхемы.

В настоящее время одной из основных тенденций развития современных силовых ключей, основанной на достижениях технологии, является объединение в едином корпусе прибора функций переключателя, его управления и защиты от перегрузок. Данные силовые приборы получили название «разумные» (Smart) или интеллектуальные (Intelligent). Они позволяют избавиться от громоздких и неэкономичных дополнительных цепей защиты.

Контрольные вопросы

1. Какую область электроники называют энергетической или силовой электроникой?
2. В чем заключается основная задача полупроводникового прибора в энергетической электронике?
3. Какими свойствами должен обладать ключ силовой электроники?
4. Приведите определение тиристора.
5. Чем отличаются друг от друга тиристоры?
6. Приведите основные типы тиристоров.
7. Поясните структуру простейшего диодного тиристора.
8. Что представляет собой динистор при обратном включении?
9. Поясните понятие – режим глубокого насыщения тиристора.
10. Нарисуйте и поясните график ВАХ динистора.
11. Нарисуйте схему включения динистора в электрическую цепь, а также эквивалентную схему, замещающую тиристор, в виде ключа.
12. Перечислите и поясните режимы работы диодного тиристора.
13. Поясните устройство и принцип действия триодного тиристора.
14. Как осуществляется управление включением триодного тиристора?
15. Нарисуйте и поясните семейство вольт-амперных характеристик триодного тиристора при различных токах управляющего электрода.
16. Поясните понятия: включить и выключить триодный тиристор.
17. Какие тиристоры называются незапираемыми?
18. Какие тиристоры называют двухоперационными тиристорами?
19. Для решения каких задач предназначены симисторы и где они применяются?
20. Нарисуйте структуру и условное обозначение на схемах симистора.
21. Нарисуйте и поясните семейство вольт-амперных характеристик симистора при различных токах управляющего электрода.
22. Какой способ регулирования напряжения и тока в нагрузке часто называют фазовым управлением?
23. Перечислите основные разновидности мощных тиристоров, их назначение и основные отличия.
24. В чем заключаются преимущества индукционных тиристоров?
25. Охарактеризуйте принцип работы и назначение полевых тиристоров с изолированным затвором МСТ.

26. Перечислите основные преимущества полевых транзисторов над биполярными при использовании их в силовых ключевых схемах.
27. Что называют статическим индукционным транзистором? Для чего они предназначены?
28. Нарисуйте и поясните типичное семейство стоковых характеристик СИТ в полевом режиме.
29. Что представляет собой и для чего предназначен биполярный транзистор с изолированным затвором IGBT?
30. Приведите условное обозначение на схемах и эквивалентную схему биполярного транзистора с изолированным затвором.
31. Какое явление называют модуляцией проводящего канала?
32. Какой важной отличительной особенностью обладают IGBT по сравнению с полевыми транзисторами?
33. Охарактеризуйте кратко назначение и структуру усилителя импульсов управления (драйвера).

