

Высшее образование

Г. Г. Шишкин, А. Г. Шишкин

Электроника



ДРОФА

Г. Г. Шишкин, А. Г. Шишкин

Электроника



Рекомендовано

*учебно-методическим объединением вузов Российской Федерации
по образованию в области радиотехники, электроники,
биомедицинской техники и автоматизации
в качестве учебника для студентов высших учебных заведений,
обучающихся по направлению 210300 — «Радиотехника»*

МОСКВА



ДРОФД

2 0 0 9

УДК 621.38(075.8)

ББК 32.85я73

Ш65

Рецензенты:

д-р техн. наук, проф. *В. Н. Кулешов*
(зав. каф. «Формирование колебаний и сигналов» МЭИ (ТУ));
канд. техн. наук, доц. *В. Н. Павлов*
(СПб. ГЭТУ «ЛЭТИ» им. В. И. Ульянова (Ленина))

Шишкин, Г. Г.

Ш65 Электроника : учеб. для вузов / Г. Г. Шишкин, А. Г. Шишкин. — М. : Дрофа, 2009. — 703, [1] с. : ил.

ISBN 978-5-358-03595-9

Учебник состоит из шести разделов и охватывает все области современной электроники — физику полупроводников и электрических переходов, физические процессы, устройство и характеристики полупроводниковых диодов, биполярных и полевых транзисторов, тиристоров, фотоэлектрических, оптоэлектронных и электровакуумных приборов, а также такие актуальные вопросы, как базовые элементы аналоговых и цифровых интегральных схем, принципы и устройства квантовой электроники, наноэлектроники и функциональной электроники. Рассматриваются проблемы шумов, радиационной стойкости и надежности приборов. Все главы снабжены контрольными вопросами для повторения материала.

Для подготовки бакалавров, магистров и инженеров (специалистов) широкого профиля, обучающихся по направлению 210300 — «Радиотехника».

УДК 621.38(075.8)

ББК 32.85я73

ISBN 978-5-358-03595-9

© ООО «Дрофа», 2009

Данный учебник предназначен для изучения дисциплины «Электроника», которая составляет основу общепрофессиональной подготовки бакалавров, магистров и специалистов (инженеров) широкого профиля по направлению 210000 «Электронная техника, радиотехника и связь» (в частности, 210300 «Радиотехника»). В связи со значительным прогрессом в области электроники, достигнутым в последнее десятилетие, произошли существенные преобразования в содержании государственного стандарта высшего профессионального образования по указанному направлению и в разработанных на его основе учебных планах и программах. Данная книга является, по существу, первым учебником, который полностью соответствует современным требованиям государственного стандарта и примерной программе по дисциплине «Электроника».

С учетом вышесказанного учебник включает ряд вопросов, которые традиционно не рассматривались в курсах, посвященных изучению электронных приборов [1,2], и охватывает все разделы современной электроники, в том числе и базовые элементы аналоговых и цифровых интегральных схем, принципы и устройства квантовой и функциональной электроники, а также наноэлектроники.

При изложении материала в основном использовался принцип объединения электронных приборов по виду рабочей среды, что обуславливает общность большинства физических явлений в изучаемых вопросах и облегчает усвоение материала студентами. В первых пятнадцати главах, содержание которых составляет основу дисциплины, рассмотрение начинается с полупроводниковых, а заканчивается электровакуумными приборами (ЭВП) сверхвысоких частот (СВЧ) с динамическим управлением и газоразрядными приборами. Последующие главы (гл. 16, 17, 19—21) посвящены оптоэлектронным приборам, включая приборы квантовой (когерентной) электроники. Рассмотрение общих вопросов, определений, относящихся к классификации электронных приборов различного типа, режимам и условиям их эксплуатации, на-

дежности, радиационной стойкости, шумам, вынесено во введение и в последний, шестой раздел книги.

Основное внимание в учебнике уделено физическим основам, принципам работы и характеристикам полупроводниковых приборов, интегральных схем (ИС), ЭВП СВЧ с динамическим управлением, электронно-лучевых приборов, газоразрядных приборов и индикаторов, оптоэлектронных приборов и приборов квантовой электроники (лазеров и мазеров). Подробно анализируются процессы, составляющие основу современной микроэлектроники, а именно явления, происходящие в электронно-дырочных переходах, контактах металл—полупроводник, гетеропереходах; рассматриваются свойства различных моделей диодов, транзисторов, основы технологии изготовления дискретных полупроводниковых приборов и элементов полупроводниковых и вакуумных ИС.

В учебник, согласно современным требованиям, впервые включены материалы по основным базовым элементам аналоговых и цифровых ИС (гл. 7—9), по наноэлектронным приборам и приборам функциональной электроники (гл. 10), фотоэлектронным приборам с зарядовой связью (гл. 16). Новацией в данной книге является также включение раздела, посвященного приборам квантовой электроники — лазерам и мазерам (гл. 17—21), а также существенное расширение сведений по плазменным панелям и жидкокристаллическим индикаторам (гл. 15).

Для практического закрепления изучаемого материала в основном тексте учебника и в приложении 1 приведены сведения о порядках величин параметров приборов, электрофизических характеристиках материалов и исследуемых процессов. Подавляющее большинство графиков имеет размерный масштаб. В теоретических разделах, а также в приложении 2 приводятся необходимые математические выкладки и расчетные соотношения, большая часть формул иллюстрируется числовыми примерами; используются определения и термины, регламентированные государственными стандартами. Для получения данных о постоянно обновляемой номенклатуре электронных приборов в учебнике даются ссылки на соответствующие каталоги, справочники и другие издания. Специальные вопросы, имеющие факультативный характер, — методы описания шумов (п. 22.3) и сведения по радиационной стойкости электронных приборов (п. 23.3) — набраны мелким шрифтом. В помощь студентам в конце каждой главы приведены контрольные вопросы для повторения изученного материала.

Для освоения дисциплины «Электроника» студенту достаточно знаний по общеобразовательным и общетехническим предметам. В результате изучения этой дисциплины студент получает базовую подготовку, необходимую для успешного освоения специальных радиотехнических курсов и последующего решения различного рода профессиональных задач, связанных с рациональным выбором электронных приборов и режимов их работы в радиотехнической аппаратуре. Подробное рассмотрение физических основ явлений, принципов работы, параметров, характеристик и моделей приборов направлено на развитие у студента умения самостоятельно решать задачи моделирования, анализа и синтеза электронных устройств при их проектировании и эксплуатации. Успехов в разработке и использовании такой техники смогут достичь лишь специалисты, имеющие прочные знания по электронике полупроводниковых приборов.

Основной текст учебника написан Г. Г. Шишкиным; главы 15, 16, 19 и приложение 2 написаны авторами совместно.

Авторы выражают признательность коллегам по кафедре «Теоретическая радиотехника» Московского авиационного института, высказавшим конструктивные замечания по содержанию рукописи, приносят искреннюю благодарность академику МАНШ, профессору СГУ, доктору физико-математических наук, заслуженному деятелю науки РФ Д. А. Усанову за ряд полезных критических замечаний и рекомендаций. Особую благодарность авторы приносят редактору книги доценту, кандидату физико-математических наук И. Я. Ицхоки за тщательность и высокий профессионализм редактирования, а также рецензентам: профессору, доктору технических наук В. Н. Кулешову и доценту, кандидату технических наук В. Н. Павлову за ценные замечания и рекомендации, способствовавшие улучшению качества учебника. Большую помощь при оформлении рукописи оказали студенты факультета радиоэлектроники Московского авиационного института А. Пашенько, П. Федоров, В. Суриков, за что авторы им благодарны.

Основные сведения об электронных приборах. Важнейшими направлениями научно-технического прогресса являются освоение передовых технологий, включая лазерную, плазменную и пучковую, и автоматизация и механизация производства. Современный этап решения задач в рамках этих направлений опирается на революцию в электронно-вычислительной технике. Для того чтобы обеспечить глубокие качественные изменения в производительных силах, создание принципиально новых видов продукции, техники и технологии, необходимо интенсивное развитие и расширение исследований в области естественных и технических наук, включая физику твердого тела, микро- и квантовую электронику и оптику, радиофизику и радиоэлектронику, от уровня развития которых зависит общий индустриальный и научный потенциал общества, включая достижения в автоматизации производства, развитии принципиально новых технологий, освоении космоса, вычислительной технике, медицине, искусстве, телевидении и т. д.

Таким образом, характеристики радиоэлектронных информационных систем, которые осуществляют накопление, передачу, хранение и переработку информации, представляемой в форме различного рода информационных сигналов, определяются уровнем развития электроники.

Электроника — это наука о взаимодействии электронов с электромагнитными полями, о методах, охватывающих исследование и разработку электронных приборов, и о принципах их использования. *Информационный сигнал* формируется за счет изменения параметров какого-либо физического процесса. Для передачи и обработки больших объемов информации наиболее часто сигналы формируются с помощью электрического тока (или напряжения) и электромагнитного излучения. Процесс обработки информации в том или ином виде осуществляется посредством электронных приборов.

Электронные приборы — это устройства, работа которых основана на использовании электрических, тепловых, оптических

ских и акустических явлений в твердом теле, жидкости, вакууме, газе или плазме. Электронные приборы используются как элементы радиоэлектронной аппаратуры, не подлежащие сборке, разборке и ремонту.

Основные наиболее общие функции, выполняемые приборами, состоят в преобразовании либо информационных сигналов, либо энергии. В качестве примеров преобразования энергии можно привести преобразование солнечной энергии в электрическую, тепловой в электрическую, преобразования переменного тока в постоянный и наоборот.

Само название «электронные приборы» связано с тем, что все процессы преобразования сигналов и энергии происходят либо только за счет движения электронов, либо при их непосредственном участии. Основными задачами электронного прибора как преобразователя информационных сигналов являются: усиление, генерирование, передача, накопление и хранение сигналов, выделение их на фоне шумов.

Наиболее часто электронные приборы классифицируются по их назначению, физическим свойствам, основным электрическим параметрам, конструктивно-технологическим признакам, по роду рабочей среды и т. д.

В зависимости от вида сигналов и способа обработки информации все существующие электронные приборы разделяются на *электропреобразовательные, электросветовые, фотоэлектрические, термоэлектрические, акустические и механоэлектрические*.

В электропреобразовательных приборах происходит преобразование электрических сигналов, в электросветовых приборах электрические сигналы преобразуются в оптические (световые), в фотоэлектрических и термоэлектрических соответственно оптические и тепловые сигналы преобразуются в электрические, в акустоэлектронных приборах электрические сигналы преобразуются в акустические. Механоэлектрические преобразуют механический сигнал в электрический.

Электропреобразовательные приборы представляют самую большую группу электронных приборов. К ним относятся различные типы полупроводниковых диодов, биполярные и полевые транзисторы, тиристоры, электровакуумные лампы — диоды, триоды, тетроды, пентоды и т. д.; газоразрядные приборы — стабилитроны, газотроны, тиратроны и т. д. К электросветовым — светодиоды, люминесцентные конденсаторы, лазеры, электронно-

лучевые трубки. К фотоэлектрическим — фотодиоды, фототранзисторы, фототиристоры, солнечные батареи и ряд других. К термоэлектрическим — полупроводниковые диоды, транзисторы, термисторы. К акустоэлектрическим — акустоэлектрические усилители и генераторы, фильтры, линии задержки на поверхностных акустических волнах и др.

По виду рабочей среды различают следующие классы приборов: полупроводниковые, твердотельные, электровакуумные, газоразрядные, хемотронные (рабочая среда — жидкость).

В зависимости от назначения и выполняемых функций электронные приборы разделяются на выпрямительные, усилительные, генераторные, переключательные, преобразовательные, индикаторные и т. д. По диапазону рабочих частот — на низкочастотные, высокочастотные и сверхвысокочастотные; по мощности — на маломощные, средней мощности и мощные.

Основные понятия о режимах и параметрах электронных приборов. Понятие *режима электронного прибора* включает совокупность условий, определяющих его работу. Если режим работы прибора соответствует требованиям нормативно-технической документации для эксплуатации прибора данного типа, то такой режим называют *типовым*. Любой режим определяется совокупностью величин, которые называют параметрами. *Параметром режима* электронного прибора называется любая величина, характеризующая режим прибора; например, *электрические параметры режима* — это токи, напряжения и т. д. Различают *электрический режим*, который определяет величины напряжений на электродах и токи в их цепях, *механический режим* — это совокупность *механических воздействий* на работающий прибор (удары, тряска и т. п.), и *климатический режим*, который определяет интервал рабочих температур, относительную влажность окружающей среды, уровень радиации и т. д. Механические и климатические воздействия на электронные приборы определяют допустимые уровни влияния окружающей среды (см. гл. 15).

Упомянутые параметры электрического режима относятся к *функциональным*, к которым в зависимости от вида энергии сигнала относятся также *электромагнитные, световые, тепловые* и другие параметры. Численные значения функциональных параметров, установленные нормативно-технической документацией, называются *номинальными значениями*. Оптималь-

ные условия работы прибора при эксплуатации, испытаниях или измерениях его параметров определяются номинальным режимом. Кроме номинальных, для электронных приборов рассматривают *предельные параметры*, которые характеризуют предельно допустимые режимы работы прибора, например, *максимально допустимые значения напряжений на электродах* прибора, *максимально допустимую мощность*, рассеиваемую электродом или прибором в целом, и т. д.

К электродам электронных приборов подключаются источники как постоянных, так и переменных напряжений. Если прибор работает при постоянных напряжениях на электродах, то такой режим работы называют *статическим*. В этом случае все параметры режима не изменяются во времени. Режим работы прибора, при котором хотя бы один из параметров режима изменяется во времени, называется *динамическим*.

Иногда параметры прибора изменяются во времени настолько медленно по сравнению с временем движения носителей в рабочем пространстве, что в каждый момент времени они несущественно отличаются от статических, т. е. для прибора остаются практически справедливыми законы статического режима. В этом случае режим работы называют *квазистатическим*.

Кроме параметров режима, различают *параметры электронного прибора*, которыми называются физические величины, характеризующие общие свойства прибора (например, коэффициент усиления, внутреннее сопротивление, междуэлектродные емкости и т. д.).

Зависимость какого-либо параметра режима или параметра прибора, который принят в качестве функции, от другого параметра, взятого в качестве аргумента, при условии, что все остальные величины остаются неизменными, называется *характеристикой электронного прибора*. Совокупность характеристик при различных фиксированных значениях независимого третьего параметра называют *семейством характеристик*. Важнейшими характеристиками электронных приборов являются *статические характеристики*, отображающие зависимость тока в цепи какого-либо электрода от напряжения на любом электроде в статическом режиме. Часто названия статических характеристик связывают с названием электродов, например анодные, эмиттерные и т. д. Помимо этого, существуют также и обобщенные названия семейств статических характеристик, без указания названий электродов, например входные, выходные и т. д.

Краткий исторический очерк развития электронных приборов. История создания электронных приборов базируется на открытиях и исследованиях физических явлений, связанных с взаимодействием свободных электронов с электромагнитными полями и веществом. Поэтому первые работы М. В. Ломоносова, Г. В. Рихмана (Россия) и Б. Франклина (США) в конце XVIII в. по исследованию электричества могут быть отнесены к началу возникновения электроники. Открытие электрической дуги академиком В. В. Петровым в 1802 г. является началом технического использования электричества. Работы как отечественных, так и зарубежных ученых в течение XIX в. создали фундамент электроники. Среди наиболее важных достижений можно отметить труды А. Ампера и М. Фарадея, установивших законы движущегося электричества и электромагнитной индукции, создание теории электромагнетизма М. Максвеллом и теории электронов Г. А. Лоренцем, экспериментальное обнаружение электромагнитных волн Г. Герцем. Первый в мире электровакуумный прибор — лампа накаливания — изобретена А. И. Лодыгиным (Россия) и усовершенствована Т. А. Эдисоном (США). Важными работами для создания электронных приборов несомненно являются открытие фотоэлектронной эмиссии русским ученым А. Г. Столетовым и термоэлектронной эмиссии Т. А. Эдисоном. Большое влияние на развитие электроники оказали работы А. С. Попова, К. Ф. Брауна, Д. Томсона, О. У. Ричардсона, А. Эйнштейна и др., выполненные в конце XIX и начале XX в. В начале XX в. были изобретены электровакуумные диоды и триоды, газотроны. В 1907 г. русский ученый Б. Л. Розинг предложил использовать электронно-лучевую трубку для приема изображений, что может считаться началом телевидения. Много сделали для развития отечественной электроники русские ученые и инженеры В. И. Коваленков, А. Д. Папалекси, М. А. Бонч-Бруевич, О. В. Лосев. В советский период значительный вклад внесли А. А. Чернышев, который выдвинул идею создания видикона, Л. А. Кубецкий — изобретатель фотоэлектронного умножителя, А. П. Константинов и С. И. Котов — авторы приемной телевизионной трубки иконоскопа и др.

Революционные открытия были сделаны в послевоенный период (1940—1960-х гг.). В 1948 г. американские ученые Д. Бардин, У. Браттейн и У. Шокли предложили биполярный транзистор. В 1950-е г. изобретены: полевой транзистор с $p-n$ -переходом, солнечные батареи, оптроны, туннельные диоды, тиристоры и др.

В 1960 г. Д. Кинг и М. Аттала создали МОП-транзистор, а в 1966 г. С. Мид разработал полевой транзистор с барьером Шоттки.

В 1960—1970-е гг. большую роль в создании полупроводниковых приборов с гетеропереходами сыграли работы академика Ж. И. Алферова, за которые ему была присуждена в 2000 г. Нобелевская премия.

Что касается приборов квантовой электроники, то впервые вопрос о квантовом взаимодействии между светом и средой был рассмотрен в 1916 г. А. Эйнштейном, который показал, что между средой, состоящей из молекул (атомов), и светом постоянно происходит обмен энергией, сопровождающийся рождением (испусканием) одних и уничтожением (поглощением) других квантов света. При этом Эйнштейн впервые теоретически обосновал существование вынужденного излучения.

В 1939—1940 гг. при анализе спектра газового разряда В. А. Фабрикант указал на возможность усиления света за счет стимулированного (вынужденного, или индуцированного) излучения, теоретически сформулировав необходимые для этого условия. В 1950-х гг., продолжая свои работы, он вместе с сотрудниками впервые получил экспериментальное подтверждение своих расчетов и опубликовал эти результаты. В 1954—1955 гг. Н. Г. Басову и А. М. Прохорову (СССР), Ч. Таунсу, Дж. Гордону и Ж. Цайгеру (США) независимо друг от друга удалось осуществить усиление и генерацию СВЧ-волн на частоте 23 870 МГц, используя пучок молекул аммиака. В 1956 г. проф. Н. Бломберген (США) создал твердотельный трехуровневый мазер, работающий в непрерывном режиме.

В 1957—1958 гг. в СССР Н. Г. Басов и А. М. Прохоров, а в США Ч. Таунс разработали теоретические основы процессов, происходящих в лазерах. За эти работы они получили Нобелевскую премию. В 1960 г. Т. Мейманом (США) был создан первый лазер, работающий на рубине, что послужило толчком к дальнейшему прогрессу в области квантовой электроники.

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ

Глава 1

ФИЗИКА ПОЛУПРОВОДНИКОВ

1.1. Равновесная концентрация свободных носителей заряда

Все твердые тела в соответствии с их электрофизическими свойствами можно разделить на металлы, полупроводники и диэлектрики. Удельное сопротивление (ρ) различных твердых тел изменяется в весьма широких пределах: для металлов $\rho < 10^{-4}$ Ом·см, для полупроводников $\rho \sim 10^{-4} - 10^{10}$ Ом·см, для диэлектриков $\rho > 10^{10}$ Ом·см. Эти различия в значениях ρ обусловлены особенностями энергетической структуры для различных типов кристаллических твердых тел. Структуры энергетических состояний полупроводников и диэлектриков (рис. 1.1) принципиально не отличаются друг от друга, все отличия обусловлены только разницей в ширине запрещенной зоны (ΔE_g): в полупроводниках обычно $\Delta E_g \lesssim 3$ эВ, а в диэлектриках $\Delta E_g > 3$ эВ.

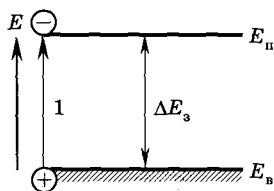


Рис. 1.1

Наибольшее применение в электронных приборах нашли полупроводниковые материалы, которые подразделяют на *собственные* (чистые, беспримесные) и *примесные*. Как в собственных, так и в примесных полупроводниках (энергети-

ческие диаграммы последних приведены на рис. 1.2) существуют два типа свободных носителей заряда — *электроны* и *дырки*. *Свободными носителями заряда* называются такие носители, кинетическая энергия которых больше их потенциальной энергии связи с атомами. Концентрация свободных носителей определяется двумя противоположными процессами — их *генерацией* и *рекомбинацией*. Генерация носителей заряда, т. е. образование свободных электронов и дырок, осуществляется при воздействии на полупроводник тепловой энергией, светом, ионизирующим облучением, пучками заряженных частиц и другими энергетическими факторами. В условиях термодинамического равновесия (при температурах $T > 0$ К) всегда присутствует тепловая генерация носителей, интенсивность которой увеличивается с ростом температуры. В собственном полупроводнике в процессе генерации образуются электронно-дырочные пары.

На энергетической диаграмме собственного полупроводника (см. рис. 1.1) этот процесс иллюстрируется стрелкой 1, которая показывает переход электрона из валентной зоны, верхняя граница которой соответствует энергии E_B , в зону проводимости (E_{π} — ее нижняя граница). В валентной зоне при переходе электрона в зону проводимости остается дырка. (Обозначим концентрацию электронов и дырок n и p соответственно.) Таким образом, в состоянии равновесия в собственном полупроводнике $n = p = n_i$, т. е.

$$n \cdot p = n_i^2, \quad (1.1)$$

где n_i — равновесная концентрация свободных носителей заряда в собственном полупроводнике при данной температуре.

В состоянии равновесия процессы генерации электронно-дырочных пар в собственном полупроводнике уравновешены об-

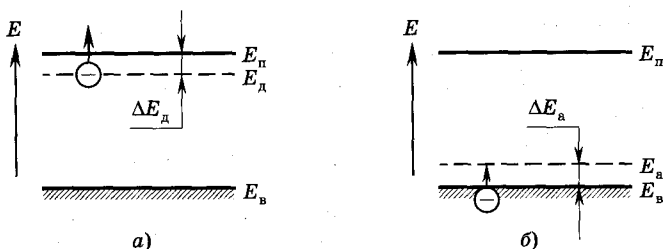


Рис. 1.2

ратными процессами рекомбинации. Равновесные концентрации электронов и дырок для собственного полупроводника с шириной запрещенной зоны ΔE_g могут быть вычислены согласно следующему выражению [1]:

$$n = p = n_i = \sqrt{N_{\text{п}} N_{\text{в}}} \exp[-\Delta E_g / (2kT)], \quad (1.2)$$

где $N_{\text{п}} = 2(2\pi \cdot m_n kT / h^2)^{3/2}$, $N_{\text{в}} = 2(2\pi \cdot m_p kT / h^2)^{3/2}$ — эффективные плотности энергетических состояний в зоне проводимости и в валентной зоне соответственно; m_n и m_p — эффективные массы электронов и дырок; $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана; $h \approx 6,6 \cdot 10^{-34}$ Дж·с — постоянная Планка; T — температура в градусах Кельвина (К).

В выражении (1.2) экспоненциальный множитель обуславливает резкое увеличение концентрации свободных носителей заряда при возрастании температуры T или уменьшении ширины запрещенной зоны ΔE_g . Влияние ширины запрещенной зоны на концентрацию носителей в собственных полупроводниках можно проиллюстрировать на примере кремния (Si) и арсенида галлия (GaAs), наиболее широко используемых в полупроводниковой технике: при $T = 300$ К $\Delta E_g = 1,12$ эВ для Si и $\Delta E_g = 1,42$ эВ для GaAs, а концентрация собственных носителей соответственно — $1,4 \cdot 10^{10}$ и $1,8 \cdot 10^6$ см⁻³. Этот пример показывает, что отличие ширины запрещенной зоны всего лишь в 1,27 раза приводит к изменению концентрации носителей на четыре порядка.

Примесные полупроводники могут быть **донорными, акцепторными и компенсированными**. В донорных полупроводниках, или в **полупроводниках *n*-типа** (они содержат пятивалентную донорную примесь, например, фосфор или мышьяк для кремния), преобладает электронная проводимость. Это значит, что концентрация свободных электронов n_{n0} , которые в данном случае называются **основными носителями**, в равновесном состоянии при не слишком высоких температурах T (таких, что $kT \ll E_g$) на много порядков превышает концентрацию собственных носителей n_i и дырок p_{n0} , которые в данном случае являются **неосновными носителями**.

При не слишком высоких температурах подавляющее число электронов в полупроводнике *n*-типа возникает из-за тепловой ионизации донорных атомов; в результате донорные атомы превращаются в положительно заряженные ионы, а электроны, оторванные от них, становятся свободными носителями заряда.

На рис. 1.2, а этот процесс иллюстрируется стрелкой и соответствует переходу электрона с донорного уровня E_d в зону проводимости. Уровень E_d образуется атомами донорной примеси в запрещенной зоне. Разность энергий $\Delta E_d = E_n - E_d$ равна энергии ионизации доноров. Из-за малости энергии ионизации (сотые доли электрон-вольта и менее) при комнатной температуре ($T = 300$ К; $kT = 0,026$ эВ) практически все атомы доноров ионизованы и концентрация основных носителей (электронов в данном случае) равна концентрации доноров $n_{n0} \approx N_d$, а концентрация неосновных носителей (дырок) определяется **законом действующих масс** $n_{n0} \cdot p_{p0} = n_i^2$ и равна

$$p_{p0} = n_i^2 / N_d. \quad (1.3)$$

В состоянии равновесия в примесных полупроводниках, так же как и в собственных, протекают одновременно процессы генерации и рекомбинации свободных носителей. В результате устанавливаются равновесные концентрации электронов и дырок. Используя выражения (1.2) и (1.3), концентрацию неосновных носителей (дырок) в донорном полупроводнике в состоянии равновесия можно определить по следующей формуле:

$$p_{p0} = (N_n \cdot N_v / N_d) \exp[-\Delta E_a / (kT)]. \quad (1.4)$$

При введении в полупроводник акцепторной примеси с концентрацией $N_a \gg n_i = p_i$ в нем будет преобладать дырочная проводимость. Такой полупроводник называют **дырочным** или **полупроводником p-типа**. Дырки в этом случае возникают за счет ионизации акцепторных атомов, т. е. в результате присоединения к ним электронов, возникающих при разрыве связей в атомах собственного полупроводника.

На энергетической диаграмме (см. рис. 1.2, б) описанный процесс соответствует переходу электрона из валентной зоны на акцепторный уровень E_a , расположенный в запрещенной зоне вблизи потолка E_v валентной зоны. В результате в валентной зоне образуются свободные уровни, а акцепторный атом превращается в отрицательный ион. Аналогично донорному полупроводнику в акцепторном из-за малости энергии ионизации при комнатной температуре практически все акцепторные атомы ионизованы и концентрация основных носителей p_{p0} (в данном случае дырок) равна концентрации акцепторов N_a , т. е. $p_{p0} \approx N_a$. Равновесную

концентрацию неосновных носителей — электронов n_{p0} — определим из аналогичного формуле (1.3) соотношения

$$n_{p0} = n_i^2 / N_a. \quad (1.5)$$

С учетом (1.2) оно приводит к выражению, «симметричному» формуле (1.4):

$$n_{p0} = (N_{\text{II}} N_{\text{V}} / N_a) \exp [-\Delta E_3 / (kT)]. \quad (1.6)$$

В полупроводниковых приборах концентрация доноров N_d и акцепторов N_a изменяется в широких пределах от 10^{13} до 10^{21} см⁻³. При большой концентрации примесных атомов из-за сильного взаимодействия между ними примесные уровни (E_d или E_a) расщепляются на подуровни, в результате чего образуется примесная зона, которая при концентрациях N_a , N_d более 10^{20} см⁻³ перекрывается с зоной проводимости для донорных полупроводников и с валентной зоной для акцепторных полупроводников. При перекрытии примесных уровней с зоной проводимости или с валентной зоной энергия ионизации примеси уменьшается до нуля и возникает частично заполненная зона. Как и в металлах, в этом случае в полупроводниках проводимость существует и при $T = 0$ К. Такие полупроводники называются *вырожденными*.

В реальных условиях в полупроводниках обычно имеются как донорные, так и акцепторные примеси. Если $N_d > N_a$, получается полупроводник n -типа, а при $N_a > N_d$ — полупроводник p -типа. При этом в первом случае важна эффективная концентрация доноров $N_d - N_a$, а во втором случае — эффективная концентрация акцепторов $N_a - N_d$. При $N_a = N_d$ полупроводник называется *компенсированным*. В нем концентрация свободных носителей такая же, как и в собственном полупроводнике.

Атомы некоторых примесей могут образовывать энергетические уровни в запрещенной зоне на значительном удалении от E_{II} и E_{V} ; такие атомы называются *ловушками*. Энергетические уровни, соответствующие донорным ловушкам, расположены выше середины запрещенной зоны, а акцепторные — ниже. Донорная ловушка нейтральна, если соответствующий ей энергетический уровень заполнен (занят электроном), и превращается в положительный ион, если уровень свободен. Акцепторные ловушки нейтральны при свободном уровне и заряжены отрицательно (отрицательные ионы) при его заполнении.

Температурная зависимость концентрации свободных носителей заряда. Концентрация носителей в примесных полупроводниках, так же как и в собственных, существенно зависит от температуры. Рассмотрим температурную зависимость концентрации электронов в кремнии на примере полупроводника n -типа (рис. 1.3). На ней можно выделить три области. При низких температурах (область 1) с ростом температуры концентрация свободных электронов ($n \approx n_n$) увеличивается, так как возрастает число ионизированных доноров. Зависимость концентрации электронов от $1/T$ определяется экспоненциальной функцией вида $\exp[-\Delta E_d/(2kT)]$, поэтому в полулогарифмическом масштабе она изображается прямой линией, тангенс угла наклона α которой пропорционален энергии ионизации доноров ΔE_d . В области 2 почти все доноры ионизованы, а концентрация собственных электронов n_i незначительна, поэтому с ростом температуры полное число свободных электронов изменяется не существенно, и их концентрацию можно считать равной концентрации доноров: $n \approx n_{n0} \approx N_d$. В области высоких температур (область 3) происходит интенсивная ионизация собственных атомов полупроводника, так что концентрация собственных носителей становится больше концентрации основных примесных носи-

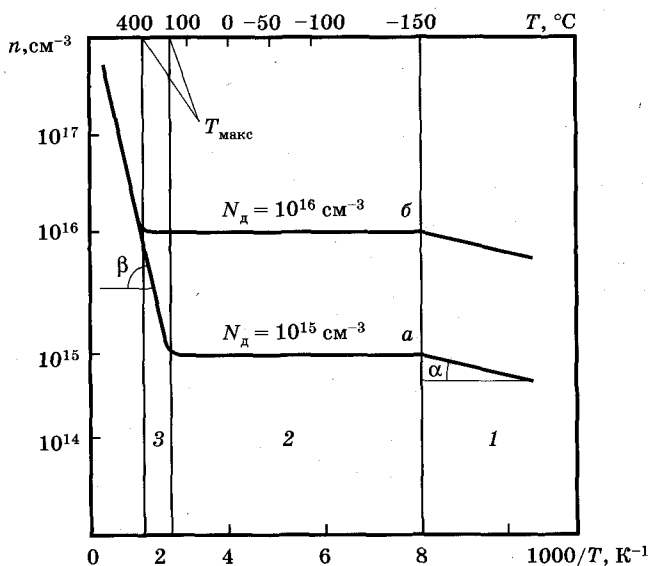


Рис. 1.3

телей, т. е. $n_i > n_{n0} \approx N_d$. В рассматриваемой области концентрация носителей определяется зависимостью $n \approx n_i \sim \exp(-\Delta E_g/(2kT))$, которая в полулогарифмическом масштабе изображается прямой линией с углом наклона β , причем $\text{tg } \beta$ пропорционален ширине запрещенной зоны ΔE_g .

Увеличение концентрации примесей приводит не только к увеличению концентрации основных носителей, но и к пропорциональному уменьшению концентрации неосновных, в соответствии с выражениями (1.3) и (1.5), что связано с увеличением вероятности их рекомбинации, пропорциональной произведению отмеченных концентраций.

Большинство полупроводниковых приборов нормально работает в температурном интервале, соответствующем области 2 на рис. 1.3. Максимальная температура в этой области $T_{\text{макс}}$ приближенно определяется из условия $n_i = N_d$ (для полупроводника n -типа). Она пропорциональна ширине запрещенной зоны и увеличивается с возрастанием концентрации примесей (см. рис. 1.3, кривые $a, б$).

Концентрация неосновных носителей в области 2, в отличие от концентрации основных носителей, сильно увеличивается с ростом температуры согласно выражениям (1.4) и (1.6) соответственно для электронного полупроводника (где дырки — неосновные носители) и для дырочного (неосновные носители — электроны). Параметры приборов, которые зависят от концентрации неосновных носителей, также будут изменяться с температурой даже в области полной ионизации примесей (область 2 на рис. 1.3), и максимальная рабочая температура таких приборов может быть заметно ниже температуры, определяемой условиями $n_i = N_d$ или $n_i = N_a$ (для электронного или дырочного полупроводников).

Уровень Ферми. Свободные носители в твердом теле заполняют энергетические состояния с существенно различной вероятностью. Согласно квантовой статистике вероятность заполнения электроном энергетического уровня с энергией E определяется функцией Ферми—Дирака $F(E)$, которая вычисляется согласно следующей формуле:

$$F(E) = 1/[1 + \exp(E - E_F)/(kT)], \quad (1.7)$$

где E_F — энергия, соответствующая уровню Ферми. В любой равновесной системе, какой бы разнородной она ни была, уровень Ферми одинаков для всех ее частей. Как показывают вычисле-

ния, в собственном полупроводнике при $m_n = m_p$ уровень Ферми лежит посередине запрещенной зоны $E_{\Phi} = E_{\Phi_i} = 0,5(E_{\Pi} + E_{\text{В}})$. В невырожденном полупроводнике n -типа ($N_{\Pi} \gg n \gg n_i$) уровень Ферми E_{Φ_n} расположен ближе к зоне проводимости, а в невырожденном полупроводнике p -типа уровень Ферми E_{Φ_p} расположен ближе к валентной зоне. При комнатной температуре ($T = 300 \text{ К}$) он лежит, как правило, ниже уровня доноров и выше уровня акцепторов для полупроводников n - и p -типа соответственно. Если в примесных полупроводниках уровень Ферми лежит в запрещенной зоне на расстоянии не менее $(2...3)kT$ от соответствующей ее границы, то концентрации электронов и дырок будут равны [2, 3]:

$$n = N_{\Pi} \exp [-(E_{\Pi} - E_{\Phi_n})/(kT)], \quad (1.7, \text{а})$$

$$p = N_{\text{В}} \exp [-(E_{\Phi_p} - E_{\text{В}})/(kT)]. \quad (1.7, \text{б})$$

С ростом температуры в примесном полупроводнике (при $m_n \approx m_p$) уровень Ферми приближается к середине запрещенной зоны, так как при этом начинает преобладать собственная проводимость над примесной. Зависимость положения уровня Ферми от температуры для кремния с различной концентрацией донорной и акцепторной примеси показана на рис. 1.4, где $E = E_{\Phi} - E_{\text{В}}$.

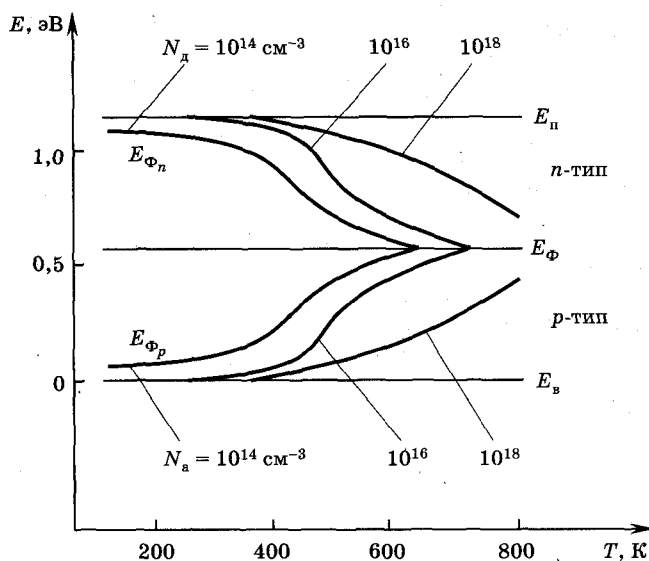


Рис. 1.4

Если $n = N_n$ или $p = N_p$ (вырожденный полупроводник), т. е. концентрация носителей соизмерима с концентрацией разрешенных состояний, то, в силу принципа Паули, электроны не могут произвольно занимать энергетические уровни. Уровень Ферми в этом случае лежит либо в запрещенной зоне на расстоянии менее $(2...3)kT$ от ее границ, либо в зоне проводимости для n -полупроводника или в валентной зоне для p -полупроводника. Для сильно вырожденных полупроводников положение уровня Ферми, так же как и концентрация основных носителей, не зависят от температуры.

1.2. Неравновесные носители заряда

Неравновесные носители в полупроводниках могут возникать под воздействием внешних электрических полей, неионизирующего электромагнитного излучения (включая оптический диапазон), ионизирующего излучения и других энергетических факторов. Указанные факторы вызывают переход электронов из валентной зоны в зону проводимости. Помимо этого, под воздействием электрического поля может происходить ударная ионизация атомов, когда электрон или дырка, ускоряясь в сильном электрическом поле до энергии, достаточной для процесса ионизации, сталкивается с атомом, что вызывает рождение электронно-дырочной пары. Часто такой процесс приводит к пробоем электрических переходов (см. гл. 2). Во всех этих случаях неравновесные носители заряда являются избыточными над равновесными носителями при данной температуре.

В большинстве типов полупроводниковых приборов наиболее распространенным механизмом создания неравновесных носителей является *инжекция* их из одной полупроводниковой области в другую под действием электрического поля, например, инжекция электронов или дырок через электрический p — n -переход (см. п. 2.2). Обычно в этих случаях интерес представляет поведение неосновных носителей. Поэтому, если превышение концентрации неравновесных носителей (Δn_p или Δp_n) над равновесными концентрациями основных носителей мало, т. е. если $\Delta n_p = n_p - n_{p0} \ll p_{p0} \approx N_a$ или $\Delta p_n = p_n - p_{n0} \ll n_{n0} \approx N_d$, то изменение избыточных концентраций Δn_p или Δp_n в областях полупроводника описывается уравнением генерации-рекомбинации. Для n -области оно имеет вид

$$d(\Delta p_n)/dt = -(\Delta p_n/\tau_p) + G, \quad (1.8)$$

где τ_p — время жизни неравновесных неосновных носителей (среднее время от момента появления неравновесного носителя до его рекомбинации); G — скорость их генерации (определяет число неравновесных носителей, возникающих в единицу времени в единице объема, измеряется в $\text{см}^{-3} \cdot \text{с}^{-1}$); $\Delta p_n / \tau_p$ — число рекомбинирующих носителей в единице объема в единицу времени. Для p -области можно записать аналогичное уравнение с соответствующей заменой обозначений.

Если рассматривать процесс после прекращения ионизации, когда $G = 0$, то решение уравнения (1.8) имеет вид

$$\Delta p_n(t) = \Delta p_n(0) \exp(-t/\tau_p), \quad (1.9)$$

где $\Delta p_n(0)$ — избыточная концентрация дырок в момент $t = 0$, когда прекращаются внешние воздействия и G становится равным нулю. Зависимость (1.9), когда $t = \tau_p$, позволяет определить время жизни как интервал, в течение которого избыточная концентрация уменьшается в e раз. Время жизни неосновных носителей τ_p для n -полупроводника и τ_n для p -полупроводника характеризует скорость изменения концентрации, поэтому быстрдействие большинства полупроводниковых приборов зависит от этого параметра. Если в начальный момент времени $t = 0$ $\Delta p_n(0) = 0$ и начал действовать внешний энергетический источник, вызывающий постоянную скорость генерации G , тогда решение уравнения (1.8) можно представить в следующем виде:

$$\Delta p_n(t) = G\tau_p[1 - \exp(-t/\tau_p)]. \quad (1.10)$$

В уравнении (1.10) параметр τ_p определяет скорость нарастания избыточной концентрации, конечное установившееся значение которой равно $G\tau_p$. Таким образом, уравнение (1.9) описывает уменьшение концентрации избыточных носителей за счет рекомбинации, а уравнение (1.10) увеличение (нарастание) избыточной концентрации за счет генерации.

При рекомбинации происходит переход электронов из зоны проводимости в валентную зону с выделением энергии, величина которой равна ширине запрещенной зоны. Переход электронов из зоны в зону может происходить либо непосредственно из зоны проводимости в валентную зону, либо ступенчато в несколько стадий через центры рекомбинации (*ловушки*), уровни которых расположены в запрещенной зоне. В первом случае при непосредственной межзонной рекомбинации электрон встреча-

ется сразу с дыркой, а во втором случае электрон сначала встречается с ловушкой, захватывается ею, переходя на уровень ловушки, затем происходит захват дырки ловушкой, что соответствует переходу электрона с уровня ловушки в валентную зону.

Если в процессе рекомбинации энергия выделяется в виде электромагнитного излучения, то рекомбинация называется *излучательной*. Если же энергия рекомбинации передается кристаллической решетке с образованием *фононов* (акустических квантов энергии) или же непосредственно другим электронам или дыркам, то рекомбинация будет *безызлучательной*.

При рекомбинации на ловушках в полупроводниках *p*-типа первая стадия заключается в переходе электронов из зоны проводимости на уровни ловушек, которые в данном случае почти все свободны, поскольку уровень Ферми расположен вблизи валентной зоны. Таким образом, в *p*-полупроводнике происходит захват неосновного носителя (электрона) ловушкой. Этот процесс является медленным из-за малой концентрации электронов. Вторая стадия процесса рекомбинации — переход электрона с уровня ловушки в валентную зону является быстрым процессом, так как он определяется столкновением дырок с ловушками, а концентрация дырок в полупроводнике *p*-типа велика. Следовательно, скорость рекомбинации носителей и их время жизни τ_n определяются первой стадией процесса, причем величина τ_n обратно пропорциональна концентрации свободных ловушек. С ростом температуры время жизни носителей увеличивается. Итак, время жизни неравновесных носителей в *p*-полупроводниках определяется временем жизни τ_n неосновных носителей — электронов, а в полупроводниках *n*-типа — временем жизни τ_p неосновных носителей — дырок. Таким образом, в полупроводнике с любым типом проводимости первой и основной стадией рекомбинации, от которой зависит время жизни неравновесных носителей, является захват неосновного носителя ловушкой. Для изменения быстродействия полупроводниковых приборов и интегральных схем используется введение специальных примесей, создающих уровни ловушек, являющихся центрами рекомбинации. Например, в кремнии для этого используются атомы золота, создающие два уровня вблизи середины запрещенной зоны, что позволяет изменять время жизни неравновесных носителей в пределах 10^{-3} — 10^{-2} с за счет изменения концентрации атомов золота.

1.3. Электропроводность полупроводников

В твердых телах свободные носители при своем движении непрерывно испытывают столкновения с атомами и ионами примесей и различными дефектами, в результате происходит рассеяние свободных носителей. При наличии электрического поля равноускоренное движение носителей происходит только в промежутках между столкновениями. После каждого столкновения свободная частица начинает заново ускоряться. При рассеянии изменяются энергия и направление движения носителей, что сопровождается возбуждением или поглощением фононов. Средняя скорость упорядоченного движения электронов и дырок, направленная вдоль электрического поля, которая называется *дрейфовой скоростью* ($v_{др}$), пропорциональна напряженности электрического поля \mathcal{E} :

$$v_{др} = \mu \mathcal{E}. \quad (1.11)$$

Коэффициент пропорциональности μ между дрейфовой скоростью $v_{др}$ и напряженностью электрического поля \mathcal{E} называется подвижностью и имеет размерность $m^2/(V \cdot c)$.

Из-за различия эффективных масс электронов m_n и дырок m_p их подвижности могут сильно различаться. Так, например, в кремнии подвижность электронов почти в три раза больше подвижности дырок. При большей подвижности носителей будет выше быстродействие полупроводниковых приборов. Подвижность носителей в полупроводниках зависит от многих факторов, важнейшими из которых являются температура, концентрация примесей, а также и напряженность электрического поля. Зависимость подвижности от температуры T определяется рассеянием носителей на узлах кристаллической решетки и на ионах примеси. При рассеянии на ионах примеси температурная зависимость подвижности имеет вид

$$\mu_{ион} = \mu_{0\ ион} (T_0/T)^{3/2}. \quad (1.12)$$

Если преобладает рассеяние на атомах решетки, то

$$\mu_{реш} = \mu_{0\ реш} (T_0/T)^{-a}. \quad (1.13)$$

В формулах (1.12) и (1.13) значения T_0 — некоторые начальные (исходные) температуры, при которых подвижность равна μ_0 . Показатель степени « a » зависит от материала и типа проводимости; так, для кремния n -типа (p -типа) $a = 2,42$ ($a = 2,2$); соответственно для Ga As $a = 1$ ($a = 2,1$).

При учете обоих из указанных выше процессов рассеяния носителей выражение для подвижности имеет следующий вид:

$$(1/\mu) = (1/\mu_{\text{ион}}) + (1/\mu_{\text{реш}}). \quad (1.14)$$

Рассеяние носителей на тепловых колебаниях кристаллической решетки полупроводника в основном доминирует при сравнительно высоких температурах и малых концентрациях примеси. При больших температурах из-за возрастания амплитуды тепловых колебаний атомов решетки увеличивается частота столкновений с ними носителей заряда, что приводит к более интенсивному их рассеянию. Таким образом, с ростом температуры подвижность носителей уменьшается вследствие их рассеяния на тепловых колебаниях решетки.

Рассеяние носителей на ионах примеси превалирует при низких температурах и высоких концентрациях примеси. Если температура мала, то тепловая скорость электрона также невелика, поэтому при движении относительно иона примеси траектория электрона сильно изменяется случайным образом, что и соответствует большему рассеянию. С ростом температуры из-за увеличения скорости электронов время их кулоновского взаимодействия с ионами примеси уменьшается, что приводит к меньшему рассеянию носителей. Итак, при рассеянии на ионах примеси подвижность носителей растет с увеличением температуры в соответствии с зависимостью $\mu_{\text{ион}} \sim T^{3/2}$.

При постоянной температуре результирующая подвижность но-

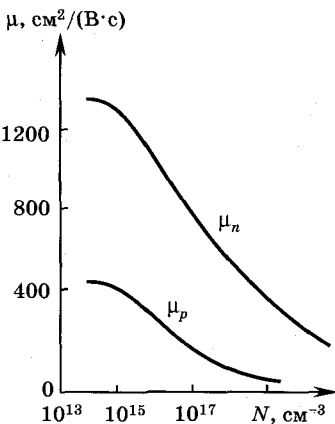


Рис. 1.5

сителей существенно зависит от концентрации примеси. Если она мала, то, как отмечалось выше, основную роль, как и в беспримесных полупроводниках, играет рассеяние на тепловых колебаниях решетки (фононах), при этом подвижность имеет максимальное значение. С ростом концентрации примеси N начинает преобладать рассеяние на ионах примеси, вследствие чего подвижность носителей уменьшается (рис. 1.5), где μ_n , μ_p — соответственно подвижности электронов и дырок.

Зависимость подвижности от напряженности электрического по-

ля носит неоднозначный характер. В слабых электрических полях, когда напряженность поля $\mathcal{E} < 10^3 \dots 10^4$ В/см, дрейфовая скорость оказывается много меньше тепловой скорости электронов. При этом поле не оказывает существенного влияния на рассеяние носителей, подвижность μ не зависит от величины \mathcal{E} , и дрейфовая скорость $v_{др}$, в соответствии с формулой (1.11), линейно возрастает с ростом напряженности поля \mathcal{E} (рис. 1.6).

В сильных электрических полях ($\mathcal{E} \geq 10^4 \dots 10^5$ В/см) скорость дрейфа носителей приближается к средней тепловой скорости, а средняя энергия направленного движения электронов становится достаточной для увеличения амплитуды колебаний атомов в узлах кристаллической решетки полупроводника, что приводит к увеличению частоты столкновений и, соответственно, к интенсификации процесса рассеяния носителей. В результате с ростом \mathcal{E} подвижность носителей падает при постоянстве дрейфовой скорости (см. рис. 1.6). В арсениде галлия зависимость подвижности более сложная по сравнению с Si и Ge, что связано с особенностями зоны проводимости. Эти особенности анализируются в последующих главах при рассмотрении процессов в некоторых типах оптоэлектронных приборов и лазеров, а также диодов Ганна.

Зная особенности поведения подвижности и концентрации в зависимости от различных физических и электрофизических условий, можно проанализировать такой важный параметр,

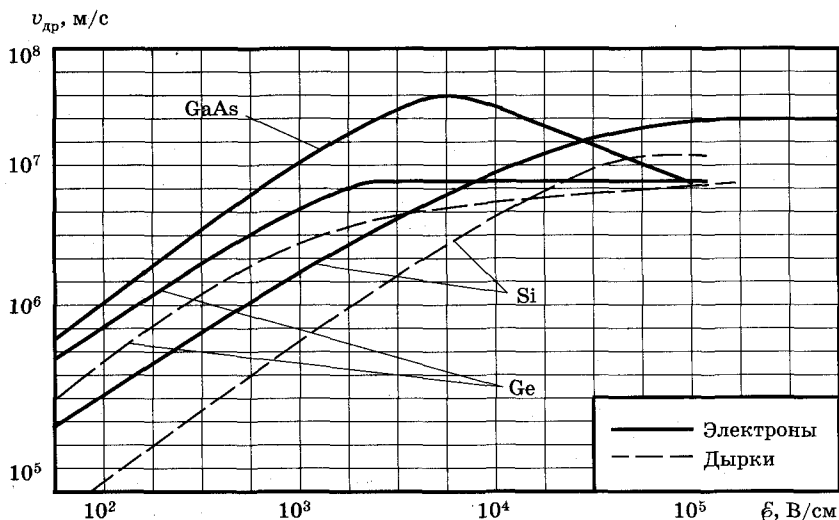


Рис. 1.6

как *электропроводность* полупроводников. Согласно обобщенному закону Ома плотность дрейфового тока $j_{др}$, протекающего через среду, напряженность поля \mathcal{E} и удельная электрическая проводимость среды σ связаны соотношением $j_{др} = \sigma \mathcal{E}$. Плотность дрейфового тока в зависимости от концентрации носителей заряда и их дрейфовой скорости $v_{др}$ равна $j_{др} = qNv_{др}$, где q — заряд частицы, N — концентрация частиц определенного вида.

Для электронов $N = n$, а для дырок $N = p$. При наличии в полупроводнике зарядов обоого вида плотность тока будет определяться суммой токов положительных и отрицательных зарядов, т. е.

$$j_{др} = q(nv_{др, n} + pv_{др, p}),$$

где $v_{др, n}$, $v_{др, p}$ — соответственно дрейфовая скорость электронов и дырок. Из этого соотношения и обобщенного закона Ома с учетом формулы (1.1) следует, что удельную электрическую проводимость полупроводников можно записать в следующем виде:

$$\sigma = q(n\mu_n + p\mu_p). \quad (1.15)$$

Зная температурные характеристики концентрации и подвижности свободных носителей, можно объяснить экспериментальную зависимость σ от температуры. Для примесных полупроводников n -типа на основе Ge и Si эти зависимости приведены на рис. 1.7. При температурах $T > T_{макс}$ электропроводность определяется произведением концентрации и подвижности *собственных* носителей, причем с ростом T подвижность носителей

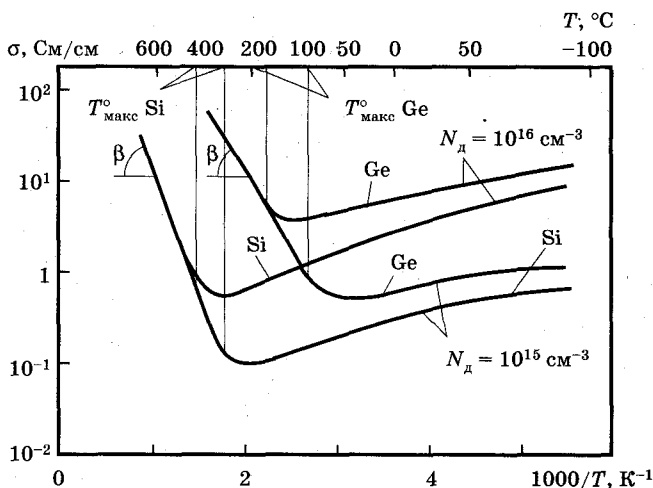


Рис. 1.7

уменьшается, а концентрация экспоненциально возрастает. Уменьшение подвижности $\mu(T)$ в этой области изменения температуры гораздо слабее, чем рост концентрации $n(T)$. В результате при $T > T_{\text{макс}}$ электропроводность полупроводника растет с увеличением температуры, и в полулогарифмическом масштабе функция $\sigma = f(1/T)$ является практически линейной с тангенсом угла наклона $\text{tg } \beta \sim \Delta E_g$. При температурах $T < T_{\text{макс}}$ концентрация собственных носителей мала ($n_i \ll N_d, N_a$) и проводимость определяется в основном концентрацией примесных носителей.

Так, например, в донорном полупроводнике при температурах, характерных для области работы большинства полупроводниковых приборов, когда все атомы примеси ионизованы, удельная электропроводность равна

$$\sigma = qn\mu_n \approx qN_d\mu_n.$$

На этом участке изменения функции $\sigma = f(1/T)$ (см. рис. 1.7) незначительное уменьшение электропроводности примесных полупроводников с ростом температуры связано лишь с уменьшением подвижности носителей при их практически неизменной концентрации.

В Ge зависимость подвижности носителей от температуры слабее, чем в Si, а величина μ в Ge больше, чем в Si. Это иллюстрирует рис. 1.7, где показаны зависимости $\sigma = f(1/T)$ для двух различных концентраций примеси N_d в Si и Ge.

1.4. Законы движения носителей заряда в полупроводниках

Направленное движение носителей заряда в полупроводниках вызвано двумя причинами: *диффузией* и *дрейфом* под действием электрического поля. Диффузия происходит из-за градиента концентрации зарядов, а наличие градиента потенциала вызывает дрейф носителей (см. п. 1.3) в направлении вектора напряженности электрического поля или против него, в зависимости от знака заряда. Наличие двух типов носителей приводит к тому, что полный ток состоит из четырех составляющих:

$$j = (j_n)_{\text{диф}} + (j_p)_{\text{диф}} + (j_n)_{\text{др}} + (j_p)_{\text{др}}. \quad (1.16)$$

Здесь j — плотность полного тока, а индексы « n » и « p » относятся соответственно к электронным и дырочным составляющим плотностям диффузионного и дрейфового токов $j_{\text{диф}}$ и $j_{\text{др}}$.

Для простоты рассмотрим одномерный случай, т. е. будем считать, что движение носителей заряда происходит только

вдоль оси x , тогда с учетом их знаков можно записать следующие выражения:

$$(j_n)_{др} = qn\mu_n \mathcal{E} = -qn\mu_n (d\phi/dx), \quad (1.17)$$

$$(j_p)_{др} = qp\mu_p \mathcal{E} = -qp\mu_p (d\phi/dx), \quad (1.18)$$

$$(j_n)_{диф} = q\mu_n \Phi_T \frac{dn}{dx} = qD_n \frac{dn}{dx}, \quad (1.19)$$

$$(j_p)_{диф} = -q\mu_p \Phi_T \frac{dp}{dx} = -qD_p \frac{dp}{dx}, \quad (1.20)$$

где $\Phi_T = \frac{kT}{q}$ — тепловой потенциал, $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана, D_n и D_p — коэффициенты диффузии электронов и дырок соответственно. Коэффициенты D_n и D_p аналогичны по своей роли подвижностям μ_n и μ_p при дрейфовом механизме движения. Связь между подвижностями и коэффициентами диффузии определяется формулой Эйнштейна

$$D = \Phi_T \mu. \quad (1.21)$$

Из сравнения выражений (1.17), (1.18) с (1.19), (1.20) вытекает, что дрейфовые составляющие токов пропорциональны концентрации носителей, а диффузионные определяются градиентами концентраций соответствующих носителей. Как видно из приведенных выражений, для вычисления токов необходимо знать распределения концентрации носителей $n(x)$ и $p(x)$. Поскольку концентрации могут зависеть не только от координаты, но и от времени, то они могут быть вычислены на основе решения уравнения непрерывности, которое вытекает из уравнения Больцмана (приложение 2, уравнение (П2.1)).

Для электронов и дырок с учетом знака заряда уравнения непрерывности можно записать в следующем виде:

$$\frac{\partial n}{\partial t} = -\frac{n - n_{p0}}{\tau_n} + \frac{1}{q} \operatorname{div} (j_n) + G, \quad (1.22)$$

$$\frac{\partial p}{\partial t} = -\frac{p - p_{n0}}{\tau_p} - \frac{1}{q} \operatorname{div} (j_p) + G, \quad (1.23)$$

где j_n , j_p — электронная и дырочная составляющие плотности полного тока.

Первый член в правых частях уравнений (1.22) и (1.23) дает изменение концентрации соответственно электронов и дырок из-за рекомбинации. Второе слагаемое определяет изменения

концентрации носителей в элементарном объеме dV из-за поступления в этот объем или ухода из него носителей. Величина G характеризует генерацию носителей. Поскольку уравнения непрерывности для электронов и дырок аналогичны, то дальнейший анализ будет проведен только для электронов, плотность тока которых согласно (1.17) и (1.19) равна

$$j_n = qn\mu_n\mathcal{E} + qD_n\left(\frac{\partial n}{\partial x}\right). \quad (1.24)$$

Подставляя (1.24) в (1.22), получим (для одномерного случая):

$$\frac{\partial n}{\partial t} = G - \frac{n - n_{p0}}{\tau_n} + D_n \frac{\partial^2 n}{\partial x^2} + \mu_n \mathcal{E} \frac{\partial n}{\partial x} + n\mu_n \frac{\partial \mathcal{E}}{\partial x}. \quad (1.25)$$

Последнее слагаемое в правой части уравнения (1.25) связано с наличием объемных зарядов внутри полупроводника. В условиях его электронейтральности $\frac{\partial \mathcal{E}}{\partial x} = 0$ и соответствующее слагаемое выпадает. Член $\mu_n \mathcal{E} \frac{\partial n}{\partial x}$ необходимо учитывать, например, в случае наличия внутреннего поля в неоднородных полупроводниках (см. гл. 7).

Если внутри полупроводника электрическое поле и генерация зарядов отсутствуют, то уравнение непрерывности в стационарном случае ($\partial n / \partial t = 0$) вырождается в уравнение диффузии

$$D_n(\partial^2 n / \partial x^2) - (n - n_{p0}) / \tau_n = 0. \quad (1.26)$$

Уравнение диффузии (1.26) описывает диффузионное движение электронов в дырочном полупроводнике с учетом рекомбинации.

При анализе работы полупроводниковых приборов часто основной интерес представляют только избыточные (неравновесные) концентрации носителей. Предположим, что в полупроводнике в области, примыкающей к некоторой плоскости $x = 0$, создается избыточная концентрация носителей $\Delta n_p(0) = \Delta p_p(0)$. Это можно реализовать, например, за счет освещения поверхности полупроводника. В результате возникает диффузия созданных избыточных носителей из области $x = 0$ в глубь полупроводника. Вследствие рекомбинации концентрация избыточных носителей будет уменьшаться по мере их продвижения в глубь полупроводника и при $x \rightarrow \infty$ $\Delta n_p(x) = \Delta p_p(x) \rightarrow 0$. Решение уравнения диффузии (1.26) для указанных граничных условий имеет вид

$$\Delta n_p(x) = \Delta n_p(0) \exp(-x/L_n). \quad (1.27)$$

Параметр L_n в формуле (1.27), называемый *диффузионной длиной* электронов в дырочном полупроводнике, определяется соотношением

$$L_n = \sqrt{D_n \tau_n}. \quad (1.28)$$

Для дырок получаются аналогичные результаты. Диффузионная длина характеризует среднее расстояние, на которое носители успевают перемещаться за время жизни. Отношение диффузионной длины к времени жизни носителей (L/τ) определяет среднюю скорость диффузии носителей. Для кремния типичные значения диффузионной длины в зависимости от времени жизни носителей составляют величины порядка 5—20 мкм. Из выражения (1.27) следует, что на расстоянии диффузионной длины, т. е. при $x = L_n$, избыточная концентрация уменьшается в $e \approx 2,718$ раз; на расстоянии $x = (3...4)L_n$ она падает в 20—50 раз, т. е. становится пренебрежимо малой по сравнению с граничной. Зная градиент концентрации избыточных носителей, вычислим плотность тока диффузии при $x = 0$:

$$(j_n)_{\text{диф}}(0) = qD_n \left. \frac{dn_p}{dx} \right|_{x=0} = \frac{qD_n \Delta n_p(0)}{L_n}. \quad (1.29)$$

При продвижении в глубь полупроводника плотность диффузионного тока уменьшается из-за рекомбинации, как и концентрация свободных носителей.



Контрольные вопросы

1. Что представляют собой зонные диаграммы и какова концентрация свободных носителей в собственных и примесных полупроводниках?
2. Какой вид имеет температурная зависимость концентрации носителей в собственных и примесных полупроводниках?
3. Что такое уровень Ферми и как зависит его положение от температуры в примесных полупроводниках?
4. Что называется подвижностью носителей и какова ее температурная зависимость в полупроводниках?
5. Дрейфовое и диффузионное движение носителей. Коэффициент диффузии.
6. Электропроводность полупроводников и ее зависимость от температуры.

7. Неравновесные носители в полупроводниках. Время жизни неравновесных носителей.
8. Основные уравнения, описывающие явления переноса носителей в полупроводниках.

Глава 2

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ. ЭЛЕКТРИЧЕСКИЕ ПЕРЕХОДЫ

2.1. Основные определения. Классификация электрических переходов

Основная масса дискретных полупроводниковых приборов, основных элементов интегральных схем, оптоэлектронных приборов и устройств других типов представляют собой сугубо неоднородные структуры. При контакте двух полупроводников с различными электрофизическими параметрами или полупроводников с металлами и диэлектриками в пограничных контактных слоях возникают электрические потенциальные барьеры, и концентрации носителей заряда внутри этих слоев могут сильно изменяться по сравнению с их значениями в объеме. Эти контактные слои называют **электрическими переходами**. Электрические переходы используются практически во всех полупроводниковых приборах. Физические процессы в электрических переходах являются определяющими в большинстве приборов.

Существует заметное число разновидностей таких переходов. Переходы между областями полупроводника с различными видами проводимости называются **электронно-дырочными** или **$p-n$ -переходами**. Характеристики этих переходов определяются распределением концентрации примесей, шириной запрещенной зоны, диэлектрической проницаемостью полупроводника и геометрией контактирующих областей.

Если концентрации примесей в контактирующих областях одинаковы, то $p-n$ -переход называют **симметричным**. В противном случае, т. е. при разных концентрациях легирующих примесей, он называется **несимметричным**. При этом если уровень легирования примесью одной области примерно на порядок или

более превышает уровень другой, то область с большей концентрацией примесных атомов называют **эмиттером**, а с меньшей — **базой**. Область с повышенной концентрацией примесей обозначается как n^+ или p^+ . Электрический переход может образоваться и при контакте областей полупроводника с одним типом проводимости, но с существенно разной концентрацией легирующей примеси, такие переходы называются электронно-электронными ($n^+ - n$) или дырочно-дырочными ($p^+ - p$).

При контакте полупроводников с различной шириной запрещенной зоны образуются **гетеропереходы**, а при контакте областей с одной шириной запрещенной зоны формируются **гомопереходы**. Поверхность, где концентрации примесей равны по величине и которая разделяет области с дырочной и электронной проводимостью, называется **металлургической границей** (МГ).

В электронных полупроводниковых приборах и интегральных схемах широко используются электрические переходы при контакте полупроводник — металл, а также переходы между металлом и диэлектриком, диэлектриком и полупроводником.

Из всего многообразия электрических переходов в электронных приборах наибольшее применение нашли $p - n$ -переходы и контакт металла с полупроводником. Далее наиболее подробно будут рассматриваться $p - n$ -переходы, являющиеся основой микроэлектроники и дискретных полупроводниковых приборов, а также контакты металла с полупроводником, широко применяемые как для изготовления полупроводниковых приборов, так и для омических контактов, особенно при изготовлении выводов приборов и различных соединений в интегральных схемах. Другие виды переходов будут описаны непосредственно при изучении конкретных типов приборов, где такие электрические переходы используются.

Различные виды $p - n$ -переходов играют важную роль в современной электронике как самостоятельные приборы (диоды) и как составляющие элементы, определяющие работу других полупроводниковых приборов.

2.2. Физические процессы в электронно-дырочных переходах

Физические процессы в равновесном переходе. В полупроводниковых приборах большое распространение получили несимметричные переходы, которые могут быть ступенчатыми

и плавными. **Ступенчатые (резкие) переходы** образуются в том случае, если на металлургической границе или в непосредственной близости от нее происходит резкое скачкообразное изменение концентрации легирующих примесей.

Плавными переходами называют такие, у которых в районе металлургической границы концентрация одного типа примесей постепенно уменьшается, а другого типа — растет. На металлургической границе в этом случае будет выполняться равенство примесных концентраций, т. е. $N_d = N_a$.

Рассмотрим сначала физические процессы в ступенчатом переходе в состоянии теплового равновесия, когда отсутствует приложенное напряжение и нет тока. На рис. 2.1 в полулогарифмическом (а) и линейном (б) масштабах (концентрация указана в см^{-3}) показано распределение концентрации атомов примеси (N_d , N_a) и свободных носителей (n , p) в кремнии. Помимо этого, на рис. 2.1, а указана концентрация собственных носителей n_i при комнатной температуре, когда все атомы примеси ионизованы. Различные масштабы на рис. 2.1 позволяют наглядно продемонстрировать сильное различие в концентрации как примесей, так и основных и неосновных носителей заряда. Поскольку

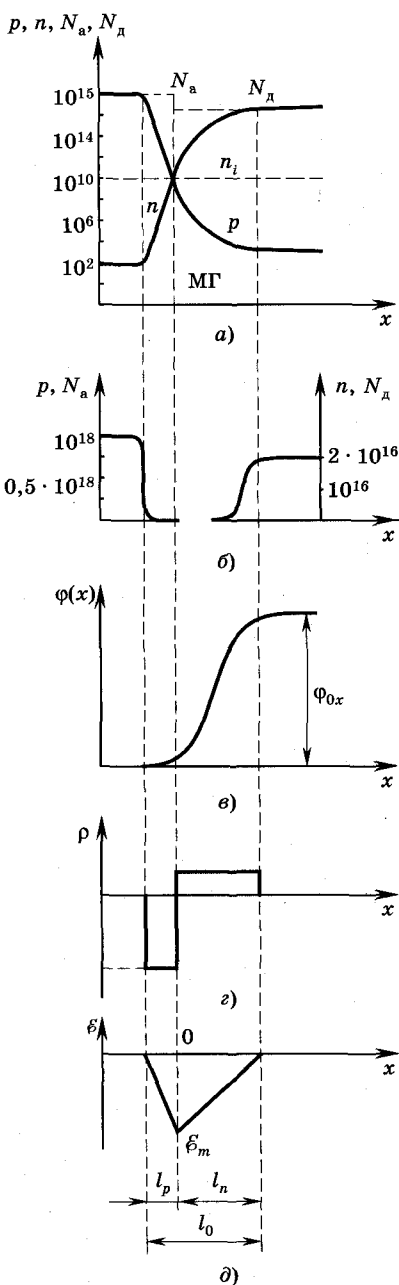


Рис. 2.1

распределение носителей заряда является неравномерным, то возникает диффузия электронов из области n в область p , а дырок, наоборот, — из области p в область n . При этом в слое p вблизи металлургической границы (МГ на рис. 2.1) окажутся избыточные электроны, которые будут рекомбинировать с дырками до тех пор, пока не наступит равновесие. В результате рекомбинации концентрация дырок уменьшится и обнажатся нескомпенсированные отрицательные заряды (ионы) акцепторов. Справа от металлургической границы обнажатся нескомпенсированные положительные ионы доноров, от которых ушли электроны. Аналогичные процессы будут и для дырок, диффундирующих из области p (слева от металлургической границы) в область n .

В рассматриваемом случае, когда $N_a \gg N_d$ и, соответственно, $p_{p0} \gg n_{n0}$, перемещение электронов играет значительно меньшую роль, чем перемещение дырок, из-за сильного различия градиентов в области контакта.

Вблизи металлургической границы образуется слой с пониженной концентрацией свободных носителей, которые образуют обедненную область. Из рис. 2.1, *a* видно, что на металлургической границе концентрации свободных носителей равны n_i .

Возникающие в окрестности металлургической границы объемные заряды ионов доноров и акцепторов создают электрическое поле, препятствующее диффузии основных носителей.

Это же электрическое поле вызывает дрейфовое движение неосновных носителей через p — n -переход, т. е. электронов (n_{p0}) из p -области в n -область и дырок (p_{n0}) в обратном направлении.

Напряженность внутреннего электрического поля нарастает до тех пор, пока оно не скомпенсирует диффузионное движение зарядов через p — n -переход. В результате как электронный, так и дырочный ток, протекающий через p — n -переход в равновесном состоянии, будет равен нулю, а уровень Ферми установится одинаковым для областей n - и p -типа.

Высота возникающего равновесного потенциального барьера определяется разностью электростатических потенциалов в p - и n -областях. Поскольку уровень Ферми устанавливается одинаковым для всей полупроводниковой структуры, то энергия равновесного энергетического барьера ($q\Phi_0$) равна разности уровней Ферми в n - и p -областях:

$$q\Phi_0 = E_{\Phi_n} - E_{\Phi_p}.$$

Из (1.7, а) и (1.7, б) с учетом (1.2), (1.3) и (1.5) следует, что

$$q\Phi_0 = \Delta E_s - kT \ln \left(\frac{N_n N_p}{n_{n0} p_{p0}} \right) = kT \ln \left(\frac{N_n N_p}{n_i^2} \right) - kT \ln \left(\frac{N_n N_p}{n_{n0} p_{p0}} \right) = kT \ln \left(\frac{p_{p0} n_{n0}}{n_i^2} \right).$$

Поскольку согласно (1.3) и (1.5) $p_{p0} n_{n0} = n_{n0} p_{p0} = n_i^2$, то

$$q\Phi_0 = kT \ln \frac{n_{n0}}{n_{p0}} = kT \ln \frac{p_{p0}}{p_{n0}}. \quad (2.1)$$

Если все примеси ионизованы, т. е. $n_{n0} = N_d$ и $p_{p0} = N_a$, то выражение (2.1) для невырожденных полупроводников можно записать в следующем виде:

$$q\Phi_0 = kT \ln \frac{N_a N_d}{n_i^2} = \Delta E_s - kT \ln \left(\frac{N_n N_p}{N_a N_d} \right). \quad (2.2)$$

Из (2.2) следует, что энергия потенциального барьера определяется, прежде всего, шириной запрещенной зоны полупроводника. С увеличением температуры потенциальный барьер уменьшается, а увеличение концентрации примесей приводит к слабому росту этого барьера. Величины ΔE_s , T , N_a , N_d влияют на значения градиента в концентрации носителей в области перехода, что и приводит к изменению потенциального барьера.

При тепловом равновесии электрическое поле в нейтральных областях полупроводника равно нулю, поэтому общий отрицательный заряд ионов акцепторов на единицу площади в p -области перехода равен $\rho'_p = qN_a^- l_p$, а положительный заряд ионов доноров в n -области равен соответственно $\rho'_n = qN_d^+ l_n$. Исходя из условия сохранения заряда

$$qN_d^+ l_n = qN_a^- l_p, \text{ или } N_d^+ l_n = N_a^- l_p, \quad (2.3)$$

получаем соотношение для длин l_n и l_p , на которых сосредоточен нескомпенсированный заряд положительных ионов доноров и отрицательных ионов акцепторов соответственно (рис. 2.1, з). Для ступенчатого перехода из уравнения Пуассона следует уравнение

$$-\frac{\partial^2 \Phi}{\partial x^2} \equiv \frac{\partial \mathcal{E}}{\partial x} = \frac{\rho(x)}{\epsilon \epsilon_0} = \frac{q}{\epsilon \epsilon_0} [p(x) - n(x) + N_d^+(x) - N_a^-(x)], \quad (2.4)$$

где $\rho(x)$ — объемный заряд, т. е. заряд, приходящийся на единицу объема полупроводника, например на 1 см^3 ; ϵ , ϵ_0 — относительная диэлектрическая проницаемость полупроводника и электрическая постоянная соответственно; $p(x)$, $n(x)$ — объемные концентрации дырок и электронов в обедненной области. Учитывая почти прямоугольное распределение зарядов в областях перехода l_n и l_p (см. рис. 2.1, *з*) и пренебрегая вследствие их малости величинами $n(x)$ и $p(x)$ по сравнению с N_d^+ и N_a^- , уравнение (2.4) можно представить в следующей форме

$$-\frac{\partial^2 \phi}{\partial x^2} \approx -\frac{q}{\epsilon \epsilon_0} N_a \text{ при } -l_p \leq x < 0,$$

$$-\frac{\partial^2 \phi}{\partial x^2} \approx \frac{q}{\epsilon \epsilon_0} N_d \text{ при } 0 < x \leq l_n. \quad (2.5)$$

В уравнениях (2.5), как и в последующих выражениях, опущены знаки «плюс» и «минус» в обозначениях ионов примеси, поскольку при комнатной температуре все атомы примеси ионизованы и, следовательно, $N_d^+ = N_d$ и $N_a^- = N_a$.

Интегрируя эти уравнения, получим напряженность электрического поля

$$\mathcal{E}(x) = \frac{-qN_a(x+l_p)}{\epsilon \epsilon_0} \text{ при } -l_p \leq x < 0,$$

$$\mathcal{E}(x) = \mathcal{E}_m + \frac{qN_dx}{\epsilon \epsilon_0} = \frac{qN_d}{\epsilon \epsilon_0}(x-l_n) \text{ при } 0 < x \leq l_p. \quad (2.6)$$

Здесь обозначено \mathcal{E}_m — максимальное (по абсолютной величине) значение напряженности электрического поля, которое достигается на металлургической границе (при $x = 0$ на рис. 2.1, *д*):

$$|\mathcal{E}_m| = \frac{qN_al_p}{\epsilon \epsilon_0} = \frac{qN_dl_n}{\epsilon \epsilon_0}. \quad (2.7)$$

Интегрируя уравнения (2.5) еще раз с граничными условиями $\phi(-l_p) = 0$ и $\phi(l_n) = \phi_0$, получим распределение потенциала ϕ_p ($-l_p < x < 0$) и ϕ_n ($0 < x < l_n$) в различных областях p — n -перехода и значение $\phi_0 = \phi_n(x=l_n) - \phi_p(x=-l_p)$, равное разности потен-

циалов, возникающей в результате диффузии подвижных носителей заряда:

$$\varphi_p(x) = 0,5|\varepsilon_m|(l_p + x)^2/l_p,$$

$$\varphi_n(x) = \varphi_0 - 0,5|\varepsilon_m|(l_n - x)^2/l_n, \quad (2.8)$$

$$\varphi_0 = \frac{1}{2}|\varepsilon_m|(l_n + l_p) = \frac{1}{2}|\varepsilon_m|l_0, \quad (2.9)$$

где $l_0 = l_n + l_p$ — полная ширина обедненной свободными носителями области. Из формулы (2.9) следует, что $\varphi_p(0) = \varphi_n(0)$ при $x = 0$, т. е. условие непрерывности потенциала. Используя выражения (2.7) и (2.9), получим для полной ширины l_0 ступенчатого перехода следующее выражение

$$l_0 = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \left(\frac{N_a + N_d}{N_a N_d} \right)}. \quad (2.10)$$

В ступенчатом несимметричном переходе, когда, например, $N_a \gg N_d$, выражение (2.10) без заметных погрешностей можно упростить:

$$l_0 = \left(\frac{2\varepsilon_0\varepsilon\varphi_0}{qN_d} \right)^{1/2}. \quad (2.11)$$

Из формулы (2.11) видно, что в существенно несимметричном случае обедненная область (p — n -переход) сосредоточена в основном в той части полупроводника, где концентрация примеси меньше, что, естественно, связано с условием электронейтральности, согласно которому полный заряд положительных ионов доноров в области перехода равняется полному заряду отрицательных ионов акцепторов.

Рассмотренная простая модель ступенчатого перехода дает достаточно точные оценки для большинства резких p — n -переходов, однако для существенно несимметричных переходов, переходов со сверхмелким залеганием, плавных переходов и т. д. в целях получения большей точности необходимо выполнять численные расчеты или проводить более строгое аналитическое рассмотрение.

Неравновесный p — n -переход. *Неравновесным переходом* называется такой p — n -переход, к которому приложено внешнее напряжение. Поскольку сопротивление обедненного слоя значительно

больше сопротивления нейтральных областей, расположенных вне перехода, то при малых токах внешнее напряжение U практически приложено к обедненному слою (p — n -переходу). Изменение высоты потенциального барьера внутри p — n -перехода будет равно приложенному напряжению.

В зависимости от полярности этого напряжения различают *прямое и обратное включение перехода*. Если напряжение U приложено плюсом к p -слою, а минусом к n -слою, то высота потенциального барьера внутри перехода уменьшится. Переход в этом случае смещен (включен) в прямом направлении. В противоположном случае, когда плюс источника подключен к n -области, а минус — к p -области, высота барьера вырастет на величину приложенного напряжения. Напряжение такой полярности является обратным (переход смещен в обратном направлении).

Следовательно, при включении внешнего напряжения высота потенциального барьера ϕ будет равна $\phi = \phi_0 - U$. Тогда для существенно несимметричного перехода, в соответствии с соотношением (2.11), ширина обедненной области может быть вычислена по формуле

$$l_0 = \sqrt{\frac{2\epsilon\epsilon_0(\phi_0 - U)}{qN_d}}. \quad (2.12)$$

Как видно из этого выражения, переход сужается при прямом включении ($U > 0$) и расширяется при обратном ($U < 0$). При обратном напряжении увеличение толщины (ширины) обедненного слоя происходит за счет смещения основных носителей под действием электрического поля. Электроны в n -области смещаются к выводу, присоединенному к положительному полюсу источника, а дырки в p -области — соответственно к отрицательному. При таком смещении происходит компенсация заряда, привносимого источником. В результате обнажается дополнительный слой ионов примесей у границ перехода, заряд которых равен заряду, привносимому источником. При прямом включении происходит обратный процесс и ширина перехода сужается из-за смещения основных носителей в сторону обедненного слоя. Проникая в обедненный слой, основные носители компенсируют часть его объемного заряда. Описанные процессы происходят за время диэлектрической релаксации ($\sim 10^{-12}$ с).

На энергетических диаграммах перехода при прямом и обратном смещении уровни Ферми E_{Φ_p} и E_{Φ_n} в областях p - и n -типа, в

отличие от равновесной диаграммы, располагаются на разной высоте (см. рис. 2.6), а разность энергий между ними равна $q|U|$.

При прямом напряжении потенциальный барьер уменьшается, равновесие нарушается и происходит диффузия электронов из n -области и встречная диффузия дырок из p -области. Из-за того, что градиенты концентрации носителей заряда в резком несимметричном переходе вследствие разной концентрации примесей в n - и p -областях иногда различаются на несколько порядков, диффузия дырок, например при $N_a \gg N_d$, будет преобладать над диффузией электронов. Вследствие диффузии увеличивается концентрация неосновных носителей в нейтральных областях, граничащих с переходом. Этот процесс называется *инжекцией* неосновных носителей. Изменение высоты потенциального барьера при инжекции приводит к изменению концентрации как основных, так и неосновных носителей. Поскольку концентрация основных носителей значительно больше, чем неосновных, то можно считать, что относительное изменение неосновных носителей (Δp_n и Δn_p) в областях инжекции существенно больше изменения основных. Концентрации избыточных инжектированных носителей Δn_p в p -области и Δp_n в n -области у границ перехода можно вычислить на основе соотношений (2.1), предполагая, что $\Delta p_n \ll n_{n0}$ и $\Delta n_p \ll p_{p0}$ и заменяя φ_0 на $(\varphi_0 - U)$, а n_p на $(n_{p0} + \Delta n_p)$. В результате этих операций получим

$$q(\varphi_0 - U) = kT \ln \frac{n_{n0}}{n_{p0} + \Delta n_p} = kT \ln \frac{n_{n0}}{n_{p0}} - kT \ln \left(1 + \frac{\Delta n_p}{n_{p0}} \right). \quad (2.13)$$

Из этого выражения с учетом формулы (2.1) получаем, что

$$\Delta n_p = n_{p0} \left(\exp \frac{U}{\varphi_T} - 1 \right). \quad (2.14)$$

Аналогичное соотношение вытекает из (2.1) и для Δp_n

$$\Delta p_n = p_{n0} \left(\exp \frac{U}{\varphi_T} - 1 \right). \quad (2.15)$$

Поделив (2.14) на (2.15) и выразив n_{p0} и p_{n0} через N_d и N_a из формул (1.3) и (1.5), получим

$$\frac{\Delta n_p}{\Delta p_n} = \frac{N_d}{N_a}. \quad (2.16)$$

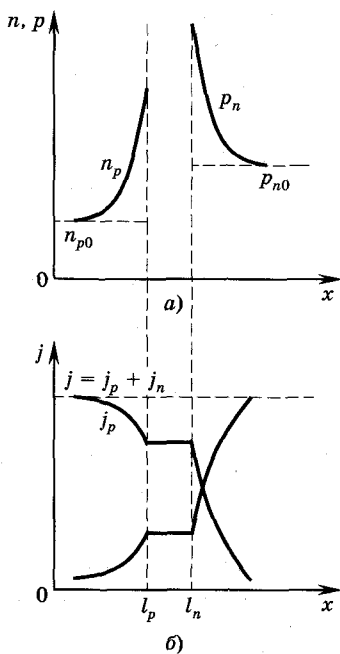


Рис. 2.2

Из (2.14) и (2.15) видно, что при инжекции наблюдается сильная зависимость концентрации неосновных носителей от приложенного напряжения. Так, увеличение напряжения U на $3\phi_T$ (на 78 мВ при $T = 300$ К) повышает концентрацию неосновных носителей (например, дырок в n -области) больше чем на порядок. На рис. 2.2, а показано изменение концентрации инжектированных носителей в областях, примыкающих к p - n -переходу, при прямом смещении.

Для несимметричного перехода при $N_a \gg N_d$ концентрация дырок, инжектированных из сильнолегированной области p (эмиттера) в слаболегированную область n (базу), намного больше концентрации электронов, инжектированных в противоположном направлении, т. е. $\Delta p_n \gg \Delta n_p$, и, следовательно,

для таких переходов можно считать, что характерна **односторонняя инжекция**. Изменение градиентов концентраций носителей вызывает соответствующие изменения электронной и дырочной составляющих плотностей полных токов j_n и j_p , распределения которых приведены на рис. 2.2, б. Отношение плотностей тока инжектированных в базу носителей и полного тока j определяет **коэффициент инжекции** γ . Так, для рассматриваемого несимметричного перехода

$$\gamma = j_p / j,$$

где $j = j_p + j_n$; j_p, j_n — плотности токов инжекции соответственно дырок и электронов. При $N_a \gg N_d$ коэффициент инжекции γ стремится к единице.

Отношение концентрации инжектированных в базу неосновных носителей к равновесной концентрации основных носителей в ней называется **уровнем инжекции** δ :

$$\delta = \frac{\Delta p_n}{n_{n0}} = \frac{\Delta p_n}{N_d},$$

причем $\delta \ll 1$ соответствует низкому уровню инжекции, а $\delta \gg 1$ — высокому.

При обратном напряжении величина потенциального барьера в области перехода увеличивается, ширина (толщина) перехода возрастает. Неосновные носители (дырки p_{n0} в n -области и электроны n_{p0} в p -области) при своем тепловом движении попадают в область перехода, где велика напряженность ускоряющего их электрического поля. Под действием этого поля неосновные носители дрейфуют в область, где становятся основными носителями, т. е. дырки из n -области дрейфуют в p -область, а электроны — из p -области в n -область. Если при инжекции (прямом смещении) происходит диффузия носителей через переход, то при обратном — дрейф носителей. В результате описанных процессов концентрация неосновных носителей у границ перехода уменьшается (рис. 2.3, а), это явление называют *экстракцией неосновных носителей*. Изменение концентрации неосновных носителей у границ перехода под действием обратного напряжения ($U < 0$) можно вычислить по формулам (2.14) и (2.15), т. е. эти формулы справедливы как при инжекции, так и при экстракции носителей. Распределение возникающих токов при обратном напряжении приведено на рис. 2.3, б.

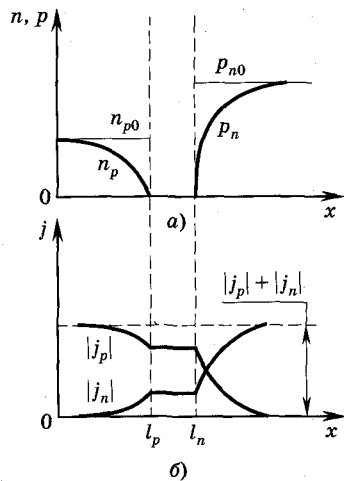


Рис. 2.3

2.3. Вольт-амперная характеристика p — n -перехода

2.3.1. Вольт-амперная характеристика идеализированного p — n -перехода.

Выражения (2.14) и (2.15) являются основными граничными условиями при вычислении идеализированной *вольт-амперной характеристики (ВАХ)*.

ВАХ идеализированного p — n -перехода может быть вычислена на основе решения уравнений (1.25) и (1.26). Идеализированный p — n -переход является упрощенной моделью реального перехода. При получении ВАХ идеализированного p — n -перехода принимаются следующие основные допущения:

— внутри p — n -перехода отсутствуют генерация, рекомбинация и рассеяние носителей;

— носители преодолевают переход мгновенно, т. е. не учитывается время их перемещения через переход, в результате токи носителей одного знака одинаковы на обеих границах перехода;

— электрическое поле существует только внутри перехода, т. е. считается, что все напряжение источника приложено к переходу, сопротивление которого много больше сопротивления прилегающих к нему областей;

— границы p — n -перехода являются безграничными плоскостями, и краевые эффекты не учитываются;

— предполагается, что изменение концентрации неосновных носителей в областях за границами перехода при небольших прямых напряжениях (малый уровень инжекции) не приводит к нарушению электрической нейтральности в этих областях;

— размеры нейтральных областей много больше диффузионной длины неосновных носителей в этих областях.

При сделанных допущениях в нейтральной области, где отсутствует электрическое поле, для вычисления параметров неравновесных носителей можно воспользоваться уравнением диффузии в форме (1.26), записанным для электронов, а также и для дырок, заменив в (1.26) соответствующим образом обозначения с n на p .

Решение уравнения диффузии для дырок с граничными условиями (2.15) и с учетом того, что при $x \rightarrow \infty p_n = p_{n0}$, имеет вид

$$\Delta p_n = p_{n0} (e^{U/\Phi_T} - 1) e^{-(x - l_n)/L_p}, \quad (2.17)$$

где $L_p = \sqrt{D_p \tau_p}$.

В результате при $x = l_n$ плотность дырочного тока, определяемого формулой (1.20), равна (с учетом того, что $\frac{d(\Delta p_n)}{dx} = \frac{dp_n}{dx}$)

$$j_p = -qD_p \left. \frac{dp_n}{dx} \right|_{l_n} = \frac{qD_p p_{n0}}{L_p} (e^{U/\Phi_T} - 1). \quad (2.18)$$

Рассматривая p -область, аналогично можно получить плотность электронного тока при $x = -l_p$:

$$j_n = qD_n \left. \frac{dn_p}{dx} \right|_{-l_p} = \frac{qD_n n_{p0}}{L_n} (e^{U/\Phi_T} - 1). \quad (2.19)$$

Распределения концентрации избыточных неравновесных носителей и токов, приведенные на рис. 2.2 для прямого смещения и на рис. 2.3 — для обратного, соответствуют общей плотности тока j через переход, постоянной во всех сечениях и равной сумме j_n и j_p , т. е.

$$j = j_n + j_p = j_0 \left(e^{U/\varphi_T} - 1 \right), \quad (2.20)$$

где

$$j_0 = \frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n} \quad (2.21)$$

есть плотность теплового тока, т. е. тока, обусловленного тепловой генерацией носителей в областях за пределами p — n -перехода.

Умножив обе части выражений (2.20) и (2.21) на площадь перехода S , получим вместо плотности тока значение тока, протекающего через p — n -переход, т. е. $I = jS$. Выражение (2.20) с учетом (2.21) описывает ВАХ идеализированного диода (рис. 2.4) и представляет собой известную формулу Шокли.

При подаче прямого смещения при $T = 300$ К для увеличения тока через переход на порядок требуется изменить напряжение всего лишь на 59,5 мВ ($\sim 2,3\varphi_T$). Из рис. 2.4, б видно, что при $|U| > 3\varphi_T$ наклон характеристики при прямом смещении в полулогарифмическом масштабе постояен, а при обратном — ток стремится к насыщению (к току $I_0 = j_0 S$).

Токи при прямом смещении, обычно соответствующие номинальным режимам работы p — n -переходов в полупроводниковых приборах, на много порядков превышают обратный ток, который для идеализированного перехода при $U > 3\varphi_T$ прак-

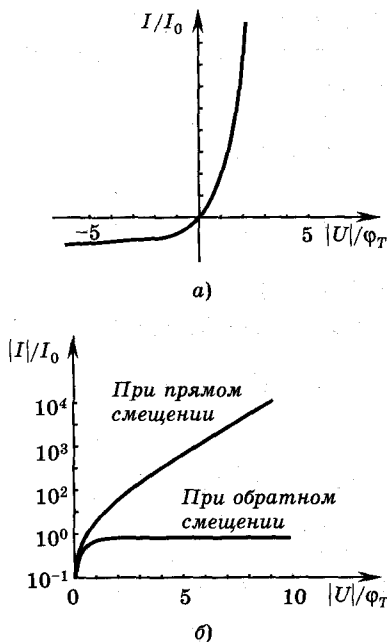


Рис. 2.4

тически равен тепловому току I_0 . Поэтому часто на графиках прямые и обратные токи изображают в разных масштабах.

На ток I_0 достаточно сильно влияет изменение температуры p — n -перехода. В случае резкого несимметричного перехода, когда $N_a \gg N_d$ и $p_{n0} \gg n_{p0}$, вторым слагаемым в выражении (2.21) можно пренебречь. Температурная зависимость теплового тока I_0 вызвана, главным образом, изменением концентрации неосновных носителей (в данном случае p_{n0} , см. выражение (1.4)), поэтому $I_0 \sim n_i^2 \sim \exp(-\Delta E_g/kT)$. Для иллюстрации сильной зависимости тока I_0 от температуры можно привести следующий пример. Если в кремнии температуру изменять от -15° до 65°C , то тепловой ток увеличится в 2^{16} раз, в германии при тех же условиях увеличение составит 2^{10} раз. В кремнии при изменении температуры от 25° до 30°C (т. е. лишь на $\Delta T = 5^\circ$), тепловой ток удваивается. В германии, чтобы получить удвоение тока, необходимо увеличить температуру от 25°C на $\Delta T = 8^\circ$. Тепловой ток I_0 уменьшается при увеличении концентрации примесей из-за снижения концентрации неосновных носителей согласно выражению (1.5). Поэтому при резко несимметричном p — n -переходе, когда $N_a \gg N_d$, основной вклад в I_0 вносят дырки, генерируемые в слабелегированной базе (n -область).

Для описания работы p — n -перехода на малом переменном сигнале используется *дифференциальное сопротивление*

$$r_{\text{диф}} = \frac{dU}{dI} = \frac{\Phi_T}{I + I_0}.$$

Для малого низкочастотного сигнала p — n -переход представляет собой линейный резистор, так как в пределах малого приращения напряжения $dU \approx \Delta U$ можно считать, что ток изменяется линейно. Величина $r_{\text{диф}}$ зависит от режима работы, т. е. от положения рабочей точки. На практике часто используют величину, обратную $r_{\text{диф}}$, которая называется крутизной ВАХ p — n -перехода.

2.3.2. ВАХ реального p — n -перехода.

Идеализованная ВАХ (формула Шокли (2.20)) учитывает инжекцию и экстракцию неосновных носителей и их диффузию в областях, прилежащих к p — n -переходу.

Формула (2.20) удовлетворительно описывает ВАХ германиевых $p-n$ -переходов при низких плотностях тока. Для $p-n$ -переходов в кремнии и арсениде галлия эта формула дает лишь качественное описание реальных характеристик.

В реальных $p-n$ -переходах существенную роль могут играть такие процессы, как рекомбинация и генерация носителей в области $p-n$ -перехода, сопротивление базы, ток утечки по поверхности, туннелирование носителей между энергетическими состояниями в запрещенной зоне, высокий уровень инжекции и пробой перехода. Оценим влияние каждого из перечисленных факторов на ВАХ.

Генерационный ток при обратном смещении. Плотность тока, обусловленного термогенерацией носителей в обедненной области (**генерационный ток**), может быть оценена следующей формулой

$$j_{\text{ген}} = \frac{qn_i l_0}{\tau_{\text{эф}}}, \quad (2.22)$$

где l_0 — ширина перехода, $\tau_{\text{эф}}$ — эффективное время жизни носителей в обедненном слое, которое определяется временем жизни электронов (τ_n) и дырок (τ_p).

Если $\tau_{\text{эф}}$ мало изменяется с температурой, то генерационный ток будет иметь примерно ту же зависимость, что и n_i . При фиксированной температуре плотность тока $j_{\text{ген}}$ пропорциональна величине l_0 , которая зависит от приложенного обратного напряжения.

В общем случае полный обратный ток при $p_{n0} \gg n_{p0}$ и $|U| > \phi_T$ можно приближенно представить суммой диффузионного тока в нейтральной области и генерационного тока в обедненной области

$$j_{\text{обр}} = q \sqrt{\frac{D_p}{\tau_p} \frac{n_i^2}{N_d}} + \frac{qn_i l_0}{\tau_{\text{эф}}}. \quad (2.23)$$

В таких полупроводниках, как германий, у которого ширина запрещенной зоны $\Delta E_g \approx 0,7$ эВ и концентрация собственных носителей при комнатной температуре равна $n_i \approx 10^{13}$ см⁻³, будет преобладать диффузионный ток, и обратный ток может быть вычислен по формуле Шокли. Если же величина n_i мала

(кремний: $n_i \approx 10^{10} \text{ см}^{-3}$, арсенид галлия: $n_i \approx 10^6 \text{ см}^{-3}$ при комнатной температуре), то генерационный ток $j_{\text{ген}}$ будет преобладать над диффузионным.

Ток рекомбинации при прямом смещении. *Рекомбинационный ток* добавляется к диффузионному инжекционному току и в этом случае, как и при обратном смещении, рекомбинационный ток пропорционален n_i . Экспериментальные исследования показывают, что в общем случае плотность тока при прямом смещении можно описать следующим выражением:

$$j \sim \exp\left(\frac{U}{a\phi_T}\right), \quad (2.24)$$

где $a = 2$, если преобладает рекомбинационный ток, $a = 1$, если преобладает ток инжекции. Если оба тока сравнимы, то $1 < a < 2$.

Влияние сопротивления базы. В реальных p — n -переходах сопротивление базы r_b составляет значения от единиц до сотен Ом. Наличие сопротивления базы приводит к тому, что внешнее напряжение от источника питания распределяется между p — n -переходом и базовой областью. Поскольку в формуле Шокли (2.20) в показателе экспоненты стоит напряжение на переходе, то при наличии сопротивления r_b вместо U надо подставить в показатель величину $U - Ir_b$, и тогда формула Шокли примет следующий вид:

$$I = I_0 \left(e^{(U - Ir_b)/\phi_T} - 1 \right). \quad (2.25)$$

Из последней формулы видно, что при малых прямых токах падение напряжения на базе Ir_b можно не учитывать. При увеличении тока падение напряжения на базе может превысить напряжение на переходе и на ВАХ появляется участок, близкий к линейному, крутизна ВАХ будет меньше и прямое напряжение для данного тока заметно больше (рис. 2.5, a , кривая 2), чем для идеализированного перехода (см. рис. 2.5, a , кривая 1). (Кривая 3 на рис. 2.5 соответствует некоторому промежуточному сопротивлению r_b .)

Для кремниевых переходов при $U < 0,5 \text{ В}$ ток очень мал и в линейном масштабе на ВАХ практически равен нулю. Если $U > 0,5 \text{ В}$, ток на ВАХ идеализированного перехода изменится

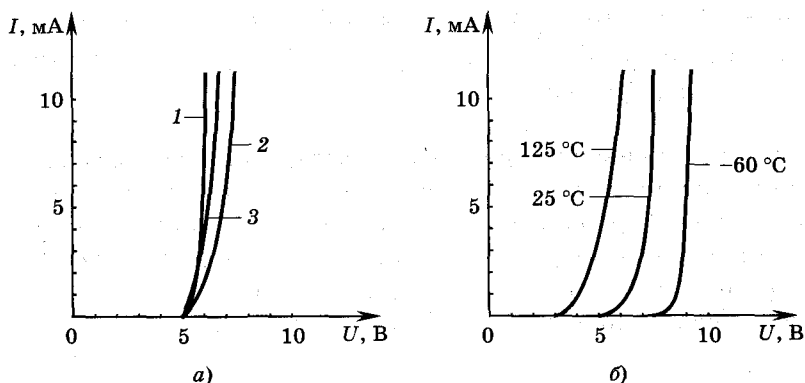


Рис. 2.5

очень резко при малом изменении напряжения (см. рис. 2.5, а). Например, при повышении напряжения от 0,51 на 0,1 В (для кривой 1) ток увеличивается в 50 раз.

При увеличении температуры подвижность свободных носителей в реальных условиях эксплуатации диодов увеличивается, и сопротивление базы заметно возрастает. На рис. 2.5, б показаны ВАХ кремниевого перехода для разных температур. При повышении температуры ВАХ смещается влево, а ее крутизна уменьшается, так как дифференциальное сопротивление растет.

При высоких уровнях инжекции наблюдается **эффект модуляции сопротивления базы**, который заключается в уменьшении r_b с ростом тока из-за увеличения концентрации носителей в базе. Особенно сильно этот эффект сказывается при малой толщине базы, когда она меньше диффузионной длины. В этом случае при увеличении тока r_b стремится к нулю.

На поверхности $p-n$ -перехода (полупроводника) из-за наличия загрязнений и влияния поверхностного заряда могут образовываться проводящие пленки и каналы, которые обуславливают появление **тока утечки**. Ток утечки возрастает при увеличении напряжения и при больших обратных напряжениях может превысить тепловой ток и ток генерации.

Пробой перехода. *Пробоем $p-n$ -перехода* называют резкое увеличение тока через переход при большом обратном напряжении, создающем в переходе большую напряженность электрического поля. Напряжение, при котором происходит лавинооб-

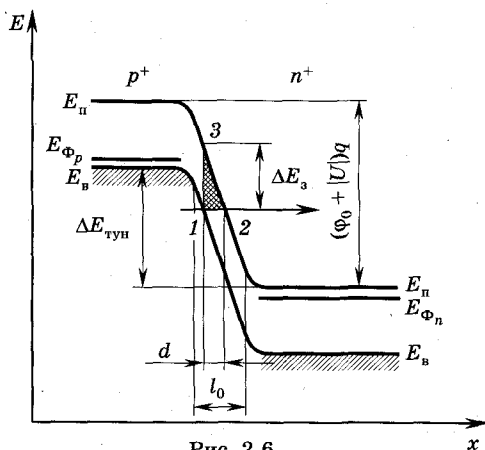


Рис. 2.6

разное нарастание тока, называется напряжением пробоя $U_{\text{проб}}$. Существует три основных механизма пробоя: **туннельный**, **тепловой** и **лавинное умножение**. Первый и третий механизмы обусловлены увеличением напряженности электрического поля в переходе, а второй — увеличением рассеиваемой мощности в переходе и, соответственно, его температуры.

В основе **туннельного пробоя** лежит туннельный эффект, связанный с переходом электронов через тонкий потенциальный барьер без изменения энергии (рис. 2.6). Толщина барьера при туннельном переходе частиц составляет величину порядка 10 нм. Такие барьеры возможны при контакте между сильнолегированными полупроводниками ($N_d, N_a > 5 \cdot 10^{18} \text{ см}^{-3}$). На энергетической диаграмме перехода при обратном смещении (см. рис. 2.6) показан туннельный переход электрона с уровня 1 в валентной зоне p -области на энергетический уровень такой же высоты в зоне проводимости n -области (положение 2). Электрон преодолевает энергетический барьер треугольной формы с максимальной высотой ΔE_s (точки 1, 2, 3). Необходимым условием туннельного перехода электронов является наличие занятых энергетических состояний в валентной зоне p -области и свободных состояний с теми же значениями энергии в n -области. На рис. 2.6 такие состояния сосредоточены в интервале энергий $\Delta E_{\text{тун}}$.

Ширина потенциального барьера d меньше толщины обедненного слоя l_0 и уменьшается при увеличении обратного напряжения. При повышении температуры высота барьера ΔE_s уменьша-

ется из-за уменьшения ширины запрещенной зоны и напряжение туннельного пробоя снижается, т. е. температурный коэффициент напряжения туннельного пробоя отрицателен.

Лавинное умножение (или **ударная ионизация**) связано с размножением носителей под действием сильного электрического поля. Носители, перемещающиеся через $p-n$ -переход при подаче большого обратного напряжения, на длине свободного пробега приобретают энергию, достаточную для образования новых электронно-дырочных пар за счет ударной ионизации атомов полупроводника. Рожденные электронно-дырочные пары, ускоряясь полем перехода, также приобретают энергию, достаточную для ионизации атомов, в результате появляются все новые электронно-дырочные пары, количество которых лавинообразно нарастает.

При большей ширине запрещенной зоны носители должны приобретать большую энергию в электрическом поле для реализации ударной ионизации, таким образом, большими энергиям ΔE_3 соответствуют и большие значения $U_{\text{проб}}$.

При возрастании температуры $U_{\text{проб}}$ увеличивается из-за уменьшения длины свободного пробега носителей, при этом увеличивается и напряженность электрического поля, необходимая для ударной ионизации. Следовательно, температурный коэффициент напряжения лавинного пробоя положителен. Таким образом, по знаку температурного коэффициента напряжения пробоя можно отличить туннельный пробой от пробоя ударной ионизации.

Тепловой пробой происходит в результате разогрева $p-n$ -перехода под действием обратного тока $I_{\text{обр}}$. Увеличение обратного напряжения вызывает повышение температуры перехода, что, в свою очередь, приводит к возрастанию тока $I_{\text{обр}}$. Количество выделенного переходом тепла, определяющего его температуру, пропорционально произведению $I_{\text{обр}} U_{\text{обр}}$. Если количество теплоты, выделенной в переходе, превышает количество отводимой от перехода теплоты, то при достаточном напряжении развивается процесс непрерывного нарастания температуры, а следовательно, и тока, т. е. происходит тепловой пробой.

Напряжение теплового пробоя тем ниже, чем больше тепловой обратный ток. Вследствие этого на характеристике возникает участок с отрицательным дифференциальным сопротивлением. В $p-n$ -переходах на основе Si и GaAs обратные токи весьма малы, и напряжение теплового пробоя в таких переходах больше

напряжения лавинного пробоя. При высоких температурах окружающей среды возможен пробой, сочетающий оба механизма.

В заключение подытожим основные причины, приводящие к различию идеализированных и реальных ВАХ $p-n$ -перехода. При обратном смещении основными физическими процессами, обуславливающими указанные различия, являются: ток термогенерации в обедненном слое, ток утечки и пробой перехода. При прямом напряжении в начальной части характеристики это ток рекомбинации-генерации, а при больших напряжениях — наличие сопротивления базы. Для иллюстрации влияния указанных причин на рис. 2.7 приведены реальные и идеализированные ВАХ кремниевого перехода, где по оси абсцисс отложено напряжение, нормированное на тепловой потенциал, а по оси ординат — относительная плотность тока (I_0 — тепловой ток, соответствующий идеализированному $p-n$ -переходу). На рис. 2.7 участок «а» соответствует преобладанию генерационно-рекомбинационного тока, «б» — преобладанию диффузионного (инжекционного) тока, участок «в» характеризуется высоким уровнем инжекции, «г» — влиянием последовательного сопротивления базы, участок «д» объясняется наличием тока термогенерации в обедненной области и обратного тока утечки, участок «е» — пробой $p-n$ -перехода.

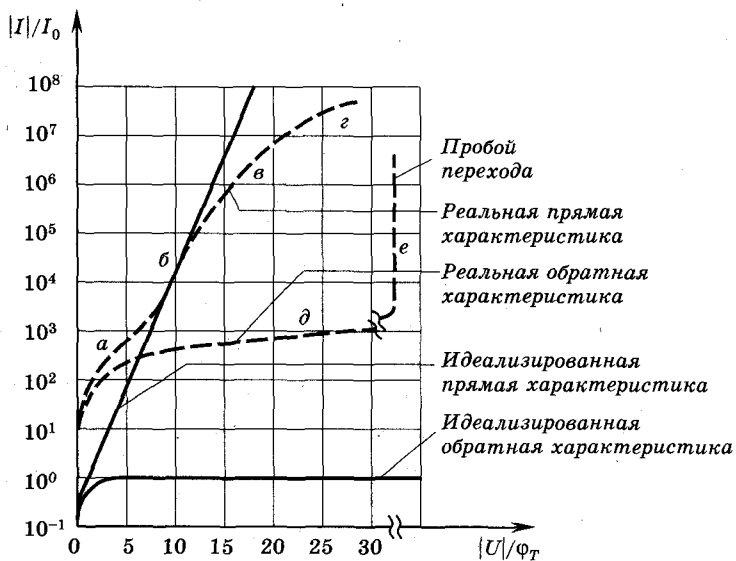


Рис. 2.7

2.4. Электрическая модель $p-n$ -перехода

Барьерная емкость. Как показано при рассмотрении физических процессов в $p-n$ -переходе (см. п. 2.1), по обе стороны от металлургической границы возникают объемные заряды ионов доноров (концентрация N_d) и акцепторов (концентрация N_a). Величины этих зарядов зависят от формы перехода, температуры и ряда других факторов, приводящих к изменению толщины обедненного слоя, в частности, от приложенного напряжения. Наличие зарядов противоположного знака в этой области приводит к появлению емкости, которая называется *барьерной*. Эта емкость оказывает влияние на работу $p-n$ -перехода при обратных напряжениях. Барьерная емкость $p-n$ -перехода определяется выражением $C_{бар} = dQ_{об}/dU$, где $dQ_{об}$ — дифференциальное приращение объемного заряда, вызванное достаточно малым изменением приложенного напряжения. ($Q_{об}$ — объемный заряд, сосредоточенный внутри $p-n$ -перехода.) Таким образом, влияние $C_{бар}$ в электрических схемах проявляется при изменении во времени напряжения на переходе. В этом случае, помимо тока, определяющего ВАХ, в $p-n$ -переходе протекает емкостный ток $I(t) = dQ_{об}/dt = (dQ_{об}/dU)(dU/dt)$. В несимметричном p^+-n -переходе со ступенчатым распределением примеси объемный заряд $Q_{об} = qSN_d l_0(U)$ определяется шириной обедненного слоя l_0 , которая зависит от приложенного напряжения смещения U .

Учитывая, что согласно (2.12) $l_0(U) = \sqrt{2\varepsilon\varepsilon_0(\varphi_0 - U)/(qN_d)}$, получаем для барьерной емкости $C_{бар}$ формулу

$$C_{бар} = \left| \frac{dQ_{об}}{dU} \right| = S \sqrt{\frac{qN_d\varepsilon_0\varepsilon}{2(\varphi_0 - U)}} = \frac{S\varepsilon_0\varepsilon}{l_0}. \quad (2.26)$$

Поскольку с ростом модуля $|U|$ при обратном включении толщина обедненного слоя l_0 возрастает, емкость $C_{бар}$ уменьшается с увеличением обратного напряжения. Зависимость емкости от приложенного напряжения называется *вольт-фарадной характеристикой (ВФХ)*. На рис. 2.8 приведены ВФХ для p^+-n -перехода со ступенча-

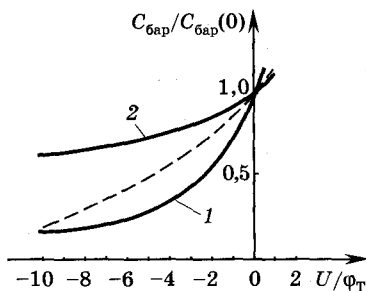


Рис. 2.8

тым (кривая 1) и линейным (кривая 2) распределениями примесей, построенные в относительных (безразмерных) координатах. Штриховая кривая соответствует некоторому другому возможному распределению примесей в переходе.

Диффузионная емкость. При подаче прямого напряжения существуют две причины, обуславливающие емкость $p-n$ -перехода: изменение зарядов в обедненном слое и изменение концентрации инжектированных носителей в нейтральных областях вблизи границы перехода в зависимости от приложенного прямого напряжения.

В результате при подаче прямого напряжения полная емкость равна $C = C_{\text{бар}} + C_{\text{диф}}$, где $C_{\text{диф}}$ — диффузионная емкость. Емкость $C_{\text{диф}}$ связана с диффузией неосновных носителей, инжектированных через переход при прямом смещении, и определяется зарядом этих носителей, накопленных за пределами области перехода.

В случае несимметричного $p^+ - n$ -перехода $C_{\text{диф}}$ определяется зарядом дырок Q_p ($Q_p \gg Q_n$), накопленным в базе. Интегрируя распределение (2.17) $\Delta p_n(x)$ (п. 2.2) для толстой базы $W_b \gg L_p$ (L_p — диффузионная длина для дырок), получим

$$Q_p(U) = I_0 \tau_{\text{эф}} \exp(U/\phi_T - 1), \quad (2.27)$$

где $\tau_{\text{эф}}$ — эффективное время жизни неосновных носителей (дырок в n -области), зависящее от скорости рекомбинации (в данном случае $\tau_{\text{эф}} \approx \tau_p$).

Для тонкой базы ($W_b \ll L_p$) и малого переменного сигнала на низких частотах диффузионная емкость определяется следующим выражением

$$C_{\text{диф}} = \frac{dQ_p}{dU} = I_0 \tau_p \frac{e^{U/\phi_T}}{\phi_T}. \quad (2.28)$$

При $U \gg \phi_T$ диффузионная емкость равна

$$C_{\text{диф}} \approx \frac{I \tau_p}{\phi_T}, \quad (2.29)$$

где ток I приблизительно равен току инжекции, что для кремниевого перехода справедливо при $U > 0,4 \dots 0,5$ В. Диффузион-

ная емкость при $U > 0,4 \dots 0,5$ В значительно превышает барьерную, для $U < 0,4 \dots 0,5$ В, наоборот, барьерная емкость больше диффузионной (рис. 2.9).

Для диффузии неосновных носителей через базу и установления заряда неосновных носителей в общем случае необходимо время порядка $\tau_{эф}$. На высоких частотах, когда $\omega \approx 4/\tau_{эф}$, диффузионная емкость уменьшается с ростом частоты, и при $\omega \gg 4/\tau_{эф}$ она стремится к нулю, поскольку в течение малого периода $T = \frac{2\pi}{\omega} \ll \tau_{эф}$ заряд не успевает изменяться синхронно с напряжением и $dQ_p/dU = 0$.

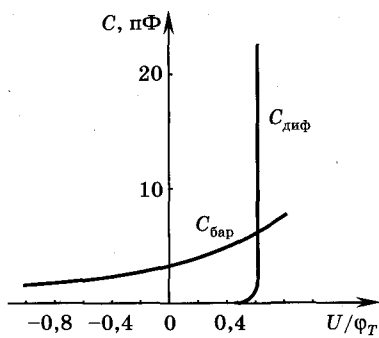


Рис. 2.9

Эквивалентная схема $p-n$ -перехода. Для многих практических применений при разработке электронных схем необходимо аналитическое описание полупроводниковых приборов, когда сами приборы заменяются электрическими моделями. Наиболее распространенным способом моделирования прибора является его эквивалентная электронная схема, представляющая прибор в виде соединения простейших элементов: резисторов, конденсаторов, индуктивностей, идеализированных диодов и т. п. Параметры указанных элементов и их взаимосвязь на постоянном и переменном токе определяются различными соотношениями. Токи и напряжения на внешних выводах вычисляются из эквивалентной схемы методом теории цепей.

Наиболее общей является модель для большого сигнала, которая пригодна для токов и напряжений, изменяющихся в любых пределах, т. е. когда связь между ними нелинейна. На рис. 2.10 приведена эквивалентная схема $p-n$ -перехода (диода), которая является одной из наиболее распространенных. На этом рисунке диод VD моделирует идеализированный $p-n$ -переход, $r_б$ — объемное сопротивление базы, $R_{обр}$ учитывает ток термогенерации в $p-n$ -пе-

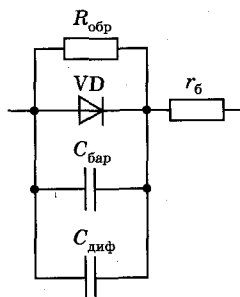


Рис. 2.10

реходе и ток утечки. Поскольку в общем случае эти токи зависят от напряжения нелинейно, то $R_{обр}$ является функцией обратного напряжения и для ее задания необходима определенная аппроксимация и дополнительные параметры. Для практических целей часто используют $R_{обр}$ в виде постоянного резистора, что соответствует линейной аппроксимации обратной ветви ВАХ в заданном диапазоне изменения $U_{обр}$. Поскольку резистор $R_{обр}$ имеет большую величину, то при прямых смещениях он практически не влияет на ток, и его можно не учитывать.

Описанная модель не учитывает эффект модуляции сопротивления базы, ток рекомбинации, пробой, частотную зависимость $C_{диф}$ и ряд других явлений. Для учета этих особенностей реального $p-n$ -перехода требуются более сложные модели.

2.5. Переходные процессы в $p-n$ -переходе

При использовании диодов ($p-n$ -переходов) в качестве переключателей и в различных импульсных схемах требуется, чтобы время перехода от прямого смещения к обратному было по возможности малым, т. е. временные интервалы переходных процессов должны быть минимальными. Процессы установления напряжения или тока в $p-n$ -переходе при воздействии импульсных сигналов называются *переходными процессами*.

Переходные процессы в $p-n$ -переходе связаны в основном с двумя явлениями:

- 1) накоплением неосновных носителей в базе $p-n$ -перехода при его прямом включении и их рассасыванием при снятии или уменьшении напряжения;
- 2) перезарядкой барьерной емкости.

Движение неосновных носителей в базе носит преимущественно диффузионный характер и в силу этого происходит достаточно медленно, поэтому накопление носителей в базе и их рассасывание влияет на инерционные свойства $p-n$ -переходов в режиме переключения.

При больших плотностях тока через переход основную роль играет накопление неосновных носителей, и перезарядкой барьерной емкости можно пренебречь.

Для малых плотностей тока импульсные процессы в переходе определяются в основном перезарядкой барьерной емкости.

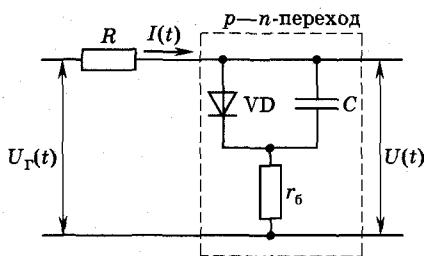


Рис. 2.11

Параметры внешней электрической цепи также могут существенно изменить временные зависимости напряжения и тока, характеризующие импульсные процессы в переходе.

Рассмотрим сначала переходной процесс включения $p-n$ -перехода в простейшей схеме, представленной на рис. 2.11. Когда на вход схемы подается импульс напряжения положительной полярности $U_{Г1}(t)$ (рис. 2.12, а) в момент времени t_1 , то ток через переход при $R \gg r_6$ изменится скачком до величины $I_{Г1} = (U_{Г1} - U_n)/R$ (рис. 2.12, б), где U_n — падение напряжения на переходе. Если $U_{Г1}(t) \gg U_n$, то ток I в течение переходного процесса практически не изменится и будет равен $I_{Г1} = U_{Г1}/R$. В этом случае внешняя цепь (генератор, сопротивление R) по отношению к переходу является источником тока. Форма прямого тока будет повторять форму положительного импульса напряжения.

При малом прямом токе можно пренебречь падением напряжения на сопротивлении базы, и напряжение на переходе U_n плавно и монотонно увеличивается по мере заряда барьерной емкости (рис. 2.12, в).

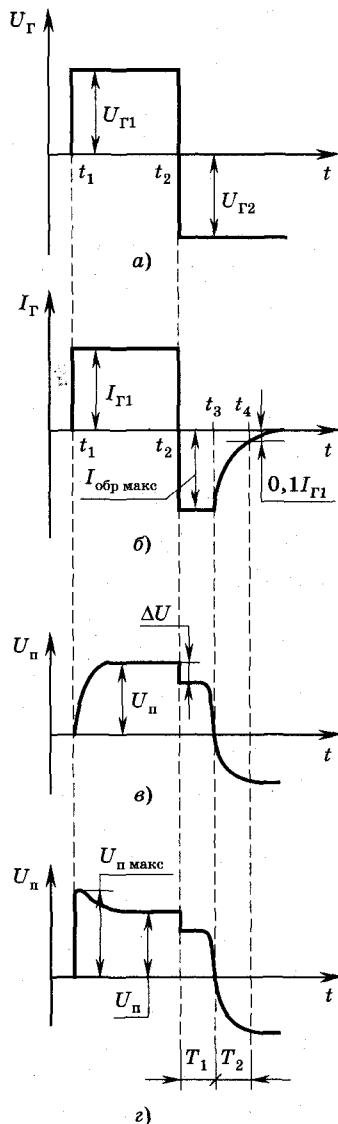


Рис. 2.12

При малых токах диффузионная емкость заметно меньше барьерной и ею можно пренебречь в течение переходного процесса.

При больших прямых токах и высоком уровне инжекции необходимо учитывать падение напряжения на сопротивлении базы и модуляцию этого сопротивления. В этом случае в момент t_1 происходит скачок напряжения $\delta U_1 = I_{Г1} r_6$. По мере заряда барьерной емкости увеличивается напряжение на $p-n$ -переходе и общее напряжение достигает значения $U_{п\text{ макс}} = U_{п} + I_{Г1} r_6$ (рис. 2.12, з). Из-за большого тока $I_{Г1}$ этот процесс происходит за очень малое время, т. е. можно считать, что напряжение изменяется от 0 до $U_{п\text{ макс}}$ скачком. Влияние диффузионной емкости при быстрых изменениях напряжений и токов пренебрежимо мало. Уменьшение напряжения относительно $U_{п\text{ макс}}$ со временем обусловлено постепенным уменьшением сопротивления базы (модуляция сопротивления базы) по мере накопления в ней инжектированных носителей. Процессы установления прямого напряжения происходят за время порядка эффективного времени жизни неосновных носителей в базе.

Рассмотрим теперь процессы, происходящие при переключении перехода, когда в момент времени t_2 напряжение генератора скачком изменяется с прямого на обратное. В результате этого происходит резкое изменение тока на величину $\Delta I = (U_{Г1} + |U_{Г2}|)/R$. При $t_3 > t \geq t_2$ через $p-n$ -переход протекает обратный ток $I_{обр.\text{ макс}} = U_{Г2}/R$ (см. рис. 2.12, б), который много больше теплового тока и тока термогенерации. Этот ток вызван движением избыточных неосновных носителей, накопленных в базе за время существования положительного прямого напряжения на переходе (за время протекания прямого тока $t = t_2 - t_1$). Скачку тока ΔI соответствует скачок напряжения $\Delta U = \Delta I r_6$ (см. рис. 2.12, в). Весь переходной процесс при переключении напряжения с прямого на обратное обычно разделяется на две стадии.

Первая стадия при $t_3 > t \geq t_2$ характеризуется **высокой обратной проводимостью**, обусловленной наличием в базе у границы перехода избыточной концентрации неосновных носителей. Обратный ток при $U_{Г2} \gg U_{п}$ постоянен и имеет большое значение. В рассматриваемом временном интервале происходит рассасывание накопленных в базе неосновных носителей за счет их перехода в эмиттер и рекомбинации. Эти процессы иллюстриру-

ются рис. 2.13, где показано распределение концентрации дырок в базе в различные моменты времени, при этом полагается, что толщина базы W_6 много больше диффузионной длины L_p . Кривая распределения дырок в базе при $t < t_2$ соответствует протеканию постоянного прямого тока $I_{\Gamma 2}$, ограниченного внешним сопротивлением R . В последующие моменты времени при $t_2 < t < t_3$ из-за

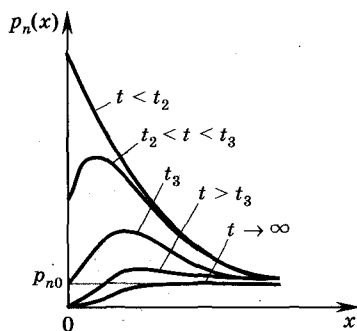


Рис. 2.13

изменения тока градиент концентрации дырок dp_n/dx изменяет знак на противоположный на границе перехода. В момент t_3 на границе базы со стороны перехода ($x = 0$) имеет $p_n(0) = p_{n0}$, где p_{n0} — исходная концентрация дырок (неосновных носителей) в n -базе. В это время напряжение на переходе становится равным нулю (см. рис. 2.12, в, з, момент $t = t_3$) и заканчивается первая стадия — стадия высокой обратной проводимости. За время $t_3 - t_2$ из базы удаляется большая часть избыточного заряда. Длительность первой стадии T_1 прямо пропорциональна времени жизни дырок в базе и зависит от соотношения прямого $I_{\Gamma 2}$ и обратного $I_{обр}$ токов через переход. Время T_1 может быть вычислено по формуле $T_1 = \tau_6 \ln(1 + I_{\Gamma 1}/I_{обр \max})$, т. е. T_1 определяется диффузионной емкостью. При $t > t_3$ происходит **восстановление обратного сопротивления** перехода, сопровождающееся уменьшением обратного тока. Ток в этой стадии процесса определяется не только переходом оставшихся избыточных неосновных носителей из базы в эмиттер, но и перезарядкой барьерной емкости. При большом обратном токе $I_{обр}$ (при малом внешнем сопротивлении R) емкость перезарядается быстро, а при малом токе $I_{обр}$ переходной процесс более длителен и $T_2 = RC_{бар}$ при $R \gg r_6$. Принято считать, что стадия восстановления обратного напряжения заканчивается при $I_{обр} = 0,1I_{\Gamma 1}$, и ее продолжительность определяется временем T_2 (см. рис. 2.12, б, в, з).

Таким образом, рассмотренные явления позволяют заключить, что длительности переходных процессов определяются значениями барьерных и диффузионных емкостей p — n -перехода.

2.6. Контакты металл—полупроводник

Электрические переходы металл—полупроводник применяются для создания диодов, транзисторов и омических контактов, при изготовлении внешних выводов полупроводниковых приборов и интегральных схем.

При непосредственном контакте металла с полупроводником высота возникающего потенциального барьера для электронов зависит от работы выхода металла и плотности поверхностных состояний.

Работой выхода называется разность энергий между уровнем Ферми и уровнем свободного пространства (вакуума) $E_{\text{вак}}$. Для металла эта величина соответствует $q\phi_m$, а в полупроводнике n -типа она равна $q(\chi + U_n)$ (рис. 2.14, а, правый), где $q\chi$ — разность энергий между энергией дна зоны проводимости $E_{\text{ц}}$ и уровнем вакуума $E_{\text{вак}}$ (электронное сродство), qU_n — разность энергий между уровнем дна зоны проводимости $E_{\text{ц}}$ и уровнем Ферми E_{ϕ_n} . Разность работ выхода металла и n -полупроводника, равная $q(\phi_m - \chi - U_n)$, определяет контактную разность потенциалов $q\phi'_n$, т. е. $q\phi'_n = q(\phi_m - \chi - U_n)$.

Оценим высоту барьера для рассматриваемого контакта. Зонные диаграммы равновесного контакта металла с p - и n -полупроводниками приведены на рис. 2.14, а, где E_{ϕ_i} — уровень Ферми для собственного полупроводника. Рассмотрим сначала соединение металла с n -полупроводником, работа выхода электронов у которого меньше, чем у металла. В этом случае часть электронов переходит из полупроводника в металл. В результате в полупроводнике появится обедненный слой, содержащий положительный заряд ионов доноров. Переход электронов будет происходить до тех пор, пока уровни Ферми в обоих материалах не сравняются ($E_{\phi_m} = E_{\phi_n}$), т. е. уровень Ферми n -полупроводника (E_{ϕ_n}) понизится относительно уровня Ферми в металле E_{ϕ_m} на величину $q\phi'_n$, равную разности соответствующих работ выхода. На такую же величину понизится уровень вакуума $E_{\text{вак}}$ (см. рис. 2.14, а, правый). В результате энергия, соответствующая высоте потенциального барьера $q\phi_n$ (рис. 2.14, а), преодолеваемого электронами при переходе из металла с уровня E_{ϕ_m} в зону проводимости полупроводника (при идеальном контакте n -полупроводника и металла, когда работа выхода из

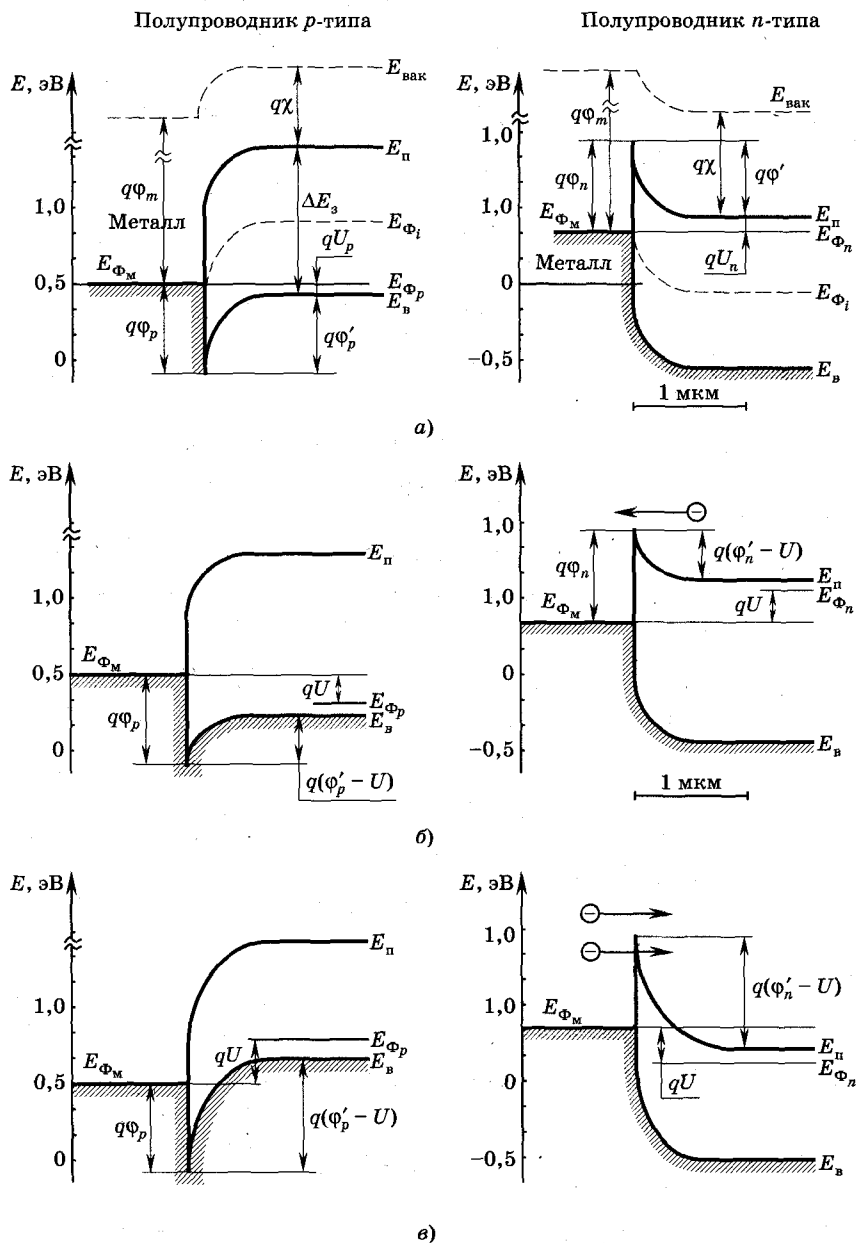


Рис. 2.14

металла больше), равна разности между работой выхода металла и электронным сродством полупроводника:

$$q\varphi_n = q(\varphi_m - \chi). \quad (2.30)$$

При идеальном контакте между металлом и полупроводником p -типа (рис. 2.14, а, левый) контактная разность потенциалов $q\varphi'_p$ и высота потенциального барьера $q\varphi_p$ определяются аналогичными выражениями (с учетом ширины запрещенной зоны ΔE_g):

$$\begin{aligned} q\varphi'_p &= \Delta E_g - q(\varphi_m - \chi - U_p), \\ q\varphi_p &= \Delta E_g - q(\varphi_m - \chi), \end{aligned} \quad (2.31)$$

где qU_p — энергия между потолком валентной зоны и уровнем Ферми E_{Φ_p} .

Таким образом, при контакте металла с полупроводником валентная зона проводимости полупроводника занимает определенное энергетическое положение по отношению к уровню Ферми металла. Если это положение известно, то оно служит граничным условием при решении уравнения Пуассона в полупроводнике, которое записывается в том же виде, что и для случая p — n -переходов (см. п. 2.1). В результате можно, как и для p — n -перехода, вычислить все параметры перехода металл — полупроводник.

Выражения (2.30) и (2.31) дают хорошее приближение при отсутствии поверхностного заряда.

Реально в n -полупроводнике часто существует достаточно большой отрицательный поверхностный заряд, удаляющий электроны из приповерхностного слоя полупроводника. Величина такого заряда определяется плотностью поверхностных состояний. В этом случае высота барьера определяется как разностью работ выхода, так и плотностью поверхностного заряда. При очень большой плотности поверхностного заряда высота барьера определяется свойствами поверхности полупроводника и не зависит от работы выхода металла.

При контакте с полупроводником p -типа отрицательный поверхностный заряд обогащает приповерхностный слой дырками, поэтому формирование обедненного слоя, необходимого для получения выпрямляющего контакта, можно получить при работе выхода электронов из металла меньшей, чем из полупроводника p -типа. В этом случае электроны из металла переходят в валентную зону полупроводника, уменьшая тем самым концентрацию дырок в приповерхностной области.

Как следует из выражений (2.30) и (2.31), высота барьера ϕ_n и ϕ_p не зависит от концентрации примесей и температуры, а определяется только типом металла и полупроводника, а в реальном контакте, как отмечалось выше, — также плотностью поверхностного заряда в полупроводнике, а следовательно, плотностью поверхностных состояний. Теоретически вычислить высоту барьера ϕ_n и ϕ_p достаточно сложно. На практике эти величины определяют экспериментальным путем. В качестве примера в таблице 2.1 приведены значения ϕ_n и ϕ_p для некоторых видов контактов.

Приведенные в таблице значения для ϕ_n и ϕ_p соответствуют концентрации доноров $N_d \approx 10^{15} \text{ см}^{-3}$, $N_a \approx 3 \cdot 10^{15} \text{ см}^{-3}$ при $T = 300 \text{ К}$. Работа выхода из алюминия (Al) и кремния (Si) n -типа при этих данных примерно одинакова и составляет $\sim 4,3 \text{ эВ}$, следовательно, образование барьера и обедненного слоя здесь обусловлено отрицательным поверхностным зарядом. Для контакта p —Si—Al отрицательный поверхностный заряд уменьшает высоту барьера.

Чем выше высота барьера, тем больше ширина обедненного слоя, которая, как и для p — n -перехода, уменьшается с ростом концентрации доноров.

В неравновесном контакте металла с n -полупроводником, когда к нему приложено внешнее напряжение U , происходит понижение потенциального барьера при подаче прямого смещения («плюс» к металлу при использовании n -полупроводников, или «минус» к металлу для p -полупроводников, см. рис. 2.14, б) и увеличение барьера при обратном напряжении («минус» к металлу при использовании n -полупроводников и, наоборот, для p -полупроводников, см. рис. 2.14, в). Прямой ток через контакт образуют электроны, движущиеся из полупроводника (стрелка

Таблица 2.1

Контакт	$\phi_n, \text{ В}$	$\phi_p, \text{ В}$
Кремний с алюминием Si—Al	0,72	0,58
Кремний с золотом Si—Au	0,8	0,34
Арсенид галлия с алюминием GaAs—Al	0,8	—
Арсенид галлия с золотом GaAs—Au	0,9	0,42

на рис. 2.14, б для n -полупроводника) с энергией большей, чем высота пониженного барьера $q(\phi'_n - U)$. Для p -области прямой ток образует электроны, переходящие из металла в p -полупроводник и преодолевающие потенциальный барьер $q(\phi'_p - U)$.

Обратный ток I_0 образуется электронами, переходящими из металла в полупроводник (стрелки на рис. 2.14, в) и преодолевающими барьер $q\phi_n$ для n -полупроводника или $q\phi_p$ для p -полупроводника. Величина тока I_0 определяется термоэмиссией электронов из металла в полупроводник и может быть вычислена по формуле

$$I_0 = SAT^2 \exp\left(-\frac{\phi_{n,p}}{\phi_T}\right), \quad (2.32)$$

где $\phi_{n,p} = \phi_n$ для n -полупроводников и $\phi_{n,p} = \phi_p$ для p -полупроводников, S — площадь контакта, A — постоянная термоэмиссии, имеющая разные значения для различных материалов; так, для кремния n -типа $A = 110 \text{ А}/(\text{см}^2 \cdot \text{К}^2)$.

С ростом ϕ_n и ϕ_p обратный ток сильно уменьшается, но при одинаковых условиях он значительно больше обратного тока p — n -перехода, где он определяется тепловым током I_0 . Отличия при комнатной температуре и $N_d = 10^{15} \text{ см}^{-3}$ составляют около 5 порядков; так, например, для контакта $\text{Al}-n\text{-Si}$ $I_0 = 2 \cdot 10^{-9} \text{ А}$, для p^+ — n -перехода $I_0 = 10^{-14} \text{ А}$ при всех прочих равных условиях. В силу этих причин различаются и ВАХ контакта металл — полупроводник и p — n -перехода. Для сравнения на рис. 2.15 приведены ВАХ для контакта $\text{Al}-(n\text{-Si})$ (кривая 1) и кремниевого p — n -перехода (кривая 2). Для контакта металл —

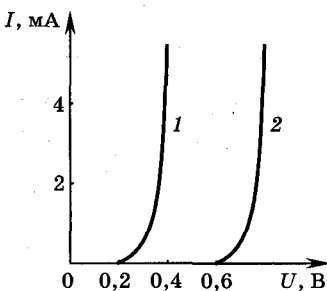


Рис. 2.15

полупроводник больший обратный ток обуславливает меньшее прямое напряжение при одинаковом прямом токе. ВАХ электрических переходов металл — полупроводник может быть получена с использованием той же формулы (2.20), что и для p — n -переходов, где ток I_0 вычисляется в соответствии с выражением (2.32). При малых прямых токах ток рекомбинации заметно меньше, чем в p — n -переходе, и ВАХ реального

перехода металл — полупроводник практически не отличается от теоретической. Однако в области больших прямых токов из-за наличия падения напряжения на нейтральной полупроводниковой области ВАХ может отличаться от теоретической кривой.

Что касается обратного тока, то в сильном электрическом поле и при концентрациях примесей $N_d > 3 \cdot 10^{15} \text{ см}^{-3}$ толщина обедненного слоя становится очень малой, вследствие чего появляется возможность туннельного перехода электронов из металла в полупроводник, и обратный ток резко увеличивается. Кроме того, в обедненном слое полупроводника происходит тепловая генерация свободных носителей, что также увеличивает обратный ток.

Переходы металл — полупроводник (контакт Шоттки) отличаются от p — n -структур отсутствием накопления неосновных носителей при переменных процессах и высоким быстродействием.

Омический контакт. Омический контакт формируется переходом металл — полупроводник и характеризуется очень малым сопротивлением и линейной ВАХ. В хорошем омическом контакте падение напряжения при пропускании через него требуемого тока должно быть достаточно мало по сравнению с падением напряжения на активной области прибора. Наиболее важной характеристикой омического контакта является дифференциальное сопротивление при нулевом смещении $R_k = (dU/dI)|_{U=0}$.

В контакте металл — полупроводник с относительно низким уровнем легирования ($N_d < 10^{17} \text{ см}^{-3}$) преобладает термоэлектронная компонента тока. Для получения малых R_k в соответствии с выражением (2.32) нужно изготавливать контакт с малой высотой потенциального барьера. При высокой степени легирования полупроводника ($N_d > 10^{19} \text{ см}^{-3}$) будет преобладать туннельная компонента тока и удельное сопротивление контакта экспоненциально зависит от параметра $\varphi_{n,p}/\sqrt{N_d}$. Поэтому для получения малых R_k нужны как высокая степень легирования, так и малая высота потенциального барьера. В широкозонных полупроводниках, например GaAs, трудно изготовить контакт с малой высотой барьера при большой работе выхода из металла, вследствие чего для изготовления омических контактов создают дополнительный высоколегированный слой на поверхности полупроводника.

2.7. Гетеропереходы

Гетеропереход в общем случае может быть определен как граница раздела между двумя различными веществами (в частности, полупроводниками) с разной шириной запрещенной зоны. Если два рассматриваемых полупроводника имеют одинаковые типы проводимости, то переход называется **изотипным** гетеропереходом, в противном случае он называется **анизотипным**. Анизотипные гетеропереходы, как и гомопереходы, разделяются на n - p - и p - n -типа и являются структурами с неосновными носителями. До настоящего времени, в противоположность гомопереходам, не существует моделей, объясняющих все физические явления в гетеропереходах, поскольку в них свойства границы раздела сильно изменяются от материала к материалу и в значительной мере зависят от технологии изготовления. Существующие модели анизотипных гетеропереходов могут рассматриваться как обобщение общепринятой модели гомопереходов. Типичные диаграммы энергетических зон двух различных полупроводников p - и n -типов до контакта и резкого p - n -гетероперехода после тесного контакта в равновесном состоянии приведены на рис. 2.16, а, б.

Оба полупроводника имеют различные значения ширины запрещенной зоны $\Delta E_{s1,2}$, относительной диэлектрической проницаемости $\epsilon_{1,2}$, работы выхода $q\phi_{1,2}$ и электронного сродства $\chi_{1,2}$ (индекс «1» на рис. 2.16 относится к p -области, «2» — к n -области).

Как отмечено в п. 2.6, электронное сродство и работа выхода определяются как энергии, необходимые для удаления электрона соответственно со дна зоны проводимости E_n (E_{n1} или E_{n2} на рис. 2.16) и с уровня Ферми E_F (E_{F1} или E_{F2} на рис. 2.16) на уровень вакуума $E_{\text{вак}}$ (см. рис. 2.16) или на расстояние от поверхности, большее радиуса действия сил зеркального изображения, но меньшее размеров образца. За счет различной ширины запрещенных зон и диэлектрических проницаемостей контактирующих полупроводников на металлургической границе перехода наблюдаются разрывы в энергетических уровнях зоны проводимости (ΔE_n) и валентной зоны (ΔE_v), причем $\Delta E_n = q(\chi_1 - \chi_2)$ определяется разностью энергий электронного сродства двух полупроводников, а $\Delta E_v = (q\chi_2 + \Delta E_{s2}) - (q\chi_1 + \Delta E_{s1}) = q(\chi_2 - \chi_1) + (\Delta E_{s2} - \Delta E_{s1})$ включает также соответствующую разность для ширины зон проводимости.

В рассматриваемом типе гетероперехода обедненные слои образуются на каждой стороне от границы раздела, и, если не учитывать влияния границы раздела, объемные заряды этих слоев

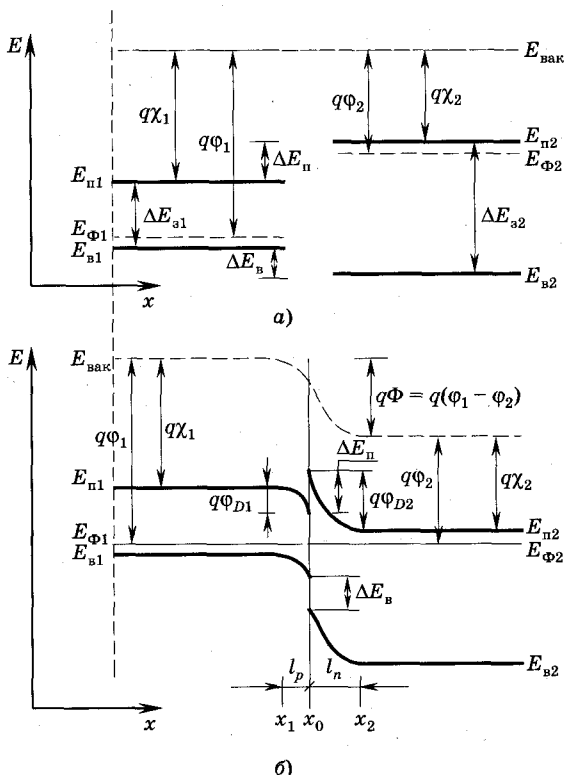


Рис. 2.16

противоположны по знаку и равны по величине, как и для гомоперехода.

Полная контактная разность потенциалов есть $\Phi = \phi_1 - \phi_2 = \phi_{D1} + \phi_{D2}$ (см. рис. 2.16, б), где ϕ_{D1} и ϕ_{D2} — электростатические потенциалы равновесного состояния соответствующих полупроводников.

Обобщая решение уравнения Пуассона для гомоперехода (см. п. 2.2), можно получить размеры обедненных областей с каждой стороны границы раздела резкого $p-n$ -гетероперехода:

$$l_p = x_0 - x_1 = \sqrt{\frac{2N_{a1}\epsilon_1\epsilon_2\Phi\epsilon_0}{qN_{d2}(\epsilon_1N_{a1} + \epsilon_2N_{d2})}}, \quad (2.33)$$

$$l_n = x_2 - x_0 = \sqrt{\frac{2N_{d2}\epsilon_1\epsilon_2\Phi\epsilon_0}{qN_{a1}(\epsilon_1N_{a1} + \epsilon_2N_{d2})}}, \quad (2.34)$$

где индекс «1» относится к p -полупроводнику, а «2» — к n -полупроводнику.

Проводя операции, аналогичные проведенным для p — n -гомоперехода (см. п. 2.4), получим выражение для барьерной емкости равновесного p — n -гетероперехода

$$|C_{\text{бар}}| = S \sqrt{\frac{qN_{a1}N_{d2}\varepsilon_1\varepsilon_2}{2(\varepsilon_1N_{a1} + \varepsilon_2N_{d2})} \cdot \frac{\varepsilon_0}{\Phi}}, \quad (2.35)$$

где S — площадь перехода.

В случае неравновесного перехода в выражениях (2.33)—(2.35) (так же как и в формулах (2.12) и (2.26) для гомоперехода) необходимо вместо Φ подставить разность $\Phi - U$, где U — приложенное к переходу напряжение.

В предположении, что вследствие разрывов краев зон на границе раздела диффузионный ток обусловлен электронами (это справедливо для рассматриваемых переходов из-за меньшего потенциального барьера для электронов, чем для дырок), ВАХ p — n -гетероперехода может быть описана следующим выражением:

$$I = A \exp(-q\phi_{D2}/kT) [\exp(qU_2/kT) - \exp((-qU_1/kT))], \quad (2.36)$$

где U_1 и U_2 — составляющие приложенного напряжения U , приходящиеся на полупроводники p - и n -типов, $A = SqXN_{d2} \times (D_{n1}/\tau_{n1})^{1/2}$, X — коэффициент пропускания электронов через границу раздела, D_{n1} и τ_{n1} — соответственно коэффициент диффузии и время жизни электронов в полупроводнике p -типа.

Первый член в квадратных скобках в формуле (2.36) определяет ток при прямом смещении, а второй — при обратном. Энергетическая диаграмма перехода при прямом смещении приведена на рис. 2.17.

Рассмотренная модель достаточно грубо описывает реальную ВАХ, однако, изменяя коэффициент A (за счет варьирования прозрачности барьера, т. е. коэффициента пропускания электронов через барьер) и соотношение между U_1 и U_2 , можно получить удовлетворительное совпадение расчетных и экспериментальных данных. Для объяснения расхождения теории и эксперимента и учета других механизмов переноса носителей были разработаны эмиссионная, эмиссионно-рекомбинационная, туннельная и туннельно-рекомбинационные модели. Однако и они не позволяют в полной мере с хорошей точностью описать ВАХ

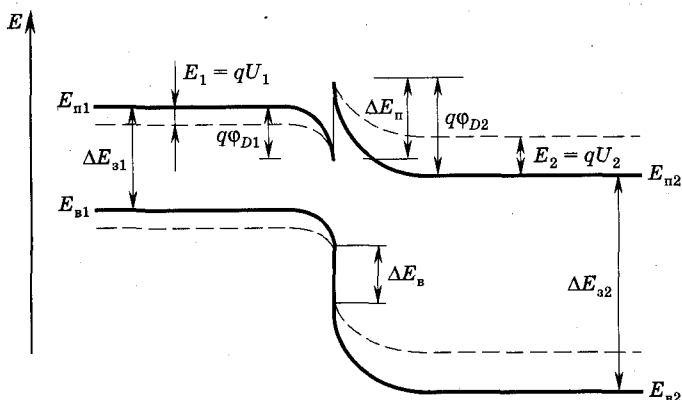


Рис. 2.17

гетеропереходов. В различных условиях и для различных переходов могут доминировать те или иные процессы или их совокупности.

Однако следует отметить, что модели на основе туннельного переноса электронов через барьер точнее и лучше других моделей описывают ВАХ.

В качестве примера на рис. 2.18 показана ВАХ прямо смещенного анизотипного перехода $(p)\text{Ge}-(n)\text{GaAs}$. Из-за использования полулогарифмического масштаба (по оси ординат — логарифмический масштаб, а по абсцисс — линейный) экспоненты вырождаются в прямые линии.

При комнатной температуре и ниже можно выделить две области на ВАХ. Так, при $U \leq 0,3$ В и при $T = 298$ К основной вклад в ток дает рекомбинационный ток. При $U \geq 0,3$ В ток обусловлен туннелированием электронов через переход.

Обратные характеристики переходов типа $(p)\text{Ge}-(n)\text{GaAs}$ при малых напряжениях линейны, т. е. $I_{\text{обр}} \sim U$, а при больших напряжениях $I_{\text{обр}} \sim U^m$, где $m > 1$.

Для других типов переходов обратные характеристики для

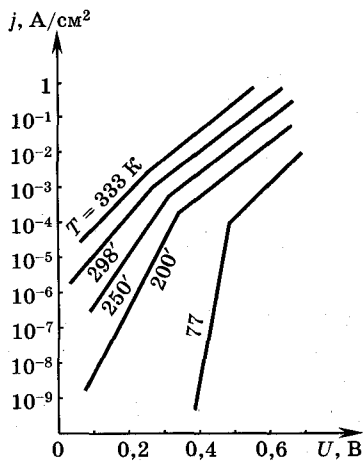


Рис. 2.18

$n-p$ и $p-p$ -гетеропереходов часто описываются соотношением типа

$$I_{\text{обр}} \sim \exp [A(\Phi - U)^{-1/2}], \quad (2.37)$$

где коэффициент A не зависит от температуры. Такое поведение характерно для туннельных токов.

Выражение для емкости анизотипного перехода в зависимости от приложенного напряжения (ВФХ) легко получить простым обобщением формулы (2.35) и метода вычисления емкости для гомоперехода (см. п. 2.2 и 2.4)

$$|C_{\text{бар}}| = \left[\frac{q\epsilon_1\epsilon_2 N_{\text{д}} N_{\text{а}} \epsilon_0}{2 \cdot (\epsilon_2 N_{\text{д}} + \epsilon_1 N_{\text{а}})} \right]^{1/2} (\Phi - U)^{-1/2} S. \quad (2.38)$$

Из этого выражения видно, что зависимость $C_{\text{бар}}^{-2}$ от приложенного обратного напряжения линейна и ее экстраполяция на ось напряжений (напряжение отсечки) дает контактную разность потенциалов, что и используется часто для определения Φ . Линейность ВФХ указывает на то, что гетеропереход является резким.

Изотипные резкие $n-n$ - и $p-p$ -гетеропереходы являются, подобно контактам Шоттки (см. п. 2.6), устройствами с основными носителями тока. Следовательно, в отличие от анизотипных гетеропереходов, в изотипных вклад неосновных носителей в электрический ток пренебрежимо мал. Энергетическая диаграмма резкого изотипного $n-n$ -гетероперехода приведена на рис. 2.19. Для резких $n-n$ -гетеропереходов, по аналогии с результатами эмиссионной теории для диодов на основе перехода металл — полупроводник (см. п. 2.6), ВАХ может быть описана следующей формулой (при условии $\phi_{D1} \ll \phi_{D2}$):

$$I = A_1 \exp(-q\phi_{D2}/(kT)) [\exp(qU_2/(kT)) - \exp(qU_1/(kT))], \quad (2.39)$$

где $U = U_1 + U_2$ — приложенное напряжение; $A_1 = SqN_{\text{д}2} \times (kT/2m_e^*)^{1/2}$; X — коэффициент пропускания электронов через поверхность раздела; m_e^* — эффективная масса электрона. Поскольку $U_1 \ll U_2$, выражение (2.39) можно упростить. В результате получим

$$I = A_1 \exp(-q\phi_{D2}/(kT)) [\exp(qU_2/(kT)) - 1]. \quad (2.40)$$

Для описания имеющегося рассогласования экспериментальных и расчетных данных для некоторых типов изотипных пе-

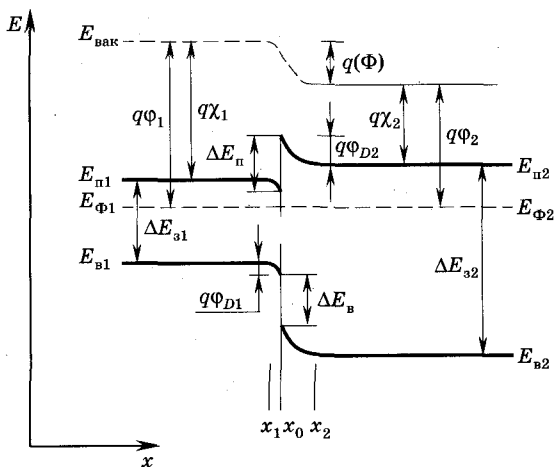


Рис. 2.19

переходов используются диффузионная модель, модель двойного диода Шоттки и туннельные модели.

В случае, когда электронная составляющая полного тока является преобладающей, ВАХ $n-n$ -гетеропереходов подобны ВАХ системы из двух диодов Шоттки, соединенных последовательно навстречу друг другу.

Для многих гетеропереходов, например $(n)\text{Ge}-(n)\text{GaAs}$, прямые ветви ВАХ могут быть описаны выражением типа $I \sim \exp(qU/\eta kT)$ (модификация формулы (2.40), где η — величина, близкая к единице). Для других случаев, в частности для переходов $\text{Ge}-\text{Si}$, более подходящей является модель, представляющая гетеропереход в виде двух барьеров Шоттки, включенных навстречу друг другу. Эта модель объясняет, например, насыщение тока, наблюдающееся в переходах $(n)\text{Ge}-(n)\text{Si}$ при некоторых уровнях легирования как германия, так и кремния (рис. 2.20). На рис. 2.20 представлены равновесные диаграммы энергетических зон и ВАХ трех гетеропереходов $(n)\text{Ge}-(n)\text{Si}$ с различной концентрацией примеси. Для других разновидностей изотипных гетеропереходов ВАХ имеют часто свои особенности.

ВФХ таких приборов при $N_{д1} \gg N_{д2}$ аналогично контакту металл — полупроводник приближенно можно описать выражением [40]

$$C = S \left[\frac{q\epsilon_2 N_{д2}}{2} \right]^{1/2} (\Phi - U)^{-1/2}, \quad (2.41)$$

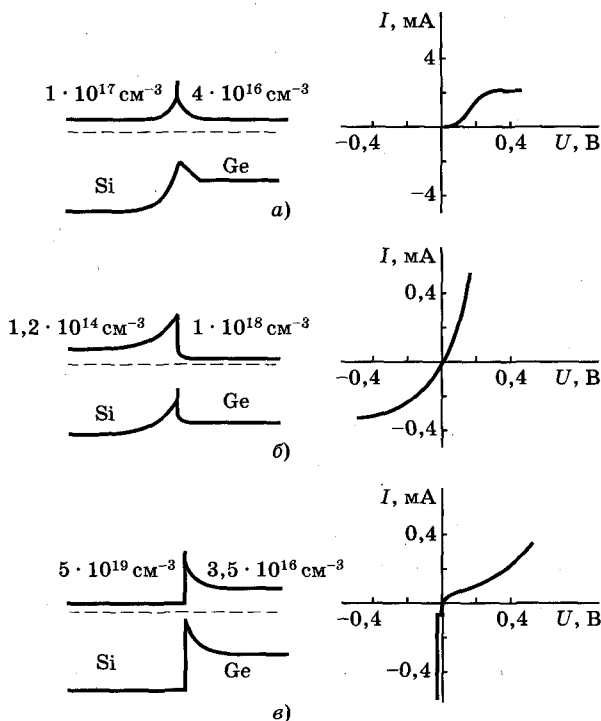


Рис. 2.20

из которого следует, что зависимость $C^{-2} = f(U)$ линейна, как и для анизотипного перехода.

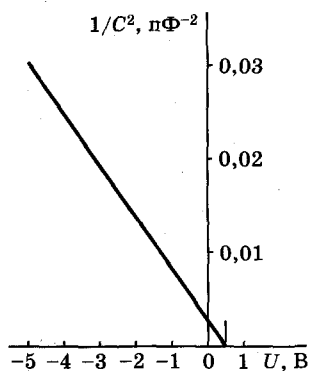


Рис. 2.21

Гетеропереходы Ge—GaAs отличаются почти точным согласованием решеток, поэтому они хороши для проверки различных теоретических моделей. Зависимость $C^{-2} = f(U)$ для такого типа переходов является линейной, и контактная разность потенциалов Φ (пересечение характеристики с осью абсцисс) равна 0,48 В для n - n -гетероперехода (n)Ge—(n)GaAs и 0,37 В — для (p)Ge—(p)GaAs. Аналогичная зависимость реализуется для многих изотипных переходов, в частности для гетероперехода (n)InP—(n)GaAs (рис. 2.21), у которого

величина Φ больше, чем в предыдущем случае, поскольку ширина запрещенной зоны InP (0,18 эВ) существенно меньше ширины запрещенной зоны Ge (0,72 эВ). Гетеропереходы широко используются в полупроводниковых приборах, в частности в светодиодах (п. 16.3), полупроводниковых лазерах (гл. 21) и др.



Контрольные вопросы

1. Виды электрических переходов и их характеристики.
2. Каковы физические процессы в равновесном переходе? Формирование обедненной области и барьерной емкости $p-n$ -перехода.
3. Объясните процессы в $p-n$ -переходе при прямом и обратном смещении: инжекция, экстракция неосновных носителей.
4. Каковы ВАХ идеализированного $p-n$ -перехода. В чем состоит идеализация ВАХ? ВАХ реального $p-n$ -перехода. Каковы физические механизмы пробоя $p-n$ -перехода?
5. ВФХ и эквивалентная схема $p-n$ -перехода.
6. Переходные процессы в $p-n$ -переходе.
7. Контакт металл — полупроводник (энергетические диаграммы и ВАХ; ее отличие от ВАХ $p-n$ -перехода).
8. Виды гетеропереходов и их энергетические диаграммы.
9. ВАХ и ВФХ изотипных и анизотипных гетеропереходов.

Глава 3

ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

3.1. Общие сведения и классификация диодов

Классификация полупроводниковых диодов. Полупроводниковые диоды относятся к электропреобразовательным приборам с двумя выводами. Несмотря на большое разнообразие и широкую номенклатуру диодов, основная масса их содержит полупроводник с одним электрическим переходом и омическими

контактами к областям полупроводника, формирующим этот переход. К диодам, которые не имеют электрического перехода, относятся диоды на основе *эффекта Ганна*. Некоторые типы полупроводниковых приборов имеют два, три и более переходов, например *p-i-n-диоды*, *диодные тиристоры* и др. (гл. 5). В качестве выпрямляющих переходов в полупроводниковых диодах применяются электронно-дырочные переходы, контакт металл — полупроводник и гетеропереходы. В диодах с *p-n*-переходом имеются два омических контакта, а в диодах на основе металл — полупроводник — один омический переход.

Обычно полупроводниковые диоды выполняются с использованием несимметричного *p⁺-n*-перехода или перехода металл — полупроводник. Слаболегированная область диода на основе *p⁺-n*-перехода, в которую преимущественно осуществляется инжекция носителей зарядов из сильнолегированной области, называется *базой диода*, а сильнолегированная область является *эмиттером*.

Рассмотренные в гл. 2 физические процессы в электрических переходах присутствуют в основной массе диодов, поэтому, несмотря на разнообразие марок и типов, в основе их лежат одни те же физические явления.

Классификация современных полупроводниковых диодов осуществляется в соответствии с их назначением, физическими свойствами, основными электрическими параметрами, конструктивно-технологическими признаками, исходным полупроводниковым материалом и т. д., что находит отражение в системе условных обозначений диодов. В основу маркировки диодов положен буквенно-цифровой код, отражающий информацию об исходном полупроводниковом материале, подклассе прибора, его назначении, порядковом номере разработки. Условные обозначения включают также классификационные признаки по параметрам, конструктивным особенностям и т. д.

В зависимости от типа полупроводникового материала различают кремниевые, германиевые, арсенид-галлиевые, селеновые и карбид-кремниевые диоды. Большинство диодов выполняется из кремния. Кремниевые диоды имеют большие значения максимальной рабочей температуры (125...150 °С) и малый обратный ток. Арсенид-галлиевые диоды обладают лучшими параметрами по сравнению с кремниевыми, но имеют более сложную технологию изготовления.

В соответствии с конструктивно-технологическими особенностями изготавливают *плоскостные, точечные и микросплавные диоды*.

По выполняемым функциям различают *выпрямительные, импульсные, преобразовательные, переключательные, детекторные диоды, стабилитроны, варикапы и параметрические диоды, светодиоды, полупроводниковые лазерные диоды, фотодиоды* и др. У некоторых диодов в названии отображаются основные физические процессы в переходе, например *туннельный диод, лавинно-пролетный диод*, или характер преобразования энергии сигнала, например *светодиод, фотодиод* (гл. 16), и т. д.

Классификацию отдельных подклассов диодов проводят в зависимости от диапазона рабочих частот — *низкочастотные, высокочастотные, сверхвысокочастотные (СВЧ)*.

Диоды различного назначения отличаются структурой, площадью и формой перехода (плоский, точечный, полусферический и др.), распределением концентрации примесей (ступенчатое, линейное и др.), режимами работы перехода, конструкцией корпуса, системой справочных параметров.

Конструкция корпуса низкочастотных мощных (силовых) выпрямительных диодов является массивной для обеспечения хорошего теплоотвода. *Высокочастотные диоды* имеют миниатюрные пластмассовые или стеклянные корпуса. Конструкция корпуса СВЧ-диодов приспособлена для размещения их в волноводных трактах. В гибридных интегральных схемах и микросборках применяются бескорпусные диоды.

Система параметров диодов. Система параметров диодов включает большое число наименований. Параметры диодов подразделяются на предельно допустимые, минимально и (или) максимально допустимые и рабочие (характеризующие).

Под предельными эксплуатационными параметрами понимаются максимально допустимые значения токов, напряжений, рассеиваемых мощностей, температур и других параметров, при которых гарантируется работоспособность диода. Превышение указанных значений приводит к выходу прибора из строя или возможному его повреждению. *Допустимые значения параметров определяют удовлетворительную работу прибора. Предельно допустимые значения параметров нельзя измерить, их можно проверять на основе опытов, испытаний (иногда разрушающих) или путем расчетов.*

Рабочее (характеризующее) значение параметра отражает определенное свойство прибора. В качестве таких параметров могут выступать электрические, тепловые, механические и другие величины. Рабочие значения параметров можно непосредственно измерить.

Многие приборы являются элементами полупроводниковых интегральных схем, т. е. они не являются самостоятельными изделиями. Для таких диодов важны только физические процессы и электрические параметры, связанные со структурой, например тепловой ток перехода, барьерные емкости, сопротивление базы, время жизни носителей и т. д. Диоды интегральных схем с использованием $p-n$ -переходов, как правило, выполняются на основе биполярных транзисторов, а диоды с использованием контакта металл — полупроводник совмещены с транзисторами в единую структуру (гл. 7).

Технология создания полупроводниковых приборов и интегральных схем. Для создания электрических переходов диодов и других полупроводниковых структур и полупроводниковых приборов с необходимыми электрическими параметрами используются различные технологические методы. К наиболее важным из них относятся: *диффузия примесей, ионное легирование, эпитаксия, сплавление, локальная диффузия* и ряд других физико-технологических приемов.

Технологический прием *сплавление* реализуется следующим образом. На очищенную поверхность полупроводниковой пластины с определенным типом проводимости (обычно n -типа) помещается таблетка или тонкая проволока металлического материала (часто акцепторного типа, например, алюминия, при использовании кремния). При нагреве до температуры $900...1000$ К алюминий плавится и растворяет прилегающий слой кремния, температура плавления которого выше. В результате рекристаллизации у поверхности пластины образуется слой кремния (Si) p -типа, насыщенный алюминием. Недостатками этого метода являются плохая воспроизводимость параметров из-за неконтролируемости процесса и низкие рабочие токи из-за невысокого пробивного напряжения для микросплавных диодов. Для устранения этого недостатка используется не таблетка, а тонкая алюминиевая проволока.

При *диффузии* примесные атомы обычно из газовой фазы проникают через поверхность в полупроводниковую пластинку за

счет теплового движения при высокой температуре (~ 1300 К). Чем больше время диффузии или чем выше температура, тем глубже атомы проникают в пластину. Параметры процесса в этом случае контролируются, что обеспечивает лучшую воспроизводимость, по сравнению с сплавлением, электрических параметров полупроводниковых структур. Концентрация введенных атомов примесей является в этом процессе пространственно неравномерной. Она максимальна у поверхности и уменьшается при удалении от нее.

Электронно-дырочный переход, сформированный диффузией, получается плоским, с большой площадью, равной площади исходной пластинки. Это позволяет получить большие рабочие токи, что необходимо в мощных силовых дискретных диодах.

В микроэлектронике при изготовлении полупроводниковых интегральных схем используют методы *локальной диффузии*, при которой диффузия примесей осуществляется через отверстие в тонкой маске из оксида кремния SiO_2 , формируемой на поверхности кремниевой пластины термическим окислением. Оксид кремния не пропускает донорные атомы фосфора и акцепторные атомы бора. Отверстие в маске образуется путем *фотолитографии*. Максимальная площадь формирующегося таким методом перехода определяется разрешающей способностью фотолитографии, задающей минимальный размер отверстия, и толщиной диффузионного слоя. Возможности современной технологии позволяют формировать переходы с площадью в несколько квадратных мкм.

Ионное легирование осуществляется ускоренными до энергий в десятки и сотни кэВ ионами примесей, бомбардирующими поверхность полупроводника. Глубина проникновения ионов составляет менее одного микрона, что позволяет получать очень тонкие слои. Легирование производится через отверстие в маске. Поскольку ионный пучок направлен строго перпендикулярно поверхности полупроводника и боковое рассеивание невелико, то размеры легированной области точно соответствуют отверстию в маске. В результате удается получить переходы с меньшей площадью (менее 1 мкм^2), чем с помощью локальной диффузии. Необходимое распределение концентрации примесей достигается регулированием энергии и плотности потока ионов за счет ускоряющего напряжения и времени облучения.

Эпитаксия — это процесс наращивания при $T \sim 1300$ К на пластину полупроводника (подложку) монокристаллического слоя толщиной $1 \dots 15$ мкм, который воспроизводит структуру подложки.

В результате формируется единичный монокристалл. Обычно материал наращиваемого слоя и подложки одинаков, но возможно применение и различных материалов с близкой кристаллической структурой, например при формировании гетеропереходов. Для создания многослойной структуры проводят несколько последовательных эпитаксий.

Для получения очень тонких (до нескольких нанометров) структур, близких к ступенчатым, используют молекулярно-лучевую технологию (нанотехнологию), осуществляемую в сверхвысоком вакууме. Молекулярные пучки получают за счет испарения компонентов пленки.

3.2. Выпрямительные диоды

По своим функциональным свойствам выпрямительные диоды можно разделить на низкочастотные и высокочастотные.

Выпрямительные низкочастотные диоды. Для преобразования переменного тока в постоянный предназначены **выпрямительные низкочастотные диоды**. Они выполняются обычно на основе плоскостного перехода. Выпрямительный диод должен иметь низкое сопротивление при прямом смещении и очень высокое при обратном. Особенности физических процессов при выпрямлении можно проиллюстрировать на примере схемы однополупериодного выпрямителя (VD) (рис. 3.1, а), где $U_{\text{ип}}$ — напряжение источника переменного тока, $R_{\text{н}}$ — сопротивление нагрузки. В положительный полупериод напряжение на нагрузке равно $U_{\text{н}} = U_{\text{ип}} - U_{\text{пр}} \approx U_{\text{ип}}$, поскольку падение напряжения на диоде $U_{\text{пр}}$ при прямом смещении мало. Ток в цепи диода равен $I_{\text{д}} = U_{\text{ип}}/R_{\text{н}}$. При обратном смещении через диод течет малый ток $I_{\text{обр}}$, поэто-

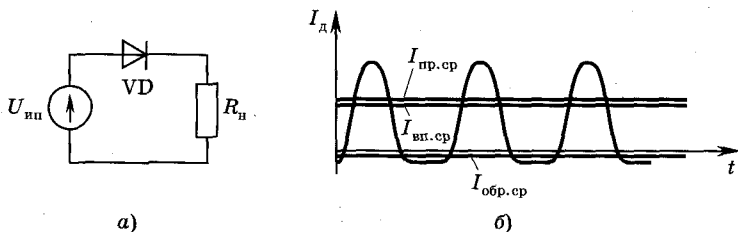


Рис. 3.1

му $U_n \approx 0$, а напряжение на диоде $U_d \approx U_{ип}$ (рис. 3.1, б). Работа выпрямителя характеризуется следующими параметрами: $I_{вп. ср} = I_{пр. ср} - I_{обр. ср}$ — средний выпрямительный ток, являющийся разностью усредненных за период прямого тока $I_{пр. ср}$ и обратного $I_{обр. ср}$ (см. рис. 3.1, б).

К другим основным динамическим параметрам относятся: $U_{пр. ср}$ — среднее значение прямого падения напряжения при заданном среднем значении прямого тока; $U_{обр. ср}$ — среднее за период значение обратного напряжения. Величины $U_{пр. ср}$ при заданном $I_{пр. ср}$, $I_{обр. ср}$ и максимальном амплитудном значении обратного напряжения $U_{обр. макс}$ являются основными электрическими параметрами диода.

Верхняя частотная граница работы выпрямительных силовых диодов обычно не превышает 50 кГц. Основными эксплуатационными параметрами выпрямительных диодов являются следующие: $U_{обр. макс}$ — максимально допустимое обратное напряжение, при котором диод может длительное время функционировать без нарушения работоспособности; обычно $U_{обр. макс} \leq (0,5...0,8) U_{проб}$ ($U_{проб}$ — напряжение пробоя); $I_{пр. макс}$ — максимально допустимый постоянный прямой ток; $U_{пр}$ — постоянное прямое напряжение при заданном прямом токе $I_{пр} = I_{пр. макс}$; $I_{обр. макс}$ — максимально допустимый обратный ток; $P_{макс}$ — максимально допустимая мощность, рассеиваемая диодом.

При протекании тока через диод происходит повышение температуры перехода. В установленном режиме подводимая к переходу и отводимая от него мощности должны быть равны и не превышать $P_{макс}$, иначе возможен тепловой пробой. В связи с этим одним из важных параметров диодов является диапазон температур окружающей среды и максимальная температура корпуса, при которых гарантируется нормальная температура диодов. Для кремниевых диодов регламентируемая температура окружающей среды от -60 до 125 °С. Для снижения температуры перехода в силовых выпрямительных диодах используются специальные радиаторы, позволяющие отводить избыток тепла.

Для кремниевых диодов с $p-n$ -переходом, имеющих наибольшее распространение, $U_{пр. ср} \approx 1...1,5$ В при $T^\circ = 20$ °С. С ростом температуры прямое напряжение $U_{пр. ср}$ уменьшается. Обратный ток таких диодов при $T^\circ = 20$ °С обычно не превышает

десятых долей мкА и увеличивается при повышении температуры. Напряжение пробоя для кремниевых диодов составляет сотни вольт и увеличивается с ростом температуры.

По максимально допустимому выпрямленному току диоды разбиты на три группы: диоды малой мощности ($I_{пр} \leq 0,3 \text{ А}$), диоды средней мощности ($0,3 \text{ А} < I_{пр} < 10 \text{ А}$) и мощные (силовые) диоды ($I_{пр} \geq 10 \text{ А}$). Предельные токи некоторых марок силовых выпрямительных диодов могут превышать 1500 А с $U_{обр. макс}$ от 100 до 4000 В и более.

Для повышения предельных токов и напряжений изготавливают **выпрямительные столбы**, состоящие из кремниевых сплавных или диффузионных последовательно соединенных диодов, количество которых может превышать полтора десятка единиц. Обратные токи и напряжения в выпрямительных столбах подбираются одинаковыми, иначе на одном из диодов обратное напряжение может оказаться выше пробивного и весь столб выйдет из строя. Для обеспечения идентичности ВАХ и параметров диодов их изготавливают в едином технологическом цикле подобно интегральным схемам.

Выпрямительные высокочастотные диоды. Диоды этого типа используются в детекторах, смесителях, преобразователях частоты, ограничителях и других устройствах и осуществляют нелинейное электрическое преобразование сигналов на частотах до сотен мегагерц.

Поскольку рассматриваемые диоды работают на очень высоких частотах, для них важны такие динамические параметры, как высокая граничная частота работы $f_{гр}$, индуктивность диода L_d , емкости диода $C_{корп}$, $C_{бар}$, сопротивление базы r_b . Параметры L_d , $C_{корп}$, $C_{бар}$, r_b и время жизни неосновных носителей в базе для высокочастотных выпрямительных диодов должны иметь возможно меньшие значения, что достигается уменьшением площади перехода за счет применения точечных контактов, легированием базы золотом, которое снижает время жизни носителей, и т. д.

Простейшая эквивалентная схема (электрическая модель) высокочастотного диода для большого сигнала, справедливая для линейных детекторов, ограничителей и т. д., приведена на рис. 3.2, а. Эта модель отличается от аналогичной модели $p-n$ -перехода наличием конденсатора $C_{корп}$, который учитывает пара-

зитные емкости выводов и корпуса диода. Резистор r_6 учитывает сопротивление базы и контактов, $r_{\text{диф}}$ — дифференциальное сопротивление перехода, R_y — сопротивление утечки, L_d определяется индуктивностью выводов диода. Емкость перехода $C_{\text{пер}} = C_{\text{диф}} + C_{\text{бар}}$. Следует отметить, что для диодов на основе контакта металл—полупроводник (диоды Шоттки) $C_{\text{диф}}$ исключается из схемы на рис. 3.2, а. При малых сигналах и прямых смещениях эквивалентная схема принимает вид, изображенный на рис. 3.2, б.

Выпрямительные высокочастотные диоды могут выполнять различные схемные функции, т. е. в этом смысле они универсальны. Поэтому их электрические параметры являются общими для всего класса диодов и не отражают специфики применения.

Как уже отмечалось, для увеличения быстродействия диодов необходимо снижать значения всех параметров, указанных на рис. 3.2. Отметим характерные значения некоторых из параметров. Емкость высокочастотных диодов, как правило, не превышает 10 пФ (указывается на заданной частоте). Время восстановления обратного сопротивления для высокочастотных диодов лежит в пределах от 0,1 до единиц нс. Минимальные значения соответствуют диодам Шоттки.

СВЧ выпрямительные диоды. Эти диоды, как и высокочастотные, используются в детекторах, преобразователях частоты, переключателях, ограничителях и т. д., но функционируют они на частотах СВЧ-диапазона от 300 МГц до 100 ГГц.

Главным требованием, предъявляемым к СВЧ-диодам, являются малые значения емкости диода (C) и сопротивления базы (r_6), которые соответственно должны быть порядка десятых долей пФ и единиц Ом. Постоянная времени Cr_6 при работе на СВЧ не должна превышать периода СВЧ-колебаний T_k ; как правило, $Cr_6 \ll T_k$. Наилучшими параметрами C и r_6 характеризуются пла-

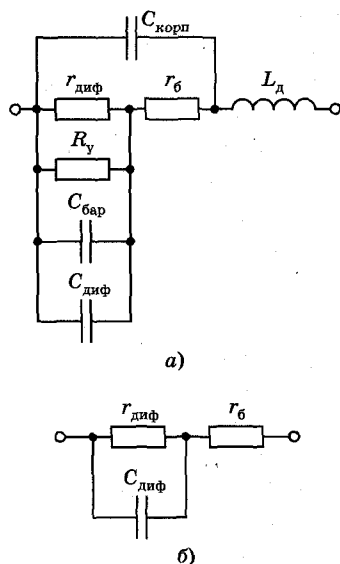


Рис. 3.2

нарно-эпитаксиальные плоскостные диоды на основе арсенида галлия. Для них при диаметре контакта порядка 1 мкм постоянная времени удовлетворяет неравенству $\tau_0 < 10^{-12}$ с, что обеспечивает работоспособность диода на частотах вплоть до 100 ГГц.

Для обеспечения малых значений паразитных емкостей и индуктивностей выводов корпуса СВЧ-диодов часто выполняются в виде элементов фидерных линий; так, например, при использовании коаксиальных линий выводы диодов имеют коаксиальную конструкцию, легко интегрируемую с СВЧ-линией; в волноводных линиях диод с контактами монтируют непосредственно в отрезке волновода. Бескорпусные диоды используют в микрополосковых линиях.

Специфика применения диодов в различных устройствах СВЧ отражается набором характерных электрических параметров. Основными специфическими параметрами *детекторных диодов* являются: *рабочая длина волны λ* , *чувствительность по току $\beta_m = I_{\text{вп}}/P_{\text{ид}}$* ($I_{\text{вп}}$ — выпрямленный ток, $P_{\text{ид}}$ — мощность СВЧ-сигнала), а также *коэффициент стоячей волны (КСВ)*. Параметр β_m измеряется в режиме короткого замыкания цепи выпрямленного тока и составляет $\beta_m \sim 1$ А/Вт при $\lambda = 3 \dots 10$ см и $P_{\text{ид}} = 0,1$ мВт.

Важнейшими характеристиками *смесительных диодов* являются *рабочая длина волны*, *потери преобразования*, *относительная шумовая температура*, *выпрямленный ток* (при заданной мощности СВЧ-сигнала), КСВ и ряд других параметров.

Потери преобразования (в дБ) определяются величиной $L_{\text{пр}} \text{ (дБ)} = 10 \lg (P_{\text{вх}}/P_{\text{п.ч}})$, где $P_{\text{вх}}$ и $P_{\text{п.ч}}$ — соответственно мощности входного и выходного СВЧ-сигналов на промежуточной частоте. Для снижения потерь в смесительных, так же как и в детекторных, диодах, надо уменьшать прямое напряжение, обратный ток, емкость, сопротивление базы и увеличивать нелинейность ВАХ. Типичные значения потерь преобразования составляют $L_{\text{пр}} \sim 5 \dots 10$ дБ. Относительная шумовая температура $t_{\text{ш}} = P_{\text{ш}}/(kT \Delta f)$, где $P_{\text{ш}}$ — мощность шумов в диоде, Δf — полоса пропускания. Параметр $t_{\text{ш}}$ важен для приемных устройств. Типичные значения $t_{\text{ш}}$ лежат в пределах 2...3 при $T = 293$ К.

Одними из самых распространенных видов СВЧ-диодов являются *переключающие диоды*, которые предназначены для коммутации сигналов в различных СВЧ-устройствах. Эти диоды включаются в СВЧ-тракт через управляющее устройство, которое задает постоянную составляющую напряжения или тока че-

рез диод. В зависимости от управляющего сигнала диод может находиться или в открытом состоянии, закорачивая тем самым СВЧ-линию, или в закрытом. В первом случае СВЧ-тракт закрыт и коэффициент отражения близок по модулю к единице, а во втором — СВЧ-линия открыта и коэффициент отражения близок к нулю.

Переключающий диод имеет очень малую барьерную емкость $C_{бар} = C_{выв}$ и при прямом смещении малое суммарное сопротивление, состоящее из сопротивления перехода ($r_{диф}$) и сопротивления базы и контактов (r_6). Обратное напряжение, прямой ток и рассеиваемая мощность должны быть достаточно большими. Чем больше допустимая рассеиваемая мощность, тем больше и максимальная мощность коммутируемого СВЧ-сигнала.

Для уменьшения $C_{бар}$ необходимо уменьшать площадь перехода, а для увеличения прямого тока, наоборот, площадь перехода делается большей. Из-за противоречивости этих требований использование $p-n$ -переходов и переходов металл — полупроводник не позволяет достичь наилучших параметров при переключении. Поэтому в качестве переключающих диодов используются $p-i-n$ -структуры, в которых между сильно легированными областями n^+ - и p^+ -типа располагается i -слой с малой проводимостью, близкой к собственной. Из-за малой проводимости i -слоя даже при нулевом напряжении он практически полностью обеднен носителями заряда. В то же время из-за высокой степени легирования n^+ - и p^+ -областей обедненный слой в них имеет очень малую протяженность. Как результат, толщина обедненного слоя $l_{об}$ приблизительно равна ширине i -слоя (единицы — десятки мкм) и барьерная емкость практически не зависит от напряжения. Следовательно, в $p-i-n$ -диодах можно получить малые, практически не меняющиеся от напряжения емкости диода (десятые доли пФ) при достаточно больших напряжениях пробоя (до 1000 В) и площадях перехода, которые обеспечивают большой допустимый прямой ток (до сотен мА).

При подаче прямого смещения за счет инжекции носителей из n^+ - и p^+ -областей i -слой заполняется электронами и дырками, которые при больших прямых токах образуют квазинейтральную электронно-дырочную плазму. Из-за большой концентрации носителей сопротивление i -слоя очень мало (менее 0,1 Ом). Для его уменьшения необходимо увеличивать время жизни носителей в i -слое. Скорость накопления и рассасывания носите-

лей зависит от скорости процесса рекомбинации и дрейфовой скорости носителей. Эти процессы определяют один из основных параметров переключаательных диодов — **время переключения**. Теоретически предельное время переключения есть величина порядка времени пролета через i -слой носителей заряда, движущихся с предельной дрейфовой скоростью, т. е. скоростью насыщения.

Основные специфические параметры переключающих диодов как приборов СВЧ:

— **потери пропускания** в открытом состоянии тракта (десятые доли дБ);

— **потери запираания** в закрытом состоянии (десятки дБ);

— **время прямого восстановления** при подаче на диоды функции включения (от нескольких нс до нескольких мкс);

— **время обратного восстановления** при подаче обратного напряжения (от нескольких нс до нескольких мкс).

Указанные параметры определяют времена переключения СВЧ-тракта с диодом из открытого в закрытое состояние или наоборот, в течение которых потери достигают заданных значений.

Характерные предельно допустимые параметры: **переключаемая в непрерывном режиме СВЧ-мощность** (до единиц Вт), **импульсная СВЧ-мощность** (до 1 кВт) и **рассеиваемая мощность**.

3.3. Импульсные диоды

Этот тип диодов используется в ключевых и цифровых схемах, а также в устройствах преобразования и формирования импульсных сигналов. Импульсные диоды должны иметь малую длительность переходных процессов, чтобы обеспечить импульсный режим работы. Современные импульсные диоды позволяют работать с сигналами длительностью до наносекундного диапазона. Основные параметры, определяющие инерционность диодов, — это барьерная емкость и скорость рассасывания накопленного в базе заряда носителей (см. п. 2.5).

Для уменьшения значений указанных параметров в импульсных диодах делают малую площадь электрического перехода и снижают время жизни неравновесных носителей заряда в базе за счет дополнительного легирования ее атомами золота, которые создают в запрещенной зоне кремния уровни ловушек. Основными исходными материалами для изготовления

импульсных диодов являются кремний и арсенид галлия. Помимо $p-n$ -перехода, в импульсных диодах используется контакт металл — полупроводник. В соответствии с используемой технологией импульсные диоды бывают точечные, сплавные, микросплавные, диффузионные, мезадиффузионные и эпитаксиально-планарные.

Основным импульсным параметром является **время восстановления обратного сопротивления** — $t_{\text{вос}}$, определяемое как временной интервал от момента подачи обратного напряжения до момента, когда обратный ток диода уменьшается до заданного значения. Для сверхбыстродействующих диодов $t_{\text{вос}} < 0,1$ мкс, а для быстродействующих — $0,1 \leq t_{\text{вос}} \leq 10$ мкс.

Помимо $t_{\text{вос}}$, другим импульсным параметром является $t_{\text{уст}}$ — **время установления прямого сопротивления** диода, которое определяет интервал времени от начала импульса прямого тока до момента уменьшения напряжения на диоде до 1,2 установившегося значения.

К специфическим импульсным параметрам относятся также следующие: $r_{\text{имп}}$ — **максимальное импульсное сопротивление**, представляющее отношение максимальной амплитуды импульса прямого напряжения на диоде к току через него; $I_{\text{вос. макс}}$ — **максимальный ток восстановления**, определяемый наибольшим обратным током после переключения напряжения на диоде с прямого на обратное; C_d — **емкость диода** при заданном обратном напряжении; $Q_{\text{пк}}$ — **заряд переключения**, переносимый обратным током после переключения диода с прямого на обратное напряжение при соответствующих значениях прямого и обратного напряжений; $U_{\text{пр. и. макс}}$ — **максимальное падение напряжения на диоде** в прямом направлении при заданной силе импульсного прямого тока.

Наименьшее время переключения имеют диоды Шоттки, в которых практически отсутствует эффект накопления неосновных носителей заряда.

К импульсным диодам относятся также **диоды с накоплением заряда (ДНЗ)**.

Импульсные ДНЗ конструируются таким образом, чтобы они запасали заряд во время протекания прямого тока, а после переключения пропускали ток в обратном направлении за короткое время. Таким образом, эти диоды отличаются тем, что в них при переключении с прямого напряжения на обратное длительность стадии восстановления обратного сопротивления ма-

ла по сравнению с длительностью стадии высокой обратной проводимости, в результате чего формируется почти прямоугольный импульс обратного тока. Для реализации таких процессов толщина базы должна быть намного меньше диффузионной длины, а на границе n -базы с n^+ -подложкой формируется тормозящее электрическое поле $n-n^+$ -перехода, не пропускающее неосновные носители (в данном случае дырки) в подложку. Носители накапливаются вблизи $p-n$ -перехода (с другой стороны базы) и рассасываются за эффективное время жизни ($\tau_{эф}$) неосновных носителей в базе, которое в диодах с накоплением заряда составляет десятки-сотни наносекунд. Диффузионная длина L_p не превышает 10 мкм при толщине базы порядка 1 мкм.

Диоды с накоплением заряда используются в генераторах наносекундных импульсов, в умножителях и делителях частоты, а также других быстродействующих схемах.

3.4. Стабилитроны

Стабилитрон — это полупроводниковый диод, предназначенный для стабилизации напряжения в электронных схемах. Стабилитроны применяются во вторичных источниках питания, ограничителях, источниках опорного напряжения и т. д. Простейшая схема стабилизации напряжения приведена на рис. 3.3, а.