Рекомендуемая литература

1. Алиев И.И. Виртуальная электроника. Компьютерные технологии в электротехнике и электронике. – М.: РадиоСофт, 2003. – 112 с.
2. Гультаев А.К. Визуальное моделирование в среде Matlab: Учебный курс. – Санкт-Петербург: КОРОНА-Принт, 2000. – 228 с.
3. Гусев В.Г., Гусев Ю.М. Электроника и микропроцессорная техника. – М.: Высшая школа, 2004. – 790 с.
4. Данилов И.А. Общая электротехника с основами электроники. – М.: Высшая школа, 2000. – 752 с.
5. Карлащук В.И. Электронная лаборатория на IBM PC. – М.: СОЛОН-Пресс, 2004. – 705 с.
6. Киселев Б.М. Matlab. Пакет Simulink // Радиомир. Ваш компьютер – 2005 – №9. – С14 – 18.
7. Лачин В.В. Электроника: Учебное пособие для втузов / В.В. Лачин. – Ростов на Дону: Феникс, 2002. – 576 с.
8. Опадчий Ю.Ф., Глудкин О.П., Гуров А.И. Аналоговая и цифровая электроника (Полный курс): Учебник для вузов / Ю.Ф. Опадчий, О.П. Глудкин, А.И. Гуров. – М.: Горячая линия – Телеком, 2002. – 768 с.
9. Партала О.Н. Цифровая электроника. – СПб.: Наука и Техника, 2001. – 224 с.
10. Петров К.С. Радиоматериалы, радиокомпоненты и электроника. СПб.; М.; Харьков; Минск: Питер, 2003. – 512 с.
11. Пресс Ф.П. Фоточувствительные приборы с зарядовой связью. – М.: Радио и связь, 1991. – 264 с.
12. Прянишников В.А. Электроника: Полный курс лекций. – 4 изд. – СПб.: КОРОНА-Принт, 2004. – 416 с.
13. Рекус Г.Г., Белоусов А.И. Сборник задач по электротехнике и электронике: Учебное пособие для вузов / Г.Г. Рекус, А.И. Белоусов. – М.: Высшая школа, 1991. – 416 с.
14. Черных И.В. Simulink: Инструмент моделирования динамических систем. – <http://matlab.ru>, раздел «Simulink».
15. Четверухин Б.М. Основы электротехники и электроники: Конспект лекций / Б.М. Четверухин, Г.С. Прокудин. – К.: Изд-во Европейского ун-та, 2002. – 149 с.

Содержание

| | |
|--|----|
| Введение | 3 |
| 1. Основные структуры полупроводниковой электроники и диоды на их основе | 7 |
| 1.1. Основные понятия и материалы полупроводниковой электроники | 7 |
| 1.1.1. Основные материалы полупроводниковой электроники и понятия их характеризующие | 7 |
| 1.1.2. Собственные полупроводники и понятия их характеризующие | 10 |
| 1.1.3. Примесные полупроводники и понятия их характеризующие | 23 |
| 1.2. Структуры полупроводниковой электроники | 32 |
| 1.3. Электронно-дырочная структура и ее свойства | 34 |
| 1.3.1. Образование и равновесное состояние электронно-дырочного перехода | 34 |
| 1.3.2. Параметры, характеризующие электронно-дырочную структуру (идеальный диод) в состоянии термо-дина- мического равновесия и при отсутствии напряжения | 43 |
| 1.3.3. Функционирование электронно-дырочной структуры при приложении к ней внешнего прямого напряжения ... | 45 |
| 1.3.4. Функционирование электронно-дырочной структуры при приложении внешнего обратного напряжения | 51 |
| 1.3.5. Вольт-амперная характеристика и основные свойства электронно-дырочной структуры (идеального диода) | 55 |
| 1.4. Переход металл – полупроводник и его свойства | 59 |
| 1.4.1. Основные понятия | 59 |
| 1.4.2. Образование контакта металл – полупроводник и процессы в нем в равновесном состоянии | 61 |
| 1.4.3. Функционирование структуры металл – полупро- водник (идеального диода Шоттки) при приложении к ней внешнего напряжения | 62 |
| 1.4.4. Омические (невыпрямляющие) контакты | 64 |
| 1.5. Полупроводниковые диоды общего применения | 66 |
| 1.5.1. Общие сведения | 66 |
| 1.5.2. Отличия вольт-амперных характеристик реальных | |