В основе работы стабилитронов, за исключением одного типа — **стабисторов**, лежит лавинный или туннельный пробой в $p-n$ -переходе, т. е. стабилитроны работают при обратном смещении. Лавинный механизм пробоя реализуется при относительно малом уровне легирования примесями базы диода. На-

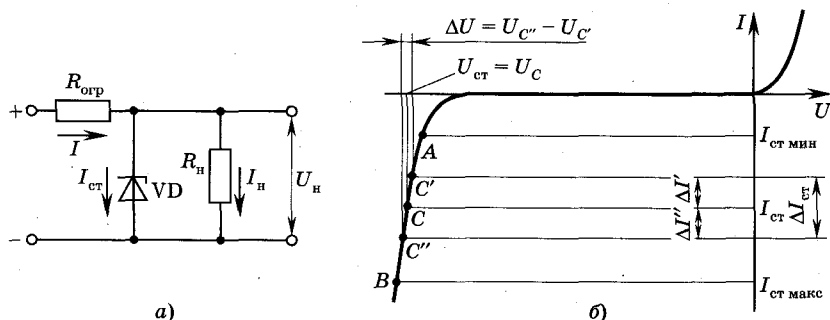


Рис. 3.3

пряжение стабилизации при лавинном пробое, как правило, превышает 6 В. При большой концентрации примесей возникает туннельный пробой, напряжение стабилизации при котором менее 6,3 В. Как следует из ВАХ (рис. 3.3, б), напряжение на участке пробоя слабо зависит от протекающего через прибор тока. Это напряжение практически равно *напряжению стабилизации* $U_{ст}$. *Минимальный ток стабилизации* $I_{ст. мин}$ соответствует началу пробоя, т. е. началу вертикального участка ВАХ (точка А на рис. 3.3, б), где дифференциальное сопротивление $r_{диф}$ резко уменьшается и становится малым. *Максимальный ток стабилизации* $I_{ст. макс}$ (точка В на рис. 3.3, б) регламентируется *допустимой мощностью рассеивания*. Напряжение стабилизации $U_{ст}$ определяется при заданном значении тока стабилизации $I_{ст}$ (точка С на рис. 3.3, б) на рабочем участке АВ стабилитрона.

Рассмотрим работу простейшей схемы стабилизации напряжения (см. рис. 3.3, а), когда напряжение на резисторе нагрузки R_n равно $U_n = U_{ст}$. Ток через ограничивающий резистор $R_{огр}$ равен $I = (E - U_{ст})/R_{огр}$, при этом ток через стабилитрон $I_{ст} = I - I_n$, а ток нагрузки $I_n = U_{ст}/R_n$ (точка С на рис. 3.3, б). Предположим, что напряжение на входе схемы изменилось на величину $\Delta U''$. В рабочей точке С дифференциальное сопротивление диода $r_{диф}$ много меньше общего сопротивления схемы, поэтому ток через стабилитрон при увеличении входного напряжения на $\Delta U''$ возрастет на величину $\Delta I'' = \Delta U''/R_{огр}$ и примет значение, соответствующее точке С''. Возрастание тока через стабилитрон приведет к увеличению общего тока I , протекающего через резистор $R_{огр}$, также на величину $\Delta I''$, что приведет к возрастанию падения напряжения на резисторе $R_{огр}$. В результате напряжение на нагрузке практически не изменится. Подобные процессы будут протекать и при изменении напряжения на $-\Delta U'$. В этом случае рабочая точка переместится в положение С', а ток уменьшится на величину $\Delta I'$. Для получения хорошей стабилизации дифференциальное сопротивление стабилитрона должно быть как можно меньше по сравнению с $R_{огр}$ и R_n . Аналогичные процессы будут происходить при изменении значения сопротивления нагрузки. Это также будет приводить к изменению общего тока в схеме и, следовательно, к изменению падения напряжения на $R_{огр}$, что восстановит первоначальное падение напряжения $U_{ст}$ на R_n .

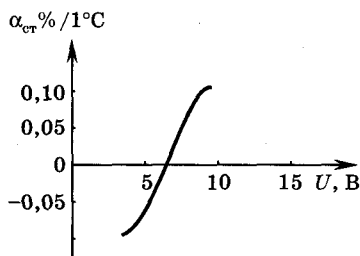


Рис. 3.4

Поскольку напряжение пробоя зависит от температуры, то температурный коэффициент напряжения стабилизации $\alpha_{ст} = (1/U_{ст})(dU_{ст}/dT)$ при $I_{ст} = \text{const}$ является одним из важных параметров стабилитрона. Температурный коэффициент напряжения стабилизации у приборов с лавинным пробоем является положительным, а с туннельным — отрицательным. На рис. 3.4 приведена кривая, из которой можно оценить разброс значений $\alpha_{ст}$ для различных типов стабилитронов.

Для кремниевых высоковольтных стабилитронов с лавинным пробоем значение коэффициента $\alpha_{ст}$ может быть значительно уменьшено при последовательном их соединении с одним или двумя $p-n$ -переходами, смещенными в прямом направлении. Поскольку прямое напряжение уменьшается при повышении температуры, то за счет этого происходит компенсация увеличения напряжения пробоя основного $p-n$ -перехода. Такие стабилитроны с термокомпенсацией называются *прецизионными*. Они обычно применяются в качестве источников опорного напряжения. Помимо стабилитронов общего назначения и прецизионных, в радиоэлектронных схемах нашли применение *импульсные* и *двуханодные стабилитроны*, а также *стабисторы*.

Импульсные стабилитроны используются как для стабилизации напряжения, так и в качестве ограничителей амплитуды импульсного напряжения или смещения уровня постоянного напряжения на величину $U_{ст}$. Импульсные стабилитроны являются быстродействующими приборами, поскольку их время переключения определяется только перезарядкой барьерной емкости из-за того, что при пробое нет накопления неосновных носителей в базе. Время нарастания тока лавины или туннельного прохождения электронов через потенциальный барьер пренебрежимо мало (< 10 пс).

Двуханодные стабилитроны изготавливаются на основе двух встречно включенных $p-n$ -переходов. Внешние выводы осуществляются от областей p -типа, которые являются анодами. При подаче напряжения любой полярности один из стабилитронов смещен в прямом, а другой в обратном направлении. Последний работает в режиме пробоя, а первый, прямосмещен-

ный, переход осуществляет термокомпенсацию подобно прецизионному стабилитрону. Двуханодные стабилитроны, помимо стабилизации напряжения, применяют в схемах двухстороннего ограничения напряжения, в системах защиты цепей от перенапряжения любой полярности.

Стабилитрон, в котором для поддержания постоянства напряжения используют прямую ветвь ВАХ, называют *стабистором*. Эти приборы имеют значительно меньшее напряжение стабилизации (< 1 В) по сравнению с лавинными и туннельными стабилитронами, поскольку работают при прямом смещении. Для увеличения напряжения стабилизации используют последовательное включение нескольких отдельных стабисторов или изготовление такого соединения внутри корпуса одного прибора, чтобы уменьшить $r_{\text{диф}}$ и сопротивления базы r_b в стабисторах применяют высоколегированный кремний. Напряжение стабилизации в этом случае составляет величину порядка 0,7 В, что соответствует участку относительно больших прямых токов для одиночного $p-n$ -перехода. Параметр $\alpha_{\text{ст}}$ отрицателен и по порядку величины лежит в пределах 0,1...0,4% /°С. Стабисторы иногда используют вместе с другими типами стабилитронов в качестве термокомпенсаторов.

3.5. Варикапы

Варикапы — это полупроводниковые диоды, использующие при своей работе зависимость барьерной емкости $p-n$ -перехода от обратного напряжения. Эта зависимость называется *вольт-фарадной характеристикой (ВФХ)*.

Варикапы применяются в схемах электронной перестройки частоты колебательного контура, в усилительных параметрических схемах, в делителях и умножителях частоты, в управляемых фазовращателях и других устройствах.

Основными исходными материалами для изготовления варикапов являются кремний и арсенид галлия. Варикапы создаются на основе эпитаксиально-планарных структур, сплавных и диффузионных технологий. Эпитаксиально-планарные варикапы имеют оптимальные параметры.

Важнейшей характеристикой варикапа является ВФХ, т. е. зависимость $C_{\text{вар}} = f(U)$, или $C_{\text{бар}} = f(U_{\text{обп}})$ (см. гл. 2, рис. 2.8). В качестве важного параметра выступает коэффициент K_C , ха-

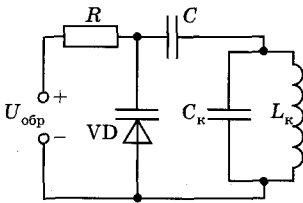


Рис. 3.5

рактически относительное изменение емкости и определяющий крутизну ВФХ:

$$K_C = (dC_{\text{бар}}/dU_{\text{обр}})/C_{\text{бар}}.$$

Величина K_C определяет относительное изменение резонансной частоты ω_0 колебательного контура (C_K — емкость, L_K — индуктивность, рис. 3.5) при совместной работе с варикапом:

$$(d\omega_0/dU_{\text{обр}})/\omega_0 = K_C/[2(1 + C_K/C_{\text{бар}})].$$

На схеме, представленной на рис. 3.5, $C \gg C_{\text{бар}}$ и управляющее напряжение U подается на варикап VD, через высокоомный резистор R , уменьшающий шунтирование варикапа и колебательного контура малым сопротивлением источника питания. При варьировании обратного напряжения происходит изменение емкости варикапа, что приводит к смещению резонансной частоты колебательного контура.

К основным специфическим электрическим параметрам варикапа относятся: *емкость при номинальном, максимальном и минимальном напряжениях*, которая измеряется на заданной частоте; *коэффициент перекрытия по емкости*; *добротность*; *частотный диапазон*; *температурные коэффициенты емкости* $TKC_B = dC_B/(C_B dT)$ и *добротности* $TKQ_B = dQ_B/(Q_B dT)$.

Электрическую модель варикапа можно получить, используя общую модель диода (см. рис. 2.10), в которой необходимо исключить элементы VD и $C_{\text{диф}}$ и добавить индуктивности вводов L_B , емкость корпуса $C_{\text{корп}}$, сопротивление $r_{\text{диф}}$, заменив r_B на $r_{\text{пот}} = r_B + r_K$, где r_K — сопротивление контактов. Для нормальной работы емкость выводов варикапа должна быть заметно меньше барьерной емкости.

Эквивалентная схема варикапа показана на рис. 3.6, а. При работе на низких частотах в эквивалентной схеме можно исключить L_B , $C_{\text{корп}}$, пренебречь сопротивлением базы r_B , в основном определяющим величину $r_{\text{пот}}$, которая мала по сравнению с емкостным сопротивлением $X_C = 1/(\omega C_{\text{бар}})$; эквивалентная схема (рис. 3.6, б) вырождается в параллельное соединение резистора R_{Π} и емкости $C_{\text{бар}}$ (здесь R_{Π} заменяет параллельное соедине-

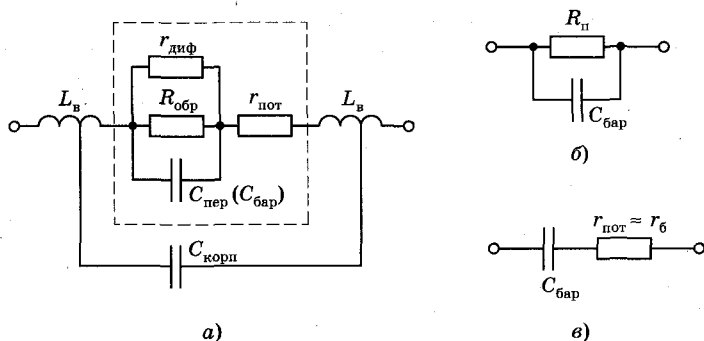


Рис. 3.6

ние $r_{\text{диф}}$ и $R_{\text{обр}}$, показанные на рис. 3.6, а). Добротность Q варикапа для низких частот вычисляется по формуле $Q_{\text{н.ч}} = \omega R_{\text{п}} C_{\text{бар}}$. С увеличением частоты ($\omega = 2\pi f$) изменяется соотношение между реактивной и активной проводимостями и $Q_{\text{н.ч}}$ будет увеличиваться (рис. 3.7, левая ветвь кривой). Для получения больших значений $r_{\text{пер}} = r_{\text{диф}}$, а следовательно, и высокой добротности $Q_{\text{н.ч}}$ при $R_{\text{у}} \gg r_{\text{диф}}$, целесообразно использовать полупроводники с большой шириной запрещенной зоны (Si, GaAs).

На высокой частоте емкостное сопротивление $X_C = 1/(\omega C_{\text{бар}})$ становится малым и в эквивалентной схеме варикапа можно не учитывать большое параллельно включенное активное сопротивление перехода. Однако сопротивление базы $r_{\text{б}} \approx r_{\text{пот}}$ может оказаться сравнимым с X_C , поэтому его исключать нельзя. В результате эквивалентная схема будет выглядеть согласно показанной на рис. 3.6, в. Для такой последовательной схемы замещения добротность вычисляется в соответствии с выражением $Q_{\text{в.ч}} = 1/(\omega r_{\text{б}} C_{\text{бар}})$.

Как видно из этой формулы, $Q_{\text{в.ч}}$ уменьшается с ростом частоты (см. рис. 3.7). Физически это означает, что уменьшается отношение реактивного сопротивления варикапа к сопротивлению потерь $r_{\text{пот}} \approx r_{\text{б}}$.

Основное применение варикапы находят на ВЧ и СВЧ, поэтому для определения их доброт-

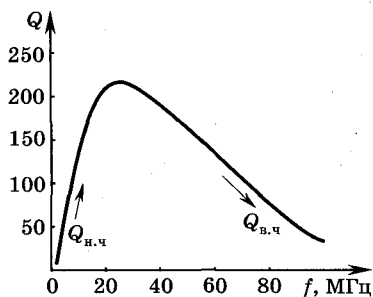


Рис. 3.7

ности используется последовательная эквивалентная схема замещения, показанная на рис. 3.6, *в*, и, кроме того, варикапы изготавливаются с малым значением r_6 .

3.6. Туннельные диоды

Работа туннельных диодов основана на явлении квантово-механического туннелирования основных носителей через потенциальный барьер, энергия которого превышает энергию носителей.

Время туннелирования (τ_t) определяется вероятностью квантово-механического перехода в единицу времени, которое пропорционально $\exp[-2k(0)W]$, где W — ширина потенциального барьера, $k(0)$ — среднее значение волнового вектора носителя заряда в процессе туннелирования, приходящееся на один носитель с нулевым поперечным импульсом и энергией, равной энергии Ферми.

Время туннелирования τ_t чрезвычайно мало, и поэтому туннельные приборы используются в СВЧ-диапазоне в качестве генераторов, переключателей, для туннельной спектроскопии и т. д.

Поскольку вероятность туннельного перехода сильно зависит от ширины потенциального барьера W , то в туннельных диодах используются p — n -переходы, образуемые вырожденными полупроводниками, т. е. полупроводниками с концентрацией примеси порядка $10^{20} \dots 10^{21} \text{ см}^{-3}$. Из-за сильного легирования уровень Ферми располагается внутри зоны проводимости (для n -области) и внутри валентной зоны (для p -области), как это показано на рис. 3.8, *а*, диаграмма 2. Ширина обедненного слоя (ширина p — n -перехода) при указанной степени легирования составляет величину ~ 100 ангстрем (Å) и менее.

Рассмотрение работы туннельных диодов для простоты проведем при температуре абсолютного нуля. В этом случае выше уровня Ферми все разрешенные энергетические состояния по обеим сторонам перехода являются свободными, а ниже уровня Ферми — заполненными электронами (рис. 3.8, *а*). В отсутствие приложенного напряжения ($U = 0$ на рис. 3.8, *а*) туннельный ток через переход не протекает, поскольку туннелирование частиц в данном случае не имеет места, так как туннельный переход происходит без изменения энергии частицы, а при $U = 0$ уровни одинаковой энергии в обеих областях или свободны (расположены выше E_{Φ_n} , E_{Φ_p}), или заняты (расположены ни-

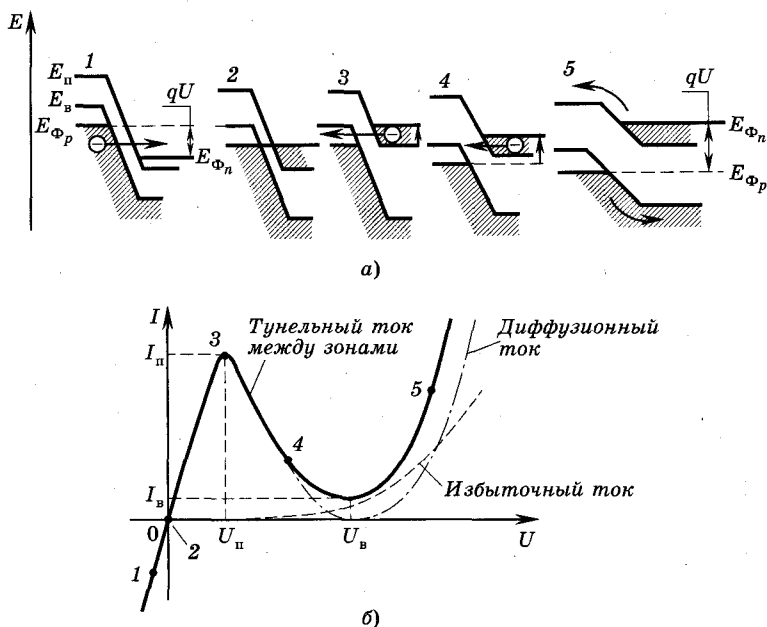


Рис. 3.8

же E_{Φ_n} и E_{Φ_p}). На ВАХ (рис. 3.8, б) рассматриваемому случаю соответствует точка 2, а на рис. 3.8, а — диаграмма 2.

При подаче напряжения происходит туннелирование электронов с занятых состояний валентной зоны p -полупроводника на свободные состояния n -полупроводника (обратное смещение, точка 1 на ВАХ рис. 3.8, б, диаграмма 1 на рис. 3.8, а) или, наоборот, с занятых состояний n -полупроводника на свободные уровни p -полупроводника (прямое смещение, точка 3 на рис. 3.8, б, диаграмма 3 на рис. 3.8, а). Для реализации указанных туннельных переходов необходимо выполнение следующих условий:

- 1) наличие заполненных состояний в области, откуда электроны туннелируют;
- 2) наличие свободных состояний с теми же значениями энергии в области, куда электроны туннелируют;
- 3) ширина потенциального барьера должна быть как можно меньше (сравнима с длиной волны де Бройля) для того, чтобы вероятность туннелирования была как можно больше;
- 4) квазиимпульс туннелирующих электронов должен сохраняться при переходе.

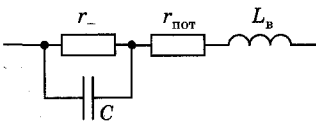


Рис. 3.9

(между точками 3 и 4 на рис. 3.8, а) представлена на рис. 3.9. Резистор r_- — дифференциальное отрицательное сопротивление $p-n$ -перехода, $r_{\text{пот}}$ — сопротивление потерь, определяемое сопротивлениями областей n и p диода

контактов и выводов, $L_в$ — индуктивность выводов, $C = C_{\text{бар}}$ — емкость диода при фиксированном напряжении на переходе, $r_{\text{пот}} \approx r_в$ (сопротивление выводов и контактов).

Импеданс приведенной на рис. 3.9 схемы равен

$$z = [r_в + r_- / (1 + \omega r_- C)^2] + i[\omega L_в - \omega C r_-^2 / (1 + (\omega r_- C)^2)]. \quad (3.2)$$

Из формулы (3.2) видно, что активная и реактивная составляющие импеданса при некоторых значениях частот $\omega = 2\pi f$ равны нулю, при этом эти частоты различны для этих составляющих полного импеданса. Частота, при которой активная составляющая равна нулю, называется *предельной резистивной частотой* f_R

$$f_R = \frac{(|r_-|/r_{\text{пот}} - 1)^{1/2}}{2\pi r_- C}. \quad (3.3)$$

Резонансная частота f_0 туннельного диода соответствует равенству нулю реактивной составляющей полного импеданса

$$f_0 = (1/2\pi)[1/(L_в C) - 1/(r_- C)^2]^{1/2}. \quad (3.4)$$

При разработке туннельных диодов закладывается условие $f_0 > f_R$. В результате возможные паразитные резонансы могут возникать только на частотах, на которых диод не обладает отрицательным дифференциальным сопротивлением. Из указанного неравенства следует ограничение на индуктивность $L_в < |r_-|/r_{\text{пот}} C$, что достигается использованием коаксиальной или волноводной табличной конструкции корпуса.

Обращенный диод. Если концентрации легирующих примесей таковы, что p - и n -области близки к вырождению либо слабо вырождены, то при малых прямых и обратных смещениях ток в прямом направлении меньше тока в обратном. Отсюда возникает название такого туннельного диода — **обращенный диод**. В равновесии уровень Ферми в обращенном диоде близок к границам зон разрешенной энергии, т. е. к потолку валентной зоны p -области и дну зоны проводимости n -области.

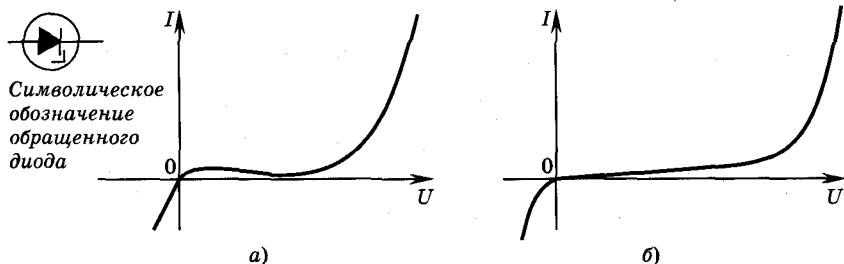


Рис. 3.10

При обратном смещении электроны легко туннелируют из валентной зоны p -области в зону проводимости n -области, что будет приводить к возрастанию туннельного тока с ростом абсолютного значения обратного напряжения. В результате обратная ветвь ВАХ обращенного диода аналогична ВАХ туннельного диода (рис. 3.10, а, б). При подаче прямого напряжения ток в обращенном диоде определяется инжекцией носителей заряда через переход подобно обычному диоду. При малых прямых напряжениях $U < 0,5$ В прямой ток в обращенном диоде значительно меньше обратного. За счет избыточного туннельного тока в обращенных диодах возможны слабые проявления туннельного эффекта при прямых смещениях (см. рис. 3.10, а). Обращенные диоды можно использовать в качестве детекторов малых сигналов СВЧ-излучения, смесителей, переключателей.

Обращенные диоды имеют хорошие частотные характеристики, поскольку их работа не сопровождается накоплением неосновных носителей, и малый $1/f$ шум (см. гл. 22). ВАХ этого типа диодов не чувствительна к влиянию температуры и облучения.

Помимо рассмотренных приборов, туннельный эффект используется также в МДП-диодах (структура металл — диэлектрик — полупроводник) при толщине диэлектрика от 10 до 50 Å, в МДМ-диодах (структура металл — диэлектрик — металл) и туннельных транзисторах на основе МДП и МДМ структур.

3.7. Лавинно-пролетные диоды

Принцип работы лавинно-пролетных диодов (ЛПД) основан на возникновении отрицательного сопротивления в результате использования процессов лавинного умножения носителей и времени их пролета через полупроводниковую структуру при наличии

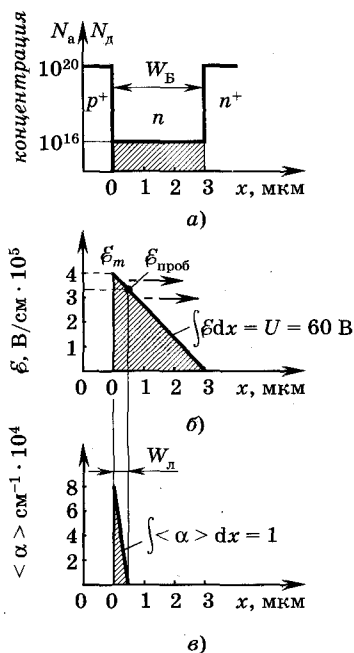


Рис. 3.11

на электродах помимо постоянного также и переменного напряжения.

Появление отрицательного сопротивления связано с фазовым сдвигом между током и напряжением. Необходимо подчеркнуть, что статическая ВАХ ЛПД не отличается от характеристик обычных диодов и, следовательно, на ней отсутствует участок с отрицательным дифференциальным сопротивлением. Отрицательное сопротивление реализуется только в динамическом режиме, когда на электродах одновременно присутствуют постоянное обратное смещение и переменное напряжение. ЛПД работают при подаче обратного напряжения, близкого к напряжению пробоя.

Для изготовления ЛПД используются различные структуры, такие, как четырехслойная структура Рида ($p^+ - n - i - n^+$), асимметричный резкий $p - n$ -переход ($p^+ - n - n^+$), симметричный $p - n$ -переход — диод с двумя дрейфовыми областями ($p^+ - p - n - n^+$), диод с двухслойной базой, диод с трехслойной базой (модифицированный диод Рида) и $p - i - n$ -структуры (диоды). Каждая из приведенных структур характеризуется определенным напряжением пробоя, размером области лавинного умножения и области дрейфа, степенью влияния объемного заряда носителей и температуры, а также динамическими характеристиками. Рассмотрение физических процессов в ЛПД проведем на основе анализа простейшей структуры — асимметричного резкого $p^+ - n$ -перехода (рис. 3.11). Отметим, что в ЛПД распределение концентраций примесей в переходах должно быть как можно ближе к ступенчатому, подобно тому, как это представлено на рис. 3.11, а. На рис. 3.11, б, в показаны распределение напряженности электрического поля ϵ при подаче обратного напряжения на структуру («минус» источника питания подключен к области p^+) и эффективного коэффициента ударной ионизации $\langle \alpha \rangle$. ЛПД работают

при подаче постоянного смещения, близкого к пробивному, при этом обедненный слой $p^+ - n$ -перехода распространяется через всю n -базу, которая является областью дрейфа носителей. Максимум напряженности электрического поля \mathcal{E}_m достигается на металлургической границе $p^+ - n$ -перехода, вблизи которой существует узкий слой $W_d = x - x_0(0)$ (см. рис. 3.11, б, в), где $\mathcal{E} > \mathcal{E}_{\text{проб}}$ и, следовательно, происходит лавинное размножение носителей с коэффициентом $\langle \alpha \rangle$. Дырки сразу же под действием поля попадают в p^+ -область, а электроны дрейфуют по направлению к n^+ -области со скоростью насыщения $v_{\text{нас}}$, которая для кремния равна 10^7 см/с при $\mathcal{E} > 10^4$ В/см, а время дрейфа $t_{\text{др}} = W_d/v_{\text{нас}}$.

Использование ЛПД или другого электронного прибора в качестве активного элемента в схеме автогенератора СВЧ-колебаний основано на том, что при определенных условиях этот прибор может представлять собой отрицательное динамическое сопротивление (ОДС). Понятие отрицательного динамического сопротивления характеризует сопротивление диода на переменном токе, когда между переменными составляющими тока i_d , текущего через прибор, и напряжения U_d , приложенного к его электродам, существует фазовый сдвиг θ такой величины, что в течение большей части периода процесс нарастания напряже-

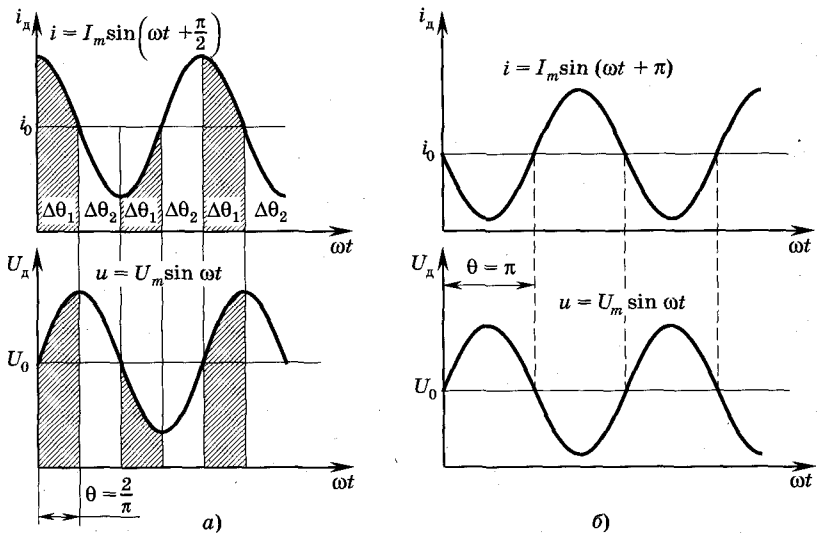


Рис. 3.12

ния сопровождается снижением тока, и наоборот. Из рис. 3.12, а легко видеть, что при $\theta = \pi/2$ промежутки времени $\Delta t_1 = \Delta\theta_1/\omega$, в течение которых наблюдается отрицательное динамическое сопротивление (на рисунке эти участки заштрихованы), чередуются с равными по величине интервалами $\Delta t_2 = \Delta\theta_2/\omega$, когда динамическое сопротивление положительно. Следовательно, в среднем за период динамическое сопротивление не является отрицательным.

Отрезки времени Δt_1 превысят по величине отрезки Δt_2 , если сдвиг фаз θ будет больше $\pi/2$, а в случае $\theta = \pi$ (см. рис. 3.12, б) диод будет представлять собой отрицательное динамическое сопротивление в течение всего периода. Таким образом, условие, при котором диод в течение большей части периода характеризуется отрицательным динамическим сопротивлением, можно записать в виде

$$3\pi/2 > \theta > \pi/2. \quad (3.5)$$

В ЛПД требуемый фазовый сдвиг между переменными составляющими тока и напряжения, определяемый формулой (3.5), образуется за счет конечного времени протекания основных физических процессов: лавинного размножения свободных носителей заряда в запирающем слое электрического перехода и их дрейфового движения в базе прибора.

Рассмотрим процессы более детально. Предположим, что ЛПД в виде структуры, представленной на рис. 3.11, а, включен в схему, содержащую источник внешнего постоянного напряжения U_0 , а также колебательный контур с параметрами L_k , C_k , R_k (рис. 3.13). Конденсатор C шунтирует источник постоянного напряжения по переменному току, $R_{\text{орп}}$ задает положение рабочей точки. Предположим, что в колебательном контуре возникли колебания с частотой ω , равной резонансной частоте контура, и амплитуда этих колебаний U_m стационарна. Работа такой схемы в качестве автогенератора в стационарном режиме возможна, если ЛПД представляет собой элемент с отрицательным динамическим сопротивлением. Тогда энергия, поступающая в контур за счет переменной составляющей тока диода, расходуется на восполнение потерь в

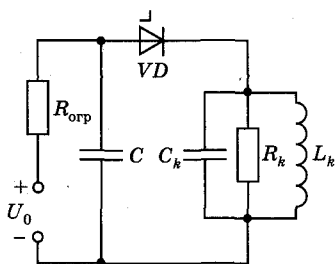


Рис. 3.13

контуре, и амплитуда этих колебаний U_m стационарна. Работа такой схемы в качестве автогенератора в стационарном режиме возможна, если ЛПД представляет собой элемент с отрицательным динамическим сопротивлением. Тогда энергия, поступающая в контур за счет переменной составляющей тока диода, расходуется на восполнение потерь в

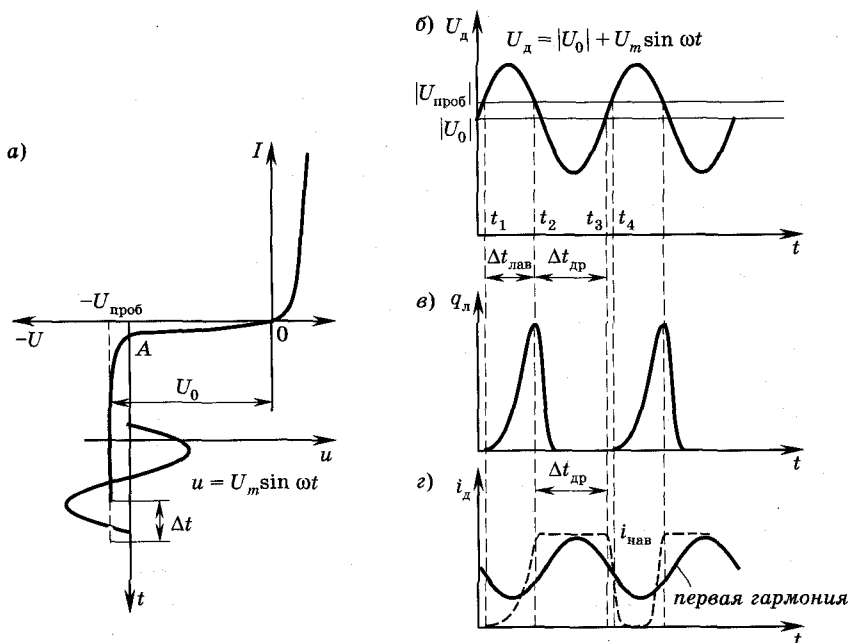


Рис. 3.14

самом контуре и в других цепях схемы и способствует таким образом поддержанию стационарной амплитуды колебаний.

Рассмотрим физические процессы в ЛПД и докажем, что в этой схеме обеспечиваются необходимые условия для поддержания колебаний стационарной амплитуды.

На рис. 3.14, а показана ВАХ диода. Напряжение U_0 обратного смещения, определяющее рабочую точку A на обратной ветви характеристики диода, по абсолютной величине меньше пробивного напряжения $|U_0| < |U_{\text{проб}}|$. На этом же рисунке показано переменное напряжение $u = U_m \sin \omega t$, возникающее на колебательном контуре, и, следовательно, на электродах диода. Суммарное напряжение на диоде $U_d = U_0 + U_m \sin \omega t$ таково, что в интервалы Δt отрицательного полупериода оно по абсолютной величине превышает напряжение $U_{\text{проб}}$.

Электрическое поле, создаваемое в диоде за счет внешних постоянного и переменного напряжений, суммируется с контактным полем в электрическом переходе. При условии $|U_d| > |U_{\text{проб}}|$ (момент t_1 на рис. 3.14, б) в диоде возникает лавинный пробой,

причем образование лавины ограничено узкой областью запирающего слоя вблизи границы p^+ - и n -областей — слоем умножения $W_{\text{л}}$ (см. рис. 3.11, *в*), где поле имеет наибольшую величину и где коэффициент умножения носителей больше единицы. Таким образом, в узком слое умножения $W_{\text{л}}$ лавинообразно нарастают концентрации электронов и дырок, и объемный заряд $q_0 = q_n + q_p$, возникающий за счет их встречного дрейфа.

Образование лавины — это ряд последовательных процессов ионизации атомов полупроводника, причем скорость генерации электронно-дырочных пар в области размножения носителей зависит от напряженности электрического поля и концентрации частиц. Поэтому число размноженных носителей и, соответственно, объемный заряд q_0 достигают максимальной величины лишь по истечению некоторого времени $\Delta t_{\text{лав}}$ после момента t_1 начала возникновения лавины (рис. 3.14, *в*). После того как напряжение на диоде снизится до значения $|U_{\text{д}}| < |U_{\text{проб}}|$ (момент t_2 на рис. 3.14, *б*), процесс размножения носителей заряда прекращается. На рис. 3.14, *в* показан процесс изменения объемного заряда электронов q_0 , причем интервал времени $\Delta t_{\text{др}}$ несколько больше $T/4$ (T — период колебаний напряжения $U_{\text{д}}$). По мере появления носителей заряда в узком слое умножения они под влиянием электрического поля уходят из этого слоя, при этом дырки и электроны дрейфуют в противоположных направлениях. В следствие того, что $W_{\text{л}} \ll W_{\text{Б}}$ (см. рис. 3.11, *а, в*), дырки в течение короткого отрезка времени уходят из запирающего слоя в p^+ -область, а сгусток вновь образовавшихся электронов дрейфует в базе в течение более длительного времени $\Delta t_{\text{др}}$ (см. рис. 3.14, *в*). Если величина поля в базе диода при любом значении $U_{\text{д}}$ превышает напряженность поля $\mathcal{E}_{\text{нас}}$, при которой наблюдается насыщение дрейфовой скорости, то электроны движутся в базе с неизменной скоростью $v_{\text{др}} = v_{\text{др. нас}}$. Достигая $n-n^+$ -перехода, электроны экстрагируются полем этого перехода (момент t_3 на рис. 3.14, *б, в*). За время движения в базе объемный заряд электронов наводит во внешней цепи ток $i_{\text{нав}}$ (рис. 3.14, *г*), близкий по форме к прямоугольному импульсу. В момент времени t_4 напряжение на диоде снова превышает пробивное напряжение и описанные процессы повторяются.

Таким образом, во внешней цепи диода наблюдаются импульсы тока, следующие друг за другом с периодом повторения T , рав-

ным периоду переменного напряжения U_d на диоде. Значение интервалов времени $\Delta t_{\text{лав}}$ и $\Delta t_{\text{др}}$ могут быть выбраны такими, что первая гармоника импульсной последовательности (сплошная линия на рис. 3.14, з) окажется в противофазе ($\theta = \pi$) с напряжением на диоде. В этом случае диод будет представлять собой отрицательное динамическое сопротивление в течение всего периода.

Если же фазовый сдвиг $\theta \neq \pi$, то условие поддержания колебаний в контуре выполняется лишь в те отрезки времени, когда диод характеризуется отрицательным дифференциальным сопротивлением. В этом случае в контур передается меньшая энергия и амплитуда колебаний уменьшается. Величина фазового сдвига зависит, в частности, от частоты колебаний ω :

$$\theta = \theta_{\text{лав}} + \theta_{\text{др}} = \omega(\Delta t_{\text{лав}} + \Delta t_{\text{др}}). \quad (3.6)$$

Отсюда следует, что при неизменных условиях образования лавины и дрейфа носителей в диоде величина угла θ может измениться при перестройке контура. Поэтому для данного диода можно определить оптимальную частоту $\omega_{\text{онт}}$ генерируемых колебаний, при которых фазовый угол θ наиболее близок к величине 180° и мощность колебаний максимальна.

Отклонение величины фазового угла θ от оптимальной может произойти и по другим причинам. Так, например, возникновение плотного объемного заряда электронов q_n в базе диода снижает потенциал в области существования этого заряда, и распределение электрического поля \mathcal{E}_k в диоде может измениться так, как это показано на рис. 3.15 сплошной линией (штриховой линией показана напряженность поля в отсутствие объемного заряда).

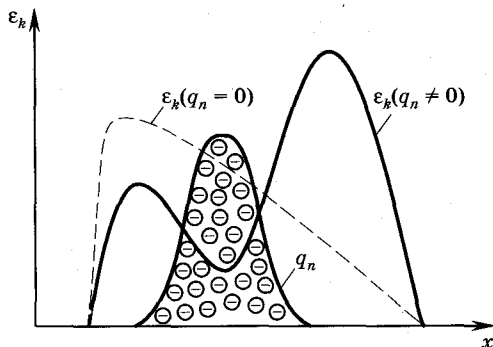


Рис. 3.15

Снижение величины поля \mathcal{E}_k в слое умножения приведет к тому, что процесс нарастания лавины прекратится ранее момента t_2 (см. рис. 3.14, б); следовательно, уменьшится величина $\theta_{\text{лав}}$, а значит, и величина фазового сдвига $\Delta\theta$.

Показанное на рис. 3.15 искажение распределения напряженности электрического поля в диоде может привести и к более существенным изменениям всех физических процессов в приборе.

Если переменное напряжение на диоде достигает значения, примерно равного удвоенному пробивному напряжению, то в лавинной области создается столь плотный заряд электронов, что напряженность поля со стороны p^+ -области понижается практически до нуля, а в области базы повышается до уровня, достаточного для возникновения процесса ударной ионизации. В результате этого процесса слой лавинного умножения смещается и формируется в области базы на фронте сгустка электронов. Генерируемые электроны и дырки практически мгновенно изменяют распределение электрического поля. Это приводит к тому, что в пространстве дрейфа образуется движущийся в направлении к n^+ -области лавинно-ударный фронт, который оставляет за собой большое количество электронов и дырок, концентрация которых столь велика, что напряженность поля здесь может понизиться практически до нуля (рис. 3.16). В результате ионизации атомов полупроводника в области базы образуется равное количество электронов и дырок (положительных и отрицательных зарядов). Такое состояние принято называть компенсированной полупроводниковой плазмой, а режим работы ЛПД — режимом захваченной плазмы.

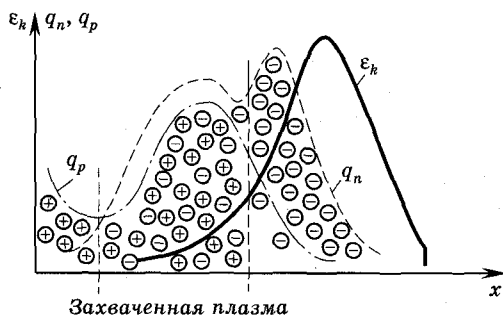


Рис. 3.16

В иностранной литературе этот режим носит наименование TRAPATT-режима. Название образовано из начальных букв английской фразы TRApped Plasma Avalanche Triggered Transit (захваченная плазма, пробег области лавинного умножения).

В этом режиме можно выделить три фазы. Первая фаза — образование лавинного ударного фронта. Лавинный фронт перемещается в диоде со скоростью v_{ϕ} , значительно превышающей дрейфовую скорость насыщения: $v_{\phi} > v_{\text{др. нас}}$. Таким образом, лавинный фронт быстро проходит через диод, оставляя его заполненным плазмой, захваченной слабым электрическим полем. Ток, текущий через прибор в этой фазе, существенно увеличивается вследствие дополнительного размножения носителей в базе, а напряжение на диоде за счет образования плазмы снижается.

Вторая фаза — период восстановления. База диода в этой фазе заполнена электронно-дырочной плазмой, а напряженность поля значительно меньше величины, соответствующей насыщению дрейфовой скорости. Поэтому дырки из области базы дрейфуют к p^+ -области, а электроны — к n^+ -области со скоростью $v_{\text{пл}}$, существенно меньшей, чем дрейфовая скорость насыщения. Происходит постепенное рассасывание плазмы. Ток, текущий через прибор в этой фазе, остается неизменным; его величина определяется подвижностью носителей μ_n и μ_p , их концентрацией p_i и n_i и напряженностью поля $\mathcal{E}_{\text{пл}}$ в области плазмы

$$j = q\mathcal{E}_{\text{пл}}(\mu_n n_i + \mu_p p_i). \quad (3.7)$$

С уходом носителей из базы диода поле в базе у переходов увеличивается со временем и постепенно наступает третья фаза, характеризующаяся высоким значением напряженности поля в диоде и предшествующая последующему образованию лавинного ударного фронта.

Из сказанного ясно, что все описанные процессы протекают за время, превышающее время нарастания лавины и дрейфа носителей в базе при пролетном режиме. Иначе говоря, период повторения процессов в режиме с захваченной плазмой существенно больше периода повторения импульсов тока в пролетном режиме. Поэтому при работе в режиме с захваченной плазмой контур настраивается на более низкую частоту и соответственно частота колебаний в этом режиме значительно меньше частоты генерации в пролетном режиме.

Помимо этого, режим с захваченной плазмой отличается более высоким значением КПД (более 50%) по сравнению с пролетным режимом (КПД ~ 10%). Это объясняется главным образом повышенной плотностью тока при малом напряжении, что характерно для большей части периода повторения процессов.

3.8. Диоды Ганна

Диод Ганна — полупроводниковый прибор, принцип работы которого основан на использовании объемных свойств полупроводника. Этот диод не содержит ни p — n -переходов, ни каких-либо других границ раздела, кроме омических контактов. Основное назначение диода Ганна — работа в усилителях и генераторах электромагнитных колебаний СВЧ.

Диод Ганна обычно выполняется из арсенида галлия n -типа в виде пластинки или шайбы, в которую с двух противоположных сторон вплавляются омические контакты. Диод помещается в герметичный керамический корпус, фланцы которого выполнены из металла и служат выводами прибора. Конструкция диода Ганна рассчитана на включение в коаксиальный или волноводный тракты.

Рассмотрим физические процессы, происходящие в диоде Ганна. В основе работы диода лежат физические явления, связанные с возникновением управляемого электрическим полем отрицательного дифференциального сопротивления в полупроводниках (GaAs, CdTe, InP, InSb и др.), зона проводимости которых имеет два минимума энергии.

Диаграмма энергетических уровней арсенида галлия показана на рис. 3.17, где для кристаллографической плоскости

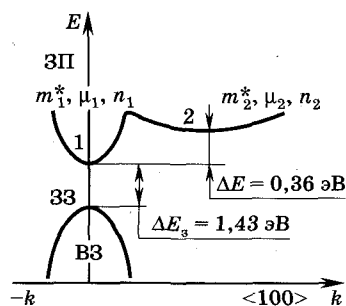


Рис. 3.17

$\langle 100 \rangle$ кристалла GaAs приведена зависимость энергии электронов E в валентной зоне (ВЗ) и зоне проводимости (ЗП) от волнового числа $k = p/h$ (p — импульс частицы, а $h = 6,6 \times 10^{-34}$ Дж·с — постоянная Планка). Как видно, зависимость $E = f(k)$ имеет в зоне проводимости два минимума, которые обычно называют долинами. Обозначим низкую (узкую) долину цифрой 1, а верхнюю (широкую) — цифрой 2. Энергетический

зазор ΔE между долинами равен 0,36 эВ. Как известно, характер зависимости $E = f(k)$ определяет величину эффективной массы частицы

$$1/m^* = \partial^2 E / \partial p^2. \quad (3.8)$$

Отсюда следует, что эффективная масса частицы m_1^* в нижней долине меньше эффективной массы частицы m_2^* в верхней долине. Для GaAs эффективная масса $m_1^* = 0,07m_0$ (m_0 — масса покоя электрона), а масса $m_2^* = 1,2m_0$. Вследствие различия эффективных масс существенно отличаются и подвижности электронов в нижней и верхних долинах: в долине 1 электроны имеют высокую подвижность ($\mu_1 = 0,8 \text{ м}^2/(\text{В} \cdot \text{с})$), а в долине 2 — низкую ($\mu_2 = 0,01 \text{ м}^2/(\text{В} \cdot \text{с})$).

Рассмотрим поведение электронов в зоне проводимости кристалла GaAs, помещенного в электрическое поле с напряженностью \mathcal{E} . Как известно, плотность дрейфового тока определяется соотношением

$$j = en\mu\mathcal{E}, \quad (3.9)$$

где e — заряд, n — концентрация и μ — подвижность электронов; \mathcal{E} — напряженность электрического поля.

Если все свободные электроны находятся в нижней долине ($n = n_1$), то

$$j = en_1\mu_1\mathcal{E}, \quad (3.10)$$

и зависимость $j = f(\mathcal{E})$ имеет вид прямой 1, показанной на рис. 3.18.

Если бы все электроны находились в верхней долине ($n = n_2$), то

$$j = en_2\mu_2\mathcal{E}, \quad (3.11)$$

и, вследствие их более низкой подвижности ($\mu_2 < \mu_1$), зависимость $j = f(\mathcal{E})$ была бы иной (прямая 2 на рис. 3.18).

В реальных условиях при слабых электрических полях электроны находятся в термодинамическом равновесии с кристаллической решеткой полупроводника и занима-

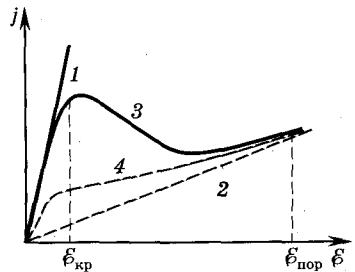


Рис. 3.18

ют в основном энергетические уровни нижней долины. С увеличением напряженности поля электроны, находящиеся в нижней долине, приобретают дополнительную энергию, и когда их энергия возрастет на величину, равную $\Delta E = 0,36$ эВ, становится возможным переход электронов в верхнюю долину. Назовем значение напряженности электрического поля, при котором начинается этот процесс, критической напряженностью электрического поля $\mathcal{E}_{кр}$. Для GaAs $\mathcal{E}_{кр} \approx 3$ кВ/см. Энергия свободных электронов в кристалле распределена в соответствии с законами статистики. При некотором значении напряженности электрического поля, называемом пороговым $\mathcal{E}_{пор}$ ($\mathcal{E}_{пор} > \mathcal{E}_{кр}$), большинство электронов из первой долины переходят во вторую, и дальнейшее возрастание поля \mathcal{E} не приводит к существенному увеличению количества электронов во второй долине (при условии, что электрический пробой в кристалле отсутствует).

Таким образом, можно считать, что для электрических полей $\mathcal{E} < \mathcal{E}_{кр}$ и $\mathcal{E} > \mathcal{E}_{пор}$ плотность тока через полупроводник описывается соответственно формулами (3.10) и (3.11). При напряженностях электрического поля $\mathcal{E}_{кр} < \mathcal{E} < \mathcal{E}_{пор}$ электроны находятся как в нижней, так и в верхних долинах и плотность тока j в кристалле определяется как сумма плотностей токов за счет дрейфа электронов в обеих долинах:

$$j = e(n_1\mu_1 + n_2\mu_2)\mathcal{E} = e\mu_{ср}\mathcal{E}, \quad (3.12)$$

где $n = n_1 + n_2$ — концентрация электронов в зоне проводимости, а средняя подвижность электронов в обеих долинах дается формулой

$$\mu_{ср} = (n_1\mu_1 + n_2\mu_2)/n. \quad (3.13)$$

Характер зависимости $j = f(\mathcal{E})$ в интервале изменения напряженности поля от $\mathcal{E}_{кр}$ до $\mathcal{E}_{пор}$ зависит от ряда условий и соотношения величин, входящих в формулу (3.12). В определенных условиях зависимость $j = f(\mathcal{E})$ может иметь вид, описываемый кривой 3 на рис. 3.18. При этом кристалл арсенида галлия в некотором интервале энергий характеризуется отрицательным дифференциальным сопротивлением, так как на этом участке с увеличением напряженности поля \mathcal{E} плотность тока j уменьшается. В иных условиях зависимость $j = f(\mathcal{E})$ может не иметь области с отрицательным дифференциальным сопротивлением (кривая 4 на рис. 3.18). Условия возникновения области с отрицательным

дифференциальным сопротивлением можно установить, дифференцируя соотношение (3.12) по ε и полагая $\partial j / \partial \varepsilon < 0$.

Анализ полученного таким образом выражения позволяет сформулировать следующие условия получения зависимости $j = f(\varepsilon)$ с областью отрицательного дифференциального сопротивления:

- 1) разность ΔE между минимальными значениями энергии в первой и во второй долине должна быть больше энергии теплового движения носителей, чтобы в отсутствие внешнего электрического поля большинство носителей находилось в нижней долине зоны проводимости;
- 2) величина интервала ΔE должна быть меньше ширины запрещенной зоны ΔE_z , так как при невыполнении этого условия произойдет электрический пробой до того, как носители начнут переходить в верхнюю долину;
- 3) подвижность электронов в верхней долине должна быть много меньше их подвижности в нижней долине.

В полупроводнике, обладающем ВАХ $j = f(\varepsilon)$ с областью отрицательного дифференциального сопротивления, произвольная флуктуация плотности тока в любой точке кристалла приводит к возникновению неустойчивости объемного заряда.

Рассмотрим физические процессы возникновения и развития такой неустойчивости объемного заряда. На рис. 3.19, а показан кристалл арсенида галлия длиной l , к которому подключен источник внешнего напряжения U . Предположим, что концентрация свободных электронов во всем объеме кристалла одинакова и все электроны находятся в нижней долине. Тогда распределение потенциала по длине кристалла линейно (прямая 1 на рис. 3.19, б), а напряженность поля $\varepsilon_1 = \partial U / \partial l$

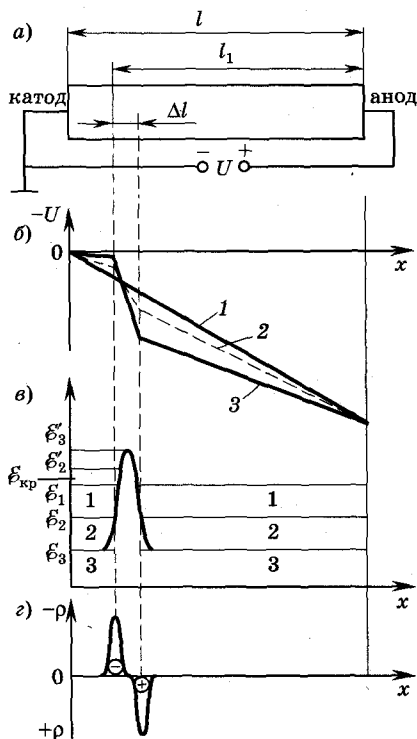


Рис. 3.19

неизменна (прямая 1 на рис. 3.19, в). Если величина напряжения U такова, что напряженность поля в кристалле $\mathcal{E} < \mathcal{E}_{\text{кр}}$, то плотность тока, текущего в полупроводнике, обусловлена дрейфом электронов с подвижностью μ_1 и определяется соотношением (3.10).

Предположим далее, что в кристалле имеется неоднородность в виде области $\Delta l = d$ с пониженной концентрацией свободных электронов и, следовательно, с более высоким удельным сопротивлением. Падение потенциала в этой области должно возрасти, и распределение потенциала вдоль кристалла в этом случае будет характеризоваться ломаной линией (линия 2 на рис. 3.19, б). На участке $\Delta l = d$ напряженность поля увеличится до \mathcal{E}'_2 , а за его пределами уменьшится до значения \mathcal{E}_2 (см. рис. 3.19, в).

Если напряженность поля \mathcal{E}'_2 в области Δl превысит величину $\mathcal{E}_{\text{кр}}$, то в этой области начнется переход электронов из нижней долины в верхнюю. Увеличение числа электронов с меньшей подвижностью ($\mu_1 > \mu_2$) приводит к увеличению электрического сопротивления этого участка, а значит, к дальнейшему изменению распределения потенциала, возрастанию напряженности поля до значения \mathcal{E}'_3 на участке Δl (ломаная линия 3 на рис. 3.19, б). Иначе говоря, на участке Δl будет формироваться область с более высокой напряженностью поля или, как говорят, домен сильного поля. С возникновением домена высокого поля напряженность поля за его пределами должна снизиться до значения \mathcal{E}_3 , так как с появлением участка с более высоким сопротивлением возрастет общее сопротивление кристалла и уменьшится плотность тока. Следовательно, в области, где возникает домен, электроны находятся преимущественно в верхней долине и обладают низкой подвижностью μ_2 , а в остальном объеме кристалла электроны находятся в нижней долине и характеризуются более высокой подвижностью $\mu_1 > \mu_2$. И в той, и в другой области под влиянием приложенной разности потенциалов возникает дрейфовое движение электронов. Электроны прикатодной области, приближаясь к домену, уменьшают скорость дрейфа и образуют в части домена, обращенной к катоду, более плотный отрицательный объемный заряд. В то же время электроны между доменом и анодом характеризуются более высокой дрейфовой скоростью. Поэтому в части домена, обращенной к аноду, образуется область, обедненная электронами. Таким образом, с развитием домена высокого поля в этой области формируется двойной электрический слой (рис. 3.19, г).

Процесс формирования домена будет развиваться до тех пор, пока не установится некоторое динамическое равновесие, определяемое законом сохранения заряда, согласно которому плотность тока в неразветвленной электрической цепи должна быть одинаковой в любом сечении. Условие постоянства плотности тока может выполняться при различ-

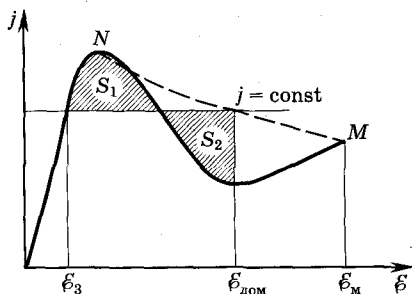


Рис. 3.20

ных значениях напряженности поля в домене $\epsilon_{\text{дом}} > \epsilon_{\text{кр}}$ и в областях кристалла вне домена ϵ_3 (см. обозначения на рис. 3.19, в) и определяется так называемым правилом равных площадей $S_1 = S_2$ (рис. 3.20). Положение прямой $j = \text{const}$ на характеристике $j = f(\epsilon)$ определяется величиной приложенного напряжения U , подключенного к кристаллу. Штриховая кривая на рис. 3.20 начинается в точке максимума N зависимости $j = f(\epsilon)$ и заканчивается при $\epsilon = \epsilon_m$. Эта кривая NM ограничивает область значений j и ϵ , где выполняется правило равных площадей.

Таким образом, в условиях динамического равновесия заканчивается формирование домена высокого поля; плотность тока в кристалле неизменна ($j = \text{const}$) и дрейфовая скорость $v_{\text{др. дом}}$ электронов в области домена и за его пределами $v_{\text{др. 1}}$ равны друг другу:

$$v_{\text{др. дом}} = v_{\text{др. 1}}$$

Итак, условие $j = \text{const}$ выполняется, когда усредненная по всем значениям ϵ дрейфовая скорость электронов в домене равна дрейфовой скорости вне домена и, следовательно, существуют условия формирования домена. Домен высокого поля со скоростью $v_{\text{др. дом}}$ перемещается по направлению к аноду. На аноде область повышенной плотности объемного заряда расформируется и домен высокого поля исчезает. По мере исчезновения домена электрическое сопротивление кристалла уменьшается, плотность тока возрастает и на неоднородности $\Delta l = d$ вновь начинается процесс формирования домена.

В рассматриваемом случае новый домен может возникнуть лишь после расформирования существовавшего домена, так как до этого момента напряженность поля ϵ_3 вне домена ниже пер-

воначальной величины ϵ_1 и тем более ниже критического значения ($\epsilon_3 < \epsilon_{кр}$). Иначе говоря, в кристалле может возникнуть и существовать лишь единственный домен. Новый домен зарождается лишь по прошествии некоторого интервала времени $t_{пр}$, необходимого для перемещения предыдущего домена от места его формирования до анода:

$$t_{пр} = l_1 / v_{др. дом}, \quad (3.20)$$

где l_1 — расстояние от места возникновения домена до анода (см. рис. 3.19, а).

Следовательно, домены высокого поля возникают и исчезают в кристалле периодически с интервалом времени $t_{пр}$ и, что особенно важно, также периодически и с тем же интервалом $t_{пр}$ изменя-

ется плотность тока, текущего через кристалл. На рис. 3.21, е показана форма тока, текущего через кристалл, а на рис. 3.21, б, в, г, д — эюры распределения поля в кристалле в различные моменты времени, отмеченные на рис. 3.21, е соответствующими буквами.

Если в цепь диодов включить колебательный контур, настроенный на частоту

$$f_{пр} = 1/t_{пр} = v_{др. дом} / l_1, \quad (3.21)$$

то такое устройство может служить генератором незатухающих колебаний.

Условия формирования домена высокого поля в кристалле не ограничиваются наличием ВАХ с областью отрицательного дифференциального сопротивления. Важную роль играют и другие факторы: геометрические размеры кристалла, концентрация носителей заряда и др.

Для стабильности периодического процесса зарождения домена высокого поля в определенной области кристалла необходимо обеспечить однородность кристалла полупроводника

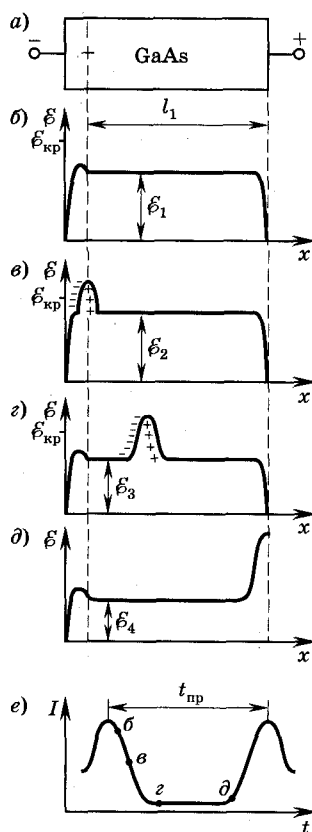


Рис. 3.21

во всем объеме, за исключением единственной, локальной области неоднородности. В качестве такой неоднородности могут служить различные дефекты кристаллической структуры, в том числе и незначительное отклонение концентрации примеси в области неоднородности. Экспериментальные исследования показывают, что в арсениде галлия домен должен стабильно зарождаться на неоднородности, образованной изменением концентрации примесей в 0,01% на участке длиной около 1 мкм. Обычно такого рода неоднородность образуется в прикатодной области кристалла при его изготовлении.

Условие существования одной неоднородности удается выполнить лишь для небольших кристаллов ($l \leq 0,1$ мм). В более длинных кристаллах образуется несколько неоднородностей, на которых домены зарождаются случайным образом. Колебание тока в таком кристалле носит не периодический, а случайный характер. С другой стороны, технологические трудности ограничивают возможности создания кристаллов длиной $l < 1$ мкм. В арсениде галлия при температуре 300 К и напряженности поля $\mathcal{E} \approx 10$ кВ/см дрейфовая скорость домена $v_{\text{др. дом}} \approx 10^7$ см/с. Полагая $l \approx l_1$, на основании (3.21) легко оценить возможные частотные пределы использования диодов Ганна как активных элементов в генераторах незатухающих колебаний. При $l_1 \approx 0,1$ мм частота $f_n \approx 1$ ГГц, а при $l_1 \approx 1$ мкм частота $f_n \approx 100$ ГГц. Важную роль в формировании домена играет концентрация n свободных носителей заряда, которая определяет скорость процесса накопления зарядов Q в области высокого поля. Этот процесс протекает во времени t по экспоненциальному закону

$$Q(t) = Q(0) \exp(t/\tau_p), \quad (3.22)$$

где τ_p — максвелловское время релаксации — определяется соотношением

$$\tau_p = \epsilon \epsilon_0 / (en \mu_{\text{ср}}). \quad (3.23)$$

Здесь $\mu_{\text{ср}}$ — средняя подвижность электронов, определяемая выражением (3.13); ϵ — диэлектрическая проницаемость кристаллической решетки полупроводника; ϵ_0 — диэлектрическая постоянная. Чтобы процесс формирования домена завершился по крайней мере за время прохождения домена по кристаллу ($t_{\text{пр}} = l_1/v_{\text{др. дом}}$), показатель степени в экспоненте в формуле

(3.22) должен быть больше единицы. С учетом (3.23) и выражения для $t = t_{\text{пр}}$ это условие запишется в виде

$$e\mu_{\text{ср}}l_1/(\epsilon\epsilon_0v_{\text{др. дом}}) > 1 \quad (3.24)$$

или

$$nl_1 > \epsilon\epsilon_0v_{\text{др. дом}}/(e\mu_{\text{ср}}). \quad (3.25)$$

Для арсенида галлия получаем

$$\epsilon\epsilon_0v_{\text{др. дом}}/(e\mu_{\text{ср}}) \approx 10^{12} \text{ см}^{-2}. \quad (3.26)$$

Таким образом, условие окончательного формирования домена в арсениде галлия можно записать в виде

$$nl_1 \geq 10^{12} \text{ см}^{-2}. \quad (3.27)$$

Отсюда, в частности, следует, что при уменьшении l_1 , т. е., как видно из формулы (3.21), при использовании диода Ганна на более высоких частотах следует применять материалы с большей концентрацией примесей.

Если условие (3.26) не выполняется, то в кристалле не возникает домена высокого поля, так как на неоднородности формируется не диполь — совокупность областей с повышенной концентрацией зарядов противоположных знаков (см. рис. 3.19, з), а слой повышенной концентрации зарядов одного знака. Периодически возникая, перемещаясь по кристаллу и исчезая на аноде, подобно доменам, такие слои создают нарастающие по амплитуде вдоль кристалла волны объемного заряда. В этих условиях диод Ганна может быть использован как усилитель сигналов с частотой $nf_{\text{пр}}$, где $n = 1, 2, 3, \dots$

Работа диода Ганна в цепи с колебательным контуром. Такая цепь, показанная на рис. 3.22, а, содержит помимо источника постоянного напряжения U_0 , шунтированного емкостью C , колебательный контур с параметрами L_0, C_0 . Резонансная частота контура $\omega_0 = (L_0C_0)^{-1/2}$ может быть равна пролетной частоте $\omega_{\text{пр}} = 2\pi f_{\text{пр}}$ или же несколько выше этой частоты.

За счет батареи U_0 в кристалле создается постоянное электрическое поле, величина которого $\epsilon_1 = U_0/l_1$ определяет рабочую точку «а» в области характеристики с отрицательным дифференциальным сопротивлением (рис. 3.22, б). В режиме стационарных колебаний переменное напряжение, развивающееся на коле-

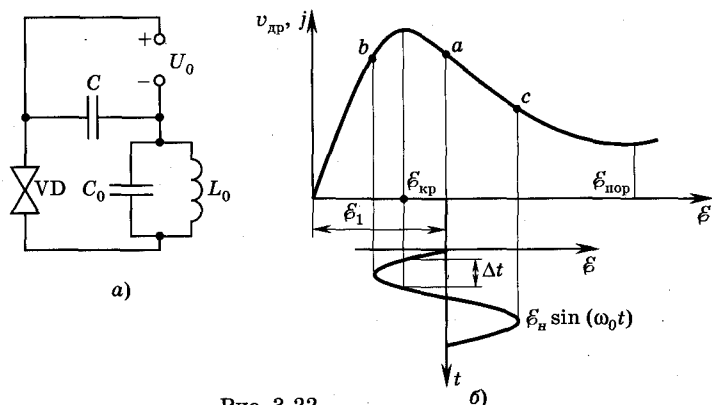


Рис. 3.22

батальном контуре, суммируется с напряжением U_0 , и в кристалле создается переменное электрическое поле $\mathcal{E}_2 = \mathcal{E}_1 + \mathcal{E}_m \sin \omega_0 t$. В соответствии с мгновенным значением рабочая точка перемещается по характеристике между точками b и c . Из рис. 3.22, b видно, что в течение некоторого интервала времени Δt напряженность поля \mathcal{E}_2 может быть меньше критического значения $\mathcal{E}_{кр}$. В эти интервалы времени домен высокого поля, если он сформировался, может разрушиться, не достигнув анода. Если же к началу периода формирование нового домена еще не наступило, то его образование начнется лишь по истечении периода Δt , т. е. с некоторой задержкой по отношению к моменту исчезновения прежнего домена. И, наконец, к началу периода процесс формирования нового домена может оказаться незавершенным: в этом случае в кристалле формируется не домен высокого поля в виде двойного слоя электрических зарядов разных знаков, а область более плотного отрицательного объемного заряда, который и перемещается по кристаллу по направлению к аноду.

В любом из этих случаев колебания в контуре не прекращаются; может измениться лишь частота колебаний или их форма. Первый случай называют *режимом с подавлением домена*, второй случай — *режимом с задержкой домена*, а третий — *режимом ограничения накопления* объемного заряда (ОНОЗ). Таким образом, помимо пролетного режима, рассмотренного нами ранее, генератор на диоде Ганна может работать и в других режимах. Наступление того или иного режима зависит от скорости формирования или рассасывания домена, скорости его перемеще-

ния по кристаллу (или от длины кристалла l при $v_{\text{др. дом}} = \text{const}$), а также величины интервала Δt , однозначно определяемого собственной частотой контура ω_0 . Поэтому важным параметром, определяющим условия существования того или иного режима, служит величина $f_0 l$, где $f_0 = \omega_0 / 2\pi$.

Отметим основные особенности различных режимов.

В *пролетном режиме* контур настраивается на частоту $f_0 = f_{\text{пр}}$. Поскольку в GaAs $v_{\text{др. дом}} \approx 10^7$ см/с, для пролетного режима произведение $f_0 l \approx 10^7$ см/с. Значения $n l$ для пролетного режима лежат в пределах $(1...3)10^{12}$ см⁻². Максимальная величина КПД не превышает 10%.

В *режиме с подавлением домена* колебательный контур настраивается на частоту выше пролетной $f_0 l > 2 \cdot 10^7$ см/с. Наивысшее значение частоты колебаний ограничивается временем разрушения домена и для арсенида галлия определяется условием $n/f_0 > 2 \cdot 10^5$ см⁻³ · с. Теоретическое значение КПД не превышает 13%.

Режим с задержкой домена наблюдается при условии, когда напряженность \mathcal{E}_2 суммарного поля снижается до величины меньше $\mathcal{E}_{\text{кр}}$ в тот момент, когда домен исчезает на аноде. Этот режим наблюдается при условии $f_0 l < 10^7$ см/с. Теоретическое значение КПД не превышает 27%.

Режим с ограниченным накоплением объемного заряда (ОНОЗ) наиболее часто используется в генераторах на диодах Ганна. В этом режиме колебательный контур настраивается на частоту $f_0 > f_{\text{пр}}$. Режим ОНОЗ определяется условием $2 \cdot 10^5$ см⁻³ · с $\gg n/f > 2 \cdot 10^4$ см⁻³ · с (здесь концентрация n записывается в см⁻³, частота — в Гц).



Контрольные вопросы

1. Каковы классификация, системы параметров и технология изготовления полупроводниковых диодов?
2. Низкочастотные выпрямительные диоды: особенности, параметры, характеристики.
3. Высокочастотные диоды: особенности, устройство, параметры, характеристики.
4. Стабилитроны: объяснить особенности ВАХ при туннельных и лавинных пробоях, ТКН и другие параметры.

5. Варикапы: ВФХ, динамический диапазон, эквивалентные схемы, параметры.
6. Импульсные диоды: особенности, импульсные параметры и их объяснение, переходные процессы.
7. Диоды Шоттки: объяснение ВАХ, особенности, параметры.
8. Туннельные диоды: объяснение ВАХ, области ВАХ, соответствующие туннельному току, эквивалентная схема, параметры.
9. ЛПД: структурные схемы, распределение концентрации и напряженности электрического поля; объяснить механизм отрицательного динамического сопротивления.
10. Параметры и характеристики физических процессов в пролетном режиме и режиме с захваченной плазмой.
11. Диоды Ганна: объяснить ВАХ и физические процессы при доменной неустойчивости.
12. Охарактеризуйте режим работы диодов Ганна.
13. Каковы основные принципы технологии полупроводниковых диодов?
14. Импульсные и переходные процессы в диодах на основе $p-n$ -переходов.

Глава 4

БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

4.1. Общие вопросы. Устройство, режимы работы транзисторов

Биполярный транзистор — это электропреобразовательный полупроводниковый прибор с одним или несколькими электрическими переходами, имеющий три или более выводов. Термин «биполярный» в названии этих транзисторов отражает тот факт, что процессы в них определяются движением носителей заряда обоих знаков (электронов и дырок). В основе работы биполярных транзисторов лежит инжекция через $p-n$ -переход неосновных носителей, заряд которых компенсируется основными носителями.

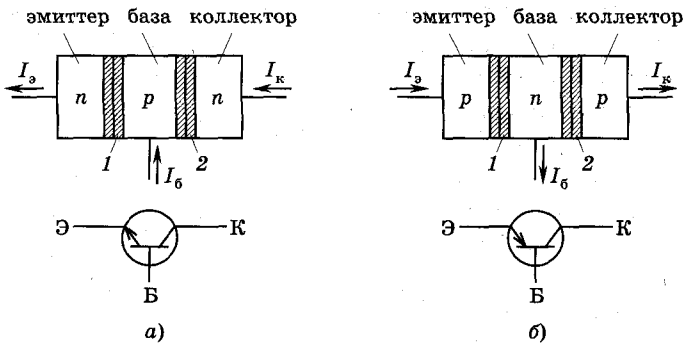


Рис. 4.1

Принципиальная структура биполярного транзистора включает три полупроводниковых области $n-p-n$ (рис. 4.1, а) или $p-n-p$ -типа (рис. 4.1, б), которые соответственно называются **эмиттером**, **базой** и **коллектором**. Так, $p-n$ -переход между эмиттером и базой (1) называется **эмиттерным**, а между базой и коллектором (2) — **коллекторным** (см. рис. 4.1, а, б). Помимо структуры транзисторов, на рис. 4.1 (внизу) приведены и их условные обозначения в схемах, где стрелка указывает направление тока при прямом смещении эмиттерного и обратном смещении коллекторного $p-n$ -перехода.

Возможны три схемы включения биполярных транзисторов: с общей базой, общим эмиттером и общим коллектором. На рис. 4.2 показаны две из них. Направления токов и полярности напряжений соответствуют нормальным условиям работы (**активному режиму**), т. е. прямому смещению эмиттерного $p-n$ -перехода и обратному смещению коллекторного перехода. Кроме этого режима возможна работа транзистора еще в трех режи-

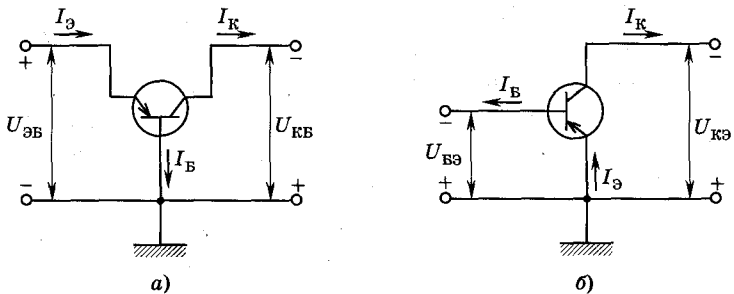


Рис. 4.2

мах: отсечки, двойной инжекции или насыщения и инверсном. В режиме отсечки оба перехода смещены в обратном направлении, в режиме двойной инжекции на оба перехода поданы прямые напряжения; в инверсном режиме коллекторный переход смещен в прямом, а эмиттерный — в обратном направлении.

По конструктивным особенностям и технологии изготовления биполярные транзисторы могут быть эпитаксиально-планарными, планарными, диффузионными, диффузионно-сплавными, сплавными и т. д.

В настоящее время транзисторы изготавливаются преимущественно из кремния. На рис. 4.3, а представлена полупроводниковая структура кремниевого эпитаксиально-планарного транзистора, характерная для большинства дискретных транзисторов.

На поверхности полупроводниковой пластины формируется тонкий диэлектрический слой SiO_2 . Сильнолегированная подложка n^+ -типа (1) вместе со слаболегированным эпитаксиальным слоем n -типа (2) толщиной $W_{\text{эп}} \sim 10$ мкм образуют коллекторную область. Области базы p -типа (3) и эмиттера n^+ -типа (4) (рис. 4.3, а, б) создаются методом диффузии или ионной имплантации. Электроды формируются тонкопленочными металлическими полосками 5, 6, 7 (рис. 4.3, а, где $a_{\text{э}}$ ширина эмиттера). Распределение концентрации примесей в направлении от поверхности (слой SiO_2) через эмиттер к коллектору приведено на рис. 4.3, б, где $N_{\text{дэ}}$, $N_{\text{аб}}$, $N_{\text{дк}}$, $N_{\text{дп}}$ — концентрация доноров в эмиттере, акцепторов в базе, доноров в коллекторе и подложке соответственно. Толщина базы $W_{\text{б}}$ современных маломощных высокочастотных транзисторов составляет 0,2...1 мкм.

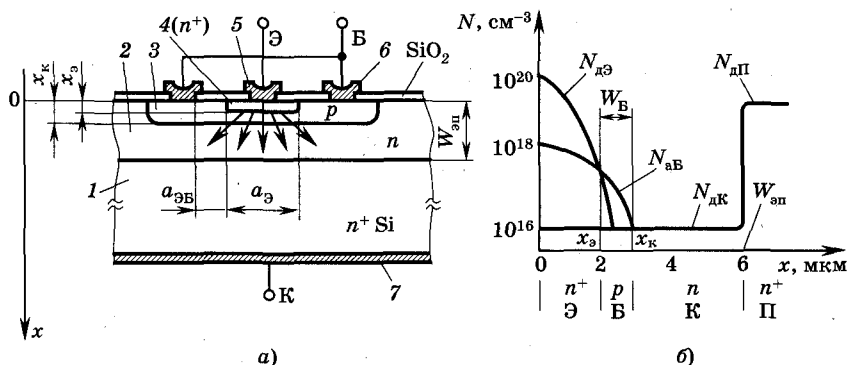


Рис. 4.3

4.2. Физические процессы в нормальном активном режиме. Коэффициенты передачи тока

В активном режиме, который является наиболее распространенным, особенно для усилительных схем, эмиттерный переход смещен в прямом направлении, а коллекторный — в обратном.

Рассмотрим транзистор $p-n-p$ -типа. (Хотя на практике чаще используют $n-p-n$ -транзисторы, дальнейшее рассмотрение будем проводить на основе $p-n-p$ -транзисторов, так как для них направление движения дырок совпадает с направлением тока, что облегчает понимание.) В этом случае дырки, инжектированные из эмиттера в базу, движутся к коллекторному переходу. Инжекцией электронов из базы в эмиттер можно пренебречь, поскольку концентрация примесей в эмиттерной области, как правило, много больше, чем в базовой. Движение инжектированных носителей через базу обусловлено как диффузией, так и дрейфом носителей. Диффузия вызвана повышением концентрации носителей из-за их инжекции в базу около эмиттерного перехода. В области, примыкающей к коллекторному переходу, под действием обратного напряжения происходит экстракция дырок. Дрейфовое движение вызвано внутренним электрическим полем в базе, возникающим из-за неравномерного распределения в ней примеси. Такие транзисторы с неоднородно легированной базой, в которой дрейфовое движение играет значительную роль, называют дрейфовыми.

Возникновение внутреннего поля можно проиллюстрировать схемой на рис. 4.4, где представлено распределение доноров в n -базе, аналогичное показанному на рис. 4.3, б. Неравномерное распределение примеси в базе, а следовательно, и основных носителей, поскольку при комнатной температуре вся примесь ионизована, вызывает диффузию электронов в направлении

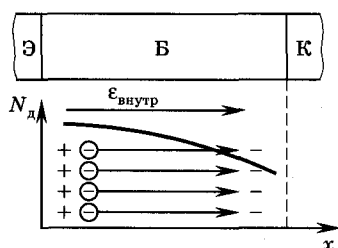


Рис. 4.4

коллектора. Из-за ухода электронов в базу со стороны эмиттерного перехода образуется избыточный нескомпенсированный заряд ионов доноров, обозначенный на рис. 4.4 крестиками «+», а со стороны коллектора ушедшие электроны образуют избыточный отрицательный заряд «-». В результате сформировавшихся избыточных зарядов и возникает внут-

реннее поле $\mathcal{E}_{\text{внутр}}$, которое будет ускоряющим для инжектированных из эмиттера дырок. Инжектированные дырки, пройдя область базы, будут втягиваться в коллектор ускоряющим электрическим полем. Часть инжектированных дырок при их движении к коллектору будет рекомбинировать в области базы, образуя базовый ток. Число рекомбинировавших носителей невелико, поскольку толщина базы мала по сравнению с диффузионной длиной дырок. В результате токи эмиттера $I_{\mathcal{E}}$ и коллектора $I_{\mathcal{K}}$ различаются незначительно и их разность равна току базы $I_{\mathcal{B}}$, т. е. $I_{\mathcal{B}} = I_{\mathcal{E}} - I_{\mathcal{K}}$. Коллекторный ток очень слабо зависит от напряжения на коллекторном переходе, поскольку при любом обратном напряжении все дырки, дошедшие до коллекторного перехода, ускоряются его полем и уносятся в коллектор. Направление токов можно проследить по схеме на рис. 4.2.

Слабое влияние коллекторного напряжения на коллекторный ток приводит к тому, что дифференциальное сопротивление коллекторного перехода $r_{\mathcal{K}} = dU_{\mathcal{KB}}/dI_{\mathcal{K}}$ очень велико, что характерно для $p-n$ -перехода, смещенного в обратном направлении. В такой ситуации в коллекторную цепь можно включить достаточно большой нагрузочный резистор $R_{\mathcal{H}}$ практически без изменения коллекторного тока. Если входной ток эмиттера увеличивается на величину $\Delta I_{\mathcal{E}}$, то приращение коллекторного тока $\Delta I_{\mathcal{K}}$ будет приблизительно тем же самым, т. е. $\Delta I_{\mathcal{E}} \approx \Delta I_{\mathcal{K}}$. Увеличение входной мощности $\Delta P_{\text{вх}}$, потребляемой в эмиттерной цепи, определяется величиной $\Delta I_{\mathcal{E}}$ и дифференциальным сопротивлением эмиттерного перехода $r_{\mathcal{E}} = dU_{\mathcal{ЭБ}}/dI_{\mathcal{E}}$, которое для прямосмещенного перехода очень мало по сравнению с сопротивлением обратносмещенного коллекторного перехода $r_{\mathcal{K}}$, т. е. $r_{\mathcal{K}} \gg r_{\mathcal{E}}$. В результате значение приращения входной мощности $\Delta P_{\text{вх}} = \Delta I_{\mathcal{E}} \Delta U_{\mathcal{ЭБ}} = \Delta I_{\mathcal{E}}^2 r_{\mathcal{E}}$ будет много меньше выделяемого на нагрузке $R_{\mathcal{H}}$ изменения выходной мощности $\Delta P_{\text{вых}} = \Delta I_{\mathcal{K}} \Delta U_{\mathcal{KB}} = \Delta I_{\mathcal{K}}^2 R_{\mathcal{H}}$, поскольку обычно $R_{\mathcal{H}} \gg r_{\mathcal{E}}$. Таким образом, схема усиливает с коэффициентом усиления $K = \Delta P_{\text{вых}}/\Delta P_{\text{вх}} = R_{\mathcal{H}}/r_{\mathcal{E}} > 1$. В активном режиме ток эмиттера $I_{\mathcal{E}}$ для $p-n-p$ -транзистора состоит из токов инжекции дырок в базу $I_{\mathcal{E}p}$ и электронов из базы в эмиттер $I_{\mathcal{E}n}$, а также тока рекомбинации в переходе $I_{\mathcal{E}рек}$, т. е. $I_{\mathcal{E}} = I_{\mathcal{E}p} + I_{\mathcal{E}n} + I_{\mathcal{E}рек}$. Из всех составляющих ток инжекции

дырок $I_{Эр}$ из эмиттера в базу определяет выходной коллекторный ток, и, следовательно, является полезным. Остальные две составляющие относятся к потерям, и их необходимо по возможности уменьшать. Полный ток коллектора I_K , помимо тока инжекции, учитывает ток рекомбинации в базе $I_{Bрек}$ и обратный ток коллекторного перехода $I_{КВО}$, который не зависит от тока эмиттера. Рекомбинацию инжектированных носителей в базе учтем введением коэффициента α — *статического коэффициента передачи тока эмиттера* в схеме с общей базой (ОБ). В результате полный ток коллектора можно записать в форме

$$I_K = \alpha I_{Э} + I_{КВО}. \quad (4.1)$$

Из выражения (4.1) следует, что

$$\alpha = (I_K - I_{КВО})/I_{Э} \approx I_K/I_{Э}. \quad (4.2)$$

В выражении (4.2) приближенное соотношение справедливо для рабочих токов I_K , которые обычно много больше $I_{КВО}$. Физически α определяется коэффициентами инжекции эмиттера $\gamma_{Э} = I_{Эр}/I_{Э}$ и переноса носителей через базу $\lambda_B = I_K/I_{Эр}$, т. е.

$$\alpha = \gamma_{Э}\lambda_B. \quad (4.2, a)$$

Коэффициент инжекции $\gamma_{Э}$ показывает, какую часть составляет полезный ток инжекции дырок из эмиттера в базу в полном токе эмиттера. Величина λ_B отражает потери инжектированных дырок при их движении через базу за счет рекомбинации. Рекомбинация определяет ток базы, который равен $I_B = I_{Э} - I_K$. Если воспользоваться соотношением (4.1), то можно получить

$$I_B = (1 - \alpha)I_{Э} - I_{КВО}. \quad (4.3)$$

Из выражения (4.3) видно, что при токе $I_{Э} = I_{КВО}/(1 - \alpha)$ ток $I_B = 0$. Рабочие токи эмиттера значительно больше $I_{КВО}/(1 - \alpha)$, тогда ток базы можно вычислить по формуле

$$I_B \approx (1 - \alpha)I_{Э} = I_{Э} - I_K = I_{Эл} + I_{Эрек} + I_{Bрек}. \quad (4.4)$$

В импульсных и цифровых интегральных схемах достаточно широко используется инверсный режим, когда в противоположность нормальному режиму роли эмиттера и коллектора меняют-

ся местами. В инверсном режиме коллекторный переход смещен в прямом направлении, а эмиттерный — в обратном. Входным током в схеме с ОБ будет коллекторный ток, а выходным — эмиттерный. Аналогично (4.1) для инверсного режима

$$I_{\text{Э}} = \alpha_I I_{\text{К}} + I_{\text{ЭВО}}, \quad (4.5)$$

где α_I — инверсный коэффициент передачи тока, $I_{\text{ЭВО}}$ — обратный ток эмиттерного перехода при $I_{\text{К}} = 0$.

Из (4.5) следует, что

$$\alpha_I = (I_{\text{Э}} - I_{\text{ЭВО}})/I_{\text{К}}, \quad (4.6)$$

причем аналогично (4.2) $\alpha_I = \gamma_{\text{К}} \lambda_{\text{БИ}}$, где $\gamma_{\text{К}}$ — коэффициент инжекции коллектора, $\lambda_{\text{БИ}}$ — инверсный коэффициент переноса.

Для большинства транзисторов $\alpha_I > \alpha$, поскольку коллекторный переход не обладает, в отличие от эмиттерного, свойством односторонней инжекции, так как концентрация примеси в коллекторной области много меньше, чем в эмиттерной (см. рис. 4.3). В результате $\gamma_{\text{К}} < \gamma_{\text{Э}}$. Помимо этого, внутреннее поле базы является тормозящим для носителей, движущихся из коллектора в эмиттер, что уменьшает $\lambda_{\text{БИ}}$, и в результате оказывается $\lambda_{\text{БИ}} < \lambda_{\text{Б}}$, кроме того, $\lambda_{\text{Б}}$ может уменьшаться из-за рекомбинации носителей в пассивной базе (окисленной поверхности полупроводника базы или на базовом контакте).

Для схемы с ОЭ входным током является ток базы $I_{\text{Б}}$, а выходным — ток коллектора $I_{\text{К}}$. В этом случае, пользуясь выражением (4.1) и учитывая, что $I_{\text{Э}} = I_{\text{К}} + I_{\text{Б}}$, можно для коллекторного тока получить следующее выражение

$$I_{\text{К}} = \alpha I_{\text{Б}}/(1 - \alpha) + I_{\text{КВО}}/(1 - \alpha). \quad (4.7)$$

Введем обозначение $\beta = \alpha/(1 - \alpha)$. Коэффициент β называется *статическим коэффициентом передачи тока базы*. Окончательно выражение (4.7) можно записать в виде

$$I_{\text{К}} = \beta I_{\text{Б}} + (1 + \beta) I_{\text{КВО}}. \quad (4.8)$$

Из этой формулы следует:

$$\beta = (I_{\text{К}} - I_{\text{КВО}})/(I_{\text{Б}} + I_{\text{КВО}}). \quad (4.9)$$

Таким образом, β есть отношение выходного коллекторного тока к входному базовому току. Высококачественные транзисторы имеют $\alpha > 0,99$, тогда $\beta \geq 100$.

В режиме насыщения происходит двусторонняя инжекция неосновных носителей через оба перехода, которые смещены в прямом направлении. В этом случае ток базы будет больше по сравнению с нормальным, или активным режимом (НАР), поскольку из-за инжекции носителей из базы в коллектор и из коллектора в базу происходит дополнительная рекомбинация носителей и $I_B > (1 - \alpha)I_E$ для схемы с ОБ или $\beta I_B > I_K$ для схемы с ОЭ.

В режиме отсечки на оба перехода подаются обратные напряжения и через переходы протекают обратные токи $I_{ЭБО}$ и $I_{КВО}$. Поскольку площадь и толщина коллекторного перехода больше, чем эмиттерного (степень легирования эмиттерной области много больше, как правило, чем коллекторной), то $I_{КВО} \gg I_{ЭБО}$.

Рассмотренные коэффициенты передачи токов зависят от всех составляющих токов, протекающих во всех цепях транзистора, поэтому α и β будут изменяться как функции тока эмиттера, напряжения на коллекторе, температуры и т. д.

4.3. Модель Эберса—Молла.

Статические характеристики биполярных транзисторов

Статические характеристики биполярных транзисторов. Биполярный транзистор есть совокупность двух встречно включенных взаимодействующих $p-n$ -переходов. Его можно представить в виде эквивалентной схемы, которая представляет собой физическую модель транзистора.

Аналитические выражения для ВАХ биполярных транзисторов можно получить на основе использования одной из таких моделей — *модели Эберса—Молла*, которая позволяет определить связь между физическими параметрами и электрическими характеристиками транзистора и отражает принципиальную равноправность его переходов. Простейший вариант этой модели для активного режима $p-n-p$ -транзистора представлен на рис. 4.5, где диоды VD_1 и VD_2 соответственно моделируют свойства эмиттерного и коллекторного переходов. Источник тока $\alpha I'_E$ учитывает передачу тока из эмиттера в коллектор, а источник $\alpha_I I'_K$ —

из коллектора в эмиттер, где α_I — инверсный коэффициент передачи тока. Токи $I'_Э$ и $I'_К$ определяются формулами ВАХ p — n -переходов (см. п. 2.3.1), т. е.

$$I'_Э = I_{Э0} [\exp(U_{БЭ}/\varphi_T) - 1],$$

$$I'_К = I_{К0} [\exp(U_{БК}/\varphi_T) - 1], \quad (4.10)$$

где величины $I_{Э0}$, $I_{К0}$ имеют смысл обратных тепловых токов соответствующих переходов, $U_{БЭ} = U_B - U_Э = -U_{ЭБ}$, $U_{БК} = U_B - U_К = -U_{КБ}$; $U_Э$, U_B , $U_К$ — потенциалы эмиттера, базы и коллектора.

Таким образом, в представленной модели VD_1 и VD_2 отображают или инжекцию, или экстракцию носителей через эмиттерный и коллекторный переходы, источник $\alpha I'_Э$ моделирует инжекцию носителей из эмиттера в базу, их перенос через базу в коллектор, а также нежелательную инжекцию носителей из базы в эмиттер. Аналогично источник $\alpha_I I'_К$ моделирует процессы при инжекции носителей через коллекторный переход и перенос зарядов через базу в эмиттер.

Из рис. 4.5 определим $I_Э$ и $I_К$, которые в НАР связаны с внутренними токами модели соотношениями

$$I_Э = I'_Э - \alpha_I I'_К, \quad I_К = \alpha I'_Э - I'_К. \quad (4.11)$$

Подставляя (4.10) в (4.11), получим выражения, позволяющие аналитически вычислить статические характеристики биполярного транзистора для любой схемы включения:

$$I_Э = I_{Э0} [\exp(U_{БЭ}/\varphi_T) - 1] - \alpha_I I_{К0} [\exp(U_{БК}/\varphi_T) - 1], \quad (4.12)$$

$$I_К = \alpha I_{Э0} [\exp(U_{БЭ}/\varphi_T) - 1] - I_{К0} [\exp(U_{БК}/\varphi_T) - 1], \quad (4.13)$$

$$I_B = I_Э - I_К = (1 - \alpha) I_{Э0} [\exp(U_{БЭ}/\varphi_T) - 1] - (1 - \alpha_I) I_{К0} [\exp(U_{БК}/\varphi_T) - 1]. \quad (4.14)$$

Выражения (4.12)—(4.14) называются **формулами Эберса—Молла**. Из них получаются выражения для различных семейств характеристик в любой схеме включения. Рассмотрим конкретные выражения для схемы с ОБ, для которой характерны заданные значения тока эмиттера и коллекторного напряжения. Поэтому характеристиками схемы с ОБ называют функции $I_К =$

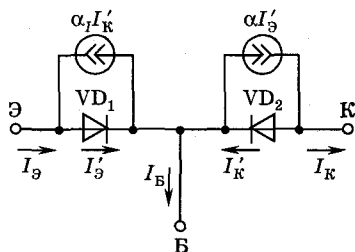


Рис. 4.5

$= f(I_{\text{Э}}, I_{\text{К}})$ и $I_{\text{Э}} = f(U_{\text{Э}}, U_{\text{К}})$. Одна из таких зависимостей $I_{\text{К}} = f(U_{\text{К}})$ с параметром $I_{\text{Э}}$ называется **семейством выходных, или коллекторных, характеристик**. Зависимость $I_{\text{Э}} = f(U_{\text{Э}})$ с параметром $U_{\text{К}}$ определяет **семейство входных, или эмиттерных, характеристик**. Формула (4.12) определяет семейство входных статических ВАХ. Семейство выходных статических ВАХ в схеме с ОБ можно получить путем исключения переменной $U_{\text{БЭ}}$ из уравнения (4.13) с учетом формулы (4.12):

$$I_{\text{К}} = \alpha I_{\text{Э}} - (1 - \alpha \alpha_I) I_{\text{КО}} [\exp(U_{\text{БК}}/\varphi_T) - 1]. \quad (4.15)$$

Семейство выходных характеристик в схеме с ОЭ можно получить из (4.13) и (4.14), сделав замену $U_{\text{БК}} = U_{\text{БЭ}} - U_{\text{КЭ}}$ и исключив переменную $U_{\text{БЭ}}$. Для $I_{\text{БЭ}} \gg I_{\text{КО}}$ получим

$$I_{\text{К}} = \beta I_{\text{Б}} [1 - (1/\alpha_I) \exp(-U_{\text{КЭ}}/\varphi_T)] / [1 + (\beta/\beta_I) \exp(-U_{\text{КЭ}}/\varphi_T)], \quad (4.16)$$

где $\beta_I = \alpha_I/(1 - \alpha_I)$.

Соотношение для входных характеристик в схеме с ОЭ получается из (4.14) также после замены $U_{\text{БК}} = U_{\text{БЭ}} - U_{\text{КЭ}}$.

В рассматриваемой простейшей модели параметры α , α_I , $I_{\text{ЭО}}$, $I_{\text{КО}}$ принимаются постоянными. Кроме того, эта модель не учитывает объемные сопротивления полупроводниковых областей, ток рекомбинации эмиттерного перехода, эффект модуляции толщины базы, эффекты высокого уровня инжекции, тока термогенерации и утечки переходов и т. д.

По аналогии с диодами эта модель является идеализированной, а реальные характеристики отличаются от теоретических так же, как реальная ВАХ от идеализированных ВАХ электрических переходов. Кроме того, рассмотренная модель определяет только статические характеристики, поэтому ее часто называют статической моделью. Таким образом, полученные аналитические выражения лишь приближенно описывают статические характеристики биполярных транзисторов.

Рассмотрим реальные семейства статических ВАХ в схемах с ОБ и ОЭ.

Входные характеристики в схеме с ОБ. Это зависимости $I_{\text{Э}} = f(U_{\text{ЭБ}})$ при постоянных значениях напряжения $U_{\text{КБ}}$ (рис. 4.6). На рис. 4.6, а изображены характеристики для малых токов и

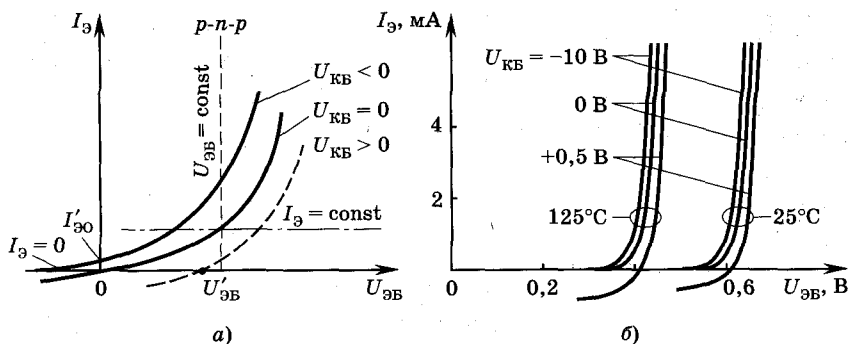


Рис. 4.6

напряжений. Для транзистора $p-n-p$ -типа положительные напряжения $U_{эб}$ соответствуют прямому включению эмиттерного перехода, а отрицательные $U_{кб}$ — обратному включению коллекторного перехода. Если $U_{кб} = 0$, то входная характеристика транзистора практически совпадает с прямой ветвью ВАХ $p-n$ -перехода. В активном режиме ($U_{эб} > 0$, $U_{кб} < 0$) характеристика смещается вверх по отношению к кривой для $U_{кб} = 0$ (см. рис. 4.6, а). Это смещение объясняется эффектом модуляции толщины базы (эффектом Эрли). Суть этого эффекта состоит в том, что при увеличении абсолютного значения $|U_{кб}|$ обедненная область коллекторного перехода расширяется, как это происходит в любом $p-n$ -переходе при увеличении обратного напряжения (см. п. 2.2.2). За счет расширения коллекторного перехода в сторону базы происходит ее сужение ($W_{Б1} > W_{Б2}$ на рис. 4.7).

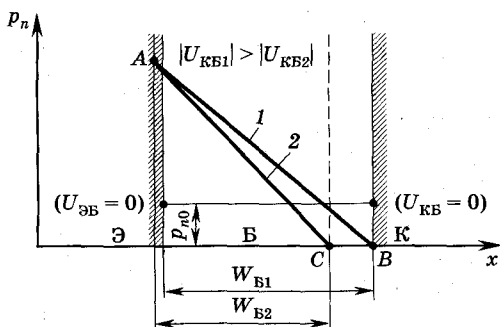


Рис. 4.7

В результате при одном и том же напряжении $U_{ЭБ}$ градиент концентрации инжектированных носителей dp_n/dx возрастает (см. рис. 4.7, прямая 2), следовательно, увеличивается и диффузионный ток инжектированных носителей, пропорциональный dp_n/dx , хотя концентрация p_n на границе и не претерпевает изменений (прямые 1 и 2 на рис. 4.7, соответствующие различным значениям $U_{КБ}$, выходят из одной точки (А) на границе эмиттерного перехода). Рост dp_n/dx увеличивает диффузионную скорость, т. е. быстрота ухода дырок из эмиттера возрастает, что и приводит к увеличению эмиттерного тока и смещению входной характеристики вверх и влево, как это показано на рис. 4.6, а при $U_{КБ} < 0$. При $U_{ЭБ} = 0$ и $U_{КБ} < 0$, хотя инжекции носителей из эмиттера в базу нет, через транзистор протекает малый ток $I'_{Э0}$ (см. рис. 4.6, верхняя кривая). Причину появления этого тока можно понять из графика на рис. 4.8, из которого видно, что за счет экстракции носителей из базы в коллектор возникает градиент неосновных носителей p_{n0} , исходно существующих в базе. За счет этого градиента происходит перенос носителей (дырок) из базы в коллектор. Для восстановления нарушенного равновесия из эмиттера в базу будет «втекать» столько дырок, сколько ушло в коллектор, что и определяет ток $I'_{Э0}$.

При подаче на коллектор положительного напряжения $U_{КБ} > 0$ и при $U_{ЭБ} > 0$ транзистор переходит в режим двойной инжекции (режим насыщения (РН)), когда помимо инжекции дырок из эмиттера происходит инжекция носителей также и из коллектора в базу. В результате градиент концентрации дырок в области базы уменьшается, хотя общее число носителей и возрастает, что приводит к уменьшению диффузионного тока, протекающе-

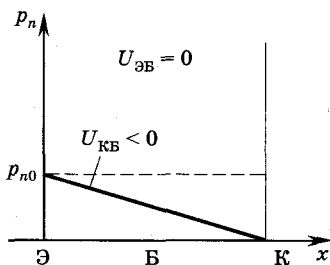


Рис. 4.8

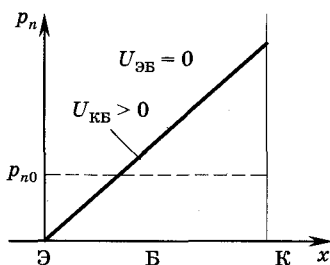


Рис. 4.9

го через базу в коллектор, и ВАХ смещается вниз относительно кривой $U_{КБ} = 0$ (штриховая кривая на рис. 4.6, а). При $U_{ЭБ} < U'_{ЭБ}$ (см. рис. 4.6, а, штриховая кривая) транзистор переходит в режим, при котором инжекция носителей из коллектора в базу преобладает над инжекцией из эмиттера в базу, и ток эмиттера изменяет направление. При $U_{ЭБ} = 0$ инжекция из эмиттера прекращается и ток эмиттера определяется инжекцией носителей из коллектора, т. е. транзистор работает в инверсном режиме. Распределение носителей в базе для этого случая дано на рис. 4.9.

Следует отметить, что изображенные на рис. 4.6, а входные характеристики соответствуют малым токам и напряжениям. Для номинальных режимов работы кремниевых и арсенид-галлиевых транзисторов в линейном масштабе значений эти токи отразить невозможно, поэтому часто характеристики для реальных приборов выглядят так, как это представлено на рис. 4.6, б.

Выходные характеристики в схеме с ОБ (рис. 4.10). Как следует из анализа физических процессов транзистора в схеме с ОБ, коллекторный ток в НАР практически равен эмиттерному и очень мало зависит от изменения напряжения $U_{КБ}$. Незначительное увеличение $I_{К}$ при увеличении обратного напряжения на коллекторном переходе связано с эффектом Эрли, т. е. при росте $|U_{КБ}|$ происходит сужение базы за счет расширения коллекторного перехода, что приводит к уменьшению интенсивности рекомбинации дырок при их движении от эмиттера к коллектору и, следовательно, к незначительному росту $I_{К}$. Коллекторный ток практически остается неизменным даже при $U_{КБ} = 0$,

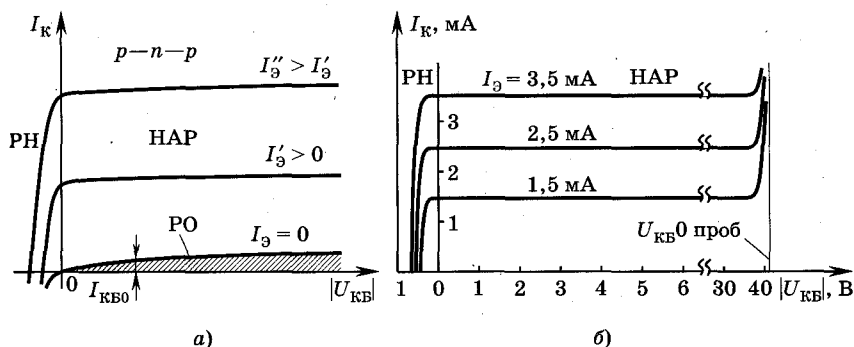


Рис. 4.10

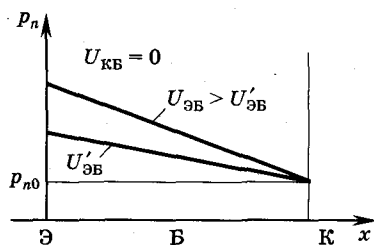


Рис. 4.11

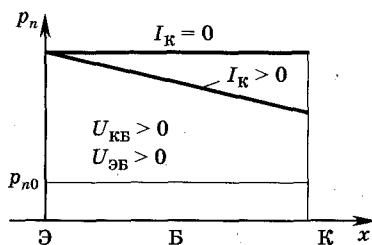


Рис. 4.12

так как избыточные инжектированные дырки продолжают извлекаться коллектором за счет контактной разности потенциалов в коллекторном переходе (рис. 4.11). При подаче на коллекторный переход положительного смещения $U_{KB} > 0$ ток коллектора падает до нуля, если плотность дырок в базе у коллектора в режиме двойной инжекции (РН) будет такой же, как и на границе эмиттерного перехода (рис. 4.12).

При увеличении коллекторного напряжения до значений, близких к напряжению пробоя коллекторного перехода, коллекторный ток начинает резко нарастать (см. рис. 4.10, б). Величина пробивного напряжения примерно такая же, как для отдельного $p-n$ -перехода (см. гл. 2). При очень узкой базе или при слабом ее легировании пробой может быть вызван проколом базы, т. е. с увеличением U_{KB} толщина базы уменьшается практически до нуля и обедненная область коллекторного перехода смыкается с обедненной областью эмиттера, вследствие чего коллектор оказывается накоротко соединенным с эмиттером, что и приводит к появлению большого тока I_K .

В режиме отсечки (РО) $I_3 = 0$ за счет подачи на эмиттерный переход обратного смещения.

Входные характеристики в схеме с ОЭ. Это зависимости $I_B = f(U_{BЭ})$ с параметром $U_{KЭ}$. При $U_{KЭ} = 0$ и $U_{BЭ} < 0$ оба перехода смещены в прямом направлении, т. е. транзистор находится в РН, когда дырки инжектируются из эмиттера и коллектора в базу. Входной ток при заданном $U_{BЭ}$ определяется инжекцией электронов из базы в коллектор и эмиттер, а также рекомбинацией дырок в базе. Этот ток имеет наибольшее значение при $U_{KЭ} = 0$ (рис. 4.13, а, кривая 1). При увеличении $|U_{KЭ}|$ до значений, больших $U_{BЭ}$ ($|U_{KЭ}| > |U_{BЭ}|$), транзистор переходит в АР.

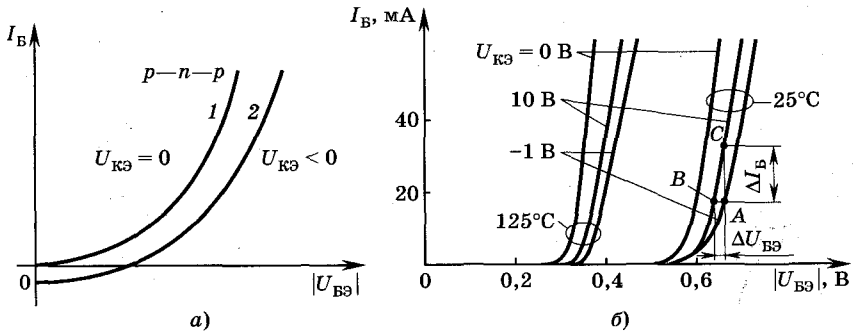


Рис. 4.13

Входной ток уменьшается из-за прекращения инжекции электронов из базы в коллектор и снижения тока рекомбинации, так как заряд дырок в базе становится меньше по сравнению с предыдущим случаем (см. рис. 4.13, а, кривая 2). При значительных напряжениях на коллекторе входной ток падает из-за уменьшения толщины базы и, следовательно, уменьшения тока рекомбинации. При больших входных токах характеристики приближаются к линейным из-за влияния сопротивления базы.

На рис. 4.13, а представлены входные характеристики для малых токов. В реальном масштабе эти токи сильно (на несколько порядков) отличаются от номинальных, поэтому входные характеристики обычно имеют вид, изображенный на рис. 4.13, б.

Выходные характеристики в схеме с ОЭ. Это зависимости $I_K = f(U_{КЭ})$ при $I_B = \text{const}$ (рис. 4.14).

Одной из характерных особенностей выходных характеристик является то, что одинаковое изменение тока базы I_B в активном режиме, когда $|U_{КЭ}| > |U_{BE}|$, вызывает неодинаковые приращения тока коллектора, т. е. характеристики неэквидистантны. Это связано с тем, что величина коэффициента передачи β зависит от тока базы (т. е. $\beta = \beta(I_B)$), который задается принудительно. Поэтому ток коллектора $I_K = [\beta(I_B)]I_B + [1]I_{КВО}$ нелинейно зависит от тока базы I_B .

Увеличение тока базы означает интенсификацию рекомбинационных процессов в области базы, т. е. происходит уменьшение коэффициента α , а поскольку $\beta = \alpha / (1 - \alpha)$, то β тоже падает.

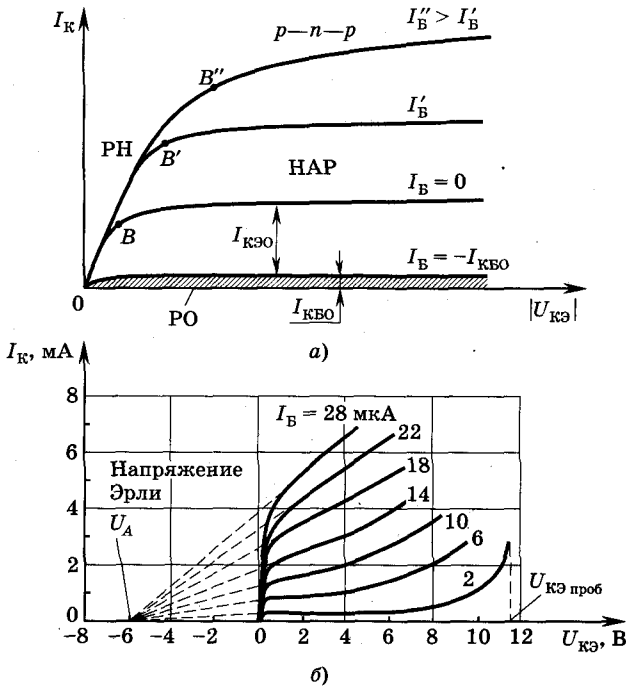


Рис. 4.14

Восходящие, крутые участки характеристик при малых напряжениях $|U_{КЭ}|$ соответствуют режиму насыщения, когда инжекция носителей в область базы происходит через оба перехода.

При $\alpha \approx \alpha_I \rightarrow 1$ выходные характеристики при $U_{КЭ} = 0$ начинаются в отличие от схемы с ОБ практически из нуля. При $U_{КЭ} = 0$ и $U_{БЭ} \neq 0$ оба перехода находятся под одним и тем же потенциалом, и через них в область базы инжектируются одни и те же токи, которые компенсируют друг друга.

Для активного режима пологий участок имеет заметно больший наклон к оси абсцисс по сравнению с выходными характеристиками для схемы включения с ОБ. Это происходит вследствие повышения β из-за уменьшения толщины базы с ростом $U_{КЭ}$ (эффект Эрли). При уменьшении толщины базы снижается ток рекомбинации в базе, а он по условиям должен быть постоянным ($I_B = \text{const}$). Для восстановления заданного I_B необходимо увеличить инжекцию носителей через эмиттерный переход, т. е. уве-

личить $U_{\text{БЭ}}$. Ток инжекции больше тока рекомбинации примерно в β раз, что и обуславливает заметное увеличение $I_{\text{К}}$ при росте $U_{\text{КЭ}}$.

Напряжение $U_{\text{А}}$, при котором пересекаются экстраполированные выходные характеристики (см. рис. 4.14, б), называется *напряжением Эрли*, которое для современных маломощных транзисторов может превышать 200...300 В. Наклон штриховых прямых на рис. 4.14, б определяет дифференциальное выходное сопротивление $r_{\text{К}} = (U_{\text{А}} + U_{\text{КЭ}})/I_{\text{К}}$ в заданной рабочей точке НАР. При больших значениях $|U_{\text{КЭ}}| \geq U_{\text{КЭ проб}}$ происходит резкое увеличение тока $I_{\text{К}}$, обусловленное пробоем. В схеме с ОЭ напряжение пробоя $U_{\text{КЭ проб}}$ значительно ниже, чем пробивное напряжение в схеме с ОБ, что связано с наличием внутренней положительной обратной связи в транзисторе. Электронно-дырочные пары, образующиеся в коллекторном переходе в результате ударной ионизации, разделяются полем перехода таким образом, что дырки перемещаются в коллектор, а электроны — в базу. Поскольку ток базы должен быть постоянным ($I_{\text{Б}} = \text{const}$), электроны накапливаются в базе, и поле их объемного заряда компенсирует заряд ионов примеси на эмиттерном переходе. В результате этого возрастает инжекция дырок в базу из эмиттера и, следовательно, увеличивается их количество, проходящее коллекторный переход, что вызывает дополнительную ионизацию атомов в переходе, и описанный процесс повторяется. Таков механизм обратной связи, вызывающей значительное увеличение коллекторного тока. Если принять коэффициент размножения носителей в коллекторном переходе равным M , то с учетом размножения коэффициент передачи тока $\beta = \alpha M / (1 - \alpha M)$. В схеме с ОЭ при пробое $\alpha M \rightarrow 1$ и $\beta \rightarrow \infty$. В схеме с ОБ пробивное напряжение больше, поскольку для возникновения пробоя требуются значительно большие значения M из-за отсутствия описанной обратной связи, присущей схеме с ОЭ.

Напряжения пробоя для схем с ОЭ и ОБ связаны эмпирической формулой

$$U_{\text{КЭ проб}} = U_{\text{КБ проб}} \sqrt[m]{1 + \beta},$$

где $m = 5$ для базы из кремния p -типа и $m = 3$ для базы n -типа.

Напряжения пробоя $U_{\text{КЭ проб}}$, $U_{\text{КБ проб}}$ измеряются при отключенной базе, когда $I_{\text{Б}} = 0$. Практически для обеспечения постоянства тока в цепи базы необходимо включить очень большое со-

противление R_B , при котором ЭДС источника питания $\mathcal{E}_B \gg U_{BЭ}$. Наличие R_B приводит к тому, что возникающий дополнительный ток рекомбинации за счет размножения носителей протекает через резистор R_B . В результате в базе накапливается меньшее число электронов, положительная обратная связь ослабевает, и напряжение пробоя увеличивается. При $R_B = 0$, когда накопление размноженных носителей минимально, напряжение пробоя максимально, а при $R_B \rightarrow \infty$ происходит максимальное накопление носителей, ток $I_B \rightarrow 0$, напряжение пробоя минимально и равно $U_{КЭ \text{ проб}}$.

На практике не рекомендуется использовать режимы с разомкнутой (отключенной) базовой цепью, поскольку низкое напряжение пробоя может привести к выходу транзистора из строя.

Влияние температуры на выходные характеристики схемы с ОЭ сильнее, чем в схеме с ОБ. Снятие выходных характеристик при различных температурах производится при $I_Э = \text{const}$ в схеме с ОБ и при $I_B = \text{const}$ в схеме с ОЭ. Поэтому в схеме с ОБ при $\alpha = \text{const}$ рост I_K при повышении температуры связан только с увеличением $I_{КБО}$. Поскольку обычно $I_{КБО} \ll \alpha I_Э$, то доля $I_{КБО}$ в коллекторном токе $I_K = \alpha I_Э + I_{КБО}$ весьма незначительна, и его можно не учитывать. В схеме с ОЭ $I_B = \text{const}$. Если допустить, что $\beta(T)$ не зависит от температуры, тогда $\beta I_B = \text{const}$, и температурная зависимость $I_K(T)$ будет определяться слагаемым $(\beta + 1)I_{КБО}$, поскольку в схеме с ОЭ $I_K = \beta I_B + (\beta + 1)I_{КБО}$. Тепловой ток коллекторного перехода, как и обычного перехода, примерно удваивается при увеличении температуры на 10°C при использовании кремния. Следовательно, при $\beta \gg 1$ прирост коллекторного тока I_K за счет величины $(\beta + 1)I_{КБО}$ может быть значительным, иногда большим исходного коллекторного тока.

4.4. Биполярный транзистор как линейный четырехполюсник. Параметры транзистора

Грамотное решение задач о рациональном создании и разработке радиоэлектронных устройств, как правило, связано с использованием различных систем параметров, в основе которых лежат так называемые физические параметры, характеризующие основные физические процессы в транзисторе. К физическим параметрам относят: коэффициенты передачи тока (α и β),

дифференциальные сопротивления переходов, объемные сопротивления областей, коэффициенты обратной связи по напряжению, емкости переходов и ряд других.

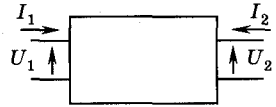


Рис. 4.15

Все перечисленные параметры так или иначе связаны с токами, напряжениями и их изменениями в цепях эмиттера, базы и коллектора. Транзистор можно рассматривать как линейный четырехполюсник (рис. 4.15), параметры которого будут определяться в основном физическими параметрами. Связь между токами и напряжениями в таком четырехполюснике может быть представлена двумя, в общем случае нелинейными, функциями, в которых в качестве независимых переменных могут выступать любые две из четырех величин (см. рис. 4.15): I_1 , I_2 , U_1 , U_2 . При малых изменениях токов и напряжений статические характеристики транзистора можно аппроксимировать линейными функциями. Функциональные зависимости переменных составляющих токов и напряжений также будут линейными. Таким образом, малые переменные составляющие, относительно сравнительно больших постоянных составляющих, можно рассматривать как малые приращения токов ΔI и напряжений ΔU .

Из большого количества функциональных зависимостей, определяемых величинами I_1 , I_2 , U_1 , U_2 , наиболее распространенными для биполярных транзисторов являются системы смешанных параметров, у которых в качестве независимых переменных выступают входной ток I_1 и выходное напряжение U_2 . Используемые функциональные зависимости имеют вид

$$U_1 = f_1(I_1, U_2) \text{ и } I_2 = f_2(I_1, U_2). \quad (4.17)$$

Вычислим полные дифференциалы для зависимостей (4.17), т. е. будем рассматривать только малые изменения токов и напряжений

$$\begin{aligned} dU_1 &= (\partial U_1 / \partial I_1) dI_1 + (\partial U_1 / \partial U_2) dU_2, \\ dI_2 &= (\partial I_2 / \partial I_1) dI_1 + (\partial I_2 / \partial U_2) dU_2. \end{aligned} \quad (4.18)$$

Введем обозначения:

$$h_{11} = \partial U_1 / \partial I_1; \quad h_{12} = \partial U_1 / \partial U_2; \quad h_{21} = \partial I_2 / \partial I_1; \quad h_{22} = \partial I_2 / \partial U_2.$$

Тогда систему уравнений (4.18) можно записать в следующем виде:

$$\begin{aligned} dU_1 &= h_{11} dI_1 + h_{12} dU_2, \\ dI_2 &= h_{21} dI_1 + h_{22} dU_2. \end{aligned} \quad (4.19)$$

В системе уравнений (4.19) дифференциалы dU_1 , dU_2 , dI_1 , dI_2 можно заменить приращениями, равными комплексным амплитудным значениям малых гармонических колебаний.

Если зафиксировать одну из зависимых переменных, то коэффициенты h_{ij} обретут вполне определенный физический смысл и примут следующий вид:

$h_{11} = dU_1/dI_1$ (при $U_2 = \text{const}$) — имеет смысл входного дифференциального сопротивления при коротком замыкании на выходе для переменной составляющей ($dU_2 = 0$);

$h_{12} = dU_1/dU_2$ (при $I_1 = \text{const}$) — коэффициент обратной связи по напряжению при разомкнутом входе для переменной составляющей тока ($dI_1 = 0$);

$h_{21} = dI_2/dI_1$ (при $U_2 = \text{const}$) — коэффициент передачи тока при коротком замыкании выхода по переменному току ($dU_2 = 0$);

$h_{22} = dI_2/dU_2$ (при $I_1 = \text{const}$) — выходная проводимость при разомкнутом входе для переменной составляющей ($dI_1 = 0$).

Конкретные значения параметров h_{ij} зависит от постоянных составляющих входного тока и выходного напряжения, от схем включения. На низких частотах влияние внутренних емкостей транзистора невелико и h -параметры являются действительными величинами. В этом случае дифференциальные h -параметры нетрудно вычислить по статическим характеристикам транзистора, заменив дифференциалы малыми приращениями.

Для примера рассмотрим схему с ОЭ, для которой входным током I_1 является ток базы I_B , а выходное напряжение $U_2 = U_{КЭ}$, входное напряжение $U_1 = U_{БЭ}$, а выходной ток $I_2 = I_K$. Параметры h_{ij} для схемы с ОЭ, т. е. $h_{11э}$, $h_{12э}$, легко определить с помощью семейства входных характеристик (см. рис. 4.13, б), а $h_{21э}$ и $h_{22э}$ вычисляются по выходным характеристикам (см. рис. 4.14).

При графическом определении h -параметров по характеристикам дифференциалы заменяются малыми конечными прира-

щениями токов (ΔI) и напряжений (ΔU). Значения приращений выбираются таким образом, чтобы нелинейностью характеристик можно было пренебречь, т. е. выбранные участки статических ВАХ должны с хорошей степенью точности аппроксимироваться линейными отрезками.

Выбирая две соседние кривые на рис. 4.13, нетрудно построить характеристический треугольник ABC , стороны которого равны приращениям $\Delta U_{BЭ}$ и ΔI_B . Приращение $\Delta U_{КЭ}$ представляет разность напряжений $U_{КЭ}$ для соседних характеристик, которая для случая, показанного на рис. 4.13, б, равна $|\Delta U_{КЭ}| = 9 \text{ В}$. Учитывая сказанное, $h_{11Э} = \Delta U_{BЭ}/\Delta I_B$ и $h_{12Э} = \Delta U_{BЭ}/\Delta U_{КЭ}$. Проводя подобные операции с семейством выходных характеристик, находим приращения ΔI_K , ΔI_B и ΔU_K и, следовательно, $h_{21Э} = \Delta I_K/\Delta I_B$ и $h_{22Э} = \Delta I_K/\Delta U_{КЭ}$.

Аналогичным образом можно вычислить h -параметры и для схемы с ОБ, в которой $I_1 = I_Э$, $I_2 = I_K$, $U_1 = U_{ЭБ}$, $U_2 = U_{КБ}$.

Рассмотрим для этого случая физический смысл h -параметров в активном режиме на низких частотах.

Входное сопротивление $h_{11Б}$ в схеме с ОБ определяется дифференциальным сопротивлением эмиттерного перехода и объемным сопротивлением базы r_B , т. е.

$$h_{11Б} = dU_{ЭБ}/dI_Э \approx \Delta U_{ЭБ}/\Delta I_Э|_{U_{КБ} = \text{const}} = r_Э + (1 - \alpha)r'_B, \quad (4.20)$$

где r'_B — эквивалентное сопротивление базы, зависящее от ее объемного сопротивления r_B и геометрии транзисторной структуры.

Сопротивление эмиттерной области обычно пренебрежимо мало из-за высокой концентрации примесей в ней. Множитель $(1 - \alpha)$ в (4.20) обусловлен тем, что в цепи базы протекает только часть переменной составляющей тока эмиттера

$$\Delta I_Э - \Delta I_K = \Delta I_Э - \alpha \Delta I_Э = (1 - \alpha) \Delta I_Э.$$

В схеме с ОЭ входным током является ток базы, поэтому входное сопротивление будет иметь иное значение, чем в схеме с ОБ:

$$\begin{aligned} h_{11Э} &= dU_{BЭ}/dI_B \approx \Delta U_{BЭ}/\Delta I_B|_{U_{КЭ} = \text{const}} \approx \\ &\approx \Delta U_{BЭ}/(1 - \alpha) \Delta I_B + r'_B = (1 + \beta)r_Э + r'_B. \end{aligned} \quad (4.21)$$

Как видно из сравнения выражений (4.20) и (4.21), входное сопротивление для схемы с ОЭ при $r'_B < r_Э$ примерно в $(1 + \beta)$ больше, чем в схеме с ОБ.

Коэффициент передачи тока $h_{21Э}$ определяется дифференцированием по I_B выражения $I_K = \beta I_B$, тогда

$$\begin{aligned} h_{21Э} &= dI_K/dI_B|_{U_{КЭ} = \text{const}} = \Delta I_K/\Delta I_B|_{U_{КЭ} = \text{const}} = \\ &= \beta + I_B(d\beta/dI_B) = \beta_D, \end{aligned} \quad (4.22)$$

где β_D — динамический коэффициент передачи тока в схеме с ОЭ.

В соответствии с известным выражением $I_K = \alpha I_Э$ и принятым направлением токов для схемы с ОБ получим

$$|h_{21Б}| = \Delta I_K/\Delta I_Э|_{U_{КБ} = \text{const}} = \alpha + I_Э(d\alpha/dI_Э) = \alpha_D, \quad (4.23)$$

α_D — динамический коэффициент передачи тока в схеме с ОБ.

Параметр $h_{12Б} = dU_{ЭБ}/dU_{КБ}|_{I_Э = \text{const}}$ определяет обратную связь (ОС) по напряжению с учетом падения напряжения на сопротивлениях базы $r'_Б$ и коллектора r_K . Физическая природа ОС обусловлена эффектом Эрли:

$$\begin{aligned} h_{12Б} &= \mu_{ОС} + r'_Б/r_K, \\ h_{12Э} &= -\mu_{ОС} + r_Э/r'_К. \end{aligned} \quad (4.24)$$

Здесь $r'_К = (U_A + U_{КЭ})/I_K$, U_A — напряжение Эрли, $\mu_{ОС}$ — статический коэффициент обратной связи по напряжению без учета $r'_Б$ и r_K . Поскольку коэффициенты α и β связаны между собой, то и все h -параметры для различных схем включения связаны между собой. Для схемы с ОЭ они выражаются через параметры схемы с ОБ следующим образом:

$$\begin{aligned} h_{11Э} &\approx h_{11Б}/(1 + h_{21Б}); \\ h_{12Э} &\approx [h_{11Б} h_{22Б}/(1 + h_{21Б})] - h_{12Б}; \\ h_{21Э} &\approx h_{21Б}/(1 - h_{21Б}); \\ h_{22Э} &\approx h_{22Б}/(1 + h_{21Б}). \end{aligned} \quad (4.25)$$

Преимуществом системы h -параметров для биполярных транзисторов является простота их измерения на переменном токе, поскольку в этом случае легко обеспечить режим холостого хода на входе из-за малости входного сопротивления и режим короткого замыкания на выходе из-за большого выходного сопротивления в активном режиме.

В результате h -параметры измеряют в режимах, близких к режимам работы транзисторов в реальных схемах.

В силу сказанного в справочниках по транзисторам низкочастотные параметры приводятся в системе h -параметров.

На очень высоких частотах и в СВЧ-диапазоне из-за влияния паразитных емкостей транзистора трудно осуществить режим холостого хода для переменных сигналов, что ограничивает применение h -параметров. В СВЧ используется специальная система S -параметров, которые формируются с помощью волновых параметров линий передачи.

Система y -параметров. Для расчета электрических схем часто целесообразно использовать другие параметры. К таким параметрам можно отнести y -параметры, при вычислении которых в качестве независимых переменных используются входное U_1 и выходное U_2 напряжения. В этом случае уравнения четырехполюсника имеют вид

$$\begin{aligned} dI_1 &= y_{11} dU_1 + y_{12} dU_2, \\ dI_2 &= y_{21} dU_1 + y_{22} dU_2, \end{aligned} \quad (4.26)$$

где $y_{11} = dI_1/dU_1$ — входная проводимость при $U_2 = \text{const}$;

$y_{12} = dI_1/dU_2$ — проводимость обратной передачи при $U_1 = \text{const}$;

$y_{21} = dI_2/dU_1$ — проводимость прямой передачи при $U_2 = \text{const}$;

$y_{22} = dI_2/dU_2$ — выходная проводимость при $U_1 = \text{const}$.

Отметим, что y -параметры достаточно просто выражаются через h -параметры:

$$\begin{aligned} y_{11} &= 1/h_{11}, \\ y_{12} &= -h_{12}/h_{11}, \\ y_{21} &= h_{21}/h_{11}, \\ y_{22} &= (h_{11}h_{22} - h_{12}h_{21})/h_{11}. \end{aligned} \quad (4.27)$$

На практике y -параметры измеряются в режимах короткого замыкания по переменному току входной и выходной цепи транзистора. Их часто используют для определения параметров полевых транзисторов (см. главу 6), которые имеют большие входные и вы-

ходные сопротивления. Для биполярных транзисторов y -параметры проще измерять на высоких частотах по сравнению с h -параметрами. Из-за малых емкостных сопротивлений на высоких частотах сильное влияние оказывают межэлектродные емкости.

Как уже отмечалось, дифференциалы токов и напряжений можно заменить комплексными амплитудами токов и напряжений. На низких частотах влияние емкостей практически отсутствует, поэтому токи и напряжения находятся в фазе и их отношения являются действительными числами.

На высоких частотах влияние емкостей приводит к тому, что между переменными токами и напряжениями происходит сдвиг по фазе. В результате входные и выходные сопротивления являются комплексными. Однако при больших входных и/или выходных сопротивлениях и на высоких частотах эти сопротивления могут носить чисто емкостный характер. В силу сказанного емкость транзистора рассматривается как один из основных параметров.

Емкости транзисторов определяются диффузионными и барьерными емкостями эмиттерного и коллекторного переходов, которые вычисляются аналогично емкостям отдельного $p-n$ -перехода (см. главу 2). При прямом напряжении ток эмиттера задает полный заряд избыточных (инжектированных) носителей, который однозначно связан с диффузионной емкостью эмиттерного перехода $C_{Э\text{ диф}}$. В соответствии с формулой (2.29) диффузионная емкость равна

$$C_{Э\text{ диф}} = I_{Э} t_{пр\text{ Б}} / \Phi_T, \quad (4.28)$$

где $t_{пр\text{ Б}}$ — среднее время пролета дырок через базу, которое должно быть меньше их времени жизни. Формула (4.28) справедлива для частот $f \ll 1/(2\pi t_{пр\text{ Б}})$.

Диффузионную емкость коллекторного перехода целесообразно рассмотреть для режима насыщения, для которого характерна двусторонняя инжекция неосновных носителей через оба перехода. При этом каждый переход, помимо инжекции носителей в базу, собирает подходящие к его границе носители, инжектированные в базу другим переходом. В режиме насыщения при том же токе эмиттера ток базы больше, чем в активном режиме, из-за инжекции электронов из базы в коллектор (для $p-n-p$ -транзисторов) и рекомбинации дырок, инжектированных из коллектора, т. е.

$$I_{Б} > (1 - \alpha)I_{Э}; \beta I_{Б} > I_{К}. \quad (4.29)$$

Неравенства (4.29) определяют условия существования режима насыщения в схеме с ОЭ. Полный заряд неосновных носителей в режиме насыщения равен сумме зарядов для активного и инверсного режимов. По сравнению с активным режимом при насыщении появляется избыточный заряд электронов в коллекторе, поскольку степень легирования базы выше, чем коллектора, этот избыточный заряд определяет диффузионную емкость коллекторного перехода

$$C_{\text{К диф}} = (\beta I_{\text{Б}} - I_{\text{К}}) \tau_{\text{К эф}} / \varphi_{\text{T}}, \quad (4.30)$$

где $\tau_{\text{К эф}}$ — эффективное время жизни неосновных носителей в коллекторе. Это выражение справедливо при частоте $f \ll 1/(2\pi\tau_{\text{К эф}})$.

4.5. Эквивалентные схемы

При рассмотрении транзистора как четырехполюсника получают параметры, зависящие сложным образом от режима работы транзистора, от частоты, температуры и т. д. Поэтому на практике для упрощения анализа свойства транзистора в режиме работы при малых переменных сигналах описывают с помощью эквивалентных электрических схем. В эквивалентных схемах транзистор заменяется электрической схемой, составленной из линейных сопротивлений, емкостей, индуктивностей, генераторов тока или напряжения, при этом физические характеристики схемы приблизительно идентичны реальному транзистору. При расчетах сначала с помощью теории цепей определяют токи и напряжения в эквивалентной схеме, а потом вычисляются другие параметры, в том числе в качестве этих параметров могут быть и параметры четырехполюсника. Разумеется, любая эквивалентная схема как модель транзистора не описывает с полной достоверностью свойства реального объекта, она является лишь неким приближением при анализе свойств транзистора.

Все эквивалентные схемы можно разделить на два класса: схемы замещения, вытекающие непосредственно из уравнений четырехполюсника, и физические эквивалентные схемы, формируемые на основе анализа физических процессов для определенных условий применения и конструкции транзистора.

Схемы замещения содержат по четыре элемента: два комплексных сопротивления, два генератора тока или два генератора напряжения. Эти эквивалентные схемы не обладают какими-

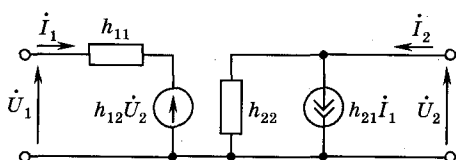


Рис. 4.16

ми-либо заметными преимуществами по сравнению с описанием транзистора с помощью уравнений четырехполюсника.

Схемы замещения могут быть Т-образные и П-образные, при этом Т-образные схемы имеют два активных элемента (генератора), а П-образные — один. Для установления связи параметров четырехполюсника и Т-образных схем замещения записываются выражения для токов и напряжений в схемах. После этого полученные значения сопоставляются с величинами токов и напряжений эквивалентного четырехполюсника. На рис. 4.16 приведена в качестве примера Т-образная эквивалентная схема для h -параметров ($\dot{U}_{1,2}$, $\dot{I}_{1,2}$ — комплексные величины гармонических напряжений и токов). Аналогичные схемы формируются для y -параметров.

Физические эквивалентные схемы создают следующим образом. Сначала выделяются некоторые части транзистора, в которых отдельно анализируются физические процессы. Простейшей эквивалентной схемой, основанной на физических соображениях, является схема, используемая в модели Эберса—Молла.

Более сложная модель, лучше приближенная к реальности, помимо четырех элементов простейшей модели содержит три резистора $r'_Э$, $r'_Б$, $r'_К$, которые учитывают влияния полупроводниковых областей эмиттера, базы и коллектора и четыре конденсатора $C_{Э\text{ бар}}$, $C_{Э\text{ диф}}$, $C_{К\text{ бар}}$, $C_{К\text{ диф}}$, которые определяют инерционные свойства эмиттерного и коллекторного переходов при работе транзистора с переменными сигналами (рис. 4.17).

На рис. 4.17 резистор $r'_Э$ исключен из-за его малой величины; значения $r'_Э$, $r'_Б \equiv r'$ и $r'_К$ могут не совпадать с объемными сопротивлениями соответствующих областей, как правило, из-за особенностей геометрии транзисторов. Наличие этих резисторов приводит к тому, что к диодам VD_1 и VD_2 приложены напряжения $U'_{БЭ}$ и $U'_{БК}$, которые меньше внешних напряжений $U_{БЭ}$ и $U_{БК}$. Поскольку диффузионные и барьерные емкости зависят от напряжений $U'_{БЭ}$ и $U'_{БК}$, то в качестве таких емкостей использу-

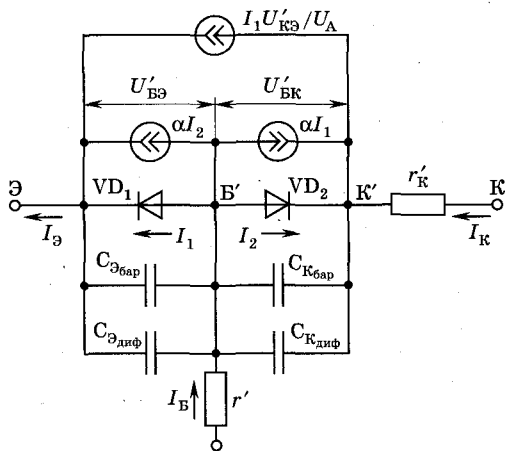


Рис. 4.17

ют либо усредненные постоянные значения, либо зависимости $C_{\text{бар}}(U)$, $C_{\text{диф}}(U)$, которые повышают точность модели. Рассмотренная модель справедлива для больших сигналов, поскольку в ней учитываются нелинейные характеристики элементов (диоды, емкости, сопротивления).

При работе на малом сигнале целесообразно использовать другие эквивалентные схемы. На рис. 4.18 представлена одна из малосигнальных моделей (Т-образная эквивалентная схема) транзистора в активном режиме, в которой диод VD_1 заменен дифференциальным сопротивлением эмиттерного перехода $r_{\text{Э}}$. Резистор $r'_{\text{Э}}$ исключен в силу его малости, источник тока $\alpha_1 I_2$ и конденсатор $C_{\text{К диф}}$ также исключены из-за незначительной величины обратного тока коллекторного перехода I_2 . Генератор

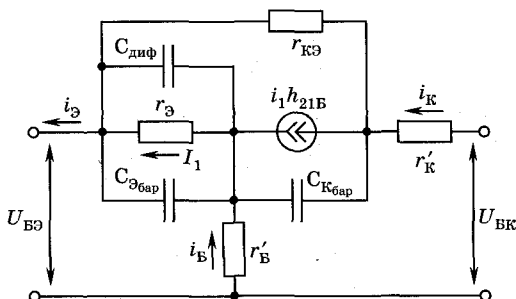


Рис. 4.18

тока $i_1 h_{21B}$ моделирует влияние динамического коэффициента передачи тока α_D (4.23). Резистор $r_{кэ} = U_A / I_{1-} = U_A / I_{к-}$, где $I_{1-} = I_{к-}$ — постоянные составляющие токов $I_1, I_{к-}$; U_A — напряжение Эрли. На схеме 4.18 строчными буквами i обозначены переменные составляющие токов. Вместо α в обозначении генератора тока использован дифференциальный параметр h_{21B} , что справедливо для малых напряжений. При заданных постоянных составляющих тока эмиттера и коллектора параметры схемы постоянны. Схема, показанная на рис. 4.18, может быть преобразована в П-образную, что часто используется для анализа и расчета усилительных схем.

4.6. Переходные и частотные характеристики биполярного транзистора

При изменении частоты сигнала или при подаче импульсных сигналов на работу транзистора и на его параметры могут существенно влиять инерционные процессы, обусловленные наличием реактивностей (в основном паразитных емкостей переходов) и конечным временем переноса носителей через область транзистора. Инерционные свойства транзистора определяют возможности его использования в конкретных электрических схемах, особенно в усилительных и генераторных устройствах, работающих на высоких частотах. При увеличении частоты время протекания физических процессов в транзисторе, вызванных изменением входного сигнала, может быть соизмеримо или превышать его период. В этом случае сопротивления конденсаторов в малосигнальных эквивалентных схемах, рассмотренных в предыдущем разделе, могут оказаться меньше дифференциальных сопротивлений эмиттерного и коллекторного переходов, объемных сопротивлений базы и коллектора. В результате этих процессов h - и y -параметры становятся комплексными величинами, зависящими от частоты. Учесть одновременно все факторы, влияющие на частотные свойства реального транзистора, очень сложно. Для того чтобы оценить частотные свойства транзистора в целом, можно допустить, что полный коэффициент передачи тока транзистора равен произведению коэффициентов передачи тока, зависящих от процессов в отдельных областях.

Рассмотрим частотную зависимость коэффициента передачи тока в схеме с ОБ h_{21B} , на которую, как отмечалось в п. 4.4,

влиять емкость цепи эмиттера, время пролета носителей заряда через базу, время пролета носителей через область объемного заряда коллекторного перехода и постоянная времени цепи коллектора. В первую очередь оценим процессы в базе, пренебрегая влиянием емкостей и явлениями в коллекторном переходе. Для простоты будем рассматривать явления в базе на примере изменения токов транзистора при подаче на его вход **функции включения**, представляющей ступенчатое изменение входного сигнала. Изменение этой функции при прохождении сигнала через транзистор связано не только с его переходными характеристиками, но также и с частотными, поскольку спектр импульсного сигнала изменяется при изменении фронтов сигнала. Предположим, что в начальный момент времени на коллектор подано постоянное обратное напряжение, а эмиттерный ток равен нулю. Не будем учитывать в коллекторном токе обратный ток термогенерации. В момент времени t_0 эмиттерный ток изменяется скачком до величины $I_{\text{Э}}$ (рис. 4.19, а), при этом для простоты считаем, что осуществляется односторонняя инжекция дырок, т. е. коэффициент инжекции $\gamma = 1$. Инжектированные в базу дырки достигают коллекторного перехода через некоторое **время задержки** t_a и при $t \geq t_0 + t_a$ коллекторный ток начинает нарастать до значения $h_{21\text{Б}}I_{\text{Э}} = \alpha I_{\text{Э}}$. Нарастание происходит постепенно, поскольку скорости отдельных носителей существенно различаются из-за того, что диффузия носителей связана со столкновениями носителей с атомами, ионами и между собой. Следовательно, скорость диффузии есть величина средняя, от-

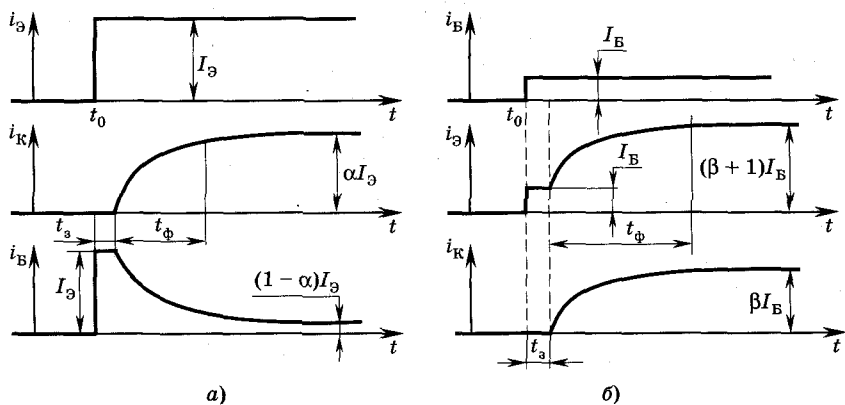


Рис. 4.19

носителем которой скорости носителей распределены по определенному закону. В результате фронт коллекторного тока имеет конечную длительность t_{ϕ} .

В течение времени t_3 , ток базы будет равен току эмиттера I_3 , затем за время t_{ϕ} он уменьшится до стационарного значения $(1 - h_{21B})I_3 = (1 - \alpha)I_3$. Одновременно с током коллектора происходит нарастание избыточного заряда в базе. Увеличение избыточного заряда и тока коллектора носит экспоненциальный характер, т. е.

$$i_K(t) = h_{21B}I_3[1 - \exp(-(t - t_0)/\tau_{КЭ})], \quad (4.31)$$

где $\tau_{КЭ}$ — постоянная времени, определяющая длительность переходного процесса коллекторного тока. Если пренебречь переходными процессами в эмиттерном и коллекторном переходах, то $\tau_{КЭ} = t_{прБ}$, где $t_{прБ}$ — время пролета носителей через базу. В случае бездрейфовых транзисторов $\tau_{КЭ} = t_D$, где t_D — среднее время диффузии носителей через базу. В общем же случае, как отмечалось ранее, структура транзистора разбивается на несколько областей в направлении от эмиттера к коллектору, тогда $\tau_{КЭ}$ может быть записана в следующем виде:

$$\tau_{КЭ} = \tau_{ЭП} + t_{прБ} + \tau_{КП} + \tau'_K, \quad (4.32)$$

где $\tau_{ЭП}$ — постоянная времени эмиттерного перехода, $t_{прБ}$ — время пролета носителей через базу, $\tau_{КП}$ — постоянная времени, которая определяется временем пролета носителей через коллекторный переход, τ'_K — постоянная времени коллектора.

Постоянная времени эмиттерного перехода $\tau_{ЭП}$ учитывает задержку нарастания тока инжекции, связанную с зарядом барьерной емкости эмиттерного перехода. В силу этого ток инжекции дырок нарастает не скачком, а экспоненциально:

$$i_3(t) = \gamma_3 I_3 [1 - \exp(-(t - t_0)/\tau_{ЭП})], \quad (4.33)$$

где γ_3 — коэффициент инжекции эмиттера, а

$$\tau_{ЭП} = C_{Эбар} r_3. \quad (4.34)$$

Время пролета носителей через коллекторный переход $t_{пр}$ может быть определено по формуле $t_{пр} = L_{обК}/v_{нас}$, где $L_{обК}$ — ширина коллекторного перехода, $v_{нас}$ — скорость насыщенная (см. п. 3.7). Из-за высокой напряженности электрического поля

скорость переноса носителей через коллекторный переход равна скорости насыщения $v_{\text{нас}}$. При движении дырок через обедненный слой они наводят в цепи коллектора ток, который начинает изменяться раньше, чем они долетят до границы слоя и перейдут в коллектор, поэтому $\tau_{\text{КП}} < t_{\text{пр}}$. Расчеты показывают, что

$$\tau_{\text{КП}} = t_{\text{пр}}/2 = L_{\text{об К}}/2v_{\text{нас}}. \quad (4.35)$$

Постоянная времени $\tau'_\text{К}$ в формуле (4.32) определяется перезарядкой барьерной емкости коллекторного перехода, т. е.

$$\tau'_\text{К} = r'_\text{К} C_{\text{К бар}}, \quad (4.36)$$

где $r'_\text{К}$ — объемное сопротивление высокоомной коллекторной области.

Зная $r_{\text{КЭ}}$, можно записать переходную характеристику $h_{21\text{Б}}(S)$ коэффициента передачи тока в схеме с общей базой в операторной форме

$$h_{21\text{Б}}(S) = h_{21\text{Б}} \exp(-St_3)/(1 + S\tau_{\text{КЭ}}), \quad (4.37)$$

где t_3 — время задержки коллекторного тока относительно фронта импульса тока эмиттера (см. рис. 4.19, а).

Используя операторную форму записи переходной характеристики и заменяя оператор S на $j\omega$ (j — мнимая единица), получим комплексную частотную характеристику коэффициента передачи в виде

$$\dot{h}_{21\text{Б}} = h_{21\text{Б}} e^{-j\omega t_3} / [1 + j(\omega/\omega_\alpha)], \quad (4.38)$$

где $\omega_\alpha = 2\pi f_\alpha = 1/\tau_{\text{КЭ}}$ — *угловая предельная частота* для схемы с ОБ.

Комплексное значение $\dot{h}_{21\text{Б}}$ может быть использовано в случае приближения малых сигналов. Полные токи в транзисторе не могут быть синусоидальными из-за вентильных свойств p — n -перехода. Амплитудно-частотная характеристика коэффициента $h_{21\text{Б}}$ представляется действительной частью выражения (4.38), т. е.

$$|h_{21\text{Б}}(\omega)| = h_{21\text{Б}} / \sqrt{1 + (\omega/\omega_\alpha)^2}. \quad (4.39)$$

Фазочастотная характеристика, полученная из (4.38), имеет вид

$$\varphi(\omega) = \omega - \arctg(\omega/\omega_\alpha). \quad (4.40)$$

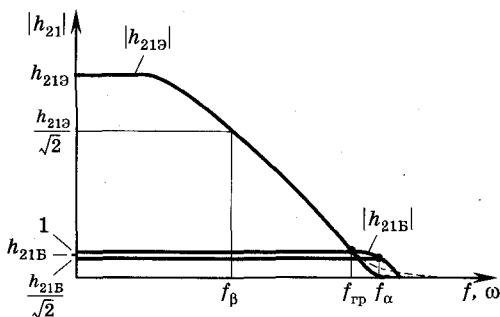


Рис. 4.20

Частотная характеристика коэффициента передачи тока $|h_{21B}(\omega)|$ показана на рис. 4.20 (нижняя кривая).

Переходные процессы в схеме с ОЭ можно проанализировать, используя рисунок 4.19, б. На вход транзистора в этом случае подается ступенька базового тока I_B . Функцию коллекторного тока можно записать в виде

$$i_K(t) = h_{21э}(t)I_B, \quad (4.41)$$

где $h_{21э}(t)$ — переходная характеристика коэффициента передачи тока ($h_{21э} \approx \beta$).

Как и ранее, полагаем $\gamma \approx 1$. Выражение для постоянной времени транзистора в схеме с ОЭ получим, учитывая связь между параметрами h_{21B} и $h_{21э}$, т. е. $h_{21э} = h_{21B}/(1 - h_{21B}) = \alpha/(1 - \alpha)$.

Переходный процесс, как и в схеме с ОБ, сохраняет экспоненциальный характер, но постоянная времени $\tau_{кэ\beta}$ в схеме с ОЭ будет много больше, чем $\tau_{кэ}$ в схеме с ОБ, в соответствии с выражением [34]

$$\tau_{кэ\beta} = \tau_{кэ}/(1 - \alpha) = (\beta + 1)\tau_{кэ}. \quad (4.42)$$

Механизм увеличения постоянной времени в схеме с ОЭ можно пояснить следующим образом. При подаче базового тока за счет поступления электронов из внешней цепи в области базы возрастает концентрация и суммарный заряд электронов. Накопление электронов приводит к тому, что их заряд компенсирует заряд ионов доноров в области эмиттерного перехода, что приводит к понижению потенциального барьера в ней и возникновению инжекции дырок из эмиттера. Инжектированные дырки будут поддерживать квазинейтральность базы, и в начальный момент $I_э = I_B$ (см. рис. 4.19, б), как и в схеме с ОБ. Инжектированные дырки до-

ходят до коллектора за время задержки t_3 (время переноса носителей через область базы), и коллекторный ток начинает увеличиваться (нижняя кривая на рис. 4.19, б). В схеме с ОБ по мере увеличения i_K ток i_B уменьшается, но в схеме с ОЭ ток базы $I_B = \text{const}$, т. е. он жестко задан, поэтому возрастание тока коллектора i_K вызывает дополнительную инжекцию дырок из эмиттера, необходимую для поддержания квазинейтральности, поскольку часть дырок уже ушла в коллектор. Одновременное возрастание токов i_K и i_3 будет происходить до тех пор, пока в базе не накопится избыточный заряд ΔQ_B , скорость рекомбинации которого будет поддерживать базовый ток постоянным, т. е.

$$\Delta Q \approx I_B \tau. \quad (4.43)$$

Следовательно, постоянная времени $\tau = \Delta Q/I_B$, т. е. время жизни носителей в базе τ равно постоянной времени транзистора в схеме с ОЭ.

Выражение для малосигнальной частотной характеристики для схемы с ОЭ можно записать в форме, аналогичной соотношению (4.38) для схемы с ОБ [34]:

$$\dot{h}_{213} = h_{213} / [1 + j(\omega/\omega_\beta)], \quad (4.44)$$

где $\omega_\beta = 1/\tau_{кэ\beta}$ — предельная частота для схемы с ОБ.

Соответственно амплитудно-частотные и фазочастотные характеристики будут иметь вид

$$|h_{213}(\omega)| = \beta / \sqrt{(1 + (\omega/\omega_\beta))^2}, \quad (4.45)$$

$$\varphi(\omega) = -\text{arctg}(\omega/\omega_\beta). \quad (4.46)$$

Граничная частота коэффициента передачи тока $\omega_{гр}$ определяется из условия уменьшения коэффициента h_{213} до единицы.

Положив $\omega = \omega_{гр}$ и $h_{213}(\omega_{гр}) = 1$ и пренебрегая единицей в подкоренном выражении (4.45), получим

$$\omega_{гр} \approx \beta \omega_\beta \approx \beta / \tau_{кэ\beta}. \quad (4.47)$$

Учитывая формулу (4.42) и равенство $\omega_\alpha = 1/\tau_{кэ}$, получаем

$$\omega_{гр} = \beta / \tau_{кэ\beta} = \frac{\beta}{(\beta + 1)\tau_{кэ}} \lesssim \frac{1}{\tau_{кэ}} = \omega_\alpha.$$

Таким образом, частота $\omega_{гр} = 2\pi f_{гр}$ близка по величине к ω_α .

Частотная характеристика $|h_{213}(\omega)|$ схемы с ОЭ изображена на рис. 4.20 (верхняя кривая).

4.7. Импульсный режим работы. Транзисторный ключ

При работе многих цифровых схем, импульсных генераторов, преобразователей импульсов, транзисторных ключей и т. д. происходит быстрое изменение токов и напряжений в больших пределах, так что в большинстве случаев проявляется нелинейность характеристик транзисторов. Такие режимы часто называют *режимами большого сигнала*.

Особенности физических и переходных процессов в режиме больших сигналов хорошо иллюстрируются на примере транзисторного ключа. При работе транзистора в ключевом режиме выходное сопротивление транзистора, непосредственно связанное с цепью нагрузки, сильно изменяется под влиянием входного управляющего импульса. При этом амплитуда входных импульсов достаточна для перевода транзистора из режима отсечки сначала в активный режим работы, а потом и в режим насыщения, а также и наоборот, т. е. из режима насыщения последовательно в режим отсечки.

Рассмотрим процессы, происходящие в транзисторе, включенном по схеме с ОЭ, при прохождении через базу импульса тока с определенной длительностью, обеспечивающего прямое смещение с последующим изменением его направления на обратное.

На рис. 4.21 изображена простейшая схема электронного ключа на основе $n-p-n$ -транзистора в схеме с ОЭ. В цепь базы включен резистор R_B , а в коллекторную цепь резистор R_K и источник постоянного напряжения E_K . Пунктирными линиями показано подключение нагрузочной емкости C_H , состоящей из

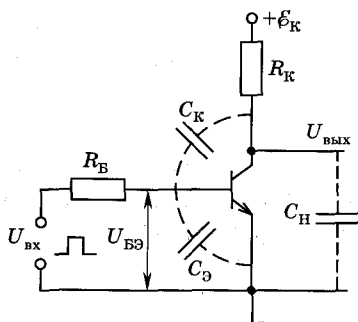


Рис. 4.21

выходной емкости ключа и входной емкости нагрузки, емкости эмиттерного и коллекторного переходов $C_{ЭВ} = C_э$ и $C_{КВ} = C_к$.

В исходном состоянии на входе существует постоянное смещение и ключ закрыт, поскольку транзистор находится в режиме отсечки, так как на оба перехода подано обратное смещение. Напряжение на выходе близко к E_K (рис. 4.22, точка С). Если на вход подать импульс положительной полярности

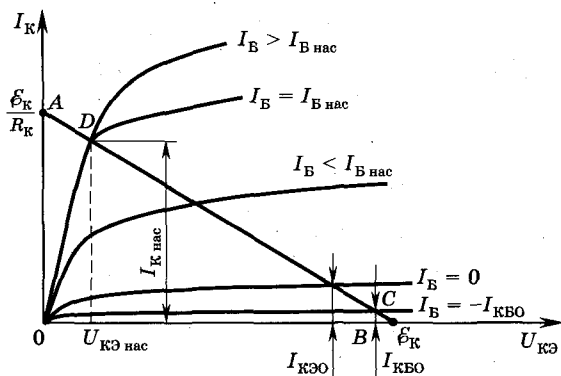


Рис. 4.22

с амплитудой ϵ_{B1} (рис. 4.23, а), то транзистор перейдет в открытое состояние, при котором он будет находиться в режиме насыщения за счет инжекции носителей как из эмиттера, так и из коллектора. В цепи базы течет ток, равный $I_{B1} = (+\epsilon_{B1} - U_{BЭ})/R_B$, где $U_{BЭ}$ — прямое падение напряжения на эмиттерном переходе, а в цепи коллектора — ток I_K создает на резисторе R_K падение напряжения $U_{RK} = R_K I_K$. Напряжение на выходе по абсолютной величине уменьшается на величину U_{RK} .

Сказанное можно проиллюстрировать с помощью графического решения уравнения для нагрузочной прямой совместно с уравнениями для выходных статических характеристик. Эти уравнения описывают токи и напряжения в схеме при наличии нагрузки в цепи коллектора. Рассмотрим этот метод решения более подробно.

Из-за падения напряжения на резисторе R_K , т. е. $U_{RK} = I_K R_K$, напряжение на коллекторе можно записать в следующем виде:

$$U_{КЭ} = \epsilon_K - U_{RK} = \epsilon_K - I_K R_K. \quad (4.48)$$

В этом уравнении две неизвестных величины ($U_{КЭ}$ и I_K), поэтому для определения параметров конкретного режима работы прибора при наличии нагрузки в его коллекторной цепи необходимо к полученному выражению (4.48) добавить еще уравнение ВАХ $I_K = f(U_{КЭ}, I_B)$. В результате получим систему из двух уравнений:

$$I_K = \epsilon_K/R_K - U_{RK}/R_K; I_K = f(U_{КЭ}, I_B). \quad (4.49)$$

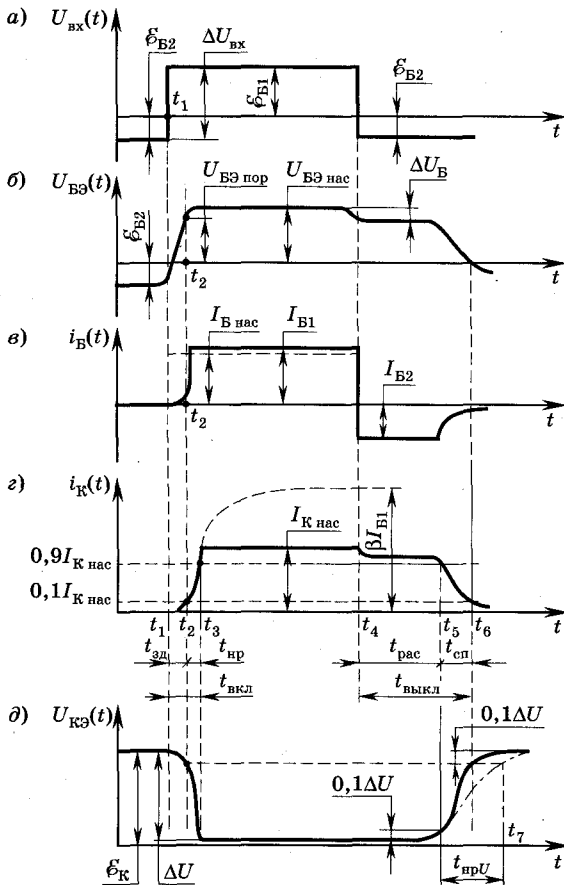


Рис. 4.23

Эту систему можно решить графически, используя семейство выходных характеристик $I_K = f(U_{КЭ})$, где в качестве параметра используется ток базы I_B . Совместное решение уравнений (4.49) определяется точкой пересечения линии нагрузки с любой выходной характеристикой $I_K = f(U_{КЭ})$.

Если выбрать выходную характеристику $I_K = f(U_{КЭ})$ при $I_B = I_{B1} = I_{Bнас}$ (см. рис. 4.22), пренебрегая влиянием R_B , и провести нагрузочную прямую, соответствующую резистору R_K , то точка D удовлетворяет решению системы (4.49). Нагрузочная прямая строится по двум точкам, одна из которых (точка B) лежит на оси абсцисс ($U_{КЭ} = \epsilon_K$), т. е. при $I_K = 0$, когда сопротивле-

ние транзистора очень велико и напряжение источника полностью приложено к коллектору. Другая точка лежит на оси ординат, когда $U_{КЭ} = 0$, т. е. транзистор закорочен или обладает бесконечно малым сопротивлением. В этом случае ток через сопротивление R_K будет равен $I_K = \mathcal{E}_K/R_K$ (точка A). Точка D пересечения нагрузочной прямой с характеристикой $I_B = I_{B \text{ нас}}$ определяет выходное напряжение ($U_{\text{вых}} = U_{КЭ \text{ нас}}$) транзистора в режиме насыщения. $U_{\text{вых}}$ должно быть как можно меньше и слабо зависимо от входного напряжения и сопротивления R_K . Эти требования наиболее полно удовлетворяются, если рабочая точка D лежит на крутом участке выходной характеристики, который соответствует режиму насыщения и $U_{\text{вых}} = U_{КЭ \text{ нас}}$ ($U_{КЭ \text{ нас}}$ — напряжение насыщения). Чтобы реализовался режим насыщения, необходимо выполнение условия $I_{B \text{ нас}} > I_{K \text{ нас}}/\beta$, где $I_{K \text{ нас}} = (\mathcal{E}_K - U_{КЭ \text{ нас}})/R_K$ — ток насыщения.

Рассмотрим более детально временные процессы перехода транзистора из закрытого (режим отсечки) в открытое состояние (режим насыщения) и снова в закрытое, т. е. проанализируем переходные процессы в транзисторном ключе на основе схемы с ОЭ. В начальном состоянии транзистор находится в режиме отсечки (точка C на рис. 4.22), оба перехода включены в обратном направлении. На базу подано отрицательное смещение \mathcal{E}_{B2} , и, следовательно, оно равно входному напряжению: $U_{\text{вх}} = -|\mathcal{E}_{B2}|$ (см. рис. 4.23), базовый i_B и коллекторный i_K токи пренебрежимо малы, выходное напряжение $U_{\text{вых}}$ близко к величине \mathcal{E}_K . В момент времени t_1 входное напряжение скачком изменяется от $-|\mathcal{E}_{B2}|$ до $\mathcal{E}_{B1} > 0$ (см. рис. 4.23, a), что приводит к перепаду входного напряжения $\Delta U_{\text{вх}} = |\mathcal{E}_{B1}| + \mathcal{E}_{B2}$. Напряжение $U_{BЭ}$ и ток базы i_B не могут измениться скачком из-за наличия инерционных процессов, связанных прежде всего с наличием паразитной емкости $C_{\text{вх}}$, которая равна сумме параллельно включенных барьерных емкостей $C_{Э}$ и C_K (см. рис. 4.21). Напряжение $|U_{BЭ}|$ нарастает примерно по экспоненциальному закону с постоянной времени $\tau_{\text{вх}} = R_B C_{\text{вх}}$, при этом полагается, что транзистор еще находится в закрытом состоянии, и его сопротивление мало влияет на величину $\tau_{\text{вх}}$. Условно считается, что транзистор находится в закрытом состоянии до тех пор, пока $I_B \leq 0,1 I_{B \text{ нас}}$, а коллекторный ток $I_K \approx 0,1 I_{K \text{ нас}}$ (см. рис. 4.23, b, z).

Если $I_B = 0,1I_{B \text{ нас}}$, то напряжение на базе равно $U_{BЭ} = U_{BЭ \text{ пор}}$. При $U_{BЭ} \geq U_{BЭ \text{ пор}}$ ключ начинает открываться, и токи базы и коллектора резко возрастают, т. е. происходит процесс включения, при этом $r_{вх} \ll R_B$. Во включенном состоянии величина базового тока равна $I_{B1} \approx \mathcal{E}_{B1}/R_B$. Таким образом, скачкообразное нарастание базового тока отстает на время задержки $t_{зд} = t_2 - t_1$ относительно подачи входного импульса (см. рис. 4.23, б).

Нарастание коллекторного тока происходит при постоянном токе базы I_{B1} , при этом в принципе необходимо учитывать время пролета носителей в базе, т. е. скачок коллекторного тока должен быть сдвинут относительно скачка базового тока на величину этого времени. Однако этим временем можно пренебречь, поскольку оно, как правило, много меньше времени задержки $t_{зд}$. Коллекторный ток нарастает по экспоненциальному закону с постоянной времени, равной времени жизни неосновных носителей в базе с учетом влияния емкости коллекторной цепи, которая перезарядается через резистор R_K , при этом сопротивлением коллекторного слоя r'_K можно пренебречь, поскольку обычно $R_K \gg r'_K$. Напряжение на эмиттерном переходе во время нарастания тока $r = i_K(t)$ изменяется мало, поэтому влиянием барьерной емкости эмиттерного перехода обычно пренебрегают. Однако следует учитывать влияние выходной емкости C_H , так как напряжение на ней изменяется на значительную величину. В момент времени t_3 коллекторный ток равен $0,9I_{K \text{ нас}}$ (см. рис. 4.23, з), а выходное напряжение (см. рис. 4.23, д) снижается до величины, близкой к $U_{KЭ \text{ нас}}$. Время нарастания $t_{нр}$ коллекторного тока и спада напряжения ($t_{сп. U}$) можно вычислить по формуле [1]

$$t_{нр} \approx t_{сп. U} = \tau_{нр} \ln [1 - 0,9I_{K \text{ нас}}/(\beta I_{B1})]^{-1}, \quad (4.50)$$

где $\tau_{нр}$ — постоянная времени нарастания коллекторного тока и спада выходного напряжения, которая определяется временем пролета и временем заряда коллекторной емкости и емкости нагрузки и может быть вычислена по формуле

$$\tau_{нр} = \beta(t_{пр} + C_{K \text{ бар}} R_K) + C_H R'_K. \quad (4.51)$$

Здесь $t_{пр}$ — суммарное время пролета носителей через базу и коллекторный переход, $C_{K \text{ бар}}$ — усредненная барьерная емкость коллекторного перехода.

Формула (4.50) дает хорошую точность в быстродействующих схемах, где емкость нагрузки имеет малую величину.

Если нагрузочная емкость велика, то время спада напряжения на выходе вычисляется по той же формуле (4.50), но оно будет больше времени нарастания тока. Суммарное время $t_{зд} + t_{нр} = t_{вкл}$ есть **время включения**. Для уменьшения величины включения $t_{вкл}$ необходимо уменьшать величины $C_{Э\text{ бар}}, C_{К\text{ бар}}, C_{Н}, t_{пр}$ и увеличивать параметр β .

По окончании процесса включения (момент $t = t_3$, см. рис. 4.23, з) ток коллектора ограничен величиной $I_{К\text{ нас}}$, но в базовой области в этот момент накоплен избыточный заряд, характерный для активного режима, а в коллекторе накапливается заряд электронов, соответствующий режиму насыщения, т. е. в рассматриваемых условиях в коллекторе транзистора по сравнению с активным режимом накапливается избыточный заряд неосновных носителей.

При $t = t_4$ входное напряжение мгновенно изменяется от значения $+\mathcal{E}_{Б1}$ до $-\mathcal{E}_{Б2}$ (см. рис. 4.23), в цепи базы появляется отрицательный ток $I_{Б2} = \mathcal{E}_{Б2}/R_1$ при $\mathcal{E}_{Б2} \gg U_{БЭ}$. Этот ток вызывает понижение напряжения $\Delta U_{БЭ}$ на базе из-за изменения напряжения на сопротивлении базы на величины $\Delta U_{БЭ} = (I_{Б1} - I_{Б2})r'_B$. Большой положительный базовый ток образуется накопленными в базе в режиме насыщения неосновными носителями (дырками). Эмиттерный и коллекторный переходы будут открыты до тех пор, пока избыточные неосновные носители у границ переходов в базе не уменьшатся до нуля. Таким образом, пока избыточные носители у границ переходов в базе существуют, ток коллектора остается равным $I_{К\text{ нас}}$, а выходное напряжение равно $U_{КЭ\text{ нас}}$. Интервал времени между t_4 и моментом, когда напряжение на коллекторе транзистора достигает уровня $U_{КЭ\text{ нас}} + 0,1 \Delta U$, называется **временем рассасывания** (см. рис. 4.23, з, д). К моменту t_5 транзистор переходит из режима насыщения в активный режим, начинается спад коллекторного тока и повышение выходного напряжения.

Время рассасывания может быть вычислено по формуле

$$t_{\text{рас}} = \tau_{\text{рас}} \ln \left\{ (I_{Б1} + I_{Б2}) / [(I_{К\text{ нас}}/\beta) + I_{Б2}] \right\}, \quad (4.52)$$

где постоянная времени рассасывания $\tau_{\text{рас}}$ близка к рассмотренной ранее в п. 4.4 величине $\tau_{Кэф}$, входящей в выражение (4.30).

Следовательно, постоянная времени $t_{\text{рас}}$ определяется в основном временем жизни дырок в коллекторе и электронов в базе. Для снижения этих времен необходимо ускорить рекомбинацию неосновных носителей, что технологически достигается легированием кремния атомами золота, которые создают уровни ловушек (центры рекомбинации) в запрещенной зоне кремния. Увеличение тока базы $I_{\text{Б2}}$ также уменьшает $t_{\text{рас}}$, поскольку при этом удаляются избыточные электроны из базы.

Интервал времени, когда выходной (коллекторный) импульс тока уменьшается от значения 0,9 его амплитуды до величины 0,1 амплитуды называется **временем спада** $t_{\text{сп}} = t_6 - t_5$ (рис. 4.23, з). Оно может быть оценено по формуле

$$t_{\text{сп}} = \tau_{\text{нр}} \ln [1 + I_{\text{К нас}} / (\beta I_{\text{Б2}})]. \quad (4.53)$$

В момент спада ток и напряжение изменяются во времени по закону, близкому к экспоненте с одинаковой постоянной времени, приблизительно равной времени нарастания $\tau_{\text{нр}}$, что и отражено в выражении (4.53).

Суммарное время $t_{\text{рас}} + t_{\text{сп}} = t_{\text{выкл}}$ называется **временем выключения**. При большой емкости $C_{\text{Н}}$ переходной процесс нарастания выходного напряжения $t_{\text{нр}U} = t_7 - t_5$ (штрихпунктирная линия на рис. 4.23, д) определяется только зарядкой нагрузочного конденсатора через резистор $R_{\text{К}}$. Это время, которое может значительно превысить время спада $t_{\text{сп}}$, аппроксимируется выражением $t_{\text{нр}U} \approx 2,3R_{\text{К}}C_{\text{Н}}$.

4.8. Разновидности биполярных транзисторов

Из всех классификационных факторов на практике наиболее часто используют классификацию по максимально допустимой мощности рассеивания и граничной частоте. В зависимости от максимально допустимой мощности рассеивания биполярные транзисторы могут быть: малой мощности ($P_{\text{макс}} \leq 0,3$ Вт), средней мощности ($0,3 < P_{\text{макс}} \leq 1,5$ Вт) и большой мощности ($P_{\text{макс}} > 1,5$ Вт). Транзисторы, имеющие граничную частоту $f_{\text{гр}} \leq 3$ МГц, относятся к низкочастотным, при $3 \text{ МГц} < f_{\text{гр}} \leq 30 \text{ МГц}$ — к транзисторам средней частоты, при $30 \text{ МГц} < f_{\text{гр}} \leq 300 \text{ МГц}$ — к высокочастотным, при $f_{\text{гр}} > 300 \text{ МГц}$ — к сверхвысокочастотным (СВЧ-транзисторам).

Низкочастотные и высокочастотные транзисторы имеют чаще всего эпитаксиально-планарную или планарную кремниевую структуру $n-p-n$ -типа. Они отличаются тем, что высокочастотные транзисторы имеют меньшие площади переходов, меньшие толщины базы и коллектора и времена жизни неосновных носителей. В силу сказанного, для низкочастотных (НЧ) транзисторов характерны емкости переходов $10 \dots 100$ пФ, а для высокочастотных (ВЧ) емкость не превышает 10 пФ, для НЧ транзисторов $t_{\text{рас}} \geq 1$ мкс, а для ВЧ $t_{\text{рас}} \leq 0,1$ мкс.

Наибольшие структурные и конструктивные особенности присущи СВЧ-транзисторам. Для увеличения граничной частоты необходимо уменьшать время пролета носителей от эмиттера до коллектора и емкости транзистора. Чтобы снизить это время, СВЧ-транзисторы изготавливаются на основе $n-p-n$ -структуры, поскольку подвижность электронов в кремнии в три раза больше, чем подвижность дырок, и, кроме того, базу делают по возможности тонкой.

Современный уровень технологии позволяет изготавливать базу толщиной менее $0,1$ мкм. Однако при этом сопротивление базы с уменьшением толщины увеличивается, что приводит к снижению величины рабочих напряжений и, следовательно, мощности. Для уменьшения влияния указанных негативных явлений увеличивают концентрацию примесных атомов в базе, но при этом возрастают емкости переходов, которые можно снизить за счет минимизации размеров областей и других элементов транзистора, включая выводы. В результате барьерные емкости переходов транзистора очень малы, поэтому на граничную частоту влияют паразитные емкости и индуктивности выводов, для уменьшения которых корпуса транзисторов изготавливают с плоскими выводами и, кроме того, часто используют бескорпусные транзисторы. Структура СВЧ-транзистора обычно содержит несколько базовых и эмиттерных областей и соответствующих слоев.

Целым рядом преимуществ по сравнению с кремниевыми СВЧ-транзисторами обладают транзисторы на основе GaAs с гетеропереходами, транзисторы на горячих электронах, особенно транзисторы с металлической базой и транзисторы с проницаемой базой.

Коротко рассмотрим особенности перечисленных приборов. Транзистор с гетеропереходом имеет широкозонный эмиттер n -типа из $\text{Al}_x\text{Ga}_{1-x}\text{As}$, базу p -типа из GaAs и коллектор n -типа из GaAs. Характерной особенностью гетеропереходов, которые образованы полупроводниками с различной шириной запрещен-

ной зоны, является наличие потенциальных скачков (барьеров) как в зоне проводимости, так и в валентной зоне (см. п. 2.7). В результате транзисторы с гетеропереходами имеют следующие преимущества:

- 1) высокую эффективность эмиттера из-за крайне малой инжекции дырок из базы в эмиттер, чему препятствует высокий потенциальный барьер в валентной зоне;
- 2) малое сопротивление базы из-за ее сильного легирования без снижения эффективности эмиттера, что также связано с наличием высоких потенциальных барьеров в области эмиттерного перехода;
- 3) лучшую переходную характеристику по сравнению с обычным транзистором из-за высокого коэффициента усиления по току и низкого сопротивления базы;
- 4) возможность работы при повышенных температурах вплоть до $T^\circ \sim 350^\circ \text{C}$. В настоящее время разработаны транзисторы с параметром $\beta \approx 350$, с граничной частотой, превышающей 10 ГГц.

Одними из перспективных в СВЧ-диапазоне являются *транзисторы на горячих электронах*. Горячими называют электроны, энергия которых превышает энергию Ферми на несколько kT ($k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана, T — температура решетки в градусах К). Горячие, иначе, быстрые электроны, формируются за счет создания больших потенциалов, ускоряющих электроны на границе между соседними областями транзистора. Было предложено и исследовано достаточно много трехэлектродных структур, подобных биполярным транзисторам, с переносом горячих электронов от эмиттера к коллектору. Основное отличие этих приборов от чисто полупроводниковых классических транзисторов заключается в способе инжекции электронов в базу. Наилучшими СВЧ-характеристиками обладает *транзистор с металлической базой* и структурой типа Si—Au—Ge, т. е. структура полупроводник—металл—полупроводник, при этом толщина золотой пленки между двумя полупроводниками составляет несколько десятков ангстрем (сотые доли микрона и меньше).

Применение таких современных перспективных технологий, как молекулярно-лучевая эпитаксия, выращивание монокристаллических металлических пленок на полупроводниках, ультрафиолетовая и рентгеновская литография, работа при низких температурах, позволяет добиться хороших СВЧ-харак-

теристик и высоких коэффициентов усиления по току в транзисторах с металлической базой.

Дальнейшим развитием подобного направления является разработка *транзисторов с проникаемой базой*, в которых металлическая пленка заменена металлической вольфрамовой сеткой с периодом порядка 0,3 мкм. Такой транзистор имеет четырехслойную структуру, включающую подложку из GaAs n^+ -типа, слой эмиттера из GaAs n -типа, фигурную вольфрамовую сетку, толщиной 0,02 мкм с шириной полоски 0,16 мкм и слой коллектора n — GaAs. Вольфрамовая сетка образует с n — GaAs барьер Шоттки высотой 0,8 В. При подаче на сетку отрицательного потенциала электроны из эмиттера при их движении к коллектору должны пройти в окрестности металлической сетки через область с отрицательным потенциалом. Посередине между металлическими полосками барьер наиболее низкий, а около границы металл — полупроводник барьер будет наиболее высоким. Этот барьер будет препятствовать проходу электронов через сетку. В результате только небольшая доля электронов преодолевает поле сетки, и в цепи коллектора протекает ток с малой плотностью (единицы А/см²). Если на базу подать положительное смещение, барьер снизится, и плотность коллекторного тока может достигать значений $\sim 10^3$ А/см². При больших положительных смещениях на базе (сетке) будет накапливаться отрицательный заряд электронов, что приведет к ограничению тока. В этих транзисторах можно получить высокие плотности управляемых токов и, как следствие, большую крутизну и высокую граничную частоту, до нескольких десятков ГГц, а в перспективе до сотни ГГц, при относительно больших значениях коэффициентов усиления (15...20 дБ).

Мощные транзисторы. При разработке мощных транзисторов приходится решать дополнительно ряд специфических проблем, которые вызваны большими напряжениями и токами коллектора. Поэтому конструкция мощного транзистора должна обеспечивать эффективный отвод рассеиваемой в нем тепловой энергии. Перегрев активных частей транзистора большой мощности при значительных размерах применяемых полупроводниковых кристаллов вызывает необходимость учета механических напряжений из-за различия температурных коэффициентов линейного расширения полупроводника и других элементов конструкции. Помимо всего, мощные транзисторы должны быть достаточно быстродействующими.

Для обеспечения большого рабочего тока в мощных транзисторах необходимо изготавливать эмиттер как можно больших размеров, причем сложной конфигурации. Поэтому обычно применяют многоэмиттерные транзисторы, содержащие большое число узких эмиттерных полосок, между которыми располагаются выводы базы. Как эмиттеры, так и отдельные базовые выводы объединяются общими выводами. Число отдельных эмиттерных полосок может быть до нескольких десятков. Площадь каждой эмиттерной полоски обычно значительно больше, чем эмиттера маломощного транзистора. Предельная ширина эмиттерной полоски ограничена эффектами вытеснения тока эмиттера на края перехода, поэтому существуют оптимальные размеры ширины полоски, которые лежат в пределах 10...20 мкм. Длина полоски ограничивается падением напряжения на ней и составляет 100...200 мкм. Последовательное включение с каждым эмиттером стабилизирующих резисторов позволяет выравнивать токи отдельных эмиттеров.

Большая часть мощных транзисторов рассчитана на работу при относительно низких напряжениях в 20...30 В, поскольку в этом случае облегчается тепловой режим. Хороший теплоотвод в мощных транзисторах достигается за счет установки полупроводникового кристалла на массивном металлическом основании корпуса часто совместно со специальным радиатором. Для уменьшения барьерной емкости и теплового сопротивления коллектора используют многоструктурные транзисторы, собранные на одном кристалле в виде матрицы отдельно параллельно соединенных транзисторов. За счет увеличения расстояния между отдельными транзисторами обеспечивается нужное тепловое сопротивление без увеличения емкости коллекторного перехода.



Контрольные вопросы

1. Каковы устройство, схемы включения, режимы работы биполярных транзисторов (БТ)?
2. Физические процессы в БТ. Коэффициенты передачи тока в различных схемах включения.
3. Статические ВАХ в схемах с ОБ и ОЭ.
4. Модель Молла—Эберса. Эквивалентные схемы БТ.
5. Высокочастотные свойства БТ.

6. Параметры БТ. БТ как четырехполюсник.
7. Работа БТ в ключевом режиме.
8. Импульсные параметры БТ.
9. Каковы разновидности и особенности БТ их работы?

Глава 5

ТИРИСТОРЫ

5.1. Общие сведения. Устройство. Режимы работы

Тиристор — это полупроводниковый прибор с тремя и более $p-n$ -переходами, ВАХ которого имеет участок с отрицательным дифференциальным сопротивлением.

При работе в схеме тиристор может находиться в двух состояниях. В одном состоянии (**закрытом**, или **выключенном**) тиристор имеет высокое сопротивление и пропускает малый ток, в другом (**открытом**, или **включенном**) — сопротивление тиристора мало и через него протекает большой ток.

Тиристоры широко применяются в радиолокации, устройствах радиосвязи, в автоматике — как приборы с отрицательной проводимостью, управляемые ключи и вентили, пороговые элементы, преобразователи энергии, триггеры, не потребляющие ток в исходном состоянии. По сравнению с биполярными транзисторами тиристоры могут обеспечить более высокий коэффициент усиления по току включения, иметь большой ток и одновременно высокое напряжение, что важно для получения хороших характеристик устройств, работающих при высоких уровнях мощности. Тиристоры обеспечивают высокий КПД преобразования энергии, обладают хорошей надежностью и долговечностью, имеют малые габариты, просты в эксплуатации.

Устройство тиристоров. В зависимости от числа выводов тиристоры делятся на **диодные**, **триодные** и **тетродные**, имеющие соответственно два, три и четыре вывода от $p-n-p-n$ -структуры. Контакт к внешнему p -слою называется **анодом** (А), а к

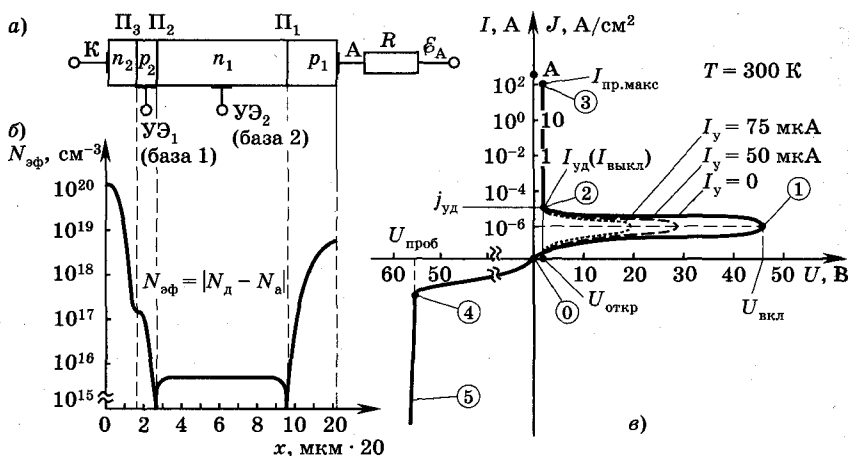


Рис. 5.1

внешнему n -слою — катодом (К) (рис. 5.1, а). Анодная p -область и катодная n -область называются соответственно p - и n -эмиттерами. Области n - и p -типа, расположенные между анодом и катодом, называются базами, а выводы от них образуют управляющие электроды (УЭ). Наиболее часто используются трехэлектродные приборы. Помимо четырехслойных структур, некоторые виды тиристоров имеют большее число полупроводниковых областей. К таким приборам относится симистор (симметричный тиристор), который может включаться при различных полярностях приложенного напряжения. Он сформирован структурой из пяти и более слоев и используется в цепях переменного тока как двусторонний ключ.

Типичное распределение эффективной концентрации атомов примеси в диффузионно-сплавном тиристоре представлено на рис. 5.1, б. На подложке n_1 методом двусторонней диффузии сформированы области p_1 и p_2 . Слой n_2 создается методом сплавления или односторонней диффузии.

Режимы работы. В зависимости от напряжения на аноде и тока, протекающего через прибор, можно выделить несколько режимов работы тиристора. Эти режимы соответствуют определенным участкам ВАХ тиристора, представленной на рис. 5.1, в. В отсутствие тока в цепи УЭ, т. е. при $I_y = 0$, ВАХ вырождается в характеристику диодного тиристора, когда цепи управляющих элект-

тродов отсутствуют или разомкнуты. Чтобы снять такую характеристику, необходимо в качестве источника электрического питания использовать генератор тока с ЭДС \mathcal{E}_A (см. рис. 5.1, а). В этом случае ток в цепи задается источником, и в зависимости от величины тока между катодом и анодом будет возникать соответствующая разность потенциалов. Выделяют пять основных режимов работы тиристора.

Режим 1 (область ВАХ 0—1) — напряжение на аноде положительно относительно катода, ток незначителен (несколько мкА). Эта область соответствует закрытому состоянию (режим прямого запираания).

Режим 2 (область 1—2) — участок характеристики с отрицательным дифференциальным сопротивлением. Начинается в точке ВАХ, когда $dU/dI = 0$. Напряжения на тиристоре в этой точке называется **напряжением включения** ($U_{\text{вкл}}$), а ток через прибор — **током включения** ($I_{\text{вкл}}$).

Режим 3 (область 2—3) соответствует открытому состоянию (режим прямой проводимости), начинается в точке 2. Напряжение в этой точке называется **напряжением в открытом состоянии** ($U_{\text{откр}}$), или (реже) **напряжением удержания** ($U_{\text{уд}}$), а ток — **током удержания** ($I_{\text{уд}}$). Параметры $I_{\text{уд}}$ и $U_{\text{откр}}$ — соответственно минимальные ток и напряжение, необходимые для поддержания тиристора в открытом состоянии.

Режим 4 (область 0—4) называется режимом обратного запираания, в котором напряжение на аноде относительно катода отрицательно.

Режим 5 (область 4—5) — режим обратного пробоя. Начинается при напряжении на аноде, равном напряжению пробоя тиристора ($U_{\text{проб}}$).

5.2. Основные физические процессы. Принцип действия

Рассмотрим основные процессы на примере диодного тиристора, ВАХ которого соответствует рис. 5.1, в при $I_y = 0$.

В режиме **обратного запираания** переходы Π_1 и Π_3 (см. рис. 5.1, а) смещены в обратном направлении, а переход Π_2 — в прямом. Если распределение примесей в областях тиристора соответствует рис. 5.1, б, то переход Π_1 будет значительно толще, чем Π_3 (см. п. 2.2), и большая часть напряжения, создаваемого на тири-

торе протекающим током I , падает на переходе Π_1 . В зависимости от размеров области n_1 и напряжения на ней толщина обедненного слоя занимает либо часть, либо всю область n_1 . В последнем случае происходит смыкание переходов Π_1 и Π_2 . Ток на участке 0—4 ВАХ определяется сопротивлениями обратносмещенных переходов Π_1 и Π_3 . Допустимое падение напряжения на тиристоре ограничивается ударной ионизацией в переходе Π_1 и смыканием базы n_1 . Напряжение пробоя в тиристоре ниже напряжения пробоя обратносмещенного изолированного эквивалентного $p-n$ -перехода Π_1 . Это обусловлено тем, что в тиристоре переход Π_1 связан с соседним переходом Π_2 и образует транзистор $p_1-n_1-p_2$ с разомкнутой базой, включенный по схеме с общим эмиттером (ОЭ). Напряжение пробоя ($U_{\text{проб}}$) уменьшается из-за влияния этого транзистора (см. п. 4.3).

В режиме прямого запираия напряжение на аноде положительно по отношению к катоду, переходы Π_1 и Π_3 смещены в прямом направлении, а переход Π_2 — в обратном. Падение напряжения между анодом и катодом тиристора равно сумме падений напряжения на переходах, т. е. $U_A = U_1 + U_2 + U_3$. Большая часть приложенного напряжения падает на переходе Π_2 и лишь незначительная — на переходах Π_1 и Π_3 . Для понимания характеристик в рассматриваемом режиме воспользуемся двухтранзисторной моделью тиристора, в которой тиристор рассматривается как соединение $p-n-p$ и $n-p-n$ -транзисторов. Коллектор каждого из этих транзисторов соединен с базой другого (рис. 5.2). Здесь

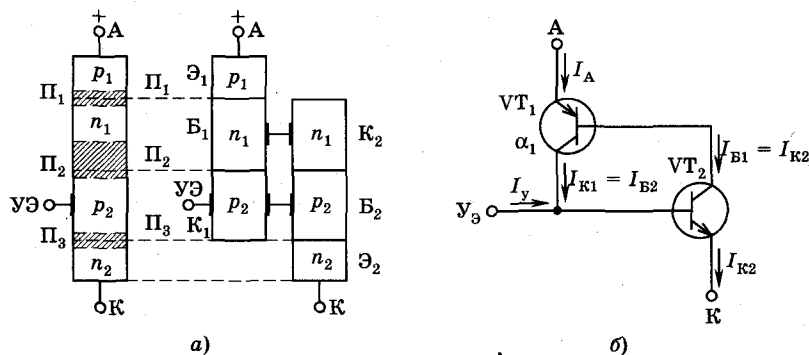


Рис. 5.2

представлена двухтранзисторная модель для триодного тиристора (рис. 5.2, а — структурная схема, рис. 5.2, б — схемное включение). Переход Π_2 является коллектором для дырок и электронов, инжектируемых соответственно из области p_1 и n_2 . Полный ток, протекающий через переход Π_2 , определяется токами инжекции этих носителей и собственным обратным током (см. п. 2.3).