| | |
|--|-----------|
| диодов от идеальных | 67 |
| 1.5.3 Выпрямительные диоды | 73 |
| 1.5.4 Стабилитроны | 74 |
| 1.6. Полупроводниковые диоды специального назначения | 75 |
| 1.6.1. Варикапы | 75 |
| 1.6.2. Импульсные диоды и высокочастотные диоды | 76 |
| Контрольные вопросы | 79 |
| 2. Биполярные и полевые транзисторы | 82 |
| 2.1. Общие сведения о биполярных транзисторах | 83 |
| 2.1.1. Назначение и структура биполярного транзистора | 83 |
| 2.1.2. Схемы включения биполярного транзистора | 85 |
| 2.1.3. Электрические режимы работы биполярного транзистора | 86 |
| 2.1.4. Что значит объяснить принцип действия транзистора ... | 89 |
| 2.2. Принцип действия биполярного транзистора в различных электрических режимах и характеристики его работы | 89 |
| 2.2.1. Функционирование биполярного транзистора в активном режиме | 89 |
| 2.2.2. Токи биполярного транзистора в активном режиме | 95 |
| 2.2.3. Признаки работы биполярного транзистора в активном режиме | 98 |
| 2.2.4. Функционирование биполярного транзистора в режиме насыщения | 101 |
| 2.2.5. Признаки работы биполярного транзистора в режиме насыщения | 107 |
| 2.2.6. Функционирование биполярного транзистора в режиме отсечки | 109 |
| 2.2.7. Признаки работы биполярного транзистора в режиме отсечки | 110 |
| 2.2.8. Статические характеристики биполярного транзистора ... | 111 |
| 2.2.9. Температурные и частотные свойства биполярного транзистора | 119 |

| | |
|--|------------|
| 2.3. Полевые транзисторы | 122 |
| 2.3.1. Общие сведения о полевых транзисторах | 122 |
| 2.3.2. Устройство и принцип действия полевых транзис- тов с управляющим переходом | 126 |
| 2.3.3. Устройство и принцип действия полевых транзис- торов с изолированным затвором | 133 |
| Контрольные вопросы | 140 |
| 3. Основы микроэлектроники | 143 |
| 3.1. Общие сведения о микроэлектронных изделиях | 144 |
| 3.1.1. Тенденции развития элементной базы электроники и их воплощение в поколениях электронной аппаратуры | 144 |
| 3.1.2. Понятие об интегральной микросхеме | 146 |
| 3.1.3. Принципы изготовления интегральных микросхем | 147 |
| 3.1.4. Особенности элементов электрических схем интегрального исполнения | 151 |
| 3.2. Особенности и примеры схемных решений типовых узлов микроэлектронных изделий | 157 |
| 3.2.1. Необходимость нового подхода к созданию схем | 157 |
| 3.2.2. Особенности схем установки рабочей точки во входной цепи усилительного каскада интегрального исполнения на биполярном транзисторе в СОЭ | 158 |
| 3.2.3. Использование «эталонов» тока и напряжения – харак- терная черта каскадов интегрального исполнения | 160 |
| 3.2.4. Использование повторителей напряжения в ИМС | 166 |
| 3.2.5. Дифференциальный усилитель – важнейшее звено интегральных схем | 171 |
| 3.2.6. Элемент флэш-памяти | 176 |
| 3.3. Датчики | 182 |
| 3.3.1. Общие сведения о датчиках | 182 |
| 3.3.2. Полупроводниковые датчики температуры и тензодатчики | 183 |
| 3.3.3. Полупроводниковые датчики магнитного поля | 184 |
| Контрольные вопросы | 187 |
| 4. Оптоэлектронные полупроводниковые приборы и компоненты средств отображения информации | 189 |
| 4.1. Общие сведения об оптическом излучении | 190 |