При разомкнутой цепи управляющего электрода токи инжекции через переходы Π_1 и Π_3 в основном ограничены малым прямым напряжением на них из-за большого сопротивления обратносмещенного перехода Π_2 . Пользуясь соотношениями п. 4.2, можно составить выражение для тока базы транзистора VT_1 (см. рис. 5.2, б):

$$I_{B1} = (1 - \alpha_1)I_A - I_{K01}. \quad (5.1)$$

Здесь α_1 — статический коэффициент передачи тока эмиттера транзистора VT_1 ; I_{K01} — обратный ток перехода коллектор — база транзистора VT_1 . Ток I_{B1} протекает через коллектор транзистора VT_2 (см. рис. 5.2, б) и $I_{B1} = I_{K2}$.

С другой стороны, коллекторный ток I_{K2} транзистора VT_2 можно определить через ток катода I_K из известного выражения (см. п. 4.2):

$$I_{K2} = \alpha_2 I_K + I_{K02}, \quad (5.2)$$

где α_2 — статический коэффициент передачи тока эмиттера транзистора VT_2 ; I_{K02} — обратный ток коллекторного перехода транзистора VT_2 . Если в цепи управляющего электрода протекает ток I_y , то ток катода равен $I_K = I_A + I_y$. Учитывая это равенство и приравняв выражения (5.1) и (5.2), получим

$$I_A = \frac{\alpha_2 I_y + I_{K0}}{1 - (\alpha_1 + \alpha_2)}, \quad (5.3)$$

где $I_{K0} = I_{K01} + I_{K02}$.

Если напряженность электрического поля в переходе Π_2 достаточна для размножения носителей за счет ударной ионизации (см. п. 2.3), а $I_y = 0$, то соотношение (5.3) можно записать в виде

$$I_A = I_{K0} / [1 - M(\alpha_1 + \alpha_2)], \quad (5.4)$$

где M — коэффициент размножения носителей.

Поскольку $M = f(U)$ и $\alpha_{1,2} = f_{1,2}(U)$, то формулы (5.3) и (5.4) описывают статическую ВАХ тиристора при напряжениях на аноде, меньших или равных $U_{\text{вкл}}$. Коэффициенты α_1 и α_2 сильно зависят от величины тока. На участке 0—1 ВАХ (см. рис. 5.1, в), пока ток и напряжение анода невелики, $(\alpha_1 + \alpha_2) < 1$ и анодный ток определяется током $I_{\text{к0}}$. С повышением напряжения на аноде возрастает прямое напряжение на эмиттерных переходах Π_1 и Π_3 , что приводит к увеличению инжекции через эти переходы и росту α_1 и α_2 . Дырки, инжектированные из области p_1 , проходят через переход Π_2 и повышают потенциал базы p_2 за счет неравновесного положительного заряда, что увеличивает инжекцию электронов из области n_2 . Эти электроны, попадая в базу n_1 , снижают ее потенциал, увеличивая тем самым инжекцию из области p_1 . В результате в тиристоре возникает положительная обратная связь, приводящая к лавинообразному увеличению анодного тока при некотором напряжении $U = U_{\text{вкл}}$.

При $U = U_{\text{вкл}}$ выполняется соотношение $(\alpha_1 + \alpha_2) \rightarrow 1$. При лавинообразном нарастании анодного тока через переход Π_2 протекает значительный ток, и его сопротивление резко падает. Напряжение источника питания перераспределяется таким образом, что падение напряжения на резисторе в анодной цепи (см. рис. 5.1, а) возрастает, а напряжение на аноде тиристора уменьшается в соответствии с соотношением $U_A = \mathcal{E}_A - IR$. Тиристор переходит в режим, соответствующий участку 1—2 ВАХ (см. рис. 5.1, в), который в области точки 2 является неустойчивым. Переход из этого режима в режим прямой проводимости (участок 2—3 ВАХ) происходит, как правило, скачком. На участке 2—3 ВАХ переход Π_2 из-за неравновесного заряда, накопленного в базах, открывается и переходит в режим насыщения. В этом режиме проводимость тиристора велика, а падение напряжения на нем мало. Оно определяется суммой падений напряжений на трех прямосмещенных p — n -переходах Π_1 , Π_2 , Π_3 , падением напряжения на базах, p_1 - и n_2 -областях и выводах (см. рис. 5.1, а). Суммарное падение напряжения на включенном тиристоре составляет 1...2 В.

Рассмотрим особенности включения тиристора при наличии размножения носителей в переходе Π_2 . Если предельные величины α_1 и α_2 малы, напряжение на переходе Π_2 может достигнуть величины $U_{\text{проб}}$. В этом случае носители заряда, переме-

щающиеся через переход Π_2 , приобретут энергию, достаточную для лавинного умножения. Дырки, рожденные в переходе, добавляются к дыркам, инжектированным через переход Π_1 , и перемещаются к переходу Π_3 , электрическое поле которого является ускоряющим для них. Электроны двигаются в противоположном направлении, т. е. к аноду. В результате рождения в переходе Π_2 пар зарядов проводимость его увеличивается, а сопротивление падает, что приводит к уменьшению падения напряжения U_2 на нем и увеличению напряжения на переходах Π_1 и Π_3 . Это, в свою очередь, увеличивает инжекцию дырок и электронов из областей p_1 и n_2 , т. е. размножение носителей в переходе Π_2 идет еще более интенсивно, и далее процесс повторяется. Таким образом, число носителей, перемещающихся через прибор, лавинообразно увеличивается, коэффициенты α_1 и α_2 растут, произведение $M(\alpha_1 + \alpha_2)$ в (5.4) приближается к единице и происходит включение тиристора.

Для коэффициента умножения M в зависимости от приложенного напряжения U и напряжения пробоя $U_{\text{проб}}$ перехода Π_2 можно воспользоваться эмпирическим соотношением вида $M = [1 - (U/U_{\text{проб}})^m]^{-1}$, где принято $U = U_{\text{вкл}}$. Тогда $M(\alpha_1 + \alpha_2) = (\alpha_1 + \alpha_2)[1 - (U_{\text{вкл}}/U_{\text{проб}})^m]^{-1} = 1$. Напряжение включения $U_{\text{вкл}}$, вычисленное из последнего выражения, определяется следующей формулой:

$$U_{\text{вкл}} = U_{\text{проб}}(1 - \alpha_1 - \alpha_2)^{1/m}, \quad (5.5)$$

где m — некоторое число, зависящее от параметров и устройства тиристора.

Из (5.5) видно, что $U_{\text{вкл}}$ всегда меньше напряжения пробоя перехода Π_2 , что обусловлено наличием положительной обратной связи в тиристорной структуре. При $\alpha_1 = 0,4$, $\alpha_2 = 0,15$, $U_{\text{проб}} = 66$ В и $m = 4$ напряжение включения $U_{\text{вкл}} = 45$ В.

После того как тиристор включился, все три перехода смещены в прямом направлении, т. е. оба эквивалентных транзистора находятся в режиме насыщения. Участок 2—3 ВАХ определяется свойствами трех последовательно соединенных и прямосмещенных p — n -переходов.

Рассмотрим теперь механизм переключения триодного тиристора при подаче прямого смещения на управляющий элект-

род УЭ₁ (см. рис. 5.1, а), когда напряжение на переходе П₂ меньше $U_{\text{проб}}$. Если в цепи управляющего электрода протекает ток I_y , будет происходить увеличение инжекции через переход П₃, рост α_2 и т. д., т. е. при меньшем анодном напряжении ток анода достигнет значения, при котором $\alpha_1 + \alpha_2 \geq 1$, и тиристор включается. Таким образом, тиристор представляет собой управляемый ключевой прибор (см. рис. 5.1, в).

При работе тиристора в цепях переменного тока и напряжения необходимо учитывать динамические процессы. Рассмотрим особенности включения тиристора в режиме малых сигналов. Будем считать, что переменный сигнал подается в цепь управляющего электрода.

Дифференцируя соотношение (5.3) и используя вместо статических коэффициентов α_1 и α_2 дифференциальные коэффициенты передачи тока $\tilde{\alpha}_1$ и $\tilde{\alpha}_2$, в режиме малого сигнала ($\tilde{\alpha} = h_{21B}$, см. п. 4.4) получим следующее равенство:

$$dI_A/dI_y = \tilde{\alpha}_2/[1 - (\tilde{\alpha}_1 + \tilde{\alpha}_2)]. \quad (5.6)$$

Из этого выражения следует, что условие включения тиристора $dI_A/dI_y \rightarrow \infty$ выполняется, если сумма малосигнальных (дифференциальных) коэффициентов $\tilde{\alpha}$ стремится к единице, т. е. $\tilde{\alpha}_1 + \tilde{\alpha}_2 = 1$. Коэффициенты $\tilde{\alpha}_1$ и $\tilde{\alpha}_2$ зависят от токов и напряжений на переходах (см. п. 4.4), например, согласно выражению (4.23) $\tilde{\alpha}_2 = \alpha_2 + I_1 \left(\frac{d\alpha_2}{dI_1} \right)$, т. е. $\tilde{\alpha}_2 = f(I_1)$. Вид функции α_2 показан на рис. 5.3, б.

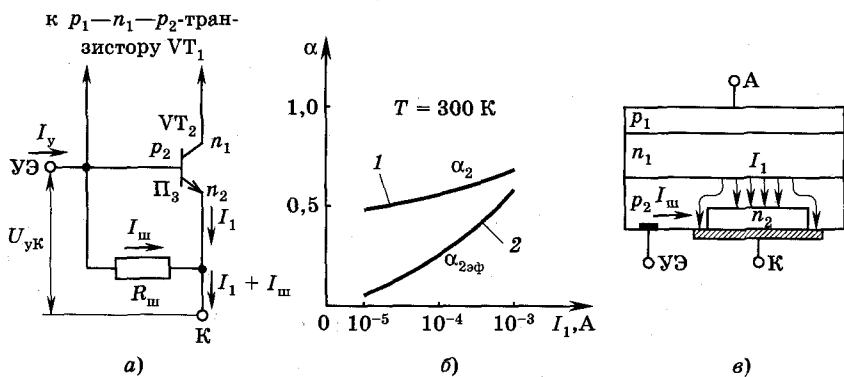


Рис. 5.3

Подача положительного сигнала $U_{yк}$ на УЭ (рис. 5.3, а) будет вызывать увеличение α_2 за счет тока I_y , так что сумма $\alpha_1 + \alpha_2 \rightarrow 1$.

Наиболее простым способом увеличения дифференциального коэффициента $\tilde{\alpha}_2 = \alpha_2 + I_1 \left(\frac{d\alpha_2}{dI_1} \right)$ при нарастании тока в цепи управляющего электрода является шунтирование перехода Π_3 посредством добавочного резистора $R_{ш}$ (см. рис. 5.3, а). При малых напряжениях $U_{yк}$ ток I_y в основном протекает через резистивный шунт $R_{ш}$, минуя базу, и статический коэффициент α_2 уменьшается. По мере увеличения управляющего напряжения $U_{yк}$ растет и доля тока, втекающего в базу VT_2 , инжекция носителей через переход Π_3 становится больше, ток I_1 увеличивается, и коэффициент α_2 возрастает (см. рис. 5.3, б). Для транзистора $n_2-p_2-n_1$ при наличии шунта (см. рис. 5.3, а) коэффициент передачи тока $\alpha_{2эф} = \alpha_2 I_1 / (I_1 + I_{ш})$. Ток через переход Π_3 $I_1 \approx I_0 \exp(U_{yк}/\varphi_T)$, а через шунт $I_{ш} = U_{yк}/R_{ш}$. Если допустить, что α_2 зависит от I_1 согласно кривой 1 на рис. 5.3, б, то зависимость $\alpha_{2эф}$ от I_1 будет соответствовать кривой 2 (для $R_{ш} = 5 \cdot 10^3$ Ом). В результате $\alpha_{2эф}$ увеличивается за счет одновременного возрастания α_2 и произведения $I_1 \left(\frac{d\alpha_2}{dI_1} \right)$, что и вызывает переключение тиристора. На практике шунт формируется за счет частичного перекрытия катодным контактом области p_2 (рис. 5.3, в). Такой тиристор называется **тиристором с закороченным катодом**. Следовательно, включением тиристора можно управлять, изменяя напряжение $U_{yк}$ (ток I_y) (см. рис. 5.1, в).

5.3. Переходные процессы и импульсные свойства тиристорov

Включение тиристорov осуществляется в основном либо с помощью управляющего электрода (рис. 5.4, а), либо изменением анодного напряжения (рис. 5.4, б). Для триодных тиристорov более распространен первый способ. Время включения $t_{вкл}$ складывается из двух составляющих: времени задержки $t_{зд}$ и времени нарастания $t_{нр}$; $t_{зд}$ — это время, отсчитываемое от начала действия управляющего импульса t_1 до момента, при кото-

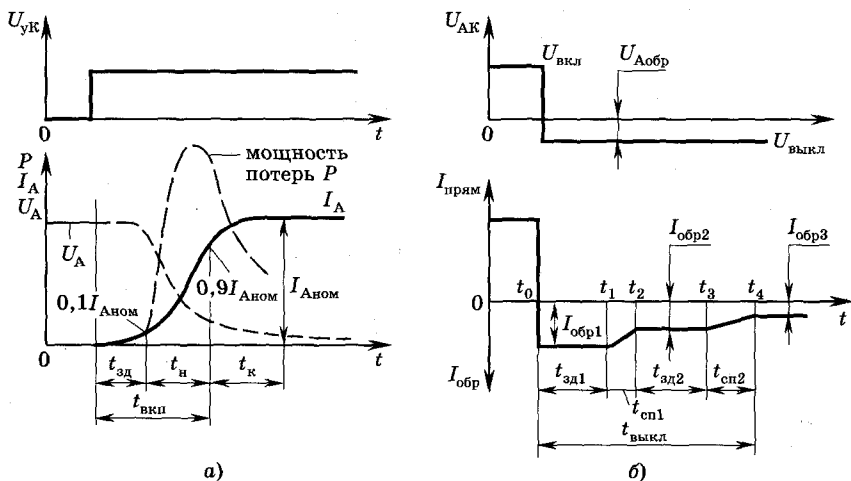


Рис. 5.4

ром анодный ток равен $0,1I_{A \text{ ном}}$ ($I_{A \text{ ном}}$ — конечное значение анодного тока, определяемое полным сопротивлением цепи) (см. рис. 5.4, а); $t_{нр}$ — время нарастания тока от $0,1I_{A \text{ ном}}$ до $0,9I_{A \text{ ном}}$. Временной интервал изменения тока от $0,9I_{A \text{ ном}}$ до $I_{A \text{ ном}}$ обозначается $t_к$.

Время задержки и нарастания для тиристора с резистивной нагрузкой определяется следующими процессами. При подаче положительного импульса на УЭ возникает инжекция электронов из катодной области n_2 в базу p_2 (рис. 5.5, а), причем в начальный момент эта инжекция происходит в непосредственной близости от контакта управляющего электрода. Пройдя базу p_2 , электроны втягиваются электрическим полем перехода Π_2 и выбрасываются в n_1 -базу, сообщая ей отрицательный заряд и

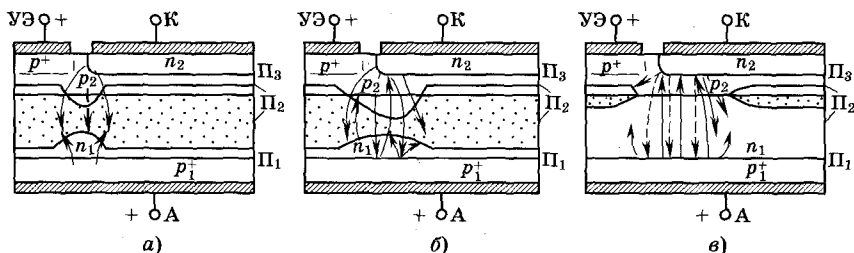


Рис. 5.5

тем самым увеличивая инжекцию дырок из области p_1 . Дырки из p_1 пересекают широкую базу n_1 , переход Π_2 , базу p_2 и, достигая перехода Π_3 , будут увеличивать инжекцию электронов (рис. 5.5, б). Это продолжается до тех пор, пока ток I_A не достигнет значения тока удержания $I_{уд}$ (см. рис. 5.1, в), после чего прибор включается. Поскольку этот процесс является циклически нарастающим, инжекция носителей увеличивается с обоих эмиттеров, что в конечном счете приводит к образованию шнура высокопроводящей электронно-дырочной плазмы в небольшой области катода вблизи управляющего электрода (рис. 5.5, в). В этих условиях проводимость этой области растёт, даже если выключить импульс тока в цепи УЭ.

Время задержки определяется явлениями, происходящими в течение первых двух стадий рассмотренных процессов (см. рис. 5.5, а и б). Как показывает анализ процессов включения, в начальной стадии, на этапе задержки, основную роль играет транзистор $n_2-p_2-n_1$. Транзистор с широкой базой $p_1-n_1-p_2$ при малых плотностях тока не вносит вклада в процесс включения до тех пор, пока анодный ток не достигает 10% его конечного значения. В этих условиях время задержки с учетом перезарядки барьерной емкости катодного перехода вычисляется по формуле

$$t_{зд} = \tau_2 \left[\frac{\alpha_2}{1 - \alpha_2} \ln \left| 1 - \frac{K(1 - \alpha_2)}{10\alpha_2} \right| \right]^{-1} + \int_0^{U_y} \frac{C(U)}{I_y} dU, \quad (5.7)$$

где $\tau_2 = W_{p2}^2 / (2D_n)$ — время пролета электронов через базу p_2 с шириной W_{p2} , D_n — коэффициент диффузии электронов; $K = \frac{I_A}{I_y}$; $C(U)$ — удельная барьерная емкость перехода Π_3 (см. рис. 5.1, а); U_y, I_y — напряжение и ток на управляющем электроде, $U_y = U_{yк}$.

Формула (5.7) показывает, что время задержки уменьшается с увеличением тока I_y и уменьшением времени τ_2 более чувствительного (по сравнению с $p_1-n_1-p_2$) транзистора $n_2-p_2-n_1$ при подаче на его базу импульса тока управления (см. рис. 5.2).

Второй член в (5.7), как правило, играет значительно меньшую роль, чем первый. Если взять структуру, показанную на рис. 5.1, а, которая является поперечным разрезом тиристора

на рис. 5.5, б, то при $W_{p2} \approx 2 \cdot 10^{-3}$ см, $D_n = 22$ см²/с (Si), $\alpha_2 \approx 0,7$; $I_A = 10^{-3}$ А (см. рис. 5.3, б), $K = \frac{I_A}{I_y} = 20$, то суммарное время задержки, согласно (5.7), составит $t_{зд} = 0,05 \cdot 10^{-6}$ с.

Время нарастания $t_{нр}$ определяется как время, отсчитываемое с момента, когда I_A достигает значения $0,1I_{A\text{ном}}$ и когда дальнейший рост I_A обусловлен процессами формирования проводящего шнура и началом его бокового распространения (см. рис. 5.5, б).

Время включения $t_{вкл} = t_{зд} + t_{нр}$, как показывает анализ, приближенно равно среднему геометрическому времени диффузии в n_1 - и p_2 -областях или

$$t_{вкл} = \sqrt{\tau_1 \tau_2}, \quad (5.8)$$

где $\tau_1 = W_{n1}^2 / 2D_n$, $\tau_2 = W_{p2}^2 / 2D_p$, W_{n1} , W_{p2} — соответственно толщины баз n_1 и p_2 .

Для $W_{n1} = 140$ мкм, $W_{p2} = 20$ мкм, $D_n = 22$ см²/с, $D_p = 12$ см²/с время включения равно $t_{вкл} \approx 0,9$ мкс.

В мощных тиристорах эмиттеры p_1 и n_2 имеют большую площадь. Эмиттеры малых размеров используются только в быстродействующих приборах.

Как было показано, в течение времени $t_{вкл}$ начинает проводить только небольшая область вблизи управляющего контакта. Для мощных тиристоров эта область высокой проводимости за счет диффузии и боковых электрических полей постепенно распространяется вдоль катода с некоторой скоростью, и за время t_x (см. рис. 5.4, а) вся поверхность эмиттера n_2 будет инжектировать электроны.

Если допустить, что анодный ток и напряжение при включении тиристора изменяются во времени в соответствии с некоторой функцией, то мгновенно рассеиваемая мощность в проводящей части тиристора (см. рис. 5.4, а) может быть записана в форме $P \approx U_A f(t, r) dI_A/dt$, где U_A — стационарное напряжение между анодом и катодом; $f(t, r)$ — функция времени t и радиуса r управляющего электрода (УЭ). Повышение температуры (ΔT) из-за рассеяния мощности P в наиболее горячей точке равно

$$\Delta T = \left(\frac{1}{\rho c_{уд}} \int_0^\infty P dt \right) \sim (dI_A/dt), \text{ где } \rho, c_{уд} \text{ — плотность и удель-}$$

ная теплоемкость полупроводника, из которого сделан тиристор (обычно Si).

Следовательно, для предотвращения перегрева необходимо не превышать предельную скорость нарастания тока («эффект dI_A/dt »). Предельное значение dI_A/dt зависит от размера первоначально включенной области и скорости распространения включенного состояния вдоль катода. Следовательно, для уменьшения dI_A/dt необходимо увеличивать площадь первоначального включения, что осуществляется за счет выполнения управляющего электрода в виде гребенки; использовать включение, индуцированное полем (эффект dU/dt см. ниже); применять специальный инжектирующий управляющий электрод и т. д.

Рассмотрим некоторые особенности включения тиристора по анодной цепи путем подачи импульса напряжения (рис. 5.4, б). Быстро изменяющееся во времени напряжение dU/dt вызывает ток смещения $I_c = \frac{d(CU)}{dt}$, протекающий через барьерную емкость C перехода Π_2 . Ток смещения I_c выполняет роль управляющего тока I_y , подаваемого в базу p_2 . Как показывает анализ, условие включения в этом случае будет таким же, как и при подаче тока (напряжения) управления, т. е. сумма $\tilde{\alpha}_1 + \tilde{\alpha}_2$ должна быть равна единице. При больших I_c коэффициенты $\tilde{\alpha}_1$ и $\tilde{\alpha}_2$, зависящие от тока, достигают величин, достаточных для включения тиристора. Рассмотренное явление в литературе часто называют «эффектом dU/dt ». За счет этого эффекта в динамическом режиме напряжение включения может быть существенно уменьшено. Уменьшение напряжения включения зависит как от амплитуды импульса анодного напряжения, так и от скорости его нарастания. В мощных тиристорах напряжение включения должно быть большим, поэтому принимают меры по ослаблению эффекта dU/dt . Наиболее радикальным способом ослабления эффекта dU/dt является использование шунтирования перехода Π_3 (которое было рассмотрено выше) с помощью добавочного резистора. В этом случае через переход Π_3 может протекать лишь незначительная часть тока смещения перехода Π_2 и его влияние на коэффициент $\alpha_2 n_2 - p_2 - n_1$ -транзистора ослабляется. В тиристорах с шунтированным переходом Π_3 устойчивость к эффекту dU/dt увеличивается в 100 раз и более (с 20 до 10^4 В/мкс).

Рассмотрим теперь процесс выключения тиристора, находящегося в проводящем состоянии. Выключение тиристора можно осуществить путем подачи запирающего импульса напряжения U_{AK} на анод (см. рис. 5.4, б), разрывом цепи управляющего электрода или же подачей на него обратного смещения. Наиболее быстрое выключение достигается при одновременном изменении направления анодного и базового токов. Рассмотрим случай, когда меняется направление только анодного тока. Выключение тиристора осуществляется путем изменения полярности анодного напряжения. Переходы Π_1 и Π_3 из прямосмещенных при перемене полярности анодного напряжения становятся обратносмещенными. После изменения полярности анодного напряжения значение тока $I_{обр}$ в течение времени $t_0 - t_1$ (см. рис. 5.4, б), пренебрегая падением напряжения на включенном тиристоре, можно выразить следующим образом: $I_{обр} = U_{A\text{обр}}/R$ ($U_{A\text{обр}}$ — амплитуда импульса обратного анодного напряжения; R — сопротивление цепи при омической нагрузке).

В течение некоторого времени, называемого *временем задержки* $t_{зд1}$, обратный ток $I_{обр}$ не изменяется. Это обусловлено тем, что в p_2 -базе во время пребывания тиристора во включенном состоянии накоплен большой заряд избыточных носителей, который и определяет ток, текущий через прибор. За время $t_{зд1}$ происходит рассасывание накопленного заряда, величина которого пропорциональна прямому току, причем градиент неравновесных носителей у переходов Π_1 и Π_3 все это время не меняется, хотя число избыточных носителей в p_2 - и n_1 -базах уменьшается под воздействием положительного потенциала, приложенного к катоду (явления аналогичны процессам в транзисторном ключе, см. п. 4.7). В конце интервала $t_{зд1}$ избыточные (неравновесные) электроны в основном выведены из p_2 -базы. У перехода Π_3 в течение времени $t_2 - t_1 = t_{сп1}$ формируется область объемного заряда, расширяющаяся под действием обратного смещения, что вызывает увеличение сопротивления перехода Π_3 и уменьшение тока $I_{обр}$. В n_1 -базе, по мере того как отрицательный анод захватывает дырки, их количество пополняется за счет инжекции через переход Π_2 и p_2 -базу. Эти процессы происходят в интервале $t_{зд2}$. В соответствии с распределением примеси, приведенным на рис. 5.1, б, инжекция дырок из p_2 -базы преобладает

над инжекцией электронов из n_1 -базы. Задержка $t_{зд2}$ определяется временем жизни неосновных носителей (τ_p) и временем их рассасывания в n_1 -базе. При этом градиент концентрации неосновных носителей в интервале $t_2 - t_3$ остается постоянным и ток $I_{обр2}$ не изменяется. В течение фазы $t_{сп2}$ удаляются оставшиеся неосновные носители в основном из широкой n_1 -области и происходит восстановление закрытого состояния тиристора (Π_1 и Π_3 смещены в обратном направлении).

Следовательно, на основании сказанного можно заключить, что процесс выключения начинается с восстановления запирающих свойств перехода Π_3 . Полное время задержки ($t_{зд1} + t_{зд2}$) зависит от эффективного времени жизни неравновесных носителей в p_2 - и n_1 -базах, от значений прямого и обратного токов, протекающих через тиристор перед моментом подачи запирающего импульса и сразу после подачи $U_{Аобр}$. При $t \geq t_4$ ток через тиристор ($I_{обр3}$) меньше тока удержания $I_{уд}$ (см. рис. 5.1, в) и тиристор закрыт. Время выключения представляется суммой $t_{выкл} = t_{зд1} + t_{зд2} + t_{сп1} + t_{сп2}$.

5.4. Разновидности тиристоров. Параметры и модели тиристоров

Имеется много разновидностей тиристоров. Коротко рассмотрим наиболее распространенные из них.

Симисторы — это переключающие приборы, которые могут работать как при положительном, так и при отрицательном напряжении, т. е. при работе в цепях переменного тока они включаются в положительные и отрицательные полупериоды. На рис. 5.6 изображена типичная структура симистора. Принцип работы такой структуры заключается в следующем. Предположим, что электрод T_2 имеет отрицательный потенциал по отношению к заземленному электроду T_1 , а на управляющий электрод УЭ подан положительный импульс, вызывающий инжекцию электронов из n_1 в p_1 . Эти инжектированные электроны перемещаются в область n_2 и понижа-

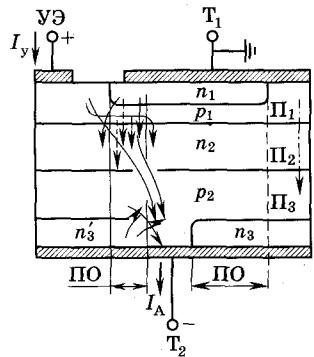


Рис. 5.6

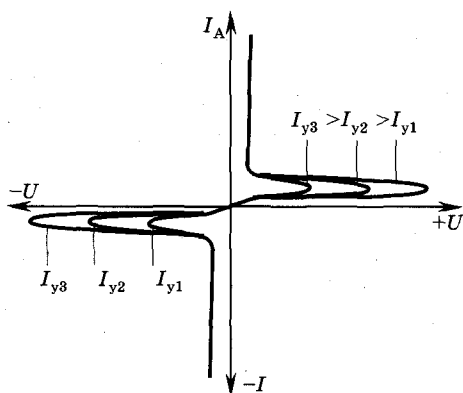


Рис. 5.7

ют ее потенциал, тем самым вызывая инжекцию дырок из p_1 в n_2 , которые за счет диффузии перемещаются к переходу Π_3 и области p_2 . В результате этого процесса возникает боковое напряжение смещения в перекрытой области (ПО) — части вспомогательной области n'_3 , перекрываемой основной областью n_1 (см. рис. 5.6). В результате бокового поля, вызванного дырками, начинается дополнительная инжекция электронов из перекрытой части области n'_3 . Этот процесс развивается таким образом, что проводящая область (включенная область) распространяется вправо, вызывая инжекцию электронов из перекрытой части ПО области n_3 . Таким образом, эквивалентный тиристор, образованный структурой $p_1-n_2-p_2-n_3$ с переходами Π_1 , Π_2 и Π_3 , включается и ток через прибор протекает в направлении от T_1 к T_2 .

В другую половину периода переменного тока контакт T_2 положителен по отношению к T_1 . Подача положительного импульса на управляющий электрод позволяет включить эквивалентный тиристор $p_2-n_2-p_1-n_1$. В этом случае ток течет в направлении от T_2 к T_1 . Вольт-амперная характеристика симистора представлена на рис. 5.7.

Имеется достаточно большое количество структурных модификаций симисторов. На рис. 5.8 представлена одна из них — семислойная структура с центральным расположением электродов, обладающая хорошими номинальными характеристиками.

В мощной технике широко распространены *тиристоры с обратной проводимостью*, которые представляют собой многослойные

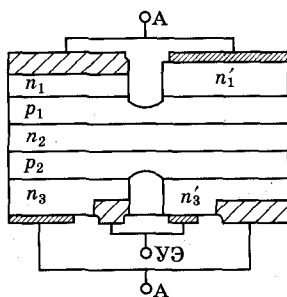


Рис. 5.8

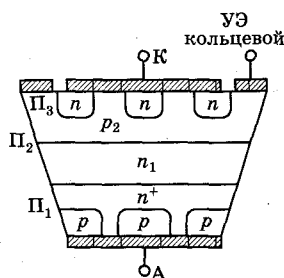


Рис. 5.9

приборы с тремя выводами. Тиристоры с обратной проводимостью отличаются от обычных приборов тем, что они имеют закороченные катод (см. п. 5.2, рис. 5.3, в) и анод, т. е. в них шунтированные участки находятся не только на катоде, но и на аноде (рис. 5.9). Поэтому через эти приборы при обратном смещении протекают большие обратные токи даже при малых напряжениях. При прямом смещении ВАХ имеет тот же вид, что и для обычных тиристоров. Мощные кремниевые тиристоры с обратной проводимостью могут работать при температурах переходов 150...200 °С, в отличие от обычных тиристоров, работающих до 125 °С. Большой температурный диапазон в тиристорах с обратной проводимостью позволяет пропускать через них большие токи в открытом состоянии. Этот тип приборов обладает малым временем восстановления, так как они по сравнению с обычными тиристорами допускают более высокий уровень легирования золотом или платиной.

Двухоперационный (выключаемый) тиристор представляет собой четырехслойную структуру, которая включается подачей положительного напряжения на управляющий электрод, а выключается подачей на этот электрод отрицательного импульса. (Обычный тиристор, как правило, выключается снижением анодного тока ниже тока удержания или за счет включения анодного тока противоположного направления.) Эти тиристоры широко применяются в импульсных генераторах и высокоскоростных мощных устройствах, так как выдерживают большие напряжения в закрытом состоянии. Конструкция их подобна конструкции обычных тиристоров.

Полевые тиристоры. Устройство прибора этого типа представлено на рис. 5.10. Полевые тиристоры состоят из $p^+ - n - p^+$ -ди-

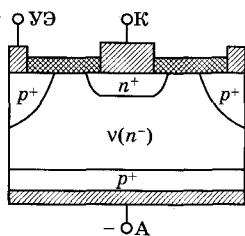


Рис. 5.10

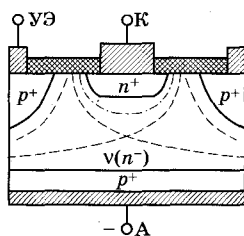


Рис. 5.11

ода, выводами которого являются анод и катод прибора, и управляющего электрода; $v(n^-)$ — это область с незначительной концентрацией донорной примеси, обладающая проводимостью, близкой к собственной проводимости полупроводника (см. п. 1.1). Конструкция УЭ аналогична той, что используется в затворах мощных полевых транзисторов с управляющим $p-n$ -переходом (см. гл. 6). Если на анод подано положительное по отношению к катоду напряжение и управляющий переход открыт, то дырки со стороны анода, а электроны со стороны катода инжектируются в v -базу, снижая ее сопротивление. Прибор открыт. Напряжение между катодом и анодом мало, типичное прямое напряжение на приборе составляет $1...1,5$ В.

Если управляющий p^+-v -переход находится под обратным напряжением, то дырки, идущие от анода, будут эффективно отводиться в цепь УЭ. При достаточно больших обратных напряжениях на управляющем электроде обедненный слой перехода может полностью перекрыть v -область (рис. 5.11). В результате возникает потенциальный барьер для электронов, идущих от катода, и, следовательно, этот барьер препятствует инжекции электронов из катода. Прибор обладает большим сопротивлением и находится в закрытом состоянии. Максимальное напряжение между анодом и катодом в этом состоянии возрастает с увеличением отрицательного смещения на управляющем электроде.

Полевые тиристоры устойчивы по отношению к эффекту dU/dt и функционируют при более высоких температурах по сравнению с обычными тиристорами, т. е. они устойчивы и по отношению к эффекту dI_A/dt . Используются в быстродействующих схемах, требующих малого времени включения и выключения.

Параметры тиристорov. Основными параметрами тиристорov являются: напряжение включения $U_{\text{вкл}}$, ток включения $I_{\text{вкл}}$, ток выключения (удержания) $I_{\text{выкл}}$ ($I_{\text{уд}}$), максимально допустимый ток в открытом состоянии $I_{\text{пр. макс}}$ (см. рис. 5.1, *в*), время задержки $t_{\text{зд}}$, время включения $t_{\text{вкл}}$, время выключения $t_{\text{выкл}}$, максимально допустимая скорость нарастания прямого напряжения $(dU/dt)_{\text{макс}}$, максимально допустимая скорость нарастания прямого тока $(dI/dt)_{\text{макс}}$ и некоторые другие.

Модели тиристорov. В п. 5.2 была рассмотрена двухтранзисторная модель тиристора (см. рис. 5.2). При анализе и численных расчетах схем, кроме этой, часто используют модель, в которой отражены все основные параметры тиристорov. Обычно в такой модели три перехода Π_1 , Π_2 и Π_3 $p-n-p-n$ -структуры (см. рис. 5.1, *а*) заменяют диодами, параллельно которым включают барьерные и диффузионные емкости переходов. Инжекция носителей заряда через переходы Π_1 и Π_3 (см. рис. 5.1, *а*), а также влияние управляющего электрода моделируются с помощью генераторов тока, включаемых параллельно диоду центрального $p-n$ -перехода Π_2 , между управляющим электродом и катодом подсоединяется сопротивление шунта $R_{\text{ш}}$ (см. рис. 5.3, *а*). Кроме того, иногда в эквивалентной схеме тиристора находят отражение параметры, связанные с особенностями конструкции, условиями эксплуатации, переходными процессами и т. д.



Контрольные вопросы

1. Общие сведения о тиристорах. Классификация.
2. Каковы устройство и ВАХ тиристора?
3. Эквивалентная схема тиристора и физические процессы в тиристоре.
4. Переходные процессы и работа тиристора в импульсном режиме.
5. Разновидности тиристорov, симисторы: устройство, особенности работы, области применения.
6. Полевые тиристоры: устройство, особенности работы, области использования.

Глава 6

ПОЛЕВЫЕ ТРАНЗИСТОРЫ

6.1. Общие сведения

Полевой транзистор (ПТ) — это трехэлектродный твердотельный прибор, в котором процессы преобразования сигналов осуществляются за счет переноса носителей одного типа (или электронов, или дырок).

Основные электроды полевого транзистора называются *исток (И)*, *затвор (З)*, *сток (С)*. Движение носителей происходит от истока к стоку через канал, расположенный под затвором. Исток, сток и канал — это полупроводниковые области одного и того же типа проводимости. Помимо этих областей, в большинстве полевых транзисторов имеется *подложка (П)*, на которой формируются указанные области. Изменение проводимости канала, а часто и само формирование канала, происходит за счет подачи и изменения потенциала на затворе. Затвор является управляющим электродом. Создаваемое им поле направлено перпендикулярно перемещению носителей в канале, движущихся от истока к стоку под действием продольного поля, формируемого между истоком и стоком. В большинстве случаев канал выполнен в виде слаболегированного тонкого полупроводникового слоя, с *n*- или *p*-типом проводимости, который располагается либо у самой поверхности полупроводника, либо внутри кристалла на некотором расстоянии от его поверхности. В противоположность каналу исток и сток являются сильнолегированными областями.

Существует три типа полевых транзисторов: *полевые транзисторы (ПТ) с управляющим p — n -переходом, с управляющим переходом металл — полупроводник, полевые транзисторы с изолированным затвором и структурой металл — диэлектрик — проводник (МДП-транзисторы)*.

Наибольшее распространение получили МДП-транзисторы, которые широко используются в интегральных схемах и как дискретные приборы. МДП-транзисторы применяются в переключающих схемах. Транзисторы с переходом металл — полупроводник находят применение в быстродействующих цифровых интегральных микросхемах и СВЧ-устройствах. ПТ с управляю-

щим $p-n$ -переходом используются чаще всего в качестве низко-частотных дискретных приборов.

В транзисторах с любым видом управляющего электрического перехода канал исходно технологически сформирован. Однако в МДП-транзисторах канал часто технологически не сформирован. Он образуется за счет приложения к затвору определенного напряжения и формирования поперечного к поверхности электрического поля. Такие транзисторы называются *МДП-транзисторами с индуцированным каналом*. Если же в исходном состоянии канал технологически сформирован, то такие приборы называются *МДП-транзисторами со встроенным каналом*.

Все указанные транзисторы различаются между собой структурой и способом управления проводимостью канала. При этом под способом в данном случае следует понимать не различие физических процессов в каждом типе транзистора, а физико-технологическую структурную разновидность, реализующую изменение проводимости под влиянием потенциалов на электродах. В полевых транзисторах с затворами металл — полупроводник или $p-n$ -переходом на электрический переход подается обратное напряжение, которое изменяет толщину обратносмещенного перехода. Следовательно, напряжение на затворе, меняя толщину обедненного слоя, изменяет и толщину проводящей части канала, т. е. его сопротивление и ток через него. В МДП-транзисторах напряжение на металлическом затворе через тонкий слой диэлектрика создает поперечное относительно поверхности электрическое поле в полупроводнике, которое управляет концентрацией носителей в канале.

Затвор в электрических схемах обычно является входным электродом, поэтому полевые транзисторы имеют, в отличие от биполярных, большое входное сопротивление на постоянном токе, которое определяется либо обратносмещенным электрическим переходом, либо тонким слоем диэлектрика. Поэтому полевые транзисторы, как и электронные лампы (см. гл. 11), относятся к приборам, управляемым напряжением (электрическим полем), в которых входное напряжение определяется ЭДС входного источника и не зависит от параметров самого прибора. В то время как биполярные транзисторы из-за малого входного сопротивления относятся к приборам, управляемым током.

Полевые транзисторы могут включаться по схемам с *общим истоком*, *общим затвором* или *общим стоком*. Наиболее распространенной является схема с общим истоком. Обозначения МДП-транзис-

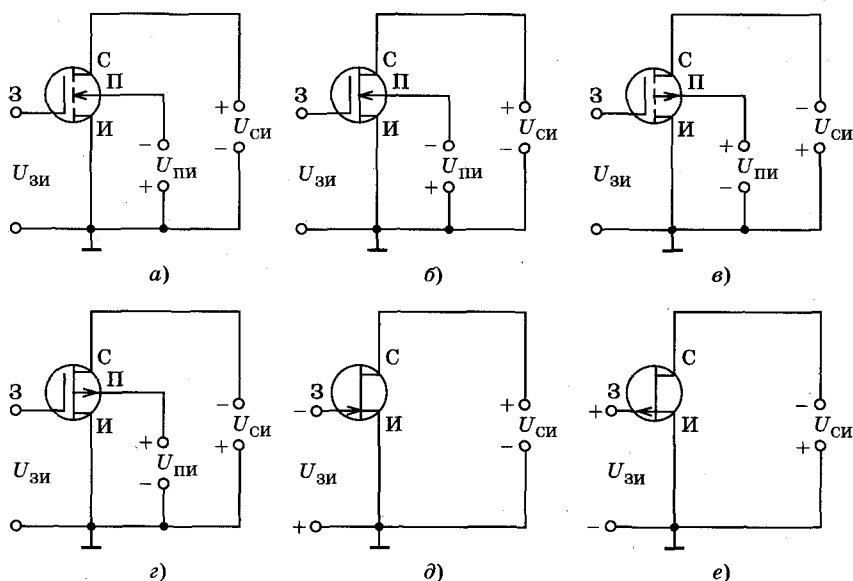


Рис. 6.1

торов, включенных по схеме с общим истоком с указанием полярности поданных на электроды напряжений, представлены на рис. 6.1. На рис. 6.1, а и 6.1, в соответственно представлены МДП-транзисторы с индуцированным каналом n -типа и p -типа, где область исток — канал — сток изображается штриховой линией, а для транзисторов со встроенными каналами n - (рис. 6.1, б) и p -типа (рис. 6.1, г) — сплошной. Условные обозначения в схеме включения с общим истоком для транзисторов с переходом металл — полупроводник и управляющим переходом изображены на рис. 6.1, д и 6.1, е, соответственно для каналов n - и p -типа.

6.2. Формирование канала в МДП-транзисторах

Устройство МДП-транзистора с индуцированным каналом представлено на рис. 6.2, а, где 1 — подложка, 2 — область истока, 3 — вывод истока, 4 — диэлектрический слой SiO_2 , 5 — затвор, 6 — индуцированный канал (или область формирования канала), 7 — сток, 8 — область стока, L — длина канала, δL — область перекрытия с 2 и 8 (см. также рис. 6.10). Если напряжение на затворе отсутствует, то в исходном состоянии канал не существует.

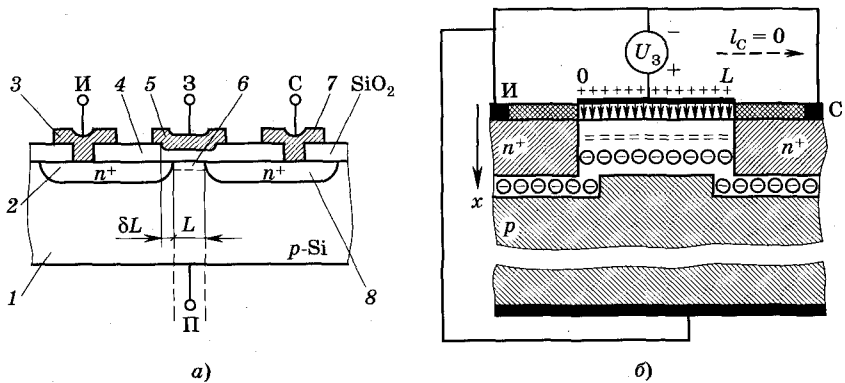


Рис. 6.2

Рассмотрим процессы в МДП-структуре, которая сформирована в области затвора и состоит из металлического слоя затвора, тонкого диэлектрического слоя, подложки p -типа проводимости и металлического электрода, образующего вывод подложки (рис. 6.2, б). Сначала предположим, что на исток и сток напряжение не подается и их цепи объединены с подложкой (см. рис. 6.2, б). Напряжение будет прикладываться между затвором и подложкой. Для простоты будем также считать, что работа выхода полупроводника и металла одинакова и контактная разность потенциалов между ними равна нулю, поверхностный заряд на границе раздела полупроводник—диэлектрик отсутствует, а концентрация атомов акцепторов в подложке одинакова по всему объему.

Рассматриваемая структура подобна конденсатору с обкладками из металла и полупроводника p -типа. Если к затвору приложить относительно подложки отрицательный потенциал (металл заряжается отрицательно), то на границе полупроводника с диэлектриком появляется индуцированный положительный заряд, который возникает за счет дырок, притянутых из объема электрическим полем, создаваемым внешним отрицательным зарядом. В диэлектрике свободных носителей заряда нет, поэтому напряженность электрического поля в нем одинакова по объему. В полупроводнике концентрация избыточных носителей максимальна у поверхности и уменьшается по мере удаления от нее в глубину объема по направлению к выводу подложки. В результате объемный положительный заряд создает в приповерхностной области электрическое поле, направленное навстрече-

чу внешнему полю, что уменьшает результирующее поле. Такой режим работы МДП-структуры называется **режимом обогащения**, поскольку в приповерхностном слое полупроводника концентрация дырок $p_{\text{пов}}$ будет больше концентрации атомов акцепторов ($p_{\text{пов}} > N_a$), чем в глубине полупроводника, где $p_0 = N_a$. Напряженность поля спадает по экспоненте $\exp(-x/L_D)$ по мере удаления от поверхности (координата x направлена от поверхности в глубь полупроводника).

Параметр $L_D = [(\epsilon_0 \epsilon_F T / q N_a)]^{1/2}$ называется **дебаевской длиной экранирования**. Толщина обогащенного слоя порядка L_D . При $x > (2 \dots 3)L_D$ электрическое поле практически отсутствует из-за экранировки внешнего поля избыточным зарядом дырок $p_{\text{пов}}$.

При подаче на затвор положительного потенциала в приповерхностном слое p -полупроводника реализуются **режимы обеднения и инверсии** (см. рис. 6.2, б). В режиме обеднения электрическое поле вытесняет дырки от поверхности полупроводника вглубь и их концентрация в этой области уменьшается, но концентрация неосновных носителей (в данном случае электронов) начинает возрастать (рис. 6.3, а, б). Однако концентрация их остается низкой, поэтому у поверхности полупроводника образу-

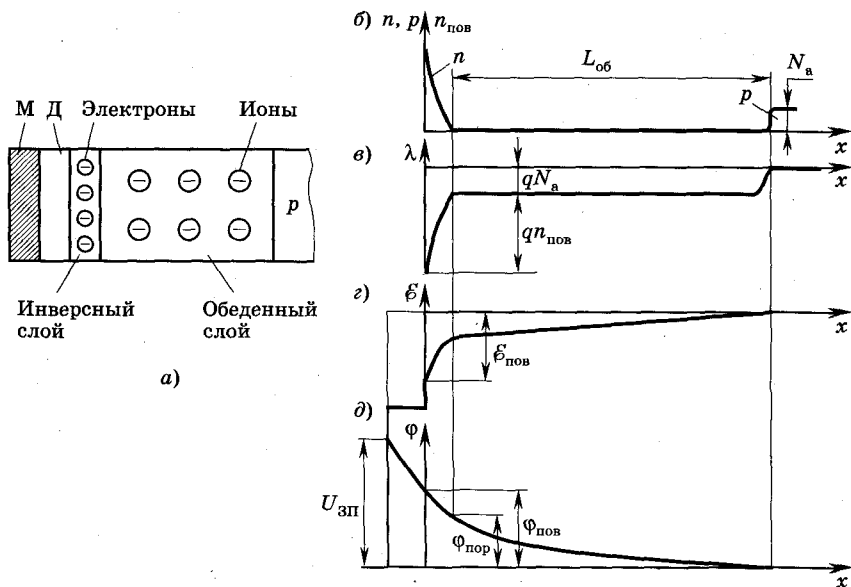


Рис. 6.3

ется обедненный основными носителями слой $L_{об}$, в котором $p_{пов}$ и $n_{пов}$ меньше N_a и существует отрицательный заряд некомпенсированных акцепторных ионов с приблизительно постоянной объемной плотностью λ (рис. 6.3, в). Напряженность электрического поля \mathcal{E} вне обедненного слоя при $x > L_{об}$ равна нулю (рис. 6.3, г). Из-за наличия отрицательного объемного заряда возникает **поверхностный потенциал** $\varphi_{пов}$ (рис. 6.3, д), который определяется разностью потенциалов между поверхностью на границе диэлектрик—полупроводник и координатами $x \geq L_{об}$, где отсутствует электрическое поле. Толщина обедненного слоя $L_{об}$ может быть определена на основе решения уравнения Пуассона, в результате получается выражение, аналогичное формуле (2.11) для несимметричного p — n -перехода (см. [2], п. 1.8):

$$L_{об} = \epsilon_0 \mathcal{E} / (qN_a) = [2\epsilon_0 \epsilon \varphi_{пов} / (qN_a)]^{1/2}, \quad (6.1)$$

где \mathcal{E} — напряженность электрического поля.

Тогда поверхностный потенциал равен

$$\varphi_{пов} = \epsilon_0 \mathcal{E}^2 / (2\epsilon qN_a). \quad (6.2)$$

Поверхностные концентрации дырок и электронов связаны со значением поверхностного потенциала $\varphi_{пов}$ и могут быть вычислены согласно следующим выражениям:

$$p_{пов} = N_a \exp(-\varphi_{пов}/\varphi_T), \quad n_{пов} = (n_i^2/N_a) \exp(-\varphi_{пов}/\varphi_T). \quad (6.3)$$

По формулам (6.3) концентрацию зарядов в приповерхностном слое можно вычислять и для других режимов (обогащения и инверсии).

Режим обеднения существует при $n_{пов} < N_a$, когда поверхностный потенциал $\varphi_{пов}$ превышает пороговое значение

$$\varphi_{пор} = 2\varphi_T \ln(N_a/n_i). \quad (6.4)$$

При $N_a = 10^{16} \text{ см}^{-3}$ и $T = 300 \text{ К}$ для Si $\varphi_{пор} = 0,7 \text{ В}$.

Величина $\varphi_{пор}$ достигается при соответствующем **пороговом напряжении** $U_{пор}$ на затворе. При $U_a = U_{пор}$ концентрация электронов в приповерхностном слое $n_{пов} = N_a$. Если же $U_a > U_{пор}$ в МДП-структуре реализуется **режим инверсии**, при котором $n_{пов} > N_a$ (см. рис. 6.3, а, б, в), т. е. у поверхности образуется хорошо проводящий инверсный слой с типом проводимости, противо-

положительным типу проводимости подложки, который выполняет роль канала (см. рис. 6.2, а, б). Под инверсным слоем располагается обедненная область, а далее подложка р-типа (см. рис. 6.2, а, б). На рис. 6.3 (б, в, г, д) показаны распределения в указанных слоях (областях) концентрации свободных носителей p , n , объемного заряда ионов акцепторов $\lambda = qN_a$, напряженности \mathcal{E} и потенциала электрического поля ϕ . Скачок напряженности электрического поля (см. рис. 6.3, г) на границе раздела Si—SiO₂ ($x = 0$) обусловлен различием диэлектрической проницаемости Si и SiO₂.

По мере удаления от поверхности ($x \geq 0$) напряженность электрического поля и концентрация электронов уменьшается, как и в режиме обогащения и обеднения, по закону $\exp(-x/L_D)$. При $N_a = 10^{16}$ см⁻³ дебаевская длина равна $L_D = 0,04$ мкм. С увеличением $|U_3|$ растет и напряженность электрического поля $|\mathcal{E}_{\text{пов}}|$ в инверсном слое, в то время как в обедненном слое она практически не изменяется.

В реальной МДП-структуре на границе Si—SiO₂ существует положительный поверхностный заряд с плотностью $Q_{\text{пов}}$. Этот заряд обусловлен тем, что структура поверхности полупроводника характеризуется большим числом дефектов, примесей и адсорбированных атомов различных веществ, формирующих в приповерхностном слое энергетические уровни, расположенные в запрещенной зоне. Эти уровни образуют поверхностные ловушки, которые, захватывая подвижные носители, превращаются в положительные и отрицательные ионы. Поверхностный заряд, образуемый ловушками, непостоянен, так как число заряженных ловушек изменяется в зависимости от напряженности электрического поля у поверхности.

В кремнии, покрытом SiO₂, вблизи границы раздела существует тонкий переходной слой, содержащий дефекты типа кислородных вакансий, образующих поверхностный положительный заряд ионов Si⁺ ($Q_{\text{пов}} = 10^{10} \dots 10^{12}$ см⁻²), который в рассматриваемой структуре является преобладающим. Его влияние в МДП-транзисторах сказывается наиболее сильно.

Поверхностный заряд $Q_{\text{пов}}$ создает собственное поле, направленное в ту же сторону, что и поле, обусловленное положительным напряжением на затворе, в результате этого происходит уменьшение порогового напряжения $U_{\text{пор}}$.

6.3. Общие принципы управления проводимостью канала в полевых транзисторах.

Статические вольт-амперные характеристики

Структурные схемы основных типов ПТ изображены на рис. 6.2, а, 6.4, а и 6.5, а. В МДП-транзисторе с изолированным затвором (см. рис. 6.2, а) канал образуется или за счет подачи

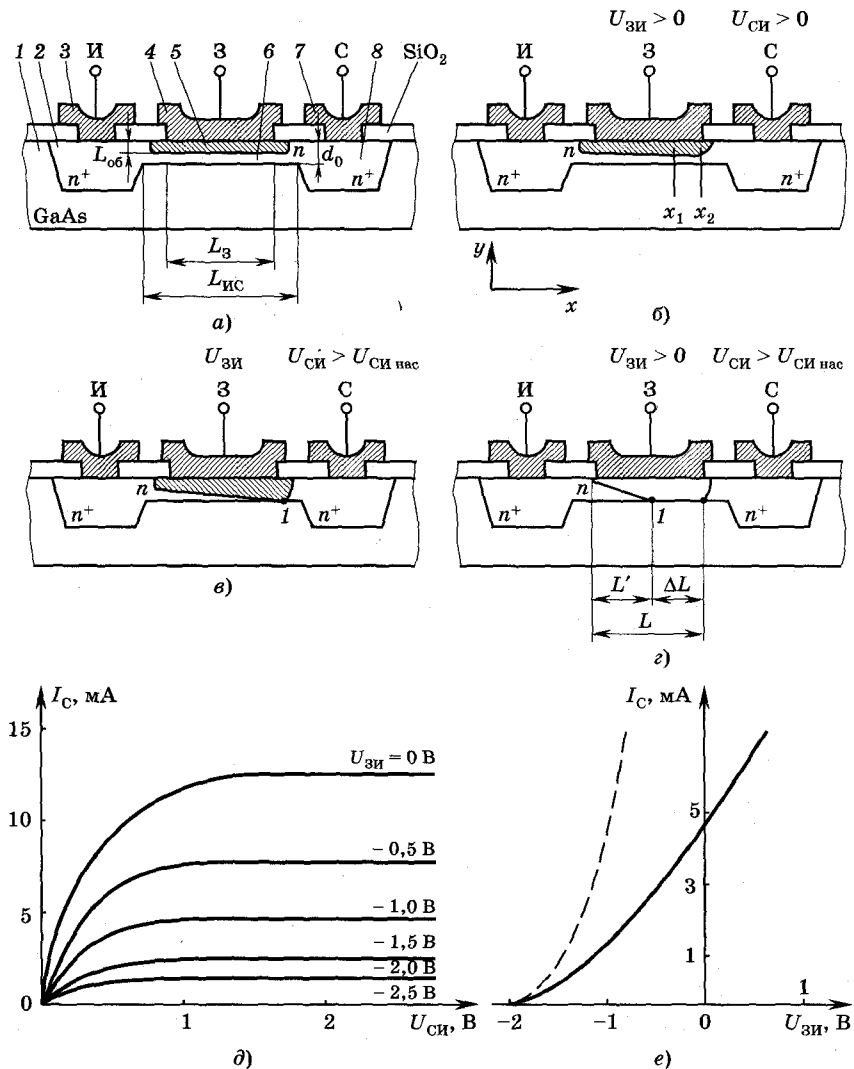


Рис. 6.4

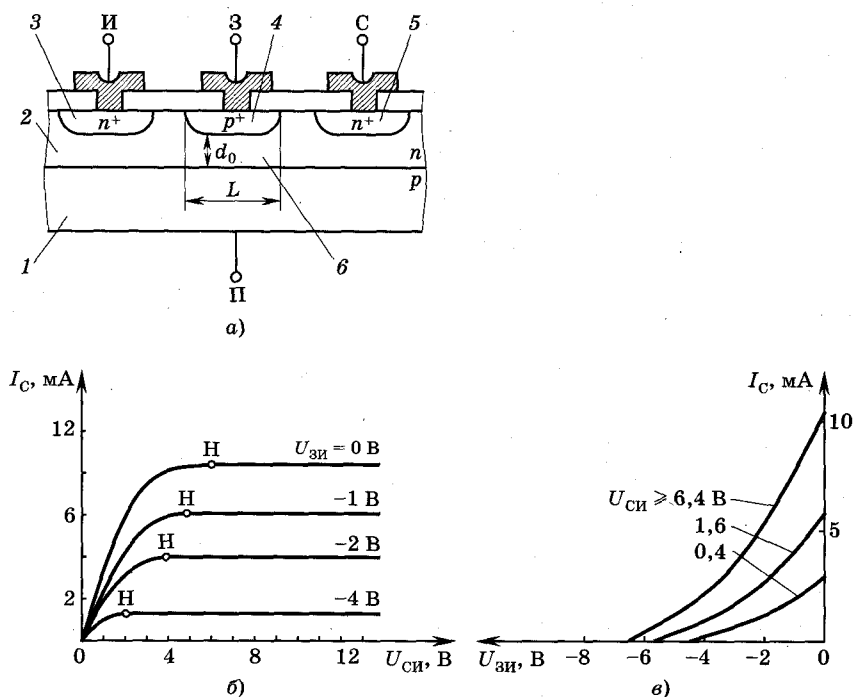


Рис. 6.5

на затвор напряжения $U_3 > U_{3\text{пор}}$, или же формируется технологически. Транзистор с затвором на основе контакта металл — полупроводник, или контакта Шоттки (ПТШ), показан на рис. 6.4, а, где приняты следующие обозначения: 1 — подложка, 2 — область истока, 3 — исток, 4 — затвор, 5 — обедненный слой, 6 — канал, 7 — сток, 8 — область стока. Транзистор с управляющим p — n -переходом (ПТУП) показан на рис. 6.5, а, где цифры имеют следующие значения: 1 — подложка, 2 — эпитаксиальный слой, в котором формируется канал, 3 — область истока, 4 — область затвора, 5 — область стока, 6 — канал. Выбранные структуры транзисторов все имеют канал n -типа. Рассматриваемые физические процессы не зависят от типа канала.

Во всех ПТ с конкретной структурой проводимостью канала можно управлять изменением напряжения на затворе, на стоке и на подложке. Последний случай реализуется обычно в МДП-транзисторах. Мы будем рассматривать схему включения с общим истоком, для которой подложка может быть либо соедине-

на с истоком, либо подключена к независимому источнику. Далее рассмотрим случай, когда исток и подложка соединены, т. е. находятся под одним и тем же потенциалом.

Если на затвор любого типа транзистора подать напряжение U_3 , то для каждого значения U_3 формируется исходная проводимость (исходное сопротивление) канала, определяемая концентрацией свободных носителей в нем и его размерами (структурой). Эта исходная проводимость при заданных напряжениях на затворе для каждого типа транзистора определяет токи через канал при изменении потенциалов между стоком и истоком в некоторых пределах $U_{СИ} \leq U_{СИ\text{нас}}$. Рассмотрим эти процессы более подробно.

Изучим сначала влияние изменения напряжения $U_{ЗИ}$ при $U_{СИ} = \text{const}$ на работу ПТШ и ПТУП. В транзисторах этих типов электрический переход должен быть смещен только в обратном направлении, поэтому полярность напряжения на затворах для n -канала является отрицательной, т. е. $U_{ЗИ} < 0$, а для p -канала $U_{ЗИ} > 0$. В МДП-транзисторах (МДПТ) полярность напряжений обратная, т. е. для n -канала $U_{ЗИ} > 0$, а для p -канала $U_{ЗИ} < 0$. Управление проводимостью осуществляется за счет изменения толщины обедненного слоя (переходов), а следовательно, и поперечного сечения канала. Чем больше напряжение на затворе по абсолютной величине при $U_{СИ} = \text{const}$, тем шире переход и уже канал (см. рис. 6.4, б, в и 6.6, б, в). Форма обедненного слоя и, следовательно, канала для ПТУП и ПТШ подобны. Следовательно, с увеличением $|U_3|$ при постоянном напряжении между стоком и истоком ($U_{СИ} = \text{const}$) ток стока будет уменьшаться (см. рис. 6.4, е и 6.5, в) из-за уменьшения сечения канала.

В МДП-транзисторах изменение напряжения $U_{ЗИ}$ при сформированном канале изменяет сечение, а также проводимость канала. Например, в МДПТ с n -каналом при увеличении $U_{ЗИ}$ ток стока будет возрастать (см. рис. 6.6, е).

Перейдем теперь к рассмотрению влияния $U_{СИ}$ при $U_{ЗИ} = \text{const}$.

Если изменять напряжение на стоке на растущем участке ВАХ, где $U_{СИ} < U_{СИ\text{нас}}$, то в этой области ток стока как функция $U_{СИ}$ изменяется почти линейно, при этом стокковые характеристики $I_C = f(U_{СИ})$ идут веерообразно для различных значений $U_{ЗИ}$ (см. рис. 6.4, д; 6.5, б; 6.6, д). Такая «веерообразность» характеристик объясняется тем, что канал представляет собой по-

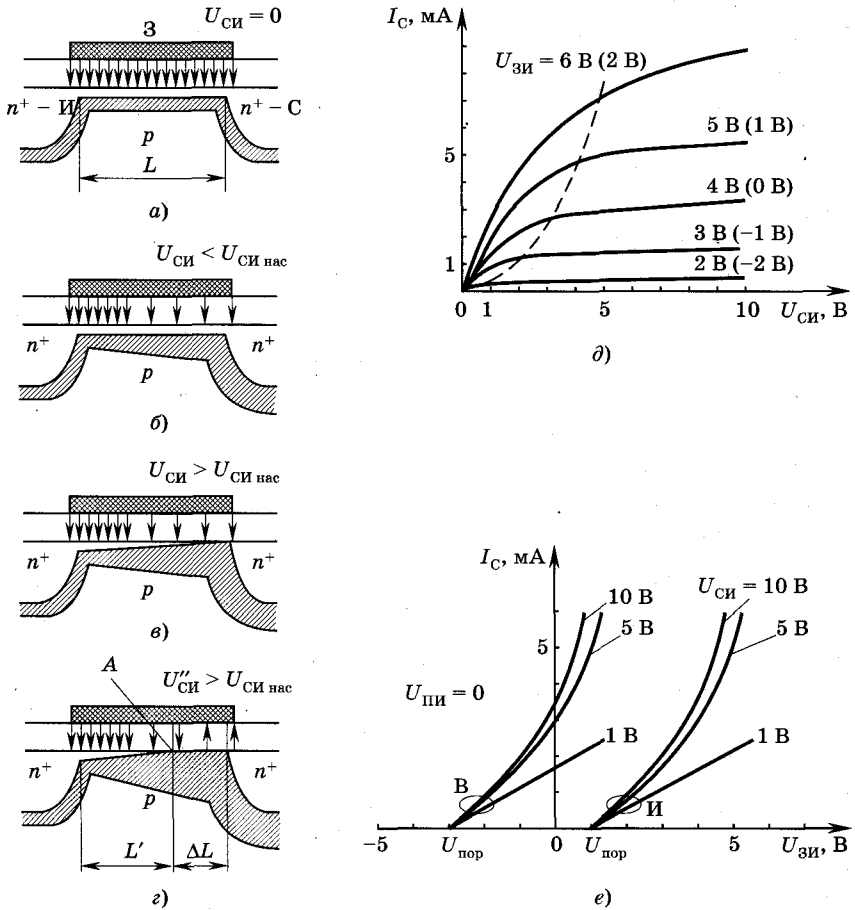


Рис. 6.6

лупроводниковое сопротивление, которое, как показано ранее, различно для разных $U_{ЗИ}$. Рост же тока стока при фиксированном $U_{ЗИ}$ на рассматриваемом участке связан с тем, что к каналу, представляющему полупроводниковое сопротивление, прикладывается большее напряжение, т. е. все происходит согласно закону Ома.

Однако при приближении напряжения на стоке к $U_{СИ\text{нас}}$ ситуация изменяется. При $U_{СИ} \approx U_{СИ\text{нас}}$ происходит перекрытие канала обедненной областью p - n -перехода, которое называют **отсечкой канала**. Во всех типах транзисторов это перекрытие осу-

ществляется со стороны области стока. В этой области обратное напряжение на электрическом переходе максимально и, следовательно, ширина обедненной области наибольшая. Поскольку расширение перехода происходит в стороны у канала и подложки, то при $U_{\text{СИ}} = U_{\text{СИ нас}}$ канал перекрывается.

Электрофизические параметры электрических переходов различаются для разных типов ПТ, поэтому и значения напряжения отсечки будут отличаться. В ПТШ и ПТУП отсечка канала происходит на границе области затвора и стока (см. рис. 6.4, в, з, точка 1), в то время как в МДП-транзисторах перекрытие канала осуществляется за счет расширения обедненной области перехода между стоком и подложкой (см. рис. 6.6, в, з). Необходимо отметить, что, в отличие от ПТУП и ПТШ, в МДП-транзисторах свойствами сформированного канала можно управлять изменением напряжения на стоке и/или на подложке (см. ВАХ на рис. 6.7), т. е. можно сказать, что напряжение на затворе формирует канал с первоначально заданными характеристиками, а потом управление проводимостью осуществляется за счет изменения напряжения $p-n$ -перехода сток—подложка.

Сказанное иллюстрируют соответствующие рисунки: на рис. 6.4, б, в, з для ПТШ и на рис. 6.6, а, б, в, з для МДП показаны форма канала и структура перехода на границах стока, канала и истока с подложкой для МДП-транзистора при различных значениях $U_{\text{СИ}}$ и $U_{\text{ЗН}} = \text{const}$. При дальнейшем (после $U_{\text{СИ нас}}$) увеличении абсолютного значения напряжения на стоковом переходе граница области отсечки канала во всех типах транзисторов будет продвигаться в направлении области истока (см. рис. 6.4, з; 6.6, з). Следовательно, между каналом и стоком размещается все большая величина обедненной области ΔL (см. рис. 6.4, з; 6.6, з), что эквивалентно увеличению сопротивления, включенного между каналом и стоком. Увеличение сопротивления этой обедненной области будет примерно пропорционально увеличению напряжения $U_{\text{СИ}}$. В результате при $U_{\text{СИ}} > U_{\text{СИ нас}}$ ток стока изменяется незначительно и стоковые ВАХ $I_{\text{С}} = f(U_{\text{СИ}})$ идут под небольшим углом к оси абсцисс (см. рис. 6.4, д для ПТШ, 6.5, б для ПТУП и 6.6, д для МДП-транзисторов).

На рис. 6.4, д и 6.5, д дано семейство стоковых (выходных) характеристик соответственно для ПТШ и ПТУП. При увеличении $|U_{\text{ЗН}}|$ проводимость канала падает и характеристики смещаются вниз; в МДП-транзисторах ситуация обратная: с ростом $|U_{\text{ЗН}}|$ уве-

личивается проводимость канала, что ведет к увеличению тока, и характеристики смещаются вверх (см. рис. 6.6, δ , где значения напряжения в скобках относятся к встроенному каналу).

Описанные процессы хорошо отражают свойства ПТ с так называемым *длинным каналом*. Канал называется длинным, если его продольные размеры (L) существенно больше поперечных. Если же отношение длины канала к его толщине не слишком велико, то канал считается коротким.

Реальные характеристики полевых транзисторов с *короткими каналами* в области насыщения могут заметно отличаться от характеристик с длинным каналом. Одной из главных причин такого отличия является зависимость дрейфовой скорости и подвижности носителей тока от величины электрического поля. Дрейфовая скорость при больших напряжениях поля насыщается. В кремнии дрейфовая скорость монотонно увеличивается при возрастании электрического поля и достигает скорости насыщения ($v_{\text{нас}} \approx 10^7$ см/с) при напряженностях поля, больших, чем $5 \cdot 10^4$ В/см. В GaAs и InP зависимость скорости дрейфа от напряженности поля $v = f(\mathcal{E})$ немонотонна. Дрейфовая скорость сначала достигает максимума ($v_{\text{нас}} \approx 2 \cdot 10^7$ см/с) при полях $\mathcal{E} \sim (1 \dots 2) 10^4$ В/см и затем уменьшается до $\sim (6 \dots 8) 10^6$ см/с при $\mathcal{E} \sim (5 \dots 6) 10^4$ В/см, оставаясь далее почти постоянной.

Рассмотрим особенность физических процессов и поведение стоковых ВАХ в области насыщения для кремниевого короткоканального ПТШ. Описываемые ниже процессы характерны и для других типов полевых транзисторов. Для объяснения процессов будем пользоваться рис. 6.4, δ , считаем, что в данном случае ПТШ имеет короткий канал.

При малых напряжениях на стоке канал, как уже отмечалось ранее, подобен резистору. При больших напряжениях темп увеличения дрейфовой скорости электронов с ростом напряжения уменьшается и ВАХ канала, т. е. кривая $I_C = f(U_{\text{СИ}})$ отклоняется вниз от начальной прямой. При еще больших напряжениях $U_{\text{СИ}} \approx U_{\text{СИ нас}}$ дрейфовая скорость носителей достигает скорости насыщения, что вызывает насыщение тока стока.

Обедненная область под затвором в ПТШ и ПТУП (см. рис. 6.4, δ) действует как слой изолятора, подобно подзатворному диэлектрику в МДП-транзисторах, что сокращает высокопроводящую область канала, через которую протекает стоковый

ток. Ширина обедненного слоя определяется величиной приложенных напряжений. Потенциал вдоль канала увеличивается от нуля на истоке до $U_{СИ}$ на стоке. Следовательно, значение локального обратного смещения перехода затвор—канал и локальная ширина обедненного слоя также увеличиваются по мере приближения к стоку (ситуация подобна рис. 6.4, б). В любом сечении канала должен выполняться закон сохранения полного тока. Поэтому результирующее уменьшение ширины проводящего канала должно быть скомпенсировано соответствующим увеличением продольного электрического поля и дрейфовой скорости электронов, чтобы величина полного тока оставалась неизменной по длине канала. При $U_{СИ} = U_{СИ\text{нас}}$ напряженность поля в канале со стороны стока достигает критического значения, а дрейфовая скорость электронов — скорости насыщения, ширина канала становится минимальной, и ток транзистора начинает насыщаться.

С увеличением напряжения стока ($U_{СИ} > U_{СИ\text{нас}}$) обедненная область расширяется к стоку, а сечение канала S_1 , проходящее через координату x_1 (см. рис. 6.4, б), где электроны достигают скорости насыщения, смещается в сторону истока, т. е. в противоположном направлении. При этом, в отличие от длинного канала, полного перекрытия короткого канала не происходит. Падение напряжения между истоком и сечением S_1 уменьшается, и, в соответствии с законом сохранения полного тока, увеличивается ток инжекции электронов из электронейтральной части канала со стороны истока в область, где происходит насыщение дрейфовой скорости. В результате этого ВАХ транзистора на участке насыщения имеет положительный наклон, т. е. происходит возрастание тока I_C с увеличением напряжения $U_{СИ}$ в области насыщения.

По мере удаления от сечения S_1 к стоку потенциал вдоль канала возрастает, размер, занятый обедненной областью, увеличивается, а проводящая область канала сужается. Поскольку в этой области скорость электронов $v_{др} = v_{нас} = \text{const}$, т. е. не зависит уже от напряженности электрического поля, то для компенсации этого сужения канала и обеспечения сохранения полного тока концентрация электронов в рассматриваемой области увеличивается и становится больше концентрации доноров, т. е. эта часть канала оказывается отрицательно заряженной. В не-

котором сечении в области стока, где сечение канала S_2 , проходящее через x_2 , равно S_1 , концентрация электронов равна концентрации доноров. Указанный отрицательный заряд области канала компенсируется положительно заряженным слоем, расположенным за слоем S_2 ближе к стоку, где наблюдается некоторый дефицит электронов. Следовательно, часть напряжения стока $U_{СИ}$, превышающая $U_{СИ\text{нас}}$, падает на образовавшемся дипольном слое, который расширяется в стоковой области канала при дальнейшем росте $U_{СИ}$. Падение напряжения на дипольном слое приводит к большему отклонению ВАХ в области насыщения к оси абсцисс.

Если увеличить отрицательное напряжение на затворе (абсолютную величину $|U_{ЗИ}|$), то ширина обедненной области возрастает и происходит сужение канала, что увеличивает сопротивление на линейном участке стоковой характеристики $I_C = f(U_{СИ})$ (при малых $U_{СИ}$) транзистора. При этом уменьшается напряжение $U_{СИ\text{нас}}$, при котором в наиболее узкой части канала достигается критическое поле, соответствующее насыщению скорости. В силу сказанного характеристика смещается вниз, а участок насыщения характеристики транзистора с увеличением $|U_{ЗИ}|$ начинается при меньших напряжениях на стоке $U_{СИ}$ (см. рис. 6.5, б и 6.6, д).

На рис. 6.5, б переход в режиме насыщения обозначен буквой «Н», а на рис. 6.6, д соответствующая ВАХ дана штриховой кривой. В скобках здесь указаны значения напряжения на затворе, относящиеся к МДП-транзисторам со встроенным каналом, где реализуется режим обеднения канала (малые токи I_C , $U_{ЗИ} < 0$) и обогащение канала носителями, втянутыми в существующий канал за счет напряжения $U_{ЗИ}$ (большие токи I_C , $U_{ЗИ} > 0$).

Подчеркнем еще раз: в транзисторах с коротким каналом полного перекрытия канала не происходит.

Помимо рассмотренных выходных стоковых характеристик $I_C = f(U_{СИ})$ при $U_{ЗИ} = \text{const}$, широко используются передаточные (стокозатворные) характеристики $I_C = f(U_{ЗИ})$ при $U_{СИ} = \text{const}$ (см. рис. 6.4, е; 6.5, в; 6.6, е). На рис. 6.6, е приведены семейства $I_C = f(U_{ЗИ})$ для МДП-транзисторов со встроенным каналом, обозначенные индексом В, и с индуцированным каналом (И), а

на рис. 6.4, *e* и рис. 6.5, *e* — аналогичные характеристики соответственно для ПТШ и ПТУП. При $U_{СИ} < U_{СИ\text{нас}}$ (крутые участки выходных характеристик) для всех типов транзисторов рассматриваемые стокзатворные характеристики близки к линейным, поскольку относятся к однородным каналам, когда в них нет перекрытия. В этом случае каналы во всех типах транзисторов ведут себя подобно полупроводниковым резисторам, что хорошо видно на рис. 6.6, *д* при $U_{СИ} = 1$ В.

При значениях $U_{СИ}$, соответствующих пологой области выходных характеристик, когда $U_{СИ} > U_{СИ\text{нас}}$, передаточные характеристики описываются квадратичной зависимостью. Квадратичная зависимость стокзатворных характеристик объясняется тем, что ток стока пропорционален объемной плотности заряда электронов $Q_n = C_{уд}(U_{ЗИ} - U_{пор})$ в канале и напряжению на перекрытой части канала $U_{СИ\text{нас}} = U_{ЗИ} - U_{пор}$, т. е. $I_C \sim Q_n$, $U_{СИ\text{нас}} \sim (U_{ЗИ} - U_{пор})^2$ (более подробно см. п. 6.4).

Квадратичная зависимость $I_C = f(U_{ЗИ}) \sim U_{ЗИ}^2$ характерна для транзисторов с длинным каналом (см. рис. 6.4, *e*, штриховая линия). Для транзисторов с коротким каналом передаточная характеристика квадратична только при напряжениях $U_{ЗИ}$ вблизи порогового (см. рис. 6.4, *e*, сплошная линия). С ростом напряжения $U_{ЗИ}$ на непокрытой части канала увеличивается напряженность продольного электрического поля. В результате подвижность электронов снижается, дрейфовая скорость приближается к скорости насыщения и перестает зависеть от напряжения $U_{ЗИ}$, в результате характеристика становится линейной.

Для МДП-транзисторов при наличии отдельного вывода от подложки, возможно использование также семейства стокзатворных характеристик, измеренных при $U_{СИ} = \text{const}$, но при разных значениях напряжения на подложке $U_{ПИ}$, измеряемого относительно истока (рис. 6.7). Напряжение $U_{ПИ}$ изменяет обратное смещение на переходе сток—подложка и, следовательно, параметры канала, подобно тому, как это показано на рис. 6.6, *б*, *в*, *г*.

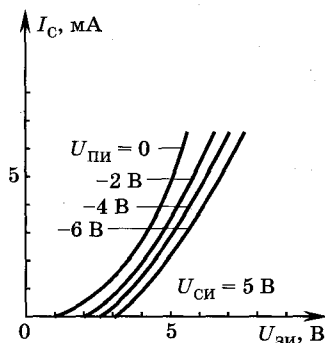


Рис. 6.7

6.4. Моделирование полевых транзисторов

Теоретически статические характеристики и параметры транзисторов могут быть определены совместным решением уравнений Пуассона и непрерывности, с помощью которых вычисляются распределение напряженности электрического поля и концентрация носителей заряда. Поскольку при анализе необходимо учитывать как продольную, так и поперечные составляющие электрического поля в канале транзисторов, то одномерное приближение не подходит для данного случая и необходимо рассматривать по крайней мере двумерную задачу, которая аналитически не решена. В результате приходится использовать численные компьютерные методы. Однако для практических целей моделирования, расчета схем и определения параметров транзисторов необходимы простые аналитические выражения ВАХ, которые можно получить только, если использовать ряд упрощающих допущений в реальной картине протекания физических процессов в транзисторах. Наиболее важными являются следующие допущения:

- 1) полагается, что продольная составляющая напряженности электрического поля много меньше поперечной, т. е. $\epsilon_y \ll \epsilon_x$ — приближение плавного канала;
- 2) движение носителей в канале считается чисто дрейфовым, т. е. диффузионной компонентой пренебрегается.

Другие допущения: подвижность носителей в канале не зависит от напряженности электрического поля; поверхностный заряд и контактная разность потенциалов металл — полупроводник равна нулю; обратные токи p — n -переходов малы по сравнению с током канала; эффект модуляции длины канала отсутствует; ударная ионизация в стоковом p — n -переходе отсутствует; сопротивления областей истока и стока пренебрежимо малы по сравнению с сопротивлением канала; поверхностный потенциал $\phi_{\text{пов}}$ у истока при $U_{\text{зи}} > U_{\text{пор}}$ равен $\phi_{\text{пор}} = \text{const}$, а у стока $\phi_{\text{пов}} = \phi_{\text{пор}} + U_{\text{си}}$ для крутой части ВАХ и $\phi_{\text{пов}} = \phi_{\text{пор}} + U_{\text{си на с}}$ в области перекрытия для пологой части выходной характеристики.

Первое приближение, когда $\epsilon_y \ll \epsilon_x$, приводит к тому, что выражение для ВАХ будет иметь приемлемую точность только для крутого участка стоковой характеристики при $U_{\text{си}} < U_{\text{си на с}}$. Приближение дрейфовой скорости выполняется при большой

концентрации носителей в канале, т. е. когда $U_{\text{зи}} - U_{\text{пор}} \gg \Phi_T$. Если $U_{\text{зи}} \approx U_{\text{пор}}$, то основной будет не дрейфовая, а диффузионная составляющая тока. При выводе приближенной ВАХ вычисляется поверхностная плотность заряда электронов в канале из уравнения нейтральности $Q_3 + Q_{\text{пов}} + Q_n = 0$, где Q_3 — плотность заряда в затворе, $Q_{\text{пов}}$ — плотность поверхностного заряда, Q_n — плотность некомпенсированных ионов примеси и подвижных носителей в канале. Используя полученное выражение, можно определить сопротивление канала как функцию напряжения U и далее получить формулу для ВАХ.

Для упрощения анализа и расчета на практике применяются аппроксимации ВАХ, простейшая из которых имеет вид

$$I_C = K[(U_{\text{зи}} - U_{\text{пор}})U_{\text{си}} - U_{\text{си}}^2/2] \quad (6.5)$$

при $U_{\text{си}} < U_{\text{си на с}} \approx U_{\text{зи}} - U_{\text{пор}}$,

$$I_C = (K/2)(U_{\text{зи}} - U_{\text{пор}})^2 \quad (6.6)$$

при $U_{\text{си}} \geq U_{\text{си на с}}$,

$$\text{где } K = b\mu_n \epsilon_0 \epsilon_d / (Ld_d), \quad (6.7)$$

где μ_n — подвижность электронов; d_d — толщина подзатворного диэлектрика для МДП-транзистора или толщина n -слоя под затвором для ПТШ или ПТУП; L — длина затвора (канала); b — ширина затвора; ϵ_d — относительная диэлектрическая проницаемость диэлектрика в МДП-транзисторе.

Поясним физический смысл $U_{\text{пор}}$ для ПТШ и ПТУП. Для ПТШ при напряжении на затворе, равном пороговому, граница обедненного слоя достигает подложки. Толщина канала и, следовательно, ток стока становятся равными нулю и при $U_{\text{зи}} < U_{\text{пор}}$ транзистор закрыт. Пороговое напряжение в этом случае, при условии, что $L_{\text{об}}(U_{\text{пор}}) = d_0$ (d_0 — толщина канала), может быть вычислено по формуле

$$U_{\text{пор}} = \Phi_{03} - [qN_K d_0^2 / (2\epsilon_0 \epsilon)], \quad (6.8)$$

где Φ_{03} — равновесная высота потенциального барьера контакта Шоттки; N_K — концентрация атомов примеси в канале; ϵ — относительная диэлектрическая проницаемость полупроводника. Аналогичная ситуация будет наблюдаться и для ПТУП.

Ток стока I_C в соответствии с выражением (6.5) достигает максимума при $U_{СИ} = U_{СИ\text{нас}}$, которое можно приближенно найти, дифференцируя указанное выражение и приравнявая производную нулю. Формулы (6.5) и (6.6) задают простую математическую модель полевых транзисторов. Параметры K , $U_{пор}$ подбираются из условия наилучшего совпадения расчетных ВАХ с экспериментальными.

Для учета эффекта модуляции длины канала в режиме насыщения необходимо в выражение (6.7) подставить вместо L эффективную длину канала $L' = L - \Delta L$ (см. рис. 6.4, *з*), где длина перекрытого участка канала ΔL оценивается по формуле

$$\Delta L = \sqrt{2\varepsilon_0\varepsilon(U_{\text{нас}} - U_{СИ\text{нас}})/(qNa)}. \quad (6.9)$$

Ввиду того что в формулу (6.7) для коэффициента K следует вместо L подставлять эффективную длину канала L' , нелинейно зависящую от напряжения $U_{СИ}$, на практике часто вместо формулы (6.6) используют ее линейную аппроксимацию в режиме насыщения

$$I_C = (K/2)(U_{ЗИ} - U_{пор})^2[1 + \lambda'(U_{СИ} - U_{СИ\text{нас}})], \quad (6.10)$$

где K уже не зависит от длины перекрытого участка канала ΔL , а λ' — коэффициент модуляции длины канала.

6.5. Полевой транзистор как линейный четырехполюсник. Параметры транзисторов. Эквивалентные схемы

В режимах работы с малыми амплитудами любой тип ПТ, как и биполярный транзистор, можно представить в виде линейного четырехполюсника. Из-за высокого входного сопротивления полевых транзисторов наиболее подходящей как с позиций измерений, так и использования является система y -параметров, смысл которых дан в п. 4.4 (см. формулы (4.26) и обозначения к ним).

Поскольку рассмотрение ведется для переменных сигналов из-за дифференциального характера y -параметров, то целесообразно сначала рассмотреть эквивалентные схемы ПТ, что позволит наглядно проиллюстрировать значения параметров транзисторов и их частотные свойства.

Малосигнальная эквивалентная схема МДП-транзистора показана на рис. 6.8, *а*. Источники тока на этом рисунке $SU_{ЗИ}$ и $S_{П}U_{ПИ}$ моделируют усилительную способность транзистора, где

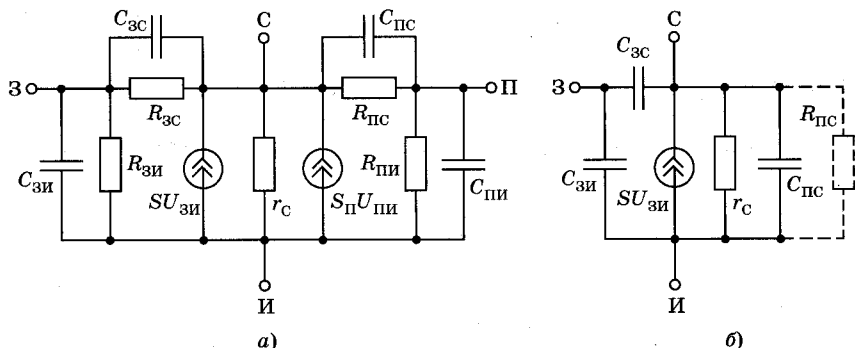


Рис. 6.8

$S_{\Pi} = \left. \frac{dI_C}{dU_{\Pi И}} \right|_{U_{С И}, U_{З И} = \text{const}}$. Сопротивления $R_{З И}$ и $R_{З С}$ определяются

сопротивлением диэлектрика затвора относительно истока и стока. Поскольку эти сопротивления очень велики ($\sim 10^{13} \dots 10^{14}$ Ом), то ими обычно пренебрегают. На рис. 6.8, а обозначено: $R_{П И}$ и $R_{П С}$ — сопротивления при обратном включении $p-n$ -переходов истока и стока; $C_{П И}$ и $C_{П С}$ — барьерные емкости тех же переходов, которые обычно составляют десятые доли пФ; $C_{З И}$ и $C_{З С}$ — емкости перекрытия $C_{\text{пер}}$ металлического электрода затвора относительно областей истока и стока, показанные на рис. 6.10. Помимо емкости перекрытия, емкость $C_{З И}$ часто включает и емкость затвор—канал ($C_{З К}$), т. е. $C'_{З И} = C_{З К} + C_{\text{пер}}$; r_c — сопротивление канала. Если исток соединен с подложкой, то $R_{П И}$ и $C_{П И}$ оказываются замкнутыми и источник тока $S_{\Pi} U_{\Pi И}$ отсутствует.

При частоте сигнала много меньшей предельной частоты крутизны f_S и при исключении сопротивлений диэлектрика $R_{З И}$ и $R_{З С}$ эквивалентная схема существенно упрощается и принимает вид, изображенный на рис. 6.8, б. Упрощенная эквивалентная схема ПТУП и ПТШ представлена на рис. 6.9. Она во многом аналогична схеме на рис. 6.8, б с той разницей, что на рис. 6.9 емкости $C_{З И}$ и $C_{З С}$ определяются емкостями обедненных

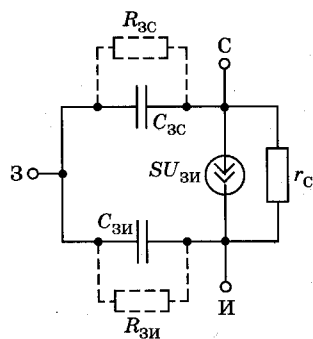


Рис. 6.9

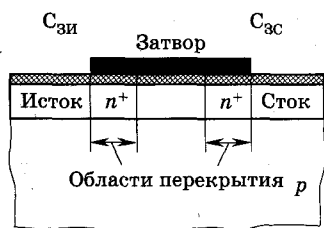


Рис. 6.10

слоев соответствующих электрических переходов, а не подзатворным диэлектриком, как у МДП-транзисторов. Природа $C_{зи}$ и $C_{зс}$ для МДП-транзисторов поясняется на рис. 6.10. Технологически не удается выполнить электрод затвора точно между слоями n^+ стока и истока, тогда между краями затвора и этими слоями образуются емкости перекрытия $C_{зи}$ и

$C_{зс}$. Инерционность полевых транзисторов по отношению к быстрым изменениям управляющего напряжения $U_{зи}$ обусловлена перезарядкой емкости затвора и межэлектродных емкостей.

Используя приведенные упрощенные эквивалентные схемы, получим приближенные значения y -параметров для гармонических сигналов малой амплитуды. Для схемы с общим истоком (см. рис. 6.9, б) входная проводимость на переменном сигнале с частотой ω равна

$$y_{11и} = \left. \frac{dI_3}{dU_{зи}} \right|_{U_{си}U_{пи} = \text{const}} = -j\omega C_{зи}. \text{ Здесь емкость } C_{зи} \text{ прибли-$$

зительно равна емкости затвор — канал.

Остальные параметры определяются формулами:

$$y_{12и} = \left. \frac{dI_3}{dU_{си}} \right|_{U_{зи}U_{пи} = \text{const}} = -j\omega C_{зс} \text{ — проводимость обратной пе-$$

редачи;

$$y_{21и} = \left. \frac{dI_c}{dU_{зи}} \right|_{U_{си}U_{пи} = \text{const}} = S \text{ — проводимость прямой передачи,}$$

определяющая статическую крутизну сток-затворной характе-

ристики транзистора $S = \left. \frac{dI_c}{dU_{зи}} \right|_{U_{си}U_{пи} = \text{const}} ;$

$$y_{22и} = \left. \frac{dI_c}{dU_{си}} \right|_{U_{зи}U_{пи} = \text{const}} = (1/R_i) - j\omega C_{сп} \text{ — выходная проводи-$$

мость, $R_i = \left. \frac{dU_{си}}{dI_c} \right|_{U_{зи}U_{пи} = \text{const}} \text{ — внутреннее сопротивление, } C_{сп} \text{ —}$

барьерная емкость стокового перехода (перехода сток—подложка МДП-транзисторов).

В приведенных формулах буква «И» в индексах обозначает исток.

На практике в основном используются такие малосигнальные параметры, как S , R_i и μ_y . Последний параметр $\mu_y = \left. \frac{dU_{СИ}}{dI_{ЗИ}} \right|_{I_C = \text{const}}$ является коэффициентом усиления транзистора по напряжению. Он может быть выражен через крутизну S и внутреннее сопротивление R_i :

$$\mu_y = SR_i. \quad (6.11)$$

Крутизна сток-затворной характеристики S для пологого участка стоковой характеристики может быть получена дифференцированием выражения (6.6) по $U_{ЗИ}$ в следующем виде:

$$S = K(U_{ЗИ} - U_{пор}). \quad (6.12)$$

Из формулы (6.12) следует, что при $U_{ЗИ} - U_{пор} = 1$ В параметр K численно равен крутизне, поэтому K называется удельной крутизной. Используя последнее выражение и формулу (6.6), установим связь крутизны с рабочим током:

$$S = \sqrt{2KI_C}. \quad (6.13)$$

Реальные значения S для полевых транзисторов составляют от десятых долей до нескольких мА/В.

Внутреннее сопротивление $R_i = r_C$ на пологом участке ВАХ обусловлено зависимостью длины канала от стокового напряжения. Увеличение напряжения $U_{СИ}$ сопровождается возрастанием ширины стокового перехода ΔL и соответственно уменьшением длины канала L' , при этом удельная крутизна K и ток стока тоже увеличиваются. Это явление подобно эффекту Эрли (см. гл. 4). Поэтому сопротивление R_i МДП-транзистора определяется выражением, подобным формуле для коллекторного сопротивления r_K . Зависимость $R_i = r_C$ от I_C такая же, как и зависимость r_K от I_K у биполярных транзисторов.

Из стоковых характеристик видно, что R_i тем больше, чем выше $U_{СИ}$. При $U_{СИ} = 0$ получается наименьшее значение внутреннего сопротивления $R_i = R_{i0}$, где

$$R_{i0} = 1/[K(U_{ЗИ} - U_{пор})] = 1/S. \quad (6.14)$$

Наибольшего значения сопротивление R_i достигает в пологой части стоковой характеристики, где оно составляет десятки и сотни кОм.

Помимо рассмотренных параметров, тесно связанных с полевыми транзисторами как четырехполюсниками, на практике имеют большое значение и такие параметры, как обратные токи истокового и стокового переходов и обратные токи затвора в ПТУП, а также напряжение пробоя подзатворного диэлектрика и рассмотренные выше паразитные емкости транзисторов.

6.6. Частотные и импульсные свойства полевых транзисторов

Как и в биполярных транзисторах, частотные свойства полевых транзисторов определяются временем пролета носителей в канале и паразитными емкостями, присущими конкретному типу и конкретной структуре транзистора. Среднее время пролета $t_{пр}$ электронами n -канала складывается из времени пролета канала L' и участка перекрытия ΔL .

В n -канале электроны перемещаются со средней дрейфовой скоростью

$$v_{др} = \mu_n \mathcal{E} = \mu_n (U_{зи} - U_{пор}) / L, \quad (6.15)$$

где средняя напряженность \mathcal{E} продольного электрического поля в канале дается формулой $\mathcal{E} = (U_{зи} - U_{пор}) / L$, а $L = L' + \Delta L$.

В области перекрытия электроны перемещаются со скоростью насыщения $v_{нас}$, в результате общее среднее время пролета частиц от истока до стока будет равно

$$t_{пр} = L' / v_{др} + \Delta L / v_{нас} \quad (6.16)$$

или с учетом (6.15)

$$t_{пр} = (L')^2 / [\mu_n (U_{зи} - U_{пор})] + \Delta L / v_{нас}. \quad (6.17)$$

Наличие времени пролета $t_{пр}$ приводит к зависимости крутизны S от времени и частоты (см. [34]), т. е. S становится комплексной величиной вида

$$\dot{S} = S / [1 + j(f/f_S)], \quad (6.18)$$

где $f_S = 1 / (2\pi t_{пр})$ — предельная частота крутизны, на которой $|S|$ уменьшается в $\sqrt{2}$ раза при $U_{си} = \text{const}$ по сравнению со статической крутизной S . При $f \ll f_S$ крутизна $\dot{S} = S \approx \text{const}$. В мало-сигнальной модели полевого транзистора время пролета моделируется постоянной времени $t_{пр} = R_i C_{зк}$, где $C_{зк}$ — емкость затвор—канал, не показанная на рис. 6.8.

Рассмотрим роль емкости затвор—сток $C_{ЗС}$, которая включена в цепь обратной связи. Полная входная емкость $C_{вх}$ определяется емкостями $C_{ЗИ}$ и $C_{ЗС}$. При наличии усиления в каскаде емкость $C_{ЗС}$ сильно увеличивает входную емкость. Ток, протекающий через конденсатор обратной связи $C_{ЗС}$, создает дополнительное напряжение на затворе, которое складывается с входным напряжением, т. е. возникает обратная связь по напряжению, которая на высоких частотах может привести к самовозбуждению усилительного каскада. Выходная емкость транзистора, включенного в усилительный каскад при наличии емкости нагрузки $C_{н}$, будет равна $C_{вых} = C_{СИ} + C_{н}$, где емкость $C_{СИ} = C_{СП}$, т. е. $C_{СП}$ в основном определяется емкостью обратносмещенного $p-n$ -перехода сток—подложка, но поскольку в рассмотренных условиях подложка соединена с истоком, то можно считать что емкость сток—исток и сток—подложка эквивалентны. Анализ показывает (см. [1], п. 5.7), что граничная частота усилителя $f_{гр}$ связана с предельной частотой крутизны соотношением

$$f_{гр} = f_S C_{ЗК} / C_{вых} = S / (2\pi C_{вых}). \quad (6.19)$$

Здесь $f_{гр}$ — частота, на которой модуль коэффициента усиления по напряжению равен 1, а емкость $C_{ЗК} = C_{ЗИ} - C_{пер}$ ($C_{пер}$ — емкость перекрытия затвор—исток).

Таким образом, если $C_{вых} \gg C_{ЗИ}$, то $f_{гр} \ll f_S$. Обычно эти соотношения справедливы для дискретных транзисторов. В интегральных схемах $C_{вых}$ и $C_{ЗИ}$ могут быть соизмеримы и $f_{гр}$ приближается по значению к f_S . Упрощенную модель (эквивалентную схему на рис. 6.8, б) в этом случае применять нельзя.

Импульсный режим работы транзистора широко применяется в цифровых устройствах, преимущественно в полевых транзисторах интегральных схем, и будет рассмотрен в гл. 9 на примере логического элемента — инвертора.

6.7. Разновидности полевых транзисторов. Силовые комбинированные транзисторы

К достижениям силовой электроники последних лет относятся разработки таких новых типов транзисторов, как транзисторы со статической индукцией (СИТ и БСИТ) и биполярные транзисторы с изолированным затвором (БИТЗ). Эти типы транзисторов могут коммутировать токи свыше 500 А и напряжения до

2 кВ. В отличие от тиристорov (см. гл. 6) эти приборы имеют лучшее управление, высокое быстродействие и малое потребление тока по цепи управления.

Биполярные транзисторы с изолированным затвором (БИТЗ) представляют собой удачное сочетание входного полевого транзистора с изолированным затвором и вертикальным каналом с выходным биполярным $n-p-n$ -транзистором. Имеются много разновидностей таких приборов, однако наибольшее распространение получили приборы, которые в зарубежной литературе имеют название Insulated Gate Bipolar Transistor (IGBT).

Структура IGBT включает два биполярных транзистора — $n-p-n$ (VT_1) и $p-n-p$ (VT_2) и полевой транзистор с изолированным затвором VT (рис. 6.11, а).

Как видно из приведенного рисунка, ток коллектора I_{K1} транзистора VT_1 влияет на ток базы транзистора VT_2 , а коллекторный ток I_{K2} транзистора VT_2 определяет ток базы VT_1 . В результате структура из двух транзисторов VT_1 и VT_2 имеет глубокую внутреннюю положительную обратную связь.

Вводя коэффициенты передачи тока α_1 для транзистора VT_1 и α_2 для транзистора VT_2 , получим, что $I_{K2} = I_{Э2}\alpha_2$, $I_{K1} = I_{Э1}\alpha_1$ и,

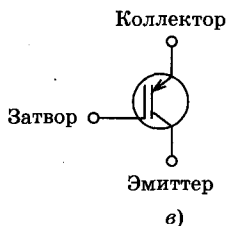
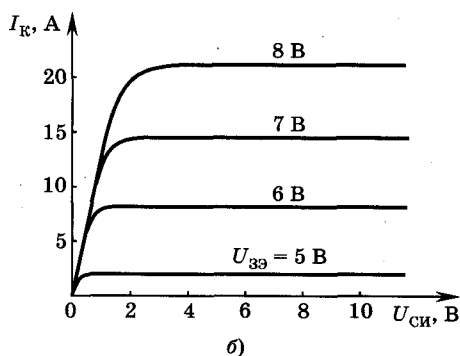
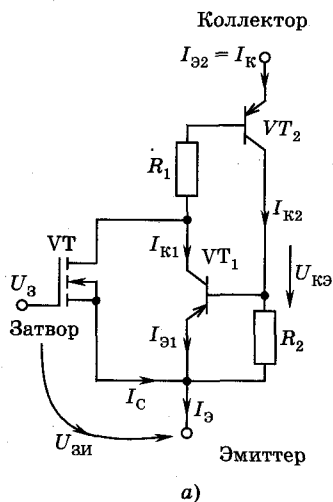


Рис. 6.11

поскольку токи I_C , I_{K1} и I_{K2} текут по параллельным ветвям, суммируясь в точке, откуда вытекает ток $I_{\text{Э}}$, то $I_{\text{Э}} = I_{K1} + I_{K2} + I_C$ (см. рис. 6.11, а). В результате ток стока I_C , полученный из последнего соотношения, равен $I_C = I_{\text{Э}}(1 - \alpha_1 - \alpha_2)$.

Поскольку ток стока полевого транзистора $I_C = SU_3$, ток IGBT транзистора равен

$$I_K = I_{\text{Э}} = SU_3 / (1 - \alpha_1 - \alpha_2) = S_3 U_3,$$

где $S_3 = S / (1 - \alpha_1 - \alpha_2)$ — эквивалентная крутизна биполярного транзистора с изолированным затвором (IGBT), которая может принимать большие значения при $\alpha_1 + \alpha_2 \rightarrow 1$. Значения α_1 и α_2 можно изменять за счет сопротивлений R_1 и R_2 (см. рис. 6.11, а). На рис. 6.11, б приведены ВАХ одного из транзисторов типа IGBT, где значение крутизны достигает 15 А/В.

Достоинством БТИЗ является значительное снижение падения напряжения на замкнутом транзисторном ключе; это объясняется тем, что в режиме насыщения последовательное сопротивление R_2 шунтируется двумя насыщенными транзисторами VT_1 и VT_2 , включенными последовательно. Условное обозначение биполярного транзистора с изолированным затвором приведено на рис. 6.11, в, где отражается гибридность этого прибора, включающего элементы условного обозначения полевого и биполярного транзисторов.

Следующей разновидностью приборов, сочетающих свойства полевых и биполярных транзисторов, являются статические индукционные транзисторы (СИТ). СИТ [35] представляют собой полевые транзисторы с управляющим $p-n$ -переходом. Они работают в режиме полевого транзистора, когда на затвор подано обратное напряжение, и в режиме биполярного транзистора, когда на затвор подано положительное смещение и затвор выполняет роль базы биполярного транзистора. По сравнению с биполярными транзисторами СИТ имеют лучшее быстродействие из-за лучшего рассасывания неосновных носителей, появляющихся в канале при прямом смещении $p-n$ -перехода (затвора). Оно обусловлено тем, что в отличие от биполярного транзистора, обратное напряжение на затворе может достигать 30 В. Время включения СИТ практически не зависит от режима работы и составляет 20...25 нс при задержке не более 50 нс. На этапе выключения происходит рассасывание накопленных в открытом состоянии неосновных носителей,

по аналогии с биполярным транзистором. В результате возникает задержка выключения на время от 20 нс до 5 мкс.

В обычных СИТ при нулевом напряжении на затворе канал находится в хорошо проводящем состоянии (нормально открытое состояние). При подаче отрицательного (обратного) напряжения между затвором и истоком проводимость сильно снижается и транзистор переходит в непроводящее состояние. Нормально открытое состояние при отсутствии управляющего сигнала затрудняет применение СИТ в качестве ключа. Этому недостатка лишены БСИТ (биполярные СИТ), в которых напряжение отсечки технологическими приемами сведено к нулю. БСИТ при отсутствии напряжения на затворе заперты аналогично биполярным транзисторам.

СИТ и БСИТ уступают БТИЗ по быстродействию и мощности управления. К достоинствам СИТ следует отнести очень малое сопротивление канала в открытом состоянии ($0,1 \dots 0,025 \text{ Ом}$).



Контрольные вопросы

1. Каковы классификация и устройство полевых и МДП-транзисторов?
2. Как происходит формирование канала в полевых и МДП-транзисторах?
3. Управление характеристиками каналов в полевых и МДП-транзисторах.
4. Объяснить различные ВАХ МДП-транзисторов.
5. Моделирование полевых транзисторов.
6. ВАХ транзисторов с управляющим $p-n$ -переходом (ПТУП) и контактом металл—полупроводник (ПТШ).
7. Параметры полевых транзисторов (ПТ).
8. Эквивалентные схемы и высокочастотные свойства полевых транзисторов.
9. Разновидности ПТ и их параметры.
10. Силовые комбинированные транзисторы и их свойства.



ИНТЕГРАЛЬНЫЕ СХЕМЫ

Глава 7

АКТИВНЫЕ И ПАССИВНЫЕ ЭЛЕМЕНТЫ ИНТЕГРАЛЬНЫХ СХЕМ

7.1. Общие вопросы. Термины и определения

Микроэлектроника — это раздел электроники, включающий создание и принципы применения качественно нового типа электронных приборов, называемых *интегральными микросхемами* (схемами).

Характерной особенностью микроэлектроники является единство физических, конструктивно-технологических и схемотехнических аспектов.

Микроэлектроника является естественным этапом развития электроники, характеризующимся непрерывным усложнением функций, выполняемых электронной аппаратурой, а также требованиями обеспечения электронными приборами высокой надежности, малых габаритов, массы, малой потребляемой мощности и т. д.

Интегральная микросхема (*интегральная схема* — ИС) как электронный прибор является совокупностью большого количества таких взаимосвязанных компонентов, как транзисторы, диоды, конденсаторы, резисторы и т. д., изготовленных одновременно в едином технологическом цикле на единой подложке. ИС выполняет определенную функцию преобразования информации.

Выполняемые ИС функции являются существенно более сложными по сравнению с функциями отдельных компонентов (транзисторов, диодов и т. д.), которые называются *элементами ИС*. Элементы ИС по своим характеристикам и параметрам отличаются от дискретных приборов и компонентов.

Для изготовления интегральных схем используются групповой метод производства, планарная технология и заключительная операция — корпусирование. Основные технологические методы и приемы, используемые в микроэлектронике, включают следующие операции: подготовительный технологический этап, эпитаксию, термическое окисление, легирование, травление, технику масок, нанесение тонких пленок, металлизацию и сборочные операции. Большинство из перечисленных технологических этапов (операций) кратко описаны в гл. 3 (п. 3.1.3).

Групповой метод производства состоит в том, что на одной полупроводниковой пластине одновременно изготавливается большое количество ИС (иногда одновременно могут обрабатываться десятки таких пластин). После проведения большинства из указанных выше операций пластина разрезается на отдельные кристаллы (*чипы*), каждый из которых является ИС.

При *планарной технологии* все элементы и их составляющие создаются в ИС через плоскость (поверхность).

Корпусирование как финальная операция изготовления ИС заключается в размещении ИС в корпусе с присоединением контактных площадок к выводам ИС.

Все интегральные схемы можно разделить на четыре типа: полупроводниковые, пленочные, гибридные и совмещенные.

В настоящее время различают следующие *полупроводниковые ИС*: биполярные, МДП (МОП — металл—окисел—полупроводник) и БИМОП. В БИМОП интегральных схемах комбинируются биполярные и МОП ИС.

В полупроводниковой ИС все элементы изготавливаются в приповерхностном слое полупроводниковой подложки.

В *пленочных ИС* элементы формируются в виде разного рода пленок, нанесенных на поверхность диэлектрической подложки. Различают *тонкопленочные ИС* (толщина пленок $\leq 1...2$ мкм) и *толстопленочные ИС* (толщина пленок $\geq 10...20$ мкм).

В *гибридных ИС* комбинируются пленочные пассивные элементы и дискретные активные элементы (транзисторы, диоды), смонтированные на одной общей диэлектрической подложке. В *совмещенных ИС* активные элементы изготовлены в полупро-

водниковом кристалле, а пассивные нанесены в виде пленок на поверхность кристалла, которая предварительно изолирована.

Характерной особенностью полупроводниковых ИС является отсутствие катушек индуктивности и трансформаторов. Кроме того, элементы биполярной ИС необходимо изолировать друг от друга, чтобы исключить их взаимодействие. Размеры кристаллов полупроводниковых ИС достигают 20×20 мм.

Функциональная сложность ИС характеризуется *степенью интеграции*, которая определяется количеством элементов (N) на кристалле. По степени интеграции различают следующие виды ИС:

$N \leq 100$ — интегральная схема;

$100 < N \leq 1000$ — ИС средней степени интеграции (СИС);

$1000 < N \leq 10^5$ — большая ИС (БИС);

$N > 10^5$ — сверхбольшая ИС (СВИС).

Другим показателем сложности ИС является *плотность упаковки* — количество элементов на единицу площади кристалла. Этот показатель в настоящее время приблизительно равен 1000 элементов на мм^2 .

Подчеркнем особенности ИС как нового типа электронных приборов.

Подобно дискретным приборам ИС представляет собой единую конструкцию, выполняет определенную функцию, удовлетворяет определенным требованиям при испытаниях и эксплуатации, поэтому ИС является специфическим типом электронных приборов.

Главной особенностью ИС как электронного прибора является самостоятельное выполнение законченных сложных функций в отличие от других электронных приборов (транзисторов, диодов, электронных ламп, за исключением электровакуумных приборов СВЧ-диапазона и т. д.), требующих наличия многих различных компонентов для выполнения аналогичных функций.

В ИС повышение функциональной сложности не ухудшает, а часто улучшает основные эксплуатационные показатели (надежность, стоимость, срок службы и т. д.). Количество технологических операций по изготовлению ИС не сильно превышает число операций при изготовлении отдельного транзистора, поэтому стоимость одного элемента ИС в сотни и тысячи раз меньше по сравнению со стоимостью дискретного компонента. Помимо этого, повышение надежности достигается отсутствием в ИС паяных и сварных соединений, присущих дискретным приборам.

Еще одной характерной особенностью ИС является предпочтительное использование, в отличие от дискретной техники, активных элементов, а не пассивных. В ИС задается не стоимость элемента, а схемы в целом, поэтому на кристалле размещается по возможности как можно большее число элементов с минимальной площадью. В ИС транзисторы и диоды имеют меньшую площадь по сравнению с пассивными элементами.

Поскольку в ИС отдельные элементы располагаются на очень малом расстоянии друг от друга (доли или единицы мкм), то различие электрофизических характеристик материала крайне незначительно, что определяет малый разброс параметров у смежных элементов, т. е. их параметры взаимосвязаны. Эта взаимосвязь не нарушается при изменении температуры, что повышает температурную стабильность ИС.

7.2. Электрическая изоляция элементов полупроводниковых ИС

Монолитная интегральная схема представляет собой небольшой кристалл полупроводника, чаще всего кремния, на котором размещается множество транзисторов и других элементов, которые необходимо изолировать друг от друга. Наиболее распространенной является изоляция p - n -переходами. Пример такой изоляции представлен на рис. 7.1, где показана структура $n^+ - p - n^+$ -транзистора в составе ИС, изолированного от соседних элементов p - n -переходом. Транзистор $n - p - n$ -типа создан в эпитаксиальном n -слое, который, в свою очередь, сформирован на подложке p -кремния. Особенности такого $n^+ - p - n^+$ -транзистора со скрытым n^+ -слоем будут описаны в следующем параграфе. Изолирующий p - n -переход создается путем диффузии акцепторов на глубину эпитаксиального слоя. В результате весь эпитаксиальный слой разбивается на отдельные n -области, изолированные посредством p -областей. Для надежной изоляции необходимо, чтобы p - n -переход между диффузионным p -слоем и n -эпитаксиальным слоем был смещен в обратном направлении. Однако эта изоляция не идеальна, так как между изолированными n -областями и p -подложкой существует ток утечки ~ 1 нА/мм² (при $T = 300$ К) и паразитная барьерная емкость изолирующего p - n -перехода. Эта емкость ($C_{бар}$) зависит от уровней легирования соответствующих областей и от напря-

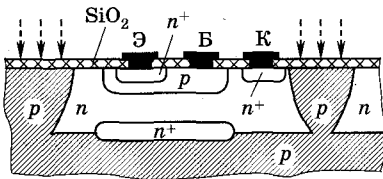


Рис. 7.1

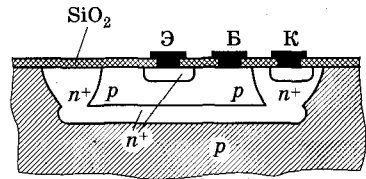


Рис. 7.2

жения смещения (при $U_{обр} = 10$ В, $C_{бар} \approx 100$ пФ/мм²). Типичное значение емкости $C_{бар}$ составляет величину порядка 1 пФ.

Весьма эффективный способ повышения плотности упаковки ИС состоит в том, чтобы взамен разделительной p^+ -диффузии использовать коллекторную изолирующую диффузию (рис. 7.2). В этом случае на p -подложке создается скрытый диффузионный n^+ -слой, поверх которого осаждается тонкий (≈ 2 мкм) слой p -типа. После этого проводится n^+ -диффузия, глубина которой выбирается таким образом, чтобы верхняя n^+ -область слилась с нижней, образовавшейся в результате обратной диффузии примеси из подложки. Верхняя n^+ -область, контактирующая с выводом коллектора, обеспечивает не только изоляцию транзисторных структур, но и создает глубокую приконтактную область, уменьшающую последовательное объемное сопротивление коллектора, что улучшает параметры интегрального транзистора (см. п. 7.3). Изоляция транзисторов друг от друга осуществляется с помощью обратнорасположенных p — n -переходов, образованных диффузионной n^+ -областью и эпитаксиальным p -слоем. В такой структуре разброс размеров ширины базы больше, чем в транзисторах, изолированных с помощью двойной диффузии (см. рис. 7.1). В результате разброс значений коэффициентов усиления по току больше и, как следствие, имеет место худшее согласование характеристик транзисторов. Кроме того, поскольку n^+ -область контактирует с p -базой, снижается напряжение пробоя коллектор—эмиттер.

Изоляция с помощью диэлектрических слоев является хотя и более трудоемкой, но более совершенной. Процесс формирования диэлектрической изоляции включает нижеследующие технологические стадии. Сначала на подложке n -типа создается диффузионный n^+ -слой, в котором формируется сетка пересекающихся ячеек. При последующем анизотропном травлении образуются V -образные канавки, стенки которых после термического

окисления покрываются оксидным слоем SiO_2 . Далее методом химического осаждения из газовой фазы на поверхность SiO_2 наносится толстый слой поликристаллического кремния. Слой получается поликристаллическим из-за того, что кремний наносится на SiO_2 (диоксид кремния), а не на чистый кремний. После этой операции кремниевые пластины монтируются на полировальном круге поликристаллическим слоем вниз и подложка n -типа осторожно сошлифовывается до вершин V -образных канавок (рис. 7.3). В результате получается матрица участков монокристаллического кремния n -типа, изолированных от поликристаллической кремниевой подложки и друг от друга. Подложка из поликристаллического кремния обеспечивает механическую прочность ИС.

Диэлектрическая изоляция имеет большие преимущества для высоковольтных и радиационно стойких ИС. Диэлектрическая прочность SiO_2 составляет величину порядка 600 В/мкм. В ИС с изолирующими $p-n$ -переходами напряжение пробоя не превышает значений ≈ 50 В.

Под действием ионизирующего излучения в кремнии может возникнуть большое количество избыточных свободных электронов и дырок, которые создают значительные приращения токов утечки $p-n$ -переходов, что сильно ухудшает параметры ИС. При использовании диэлектрической изоляции ИС более устойчивы к воздействию ионизирующего излучения из-за наличия оксидного слоя SiO_2 .

Разновидностью диэлектрической изоляции являются схемы типа «кремний-на-сапфире» (КНС). Из-за различия параметров кристаллических решеток кремния и сапфира на границе их раздела возникают значительные механические напряжения, формирующие структурные дефекты, что снижает время жизни неосновных носителей заряда в кремнии. Это ограничивает применение КНС-структур в биполярных схемах, но КНС с успехом могут быть применены в МОП-схемах, например в конфигурациях КМОП/КНС, поскольку малые паразитные емкости повышают быстродействие схем.

В настоящее время, помимо рассмотренных способов изоляции, широко применяется *изопланарная технология*, являющаяся комбинированным методом. В ее основе лежит локальное сквозное прокисление тонкого (2...3 мкм) эпитаксиального слоя кремния n -типа (рис. 7.4), который в результате оказывается

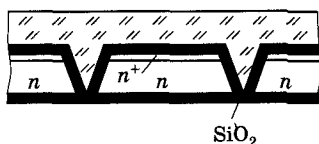


Рис. 7.3

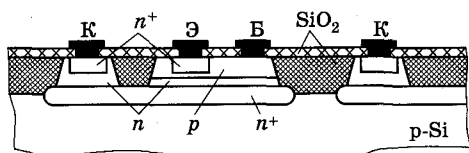


Рис. 7.4

разделенным на отдельные карманы n -типа аналогично разделительной диффузии (см. рис. 7.1), но с тем отличием, что боковые изолирующие слои являются не полупроводниковыми, а диэлектрическими.

Однако данные части карманов по-прежнему разделены встречно включенными p — n -переходами. Заполненные окислом области, отделяющие коллектор от эмиттера и базы, позволяют снизить емкость коллектор—подложка и повысить напряжение пробоя изолирующего p — n -перехода. Главное преимущество изопланарной технологии — повышение плотности упаковки, которое достигается благодаря тому, что эмиттерные n -области и базовые p -области могут непосредственно контактировать с изолирующими участками окисла.

7.3. Особенности биполярных транзисторов ИС

В полупроводниковых ИС основными элементами являются n — p — n -транзисторы. Технология других элементов приспособляется к технологии изготовления n — p — n -транзисторов, чтобы по возможности избежать дополнительных технологических операций.

Особенности биполярных транзисторов ИС рассмотрим на примере транзистора, который изолирован от других элементов ИС с помощью метода разделительной диффузии (см. рис. 7.1). На рис. 7.5, а показано распределение концентрации примесей в различных областях интегрального транзистора в направлении x от поверхности через эмиттер — базу — коллектор — скрытый слой — подложку. На рис. 7.5, б представлено распределение эффективной концентрации примеси $N_{эф} = |N_d - N_a|$, которое является в основном неравномерным, особенно в области базы, что приводит к формированию внутреннего электрического поля (см. гл. 4), т. е. биполярные транзисторы интегральных схем являются дрейфовыми. В соответствии с распределением

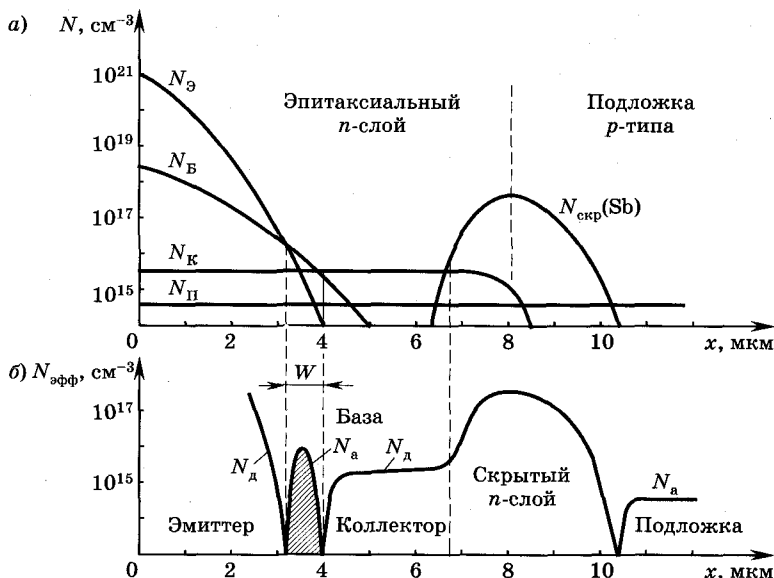


Рис. 7.5

примеси в базе электрическое поле со стороны эмиттерного перехода является тормозящим для электронов, инжектированных в базу из эмиттера, а со стороны коллекторного перехода — ускоряющим. Наличие участка с тормозящим полем в базе приводит к незначительному (на 20...30%) увеличению времени пролета электронов через базу, что, как правило, при расчетах не учитывается.

Интегральный транзистор имеет четырехслойную структуру. Наряду с рабочими электрическими переходами (эмиттерным и коллекторным), имеется паразитный третий $p-n$ -переход между коллекторным n -слоем и подложкой p -типа (см. рис. 7.1). Наличие скрытого n^+ -слоя не вносит принципиальных изменений в структуру. На подложку p -типа интегрального транзистора подается отрицательный потенциал, т. е. напряжение на переходе коллектор—подложка является всегда обратным или, в крайнем случае, равно нулю, поэтому этот переход можно заменить (моделировать) барьерной емкостью $C_{\text{КП}}$ (рис. 7.6, а). Поскольку степень легирования коллекторного перехода меньше легирования базы, то горизонтальное объемное сопротивление коллекторной области ($r_{\text{КК}}$) может сильно сказаться на инерционных свойствах транзистора из-за того,

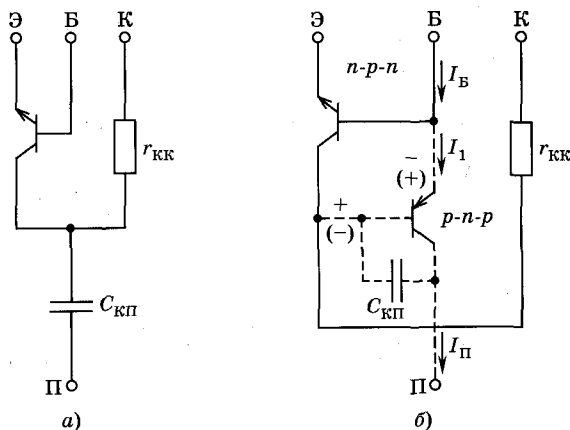


Рис. 7.6

что оно вместе с емкостью $C_{кп}$ образует цепочку с постоянной времени $\tau_к = r_{кк}C_{кп}$, подключенную к активной области транзистора (см. рис. 7.6, а). Наличие этой цепочки, которая моделирует коллектор, является основной особенностью интегрального транзистора по сравнению с дискретным. Типичные значения $r_{кк} \approx 100$ Ом без скрытого n^+ -слоя и $r_{кк} \approx 10$ Ом при его наличии.

Емкость $C_{кп}$ в эквивалентной схеме транзистора подключена параллельно емкости $C_к$ перехода коллектор—база и поэтому складывается с ней, а сопротивление $r_{кк}$ складывается с внешним сопротивлением (сопротивлением нагрузки $R_к$). Эквивалентная постоянная времени для схемы с общей базой может быть записана в следующем виде:

$$\tau_{эф} = \tau_\alpha + (C_к + C_{кп})(r_{кк} + R_к), \quad (7.1)$$

где τ_α определяется временем переноса носителей через базу.

При условии высокопроводящего n^+ -слоя и малого сопротивления $R_к$ вторым слагаемым в формуле (7.1) можно пренебречь, и тогда влияние подложки становится малосущественным.

Рассмотрим роль паразитного $p-n-p$ -транзистора, у которого эмиттер образован базой основного транзистора, база — коллектором основного транзистора, а подложка является коллектором (рис. 7.6, б): $p-n-p$ -транзистор является вертикальным, а основной $n-p-n$ -транзистор — горизонтальным.

Паразитный транзистор при обратносмещенном переходе коллектор—подложка будет работать в активном режиме, если $n-p-n$ -транзистор находится в режиме насыщения. Если же $n-p-n$ -транзистор работает в нормальном активном режиме, то паразитный $p-n-p$ -транзистор окажется в режиме отсечки, и его влияние сводится к увеличению коллекторной емкости основного транзистора на величину $C_{кп}$.

В первом случае через паразитный транзистор будет наблюдаться утечка коллекторного тока в p -слой подложки. Необходимо добиваться, чтобы коэффициент передачи тока $p-n-p$ -транзистора был по возможности малым, что автоматически достигается при сильнолегированном n^+ -скрытом слое, входящем в состав базы этого транзистора. Однако в этом случае $p-n-p$ -транзистор потребляет большой ток I_1 , который течет не через коллекторный переход, а через переход эмиттер—база паразитного транзистора.

В случае диэлектрической изоляции паразитный $p-n-p$ -транзистор отсутствует, однако емкость $C_{кп}$ сохраняется, но ее величина меньше, чем при других видах изоляции.

В микроэлектронном исполнении существуют ряд разновидностей биполярных $n-p-n$ -транзисторов, не имеющих аналогов в дискретном исполнении. Рассмотрим основные особенности некоторых из них.

Многоэмиттерные транзисторы. Этот вид транзисторов составляет основу таких цифровых ИС, как схемы ТТЛ (транзистор—транзисторная логика). Структура и схемные модели многоэмиттерного транзистора (МЭТ) представлены на рис. 7.7, где показаны три эмиттера. Количество эмиттеров может быть 5—8 и более. МЭТ можно часто рассматривать как совокупность отдельных транзисторов с соединенными базами и соединенными

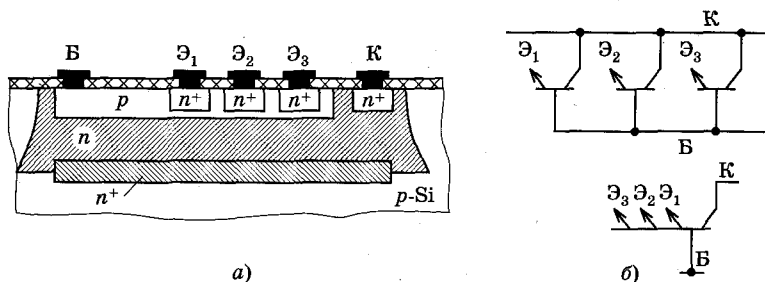


Рис. 7.7

коллекторами. Такое объединение транзисторов на одном кристалле приводит к ряду особенностей.

Одна из особенностей связана с тем, что смежные эмиттеры вместе с разделяющим их p -слоем базы формируют паразитный горизонтальный $n^+—p—n^+$ -транзистор.

При работе МЭТ может складываться ситуация, когда, например, на \mathcal{E}_1 будет отрицательный потенциал, а на \mathcal{E}_2 — положительный. В этом случае паразитный $n^+—p—n^+$ -транзистор работает в активном режиме, и через переход на границе \mathcal{E}_2 , который должен быть закрыт, будет протекать заметный ток. Для устранения или ослабления этого паразитного для МЭТ эффекта расстояние между соседними эмиттерами должно в 3...5 раз превышать диффузионную длину электронов в p -базовой области, т. е. составлять 10...15 мкм. При легировании базы золотом диффузионная длина составляет 2...3 мкм. Поскольку паразитный транзистор может попасть в инверсный режим, то инверсный коэффициент передачи тока α_I должен быть по возможности малым. В противном случае носители, инжектируемые коллектором паразитного транзистора, например \mathcal{E}_2 , достигают эмиттеров, например \mathcal{E}_1 , и аналогично предыдущему случаю через обратносмещенный переход будет протекать заметный паразитный ток, что крайне нежелательно, особенно в цифровых ИС. Для уменьшения α_I в МЭТ технологически увеличивают сопротивление пассивной базы до 200...300 Ом посредством удаления омического базового контакта от активной области транзистора на большее расстояние.

Многоколлекторные $n—p—n$ -транзисторы. Многоколлекторные транзисторы (МКТ) составляют основу схем инжекционной логики И²Л. Они по своей структуре не отличаются от структуры МЭТ. МКТ — это, по сути дела, МЭТ, используемый в инверсном режиме (рис. 7.8).

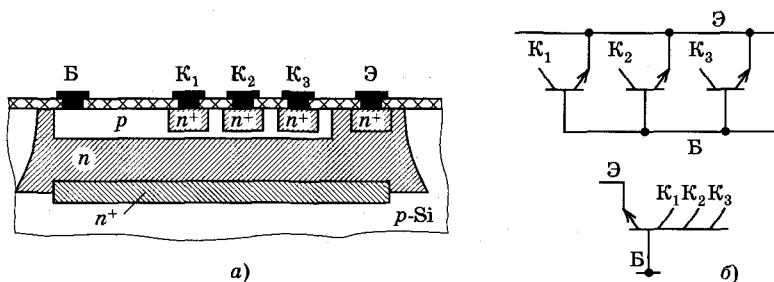


Рис. 7.8

Основным требованием при изготовлении МКТ является получение больших значений нормального коэффициента передачи тока α от общего эмиттера к коллекторам (задача, обратная МЭТ). Это достигается размещением коллекторов по возможности ближе друг к другу и совмещением скрытого n^+ -слоя с базой. В этом случае n^+ -слой, будучи частью эмиттера, обеспечит высокий коэффициент (и ток) инжекции, что увеличивает α . Оба отмеченных способа имеют конструктивно-технологические ограничения. Типичные реализованные значения $\alpha \approx 0,8 \dots 0,9$ ($\beta = 4 \dots 10$), что удовлетворяет нормальному функционированию схем И²Л (см. гл. 9), в которых β на один коллектор должен быть больше единицы. Поэтому число коллекторов не превышает 3...5. Инерционные параметры типичных МКТ имеют следующие величины: время пролета $t_{пр} \approx 5 \dots 10$ пс, предельная частота не более 20...50 МГц, коллекторная емкость по сравнению с МЭТ значительно меньше, и ее влиянием часто можно пренебречь.

Транзистор с барьером Шоттки. Используются в таких цифровых ИС, где транзистор работает в режиме насыщения. Этот тип отличается от обычных $n-p-n$ -транзисторов ИС конструкцией базового контакта (рис. 7.9, а). В этой структуре $n-p-n$ -транзистор сочетается с диодом Шоттки за счет того, что алюминиевая металлизация, обеспечивающая омический контакт с p -слоем базы, продлена в сторону коллектора. В результате образуется выпрямляющий контакт Шоттки между металлизацией и слаболегированной коллекторной областью, что эквивалентно включению диода Шоттки между базой и коллектором (рис. 7.9, б). Диод ограничивает прямое напряжение на коллекторном переходе на уровне 0,3...0,4 В (см. ВАХ контакта металл—полупроводник в гл. 2), что недостаточно для накопления избыточного заряда инжектированных неосновных носителей. Поэтому время рассасывания пренебрежимо мало, что повышает быстродействие цифровых схем.

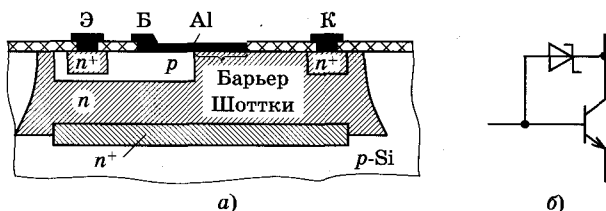


Рис. 7.9

Супербета транзистор. Эти транзисторы имеют сверхтонкую базу с шириной $0,2 \dots 0,3$ мкм, что позволяет получать коэффициенты передачи тока в схеме с ОЭ на уровне $3000 \dots 5000$ и более. Технологический процесс получения сверхтонкой базы является весьма прецизионным, что реализуется на пределе технических и технологических возможностей. Из-за тонкой базы пробивное напряжение у этих транзисторов очень мало ($1,5 \dots 2$ В). При заметно больших напряжениях реализуется эффект смыкания эмиттерного и коллекторного переходов.

Основное применение супербета транзисторов — входные каскады операционных усилителей (гл. 8).

7.4. Транзисторы ИС типа $p-n-p$

Интегральные $p-n-p$ -транзисторы заметно уступают по своим параметрам $n-p-n$ -транзисторам, в частности по коэффициенту усиления и предельной частоте. При прочих равных условиях в $p-n-p$ -транзисторах по сравнению с $n-p-n$ -транзисторами предельная частота примерно в три раза меньше, что связано с существенно меньшей подвижностью дырок по сравнению с электронами. Напомним, что в $p-n-p$ -транзисторах перенос тока осуществляется дырками. Основным структурным вариантом $p-n-p$ -транзистора является горизонтальная структура (рис. 7.10). По сравнению с ранее рассмотренным вертикальным паразитным транзистором горизонтальный имеет существенно лучшие параметры, и технология его изготовления согласуется с технологическим циклом на основе разделительной диффузии. Коллекторный слой, сформированный по этому циклу, охватывает эмиттер со всех сторон, что позволяет собирать инжектированные дырки с боковых частей эмиттерной области. Увеличению коэффициента инжекции способствует повы-

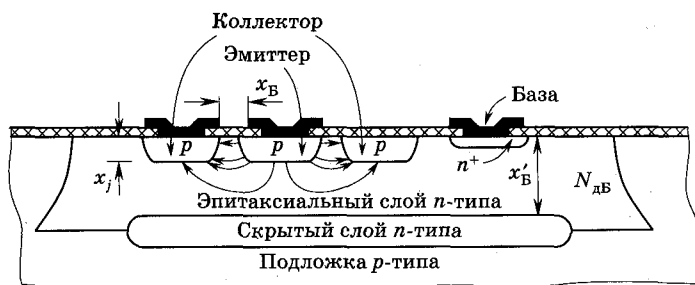


Рис. 7.10

шенная концентрация примеси на приповерхностных боковых участках p -слоев.

В однородной (эпитаксиальный слой) базе $p-n-p$ -транзистора, в отличие от $n-p-n$ -транзистора, нет внутреннего ускоряющего электрического поля, поэтому он является бездрейфовым. Непосредственно в боковом направлении из эмиттера инжектируется лишь малая часть дырок. Основная доля дырок инжектируется вниз в сторону скрытого n^+ -слоя, что увеличивает расстояние, проходимое дырками, а это приводит к возрастанию рекомбинации. В результате коэффициент передачи тока такого транзистора обычно невелик, а время пролета базы велико, емкость между базой и подложкой — большая. При таких параметрах граничная частота мала (менее 50 МГц).

Для увеличения коэффициента передачи эмиттерного тока необходимо, чтобы площадь донной части эмиттерного слоя была меньше площади боковых частей, т. е. эмиттерный слой должен быть узким. Пробивные напряжения эмиттерного и коллекторного переходов одинаковы ($U_{\text{проб}} \sim 30 \dots 50$ В) из-за однотипности эмиттерного и коллекторного слоев.

При кольцевом p -коллекторе достаточно легко сформировать многоколлекторный $p-n-p$ -транзистор. Для этого достаточно разделить кольцевой p -коллектор на необходимое количество частей и сделать от каждой из них отдельные выводы. В таком многоколлекторном транзисторе коэффициент усиления по каждому из коллекторов будет меньше по сравнению с единым коллектором примерно в число раз, равное числу коллекторов.

В вертикальных $p-n-p$ -транзисторах возможно устранение главных недостатков, присущих горизонтальным транзисторам, а именно большой ширины и однородности базы. На рис. 7.11 показана структура вертикального изолированного $p-n-p$ -транзистора. На слаболегированной подложке p -типа

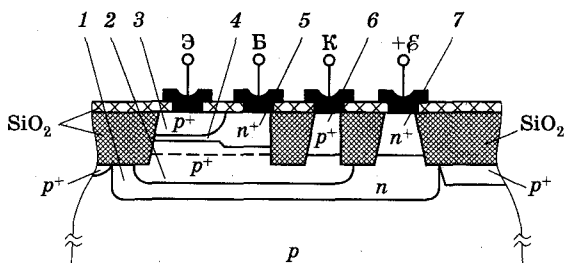


Рис. 7.11

создается скрытый слой 1 n -типа и скрытый слой 2 p^+ -типа. В эпитаксиальном n -слое формируются $p-n-p$ -транзисторы, которые располагаются в карманах p -типа. На рис. 7.11 обозначено: эмиттер — 3, активная база — 4, пассивная база — 5, коллектор 6. На скрытый слой 1 через n^+ -область подается напряжение от источника $+E$, в результате $p-n$ -переходы между слоями 1 и 2 и переход между подложкой и слоем 1 являются обратносмещенными.

Для создания ИС с вертикальными транзисторами требуется значительно больше технологических операций. Коэффициент передачи тока базы вертикального $p-n-p$ -транзистора превышает 100, а достижимая граничная частота — свыше 5 ГГц.

Как отмечалось ранее, для получения $p-n-p$ -транзисторов с улучшенными параметрами используется технология «кремний-на-сапфире — КНС», при которой $p-n-p$ -транзисторы формируются, по существу, отдельно от $n-p-n$ -транзисторов на диэлектрической подложке из сапфира. Ширина базы и уровень легирования в этом случае могут быть оптимизированы, но это достигается значительным усложнением технологии, а следовательно, и удорожанием ИС.

7.5. Интегральные диоды

Для того чтобы сформировать диод с $p-n$ -переходом, необходимо только создать диффузионную область p -типа в пластине n -типа и сделать контакты к верхней и нижней поверхностям пластины. Однако в ИС обычно формируется много диодов и транзисторов, которые должны быть изолированы друг от друга диэлектриками или обратносмещенными $p-n$ -переходами, либо другими способами изоляции (см. п. 7.2). Кроме того, отдельное изготовление диодов в групповом методе производства ИС крайне нерационально. Проще и дешевле изготавливать транзисторы. В ИС в качестве диодов используются либо эмиттерный, либо коллекторный переходы, расположенные в изолирующем кармане. Возможно использование комбинации указанных переходов. В результате диод в ИС представляет собой выполненное тем или иным способом диодное включение интегрального транзистора.

Различные варианты этого включения представлены на рис. 7.12. Кроме показанных там вариантов, в качестве диода можно использовать структуру, не имеющую эмиттерной области, своего рода транзистор без эмиттера. На рис. 7.12 «плюс»

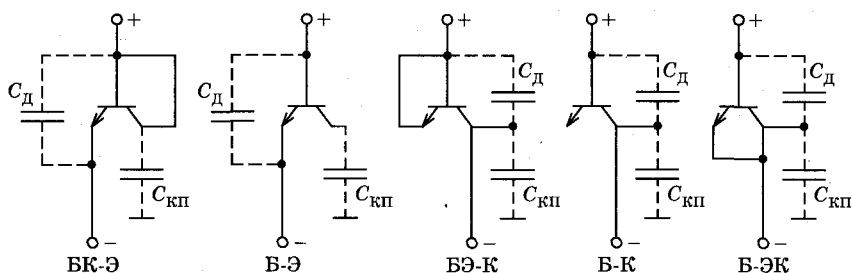


Рис. 7.12

источника подключен к аноду, а «минус» — к катоду. Аноду в буквенных обозначениях внизу рисунка соответствует буква или буквы до черточки, а катоду — после черточки. Типичные параметры интегральных диодов, изображенных на рис. 7.12, приведены в табл. 7.1, где $U_{\text{проб}}$ — напряжение пробоя соответствующего перехода, $I_{\text{обр}}$ — обратный ток, $C_{\text{д}}$ — емкость диода (емкость между анодом и катодом), C_0 — паразитная емкость между подложкой и катодом или анодом, она обычно совпадает с $C_{\text{кп}}$ и шунтирует на «землю» анод или катод; $t_{\text{в}}$ — время восстановления обратного тока, равное времени переключения диода из открытого в закрытое состояние.

Пробивное напряжение $U_{\text{проб}}$ эмиттерного перехода в 5...7 раз меньше, чем коллекторного. Эта особенность присуща всем дрейфовым транзисторам и связана с тем, что эмиттерный переход образован более низкоомными слоями, чем коллекторный, по-

Таблица 7.1
Типичные параметры интегральных диодов

Параметры	Тип диода				
	БК—Э	Б—Э	БЭ—К	Б—К	Б—ЭК
$U_{\text{проб}}$, В	7...8	7...8	40...50	40...50	7...8
$I_{\text{обр}}$, нА	0,5...1	0,5...1	15...30	15...30	20...40
$C_{\text{д}}$, пФ	0,5	0,5	0,7	0,7	1,2
C_0 , пФ	3	1,2	3	3	3
$t_{\text{в}}$, нс	10	50	50	75	100

этому ширина (толщина) эмиттерного перехода много меньше, чем коллекторного. Обратные токи определяются токами термогенерации в кремниевых переходах, поэтому они зависят от объема обедненной области и, следовательно, меньше у тех вариантов включения, где используется только эмиттерный переход, у которого наименьшая площадь и наименьшая толщина.

Емкость диода C_D зависит от площади используемых переходов и максимальна при их параллельном включении (Б—ЭК).

Паразитная емкость C_0 , совпадающая обычно с емкостью $C_{КП}$, минимальна в схеме Б—Э, где емкость коллекторного перехода $C_{КП}$ и емкость C_D соединены последовательно.

Время восстановления (переключения) t_v связано с накоплением и рассасыванием заряда подвижных носителей в базе и коллекторе и определяется эффективным временем жизни неосновных носителей в той области транзистора, где происходит их накопление при протекании прямого тока. В схеме БК—Э транзистор работает в активном режиме, накопление неосновных инжектированных из эмиттера носителей (электронов) происходит в базовой области. Эффективным временем жизни электронов в базе является среднее их время пролета через нее. Оно очень мало ($< 0,1$ нс) из-за малой толщины базы и ускоряющего внутреннего поля базы. Во всех других схемах коллекторный переход смещен в прямом направлении и через него инжектируются дырки из базы и электроны в базу.

Основное накопление носителей (дырок) будет происходить в коллекторе из-за низкой концентрации доноров и большой его толщины. В результате время жизни дырок в коллекторе, которое много больше времени пролета через базу, является большим (10...1000 нс).

Включение БК—Э используется в быстродействующих ИС, когда от диода требуется малое время восстановления обратного сопротивления. В других применениях часто наиболее удобна схема Б—Э и схема без эмиттера, не требующие создания трех выводов и имеющие минимальную площадь на кристалле.

Итак, схема БК—Э обладает минимальным временем восстановления и малым прямым напряжением, а также максимальным обратным сопротивлением, но низким напряжением пробоя. Высокое $U_{проб}$ присуще схемам БЭ—К и Б—К, но у них велико t_v .

7.6. Полевые транзисторы ИС

Полевые транзисторы с управляющим $p-n$ -переходом (ПТУП) и контактом Шоттки (ПТШ) хорошо вписываются в технологию производства биполярных ИС, и поэтому они часто изготавливаются совместно с биполярными транзисторами (БТ). Однако если не использовать дополнительных технологических приемов при совместном изготовлении ПТ и БТ, то толщина канала будет равна ширине базы $n-p-n$ -транзистора (0,3...1 мкм), что приводит к большому разбросу параметров ПТ и малому напряжению пробоя. Поэтому p -слой ПТ изготавливают отдельно от базового p -слоя $n-p-n$ -БТ, получая толщину канала более 1...2 мкм. В этом случае проводят предварительную диффузию p -слоя ПТ до базовой диффузии БТ.

Применение указанных полевых транзисторов в ИС ограничено из-за их малого быстродействия и большой площади, занимаемой на кристалле. Однако большое входное сопротивление и малый уровень шумов позволяют их применять во входных каскадах некоторых типов аналоговых ИС, где в остальных каскадах используются биполярные транзисторы.

В силу сказанного, основное применение в интегральных схемах находят МДП-транзисторы. МДП-транзисторы ИС, как правило, изготавливаются отдельно от биполярных. МДП-транзисторы по сравнению с биполярными занимают существенно меньшую площадь на кристалле, обладают крайне большим входным и выходным сопротивлениями, имеют широкий набор различных типов транзисторов на одном кристалле: со встроенными и индуцированными n - и p -каналами, различными пороговыми напряжениями и т. д. Все это позволяет разрабатывать СВИС с оптимальными параметрами, по степени интеграции намного превосходящие ИС на биполярных транзисторах. Основную роль в современной микроэлектронике играют МДП-транзисторы на основе SiO_2 , которые называются МОП-транзисторами (металл—окисел—полупроводник). Интегральные транзисторы не нуждаются в изоляции, поэтому их структура внешне не отличается от структуры дискретных приборов. Подложка всегда имеет другой тип электропроводимости, чем одинаковые по типу электропроводимости исток и сток, поэтому $p-n$ -переходы исток—подложка и сток—подложка включены встречно. При любой полярности напряжения между истоком и стоком один из $p-n$ -переходов имеет обратное включение и обеспечивает изоляцию.

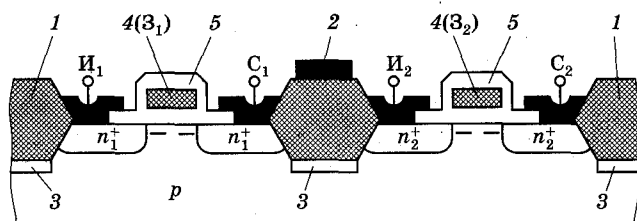


Рис. 7.13

Как показано в гл. 6, существенным фактором, влияющим на быстродействие МДП-транзисторов, являются транзисторные емкости, емкости перекрытия. Для их уменьшения используются затворы 4 (рис. 7.13) из поликристаллического кремния, которые формируются раньше n_1^+ и n_2^+ истоков I_1 и I_2 и стоков C_1 и C_2 . Сформированные поликристаллические затворы играют роль масок при последующем легировании донорами указанных областей истоков и стоков. Использование поликристаллического затвора в виде маски приводит к автоматическому совмещению («самосовмещение») краев затвора с краями истоков и стоков, т. е. практически сводится к нулю перекрытие между ними, а паразитные емкости перекрытия становятся минимальными. Над затворами 3 формируются слои SiO_2 (5), изолирующие затворы от соединительных проводников ИС. Вывод затворов 4 осуществляется на периферии. Для устранения паразитных связей между соседними транзисторами создаются сильнолегированные акцепторами слои 3 (p^+ -слои) под разделительными окисными участками 1, на границе с p -подложкой. Соединительные проводники 2 имеют положительный потенциал относительно подложки, а сама подложка соединяется с минусом источника питания ИС, чтобы $p-n$ -переходы были смещены в обратном направлении. В противном случае появляются паразитные связи и помехи по подложке, что ухудшает параметры ИС, и ее работа становится неустойчивой.

В комплементарных МОП-транзисторах на одном и том же кристалле необходимо изготавливать транзисторы с n - и p -каналами вместе (рис. 7.14). Один из транзисторов должен быть размещен в специальном кармане (на рис. 7.14 транзистор VT_p в кармане 1). Для транзистора VT_p карман должен обладать электронной проводимостью. На вывод n^+ -области 2 подается плюс от источника питания, чтобы обеспечить на $p-n$ -переходах p -канального транзис-

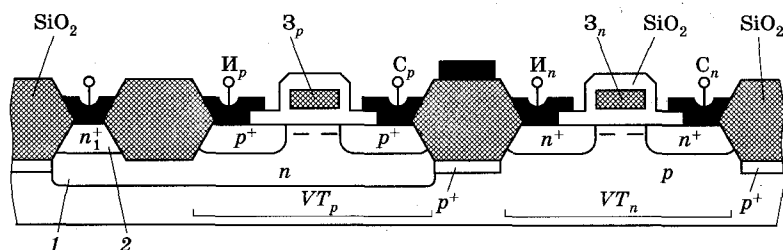


Рис. 7.14

тора обратное смещение. Подложка p -типа подключена к минусу источника электрического питания. КМОП-транзисторы на одной подложке могут быть изготовлены также с помощью КНС (кремний-на-сапфире) технологии. При использовании этой технологии на сапфировой подложке, которая имеет кристаллическую структуру, близкую к кремнию, создаются островки кремния с собственной проводимостью. Вследствие диффузии донорной или акцепторной примеси в одних островках формируются соответственно n -канальные, а в других — p -канальные транзисторы.

Помимо рассмотренной структуры на кремниевой и сапфировой подложке, существуют комплементарные транзисторы, создаваемые с помощью технологии «кремний на диэлектрике» (КНД), при которой в тонких кремниевых пленках, нанесенных на диэлектрик, формируются МОП-транзисторы. Сапфировые подложки имеют высокую стоимость, поэтому используются другие диэлектрические материалы, в частности двуокись кремния SiO_2 (структуры типа «кремний—окисел—кремний»). В таких структурах, как и в КНС-технологии, отсутствуют карманы, а соседние транзисторы для устранения между ними паразитных связей изолированы друг от друга диэлектрическими слоями. При этой технологии транзисторы располагаются на минимальном расстоянии друг от друга, что позволяет повысить степень интеграции. Влияние емкости переходов здесь также снижено. Кроме того, такие транзисторы обладают повышенной радиационной стойкостью (см. разд. 6, гл. 23).

В рассмотренных структурах ИС элементы располагаются в тонком слое у поверхности, что ограничивает рост степени интеграции, поэтому в настоящее время развиваются технологии создания многослойных структур, в которых транзисторы размещаются один под другим в несколько слоев (этажей).

7.7. Пассивные элементы ИС

Полупроводниковые резисторы. В соответствии с технологией изготовления различают два типа интегральных резисторов: *диффузионные и ионно-имплантированные (ионно-легированные)*. Диффузионные резисторы (ДР) изготавливают одновременно с формированием базовой или эмиттерной областей транзистора с применением соответствующих масок. Чаще всего для резисторов используется полоска базового слоя с двумя омическими контактами. Длина и ширина полоскового ДР ограничены. Длина не превышает размеров кристаллов, а ширина определяется возможностями фотолитографии, боковой диффузией, а также допустимым разбросом. Типичные величины максимального сопротивления, достижимые с помощью современной технологии, составляют 20 кОм. Для увеличения максимального сопротивления в 2...3 раза (до 50...60 кОм) используют зигзагообразную конфигурацию при количестве петель, обычно меньшем или равном трем, что определяется площадью, отводимой под ДР на кристалле.

При изменении температуры сопротивления резисторов, расположенных на одном кристалле, изменяются согласно, поэтому при абсолютном разбросе *температурного коэффициента сопротивления* (ТКС) ДР, равном 0,15...0,3%/°С, ТКС для отношения сопротивлений не превышает $\pm 0,01\%$. Разброс значений сопротивления относительно расчетных составляет $\pm(15...20)\%$, а разброс отношений сопротивлений сохраняется с существенно меньшим допуском ($\pm 3\%$). Для получения больших номиналов изготавливают резисторы с меньшей площадью сечения и большим удельным сопротивлением *p-слоя (пинч-резисторы)*, который делают слаболегированным. В этой разновидности ДР максимальное сопротивление может достигать значений 200—300 кОм даже при простейшей полосковой конфигурации. Однако разброс номиналов в пинч-резисторах достигает 50% из-за сильного влияния изменения толщины *p-слоя* при ТКС $\sim 0,3...0,5\%/^{\circ}\text{C}$. ВАХ пинч-резистора подобна ВАХ полевого транзистора при $U_3 = 0$, поскольку у пинч-резисторов n^+ - и *p-слой* соединены друг с другом металлизацией.

Малые значения сопротивлений ($R \leq 100$ Ом) целесообразно получать за счет использования низкоомного эмиттерного слоя, при этом удается получать минимальные номиналы в 3...5 Ом с ТКС $\approx 0,01...0,02\%/^{\circ}\text{C}$.

Ионно-легированные резисторы получают ионной имплантацией примеси. Они имеют структуру, подобную ДР, но глубина имплантированного p -слоя составляет $0,2...0,3$ мкм, что значительно меньше глубины базового слоя. Ионная имплантация обеспечивает любую малую концентрацию примеси в слое. Перечисленные факторы позволяют формировать номиналы сопротивлений в сотни килоом, с ТКС $\approx 3...5\%/^{\circ}\text{C}$ и разбросом сопротивлений меньше $\pm(5...10)\%$.

Характерной особенностью любого интегрального резистора является наличие у него паразитной емкости относительно подложки или изолирующего кармана. Совокупность резистора и паразитной емкости формирует собой распределенную RC -линию. Однако обычно пользуются Π -образными и T -образными эквивалентными схемами с сосредоточенными, а не распределенными постоянными. В простейшем случае эквивалентная схема может выродиться в простую RC -цепочку с постоянной времени $\tau = (1/2)R\bar{C}$, где \bar{C} — усредненная емкость p - n -перехода; постоянной времени τ соответствует граничная частота $f_{\text{гр}} = 1/(2\pi\tau) = 1/(\pi R\bar{C})$. Следовательно, полупроводниковый резистор ИС представляет собой активное сопротивление до частот, заметно меньших $f_{\text{гр}}$, а при $f > f_{\text{гр}}$ его сопротивление является комплексным, что может сильно сказаться на работе ИС. В ИС с диэлектрической изоляцией постоянная времени обычно в несколько раз меньше.

Интегральные полупроводниковые конденсаторы. В конденсаторах полупроводниковых ИС в качестве диэлектрика выступает обедненный слой обратносмещенного p - n -перехода, а в роли обкладок — высоколегированные полупроводниковые области или напыленные металлические пленки. Можно выделить три типа конденсаторов ИС: *диффузионные, МДП и тонкопленочные МДМ.*

Для формирования *диффузионных конденсаторов (ДК)* используются барьерные емкости обратносмещенных p - n -переходов, где хотя бы один из слоев является диффузионным (откуда и название диффузионные). Эмиттерный переход в интегральном биполярном транзисторе обладает наибольшей удельной емкостью (на единицу площади), но малым напряжением и низкой добротностью. Поэтому наиболее часто применяется коллекторный пе-

переход, который в целом формирует конденсаторы с лучшими параметрами. Поскольку емкость ДК зависит от напряжения, то они имеют нелинейную ВФХ. Такие конденсаторы находят применение в параметрических усилителях, умножителях частоты и т. д. Если на переход подать постоянное смещение, величина которого больше амплитуды переменного сигнала, то емкость перехода $C_{бар}$ будет постоянной. Такие конденсаторы пропускают переменные сигналы без искажения. Возможность изменения $C_{бар}$ с изменением напряжения позволяет использовать ДК в качестве управляемой емкости.

МОП-конденсатор принципиально отличается от ДК. Обычно при его создании над эмиттерным n^+ -слоем (нижняя обкладка) выращивается слой тонкого окисла ($\sim 0,1$ мкм). На этот слой напыляется алюминиевая верхняя обкладка конденсатора. Важным преимуществом МОП-конденсатора по сравнению с ДК является возможность его работы при любой полярности напряжения, как и для обычного конденсатора. Емкость МОП-конденсатора, как и ДК, зависит от напряжения на обкладках, поскольку эта емкость представляет последовательное соединение емкости диэлектрика и емкости обедненного слоя, который образуется в приповерхностной области полупроводника или появляется при отрицательном напряжении на Al-пленке. Если приповерхностная область обогащена электронами, что наблюдается при нулевых и положительных напряжениях, то емкость конденсатора определяется диэлектриком и имеет максимальное значение. Влияние обедненного слоя будет меньше при большой концентрации примеси полупроводника, когда емкость обедненного слоя будет максимальной. Именно поэтому в качестве второго электрода используется обогащенный эмиттерный n^+ -слой.

Тонкопленочные МДМ-конденсаторы как элементы совмещенных ИС в качестве металлических обкладок используют алюминий или тантал. В роли диэлектриков выступают окислы этих металлов, в первом случае — Al_2O_3 , а во втором — Ta_2O_5 . Диэлектрическая постоянная танталового окисла на порядок выше, чем у Al_2O_3 и других диэлектриков. Однако танталовые конденсаторы работают только на низких частотах. МДМ-конденсаторы, так же как МОП-конденсаторы, работают при любой полярности приложенного напряжения.

**Контрольные вопросы**

1. Каковы особенности ИС как электронных приборов? Классификация ИС.
2. Основные технологические приемы, используемые при изготовлении ИС. Разновидности ИС и их особенности.
3. Виды изоляции полупроводниковых ИС, достоинства и недостатки.
4. Биполярные $n-p-n$ -транзисторы ИС: структуры, особенности, свойства, эквивалентные схемы, параметры и разновидности.
5. Биполярные $p-n-p$ -транзисторы ИС: структуры, особенности, свойства, эквивалентные схемы, параметры и разновидности.
6. Интегральные диоды, диодное включение транзисторов ИС, параметры диодов.
7. Полевые транзисторы ИС: структуры, свойства, параметры, особенности.
8. Пассивные элементы ИС: структуры, особенности, параметры.

Глава 8**АНАЛОГОВЫЕ ИНТЕГРАЛЬНЫЕ СХЕМЫ****8.1. Общие сведения. Термины и определения**

Все существующие электронные схемы условно можно разделить на 2 класса: цифровые и аналоговые.

Аналоговый сигнал представляет собой непрерывно изменяющуюся во времени электрическую величину (обычно ток или напряжение), которая лежит в допустимом информативном интервале значений в любой момент времени, т. е. выходная величина и входная связаны друг с другом функциональной зависимостью $U_{\text{вых}} = f(U_{\text{вх}})$ [6,34].

Цифровой сигнал обычно характеризуется двумя устойчивыми значениями (максимальным и минимальным), при этом пере-

ход от одного значения к другому происходит в течение короткого временного интервала [6,34].

В основе аналоговых схем лежат простейшие усилительные ступени и каскады, а основу цифровых схем составляют простейшие транзисторные ключи.

На базе усилительных каскадов строятся сложные многокаскадные усилители, стабилизаторы напряжения и тока, модуляторы и детекторы, генераторы непрерывных во времени сигналов и другие схемы.

При работе любой аналоговой схемы наблюдается отклонение (разброс) выходных сигналов $U_{\text{вых}}(t)$ в некотором диапазоне, т. е. $U_{\text{вых}}(t) = U(t) \pm \Delta U(t)$. Источником отклонения $\Delta U(t)$ может быть температурный и временной дрейф параметров элементов схемы, шумы, технологический разброс параметров и т. д. Сложность получения высокой точности воспроизведения характеристик элементов при хорошей их стабильности и минимальных шумах явилась причиной отставания развития аналоговых схем по сравнению с цифровыми ИС на первых этапах становления микроэлектроники. Однако в настоящее время это отставание ликвидировано и аналоговые микросхемы используются в качестве основной элементной базы большинства аналоговых устройств. Это позволило существенно уменьшить габаритные размеры и массу этих устройств, а также потребляемую мощность и повысить точность обработки аналоговой информации. Последнее достоинство обусловлено тем, что в ИС на одной подложке сформирована совокупность элементов с взаимосогласованными характеристиками (принцип взаимного согласования цепей) и однотипные элементы имеют одинаковые параметры и взаимную компенсацию нестабильности параметров во всех диапазонах внешних допустимых воздействий.

Аналоговые ИС можно разделить на **универсальные** и **специализированные**. К универсальным аналоговым ИС относятся матрицы согласованных резисторов, диоды и транзисторы, а также интегральные операционные усилители (ОУ).

Специализированные аналоговые микросхемы выполняют некоторую определенную функцию, как то: перемножение аналоговых сигналов, фильтрацию, компрессию и т. д.

Аналогово-цифровые преобразователи (АЦП) и цифроаналоговые преобразователи (ЦАП) трансформируют аналоговую информацию в цифровую и наоборот. АЦП в основном преобразуют на-

пряжение в цифровой код. Из ЦАП наибольшее распространение получили преобразователи кода в напряжение и кода в ток.

Интегральные СВЧ-микросхемы имеют функциональную, схемотехническую и конструктивно-технологическую специфику. Их развитие стимулируется потребностями радиолокации, телевидения, авиакосмической техники и т. д., требующими массового выпуска малошумящих усилителей для приемных трактов, преобразователей частоты, переключателей СВЧ-сигналов, генераторов, усилителей мощности и т. д.

Интегральным схемам по сравнению с дискретными свойственны отличительные особенности, обусловленные спецификой их технологии. К особенностям аналоговых ИС можно отнести отмеченный ранее принцип взаимного согласования цепей и принцип схемотехнической избыточности, заключающийся в преднамеренном усложнении схемы с целью улучшения ее качества, минимизации площади кристалла и повышения технологичности. В качестве примера можно привести тот факт, что в аналоговых ИС вместо конденсатора, занимающего большую площадь, используют сложные структуры с непосредственными связями.

8.2. Источники стабильного тока, напряжения и опорного напряжения

8.2.1. Источник стабильного тока.

Под *источником стабильного тока* в данном случае понимается источник, который обеспечивает ток в нагрузке, не зависящий от падения напряжения на ней. Такое определение позволяет дать термину расширенное толкование как источника, который вырабатывает изменяющийся во времени (переменный) ток, величина которого, например, остается постоянной в течение определенных отрезков времени, или переменный ток с неизменной амплитудой.

Для получения тока со стабильным значением в источниках применяют либо биполярные транзисторы (БТ) в схеме с ОБ, работающие в активном режиме, либо полевые транзисторы (ПТ), использующие пологую часть выходной ВАХ. В обоих случаях диапазон допустимых напряжений U на нагрузке ограничен потенциалом базы для БТ или потенциалом затвора для ПТ. Для ПТ диапазон допустимых значений U уменьшается на величину порогового напряжения ($U_{\text{зи пор}}$) (см. п. 6.2).

Для многих схем источников тока, а также большинства схем дифференциальных усилителей с активной нагрузкой основой является схема **токового зеркала** (рис. 8.1 без сопротивления R_1). При полной идентичности транзисторов VT_1 и VT_2 ток, текущий через левую часть схемы, является зеркальным отражением тока в правой части, поэтому схема и называется токовым зеркалом. На практике в большинстве случаев влиянием базового тока можно пренебречь и считать, что коллекторные токи транзисторов I_{K1} и I_{K2} равны, т. е. $I_{K1} = I_{K2} \approx I_1 \approx I_2$.

В реальных ИС параметры транзисторов различаются, поэтому в паре транзисторов токового зеркала токи коллекторов не будут точно равны. В результате появится небольшое напряжение смещения (единицы мВ) пары транзисторов.

На рис. 8.1 показана схема простого источника стабильного тока, построенная на основе токового зеркала, в котором ток равен

$$I_1 = (\mathcal{E} - U_{БЭ})/R_1, \quad (8.1)$$

где \mathcal{E} — постоянное положительное напряжение питания схемы (ЭДС источника питания), $U_{БЭ}$ — напряжение между базой и эмиттером транзисторов. Ток I_2 в цепи коллектора второго транзистора равен $I_2 = I_{K2} = I_{K1} \approx I_1$, поэтому расчет этой схемы достаточно прост. При заданном токе источника I_2 и $U_{БЭ}$ и известной (выбранной) ЭДС источника питания \mathcal{E} из (8.1) определяется сопротивление R_1 . Следует обратить внимание, что ток I_2 остается приблизительно постоянным, равным заданному значению в том диапазоне значений $U_{КЭ}$, при котором транзистор VT_2 находится в активном режиме вплоть до напряжения пробоя коллекторного перехода $U_{КЭ\text{ проб}}$ (рис. 8.2). Если принять $U_{нас} \approx 0,2$ В и

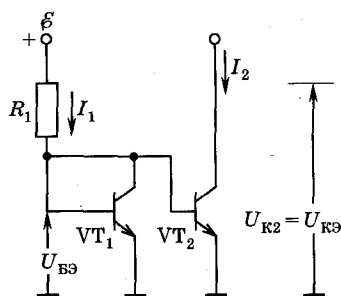


Рис. 8.1

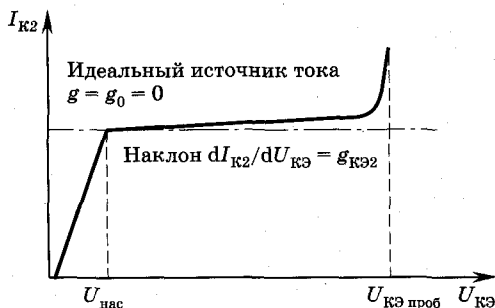


Рис. 8.2

$U_{\text{кЭ проб}} = 50 \text{ В}$, то область напряжений $U_{\text{к2}}$ от $+0,2$ до $+50 \text{ В}$ составляет рабочий диапазон источника постоянного тока.

Область изменения напряжений $U_{\text{кЭ}}$, в которой схема работает практически как источник стабильного тока, иногда называется *диапазоном линейного изменения напряжения* [37].

На рис. 8.2 сплошной линией показана как выходная характеристика источника тока (схема на рис. 8.1), так и характеристика идеального источника стабильного тока (штрихпунктирная прямая). ВАХ источника тока $I_0 = I_{\text{к2}} = f(U_{\text{кЭ}})$ в диапазоне линейного изменения напряжения имеет постоянный наклон. В этой области производная $dI_{\text{к2}}/dU_{\text{кЭ}} = g_{\text{кЭ2}}$ определяет *динамическую выходную проводимость* источника стабильного тока, а обратная величина $1/g = R$ является динамическим выходным сопротивлением. Идеальный источник стабильного тока имеет нулевую выходную проводимость $g = g_0 = 0$. ВАХ источника тока $I_0 = f(U_0)$ или $I_{\text{к2}} = f(U_{\text{кЭ}})$ позволяет определить типичные значения для выходной проводимости:

$$g = g_{\text{кЭ2}} = dI_{\text{к2}}/dU_{\text{кЭ2}} = I_{\text{к2}}/U_{\text{А}} = I_0/U_{\text{А}}, \quad (8.2)$$

где $g_{\text{кЭ2}}$ — динамическая проводимость между коллектором и эмиттером транзистора VT_2 , $U_{\text{А}}$ — напряжение Эрли (см. гл. 4), т. е. величина, обратная коэффициенту модуляции ширины базы для VT_2 . Из формулы (8.2) следует, что для относительного изменения тока $dI_{\text{к2}}/I_{\text{к2}} = dI_0/I_0 = dU_{\text{кЭ2}}/U_{\text{А}}$ при изменении выходного напряжения на 1 В ($dU_0 = dU_{\text{кЭ2}} = 1 \text{ В}$) имеем

$$(1/I_0) \cdot (dI_0/dU_0) = 1/U_{\text{А}}. \quad (8.3)$$

Поскольку в рассматриваемой на рис. 8.1 схеме ток на верхнем незаземленном выводе направлен внутрь схемы источника тока, то данный тип устройств относится к *источникам отрицательной полярности*. В обратном случае, когда ток $I_2 = I_0$ направлен в противоположную сторону, потенциал в коллекторной цепи транзистора VT_1 должен быть противоположной полярности, т. е. $-\mathcal{E}$.

Для многих применений ИС необходимо иметь источник тока с симметричным диапазоном линейного изменения напряжения, когда выходное напряжение имеет как положительные, так и отрицательные значения. В этом случае схема будет аналогична показанной на рис. 8.1, но вместо заземления эмиттеров необходимо подать на них отрицательный потенциал.

Схема источника тока для низких уровней тока по сравнению с рассмотренной включает сопротивление R_2 в эмиттерной цепи транзистора VT_1 .

Разработано достаточно много различных схем источников стабильного тока, которые являются значительно более сложными, но при этом имеют значительные преимущества по сравнению с простыми схемами токового зеркала. В качестве примера на рис. 8.3 приведена схема из нескольких источников тока отрицательной полярности (групповая схема). Здесь транзистор VT_2 с резистором R_2 служит опорным источником тока для транзисторов $VT_3...VT_6$. Транзистор VT_1 улучшает стабильность опорного тока, поскольку он уменьшает различие токов I_1 через резистор R_1 и I_2 через транзистор VT_2 [37]. Если размеры эмиттеров транзисторов $VT_2...VT_6$ подобраны таким образом, что плотности тока в них одинаковы, то и падения напряжения на резисторах $R_2...R_6$ будут одинаковыми для всех указанных транзисторов.

В рассматриваемой схеме $I_3R_3 = I_4R_4 = I_5R_5 = I_6R_6 = I_1R_2$, т. е. токи обратно пропорциональны эмиттерным сопротивлениям.

Чтобы уменьшить площадь, занимаемую транзисторами на кристалле ИС, желательно не выравнивать размеры эмиттеров. Это особенно важно при больших изменениях токов. Когда размеры транзисторов не обеспечивают равенство плотностей тока в них, то напряжения $U_{БЭ}$ транзисторов будут различаться. Для компенсации этого расхождения необходимо изменить номиналы сопротивлений в цепях эмиттеров.

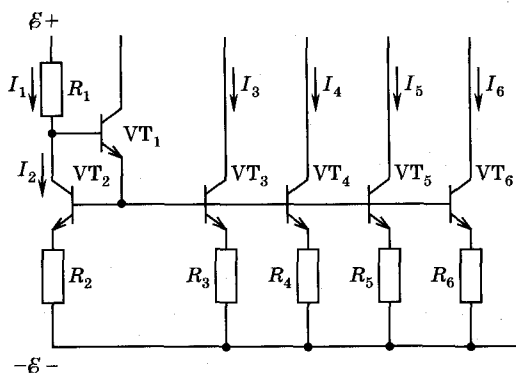


Рис. 8.3

Полевые транзисторы (как с управляющими электрически-ми переходами (ПТУП), так и МДП (МОП)-транзисторы) широко используются в ИС как стабилизаторы и источники стабильного тока. ПТУП может быть использован в качестве стабилизатора или источника тока, если он включается как диод с закороченным на исток затвором. В этом случае ВАХ такого ПТУП в диодном включении является, по сути дела, токовой характеристикой $I_{СИ} = f(U_{СИ})$ при $U_{ЗИ} = 0$. При $U_{СИ} \geq U_{отс}$ (см. гл. 6) ток $I_{СИ}$ насыщается, т. е. при увеличении напряжения $U_{СИ}$ он изменяется весьма незначительно вплоть до напряжения пробоя стокового перехода. Таким образом, ВАХ ПТ в области насыщения также формирует диапазон линейного изменения напряжения подобно характеристике, показанной на рис. 8.2. Подобно ПТУП в качестве стабилизаторов тока можно использовать и МОП-транзисторы в режиме обеднения. Существует множество схем источников тока на МОП-транзисторах, похожих на схемы рассмотренных источников на БТ.

В различных усилительных, цифровых логических и аналоговых ИС источники тока часто используются в качестве активной нагрузки. Например, токовое зеркало часто применяют как активную нагрузку в дифференциальных усилителях на БТ и ПТ (см. п. 8.3).

8.2.2. Источники напряжения.

В источниках напряжения выходное напряжение U_0 не зависит от величины нагрузки или, что одно и то же, от выходного тока. Существует два основных способа реализации источника напряжения, которые могут применяться не только по отдельности, но и совместно. В первом случае реализуются свойства транзистора преобразовывать импеданс, зависящий от усиления транзистора по току. Другой способ основывается на использовании свойств усилителя с обратной связью.

Анализ характеристик преобразования импеданса рассмотрим на примере схемы, изображенной на рис. 8.4, где база транзистора через резистор R_0 подключена к источнику напряжения $U_{вх}$. Если выходной ток I_0 возрастает на dI_0 , то базовый ток изменяется на величину $dI_B = dI_3/(\beta + 1) = dI_0/(\beta + 1)$ (см. гл. 4). Изменение падения напряжения на сопротивлении R_0 будет равно $dI_{B}R_0 = dI_0R_0/(\beta + 1)$. Падение напряжения между базой и эмиттером увеличится в соответствии с формулой $dU_{БЭ} = (dU_{БЭ}/dI_3)dI_3$,

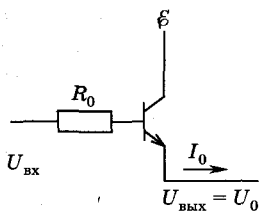


Рис. 8.4

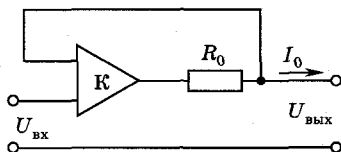


Рис. 8.5

где $dU_{БЭ}/dI_Э = r_{БЭ}$ — динамическое сопротивление перехода эмиттер — база транзистора. В активном режиме $I_Э \approx I_К = I_0 \exp(U_{БЭ}/\varphi_T)$ и, следовательно, $dI_Э/dU_{БЭ} = (r_{БЭ})^{-1} \approx dI_К/dU_{БЭ} \approx I_Э/\varphi_T$, т. е. $r_{БЭ} = \varphi_T/I_Э$.

Используя результаты этого анализа, получим для полного изменения выходного напряжения выражение

$$dU_{\text{вых}} = -dI_Б R_0 - dU_{БЭ} = -(dI_0 R_0/(\beta + 1)) - dI_0 r_{БЭ}, \quad (8.4)$$

откуда следует формула для выходного сопротивления рассматриваемой схемы:

$$r_{\text{вых}} = -dU_{\text{вых}}/dI_0 = R_0/(\beta + 1) + r_{БЭ}. \quad (8.5)$$

Из (8.5) следует, что за счет усиления (коэффициент β) значение $r_{\text{вых}}$ уменьшается по сравнению с R_0 пропорционально $(\beta + 1)$. Поскольку обычно $\beta \sim 100$, то изменение импеданса (в данном случае $r_{\text{вых}}$) существенно. Проведенный анализ справедлив для малых изменений выходного тока. В общем случае полное изменение выходного напряжения, вызываемое увеличением выходного тока от I_{01} до I_{02} , определяется формулой

$$\Delta U_{\text{вых}} = -\Delta I_0 R_0/(\beta + 1) - \varphi_T \ln(I_{02}/I_{01}), \quad (8.6)$$

где $\Delta I_0 = I_{02} - I_{01}$.

Использование отрицательной обратной связи для уменьшения выходного импеданса иллюстрируется рис. 8.5, где K — коэффициент усиления в отсутствие обратной связи, R_0 — выходной внутренний импеданс усилителя без обратной связи (для наглядности изображен отдельно от усилителя). Выходное напряжение данной схемы определяется выражением

$$U_{\text{вых}} = (U_{\text{вх}} - U_{\text{вых}})K - I_0 R_0 = U_{\text{вх}} K/(1 + K) - I_0 R_0', \quad (8.7)$$

где $R'_0 = R_0/(1 + K)$ — выходное сопротивление рассматриваемой схемы при наличии обратной связи.

Так как $R'_0 \approx R_0/K$ при $K \gg 1$, то выходное сопротивление при наличии обратной связи много меньше, чем в ее отсутствие.

Полученные выше значения динамического выходного сопротивления для источника напряжения справедливы при работе на низких частотах. На высоких частотах происходит уменьшение коэффициента усиления (см. гл. 4) и появляется фазовый сдвиг между напряжением и током, что приводит к изменению выходного импеданса, который становится комплексным, а на очень высоких частотах — чисто реактивным. Как показывают расчеты, для схемы с обратной связью выходной импеданс на высоких частотах записывается в следующем виде [37]:

$$Z_0 \approx jfR_0/f_{гр} = j\omega L_0, \quad (8.8)$$

где $L_0 = R_0/(2\pi f_{гр})$ — выходная индуктивность схемы, $f_{гр}$ — граничная частота, при которой коэффициент усиления по напряжению равен единице ($K = 1$). Источник напряжения должен иметь очень низкий динамический выходной импеданс, чтобы выходное напряжение мало изменялось при изменении выходного тока. Необходимо также стремиться к тому, чтобы у источников или стабилизаторов напряжения выходное напряжение как можно меньше зависело от напряжения источника питания. Для реализации последнего требования используют схемы с применением стабилизаторов.

8.2.3. Источники опорного напряжения с термостабилизацией.

Температурный коэффициент выходного напряжения (ТКН $= dU_{вых}/dT$) является в практическом отношении наиболее важной характеристикой источника опорного напряжения. Поскольку все электронные компоненты ИС, используемые в схемах опорного напряжения, имеют различные ТКН, основные элементы ИС подбираются так, чтобы имела место их взаимная компенсация, при которой результирующий ТКН стремится к нулю при заданной температуре. Схема на рис. 8.6, где транзисторы VT_1 , VT_2 , VT_3 находятся в диодном включении, обладает взаимной компенсацией ТКН различных компонентов. Ток от источника тока I_0 вызывает обратное смещение транзистора VT_1 , который в этом случае работает как стабилизатор. Падение напряжения на VT_1 равно напряжению пробоя перехода эмиттер — база. В результате

имеем равенство напряжений $U_e = U_{ЭБ\text{ проб}}$, которые обычно составляют 6...7 В (см. гл. 7 — диодное включение БТ ИС). При этом выходное напряжение определяется следующим выражением

$$U_0 = U_{\text{ВЫХ}} = (U_{Э2}R_2 + U_{К3}R_1)/(R_1 + R_2). \quad (8.9)$$

Напряжения на эмиттере транзистора VT_2 и на коллекторе транзистора VT_3 равны соответственно $U_{Э2} = U_e - U_{БЭ4} - U_{БЭ2}$ и $U_{К3} = U_{БЭ3}$.

Токи, текущие через транзисторы VT_4 , VT_3 и VT_2 , практически одинаковы, поскольку эти транзисторы являются интегральными и имеют идентичную конструкцию. Поэтому для этих трех транзисторов падения напряжения $U_{БЭ}$ и соответствующие ТКН $dU_{БЭ}/dT$ будут почти одинаковыми. В результате, если положить $U_{БЭ} = U_{КЭ}$, то выходное напряжение можно записать в виде

$$U_{\text{ВЫХ}} = [U_c R_2 - U_{БЭ}(2R_2 - R_1)]/(R_1 + R_2). \quad (8.10)$$

Если не учитывать температурную зависимость резисторов, то, исходя из выражения (8.10), получим для выходного напряжения

$$\text{ТКН} = dU_{\text{ВЫХ}}/dT = \frac{R_2(dU_c/dT) - (dU_{БЭ}/dT)(2R_2 - R_1)}{R_1 + R_2}. \quad (8.11)$$

Из формулы (8.11) следует, что подбором отношения сопротивлений R_1/R_2 можно получить нулевой ТКН (если числитель дроби приравнять нулю).

Для повышения опорного напряжения, а также для изоляции источника этого напряжения от нагрузки в соответствующей схеме используют усилитель с обратной связью (рис. 8.7). В этой схе-

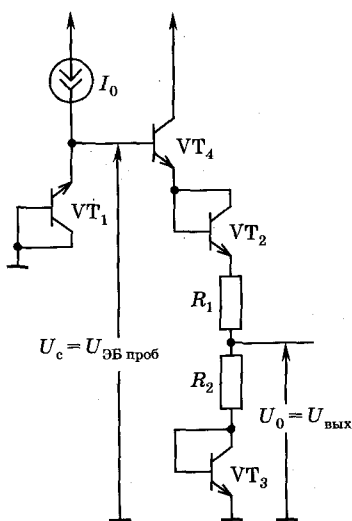


Рис. 8.6

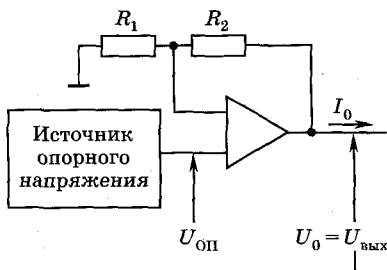


Рис. 8.7

ме часть выходного напряжения, полученного делением в отношении $R_2/(R_1 + R_2)$, подается обратно на инвертирующий вход усилителя. В усилителе это напряжение сравнивается с опорным напряжением, подаваемым на неинвертирующий вход. Из-за наличия обратной связи усилитель работает так, чтобы $U_{\text{ВЫХ}}$ удовлетворяло соотношению

$$U_{\text{ВЫХ}} = U_{\text{ОП}}(1 + R_1/R_2). \quad (8.12)$$

Отсюда следует, что подбором отношения R_1/R_2 можно увеличивать $U_{\text{ВЫХ}}$ до нужных значений, удовлетворяющих условиям реализации схемы, при этом

$$\text{ТКН} = dU_{\text{ВЫХ}}/dT = (dU_{\text{ОП}}/dT)(1 + R_1/R_2). \quad (8.13)$$

При выводе формулы (8.13) полагалось, что отношение сопротивлений R_1/R_2 не зависит от температуры.

Кроме рассмотренных схем термостабилизированных источников опорного напряжения, существует ряд схем, в которых определяющую роль играет температурная зависимость ширины запрещенной зоны полупроводника.

8.3. Дифференциальные усилители

Дифференциальные усилители (ДУ) широко применяются в аналоговых ИС различного типа: операционных усилителях (ОУ), компараторах и стабилизаторах напряжения, видеоусилителях, балансных модуляторах и демодуляторах. Кроме того, на основе ДУ построено большинство элементов цифровой эмиттерно-связанной логики (ЭСЛ). Дифференциальный усилитель является входным каскадом ОУ и других ИС, поэтому он определяет большинство важнейших рабочих характеристик ИС: напряжение смещения, входной ток смещения, входной ток сдвига, входное сопротивление, коэффициент ослабления синфазного сигнала, диапазон рабочих частот.

Существуют схемы ДУ на БТ, МОП-транзисторах и ПТУП, составных транзисторах, а также схемы с активной нагрузкой и ряд других.

8.3.1. ДУ на биполярных транзисторах.

При работе ДУ большое значение имеет одинаковость параметров элементов симметричной схемы. Базовая схема состоит только из двух транзисторов с соединяемыми эмиттерами, к которым под-

ключен идеальный источник тока, при этом на базы подается входной сигнал, а между коллекторами снимается выходной (рис. 8.8). ДУ, изображенный на рис. 8.8, состоит из двух симметричных схемных цепей, каждая из которых содержит транзистор и резистор. Помимо указанных симметричных цепей, в общей эмиттерной цепи имеется источник тока I_0 . Входным сигналом $U_{вх}$ является разность базовых потенциалов $U_{вх} = U_{Б1} - U_{Б2}$, а выходным напряжением — разность коллекторных потенциалов $U_{вых} = U_{К1} - U_{К2}$. В ДУ оба транзистора работают в активном режиме. Источник тока I_0 обеспечивает стабильность рабочей точки транзисторов, т. е. токов I_0^0 и напряжений $U_{К}^0$.

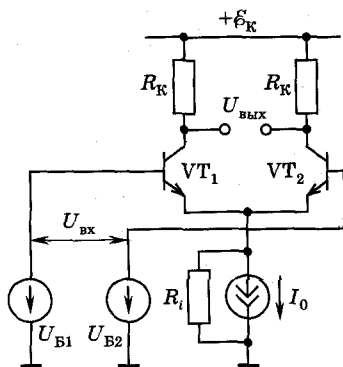


Рис. 8.8

При идеальной симметрии цепей ДУ в отсутствие входного сигнала выходное напряжение равно нулю, поскольку коллекторные потенциалы и токи транзисторов VT_1 и VT_2 одинаковы из-за симметрии цепей. Такая ситуация будет сохраняться при любом одинаковом изменении токов в обеих ветвях усилителя, т. е. в идеальном ДУ дрейф выходного напряжения отсутствует, несмотря на возможный значительный дрейф в каждом из плечей ДУ. Таким образом усилитель будет устойчив на всех частотах независимо от температуры. Любые изменения коэффициента передачи или токов утечки в усилительной цепи транзистора VT_1 компенсируются такими же изменениями в усилительном канале транзистора VT_2 .

Если на базы транзисторов VT_1 и VT_2 подаются одинаковые сигналы (*синфазные сигналы*), то $U_{вых} = 0$, поскольку в обоих транзисторах произойдут одинаковые изменения параметров и разность $U_{вых} = U_{К1} - U_{К2} = 0$.

Если же на базы подавать сигналы одинаковые по амплитуде, но противоположные по фазе (или по знаку), так называемые *дифференциальные сигналы*, то их разность $\Delta U_{Б1} - \Delta U_{Б2}$ и будет входным сигналом $U_{вх}$. В силу симметрии цепей VT_1 и VT_2 сигнал $U_{вх}$ поделится поровну между обоими эмиттерными переходами. На одном из них потенциал увеличится на $0,5U_{вх}$, а

на другом уменьшится на $0,5U_{\text{вх}}$. В результате на выходе возникнет напряжение, поскольку приращения коллекторных потенциалов $U_{\text{к1}}$ и $U_{\text{к2}}$ будут одинаковыми по величине, но противоположными по знаку. Таким образом, идеальный ДУ реагирует только на разностный (дифференциальный) сигнал; при этом потенциал эмиттеров остается неизменным, поэтому при анализе ДУ его считают заданным, а ток генератора I_0 — постоянным ($R_i \rightarrow \infty$). В реальных условиях источник тока имеет не бесконечное внутреннее сопротивление R_i , в результате возникшее приращение потенциала эмиттера $\Delta U_{\text{э}}$ вызывает изменение эмиттерного тока на величину $\Delta I_0 = \Delta U_{\text{э}}/R_i$ (при $R_i \rightarrow \infty$ приращение $\Delta I_0 \rightarrow 0$). Это изменение тока ΔI_0 делится между цепями транзисторов VT_1 и VT_2 и вызывает приращение коллекторных потенциалов $\Delta U_{\text{к1}}$ и $\Delta U_{\text{к2}}$. В случае идентичности цепей $\Delta U_{\text{к1}} = \Delta U_{\text{к2}}$ и на выходе ДУ имеется только синфазная составляющая. В противном случае $\Delta U_{\text{к1}} \neq \Delta U_{\text{к2}}$ и на выходе ДУ наряду с синфазной составляющей появится паразитная дифференциальная составляющая. Симметрия, т. е. идентичность цепей VT_1 и VT_2 ДУ, легче всего достигается в ИС. Однако даже если оба транзистора выполнены на одном кристалле, они не будут абсолютно одинаковыми, поэтому даже в отсутствие входного сигнала на выходе будет присутствовать некоторое напряжение, которое обычно называют напряжением смещения $U_{\text{см}}$; оно определяется формулой $U_{\text{см}} = \phi_T \ln(I_{01}/I_{02})$, где I_{01} и I_{02} — соответственно тепловые токи эмиттерных переходов первого и второго транзисторов.