| | |
|---|-----|
| 4.1.1. Понятия, энергетические и световые параметры оптического излучения | 190 |
| 4.1.2. Общие сведения о механизме генерации оптического излучения в полупроводниках | 194 |
| 4.1.3. Общие сведения о поглощении полупроводником оптического излучения | 200 |
| 4.2. Некогерентные и когерентные оптические излучатели | 204 |
| 4.2.1. Устройство оптического излучателя, его характеристики и параметры | 204 |
| 4.2.2. Некогерентные излучатели – светодиоды и полупро- водниковые средства отображения информации | 210 |
| 4.2.3. Некогерентные излучатели – диоды инфракрасного диапазона | 214 |
| 4.2.4 Когерентные излучатели–полупроводниковые лазеры ... | 216 |
| 4.3. Приемники оптического излучения | 222 |
| 4.3.1. Фоторезисторы | 222 |
| 4.3.2. Фотодиоды | 224 |
| 4.3.3. Специальные фотодиоды | 228 |
| 4.3.4. Фототранзисторы | 231 |
| 4.4. Компоненты оптико-электронных схем и средств отображения информации | 233 |
| 4.4.1. Оптопары и микроэлектронные реле | 233 |
| 4.4.2. Линейные и матричные датчики изображений | 234 |
| Контрольные вопросы | 239 |
| 5. Полупроводниковые ключи в энергетических (силовых) устройствах электроники | 241 |
| 5.1. Общие сведения о применении полупроводниковых ключевых приборов в энергетических (силовых) устройствах электроники | 242 |
| 5.2. Тиристоры | 244 |
| 5.2.1. Устройство и принцип действия диодного тиристора | 244 |
| 5.2.2. Устройство и принцип действия триодного тиристора ... | 251 |
| 5.2.3. Устройство и принцип действия симметричного тиристора (симистора) | 254 |
| 5.2.4. Разновидности мощных тиристоров | 257 |
| 5.3. Базовые транзисторные структуры силовых ключей | 258 |

| | |
|---|------------|
| 5.3.1. Общие сведения о базовых транзисторных структурах силовых ключей | 258 |
| 5.3.2. Полевой транзистор со статической индукцией | 260 |
| 5.3.3. Биполярный транзистор с изолированным затвором ... | 263 |
| 5.4. Общие сведения об управлении силовыми полупроводниковыми ключами | 267 |
| Контрольные вопросы | 268 |
| Рекомендуемая литература | 270 |

УЧЕБНОЕ ИЗДАНИЕ

**ОСНОВЫ ЭЛЕКТРОТЕХНИКИ И ЭЛЕКТРОНИКИ.
ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ ЭЛЕКТРОНИКИ**

Учебное пособие

Часть 2

Авторы: **Гоков Александр Михайлович**
Жидко Евгений Анатольевич

Ответственный за выпуск **Бондаренко Е. А.**

Ответственный редактор **Седова Л. Н.**

Редактор **Гергеша А. В.**

Корректор **Бриль В. А.**

НАВЧАЛЬНЕ ВИДАННЯ

**ОСНОВИ ЕЛЕКТРОТЕХНІКИ ТА ЕЛЕКТРОНІКИ.
ЕЛЕМЕНТИ ЗАГАЛЬНОЇ ТЕОРІЇ ЕЛЕКТРОНІКИ**

Навчальний посібник

Частина 2

Автори: **Гоков Олександр Михайлович**
Жидко Євген Анатолійович

ISBN 966-676-142-4

План 2006 г. Поз. № 123-П.

Подп. в печ. *17.05.2006.* Формат 60 × 90 1/16. Бумага MultiCopy. Печать Riso.

Усл.-печ. л. 17,25. Уч.-изд. л. 18,11. Тираж *400* экз. Зак. № *326*

Свидетельство о внесении в Государственный реестр субъектов издательского дела Дк №481 от 13.06.2001 г.

Издатель и изготовитель — издательство ХНЭУ, 61001, г. Харьков, пр. Ленина, 9а

Гоков А.М., Жидко Е.А.

ОСНОВЫ ЭЛЕКТРОТЕХНИКИ И ЭЛЕКТРОНИКИ. ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ ЭЛЕКТРОНИКИ

У Ч Е Б Н О Е П О С О Б И Е

Приведены основные сведения из общей теории электроники, фундаментальные идеи, термины и определения. Излагаются принципы построения, модели, характеристики и параметры простейших структур полупроводниковой электроники — диодов — и более сложных — транзисторов. Изучаются основы микроэлектроники и оптоэлектроники, модели, принципы построения и функционирования реальных устройств. Рассматриваются полупроводниковые изделия силовой электроники, их практические реализации и применение. Приведен широкий круг характерных практических примеров.



ИЗДАТЕЛЬСТВО **ХНЭУ**
ХАРЬКОВ 2006