Передаточная характеристика. Рассмотрим передающую и входную статические характеристики ДУ. Передающая характеристика, представляющая собой зависимость выходных токов $I_1 = I_{\text{к1}}$ и $I_2 = I_{\text{к2}}$ от входного напряжения $U_{\text{вх}}$ с учетом смещения $U_{\text{см}}$, т. е. от величины $U'_{\text{вх}} = U_{\text{вх}} - U_{\text{см}}$, представлена на рис. 8.9. Из этой характеристики видно, что ДУ является нелинейным устройством. Однако в ограниченных областях передающей характеристики $I_1 = f(U_{\text{вх}})$ или $I_2 = f(U_{\text{вх}})$ зависимость между токами и входным напряжением можно считать примерно линейной (область от $U_{\text{вх}} = -30$ мВ до $U_{\text{вх}} = 30$ мВ на рис. 8.9). Таким образом, для малых сигналов ДУ является линейным устрой-

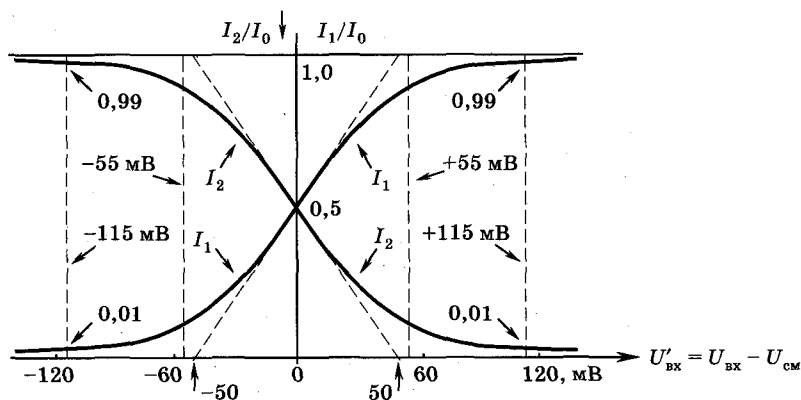


Рис. 8.9

ством. Входная характеристика $\Delta I_{\text{вх}} = f(U_{\text{вх}})$ совпадает по форме с передаточной.

Рассмотрим *некоторые из параметров ДУ*. Главным параметром ДУ является *коэффициент усиления (K)*. Поскольку потенциал эмиттеров при подаче дифференциального сигнала остается неизменным, а для переменных составляющих он вообще равен нулю, то коэффициент усиления ДУ равен коэффициенту усиления каждого плеча, ибо каждое плечо усиливает сигнал величиной $0,5U_{\text{вх}}$.

Коэффициент усиления по напряжению определяется как отношение выходного и входного сигналов, т. е. $K = U_{\text{вых}}/U_{\text{вх}}$. Выходным сигналом ДУ является переменная составляющая напряжения коллектора, которая равна $U_{\text{к}} = U_{\text{вых}} = -\alpha I_{\text{э}} R_{\text{к}}$, а входным — потенциал базы $U_{\text{вх}} = I_{\text{э}}(r_{\text{э}} + (1 - \alpha)(R_{\text{т}} + r_{\text{б}}))$, где $r_{\text{э}}$ — дифференциальное сопротивление эмиттерного перехода, $R_{\text{т}}$ — внутреннее сопротивление источника сигнала, $r_{\text{б}}$ — объемное сопротивление базы (см. гл. 4).

В результате выражение для коэффициента усиления примет вид

$$K = U_{\text{вых}}/U_{\text{вх}} = \alpha R_{\text{к}}/[r_{\text{э}} + (1 - \alpha)(R_{\text{т}} + r_{\text{б}})]. \quad (8.14)$$

В случае низкоомных источников сигнала и небольших рабочих токов вторым слагаемым в знаменателе выражения (8.14) можно пренебречь, тогда

$$K = -\alpha(R_{\text{к}}/r_{\text{э}}). \quad (8.15)$$

Входное сопротивление ДУ определяется отдельно для дифференциальной и синфазной составляющих. Для дифференциальной составляющей входное сопротивление $R_{\text{вх. д}}$ равно удвоенному входному сопротивлению каждого плеча ДУ. Поскольку каждое плечо ДУ является схемой с ОЭ, то согласно представлению транзистора как четырехполюсника в схеме с ОЭ входное сопротивление — это параметр $h_{11\beta}$ (см. гл. 4), поэтому

$$R_{\text{вх. д}} = 2h_{11\beta} = 2[(\beta + 1)r_{\text{э}} + r_{\text{б}}]. \quad (8.16)$$

При малых токах дифференциальное сопротивление эмиттерного перехода $r_{\text{э}}$ много больше, чем при больших, поэтому для увеличения $R_{\text{вх. д}}$ необходимо использовать ДУ в режиме малых токов.

Для синфазной составляющей входное сопротивление $R_{\text{вх. с}}$ определяется сопротивлением R_i источника тока I_0 . Тогда $R_{\text{вх. с}}$ можно вычислить по формуле

$$R_{\text{вх. с}} = (\beta + 1)R_i. \quad (8.17)$$

Динамический диапазон — отношение минимально возможного и максимального входных сигналов. Это отношение часто выражают в децибелах. Минимальный сигнал ограничивается собственными шумами, а максимальный — искажениями формы сигнала (нелинейные искажения). Приблизительно можно оценить максимально допустимый сигнал, пользуясь передаточной характеристикой ДУ (см. рис. 8.9)

Для работы усилителя необходима нагрузка. В рассмотренных ДУ в качестве нагрузки выступали резисторы в коллекторных цепях. Такая нагрузка называется *пассивной*. Однако нагрузка может быть и *активной (динамической)*, когда в качестве нагрузки выступают транзисторы. Необходимость использования транзисторов в качестве нагрузки вызвана тем, что для получения большого коэффициента усиления следует использовать резисторы с большим номиналом. Однако использование большого сопротивления нагрузки в ИС приводит к ряду трудностей. В ИС площадь, необходимая под резистор, пропорциональна его сопротивлению, т. е. резистор с большим сопротивлением занимает слишком много места на кристалле. Помимо этого, резистор с большим номиналом имеет и значительную паразитную емкость.

Это приводит к большому значению постоянной времени RC , что накладывает ограничения на частотную характеристику усилителя. Кроме перечисленных недостатков, для нормальной работы ДУ транзисторы всегда должны оставаться в активном режиме и не попадать в область насыщения, что вызывает ограничение входного базового напряжения. Это напряжение должно быть таким, чтобы переход коллектор—база был смещен в прямом направлении не более, чем на 0,5 В. В результате напряжение на коллекторе будет много меньше напряжения источника питания, а это приводит к значительному уменьшению диапазона изменения входного напряжения ДУ.

Из-за наличия указанных недостатков в ДУ ИС используют активную (динамическую) нагрузку. Пример ДУ на биполярных транзисторах VT_1 и VT_2 с активной нагрузкой (также из биполярных транзисторов VT_3 и VT_4) приведен на рис. 8.10, где VT_3 включен по диодной схеме. Такое сочетание VT_3 и VT_4 , как отмечалось в п. 8.2, называется токовым зеркалом. Если транзисторы VT_3 и VT_4 совершенно идентичны и напряжения база—эмиттер у них равны, то в этом случае коллекторные токи обоих транзисторов одинаковы, т. е. $I_3 = I_4$. Следовательно, любой ток через VT_4 будет «зеркальным отражением» тока через VT_3 . Анализ такой схемы показывает, что коэффициент усиления переменного напряжения такого каскада ДУ равен $K = 2gR_n$, где R_n определяют транзисторы VT_3 и VT_4 , а $g = I_1 I_2 / [(I_1 + I_2)\phi_T]$ — передаточная проводимость. Таким образом, K в ДУ с активной нагрузкой вдвое больше, чем с постоянной резистивной.

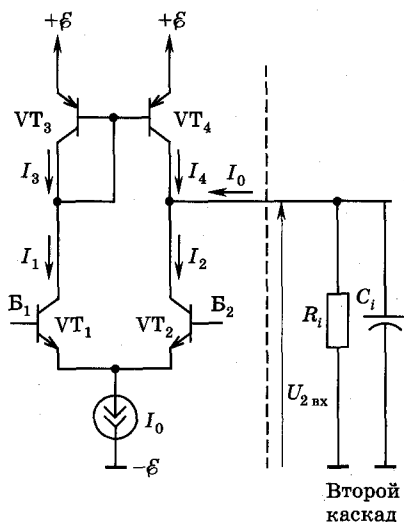


Рис. 8.10

8.3.2. ДУ на полевых транзисторах (ПТУП) и МДП (МОП)-транзисторах.

Этот тип ДУ принципиально работает так же, как и ДУ на БТ, однако ДУ на ПТ обладает по сравнению с ДУ на БТ гораздо

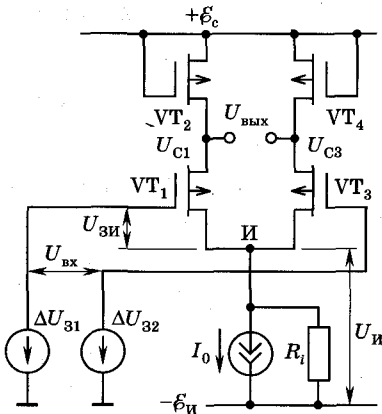


Рис. 8.11

рис. 8.11, где нагрузочные транзисторы VT_2 и VT_4 представляют собой динамические нагрузки для VT_1 и VT_3 . Нагрузочные транзисторы работают на участках насыщения ВАХ, потому что в малосигнальном приближении их сопротивления одинаковы и определяются формулой

$$R_{Cn} = dU_{Cn}/dI_{Cn} = 1/S_n, \quad (8.18)$$

где индексы $n = 2, 4$ относятся к транзисторам VT_2 и VT_4 соответственно; S_n — крутизна характеристики соответствующего транзистора, а U_{Cn} , I_{Cn} — напряжение и ток стока. Поскольку, как и в ДУ на БТ, коэффициент усиления для дифференциальной составляющей в рассматриваемом случае равен коэффициенту усиления отдельного плеча, то достаточно оценить значение K , например, для левого плеча, составленного из VT_1 и VT_2 . Выходное напряжение $U_{\text{ВЫХ}}$ определяется соотношением (см. гл. 6)

$$U_{\text{ВЫХ}} = -I_C R_C = -S R_C U_{\text{ВХ}}, \text{ откуда следует, что } K = U_{\text{ВЫХ}}/U_{\text{ВХ}} = -S_1 R_C.$$

Учитывая (8.18), получим

$$K = -S_1/S_2. \quad (8.19)$$

Коэффициент усиления ограничен значениями 5...7, поскольку он определяется размерами каналов у активного и нагрузочного транзисторов, прежде всего отношением толщин каналов, которое лимитируется возможностями технологии.

большим входным сопротивлением ($\sim 10^9 \dots 10^{12}$ Ом) и очень малым входным током смещения ($\sim 10^{-9} \dots 10^{-12}$ А). К недостаткам ДУ на ПТ с управляющими электрическими переходами можно отнести: низкую передаточную проводимость и, как следствие этого, низкий коэффициент усиления по напряжению; большое напряжение смещения пары ПТ по сравнению с парой БТ.

Схема простейшего ДУ на МДП-транзисторах с динамической нагрузкой представлена на

Анализ показывает, что коэффициент усиления для синфазной составляющей K_{cc} равен

$$K_{cc} = K / (2SR_i + 1), \quad (8.20)$$

где R_i — внутреннее сопротивление источника тока I_0 .

В МДП-транзисторных ДУ главная составляющая напряжения смещения $U_{см}$ обусловлена разбросом пороговых напряжений и удельных крутизн активных транзисторов, поэтому значения $U_{см}$ у МДП ДУ больше, чем у биполярных ДУ.

8.3.3. Дифференциальный усилитель на составных транзисторах.

Составной транзистор — это комбинация из нескольких транзисторов в ИС, соединенных между собой так, что их можно рассматривать как единое целое. Свойства составных транзисторов сильно отличаются от свойств обычных транзисторов. Наибольшее распространение получили составные транзисторы в виде так называемой *пары Дарлингтона*.

Схема ДУ с двумя парами транзисторов Дарлингтона (рис. 8.12) позволяет получать очень высокое входное сопротивление и очень низкий входной ток смещения по сравнению с обычной схемой. Несколько большее напряжение смещения из-за наличия четырех транзисторов в ДУ с транзисторами Дарлингтона является одним из недостатков этого вида ДУ. Главная особенность пары Дарлингтона, например VT_1 и VT_2 , — это исключительно большой коэффициент передачи базового тока, который равен $\beta = \beta_1\beta_2$, где β_1, β_2 — коэффициенты передачи базового тока соответственно транзисторов VT_1 и VT_2 .

Поскольку коэффициент передачи по току одиночного транзистора равен по порядку величины 10^2 , то общий коэффициент передачи по току схемы Дарлингтона есть величина порядка 10^4 . Из-за большого коэффициента передачи по току входной ток смещения (I_B) в ДУ со схемой Дарлингтона очень мал, а входное сопротивление, обратно пропорциональное току смещения (базовому току), напротив, очень велико по сравнению с обычной схемой двухтранзисторного ДУ.

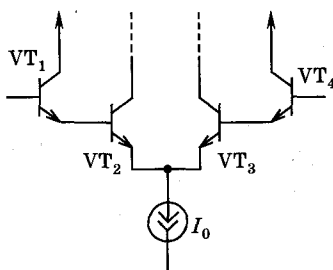


Рис. 8.12

8.4. Операционные усилители

Входные каскады *операционных усилителей (ОУ)* построены на основе ДУ: ОУ представляет собой аналоговую ИС, на выходе которой формируется напряжение U_0 , равное по величине усиленной разности $U_1 - U_2$ между двумя входными напряжениями U_1 и U_2 (рис. 8.13), т. е.

$$U_0 = K(U_1 - U_2), \quad (8.21)$$

где K — коэффициент усиления ОУ без обратной связи.

На рис. 8.13 показаны схемные обозначения ОУ (треугольником обозначен ОУ, включающий многокаскадный усилитель напряжения, содержащий от десятков до сотен транзисторов): a — основное обозначение; b и c — обозначения с указанием функционального назначения входов. В идеале ОУ чувствителен только к разности входных напряжений $U_d = U_1 - U_2$ (см. п. 8.3.1), которая называется *дифференциальным входным сигналом*, и нечувствителен к общей для обоих входов составляющей входных сигналов. Общая составляющая, которая называется *синфазным сигналом*, определяется формулой

$$U_{cc} = (U_1 + U_2)/2.$$

Если на вход ОУ подать только напряжение U_1 ($U_2 = 0$), то $U_0 = KU_1$, т. е. выходное напряжение U_0 равно усиленному *неин-*

вертированному входному сигналу U_1 , при этом нижний вход (см. рис. 8.13, b) называется *неинвертирующим*. И наоборот, если на входе есть только сигнал U_2 ($U_1 = 0$), то выходной сигнал $U_0 = -KU_2$ — *инвертированный* усиленный сигнал U_2 , а соответствующий (верхний) вход называется *инвертирующим* (см. рис. 8.13, b). В общем случае схема ДУ должна включать и подключение источника питания, однако для упрощения схемы подключения источника питания обычно не изображаются.

Диапазон изменения напряжений на входе для обеспечения правильного функционирования ОУ ограничивается таким образом, чтобы это напряжение не было мень-

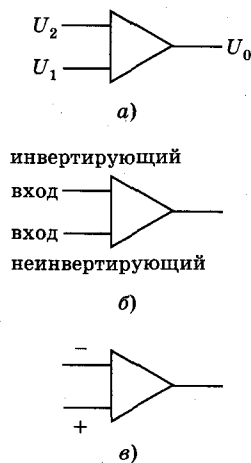


Рис. 8.13

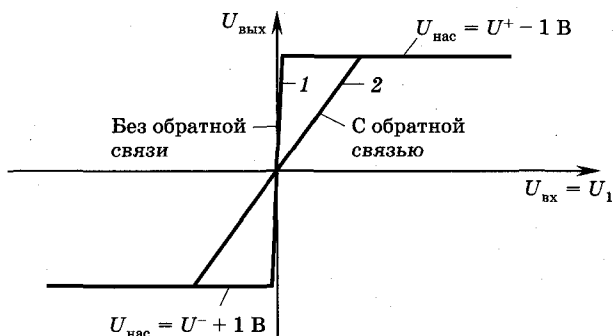


Рис. 8.14

ше $-0,5$ В. Напряжение U_0 на выходе ОУ не может превысить напряжения источника питания. При этом максимальное значение U_0 всегда по крайней мере на 1 В меньше абсолютной величины питающего напряжения. При низкоомной нагрузке U_0 заметно меньше этой максимальной величины: большинство ОУ питаются от двоякого источника питания с отрицательным и положительным полюсами, подключенными к специальным выводам ОУ (см. рис. 8.13, в).

Идеальная передаточная характеристика ОУ (рис. 8.14) представляет собой ломаную линию, состоящую из линейно восходящей (область усиления) и горизонтальной (область насыщения). В области насыщения выходное напряжение ограничено напряжением питания и, как только что отмечено, ниже его на $1 \dots 2$ В (см. рис. 8.14). Коэффициент усиления K в ОУ очень велик, особенно при работе на низких частотах, где он достигает значений $10^5 \dots 10^6$, поэтому ширина линейной зоны (диапазон изменения входных напряжений) весьма незначительна (см. рис. 8.14, кривая 1). При напряжении питания ± 10 В и указанных коэффициентах усиления диапазон изменения входных напряжений, соответствующий линейному участку, составляет $0,02 \dots 0,2$ мВ, т. е. достаточно малые входные сигналы < 1 мВ вызывают уже насыщение усиления и при больших значениях входных сигналов возникают нелинейные искажения. По этой причине и ряду других на практике обычно применяют ОУ с обратной связью (ОС) (рис. 8.15, а). В этом случае часть выходного напряжения через элемент отрицательной обратной связи Z_2 подается на инвертирующий вход. Величина напряжения, поступающего на вход за счет ОС, определяется делителем напряжения Z_1, Z_2 . Ко-

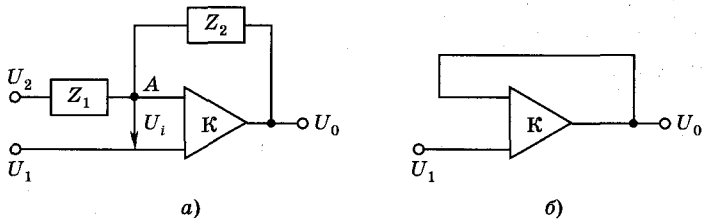


Рис. 8.15

эффицент, показывающий, какая часть напряжения поступает с выхода на вход, называется **коэффициентом обратной связи** (F). Для схемы простого делителя напряжения на рис. 8.15, а $F = Z_1/(Z_1 + Z_2)$, при этом дифференциальная составляющая напряжения на входе ОУ $U_i = U_1 - (U_2^1 + FU_0)$, где U_2^1 изменение U_2 из-за наличия ОС. Коэффициент усиления ОУ с отрицательной обратной связью K_{oc} равен

$$K_{oc} = K/(1 + FK), \quad (8.22)$$

где K — коэффициент усиления ОУ без обратной связи.

При $FK \gg 1$ соотношение (8.22) принимает вид $K_{oc} \approx K/FK \approx 1/F$, т. е. K_{oc} не зависит от K , а определяется главным образом параметрами ОС. Для схемы на рис. 8.15, а

$$K_{oc} = 1/F = (Z_1 + Z_2)/Z_1 = 1 + Z_2/Z_1. \quad (8.23)$$

Наличие отрицательной обратной связи значительно расширяет динамический диапазон входных напряжений, при этом передаточная характеристика имеет вид кривой 2 на рис. 8.14.

Коэффициенты усиления ОУ без ОС обычно сильно отличаются друг от друга (порой в десяток раз) даже в пределах партии однотипных ОУ. Кроме того, K сильно зависит от частоты входного сигнала и может изменяться от 10^6 на частоте 10 Гц до единицы и менее на частотах в несколько МГц. Колебания напряжения питания и температуры также сильно влияют на величину K .

На практике достаточно часто применяют ОУ с ОС при отсутствии делителя напряжений, когда инвертирующий вход закорачивается на выход (рис. 8.15, б). Коэффициент усиления в данном случае равен +1. Схема имеет наименование **повторителя напряжения**, так как выходное напряжение ОУ повторяет входное. Повторитель напряжения имеет очень большое входное сопротивление

и весьма малое выходное, поэтому он используется для согласования относительно низкоомной нагрузки с высокоомным источником сигнала без заметного уменьшения амплитуды сигнала.

Существует достаточно много разновидностей включения ОУ, которые позволяют получать стабильные параметры, устойчивые в широком диапазоне изменения входных величин и окружающих условий, что определяется конкретным функциональным назначением ОУ. Они находят очень широкое применение в различных устройствах и схемах. Отметим только некоторые из них: преобразователь ток — напряжение; преобразователь напряжение — ток; источник тока, управляемый напряжением; вычитающий усилитель; активный фильтр низких частот — интегратор; прецизионный детектор или выпрямитель; логарифмический преобразователь; стабилизатор положительного напряжения; фазовращатель с высоким входным и низким выходным сопротивлениями, мультиплексор аналоговых сигналов, генератор импульсов и т. д.



Контрольные вопросы

1. Особенности аналоговых ИС, применение, классификация.
2. Основные схемы, принципы работы, параметры источников стабильного тока, стабильного напряжения и опорного напряжения.
3. Схемные разновидности, особенности, принципы работы, параметры и характеристики дифференциальных усилителей.
4. Операционные усилители: схемы, коэффициенты усиления, характеристики, параметры, области применения.

Глава 9

ЦИФРОВЫЕ ИНТЕГРАЛЬНЫЕ СХЕМЫ

9.1. Особенности цифровых интегральных схем

В цифровых интегральных схемах все транзисторы работают в ключевом режиме, т. е. находятся либо в закрытом состоянии (режим отсечки), либо в открытом (режим насыщения). Состоя-

ния отсечки и насыщения имитируют логические единицу и ноль (высокий и низкий логические уровни) соответственно. В процессе переключения транзисторы быстро переходят из одного состояния в другое, минуя активный режим. С точки зрения цифровой техники схема ключа выполняет функцию простейшего логического элемента — *инвертора* и в статическом режиме может находиться в одном из указанных двух состояний.

Специфику и различия цифровых и аналоговых схем можно понять на основе рассмотрения передаточных характеристик. На рис. 9.1 изображены две передаточные характеристики, где в качестве входных и выходных величин приняты соответствующие напряжения. Кривая 1 на этом рисунке относится к инвертирующим схемам, у которых низким входным напряжениям соответствуют высокие выходные. Зависимость 2 является характеристикой неинвертирующих схем, у которых малым входным напряжениям соответствуют малые выходные. Характеристики 1 и 2 свойственны как простейшим цифровым, так и аналоговым схемам, однако использование этих характеристик в обоих случаях различается принципиально.

Входные и выходные сигналы в транзисторном ключе (цифровая схема) принимают два значения: ключ разомкнут, чему соответствует точка А на рис. 9.1: $U_{\text{вых}} = U_{\text{вых}A}$, $U_{\text{вх}} = U_{\text{вх}A}$; ключ замкнут — точка В: $U_{\text{вых}} = U_{\text{вых}B}$, $U_{\text{вх}} = U_{\text{вх}B}$. Входные и выходные напряжения имеют только два значения, соответствующие разомкнутому и замкнутому состоянию. Форма передаточной характеристики между

точками А и В для цифровой схемы не имеет значения. Ее выходные и входные величины остаются без изменения. В результате этих особенностей цифровые схемы мало чувствительны к разбросу параметров, к их температурной зависимости, к высоким электромагнитным помехам, что иллюстрируется на рис. 9.1 изменением входного сигнала $\Delta U_{\text{вх}}$ (шумы, наводки) около точки В. Эти небольшие изменения не отражаются на величине выходного сигнала.

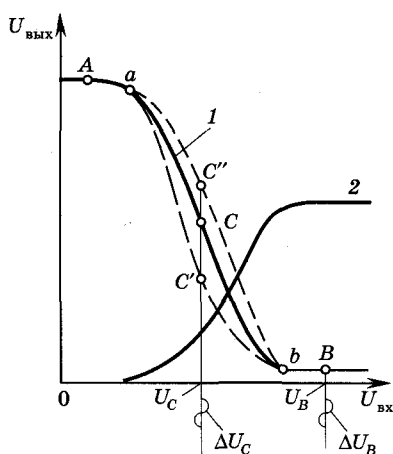


Рис. 9.1

В аналоговых усилительных схемах входные и выходные сигналы принимают любые значения между точками A и B и связаны друг с другом функциональной зависимостью $U_{\text{вых}} = f(U_{\text{вх}})$ (см. гл. 8). В этом случае любые отклонения на рассматриваемом участке (штриховые линии на рис. 9.1) будут отражаться на работе схемы, ибо будут изменяться как входные, так и выходные значения. Например, при одном и том же входном сигнале $U_{\text{вх}C} = U_C$ выходной сигнал может принимать значения $U_{\text{вых}C'}$ и $U_{\text{вых}C''}$, а рабочая точка C перемещается в положение между C' и C'' .

9.2. Элементарные (базовые) цифровые схемы на биполярных транзисторах

Простейшим схемным элементом цифровых интегральных схем является транзисторный ключ. Основные свойства ключей на БТ, их переходные и импульсные параметры были рассмотрены в гл. 4. Теперь будут рассматриваться более сложные схемы, нагрузочные параметры транзисторных ключей и другие виды базовых цифровых схем. Отдельные ключи используются в основном в аналоговых схемах. Для цифровых схем характерна совместная работа нескольких ключей. В реальных ИС часто один источник (\mathcal{E}_B) входного базового напряжения U_B управляет несколькими ключами, у которых участки база — эмиттер оказываются включенными параллельно. Для схемы двух параллельно включенных ключей (рис. 9.2) общий ток $I_B = I_{B1} + I_{B2}$ делится между базами пополам, если $I_{B1} = I_{B2}$. При неидентичности ключей их базовые токи могут сильно отличаться.

Анализ на основе модели Молла—Эберса (см. гл. 4) показывает, что входное напряжение ключа является функцией тока коллектора I_K , инверсного коэффициента передачи тока β_I , теплового тока I_{K0} . Параметры β_I и I_{K0} в ИС имеют разброс. Коллекторные токи I_K в сложных ключевых схемах могут сильно различаться. При одинаковых значениях входного базового напряжения U_B

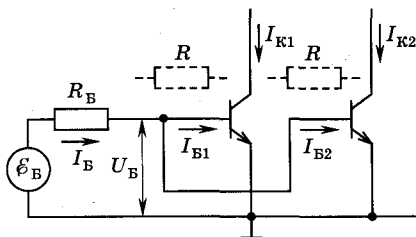


Рис. 9.2

в транзисторе с большим током I_K ток базы I_B будет меньше, степень насыщения этого транзистора тоже меньше. При большой разнице в токах I_K этот транзистор вообще окажется ненасыщенным, что нарушает работу ключа. Выравнивание токов I_B можно достигнуть за счет уменьшения крутизны входной характеристики. Для этого в цепь базы каждого транзистора последовательно включают одинаковые резисторы (на рис. 9.2 эти резисторы изображены пунктиром). В последовательной цепочке каждый из ключей, с одной стороны, управляется предыдущим и в то же время управляет последующим ключом (рис. 9.3). В результате для последовательной цепочки характерно чередование открытых и закрытых ключей.

Вариантом ключа с хорошим быстродействием является *переключатель тока*, который представляет собой симметричную схему (рис. 9.4), в которой заданный ток I_0 протекает через одну из ветвей в зависимости от потенциала на одном из базовых входов. На рис. 9.4 к первому базовому входу подключен источник сигнала $U_{вх}$, а на второй вход подается напряжение от источника постоянного напряжения \mathcal{E}_0 , который является опорным источником. Источник I_0 обеспечивает ток, соответствующий нормальному режиму. Работа схемы сводится к следующему. Если $U_{вх} = U_{вх}^0 = 0$, то транзистор VT_1 закрыт, а VT_2 — открыт. На выходе VT_2 напряжение будет малым, соответствующим насыщенному транзистору (логический ноль (0)). При подаче на вход VT_1 напряжения, соответствующего логической единице (1),

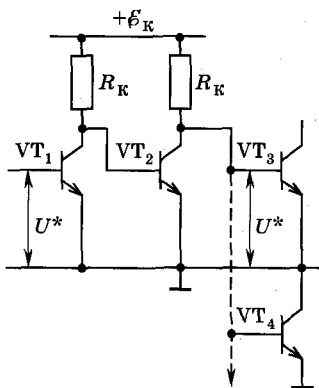


Рис. 9.3

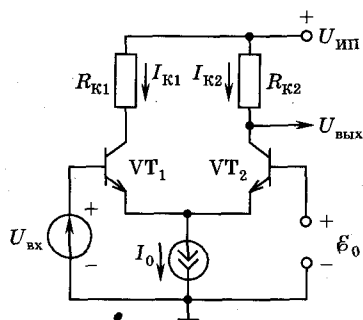


Рис. 9.4

т. е. $U_{\text{вх}} = U_{\text{вх}}^1$, VT_1 открывается, а VT_2 закрывается, т. е. на выходе второго транзистора будет большое напряжение (логическая 1 и $U_{\text{вых}} = U_{\text{вых}}^1$).

В интегральных схемах обычно используются последовательные цепочки переключателей тока, при этом необходимо между отдельными переключателями включать согласующие схемы (элементы), которые называются схемами смещения уровня. В простейшем случае смещение уровня может осуществляться за счет включения ЭДС между смежными переключателями.

9.3. Простейшие инверторные (ключевые) схемы на МДП-транзисторах

МДП-транзисторные ключи можно разделить на три типа: с резистивной нагрузкой (рис. 9.5, а), с динамической нагрузкой (см. рис. 9.7, а) и комплементарные ключи (см. рис. 9.8, а).

Рассмотрим процессы переключения *n*-канального МДП-ключа с резистивной нагрузкой (см. рис. 9.5). Если на затвор подать напряжение $U_{\text{вх}} = U_3^+ < U_{\text{пор}}$, то ключ закрыт, ток стока равен нулю, а выходное напряжение $U_{\text{вых}} = U^1 = \xi_C$ (точка А на рис. 9.5, б), что соответствует логической единице 1. При подаче на затвор напряжения $U_{\text{вх}} = U_3^+ > U_{\text{пор}}$ ключ переходит в открытое состояние. Входное напряжение должно быть достаточно большим, чтобы напряжение $U_{\text{вых}} = U_{\text{ост}} = U^0$ (точка В на рис. 9.5, б — ло-

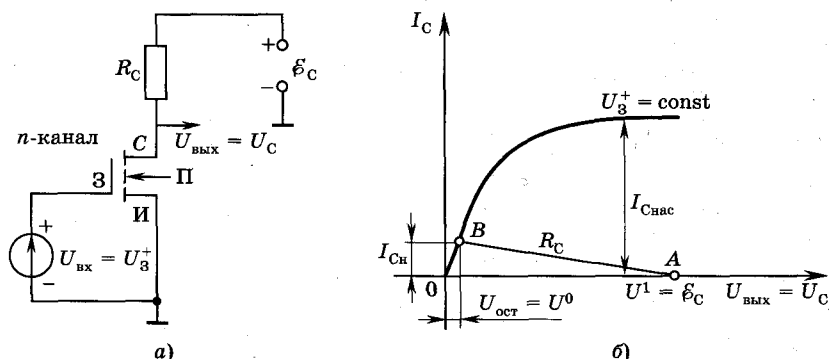


Рис. 9.5

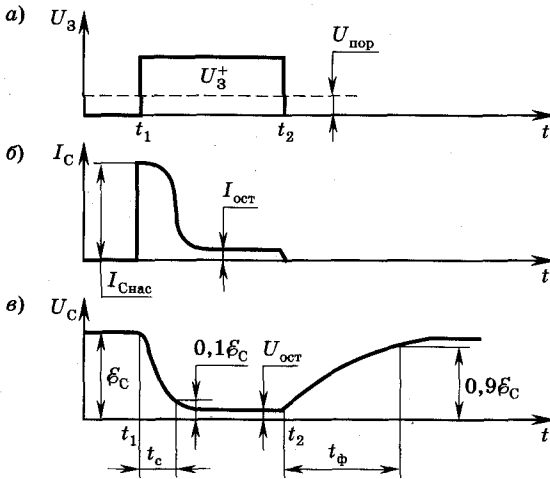


Рис. 9.6

гический 0 (ноль)) было по возможности меньшим, тогда рабочий ток $I_{СИ}$ определяется внешними элементами схемы, т. е.

$$I_{СИ} = (\epsilon_C - U_{ост})/R_C \approx \epsilon_C/R_C. \quad (9.1)$$

Для уменьшения $U_{ост}$ необходимо увеличивать сопротивление R_C и напряжение U_3^+ и, принципиально, $U_{ост} = U^0$ можно сделать сколь угодно малым. В то время как в биполярных ключах $U_{ост}$ принципиально ограничено напряжением $U_{КЭ}$.

На рис. 9.6 приведены временные диаграммы входного (а) и выходного (в) напряжений, а также тока стока (б). При поступлении входного импульса (при $U_3^+ > U_{пор}$) в момент t_1 транзистор открывается. Время формирования канала и установления тока равно по порядку величине времени пролета электронов через канал, которым можно пренебречь при достаточно большой нагрузочной емкости и считать, что ток стока I_C устанавливается мгновенно, достигая значения $I_{C\text{нас}}$ (см. рис. 9.6, б), соответствующего статической ВАХ при $U_C = \epsilon_C$. Ток I_C разряжает выходную емкость, и выходное напряжение понижается, при этом ток разряда $I_p = I_C(U_{\text{вых}}) - I_R$, где $I_C(U_{\text{вых}})$ определяется стоковой ВАХ ($U_{\text{вых}} = U_C$), а $I_R = (\epsilon_C - U_{\text{вых}})/R_C$. Напряжение на выходной емкости может быть вычислено из выражения

$$C_{\text{вых}}(dU_{\text{вых}}/dt) = -I_p(U_{\text{вых}}). \quad (9.2)$$

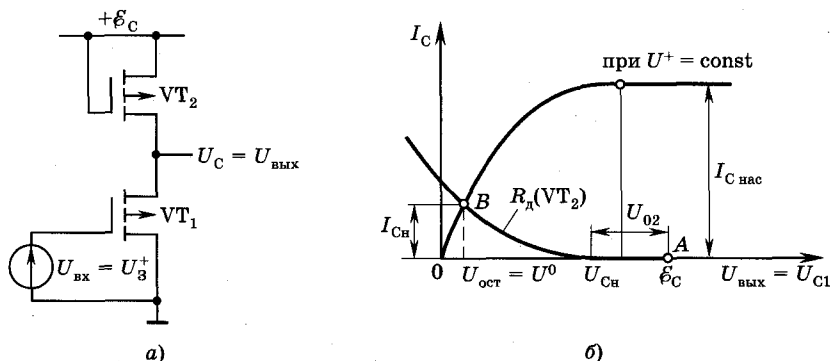


Рис. 9.7

Время спада t_c отсчитывается по уровням 0,1 и 0,9 (рис. 9.7, б) выходной амплитуды импульса ε_C . Анализ переходного процесса, включающий интегрирование уравнения (9.2) с учетом особенностей параметров схемы, позволяет вычислить время спада по формуле

$$t_c = \xi \varepsilon_C C_{\text{вых}} / I_{C \text{ нас}}, \quad (9.3)$$

где коэффициент $\xi = 1 \dots 2$ зависит от многих факторов (при оценках обычно принимают $\xi = 1,5$).

В момент выключения ($t = t_2$) канал мгновенно исчезает, поэтому $I_C = 0$. Выходная емкость заряжается через резистор R_C , и напряжение $U_{\text{вых}}$ возрастает до величины ε_C при времени нарастания (см. рис. 9.6, в) $t_n = t_\phi = 2,2 R_C C_{\text{вых}}$.

Для снижения t_n необходимо уменьшить R_C с одновременным пропорциональным увеличением тока $I_{C \text{ нас}}$, чтобы сохранить малое значение $U_{\text{ост}}$. Поскольку в цифровых ИС выход логического элемента соединяется со входами аналогичных элементов, то выходная емкость $C_{\text{вых}} = C_n$ для каждого элемента определяется емкостью затвор—исток (пропорциональной длине канала), которая может составлять сотые доли пикофарады. Более высокое быстродействие (примерно на порядок) достигается на комплементарных транзисторах с $t_c \approx t_n \leq 0,5$ нс.

В МДП-ключах с динамической нагрузкой вместо резистора используются транзисторы. В схеме на рис. 9.7, а роль динамической нагрузки выполняет нагрузочный транзистор VT_2 , у кото-

рого затвор соединен со стоком, при этом VT_1 является активным транзистором. При соединении затвора со стоком $U_{зи2} = U_{си2}$ справедливо неравенство $U_{зи2} - U_{пор} < U_{си2}$ (где $U_{зи2}$ — напряжение затвор—исток VT_2 , $U_{си2}$ — напряжение сток—исток VT_2), т. е. VT_2 работает на пологом участке выходной характеристики. ВАХ VT_2 с закороченным на сток затвором можно построить по семейству стоковых характеристик, используя равенство $U_{зи2} = U_{си2}$, т. е. сначала задается одно значение $U_{си}$ по оси абсцисс исходного семейства выходных ВАХ обычного VT_2 и проводится вертикальная линия до пересечения с той кривой, у которой $U_{зи2} = U_{си2}$, в результате чего получается одна точка (см. рис. 9.7, б, точка В). После этого аналогичная операция реализуется для других значений $U_{си2}$ и $U_{зи2}$, вплоть до значения $U_{си2} = U_{ин}$, где $U_{ин}$ — ЭДС источника питания. Полученная ВАХ позволяет построить нагрузочную линию R_d (VT_2) транзистора VT_1 (см. рис. 9.7, б), аналогичную нагрузочной прямой при резистивной нагрузке R_c (см. рис. 9.5, б). Однако эта характеристика будет начинаться не в точке А, где $U_{си} = \varepsilon_c$, а в точке $U_{сн} = \varepsilon_c - U_{02}$ (где $U_{02} = U_{зи\ пор2}$ — пороговое напряжение для VT_2). В открытом состоянии ключа рабочая точка лежит на квазилинейном участке характеристики активного транзистора VT_1 (точка В на рис. 9.7, б). Остаточное напряжение $U_{ост}$ в этой точке (обычно $U_{ост} \sim 50 \dots 100$ мВ) в несколько раз меньше соответствующего значения для МДП-ключа с резистивной нагрузкой (300 мВ и более).

Если в качестве нагрузки использовать МДП-транзистор с типом канала, противоположным основному активному, то такая схема называется **комплементарным ключом (инвертором)** (рис. 9.8, а). Подложка нагрузочного p -канального транзистора VT_2 подключается к точке с наибольшим потенциалом, а подложка основного n -канального транзистора VT_1 подключается к точке с наименьшим потенциалом. В результате исключается отпирание изолирующих p — n -переходов, которые обеспечивают изоляцию каналов МДП-структур от подложки. Если $U_{вх} = 0$, то $U_{зи1} = 0$, а $U_{зи2} = -U_{ин} < 0$, следовательно, p -канальный транзистор VT_2 — открыт, а n -канальный транзистор VT_1 — закрыт (для него $U_{ин} = \varepsilon_c > |U_{пор2}|$). В результате ток в общей цепи $I_{ост}$

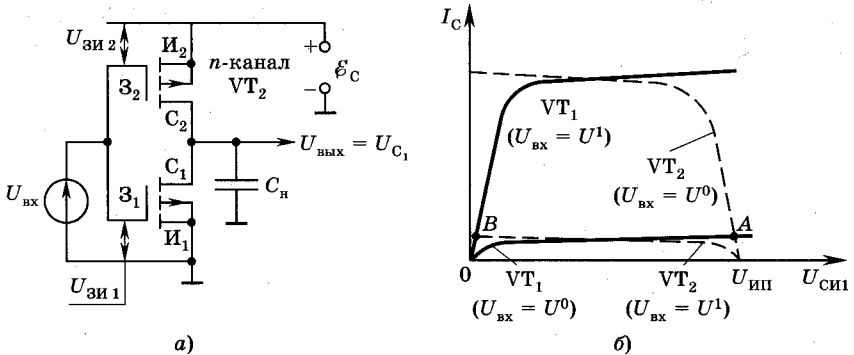


Рис. 9.8

определяется сопротивлением закрытого транзистора VT_1 . Ток $I_{\text{ост}}$ имеет очень малые значения, так как сопротивление закрытого транзистора VT_1 очень велико, а напряжение $|U_{\text{СИ}2}|$ — очень мало.

Величину $I_{\text{ост}}$ можно определить, если найти точку пересечения выходной (стоковой) характеристики закрытого VT_1 и выходной характеристики открытого VT_2 при известном $U_{\text{ЗИ}2}$ (точка A на рис. 9.8, б, где $U_{\text{СИ}} = U_C \approx U_{\text{ип}}$). Таким образом, VT_2 выступает как нелинейная нагрузка для VT_1 .

Если же на вход подать положительный сигнал величиной $U_{\text{вх}} = \varepsilon_C$, то $U_{\text{ЗИ}1} = \varepsilon_C > U_{\text{пор}1}$ и $U_{\text{ЗИ}2} = 0$. В этом случае VT_1 — открыт, а VT_2 — закрыт, при этом ток $I_{\text{ост}}$ в общей цепи остается на том же уровне, что и в предыдущем случае, хотя транзисторы переключились (точка B на рис. 9.8, б). Следовательно, комплементарные ключи потребляют крайне малую мощность в обоих состояниях, что является их важнейшим достоинством.

Однако оба устойчивых состояния сильно различаются по уровню выходного напряжения. В первом рассматриваемом случае при $U_{\text{вх}} = 0$ выходное напряжение $U_{\text{вых}} = \varepsilon_C = U_{\text{ип}} = U^1$, а во втором — при $U_{\text{вх}} = \varepsilon_C = U_{\text{ип}}$ выходное напряжение $U_{\text{вых}} = U_{\text{ост}} = U^0$. Остаточное напряжение $U_{\text{ост}}$ может иметь очень малые значения (единицы микровольт и менее), что является также достоинством рассматриваемых ключей.

Кроме полезных сигналов, на входы ключей всегда воздействуют внешние электромагнитные помехи. Ключи (инверторы)

не должны реагировать на помехи. Нечувствительность ключей к паразитным сигналам и шумам называется *помехозащищенностью* или *помехоустойчивостью*. Помехоустойчивость измеряется максимальной абсолютной величиной сигнала (обычно в вольтах), которая не вызывает ложного переключения.

9.4. Бистабильные схемы и триггеры

Бистабильные схемы и триггеры строятся на основе цепочек ключей, у которых присутствуют не только прямые связи между ключами, но и положительные обратные связи. В качестве элемента памяти наиболее часто используются бистабильные ячейки, которые имеют два устойчивых состояния, соответствующих логическому 0 или логической 1. Если в последовательной цепочке ключей выделить пару соседних ключей, изолировав рассматриваемую пару от предыдущих и последующих звеньев цепочки и соединив выход последующего (второго) ключа Q_2 со входом предыдущего (первого) и, наоборот, выход первого Q_1 со входом второго (рис. 9.9, а), мы получим схему, которая называется *бистабильной ячейкой (БЯ)* (условное обозначение БЯ дано на рис. 9.9, б). Бистабильная ячейка вместе с цепью управления составляют схему любого триггера. Хотя схеме бистабильной ячейки соответствует симметричная конфигурация, электрически она является асимметричной, поскольку один из ключей заперт, а второй открыт и насыщен. Цель управления бистабильной ячейкой состоит в том, чтобы с помощью внешних сигналов задавать то или иное из двух устойчивых состояний.

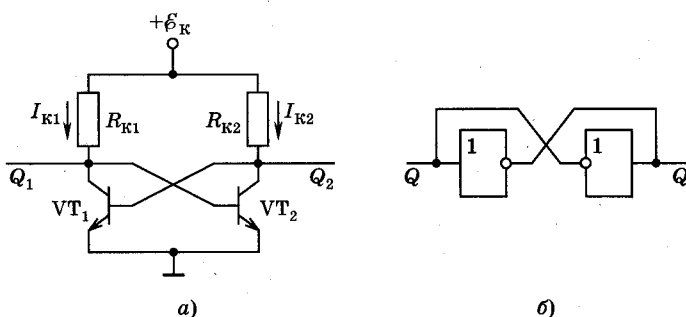


Рис. 9.9

Процессы переключения, т. е. переход транзистора из одного в другое состояние, рассмотрим на примере одной из распространенных схем триггера — триггера Шмитта (рис. 9.10). В основе триггера Шмитта лежит схема переключателя тока, рассмотренная ранее в п. 9.2, в которую включен делитель напряжения (R_1, R_2).

Предположим, что в исходном состоянии транзистор VT_1 заперт, т. е. $U_{K1} = \mathcal{E}_K$, а VT_2 находится в активном режиме, т. е. $U_{K2} = \mathcal{E}_K - I_0 R_K > U_{B2} = \mathcal{E}_K R_2 / (R_1 + R_2)$. Пусть в какой-то момент входной сигнал $U_{вх}$

превышает напряжение отпираания на некоторую малую величину, тогда появляется некое малое приращение тока ΔI_{K1} , которое вызывает следующие приращения потенциалов:

$$\Delta U_{K1} = -\Delta I_{K1} R_K, \Delta U_{B2} = R_2 \Delta U_{K1} / (R_1 + R_2), \Delta U_{\mathcal{E}} = \Delta U_{B2}.$$

Изменение $\Delta U_{\mathcal{E}}$ приведет к дополнительному приращению коллекторного тока первого транзистора $\Delta I'_{K1} > \Delta I_{K1}$, т. е. в схеме развивается лавинообразный процесс, в результате которого ток I_0 переходит в транзистор VT_1 и он открывается, а VT_2 — закрывается.

Входное напряжение $U_{вх}$, которое вызывает скачкообразное переключение VT_1 и VT_2 называют *напряжением (порогом) срабатывания триггера*. Чтобы вернуть триггер в исходное состояние, нужно уменьшить входной потенциал до значения, близкого к первоначальному значению U_{B2} , при котором начинает отпираться VT_2 .

Входное напряжение $U_{вх}$, при котором происходит скачкообразное переключение тока из VT_1 в VT_2 , называется *напряжением (порогом) отпускания триггера*. Анализ показывает, что напряжение отпускания меньше напряжения срабатывания.

Выходной сигнал в рассматриваемом триггере снимается с коллектора VT_2 , который не охвачен обратной связью. Кроме того, VT_2 работает в активном (ненасыщенном) режиме. В результате длительность переходных процессов из-за практического от-

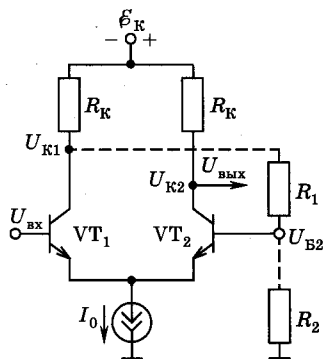


Рис. 9.10

сутствия накопления и рассасывания носителей существенно меньше, чем у бистабильной ячейки, где транзисторы работают в режиме насыщения.

9.5. Логические элементы на биполярных транзисторах

Электронные схемы, выполняющие простейшие логические операции, называются *логическими элементами (ЛЭ)*. Логические элементы используются в большинстве цифровых микросхем и во многом определяют их параметры. Логические функции представляют собой операции над логическими переменными (величинами) A , B , C и т. д. Логические переменные характеризуются двумя взаимоисключающими значениями: включено и выключено, истинно и ложно, есть и нет и т. д. Для операций с такими логическими величинами используется двоичный код, в котором величина $A = 1$, а $\bar{A} = 0$ (через \bar{A} обозначено «не A »), т. е. каждая логическая переменная принимает только одно значение: 1 или 0.

Все логические функции с любым числом логических переменных можно получить с помощью трех основных операций:

- 1) логического отрицания (инверсии, т. е. операции НЕ), которое записывается в виде $B = \bar{A}$;
- 2) логического сложения (дизъюнкции, т. е. операции ИЛИ), которое записывается для двух переменных в форме $C = A + B$, причем $C = 1$ при $A = 1$ или $B = 1$, или же при $A = B = 1$;
- 3) логического умножения (конъюнкции, т. е. операции И), которая для двух переменных представляется как $C = A \cdot B$, т. е. $C = 1$, когда $A = B = 1$.

Комбинация логических операций НЕ и ИЛИ дает более сложную функцию ИЛИ—НЕ, записываемую как $C = \overline{A + B}$. Здесь значения C противоположны ее значениям для операции ИЛИ. Сочетание операций НЕ и И образует функцию И—НЕ и записывается как $C = A \cdot B$. Схемные ЛЭ реализуют обычно одну или несколько функций: НЕ, И, ИЛИ, И—НЕ, ИЛИ—НЕ, условные обозначения которых представлены на рис. 9.11. Соединяя соответствующим образом эти схемные ЛЭ, построенные в основном на рассматриваемых выше ключах (инверторах), можно получить микросхему, выполняющую любую более сложную

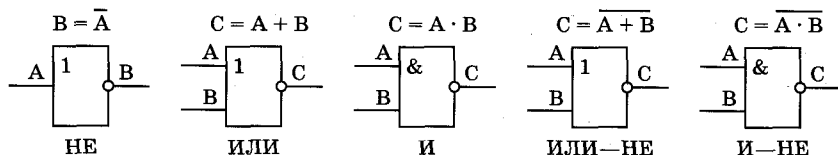


Рис. 9.11

операцию. В ЛЭ логические 0 и 1 представляются напряжением или уровнем нуля U^0 и напряжением или уровнем единицы U^1 . Разность уровней единицы и нуля называют **логическим перепадом** $U_{л} = U^1 - U^0$.

ЛЭ подразделяются по режиму работы на статические и динамические. Статические ЛЭ могут работать в статическом и импульсном режимах, а динамические ЛЭ работают только в импульсном режиме. Статические ЛЭ наиболее широко используются в современных ИС. Наибольшее распространение получили ЛЭ на биполярных и МДП-транзисторах.

Схемные элементы интегральных логических элементов называют транзисторными логиками. Наиболее распространенными транзисторными логиками на биполярных транзисторах являются: **транзисторно-транзисторная логика, эмиттерно-связанная логика, интегральная инжекционная логика.**

Отличительным признаком ЛЭ **транзисторно-транзисторной логики (ТТЛ)** является многоэмиттерный транзистор (МЭТ) (см. гл. 7) VT_1 , включенный во входную цепь ЛЭ (рис. 9.12). В базовой цепи VT_1 включен резистор R_1 и выходной инвертор (ключ) на VT_2 с резистором R_2 в коллекторной цепи. Простейший элемент ТТЛ, приведенный на рис. 9.12, выполняет логическую функцию И-НЕ ($C = \overline{A \cdot B}$), при этом МЭТ VT_1 выполняет логическую операцию И над переменными А и В. Такие элементы используются в больших интегральных схемах (БИС). Однако схема, приведенная на рис. 9.12, не нашла широкого применения вследствие низкой помехоустойчивости, малой нагрузочной способности, низкого быстродействия при работе на емкостную нагрузку и жестких требований к параметрам элементов.

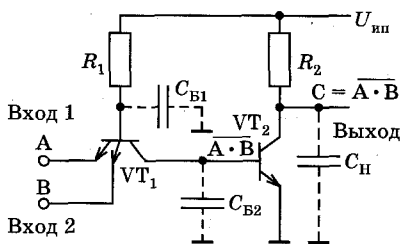


Рис. 9.12

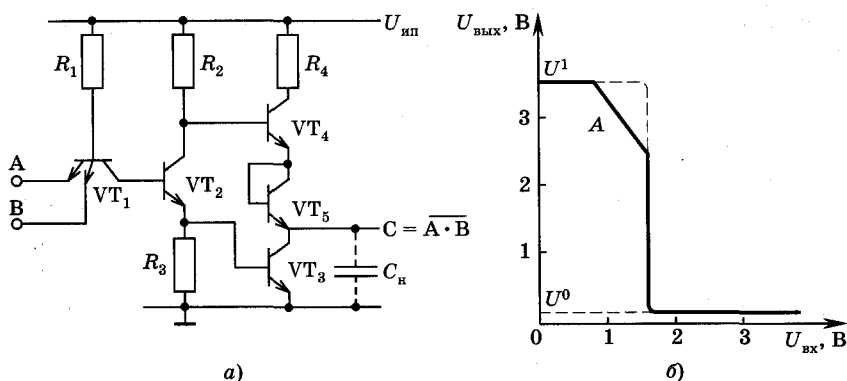


Рис. 9.13

Базовым элементом для микросхем малой и средней интеграции, а также в выходных каскадах БИС стал элемент со сложным инвертором (рис. 9.13, а). Этот элемент, как и предыдущий, реализует логическую функцию И—НЕ. Здесь VT_1 и R_1 выполняют ту же роль, что и в схеме на рис. 9.12. Остальная часть схемы представляет собой сложный инвертор, где VT_2 и R_2, R_3 — промежуточный каскад, а VT_3, VT_4, VT_5 и резистор R_4 — выходной каскад. Транзистор VT_5 используется в диодном включении ($U_{\text{БК}} = 0$). С коллектора и эмиттера VT_2 на базы VT_4 и VT_3 подаются управляющие сигналы, которые обеспечивают противофазное переключение этих транзисторов, т. е. если один из них включен, то другой выключен. Рассмотрим процессы в этой схеме. Предположим, на одном или нескольких входах $U_{\text{вх}} = U^0 = 0$, т. е. подан логический 0. Тогда соответствующие эмиттеры МЭТ VT_1 находятся под нулевым напряжением, их эмиттерные переходы смещены в прямом направлении. Анализ показывает, что в такой схеме коллекторный переход находится в насыщении. При этом ток в коллекторной цепи и напряжение на коллекторе $U_{\text{КЭ}}$ и, следовательно, на базе транзистора VT_2 близки к нулю. Напряжение $U_{\text{КЭ}}$ при $I_{\text{К}} = 0$ равно $U_{\text{КЭ}} = \varphi_T \ln(1 + 1/\beta_I)$ (где β_I — инверсный коэффициент передачи тока) и обычно не превышает $(2 \dots 3)\varphi_T$, поэтому транзистор VT_2 закрыт, VT_3 также закрыт, а VT_4 — открыт, поскольку в его базу втекает ток, задаваемый резистором R_2 . Поскольку VT_2 закрыт,

то напряжение на выходе соответствует напряжению высокого уровня, которое можно оценить по формуле

$$U^1 = U_{\text{ип}} - 2U'_{\text{БЭ}}, \quad (9.4)$$

где $2U'_{\text{БЭ}}$ — падение напряжения на эмиттерных переходах транзисторов VT_4 и VT_5 (величина $U'_{\text{БЭ}}$ обычно порядка 0,45...0,5 В). Через эти переходы протекает выходной ток, который является входным током последующих нагрузочных элементов.

При подаче U^1 на все входы VT_1 транзистор VT_2 открывается коллекторным током входного транзистора VT_1 и переходит в режим насыщения. Напряжение на его коллекторе понижается и транзистор VT_4 закрывается, а VT_3 открывается эмиттерным током VT_2 и насыщается. Выходное напряжение $U_{\text{вых}} = U^0$ определяется напряжением насыщения VT_3 . Назначение транзистора VT_5 заключается в том, чтобы поддерживать закрытое состояние VT_4 при понижении выходного напряжения.

Через резистор R_3 протекает базовый ток VT_3 во время процесса рассасывания. Резистор R_4 служит для ограничения импульсного тока VT_4 при переключении из состояния $U_{\text{вых}} = U^0$ в состояние $U_{\text{вых}} = U^1$.

Нагрузочная способность элемента ТТЛ со сложным инвертором в состоянии $U_{\text{вых}} = U^1$ за счет VT_4 больше примерно в β_4 раз по сравнению с простейшим элементом ТТЛ на рис. 9.12. Выходная характеристика элемента ТТЛ со сложным инвертером (рис. 9.13, б сплошная линия) подобна по форме выходным характеристикам обычных инверторов (ключей), но идет выше, поскольку в этом случае $U_{\text{вых}}$ больше из-за более высокого напряжения источника питания $U_{\text{ип}}$. Однако эта характеристика идет более полого при некоторых значениях $U_{\text{вх}}$ (участок А на рис. 9.13, б).

Пологий участок А выходной (передаточной) характеристики можно скорректировать за счет замены резистора R_3 более сложной корректирующей цепью, включающей, помимо резисторов, также и добавочный транзистор. В результате выходная характеристика на рис. 9.13, б примет на участке А форму, соответствующую пунктирной кривой.

Логический элемент со сложным инвертором занимает большую площадь на кристалле, чем элемент на рис. 9.12, и потребля-

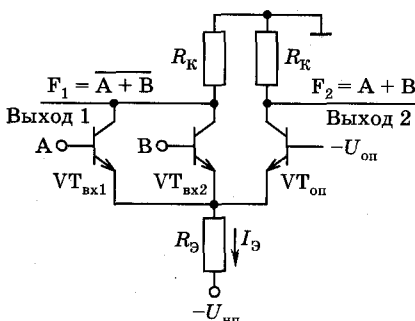


Рис. 9.14

ет большую мощность, поэтому его применение ограничено цифровыми ИС малой и средней степени интеграции. Для повышения быстродействия элементов ТТЛ в них используют транзисторы с диодом Шоттки или ЛЭ *эмиттерно-связанной логики (ЭСЛ)*, отличительной особенностью которых является использование переключателя тока (см. гл. 8)

с транзисторами, работающими в активном режиме, что исключает режим насыщения и связанное с ним рассасывание носителей, уменьшает временную задержку и повышает быстродействие.

Простейшими элементами ЭСЛ являются элементы *малосигнальной эмиттерно-связанной логики (МЭСЛ)* (рис. 9.14), которые используются во внутренних цепях средних интегральных схем (СИС) или БИС. Приведенная схема содержит два входных транзистора $VT_{вх1}$ и $VT_{вх2}$ (в действительности их может быть больше) и генератор тока в форме токозадающего резистора $R_э$. Эмиттеры всех транзисторов соединены (связаны), что и определяет название логики ЭСЛ. Особенностью применения МЭСЛ является использование отрицательного напряжения питания $-U_{нп}$, что ослабляет влияние его изменения на уровни U^0 и U^1 . В схеме рис. 9.14 на инверсном выходе 1 реализуется логическая функция ИЛИ—НЕ $\rightarrow F_1 = \overline{A + B}$, а на прямом выходе 2 — функция ИЛИ $\rightarrow F_2 = A + B$. Опорное напряжение $U_{оп}$, входные и выходные напряжения имеют отрицательную полярность из-за отрицательной полярности напряжения источника питания $U_{нп}$. При $U_{вх} = U^0 < -U_{оп}$ транзисторы $VT_{вх1}$ и $VT_{вх2}$ закрыты, а $VT_{оп}$ — открыт. Напряжение на инверсном выходе 1 равно U^1 и определяется падением напряжения на резисторе в коллекторной цепи опорного транзистора при протекании выходного тока $I_{вых}$ в нагрузку, т. е. $U^1 = -I_{вых}^1 R_к = nI_э R_к / (\beta + 1)$ (β — коэффициент передачи тока последнего ЛЭ). Это напряжение снижается с ростом числа нагрузок (n), что ограничивает нагрузочную способность этого элемента ЭСЛ в состоянии $U_{вых} = U^1$. В коллекторной цепи открытого опорного транзистора $VT_{оп}$ про-

текает ток $I_K \approx I_\Phi = -(U_{ин} - U_{оп} - U'_{БЭ})/R_\Phi$, следовательно, напряжение на прямом выходе 2 равно

$$U_{вых2}^0 \approx -(U_{ин} - U_{оп} - U'_{БЭ})R_K/R_\Phi. \quad (9.5)$$

Если $U_{вх} = U^1$, то соответствующие входные транзисторы, куда подано это напряжение, открываются, а $VT_{оп}$ закрывается, тогда

$$U_{вых1} = U_{вых1}^0 \approx -I_\Phi R_K = -(U_{ин} - U_{вх}^1 - U'_{БЭ})R_K/R_\Phi. \quad (9.6)$$

Из (9.6) видно, что напряжение низкого уровня $U_{вых1}^0$ (логический 0) на инверсном выходе уменьшается при росте входного напряжения из-за роста тока I_Φ , при этом на прямом выходе $U_{вых} = U_{вых1}^1$.

Логический перепад U_Δ для МЭСЛ составляет всего 0,3...0,5 В. Из-за такого малого логического перепада элементы МЭСЛ имеют сравнительно низкую помехоустойчивость. Потребляемая мощность практически одинакова для обоих состояний ЛЭ, поскольку ток I_Φ мало изменяется при переключении.

Для повышения быстродействия элементов МЭСЛ необходимо уменьшать емкости $p-n$ -переходов, паразитные емкости проводников, сопротивление базы, ограничивать число нагрузок и увеличивать граничную частоту транзисторов.

В случае повышенной емкости нагрузки ($C_H > 10$ пФ) и большом числе нагрузок (> 10) в быстродействующих ИС применяют более сложные элементы ЭСЛ (рис. 9.15). Схема на рис. 9.15 по сравнению с рассмотренной ранее содержит дополнительно

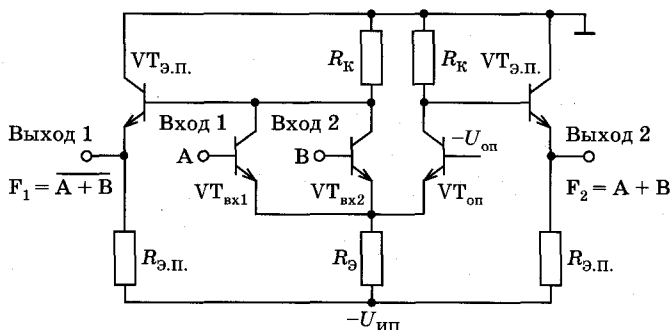


Рис. 9.15

два выходных эмиттерных повторителя $VT_{Э.П}$ с резисторами $R_{Э.П}$. Принцип работы этой схемы такой же, как и ранее рассмотренной, однако их основные свойства могут сильно различаться. Из-за большего напряжения питания элементы этой ЭСЛ по сравнению с рассмотренными в МЭСЛ имеют больший логический перепад, лучшие помехоустойчивость и нагрузочную способность, большие допустимую емкость нагрузки, потребляемую мощность и занимаемую на кристалле площадь. Элементы ЭСЛ с эмиттерными повторителями применяются в сверхбыстродействующих цифровых микросхемах малой и средней степени интеграции.

Многих из перечисленных недостатков лишены логические элементы, построенные на основе *интегральной инжекционной логики (И²Л)*, которые не имеют аналогов в дискретных транзисторных схемах и выполняются только в интегральном исполнении.

Характерной особенностью схем И²Л является индивидуальное питание базы каждого транзистора от индивидуального генератора тока. В конструкциях БИС с инжекционным питанием используются сложные биполярные структуры, в которых совмещаются области многоколлекторных транзисторов типов $n-p-n$ и $p-n-p$ (см. гл. 7). В силу такого совмещения и переплетения большого количества элементов ИС для ЛЭ И²Л не существуют (не используются) принципиальные электрические схемы. Как правило, их работа описывается с помощью эквивалентных схем. Индивидуальные генераторы тока реализуются чаще всего с помощью МКТ, включенных по схеме с ОБ. На

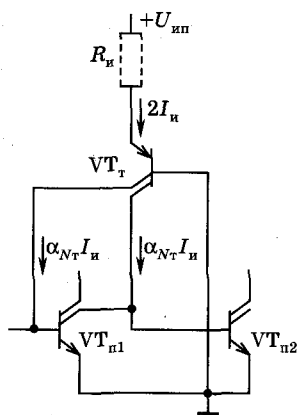


Рис. 9.16

рис. 9.16 изображена эквивалентная схема двух последовательно включенных ЛЭ [6], где VT_T двухколлекторный токозадающий транзистор $p-n-p$, подсоединенный через токозадающий резистор $R_и$ к цепи питания, являющейся общей для всего кристалла БИС. Токозадающий транзистор обычно является многоколлекторным с числом коллекторов n , достигающим до нескольких сотен.

Транзистор VT_T задает постоянные токи $\alpha_{N_T} I_и$ в базы $n-p-n$ -транзисторов VT_{n1} и VT_{n2} , где $I_и$ — ток инжектора, приходящийся на один логический эле-

мент. Транзисторы $VT_{п1}$ и $VT_{п2}$ называются переключательными и включены по схеме с ОЭ, а токозадающий VT_T — по схеме с ОБ.

Термин «инжекционное питание» означает, что питающие токи образуются за счет инжекции неравновесных дырок в эмиттерную область инжектора через инжекторный $p-n$ -переход, напряжение на котором составляет $0,7 \dots 0,8$ В (при $T^\circ = 25^\circ \text{C}$); токи питания $I_{п}$ одинаковы и равны

$$I_{п} = \alpha_{N_T} I_{п}.$$

Если на вход $VT_{п1}$ подано напряжение $U_{вх1} = U_0 \approx 0$, определяемое насыщенным (открытым) транзистором предыдущего логического элемента, то ток $I_{п}$ этого элемента ответвляется во внешнюю цепь (в цепь предыдущего ЛЭ), транзистор $VT_{п1}$ закрыт и его ток базы равен нулю. Коллекторный ток $VT_{п1}$ тоже равен нулю и весь ток $I_{п}$ генератора тока течет в базу транзистора $VT_{п2}$ (он открыт и его прямое напряжение база—эмиттер равно $0,5 \dots 0,7$ В). На выходе $VT_{п1}$ (коллектор этого транзистора) будет такое же напряжение. Это напряжение соответствует U^1 , т. е. при $U_{вх1} = U^0$ на выходе транзистора $VT_{п1}$ будет $U_{вых} = U^1$.

При $U_{вх1} = U^1$ ток $I_{п}$ втекает в базу $VT_{п1}$, поскольку $U_{вх}^1$ определяется коллекторным напряжением закрытого транзистора предыдущего ЛЭ. Транзистор $VT_{п2}$ открыт, и ток в его коллекторной цепи равен тоже $I_{п}$, поскольку он вытекает из входа $VT_{п2}$, т. е. в этом случае для $VT_{п1}$ мы имеем $I_B = I_K = I_{п}$, транзистор $VT_{п1}$ находится в режиме насыщения, и напряжение на его выходе равно $U_{вых1} = U^0$. Аналогичная ситуация будет и для второго выхода $VT_{п1}$, если к нему подключены последующие ЛЭ. Таким образом, логическая операция инверсии выполняется по всем выходам.

Входная характеристика рассмотренного элемента $U_{вэ} = f(I_B)$ при заданном токе инжектора аналогична по форме, а при $I_{п} = 0$ совпадает с характеристикой обычного транзистора. При увеличении $I_{п}$ характеристика, сохраняя форму, сдвигается по оси абсцисс в левую сторону. Рабочая область характеристик располагается при $I_B \leq 0$. При заданном токе инжектора значение $U_{вэ} = U^1$ для $I_B = 0$. Структура, в которой реализована эквивалентная схема на рис. 9.16, имеет низкие значения коэффици-

ента передачи $\beta_{\Pi} = 5 \dots 10$ переключающего транзистора из-за его инверсного включения. Минимальный логический перепад для рассматриваемого элемента составляет $U_{\text{л}} = (5 \dots 6)\varphi_{\text{T}}$.

К каждому выходу ЛЭ может быть подсоединен только один нагрузочный элемент, следовательно, нагрузочная способность равна числу коллекторов переключающего транзистора. Число коллекторов (n) ограничено коэффициентом β , который уменьшается пропорционально $1/n$, и ростом сопротивления базы, что приводит к увеличению падения напряжения на нем с увеличением n , а это вызывает неравномерное смещение эмиттерного p — n -перехода. Типичные значения $n = 2 \dots 4$, но для специальных структур эта величина может быть $10 \dots 20$.

Средняя мощность $P_{\text{ср}}$, потребляемая элементом И²Л, равна $P_{\text{ср}} = U'_{\text{ИЭ}} I_{\text{и}}$, где $U'_{\text{ИЭ}}$ — прямое напряжение на инжекторном p — n -переходе. Величина $P_{\text{ср}}$ определяется токозадающим резистором $R_{\text{и}}$, который обычно размещают вне микросхемы. Изменяя $R_{\text{и}}$ или $U_{\text{ин}}$, можно в пределах трех-четырех порядков регулировать ток питания ЛЭ, мощность и быстродействие.

Минимальная средняя задержка, т. е. максимальное быстродействие, определяется временем рассасывания носителей в областях переключающего транзистора. Для рассматриваемой структуры наибольший заряд неосновных неравновесных носителей (дырок) накапливается в высокоомном эмиттерном слое вследствие инжекции дырок из базовой области в режиме насыщения переключающего транзистора. Минимальная средняя задержка определяется эффективным временем жизни дырок и составляет $10 \dots 20$ нс. Низкое быстродействие таких ЛЭ является их главным недостатком. В связи с этим было предложено большое число новых структур И²Л с повышенным быстродействием, например структура элемента с диодами Шоттки двух типов и структуры, создаваемые путем многократных операций совмещения. Такие структуры перспективны для использования в цифровых СБИС, поскольку обеспечивают минимальную задержку, не превышающую 1 нс, и энергию переключения $\sim 0,03 \dots 0,05$ пДж в режиме малых токов инжектора. Элементы И²Л отличаются от других элементов на биполярных транзисторах малой потребляемой мощностью и малой занимаемой площадью на кристалле при большой средней задержке и плохой помехоустойчивости.

9.6. Логические элементы на полевых транзисторах

Очень большое входное сопротивление МДП-транзисторов позволяет создавать ЛЭ динамического типа с малой площадью на кристалле и малой потребляемой мощностью при относительно низком быстродействии. ЛЭ арсенид-галлиевых ИС на МЭП-транзисторах (транзисторах с контактом металл—полупроводник) с каналами n -типа обладают сверхвысоким быстродействием. В основе ЛЭ на ПТ лежат инверторы (ключи) на n -канальных МДП-транзисторах и комплементарных транзисторах (раздел 9.3). Рассмотрим логические элементы И—НЕ и ИЛИ—НЕ на n -канальных и комплементарных транзисторах.

В отличие от элементарного инвертора (ключа) (см. рис. 9.5) в ЛЭ И—НЕ вместо одного включено m активных транзисторов, которые при той же структуре и напряжениях дают в m раз меньший ток. На рис. 9.17 изображен ЛЭ И—НЕ с двумя последовательно включенными активными транзисторами VT_{01} и VT_{02} и один пассивный VT_{Π} [6]. Если на вход 1 или вход 2, или на оба входа одновременно подать напряжение низкого уровня U^0 , то либо один из соответствующих транзисторов, или оба транзистора VT_{01} и VT_{02} закрыты, ток через пассивный транзистор VT_{Π} не протекает и на выходе устанавливается напряжение высокого уровня $U^1 = U_{\text{ин1}}$. Если же на входы поступает напряжение U^1 , то все активные транзисторы открыты и на выходе устанавливается напряжение U^0 .

Рассматриваемые логические элементы имеют приблизительно те же характеристики и параметры, что и инвертор на рис. 9.5, если ввести эффективную удельную крутизну активного транзистора $K_{\text{эф}} = K_{\text{a}}/m$, где K_{a} — удельная крутизна отдельного транзистора.

Передаточные характеристики ЛЭ определяются отношением $K_{\Pi}/K_{\text{эф}}$ (K_{Π} — эффективная удельная крутизна пассивного транзистора). Чтобы сохранить эти характеристики неизменными, ЛЭ должен иметь параметр K_{Π}/K_{a} в m раз меньше, чем у инвертора. На практике это отношение уменьшают за счет K_{Π} , однако при этом пропорционально m увели-

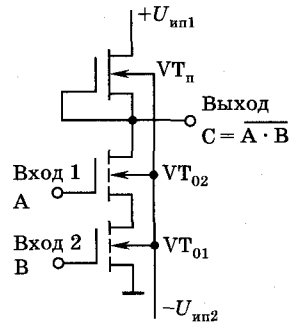


Рис. 9.17

чивается время переключения, т. е. быстродействие получается хуже, чем у отдельного инвертора. Быстродействие можно сохранить на уровне инвертора, но при этом повышается уровень логического нуля U^0 и снижается помехоустойчивость.

Для реализации логической операции ИЛИ—НЕ применяется параллельное включение активных транзисторов. Если хотя бы на один из входов подается напряжение U^1 , то соответствующий активный транзистор открыт и на выходе устанавливается потенциал U^0 . При $U_{\text{вых}} = U^0$ на всех выходах все активные транзисторы закрыты и на выходе устанавливается напряжение U^1 , равное напряжению питания.

Передаточная характеристика, напряжения U^0 , U^1 и помехоустойчивость будут такими же, как у инвертора при $U_{\text{вх}} = U^0$ на одном из входов и изменяющемся напряжении на другом. Если на обоих входах напряжение изменяется одновременно, то U^0 уменьшается и помехоустойчивость растет.

В отличие от элемента И—НЕ быстродействие ИЛИ—НЕ выше и мало зависит от числа его входов, поскольку емкости транзисторов ЛЭ составляют лишь малую часть общей емкости $C_{\text{н}}$.

ЛЭ на комплементарных транзисторах И—НЕ (рис. 9.18, а) с последовательным и параллельным включением соответственно n -канальных и p -канальных транзисторов соответственно имеют характеристики и параметры, близкие к инвертору, представленному на рис. 9.8, а, эффективная удельная крутизна транзисторов которого $K_{n\text{эф}} = K_n/m$, $K_{p\text{эф}} = mK_p$ (индекс n относится к n -канальному, а p — к p -канальному МДП-транзистору). При тех же геометрических размерах транзисторов, что и в

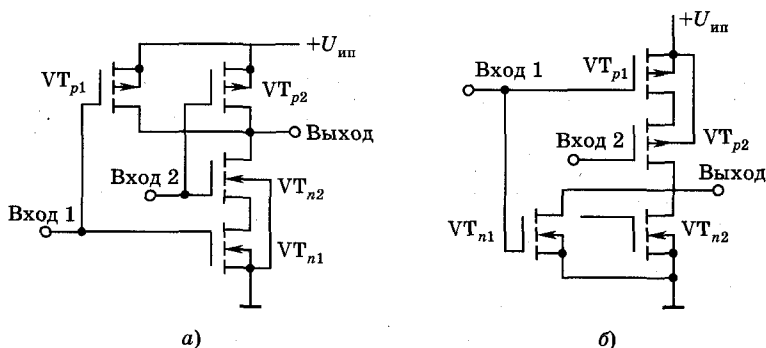


Рис. 9.18

инверторе, ток, задаваемый n -канальными транзисторами в открытом состоянии, уменьшается в m раз, а ток, задаваемый p -канальными транзисторами, увеличивается в m раз. С ростом m отношение $K_{n\text{эф}}/K_{p\text{эф}}$ уменьшается и помехоустойчивость $U_{\text{н}}^1$ падает. С изменением m средняя временная задержка изменяется сравнительно незначительно по сравнению с элементом на n -канальных транзисторах. Для $m \geq 5$ средняя задержка растет пропорционально m .

Функция ИЛИ—НЕ реализуется с помощью параллельного включения n -канальных и последовательного включения p -канальных транзисторов (рис. 9.18, б). В этом логическом элементе параметры $K_{n\text{эф}}$ и $K_{p\text{эф}}$ в сравнении с инвертором равны $K_{n\text{эф}} = mK_n$, $K_{p\text{эф}} = K_p/m$. С ростом m уменьшается помехоустойчивость по логическому нулю $U_{\text{н}}^0$, в противоположность предыдущему случаю, когда уменьшается помехоустойчивость $U_{\text{н}}^1$ по логической единице. Средняя временная задержка увеличивается пропорционально m , т. е. сильнее по сравнению с элементом И—НЕ. Последние с точки зрения быстродействия являются более предпочтительными, чем ЛЭ ИЛИ—НЕ.

Помимо рассмотренных ЛЭ на МДП-транзисторах используются также так называемые *динамические схемы*, в которых происходит кратковременное запоминание информации с использованием конденсаторов, сформированных емкостями самих транзисторов с индивидуальными каналами.

В логических элементах сверхбыстродействующих ИС используются МЭП-транзисторы на основе GaAs.

9.7. Элементы полупроводниковых запоминающих устройств

Цифровые полупроводниковые ИС памяти используются в *оперативных (ОЗУ) и постоянных (ПЗУ) запоминающих устройствах*. ПЗУ хранят информацию при отключении источника питания, тогда как в ОЗУ она теряется. *Статические ОЗУ* памяти могут хранить информацию в течение длительного времени, а *динамические ОЗУ* — ограниченное время. Статические ОЗУ обладают максимальным быстродействием, а динамические ОЗУ обеспечивают максимальную информационную емкость и минимальную потребляемую мощность. Большая часть БИС памяти создаются на МДП-транзисторах, а ИС памяти — на биполярных

транзисторах, которые обладают повышенным быстродействием, но меньшей информационной емкостью.

Оперативные ЗУ состоят из накопителя и схем управления. Данные, которые необходимо запомнить, хранятся в накопителе. Схемы управления включают усилители, разного рода ключи, коммутаторы, дешифраторы и т. д.

Накопитель состоит из элементов памяти в основном на базе бистабильных ячеек (см. п. 9.4), каждая из которых хранит один бит информации, соответствующей хранению логических 0 и 1.

Чаще всего бистабильная ячейка представляет собой симметричный триггер, содержащий два инвертора с перекрестными обратными связями; выход первого инвертора соединен со входом второго, а выход второго — со входом первого.

В ОЗУ используется достаточно много типов запоминающих ячеек, некоторые из них, наиболее распространенные, представлены на рис. 9.19.

Ячейка на МДП-транзисторах с p -каналами представляет собой триггер (транзисторы $VT_1—VT_4$) с управляющими ключами VT_5 и VT_6 , соединенных шинами столбца Y' и Y'' (рис. 9.19, а). При отсутствии выборки напряжение на шине X близко к нулю, транзисторы VT_5 и VT_6 закрыты, триггер отключен от шин столбца и элемент памяти хранит ранее записанную информацию.

При записи информации на одну из шин столбца подают напряжение U^0 , а на другую — напряжение U^1 , после этого на адресную шину X поступает положительный импульс с амплитудой, близкой к напряжению источника питания $U_{ин}$, который открывает транзисторы VT_5 и VT_6 и в точках A и B устанавлива-

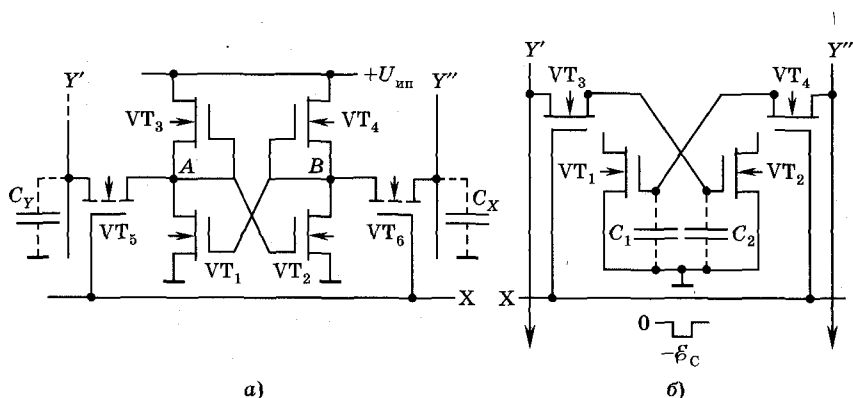


Рис. 9.19

ются такие же напряжения, что и на шинах Y' , Y'' , и триггер находится в необходимом состоянии.

В режиме считывания при поступлении на шину X импульса выборки VT_5 и VT_6 отпираются и на шинах столбца устанавливаются напряжения, соответствующие состоянию триггера (U^0 на одной из шин и U^1 на другой), которые воспринимаются усилителем считывания. Таким образом, импульс на адресной шине в обоих режимах играет роль тактового импульса.

На рис. 9.19, б изображена запоминающая ячейка динамического типа, в которой информация сохраняется с помощью конденсаторов C_1 и C_2 , сформированных транзисторами. Алгоритм записи и считывания аналогичен предыдущему случаю.

При записи на шины Y' и Y'' поданы соответственно уровни 0 и $-E_C = U^1$. Уровень $-E_C$ через ключ VT_4 поступает на затвор VT_1 , который будет открыт. На затвор VT_2 подается уровень 0 и он будет закрыт. На емкостях C_1 и C_2 напряжения будут иметь значения соответственно $U_{C1} = -E_C$, $U_{C2} = 0$. Остаточный ток закрытого VT_2 мал, и конденсатор C_1 будет разряжаться очень медленно. Следовательно, U_{C1} и U_{C2} будут сохраняться длительное время.

Для поддержания напряжения на емкости постоянным при ее неизбежном разряде при считывании осуществляют регенерацию, т. е. периодически производят запись того же кода. Динамические запоминающие ячейки из-за отсутствия источника питания в режиме хранения не потребляют мощности, поэтому они экономичнее статических.

Запоминающие ячейки на МДП(МОП)-транзисторах экономичнее и компактнее по сравнению с ячейками на биполярных транзисторах. Однако последние обладают лучшим быстродействием, чем МДП-ячейки.

Среди БТ наибольшее распространение получили ячейки памяти статического типа. Пример такой ячейки на основе МЭТ VT_1 и VT_2 изображен на рис. 9.20 [6]. Шина строки X' выполняет также роль

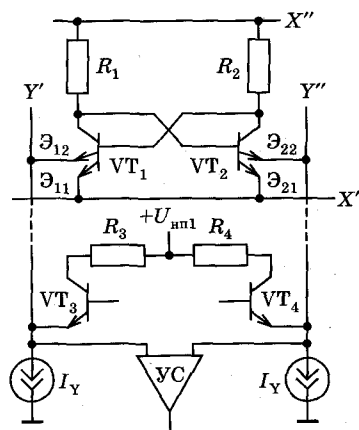


Рис. 9.20

линии питания, на нее подается минусовой потенциал. Эмиттеры \mathcal{E}_{12} и \mathcal{E}_{22} соединены с шинами столбца Y' и Y'' и применяются для записи и считывания.

Вторая шина строки X'' , которая также используется как шина питания, на нее подается плюс. Транзисторы VT_3 и VT_4 вместе с генераторами I_Y и резисторами не входят в элемент памяти и служат для его управления.

В режиме хранения при одном устойчивом состоянии VT_1 открыт и насыщен, а VT_2 закрыт, т. е. $U_{KЭ1} = U_{KЭ \text{ нас}} \approx 0,1 \text{ В}$, $U_{KЭ2} \approx \approx U'_{\text{ип}} = U_{X''} - U_{X'}$. В другом устойчивом состоянии, наоборот, VT_1 закрыт, а VT_2 насыщен, т. е. $U_{KЭ1} = U'_{\text{ип}}$, $U_{KЭ2} = U_{KЭ \text{ нас}}$. На базы VT_3 и VT_4 подают одинаковые напряжения. На шинах Y' и Y'' устанавливаются также одинаковые напряжения, при этом в управляющих эмиттерах \mathcal{E}_{12} , \mathcal{E}_{22} токи практически отсутствуют.

В режиме считывания на шине X' повышается напряжение, на шине X'' также повышается напряжение на такую или большую величину. При VT_1 открытом, а VT_2 закрытом напряжение U_{B2} на VT_2 увеличивается так же, как и на шине X' . Ток в эмиттере \mathcal{E}_{21} равен нулю, поэтому напряжение на шине Y'' ($U_{Y''}$) не изменяется. В управляющем эмиттере \mathcal{E}_{12} транзистора VT_1 появляется ток считывания. Напряжение $U_{Y'}$ повышается, а эмиттерный переход VT_3 запирается. Напряжение U_{B1} транзистора VT_1 в первый момент скачком изменяется, управляющий эмиттерный переход отпирается и U_{B1} начинает изменяться, после чего управляющий эмиттерный переход отпирается и U_{B1} начинает изменяться с той же скоростью, что и $U_{Y'}$. На шинах Y' и Y'' возникает разность напряжений, поступающая на усилитель считывания.

Задержка между поступлением импульса выборки на шину X' и моментом срабатывания усилителя считывания (время считывания) определяется процессом заряда емкости шины C_Y током элемента памяти. При VT_1 насыщенном, а VT_2 закрытом в режиме записи одновременно с подачей импульса выборки на шину X' повышается напряжение на базе транзистора VT_4 . В результате VT_4 запирается и большой ток генератора I_Y течет

через эмиттер \mathcal{E}_{22} , переводя VT_2 в режим насыщения, при этом напряжение на коллекторе понижается. Транзистор VT_1 в результате запирается. На \mathcal{E}_{12} будет обратное напряжение из-за роста напряжения на базе VT_3 и на шине Y' .

Недостатком рассмотренной запоминающей ячейки является относительно низкое быстродействие из-за процесса рассасывания избыточных неосновных носителей в транзисторе, переключающемся из режима насыщения в закрытое состояние в режиме записи. Время считывания имеет заметную величину из-за малого тока считывания $I_{сч}$.

Для увеличения $I_{сч}$ подключают генератор тока к шинам X' и X'' , а ускорение времени рассасывания достигается введением в схему диодов, включаемых параллельно R_1 и R_2 .

Из известных биполярных структур для СБИС наиболее пригодны структуры с инжекционным питанием.



Контрольные вопросы

1. В чем состоят особенности цифровых ИС?
2. Инверторы на МДП-транзисторах: разновидности, особенности, быстродействие, помехоустойчивость, принципы работы.
3. Базовые цифровые схемы на биполярных транзисторах: особенности, принципы работы, параметры и характеристики.
4. Вистабильные схемы и триггеры: схемные разновидности, принципы работы, выполняемые функции, параметры и характеристики.
5. Логические элементы на биполярных транзисторах: основные логические операции и их схемные реализации, принципы функционирования, достоинства и недостатки, характеристики.
6. Логические элементы на полевых транзисторах: основные логические операции и их схемные реализации, принципы функционирования, преимущества и недостатки, характеристики.
7. Элементы полупроводниковых запоминающих устройств: основные схемные реализации, области применения, достоинства, недостатки, характеристики.

Глава 10

**НАНОЭЛЕКТРОНИКА
И ФУНКЦИОНАЛЬНАЯ ЭЛЕКТРОНИКА****10.1. Общие положения. Возможности
наноэлектроники и функциональной электроники**

Основные тенденции развития микроэлектроники с момента ее зарождения и до настоящего времени связаны с увеличением степени интеграции и быстродействия ИС. Однако такие тенденции имеют определенные пределы, обусловленные возможностями технологии, а также ограничениями, связанными с физическими и электрическими принципами преобразования информации и энергии, заложенными в работу современных ИС. Для преодоления этих затруднений в технологии начинают широко применяться рентгеновская и лазерная литографии, которые позволяют получить разрешение при создании элементов с размерами менее 10 нм ($1 \text{ нм} = 10^{-9} \text{ м} = 10^{-3} \text{ мкм}$). В США прогнозируется, что в ближайшие десять лет будут созданы МДМ-транзисторы с шириной затвора в 20 нм. Разработанные технологии разрешают разместить на кристалле 10^9 элементов. Уменьшение элементов до размеров порядка нескольких десятков, пусть даже сотен, периодов кристаллической решетки кристалла ИС существенно изменяет физические процессы (основы) работы приборов в ИС. При таких характерных размерах начинают проявляться и преобладать волновые свойства электронов, перемещающихся через структурированные барьеры, существующие в кристаллической решетке. Размеры этих барьеров оказываются одного порядка с длиной волны де Бройля для электрона $\lambda_d = h/mv$, где h — постоянная Планка, m — масса электрона, v — скорость его движения.

Из-за проявления волновых свойств в наноразмерных элементах основными физическими процессами, которыми обусловлен перенос носителей, являются **интерференция** и **дифракция электронных волн**, квантовые энергетические ограничения при движении носителей заряда, связанные с малыми размерами объекта, и **туннелирование** электронов через пространственно узкие потенциальные барьеры. Все перечисленные эффекты и со-

ставляют базу **нанозлектроники**, представляющей очередной этап и направление развития микроэлектроники.

Функциональная электроника, функциональная микроэлектроника, молекулярная электроника — различные названия еще одного перспективного направления микроэлектроники, которое связано с получением непрерывных комбинированных сред с заданными свойствами для создания различных электронных приборов на основе использования физических принципов и явлений, позволяющих получить компоненты со сложным функциональным схемотехническим назначением.

В отличие от ИС, где используется технологическая интеграция, в функциональной электронике не существует простых элементов типа диодов, транзисторов, резисторов и т. д.

Интегральные схемы в ближайшие годы достигнут своих предельных (критических) параметров, обусловленных физическими ограничениями и возможностями технологии. Появятся ИС с топологическими нормами порядка 0,1 мкм; скорость внешнего обмена информацией ограничится 3 ГГц из-за проблемы межсоединений, хотя частотный диапазон транзисторов будет существенно выше; плотность упаковки будет на уровне 10^{10} эл/см² (элементов/см²), а для микропроцессоров — $2 \cdot 10^8$ эл/см². Такие параметры ИС уже сейчас недостаточны для многих систем оперативного распознавания образов, искусственного интеллекта, для разработки устройств параллельной обработки информации и т. д. Разработчики ИС активно ищут способы преодоления технологических и физических барьеров. Одно из направлений базируется на разработке трехмерных ИС. На этом пути необходимо преодолеть следующие трудности: решить проблему взаимных помех элементов, разработать методы проектирования схем со сложной трехмерной топологией, обеспечить низкую цену, сравнимую с планарными ИС. Однако и в **трехмерной электронике** проблема межсоединений тоже не может быть решена полностью, а порой даже усложняется.

Любая микросхема — это совокупность очень большого числа искусственно созданных за счет технологических процессов локализованных статических неоднородностей в кристалле полупроводника (области эмиттера, коллектора, базы транзистора, резистор, изолирующие области и т. д.), которые соединяют между собой в соответствии со схемотехническими решениями, позволяющими обрабатывать, хранить и генерировать информацию. В семидесятых годах прошлого столетия возникла идея

использовать в микроэлектронике динамические неоднородности и физические принципы интеграции не только числа элементов, но числа функций, выполняемых электронным прибором, что и является задачей функциональной электроники.

К динамическим неоднородностям относятся ансамбли заряженных частиц, домены (электрические и магнитные), волновые неоднородности (поверхностные акустические волны, магнитные статические волны) и ряд других.

В зависимости от типа используемой среды, той или иной комбинации физических полей или явлений различают такие направления в функциональной электронике, как *функциональная акустоэлектроника, функциональная магнитоэлектроника, функциональная оптоэлектроника, функциональная диэлектрическая электроника, молекулярная электроника (молетроника)* и т. д.

В кремниевых, арсенид-галлиевых ИС и других устройствах, построенных на основе схмотехнических принципов, информационные сигналы передаются частями по линиям межсоединений, что ухудшает помехоустойчивость и надежность изделий. В устройствах функциональной электроники полезные сигналы обрабатываются не по частям (не по байтам), а одновременно по всей структуре. При этом существует возможность создания устройств, в которых можно обрабатывать одновременно как аналоговые, так и цифровые сигналы, что позволяет достигнуть скорости передачи и обработки информации более 10^{15} операций в секунду.

Межсоединения в современных ИС занимают до 60% площади кристалла, что приводит к наличию больших значений паразитных параметров (емкостей, волновых сопротивлений, дефектов и т. д.), которые, естественно, ухудшают характеристики ИС, особенно частотные. В функциональной электронике межсоединения играют незначительную, вспомогательную роль.

Для всех устройств функциональной электроники можно выделить ряд характерных элементов и явлений, к которым можно отнести:

- 1) динамические неоднородности определенного типа (поверхностные акустические волны (ПАВ), зарядовые пакеты в приборах с зарядовой связью (ПЗС), магнитостатические волны (МСВ));
- 2) континуальные среды, в качестве которых наиболее часто выступают кристаллические твердые тела, но перспективны также различного рода полимеры и биологические среды;

- 3) генераторы динамических неоднородностей, обеспечивающие их ввод в информационный канал;
- 4) устройства управления динамическими неоднородностями;
- 5) детекторы вывода и считывания информации, преобразующие заложенную в динамических неоднородностях информацию в двоичный код, что позволяет использовать хорошо развитые методы ее цифровой обработки.

10.2. Особенности нанозлектронных приборов

Квантоворазмерные структуры. Коротко остановимся на описании основных квантовых явлений в наноразмерных структурах, относящихся к объектам, физические свойства которых определяются *квантоворазмерными эффектами*, поскольку в рассматриваемом случае роль характерного физического параметра играет дебройлевская длина волны λ_d . Если геометрический размер структуры соизмерим или меньше λ_d , то проявляются *эффекты размерного квантования*, в частности возникают дискретные энергетические спектры электронов, причем положение каждого из квантовых уровней зависит от свойств материала и его геометрических размеров [43].

Если движение электрона в твердом теле ограничено по одной координате (размерность $2D$ — свободное движение по двум координатам), то такие структуры называются *квантовыми ямами*, а электронный газ в них можно считать двумерным. При ограничении движения по двум координатам (размерность $1D$ — одна координата) структуры называются *квантовыми нитями*. При размерности $0D$ (движение электрона ограничено по всем трем координатам) структуры именуется *квантовыми точками*.

Для полупроводниковых материалов заметный интерес представляют не только одиночные структуры типа квантовых ям, но и их последовательность, для которой электронные волновые функции электрона не проникают из одной ямы в другую. Характеристики такой структуры, состоящей из многих ям, имеют ярко выраженный резонансный характер, что важно для увеличения поглощения квантовыми структурами или числа активных центров.

При перекрытии квантовых ям волновыми функциями электронов электроны имеют возможность туннелировать из одной ямы в другую и структура будет вести себя как единое целое, такая структура называется *сверхрешеткой*, представляющей систе-

му взаимодействующих квантовых ям. Свойствами одномерной сверхрешетки, состоящей из квантовых ям шириной a , разделенных потенциальным барьером шириной b и высотой U , можно управлять варьированием параметров ямы и периода сверхрешетки $A = a + b$. Поведение электрона в такой сверхрешетке подобно перемещению его в одномерном кристалле с периодом решетки A .

Полупроводники наиболее подходят для наблюдения квантоворазмерных эффектов, поскольку длины дебройлевских волн для электронов невырожденного полупроводника при комнатной температуре имеют порядок в десятки нанометров, что достаточно для наблюдения интерференционных эффектов. Кроме того, технология полупроводниковых материалов и структур позволяет выращивать структуры с точностью до одного молекулярного слоя, повторяя наращивание много раз. Указанные свойства полупроводников и структур на их основе делают их основными объектами для использования в наноэлектронике.

В качестве примера структур, где формируются квантовые ямы и барьеры, через которые перемещается двумерный электронный газ, могут выступать пленочные структуры, МДП-структуры и гетероструктуры. В МДП-структуре потенциальный барьер формируется с одной стороны границей диэлектрика, где образуется контактная разность потенциалов с полупроводником, с другой стороны роль второй граничной стенки выполняет электростатический потенциал на границе инверсного и обедненного слоев полупроводника (см. п. 6.2). Изменением напряжения на затворе можно регулировать параметры потенциальной ямы, т. е. уровни энергии размерного квантования.

Потенциальная яма формируется на границе гетероперехода (см. п. 2.7) за счет разрывов зоны проводимости для электронов и валентной зоны для дырок. Реализовать квантовую яму, близкую к прямоугольной, наиболее просто с помощью двойной гетероструктуры (см. п. 21.2). Тонкий слой узкозонного полупроводника, располагаемого между двумя широкозонными, представляет собой прямоугольную квантовую яму, в которой движение носителей в плоскости (например, XU) происходит без ограничений, а в направлении Z , перпендикулярном гетерограницам, это движение ограничено, вследствие чего и возникает дискретный энергетический спектр.

В структурах, где движение электрона ограничено по двум координатам (квантовые нити), формируются двумерные по-

тенциальные ямы, что приводит к соответствующему квантованию энергии электрона.

При ограничении движения электрона по трем направлениям (координатам) свободное движение электрона невозможно и структура называется, как отмечалось выше, квантовой точкой с размерностью 0. Электроны заперты в трехмерной потенциальной яме и в соответствии с этим условием происходит квантование их энергии. Дискретный спектр разрешенных значений энергии такой квантоворазмерной структуры подобен спектру атомов и молекул, т. е. квантовые точки представляют собой искусственные атомы, в которых, в отличие от естественных, можно изменять и формировать энергетический спектр с помощью изменения параметров трехмерной потенциальной ямы. Квантовые точки можно сформировать за счет самоорганизации при эпитаксиальном выращивании структур с различными параметрами решетки. При монослойном эпитаксиальном наращивании арсенида индия (InAs) из твердого раствора ($\text{In}_{1-x}\text{Ga}_x\text{As}$) при определенных технологических условиях возможно образование как отдельных неупорядоченных квантовых точек, так и их определенное упорядочивание. Поперечные размеры таких квантовых точек составляют единицы и десятки нанометров. Граница между указанными материалами должна быть резкой и высокого качества, чтобы происходили интерференционные эффекты, определяющие квантование энергетических уровней.

Таким образом, в квантовых ямах происходит ограничение движения электронов по одной координате за счет формирования по этому направлению потенциального барьера; в квантовых нитях потенциальные барьеры существуют по двум координатам, поэтому свобода движения остается только по одной координате и, наконец, в квантовых точках наличие трехмерных потенциальных барьеров ограничивает движение электронов по всем трем координатам. Движение электронов в области потенциальных барьеров (и ям), сопоставимых по размерам с длиной волны де Бройля, приводит к квантованию энергии электронов, т. е. квантоворазмерные структуры формируют уровни, минизоны и подзоны размерного квантования, что в конечном счете используется для увеличения быстродействия нанoeлектронных приборов. Рассмотренные квантоворазмерные структуры в нанoeлектронных приборах создаются различными методами, некоторые из которых будут рассмотрены на примере конкретных устройств.

Типы и характеристики нанозлектронных приборов. Полевые транзисторы. Новые эффекты, реализуемые в нанозлектронных приборах, позволяют создавать приборы нового поколения. В 2004 г. фирма *Intel* выпустила на рынок МДП-нотранзистор первого поколения с размером затвора в 50 нм, к 2015 г. планируется создать транзистор с шириной затвора до 20 нм, что позволит достигнуть плотности упаковки элементов в ИС до 10^9 транзисторов/см² при быстродействии 30...40 ГГц (рис. 10.1). Уменьшение элементов транзистора (длины канала, толщины канала и подложки) приводит к квантованию энергии электрона в канале, что увеличивает скорость переноса и величину туннельного тока. Помимо этого, подвижность носителей в канале и, следовательно, быстродействие могут быть увеличены за счет небольшого принудительного растяжения кристаллической решетки кремния (примерно на 1%). Для создания в кремнии механического напряжения его связывают с кремний-германиевым сплавом, имеющим несколько больший шаг кристаллической решетки. Чтобы получить бездефектные подложки и слои, используют новые подходы в КНД-технологии (технологии «кремний-

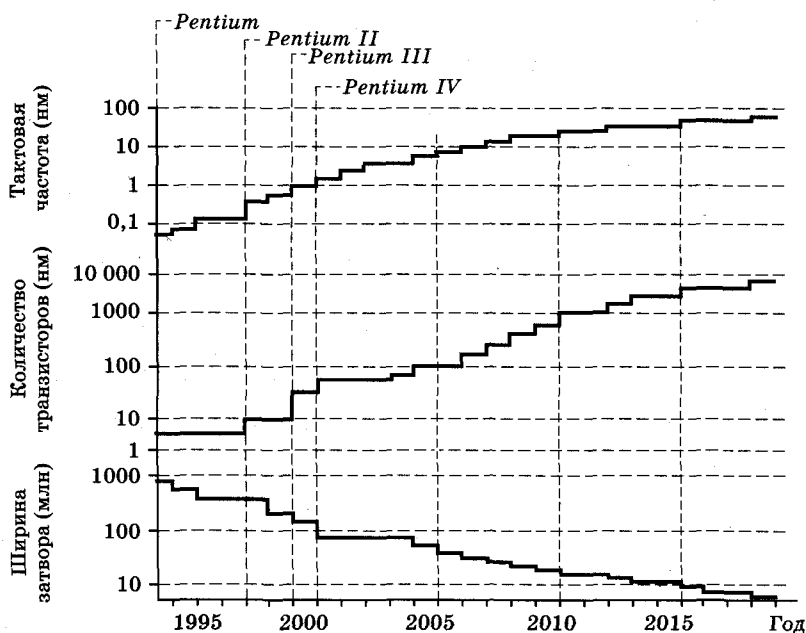


Рис. 10.1

на-диэлектрике») (см. гл. 7). Одним из перспективных способов считается следующий: сначала формируется изолирующий слой оксида непосредственно на поверхности кремниевой подложки, затем пластину переворачивают окисленной поверхностью вниз и накладывают на необработанную подложку, после чего большую часть оказавшегося сверху кремния аккуратно удаляют. В результате остается тонкий слой кремния над изолирующей прослойкой оксида, лежащего на массивной подложке, обеспечивающей механическую прочность. Для такого формирования слоев используют метод *Smart-Cut*, при котором окисленную поверхность первой подложки бомбардируют ионами водорода. Поскольку ионы водорода повреждают материал (создают дефекты) в основном в момент остановки, то внутри кремния образуется очень хрупкий слой. После соединения обработанной подложки с массивной кремниевой пластинкой ее верхняя часть легко откалывается. Остаточную шероховатость на оставшемся кремнии полируют, в результате формируется более тонкий слой кремния на оксиде.

В многослойных структурах механизм туннелирования носит резонансный характер, при котором туннельный переход через несколько (систему) потенциальных барьеров возможен только при определенной энергии электронов. Такая многослойная структура формирует потенциальные барьеры, подобные сверхрешетке. Примером такой многослойной структуры может служить структура транзистора с резонансным туннелированием (рис. 10.2, а). ВАХ рассматриваемых многослойных структур имеют резонансный характер с наличием максимумов и минимумов. Ток в максимумах ВАХ соответствует напряже-

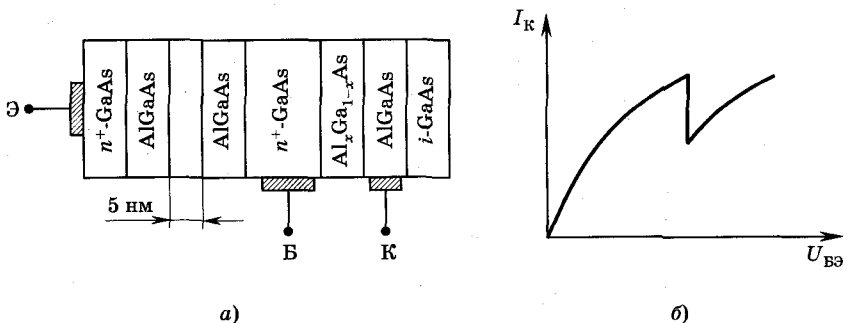


Рис. 10.2

ниям, при которых резонансные энергии кратны энергии Ферми E_F . Поскольку функция распределения электронов по энергии не является дельта-функцией, т. е. электроны движутся с разными скоростями, и длины их свободного пробега различны, то электроны под воздействием приложенного напряжения могут туннелировать через несколько слоев до момента столкновения. В результате многократного некогерентного туннелирования происходит расширение пиков ВАХ и увеличение напряжений, соответствующих пикам. ВАХ транзистора, изображенного на рис. 10.2, а, приведена на рис. 10.2, б. Эмиттер этого транзистора имеет сложную структуру с двумя потенциальными барьерами, между которыми размещается квантовая яма с шириной в несколько нм ($\sim 3 \dots 6$ нм). Такая структура обеспечивает резонансное туннелирование, что приводит к крайне малому разбросу энергии электронов, т. е. за счет наличия потенциальных барьеров большое число электронов эмиттерной области задерживаются в ней и в базу поступают электроны примерно с одинаковыми скоростями (моноскоростной поток). Из-за наличия рассмотренных процессов ВАХ такого транзистора имеет па-

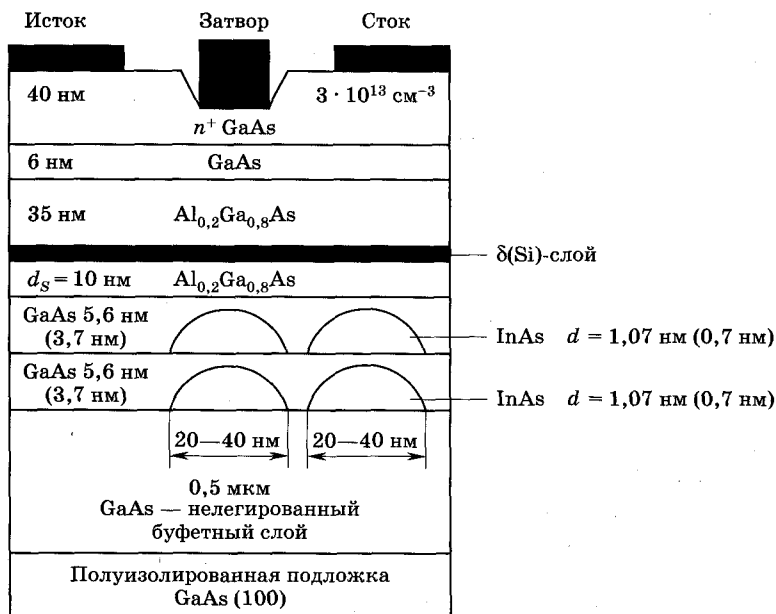


Рис. 10.3

дающий участок — участок с отрицательной дифференциальной проводимостью. Особенностью этого транзистора является то, что потенциальные барьеры в нем создаются за счет гетеропереходов на границах контактов арсенид-галлиевых слоев и слоев на основе тройных соединений AlGaAs. Эти переходы позволяют создавать узкие потенциальные барьеры, облегчающие туннелирование электронов.

Достаточно перспективным ПТ для работы в ВЧ и СВЧ схемах является гетеротранзистор, структура которого изображена на рис. 10.3. В этом типе транзисторов создаются не только квантовые ямы, но и квантовые точки. Последние вкрапляются в область протекания тока, и их поверхностная концентрация $\sim 3 \cdot 10^{10} \text{ см}^{-2}$. Стоковый ток определяется состояниями электронов, движущихся через квантовые ямы, и электронами, локализованными в квантовых точках, свобода которых ограничена трехмерными потенциальными барьерами. Влияние и роль квантовых точек проявляется в характере выходной ВАХ рассматриваемого транзистора, которая имеет двухступенчатую форму (см. рис. 10.2, б).

Разновидности нанозлектронных приборов. В настоящее время разработаны и разрабатываются методы создания различного типа нанозлектронных приборов на квантоворазмерных эффектах, работающих в различных областях длин волн, включая весь оптический диапазон, например инжекционные лазеры на гетероструктурах (см. п. 21.3). Одним из перспективных направлений нанозлектроники является создание одноэлектронных приборов, например *одноэлектронных транзисторов*. Если создать структуру из нескольких областей, формирующих туннельные переходы с малой собственной емкостью, то возможно перемещение через эти переходы одного или отдельных электронов. Нужные переходы можно сформировать между затвором, стоком и истоком. Из существующих нанозлектронных приборов можно отметить *квантовые интерферометры, квантовые каскадные лазеры, квантовые вентили (инверторы)*.

10.3. Приборы с зарядовой связью

Приборы с зарядовой связью (ПЗС) используются в самых разнообразных устройствах: в запоминающих устройствах (ЗУ) с большой емкостью памяти, в управляемых линиях задержки,

в оптоэлектронных приемниках изображений, в формирователях видеосигналов, в различных фильтрах, в спектроанализаторах и т. д. Широкое распространение ПЗС в телевидении, радиолокации, системах связи и обработки информации обусловлено простотой их конструктивных решений и технологий изготовления. Эти приборы имеют высокую степень интеграции, хорошую надежность и чувствительность, могут работать в инфракрасном, видимом и ультрафиолетовом диапазонах длин волн.

ПЗС представляют собой ансамбли МДП-конденсаторов, расположенных на минимально возможном расстоянии друг от друга. Названия электродов в ПЗС аналогичны МДП-транзисторам, верхние металлические электроды называются затворами, ввод информации обычно происходит через электрод истока, а вывод через сток. МДП-конденсаторы сформированы на общей монокристаллической полупроводниковой подложке. Расположенные на слое диэлектрика полоски затворов (1) (рис. 10.4, а) образуют регулярную линейную систему или плоскую матрицу. В большинстве ПЗС подложка (2) изготавливается из высокоомного кремния, затворы — из алюминия или поликристаллического кремния, в качестве диэлектрика (3) используется SiO_2 , толщина слоя которого 0,1 мкм.

Типичные размеры электродов: длина $L = 5$ мкм, ширина 40 мкм, зазор между электродами $l \leq \dots 2$ мкм, число электродов $\sim 10^3$ в линейном и $10^4 \dots 10^6$ в матричном ПЗС.

Затворы присоединяют к управляющим шинам, на которые относительно подложки подаются импульсные управляющие потенциалы. В зависимости от назначения ПЗС используются различные схемы организации тактового питания и взаимного расположения затворов. Далее для простоты будет рассматриваться трехтактное управление. При использовании подложки p -типа (см. рис. 10.4) управляющие напряжения имеют положительную полярность.

На рис. 10.4, б из всей совокупности показана только одна триада МДП-конденсаторов, где каждый из указанных затворов подсоединен к своей шине управления (рис. 10.5). Принцип действия ПЗС основан на накоплении и хранении зарядовых пакетов $Q_{\text{н}}$ в потенциальных ямах под затворами и на их перемещении в подложке между соседними конденсаторами при изменении импульсных напряжений на управляющих шинах. Таким образом, между соседними элементами реализуется зарядовая связь, откуда возникло и само название ПЗС.

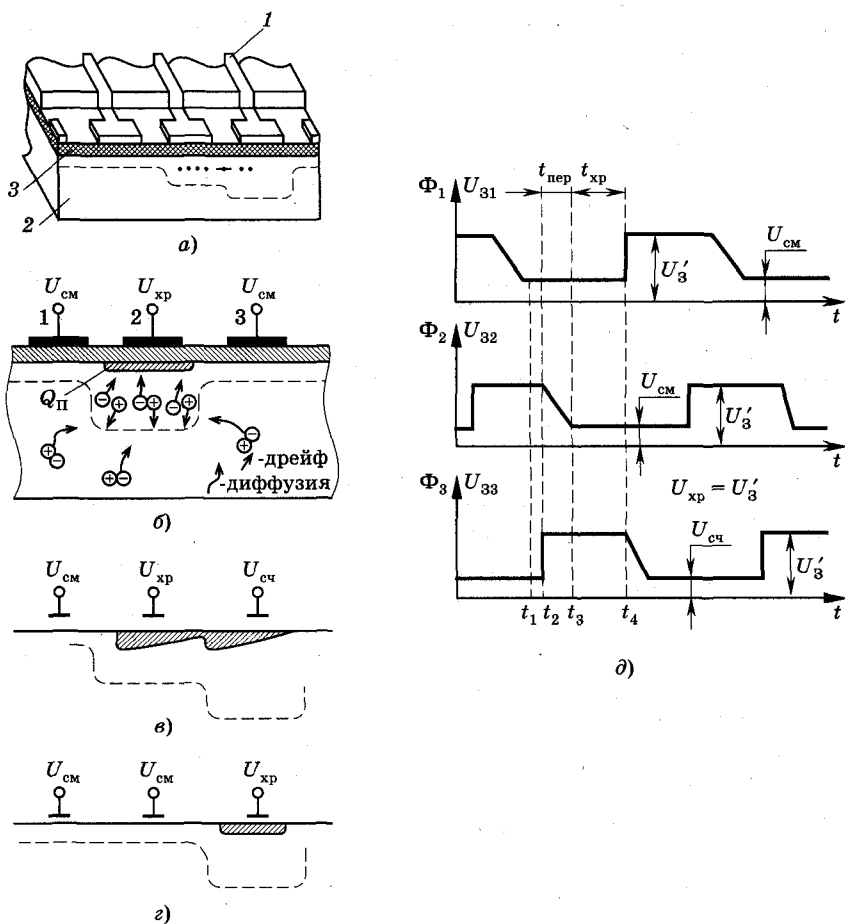


Рис. 10.4

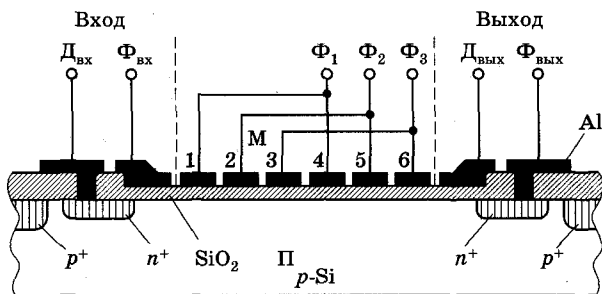


Рис. 10.5

В ПЗС затворы на основе алюминиевых или поликремниевых пленочных полупроводников присоединяют, как уже отмечено, к управляющим шинам, на которые относительно заземленного электрода подложки подают импульсные управляющие напряжения. В рассматриваемом приборе имеется три управляющих шины Φ_1 , Φ_2 , Φ_3 (см. рис. 10.5), поэтому он называется трехтактным.

Рассмотрим процессы записи, хранения и переноса информации. При записи информации напряжение в течение импульса, например, на втором затворе (см. рис. 10.4, б, в момент времени $t = t_1$), должно быть больше порогового (см. п. 6.2) $\Phi_2 = U = U_{\text{хр}} > U_{\text{пор}}$, а на соседних затворах напряжения должны удовлетворять соотношению Φ_1 , $\Phi_3 = U_{31}$, $U_{33} = U_{\text{см}} < U_{\text{пор}}$, где $U_{\text{хр}}$ — напряжение хранения, $U_{\text{см}}$ — напряжение смещения. В результате в полупроводниковой подложке под затвором образуется потенциальная яма (штриховая линия на рис. 10.4, б) для электронов. В течение времени порядка времени диэлектрической релаксации из-под затвора 2 под действием поля удаляются дырки, а электроны еще не успевают заполнить потенциальную яму, т. е. под затвором 2 оказывается сформированным обедненный слой. Глубина потенциальной ямы максимальна на границе полупроводника с диэлектриком, поэтому в первую очередь здесь начинают накапливаться электроны за счет таких процессов, как перемещение носителей из соседней МДП-структуры, неконтролируемая тепловая генерация в обедненном слое и на поверхности полупроводника, а также из-за диффузии электронов из подложки. Зависимость поверхностного потенциала от величины зарядового пакета $Q_{\text{п}}$ при заданном напряжении затвора приблизительно линейная:

$$\Phi_{\text{нов}} = -Q_{\text{п}}/C_{\text{д}}, \quad (10.1)$$

где $C_{\text{д}}$ — подзатворная емкость, определяемая диэлектрическим слоем.

При постоянном значении $Q_{\text{п}}$ поверхностный потенциал возрастает также примерно линейно при увеличении напряжения затвора.

Накопленный заряд (см. рис. 10.4, б, заштрихованная область $Q_{\text{п}}$ под затвором 2) будет храниться под этим затвором до тех пор, пока $U_{\text{хр}} = U'_3 = U_{32} > U_{31}$, U_{33} , при этом U_{31} , $U_{33} = U_{\text{см}} < U_{\text{пор}}$. (Под первым и третьим затворами заряда нет.)

Рассмотрим теперь процесс переноса зарядового пакета (считывание информации). Пусть в момент времени t_2 (см. рис. 10.4, ∂) напряжение на втором затворе начинает уменьшаться $\Phi_2 = U_{32} = U_{см} < U_{пор}$, а на третьем затворе U_{33} скачком нарастает до значений напряжений считывания $U_{сч} = U_{32} > U_{пор}$, при этом $U_{31} = U_{см} < U_{пор}$ и $U_{сч} = U'_3$. Под третьим затвором практически мгновенно формируется пустая потенциальная яма.

Отметим, что с учетом конечности времени, которое требуется для перетекания заряда из-под одного затвора под другой, задний фронт тактовых импульсов специально делают достаточно пологим.

Перенос зарядовых пакетов осуществляется благодаря наличию краевых эффектов, поскольку размеры потенциальной ямы в направлении переноса зарядов превышают размеры затвора (см. рис. 10.5).

Поскольку при $t = t_2$ в третьем элементе электронов нет, а во втором накоплен заряд $Q_{п2}$, то при $U_{32} = U_{33} = U'_3$ (см. рис. 10.4, ∂) согласно выражению (10.1) поверхностный потенциал под затвором 3 будет значительно выше, чем под затвором 2. В результате возникает продольное электрическое поле, ускоряющее электроны в сторону третьего элемента. На рис. 10.4, $a, в$ показана ситуация, иллюстрирующая изменение величины заряда под вторым и третьим затворами в процессе перетекания зарядов (момент времени между t_2 и t_3). При $t = t_3$ перенос зарядового пакета из второго элемента в третий заканчивается (см. рис. 10.4, $з$).

Для хранения и переноса одного зарядового пакета в данном случае необходимо наличие трех элементов (трех МДП-конденсаторов). В момент $t = t_4$ на шину Φ_1 (см. рис. 10.5), т. е. на затворы 1 и 4, подается напряжение высокого уровня $\Phi_1 = \Phi_4 = U_3 > U_{пор}$, и пакет заряда из третьей МДП-ячейки перемещается в четвертую, а первая заполняется зарядом из предыдущей и т. д.

Таким образом, интервал времени от t_2 до t_3 соответствует времени переноса заряда $t_{пер}$, т. е. $t_{пер} = t_3 - t_2$, а интервал $t_4 - t_3 = t_{хр}$ — это время хранения заряда. Для преобразования входных сигналов (напряжений) в сигнальные зарядовые пакеты и, наоборот, выходных зарядовых пакетов в напряжения используются устройства ввода и вывода информации.

Устройство ввода на рис. 10.5 состоит из области n^+ -типа, имеющей омический контакт с электродом $D_{вх}$, и входного за-

твора $\Phi_{\text{вх}}$. Область n^+ -типа образует вместе с подложкой входной диод на основе перехода $n^+ - p$. При подаче на вход (на $D_{\text{вх}}$) сигнала отрицательной полярности, а на $\Phi_{\text{вх}}$ — положительно-управляющего напряжения входной диод $n^+ - p$ смещается в прямом направлении. В результате из n^+ -области под входной затвор инжектируется зарядовый пакет, который затем переносится под первый затвор, где $\Phi_1 = U_{31}$. Достоинством описанного способа ввода электрического сигнала является высокое быстродействие (несколько наносекунд).

Выходное устройство (см. рис. 10.5) содержит область n^+ -типа, омический контакт $D_{\text{вых}}$ с этой областью и выходной затвор $\Phi_{\text{вых}}$; n^+ -область и подложка образуют выходной диод, который смещают в обратном направлении, при этом обратное напряжение на этом диоде превышает максимальное напряжение на выходном электроде $\Phi_{\text{вых}}$. Если на выходной затвор подается импульс положительной полярности, а в последнем элементе (6), где $U_{36} = \Phi_3$, к этому моменту времени накоплен зарядовый пакет, то он сначала переместится в потенциальную яму под выходным затвором ($D_{\text{вых}}$), а затем в более глубокую потенциальную яму области n^+ -типа и после этого в выходную цепь. Обычно к выходу подключают чувствительный усилитель на МДП-транзисторах, который создается на этой же подложке. При неразрушаемом считывании информации на основе МДП-транзисторов используются датчики поверхностного потенциала, величины которых определяются значением зарядового пакета.

Основными параметрами ПЗС являются: рабочая амплитуда управляющих напряжений, минимальная величина зарядового пакета, максимальная и минимальная тактовые частоты, эффективность переноса зарядового пакета, рассеиваемая мощность.

Рабочая амплитуда управляющих напряжений на затворах должна обеспечивать требуемую величину зарядового пакета и полное смыкание обедненных слоев соседних элементов (см. рис. 10.4), чтобы под их затворами образовывалась общая потенциальная яма при переносе зарядового пакета. Типичные значения амплитуды управляющих напряжений 10...20 В.

Максимальная величина зарядового пакета $Q_{\text{п макс}}$ пропорциональна амплитуде управляющего напряжения и площади затвора. Например, для затвора с размерами 10×20 мкм, толщиной диэлектрика $d = 0,1$ мкм и $\phi_{\text{пов}} = 5$ В получим $Q_{\text{п макс}} = 0,35$ пКл.

Минимальная тактовая частота $f_{T \text{ мин}}$ обратно пропорциональна максимально допустимому времени хранения зарядового пакета в одном элементе. Время хранения ограничено из-за неконтролируемого накопления электронов в потенциальных ямах под затвором вследствие тепловой генерации носителей заряда в обедненном слое и на границе полупроводника с диэлектриком, а также диффузии из подложки. Для уменьшения $f_{T \text{ мин}}$ (увеличения времени хранения) уменьшают концентрацию объемных центров рекомбинации и рабочую температуру. Типичные значения $f_{T \text{ мин}} = 30 \dots 300$ Гц.

Максимальная тактовая частота $f_{T \text{ макс}}$ обратно пропорциональна допустимому времени переноса, за вычетом времени хранения.

Эффективность переноса $\eta = Q_{\text{п}, i+1} / Q_{\text{п}, i} < 1$, где $Q_{\text{п}, i}$ и $Q_{\text{п}, i+1}$ — зарядовые пакеты в i -м элементе до переноса и в следующем $(i+1)$ -м после переноса соответственно. Для сложных устройств на ПЗС $\eta = 0,999 \dots 0,99999$, т. е. $1 - \eta = 10^{-3} \dots 10^{-5}$.

Потери заряда возрастают на больших тактовых частотах f_T из-за уменьшения времени переноса зарядовых пакетов, поскольку чем больше требуемая эффективность переноса, тем большее время переноса необходимо и тем ниже $f_{T \text{ макс}}$. На низких частотах потери заряда вызваны захватом части электронов поверхностными ловушками. Если зарядовый пакет переносится через незаполненные до этого элементы, то потери — возрастают, так как в них число незаполненных поверхностных ловушек оказывается больше, чем в заполненных перед этим элементах. Для уменьшения потерь заряда на ловушках создают фоновый заряд, который вводится во все элементы за счет подачи некоторого положительного минимального значения напряжения $U_{3 \text{ мин}} = U_{\text{см}}$ ($U_{\text{см}}$ — напряжение смещения, см. рис. 10.4, б, в, г, д). При этом в элементах постоянно будет существовать фоновый заряд, заполняющий ловушки, что и уменьшает потери заряда при переносе.

Рассеиваемая мощность элементов ПЗС ($P_{\text{рас}}$) существует практически только в режиме переноса зарядового пакета и увеличивается пропорционально тактовой частоте f_T , амплитуде управляющего напряжения и величине зарядового пакета. Для элементов ПЗС $P_{\text{рас}} \leq 1$ мкВт. Малая рассеиваемая мощность — одно из основных достоинств ПЗС.

Помимо рассмотренных трехтактных ПЗС с поверхностным переносом заряда, в настоящее время существуют более совер-

шенные ПЗС, к которым относятся ПЗС со скрытым каналом переноса зарядовых пакетов (рис. 10.6), двухтактным управлением, на МНОП-структурах и структурах с плавающим затвором. Эти приборы имеют более простую технологию изготовления, меньшее расстояние между затворами, в них устранено пересечение линий металлизации для соединения затворов и т. д. Время хранения записанной информации в этих структурах доходит до нескольких десятков тысяч часов. Важным достоинством ПЗС с объемным (скрытым) каналом является низкий уровень шумов за счет устранения взаимодействия зарядовых пакетов с поверхностными состояниями.

Помимо этого, в объемных ПЗС больше подвижность электронов, что увеличивает эффективность переноса и тактовую частоту прибора. В этом приборе движение зарядовых пакетов ограничено в пределах объемного (скрытого) канала, расположенного под границей с окислом (см. рис. 10.6) и формируемого внутри тонкого (толщина около 4 мкм) приповерхностного слоя n -типа, расположенного на кремниевой подложке p -типа. Концентрация доноров в n -слое $N_d \approx 2 \cdot 10^{15} \text{ см}^{-3}$, в то время как в подложке концентрация акцепторов $N_a \approx 2 \cdot 10^{14} \text{ см}^{-3}$. Входная и выходная n^+ -области размещаются на границах n -слоя. Если подложка и затворы (металлические электроды) заземлены, входная цепь разомкнута, а к выходной области подключен источник питания (например, +20 В), смещающий p — n -переход на границе с подложкой в обратном направлении, то под затворами образуются не только приповерхностные обедненные области, но и обедненная область p — n -перехода. Обе обедненные области при относительно больших напряжениях на n -слое перекрываются (смыкаются) в вертикальном направлении. Распределение потенциала,

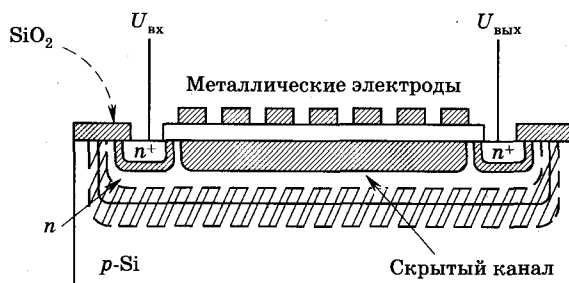


Рис. 10.6

соответствующее минимуму энергии электронов, имеет максимум на некотором расстоянии от поверхности (~ 3 мкм) внутри n -слоя. Сигнальные электроны, введенные в обедненную область, будут смещаться полем в область минимума потенциальной энергии на глубину ~ 3 мкм, т. е. аналогично ПЗС с поверхностным зарядом в этой потенциальной яме будет накапливаться заряд. Глубину потенциальной ямы можно менять напряжением на затворе, а перемещать заряды — за счет изменения напряжения на затворах, так же как в ПЗС с поверхностным каналом. Недостатком ПЗС со скрытым каналом является значительно меньшая величина максимального зарядового пакета, что обусловлено увеличением расстояния от затвора до места накопления заряда.

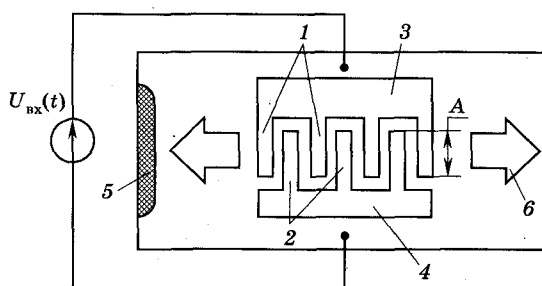
10.4. Элементы акустоэлектроники

Акустоэлектрические микросхемы используют электрические и высокочастотные акустические сигналы. Они изготавливаются технологическими методами микроэлектроники, но не содержат традиционных элементов ИС: транзисторов, диодов, резисторов и т. д. Использование в акустоэлектронных приборах электрических и акустических сигналов требует применения в этих приборах преобразователей электрических сигналов в акустические и акустических в электрические, а также наличия акустических трактов (звукопроводов). Для преобразования одного вида сигналов в другой применяют прямой и обратный пьезоэлектрические эффекты. Акустический сигнал представляет собой волну упругих механических возмущений, распространяющихся в твердом теле со скоростью звука порядка 10^5 м/с. Для преобразования и обработки сигналов наиболее важным является использование в акустоэлектронных приборах поверхностных акустических волн (ПАВ), распространяющихся в поверхностном слое пьезокристалла.

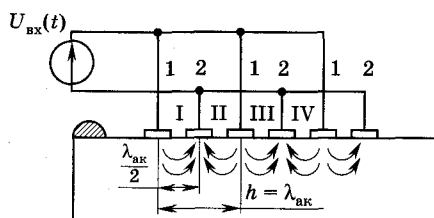
Толщина слоя, в котором распространяется (существует) ПАВ, имеет значение порядка длины акустической волны $\lambda_{\text{ак}} = v_{\text{ак}}/f$, где $v_{\text{ак}}$ и f — соответственно скорость распространения и частота акустической волны. Для основных материалов, применяемых в акустоэлектрических приборах, таких, как SiO_2 (пьезокварц), LiNbO_3 (ниобат лития), $\text{Bi}_{12}\text{GeO}_{20}$ (германат висмута), пьезополупроводники (GaAs и InSb), скорость акустической волны лежит в пределах $(1,6 \dots 4)10^5$ см/с. При частоте $f = 10^7 \dots 10^{10}$ Гц длина акустической волны составляет $\lambda_{\text{ак}} = 100 \dots 0,1$ мкм.

Наибольшее распространение в радиоэлектронике и других областях техники получили: из пассивных акустоэлектронных приборов — линии задержки и полосовые фильтры, а из активных — устройства, в которых используется также и усиление сигналов.

В линиях задержки акустоэлектронный прибор применяется как замедляющая система (см. п. 13.3), в которой происходит сильное уменьшение скорости распространения сигнала за счет его преобразования из электромагнитного (электрического) в акустический. Поскольку скорость акустической волны, в частности ПАВ, примерно на пять порядков меньше, чем у электромагнитной волны, это позволяет создавать малогабаритные интегральные линии задержки (ЛЗ) электромагнитных сигналов. Конструктивно ЛЗ состоит из входного преобразователя, звукопровода и выходного преобразователя. Наибольшее распространение нашли встречно-штыревые преобразователи, устройство которых иллюстрирует рис. 10.7, где 1 и 2 — штыри, объединенные соответственно шинами 3 и 4; 5 — поглощающее покрытие, 6 — направление распространения ПАВ. Расстояние между соседними штырями, отходящими от одной шины, назы-



а)



б)

Рис. 10.7

вается пространственным шагом h , который определяет частоту акустического синхронизма $f_0 = v_{\text{ак}}/h$, т. е. на частоте f_0 длина акустической волны $\lambda_{\text{ак}} = h$ (см. рис. 10.7, б). При этих соотношениях электрический сигнал преобразуется в ПАВ наиболее эффективно, механизм этого преобразования можно уяснить из рис. 10.7, б. Входной сигнал $U_{\text{вх}}(t)$ формирует электрические поля в зазорах между штырями у поверхности кристалла. Тангенциальные (продольные — в направлении z) составляющие векторов напряженности электрического поля в соседних зазорах имеют противоположные направления, что вызывает упругие возмущения в кристалле, которые формируют ПАВ. Возбужденная ПАВ распространяется вдоль преобразователя в обе стороны. Для получения однонаправленной волны одна из двух волн устраняется с помощью поглощающего покрытия 5 (см. рис. 10.7, б). Возможно создание и однонаправленных преобразователей, но они сложнее и занимают большую площадь на подложке. В ЛЗ на ПАВ задержка в 1 мкс соответствует длине кристалла всего в 1...2 мм. За счет использования сложной конфигурации акустических спиральных волноводов и отражателей возможно увеличение задержки вплоть до 1 мс. Полоса пропускания (Δf) акустических линий задержки может быть достаточно широкой, вплоть до $\Delta f = f_0$. Из-за неполного согласования входных и выходных преобразователей ПАВ с источником электрического сигнала и нагрузкой, а также из-за потерь части энергии ПАВ в звукопроводе суммарные потери в ЛЗ составляют несколько дБ при времени задержки $t_{\text{зд}} \sim 1$ мкс и 40...50 дБ при $t_{\text{зд}} \sim 1$ мс. При изготовлении преобразователей с минимальным шагом h рабочие частоты определяются возможностями литографии. Использование субмикронной литографии позволяет получить рабочие частоты ЛЗ вплоть до 10...15 ГГц. Элементы акустического тракта (переизлучатели, отражатели, ответвители, делители, волноводы) предназначены для изменения направления распространения ПАВ, разветвления акустического сигнала, уменьшения апертуры (локализация волны), связанной с длиной перекрытия штырей A (см. рис. 10.7, а), и т. д. Ряд из перечисленных операций можно выполнить с помощью многополосковых ответвителей, представляющих собой систему параллельных проводников, перпендикулярных направлению распространения ПАВ.

Простейшие акустоэлектрические полосовые фильтры содержат входной многоэлектродный преобразователь и выходной преобразователь с малым числом электродов, полоса пропускания которого значительно шире, чем входного. Амплитудно-частотная характеристика (АЧХ) такого фильтра определяется входным преобразователем. Форма АЧХ искажается из-за конечной длины входного преобразователя, пространственной расходимости ПАВ при ее распространении к выходному преобразователю, отражений и других факторов. Для получения оптимальной АЧХ применяют фильтры со значительно более сложной структурой. В фильтрах, использующих избирательные свойства преобразователей, не удается получить малую полосу пропускания. Узкополосные фильтры выполняются на основе резонаторов ПАВ. Резонаторы используют отрезки звукопроводов с отражателями с двух противоположных сторон. По длине резонатора должно укладываться целое число акустических полуволн. В такой резонатор помещают широкополосный преобразователь. В активных акустоэлектронных приборах, в частности усилителях, осуществляется взаимодействие свободных электронов пьезополупроводника (GaAs, InSb) с ультразвуковыми волнами с частотой $10^7 \dots 10^{13}$ Гц. Если под действием переменного электрического поля возбуждается акустическая волна, а постоянное электрическое поле создает дрейф электронов в направлении распространения волны, то при скорости электронов, большей скорости волны, возможно усиление, а при скорости электронов, меньшей скорости волны, происходит ослабление. Аналогичный механизм реализуется в электровакуумных приборах СВЧ с динамическим управлением (см. гл. 13). Подвижность электронов в GaAs и InSb много больше, чем в Ge и Si и других полупроводниках, поэтому в них проще реализовать указанный механизм.

10.5. Элементы СБИС на цилиндрических магнитных доменах

Магнитные свойства тонких эпитаксиальных пленок на основе гранатов, имеющих химическую формулу $R_3Fe_5O_{12}$, где R — редкоземельный элемент (Y, Cd и др.), используются для запоминания и обработки информации. В качестве носителей информации в таких пленках выступают цилиндрические магнитные

домены (ЦМД). Информационная емкость СБИС определяется диаметром ЦМД, который имеет размеры порядка микрона. Магнитные СБИС применяются как различного рода запоминающие устройства (ЗУ). При малой толщине пленки магнитные домены занимают все поперечное сечение пленки, а их векторы намагниченности перпендикулярны ее поверхности (рис. 10.8, а). В исходном состоянии домены имеют форму полос, при этом соседние полосы (заштрихованные и белые на рис. 10.8, а) имеют

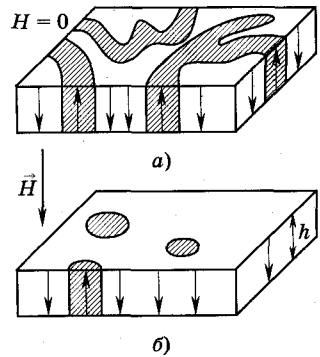


Рис. 10.8

противоположные направления вектора намагниченности (показаны стрелками), так что результирующая намагниченность пленки равна нулю. При приложении перпендикулярного по отношению к поверхности внешнего магнитного поля напряженностью \vec{H} (рис. 10.8, б) домены, у которых вектор их намагниченности совпадает с направлением внешнего поля, расширяются, а домены с противоположным направлением вектора намагниченности сужаются. При напряженности внешнего магнитного поля больше некоторой величины $H_{\text{мин}}$ полосовые домены превращаются в цилиндрические (см. рис. 10.8, б), диаметр которых уменьшается с ростом H и при некотором $H > H_{\text{макс}}$ ЦМД резко исчезают (коллапсируют), т. е. такие домены существуют при $H_{\text{мин}} < H < H_{\text{макс}}$.

Таким образом, для существования ЦМД необходимо иметь определенное поле смещения, нужная напряженность которого создается с помощью постоянного магнита и имеет значения 2000...3000 А/м.

Помимо зависимости от H , на диаметр ЦМД (d , см. рис. 10.9) сильно влияют материал и толщина пленки (h , см. рис. 10.8, б), обеспечивающая минимальный диаметр $d_{\text{мин}} \approx h \approx 0,3$ мкм. Между соседними доменами существуют магнитостатические силы отталкивания. В неоднородном магнитном поле ЦМД движется в сторону убывания H . Однако скорость такого движения невысока ($\sim 10^3$ см/с), что является одной из причин низкого быстродействия. При использовании ЦМД в элементах памяти необходима фиксация положения домена, что будет соответ-

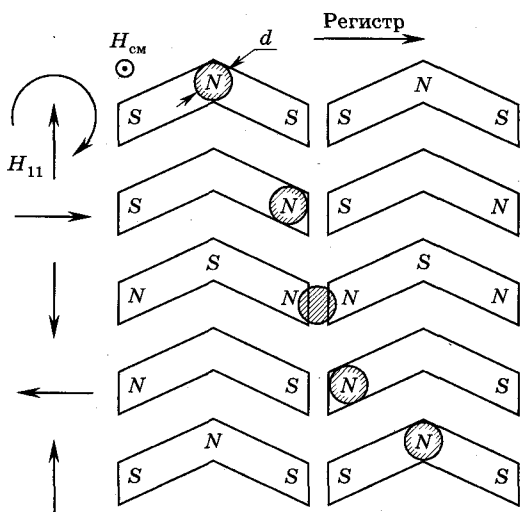


Рис. 10.9

воватъ хранению лог. 1, а отсутствие домена — хранению лог. 0. Фиксация ЦМД осуществляется с помощью магнитостатических ловушек, представляющих собой области с низким значением потенциальной энергии ЦМД (потенциальные ямы). Наиболее часто используются ловушки, создаваемые тонкими (десятые доли мкм) ферромагнитными пленками из пермаллоя (80% Ni, 20% Fe), наносимыми на поверхность, на которой формируется ЦМД. Пермаллоевая пленка ослабляет внешнее магнитное поле и «закрывает» магнитный поток, что притягивает к ней ЦМД. Возможно получение ловушек с помощью проводящей петли с током, создающим магнитное поле с направлением, противоположным вектору поля смещения.

Систему пермалловых ловушек, расположенных на расстоянии меньше диаметра домена, можно использовать в качестве регистра и перемещать ЦМД по определенному адресу. Для этого применяют управляющее магнитное поле, вектор напряженности которого направлен параллельно поверхности и вращается в плоскости пермалловых ловушек (аппликаций). Вращающееся поле можно создать с помощью двух взаимно перпендикулярных катушек, в которых токи сдвинуты по фазе на 90° . Рассмотрим процесс перемещения ЦМД на примере элемента регистра из двух шевроновых пермалловых аппликаций. На рис. 10.9 показано положение ЦМД (заштрихованный

кружок) в различные фазы вращающегося управляющего магнитного поля. Каждый ряд соответствует фазовому сдвигу $H_{\text{упр}} = H_{\parallel}$ на четверть периода вращения. Под влиянием $H_{\text{упр}}$ в пермалловых пленках образуются магнитные полюса, которые создают дополнительное магнитное поле с составляющей, направленной по или против вектора смещения ($H_{\text{см}}$). В нулевой фазе ($t = t_0 = 0$) ЦМД располагается в верхней части шеврона, где результирующее поле смещения минимально и где в этот момент формируется (располагается) северный полюс, который притягивает ЦМД. Через четверть периода ($t = T/4$) северный полюс и соответственно ЦМД перемещаются к правой стороне шеврона и т. д. В результате через период ($t = T$) ЦМД займет верхнее положение в соседней ловушке (последний ряд на рис. 10.9). Таким образом, время сдвига на один разряд регистра равно периоду вращающегося поля. Типичная частота вращения магнитного поля составляет величину порядка 100 кГц при напряженности — 2000...4000 А/м. Частота ограничена индуктивностью катушек и мощностью потерь. Скорость перемещения доменов достигает сотен метров в секунду, плотность записи информации может быть $10^4 \dots 10^5$ бит/мм, а скорость записи $10^5 \dots 10^6$ бит/с. Для считывания информации применяют петли, изготовленные из полупроводника, обладающего магниторезистивным эффектом. Эти петли наносят на основную пленку, и при пропускании через них тока магнитное поле в петле будет изменяться в момент прохождения под ней ЦМД. Это вызовет изменение сопротивления петли и соответственно тока в ней, что соответствует лог. 1.

Повышение информационной емкости устройств на ЦМД ограничивается магнитостатическим взаимодействием между ЦМД, которое требует необходимого минимального расстояния между ними. Для нормальной работы расстояние между ЦМД в соседних элементах памяти должно быть не менее $4d$, где d — диаметр ЦМД. В результате площадь элемента памяти получается большой, а информационная емкость устройства оказывается недостаточной для современных СВИС. Этот недостаток в значительной мере устранен при кодировании информации с помощью более мелких объектов, так называемых *пар вертикальных блоховских линий (ВБЛ)*, заключенных в доменной границе.

**Контрольные вопросы**

1. Каковы возможности и перспективы нано- и функциональной электроники?
2. Объяснить, что такое квантовые ямы, квантовые нити и квантовые точки.
3. В каких наноэлектронных приборах используются квантовые ямы, квантовые нити и квантовые точки?
4. Резонансное туннелирование и его использование в наноэлектронных приборах.
5. Каковы особенности наноэлектронных приборов?
6. Приборы с зарядовой связью: основные принципы формирования, хранения и переноса зарядовых пакетов.
7. Параметры и характеристики ПЗС.
8. Разновидности ПЗС.
9. Основные элементы акустоэлектроники и их характеристики, достоинства, области применения, принципы работы.
10. Элементы СБИС на цилиндрических магнитных доменах.

ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ
С ЭЛЕКТРОСТАТИЧЕСКИМ
И ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ

Глава 11

**ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ
С ЭЛЕКТРОСТАТИЧЕСКИМ УПРАВЛЕНИЕМ**

11.1. Общие сведения

Электривакуумным прибором называется устройство, в котором рабочее пространство, изолированное газонепроницаемой оболочкой (баллоном), имеет высокую степень разрежения или заполнено специальной средой (парами или газами) и действие которого основано на электрических явлениях, связанных с движением заряженных частиц в вакууме или газе. В соответствии с характером рабочей среды электривакуумные приборы подразделяются на *электронные* и *ионные (газоразрядные)*.

В *электронном электривакуумном приборе (ЭВП)* электрический ток обусловлен движением только свободных электронов в вакууме. Семейство электронных вакуумных приборов весьма обширно и объединяет такие группы приборов, как *электронные лампы, электронно-лучевые приборы, электривакуумные фотоэлектронные приборы* и др.

Принцип действия *ионного электривакуумного прибора* основан на использовании свойств электрического разряда в газе или парах металлов. Эти приборы называют также *газоразрядными (ГРП)*. К ним относятся газоразрядные приборы дугового, тлеющего, высокочастотного разрядов и др.

Любой вакуумный электронный и газоразрядный прибор состоит из системы электродов, предназначенных для управления физическими процессами внутри баллона, отделяющего внешнюю среду от рабочего внутреннего пространства прибора.

В каждом типе электровакуумных и газоразрядных приборов создаются свои специфические системы электродов. Однако во всех типах электровакуумных и большинстве типов газоразрядных приборов имеются: *катоды* — электроды, испускающие (*эмитирующие*) электроны, и *аноды* — электроды, собирающие (*коллектирующие*) электроны. Для управления потоками заряженных частиц во многих приборах используются *управляющие электроды*, выполненные в виде сеток или профилированных пластин, и специальные электромагнитные элементы конструкции (катушки). Конструкции электродов очень разнообразны и определяются назначением приборов и условиями их работы. В приборах отображения информации в наглядной (визуальной) форме (электронно-лучевые трубки — ЭЛТ, индикаторы и другие приборы) широко используются специальные конструктивные элементы — *экраны*, с помощью которых энергия электронного потока или электрического поля преобразуется в оптическое излучение (свечение) тела.

Баллоны электровакуумных и газоразрядных приборов изготавливаются самой разнообразной формы из стекла, металла, керамики, а также из различных комбинаций этих материалов. Выводы от электродов делаются через цоколь, торцевые и боковые поверхности баллонов.

11.2. Основы эмиссионной электроники

Работа выхода. Чтобы сформировать поток свободных электронов, перемещающихся в вакууме или газе под действием электрических и магнитных полей, необходимо обеспечить выход электронов из твердого тела (чаще всего металла, полупроводника). Испускание электронов твердым телом называется эмиссией и осуществляется путем подведения к телу энергии от внешнего источника. Энергия, равная разности энергии E_0 электрона, покоящегося в свободном пространстве на расстоянии, где силами, действующими на электрон со стороны поверхности твердого тела, можно пренебречь, и энергии E_{Φ} , соответствующей уровню электрохимического потенциала системы электронов в твердом теле (уровню Ферми, см. п. 1.1), т. е. $A = E_0 - E_{\Phi}$, называется рабо-

той выхода (см. также п. 2.6). Работа выхода обычно выражается в электрон-вольтах (эВ).

Работа выхода электрона складывается в основном из работы по преодолению силы, действующей на электрон со стороны *двойного слоя*, и *силы зеркального отображения*. *Двойной электрический слой* образуется вылетевшими с поверхности катода электронами и положительными ионами решетки материала катода, эмитирующего (испускающего) электроны. Двойной слой образует тормозящее электрическое поле для электронов, вылетающих с поверхности катода. При удалении от поверхности тела на электрон действует удерживающая кулоновская сила, возникающая между удаляющимся электроном и наведенным в теле катода зеркально расположенным положительным электрическим зарядом (сила зеркального отображения). Работа выхода большинства чистых металлов, используемых в качестве катодов, лежит в интервале 1,8 (Cs) ... 5(Re) эВ.

Для уменьшения работы выхода на поверхность металлической основы (керна) наносят вещество с меньшей работой выхода, электроны которого переходят в kern. Вследствие этого на поверхности катода появляются положительные ионы, которые вместе с электронами, ушедшими в kern, формируют внутренний двойной электрический слой. Электрическое поле, создаваемое этим слоем, ускоряет вылетающие электроны, т. е. уменьшает работу выхода электронов. Например, при нанесении одноатомного слоя бария на поверхность вольфрама (W) работа выхода уменьшается с 4,5 эВ (чистый W) до 1,56 эВ (активированный W).

В зависимости от вида энергии, подводимой к веществу, различают *термо-, фото-, вторичную и электростатическую электронные эмиссии частиц*.

Термоэлектронная эмиссия осуществляется за счет нагрева вещества. С ростом температуры вещества увеличивается энергия, получаемая электронами, и возрастает вероятность совершения ими работы выхода и покидания металла.

Зависимость плотности тока термоэмиссии j_e от температуры тела (катода) T определяется следующим выражением:

$$j_e = A_0 T^2 e^{-\frac{E_0 - E_\phi}{kT}}, \quad (11.1)$$

где A_0 — константа, зависящая от материала катода; для различных веществ $A_0 = 10 \dots 300 \text{ А}/(\text{см}^2 \cdot \text{К}^2)$; k — постоянная Больцмана; T — абсолютная температура.

Электростатическая (автоэлектронная) эмиссия обусловлена сильным электрическим полем, действующим на поверхность катода. При большом положительном относительно катода потенциале электрода, расположенного рядом с катодом, у поверхности последнего происходит значительное уменьшение величины и ширины энергетического барьера, который нужно преодолеть электрону для выхода во внешнее пространство (в вакуум). В результате при некоторой величине напряженности электрического поля \mathcal{E} может возникнуть значительный ток эмиссии за счет туннельного перехода через барьер электронов с поверхности катода во внешнее пространство. Это явление называется **электростатической (автоэлектронной) эмиссией**. Напряженности электрических полей, необходимые для этого вида эмиссии, должны достигать значения порядка 10^6 В/см и более. Эта величина существенно (примерно на два порядка) меньше напряженности поля внутри двойного слоя, задерживающего электроны, которые преодолевают потенциальный барьер вследствие туннелирования.

Вторичная электронная эмиссия происходит при бомбардировке поверхности катода потоками быстрых заряженных частиц. Если используются электронные потоки, то отношение общего числа вторичных (выбитых с поверхности тела) электронов n_2 к числу первичных (падающих на эту поверхность из электронного потока) электронов n_1 называют коэффициентом вторичной эмиссии σ . Для большинства металлов и полупроводников максимальное значение $\sigma \sim 1$. У сложных соединений, включающих элементы с малой работой выхода, σ может достигать нескольких единиц.

Фотоэлектронная эмиссия обусловлена действием внешнего электромагнитного излучения, действующего на поверхность вещества. Катод, использующий фотоэлектронную эмиссию, называется **фотоэлектронным** или **фотокатодом**. Механизм фотоэлектронной эмиссии объясняется законами фотоэффекта А. Г. Столетова и уравнением А. Эйнштейна. Согласно закону Столетова, фототок I_Φ пропорционален световому потоку Φ , облучающему тело:

$$I_\Phi = k\Phi. \quad (11.2)$$

Кинетическая энергия эмитированных электронов определяется частотой ν падающего света (или энергией фотонов $h\nu$) и в соответствии с законом Эйнштейна вычисляется из уравнения

$$\frac{mv^2}{2} = h\nu - A_{\text{вых}}, \quad (11.3)$$

где h — постоянная Планка; $A_{\text{вых}}$ — работа выхода; v — скорость эмитированных электронов.

Частота падающего света ($\nu = \nu_{\text{кр}}$), при которой кинетическая энергия электронов равна нулю ($h\nu_{\text{кр}} = A_{\text{вых}}$), называется *пороговой частотой фотоэлектронной эмиссии*, соответствующей длинноволновой, или красной, границе внешнего фотоэффекта. Поскольку различные вещества имеют различную работу выхода, фотоэлектронная эмиссия для разных фотокатодов возникает при характерной для данного вещества критической частоте $\nu_{\text{кр}}$. Чувствительность фотокатода оценивается отношением числа эмитированных электронов к числу падающих фотонов (квантовый выход).

Катоды электронных ламп. В большинстве электровакуумных приборов применяются *термоэлектронные катоды (термокатоды)*, использующие термоэмиссию. Наиболее распространены *однородные металлические, активированные металлические, полупроводниковые и металлополупроводниковые термокатоды*. По способу нагрева термокатоды делятся на прямонакальные и подогревные.

Прямонакальные катоды выполняются различной формы из тонкой проволоки (или ленты), которую закрепляют в массивных держателях, подсоединяемых к источнику тока накала. *Подогревные катоды (косвенного канала)* содержат изолированную нить подогрева (подогреватель) и собственно катод, выполняемый обычно в виде металлического цилиндра с активированной внешней поверхностью. Прямонакальные катоды, как правило, запитываются постоянным током накала, а подогревные могут нагреваться и переменным.

Прямонакальные катоды из чистых металлов используются относительно редко. Они применяются в электрометрических мощных лампах и электронных лампах с высоким анодным напряжением, так как являются наиболее стойкими к разрушению под действием бомбардировки ионов, возникающих за счет ионизации остаточных газов в рабочем пространстве.

Активированные металлические катоды имеют достаточно много конструктивных модификаций. В любом из них поверхность вольфрама или молибдена активируется металлами с малой работой выхода: барием, торием, стронцием и др. Для увеличения активированной поверхности и, следовательно, тока эмиссии поверхностные слои или весь катод изготавливаются: из порис-

того вольфрама; путем прессования или спекания порошков окиси металла (никеля, вольфрама и др.) и карбонатов щелочноземельных металлов; путем нанесения гексаборида лантана или гексаборида бария на молибденовую или танталовую подложку. Активированные металлические катоды применяются в самых разнообразных электронных приборах. Они обладают существенно большей плотностью тока эмиссии (до 10 А/см^2) и имеют меньшую рабочую температуру ($T_{\text{раб}} \sim 1500 \dots 1700 \text{ }^\circ\text{C}$), чем однородные металлические ($j_e < 1 \text{ А/см}^2$, $T_{\text{раб}} > 2000 \text{ }^\circ\text{C}$), многие из них стойки к ионной бомбардировке, отравлению газами, хорошо работают в сильных электрических полях.

В полупроводниковых и металлополупроводниковых термокатадах в качестве эмитирующего слоя используются полупроводники. Наибольшее распространение в электронных и газоразрядных приборах получил *оксидный катод*, представляющий собой никелевый или вольфрамовый kern с нанесенным полупроводниковым слоем из смеси окислов бария, стронция и кальция. Хорошими характеристиками обладают *оксидно-бариевые* и *оксидно-ториевые катоды*, отличающиеся высокой стойкостью к ионной бомбардировке и отравлению газами, восстановлением эмиссии после отравления. Чисто оксидные катоды этими свойствами не обладают. Общим положительным свойством всех оксидных катодов является малая рабочая температура ($T_{\text{раб}} < 1000 \text{ }^\circ\text{C}$), высокая эмиссионная способность, достигающая в импульсном режиме величины $j_e \sim 150 \text{ А/см}^2$.

Для оценки свойств термокатодов используются следующие параметры: *плотность эмиссионного тока* j_e ; *рабочая температура катода* $T_{\text{раб}}$; *долговечность*, или *срок службы катода*; *эффективность катода* H . Последний параметр характеризует отношение тока эмиссии j_e к мощности P_H , подводимой к катоду для его подогрева.

Примерная эффективность катодов: однородных металлических — $5 \dots 15 \text{ мА/Вт}$; активированных — $20 \dots 70 \text{ мА/Вт}$; полупроводниковых — $80 \dots 200 \text{ мА/Вт}$.

Фотокатоды чаще всего выполняются в виде тонкого светочувствительного слоя металла с окисленной поверхностью, на которую осаждена тонкая пленка цезия. Наиболее распространены кислородно-цезиевые и сурьмяно-цезиевые фотокатоды, причем последние обычно наносятся на тонкую никелевую пленку — подложку.

11.3. Электронно-управляемые лампы

Электронные приборы осуществляют процесс преобразования одного вида энергии, чаще всего электрической, в другой. В электронно-управляемых лампах для управления электронным потоком используется различное количество электродов. Некоторые из них являются проницаемыми для электронного потока — их называют обычно сетками. Электронно-управляемые лампы могут иметь два (**диоды**), три (**триоды**), четыре (**тетроды**), пять (**пентоды**) и более электродов. Триоды содержат одну сетку, тетроды — две, пентоды — три и т. д.

Диоды. В диоде, имеющем анод и катод, обычно потенциал анода положителен относительно катода, поэтому электроны, эмитированные катодом, перемещаются к аноду и создают ток, равный I_a во внешней замкнутой цепи. Этот ток называют **анодным**. Если потенциал катода выше потенциала анода, то ток анода близок к нулю, т. е. диод проводит ток в одном направлении. Свойство односторонней проводимости используется для выпрямления переменного тока и преобразования ВЧ- и СВЧ-колебаний.

Величина анодного тока определяется анодным напряжением U_a , током эмиссии катода и рядом других факторов. Эмитированные катодом (К) отрицательно заряженные частицы (электроны) формируют отрицательный объемный заряд, создающий в пространстве между катодом и анодом (А) в направлении x (рис. 11.1, а, кривая 1) отрицательный потенциал. При положительных напряжениях U_a отрицательный потенциал, обусловленный объемным зарядом электронов (потенциальный барьер), может сохраняться только вблизи катода (см. рис. 11.1, а, кривая 2). В этом случае отрицательный потенциал (барьер) преодолевают только быст-

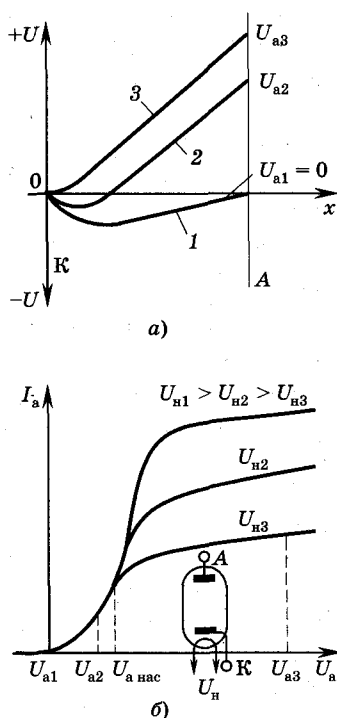


Рис. 11.1

рые (энергичные) электроны, медленные задерживаются. Быстрые электроны, преодолевшие барьер, создают анодный ток, который меньше тока эмиссии (11.1). При увеличении U_a потенциальный барьер у катода уменьшается и сдвигается к катоду; все больше электронов преодолевают барьер, и анодный ток I_a увеличивается.

В рассматриваемом случае величина тока I_a ограничена отрицательным барьером (потенциалом). Такой режим работы диода называется *режимом ограничения тока объемным зарядом*.

При некотором анодном потенциале $U_a > U_{a \text{ нас}}$ во всех точках междуэлектродного пространства потенциал становится положительным относительно катода (см. рис. 11.1, а, кривая 3), т. е. положительный потенциал анода полностью компенсирует отрицательный потенциал, формируемый объемным зарядом эмитированных электронов. В этом случае все электроны, эмитированные катодом в единицу времени, достигают анода и наступает режим насыщения.

Зависимость тока I_a от напряжения U_a ($I_a = f(U_a)$) при неизменной температуре катода, которая поддерживается постоянным напряжением нити канала U_n , нагревающей катод для обеспечения термоэмиссии, называется *анодной характеристикой*.

Из семейства анодных характеристик, показанных на рис. 11.1, б для разных значений напряжения накала U_n катода, видно, что при $U_a > U_{a \text{ нас}}$ анодный ток замедляет свой рост и достигает тока насыщения, равного току эмиссии. Из формулы (11.1) для плотности тока эмиссии следует, что его величина определяется температурой катода, т. е. растет с увеличением напряжения накала U_n . (С ростом U_n увеличивается и величина $U_{\text{нас}}$.)

Теоретический анализ показывает, что при $U_a < U_{a \text{ нас}}$ анодная характеристика диода описывается так называемым «законом степени 3/2 (трех вторых)», т. е. $I_a = GU_a^{3/2}$, где G — коэффициент, зависящий от конструкции электродов и прибора в целом. Следует отметить, что при нулевом (и даже некотором отрицательном) анодном напряжении U_a анодный ток I_a хоть и весьма мал, но не точно равен нулю (см. рис. 11.1, б), как это следует из закона трех вторых. При выводе этого закона не учитываются разброс электронов, вылетающих из катода, по скоростям и ряд других факторов, приводящих к тому, что некоторые наиболее энергичные электроны преодолевают потенциаль-

ный барьер и достигают анода. Однако отрицательная часть анодной характеристики диода практической роли не играет.

Диоды используются для детектирования и ограничения сигналов, выпрямления переменного напряжения и т. д. Для усиления и других более сложных преобразований сигналов в основном применяются триоды, тетроды и пентоды. С этой целью в этих приборах используются одна, две или три сетки.

Триоды. В триодах между катодом и анодом располагается сетка (С). При изменении потенциала сетки, который может быть как положительным, так и отрицательным относительно катода, происходит сильное изменение распределения электрического поля (потенциала) и, соответственно, объемного заряда в междуэлектродном пространстве, особенно в прикатодной области (рис. 11.2, а).

В случае больших отрицательных напряжений на сетке (даже при положительных (но не слишком больших) потенциалах анода) все электроны могут быть возвращены к катоду, т. е. в этом случае во всем пространстве между катодом и сеткой существует тормозящее электроны поле, сформированное отрицательным потенциалом сетки и объемным зарядом электронов в областях, примыкающих к катоду (пространство катод (К) — сетка (С), см. рис. 11.2, а, кривая 1). Ток анода в этих условиях равен нулю.

Наименьший по модулю отрицательный потенциал сетки, при котором $I_a = 0$, называют **напряжением запирания** U_{c0} для заданного напряжения анода.

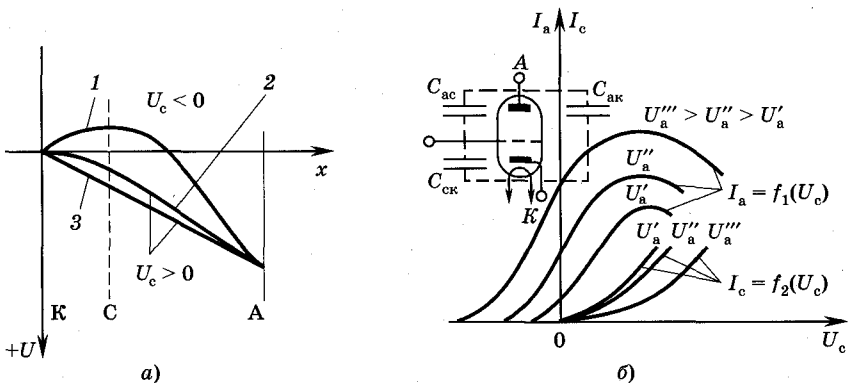


Рис. 11.2

Если напряжение на сетке U_c по модулю меньше напряжения запирающего, т. е. $|U_c| < |U_{c0}|$, то электрическое поле между витками сетки в пространстве катод—сетка делается ускоряющим, и часть электронов, эмитированных катодом в направлении таких областей, устремляется к аноду. Возникает анодный ток, который возрастает по мере снижения отрицательного потенциала сетки. При положительных потенциалах на сетке распределение потенциалов показано на рис. 11.2, а (кривые 2 и 3). Часть электронов, перемещающихся в непосредственной близости от витков, притягивается сеткой и образует сеточный ток I_c , в то время как основная доля электронного потока попадает на анод, создавая ток I_a .

Таким образом, катодный ток I_k разветвляется на два тока — ток сетки и ток анода, т. е. $I_k = I_a + I_c$. Как правило, $I_c < I_a$ из-за меньшей площади поверхности сетки по отношению к аноду (см. ниже).

Для сравнения влияния потенциалов сетки и анода на потенциальный барьер у катода и, следовательно, на число электронов, преодолевающих барьер, триод заменяют эквивалентным диодом, анод которого располагают на месте сетки. Анодное напряжение эквивалентного диода, при котором катодные токи триода и диода равны, называют *действующим напряжением* U_d . Равенство катодных токов означает, что напряженности электрического поля в области катода у триода и эквивалентного диода равны, а это означает, что равны и заряды на поверхности катода, наведенные приложенными к электродам потенциалами.

Заряды, индуцированные на катоде триода и диода, определяются междуэлектродными емкостями. Междуэлектродная емкость диода определяется площадью поверхности анода и катода и расстоянием между этими электродами. В триоде, помимо емкости анод—катод $C_{ак}$, существуют емкости сетка—катод $C_{ск}$ и анод—сетка $C_{ас}$. Емкость $C_{ск}$ называется входной, а емкость $C_{ас}$ — проходной.

Отношение емкостей $C_{ак}/C_{ск}$ определяет один из основных параметров триода — *проницаемость* D , которая характеризует степень проникновения поля анода в пространство катод—сетка (прикатодное пространство) и, соответственно, влияет на величину анодного тока.

Емкости $C_{ск}$ и $C_{ас}$ определяются конструкцией сеток, в частности их густотой и расположением между анодом и катодом.

Итак, при постоянном напряжении накала (постоянной температуре катода, постоянном токе эмиссии) анодный I_a и сеточный I_c токи зависят от напряжения на аноде и напряжения на сетке U_c , т. е. $I_a = f_1(U_a, U_c)$ и $I_c = f_2(U_a, U_c)$.

Если одну из независимых переменных выбрать в качестве параметра (так называемого *параметра режима*), то указанные функции можно представить в виде следующих четырех зависимостей: семейство анодно-сеточных характеристик $I_a = f_1(U_c)$ при $U_a = \text{const}$ и сеточных (входных) характеристик $I_c = f_2(U_c)$ при $U_a = \text{const}$ (рис. 11.2, б), а также семейства анодных (выходных) характеристик $I_a = f_3(U_a)$ (рис. 11.3) и сеточно-анодных характеристик $I_c = f_4(U_a)$ при $U_c = \text{const}$ (одна из кривых этого семейства показана на рис. 11.3 штриховой линией).

Увеличение анодного напряжения вызывает смещение начала анодно-сеточных характеристик влево из-за большего проникновения электрического поля в прикатодную область. При отрицательных напряжениях на сетке ток в анодной цепи появляется лишь при таких положительных напряжениях анода, при которых действующее напряжение $U_d > 0$.

При $U_c > 0$ появляется ток I_c , и анодный ток I_a уже не будет равен катодному, так как $I_k = I_a + I_c$, т. е. катодный ток, распределяется между анодом и сеткой. Анодно-сеточная характеристика идет заметно более полого, чем зависимость катодного тока от потенциала анода и сетки.

При $U_c > U_0$ ускоряющее поле существует как между сеткой и катодом, так и в пространстве анод—сетка. Электроны, пролетая между витками сетки, подвергаются воздействию положительно заряженных витков, и их траектории искривляются. Ток сетки образуется только электронами, которые попадают на витки сетки (*режим перехвата*). В этом режиме отношение I_a/I_c зависит от соотношения площадей поверхностей сетки и анода. Площадь поверхности витков сетки обычно меньше площади анода, поэтому ток анода больше тока сетки.

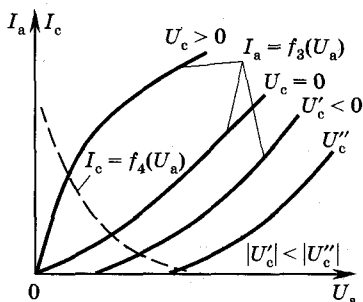


Рис. 11.3

Одновременный рост сеточного I_c и анодного I_a токов при увеличении напряжения U_c обусловлен в основном изменением катодного тока из-за увеличения действующего напряжения. Когда $U_c > U_a$, часть электронов перехватывается сеткой, другая часть электронов, пролетевших между витками по искривленным траекториям, попадает в тормозящее поле между сеткой и анодом. Это тормозящее поле преодолевают не все электроны, часть из них возвращается назад к сетке. В *режиме возврата* при увеличении U_c ток I_a уменьшается, а ток I_c возрастает, поскольку чем меньше анодное напряжение, тем большее число электронов возвращается к сетке (см. рис. 11.2, б). Переход от режима возврата к режиму перехвата зависит от соотношения расстояний катод—сетка и катод—анод, а также от густоты сетки и может происходить при различных значениях отношения U_a/U_c . Для большинства триодов это значение колеблется в пределах от 0,2 до 1,2. Отношение токов $k_T = I_a/I_c$ называется *коэффициентом токораспределения*.

Рассмотрим характеристики $I_a = f_3(U_a)$ и $I_c = f_4(U_a)$ при $U_c = \text{const}$ (см. рис. 11.3). При $U_c < 0$ ток сетки $I_c = 0$ и $I_a = I_k$, при этом анодный ток возникает только при некотором положительном потенциале на аноде U_{a0} . Величина этого потенциала определяется из соотношения положительного анодного и отрицательного сеточного напряжений, при которых действующее напряжение равно нулю: $U_d = 0$. Чем выше отрицательный потенциал на сетке, тем больше характеристики сдвигаются вправо. При $U_c > 0$ анодные характеристики начинаются из нуля и имеют два достаточно ярко выраженных участка — крутой и более пологий. Более крутой участок характеристик соответствует режиму возврата электронов к сетке, т. е. $U_c > U_a$. В пространстве между сеткой и анодом существует объемный заряд, обусловленный возвращающимися к сетке электронами. Анодное поле сильно влияет на объемный заряд (на траекторию этих электронов), поэтому увеличение напряжения U_a приводит к резкому возрастанию тока I_a и соответственному уменьшению тока I_c . При $U_a > U_c$ образуется ускоряющее поле в промежутке сетка—анод и наступает режим прямого перехвата, при котором возрастание I_a с увеличением U_a происходит медленнее, так как оно теперь обусловлено только влиянием анодного напряжения на объемный заряд у катода. В режиме перехвата сеточный ток также изменяется мало, потому что в этом случае число элект-

ронов, попадающих на сетку, в основном определяется площадью сетки. С уменьшением положительного напряжения на сетке кривая $I_c = f(U_a)$ смещается вниз, так как из-за уменьшения U_d снижается катодный ток.

Триоды в основном используются для преобразования информации, в частности для усиления и генерации электрических колебаний. Триод, как, впрочем, и другие электронные приборы, может осуществлять необходимые преобразования, если он включен в определенную электрическую цепь (схему). В этом случае в анодной, сеточной или катодной цепях включены резисторы, катушки индуктивности, колебательные контуры и другие элементы. Наличие этих элементов приводит к тому, что изменения напряжения на каком-либо из электродов вызывает изменение как токов в цепях электродов, так и напряжений на всех электродах.

Триоды имеют ряд существенных недостатков, в частности, относительно малый коэффициент усиления μ (см. ниже) и большую проходную емкость (емкость сетка—анод C_a). Эта емкость и определяемая ею проводимость создают в триоде обратную связь между входной и выходными цепями, что может привести к возбуждению усилителя на высоких частотах или к ухудшению характеристик устройства. От этих недостатков в значительной мере свободны такие приборы, как тетроды и пентоды, имеющие соответственно две и три сетки для управления параметрами электронного потока.

Тетроды. Пентоды. Наличие одной сетки (тетрод) или двух (пентод) между анодом и управляющей сеткой создает электростатический экран, снижает проходную емкость и сильно изменяет распределение потенциалов (рис. 11.4, а и рис. 11.5, а).

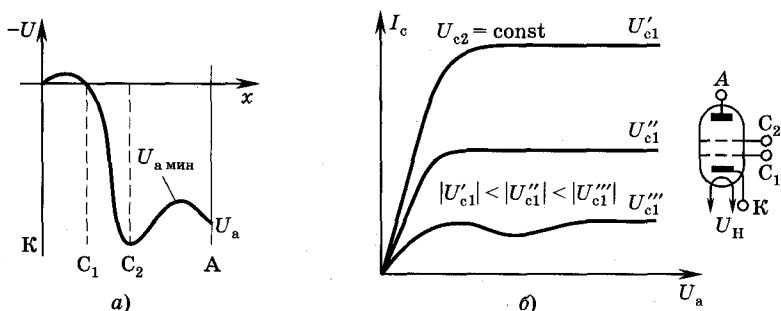


Рис. 11.4

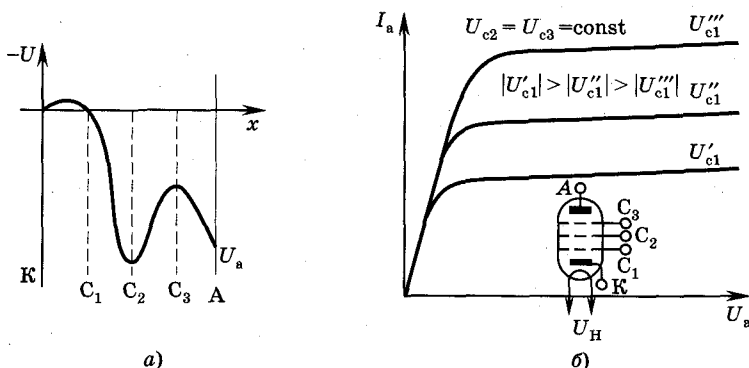


Рис. 11.5

Токи анода и сеток в тетроре и пентоде являются функциями трех (тетрод) или четырех (пентод) напряжений — соответственно U_a , U_{c1} и U_{c2} (тетрод) и U_a , U_{c1} , U_{c2} и U_{c3} (пентод) (U_{c2} — напряжение на 2-й сетке, называемой экранирующей, а U_{c3} — напряжение на 3-й сетке, называемой защитной). Поэтому семейство характеристик тетрода или пентода представляет собой зависимость любого из токов (анодного, сеточных) от какого-либо из напряжений (U_a , U_{c1} , U_{c2} , U_{c3}) (при постоянных остальных напряжениях, одно из которых выбрано в качестве параметра семейства). Типичные анодные характеристики тетрода и пентода приведены на рис. 11.4, б и 11.5, б.

Для оценки свойств электровакуумных приборов так же, как и транзисторов, на практике используют различные параметры.

Основными параметрами диодов являются крутизна анодной характеристики S , внутреннее сопротивление R_i , сопротивление постоянному току R_0 , междуэлектродные емкости, предельно допустимое обратное напряжение $U_{обр}$ и максимальная мощность, рассеиваемая анодом.

Крутизна анодной характеристики диода $S = dI_a/dU_a$ при постоянном напряжении накала (температуре) катода, т. е. при $U_H = \text{const}$, показывает, на сколько изменяется анодный ток при изменении анодного напряжения на 1 В. Дифференциальное (внутреннее) сопротивление диода $R_i = 1/S$ можно рассматривать как сопротивление переменному току.

Сопротивление лампы при постоянном токе (статическое сопротивление) $R_0 = U_a/I_a$, где U_a и I_a — напряжение и ток в заданной точке анодной характеристики.

У триода дополнительно появляются такие параметры, как крутизна анодно-сеточной (S) и сеточной (S_c) характеристик, статический коэффициент усиления μ , проницаемость ламп D .

В тетрадах из-за наличия экранирующей и защитных сеток появляются параметры, связанные с изменениями токов и напряжений на этих сетках, влияющие на параметры по анодной цепи и цепи управляющей сетки.

Крутизна анодно-сеточной характеристики триода $S = dI_a/dU_c$ при $U_a = \text{const}$ определяет изменение анодного тока триода при изменении напряжения на сетке на 1 В при постоянном анодном напряжении. Для триодов $S = 1 \dots 50$ мА/В.

Крутизна сеточной характеристики $S_c = dI_c/dU_c$ определяет изменение сеточного тока при изменении напряжения на сетке на 1 В при $U_a = \text{const}$ и $U_n = \text{const}$.

Внутреннее сопротивление триода $R_i = dU_a/dI_a$ определяется при $U_c = \text{const}$ и $U_n = \text{const}$, а для тетродов и пентодов должны быть постоянными напряжения на 2-й и 3-й сетках, т. е. $U_{c2} = \text{const}$ и $U_{c3} = \text{const}$ (для триода $R_i = 1 \dots 100$ кОм; для тетрода $R_i = 10 \dots 50$ кОм; для пентода $R_i = 10 \dots 100$ кОм).

Величину $1/S_c = R_{i_c}$ называют **входным сопротивлением лампы** (для триода ($R_{i_c} = 100$ МОм)).

Статический коэффициент усиления μ определяет относительное влияние и сеточного, и анодного напряжений на анодный ток (а не на катодный), т. е. $\mu = -dU_a/dU_c = -SR_i$ при $I_a = \text{const}$. Постоянство анодного тока ($I_a = \text{const}$) при изменении как анодного напряжения U_a , так и сеточного U_c возможно, если изменение анодного тока dI_a при увеличении или уменьшении анодного напряжения на величину dU_a компенсируется соответствующим изменением сеточного напряжения на величину dU_c . Отношение изменений dU_a и dU_c , необходимых для поддержания тока $I_a = \text{const}$, и определяет μ . Знак «минус» в формуле для μ означает, что указанная компенсация изменения анодного тока происходит таким образом, что изменение U_a должно быть противоположно по знаку изменению U_c . Для триодов $\mu = 1 \dots 100$, для тетродов $\mu = 100 \dots 700$, а для пентодов — несколько тысяч.

Отметим, что S , R_i и μ определяют влияние напряжений на электродах на анодный ток, а не на катодный. Влияние напряжений U_c и U_a на катодный ток определяется проницаемостью

лампы $D = -dU_c/dU_a$ при $I_k = \text{const}$. При $U_c < 0$ $I_k = I_a$, так как $I_c = 0$ и в этом случае $\mu = 1/D$.

Емкость $C_{ск}$ (управляющая сетка—катод) называют входной, емкость $C_{ак}$ (анод—катод) — выходной, а емкость $C_{ас}$ (анод—сетка) — проходной.

Помимо перечисленных, в тетродах и пентодах иногда используют такой параметр, как статический коэффициент усиления $\mu_{с2}$, отображающий относительное влияние напряжений U_a и $U_{с2}$ на анодный ток лампы:

$$\mu_{с2} = -dU_a/dU_{с2} \text{ при } I_a, U_{с1}, U_{с3} = \text{const}.$$

В некоторых случаях нагрузка может быть включена в цепь экранирующей сетки, при этом используют статические параметры, характеризующие зависимость тока $I_{с2}$ от напряжений на электродах лампы.

В современных многоэлектродных лампах для усиления сигналов с широким спектром частот при их конструировании стремятся уменьшить междуэлектродные емкости и увеличить крутизну анодно-сеточной характеристики. Для этого между катодом и управляющей сеткой иногда вводят специальную катодную сетку, на которую подается небольшое положительное напряжение, или же используют вторичную эмиссию со специальных дополнительных электродов. Среди ламп особого назначения, обладающих высокой экономичностью, повышенной стойкостью к ударам и вибрациям и высокой надежностью, можно выделить миниатюрные приемно-усилительные стержневые лампы, лампы со штампованными рамочными электродами, нувисторы.

В *стержневых лампах* сетки выполнены в виде жестких металлических стержней, расположенных параллельно катоду. Выводы от электродов стержневых ламп выполняются из гибких проволочек. Для уменьшения междуэлектродной емкости вывод анода выполняется в верхней части баллона.

Нувисторы — это сверхминиатюрные металлокерамические лампы повышенной надежности. Они устойчивы к механическим воздействиям и работают в широком диапазоне температур ($-60 \dots 200$ °С).

В *электрометрических лампах*, которые обладают очень большим входным сопротивлением и применяются для измерения сверхмалых токов (до 10^{-15} А), управляющую сетку для уменьшения токов утечки крепят с помощью кварцевых или керамических изоляторов.

11.4. Мощные генераторные и модуляторные лампы

Одно из основных применений электронных ламп — использование их в мощных усилителях, генераторах и модуляторах. Электронные лампы должны удовлетворять следующим требованиям: обладать большой допустимой мощностью рассеяния на аноде $P_{a \text{ макс}}$ и сетках $P_{c \text{ макс}}$, иметь высокие рабочие напряжения U_a , пропускать большие анодные токи I_a , иметь малые междуэлектронные емкости, большие статические коэффициенты усиления, «правые» анодно-сеточные характеристики (незначительные напряжения запираия), хорошую линейность характеристик и т. д.

Кроме перечисленных требований, обращается большое внимание на качество сигнала, которое определяется оптимальным сочетанием широкой полосы частот, высокого коэффициента усиления и хорошей линейности характеристик электронных приборов. Следовательно, при разработке одним из основных направлений повышения качества сигналов радиоэлектронных систем является улучшение линейности анодно-сеточных характеристик электронных ламп. Лампы с хорошей линейностью характеристик позволяют повысить и КПД. Для получения большого КПД форма нижнего искривленного участка анодно-сеточной характеристики (рис. 11.6, а) подбирается таким образом, чтобы начальный ток, т. е. ток в рабочей точке при отсутствии сигнала, был как можно меньшим. Искажения сигнала из-за взаимной модуляции будут малы, если напряжение смещения вы-

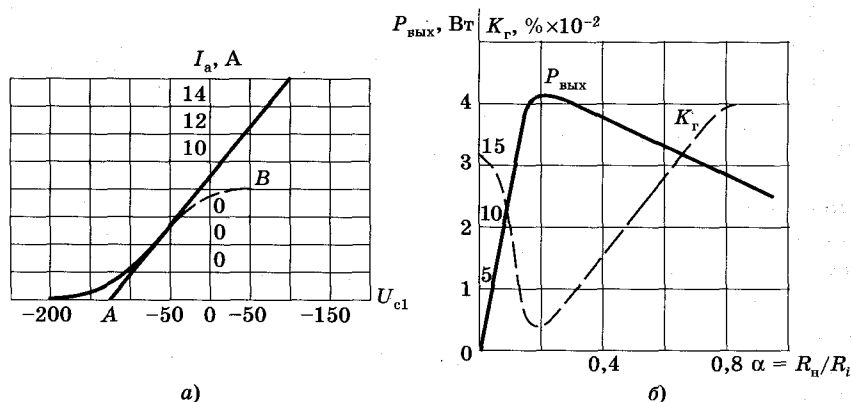


Рис. 11.6

брано вблизи точки пересечения продолжения линейного участка анодно-сеточной характеристики с осью сеточного напряжения (точка A на рис. 11.6, a). Очевидно, что всякие отклонения характеристик от линейности приводят к искажению формы сигнала в выходной цепи по сравнению с его формой на входе. Основными причинами, приводящими к отклонению характеристик от идеальной (линейной) формы в начальной области, являются следующие:

- 1) наличие распределения электронов по скоростям и углам при их вылете с катода;
- 2) неоднородность электрического поля у катода, обусловленная конструкцией управляющей и экранной сеток.

В области больших анодных токов при напряжениях на первой сетке, близких к нулю (см. рис. 11.6, a , кривая B), искривление характеристик связано с уменьшением объемного заряда у катода и изменением отношения тока экранной сетки к анодному току (I_{c2}/I_a). Изменение этого отношения вызвано следующими причинами:

- 1) при потенциале анода, меньшем потенциала второй сетки, реализуется режим возврата, т. е. основная доля электронов не может преодолеть влияние тормозящего поля в пространстве между второй сеткой C_2 и анодом A и возвращается на экранную сетку;
- 2) при большой энергии первичных электронов возникает вторичная эмиссия с экранной сетки, приводящая к увеличению тока I_a ;
- 3) при больших плотностях анодного тока часть электронов попадает на вторую сетку из-за наличия расталкивающих сил объемного заряда.

Уменьшение влияния рассмотренных причин на линейность анодно-сеточных характеристик достигается тщательной и кропотливой оптимизацией конструкции всех электродов и прибора в целом. Кроме того, проводится оптимизация режима работы каскадов, так как нагрузки усилителей и генераторов могут существенно влиять на величину нелинейных искажений сигнала. Для получения максимальной мощности в нагрузке при минимальном коэффициенте нелинейных искажений сопротивление нагрузки R_n следует выбирать из условия $R_n = \alpha_n R_i$, где коэффициент нагрузки $\alpha_n = 0,05 \dots 0,15$. Для оценки нелинейных искажений обычно используется коэффициент гармоник k_p .

В первом приближении k_{Γ} может быть определен отношением амплитудных значений I_{m2} и $I_{m a}$ токов второй и первой гармоник соответственно, т. е.

$$k_{\Gamma} = I_{m2}/I_{m a}.$$

Расчеты и анализ работы усилителей и генераторов показывают, что существует оптимальное значение $\alpha_{\text{н опт}}$, при котором мощность $P_{\text{вых}}$ максимальна, а коэффициент k_{Γ} — минимален, что иллюстрируется зависимостями на рис. 11.6, б.

Получение большой колебательной мощности $P_{\text{к}}$ является также важнейшим требованием, предъявляемым к мощным лампам. Предельные значения колебательной мощности определяются допустимой мощностью, рассеиваемой анодом, т. е. $P_{\text{а макс}}$. В зависимости от $P_{\text{к}}$ различают электронные лампы малой, средней и большой мощностей. Выпускаемые промышленностью мощные триоды и тетроды с линейными характеристиками обеспечивают уровни мощности в непрерывном режиме вплоть до 25 кВт на частотах до 500 МГц, а на частотах порядка 1 МГц — до 500 кВт и более. Коэффициенты усиления по мощности могут достигать 25...30 дБ.

В устройствах малой и средней мощностей обычно используются пентоды и тетроды, а в мощных генераторах и усилителях — триоды и тетроды.

Высокие уровни колебательной мощности реализуются при больших амплитудах колебаний анодного тока. В результате при работе генераторов и модуляторов мгновенные напряжения на управляющей сетке могут принимать большие положительные значения, т. е. лампы работают со значительными сеточными токами и имеют правые анодно-сеточные характеристики. При больших амплитудах переменного сеточного напряжения в лампах последовательно реализуются режимы отсечки, перехвата и возврата. В режиме возврата резко возрастают токи (особенно в цепи управляющей сетки) и потери мощности $P_{\text{с}}$ в цепях сеток. В этом случае наблюдается так называемый перенапряженный режим. Режиму перехвата соответствует недонапряженный режим, а в области перехода от перехвата к возврату электронов — критический режим. Таким образом, чем больше коэффициент токораспределения, тем меньше ток и мощность потерь в цепи сетки. Для обеспечения правых характеристик управляющие сетки ламп должны быть густыми, т. е. пронизаемость — малой, а статический коэффициент усиления — большим.

Модуляторные триоды усиливают низкочастотные сигналы и работают при отрицательных смещениях на сетке, т. е. имеют левые анодно-сеточные характеристики, когда сеточные токи отсутствуют. По сравнению с генераторными эти приборы имеют редкую сетку и, следовательно, незначительный статический коэффициент усиления.

В импульсных генераторных и модуляторных лампах значение импульсной мощности ограничивается максимально возможной величиной эмиссионного тока катода и максимально допустимыми напряжениями между электродами (напряжениями пробоя). При импульсной работе лампы могут обеспечивать мощность в 10^3 и более раз большую по сравнению с непрерывным режимом.

Конструкция мощных электронных приборов требует специальных мер отвода тепловой энергии от анода, чтобы не произошло его разрушения. В зависимости от способа отбора тепловой энергии от анода триоды делятся на следующие типы: с лучистым, воздушным, водяным, испарительным или кондуктивным охлаждением и охлаждением с помощью тепловых трубок.

В триодах с *воздушным охлаждением* мощностью до 100 кВт катод и сетка располагаются внутри анода. Анод, выполняемый обычно из меди, конструктивно объединен со стеклянным или керамическим баллоном и составляет его часть. Для увеличения теплоотвода поверхность анода увеличивают с помощью ребристого радиатора и лампу помещают в герметизированный объем, через который принудительно прогоняют с помощью вентилятора очищенный от пыли, грязи и паров масла воздух, охлаждающий, помимо анода, баллон и другие элементы конструкции.

В триодах с *водяным и испарительным охлаждением* анод с частью баллона помещают в закрытый герметичный сосуд, через который под давлением в $(2...3) \cdot 10^5$ Па пропускаются потоки воды. В лампах с испарительным охлаждением анод охлаждается за счет отбора теплоты парообразования.

Мощные лампы делают разборными для возможности ремонта и замены отдельных деталей. Эти приборы при эксплуатации требуют специальной технологии обслуживания.

В триодах, предназначенных для работы в усилителях напряжения и мощности импульсных сигналов и сигналов высокой и сверхвысокой частоты, предусмотрены специальные меры для уменьшения паразитных индуктивностей сеточных и катодных

выводов, междуэлектродных емкостей. С этой целью уменьшают размеры электродов и применяют баллоны и выводы электродов специальной конструкции. Так, например, делают дисковые выводы, обладающие малой индуктивностью и обеспечивающие включение триода в ВЧ- и СВЧ-тракты, между анодом и сеткой располагают специальные экраны, которые улучшают экранирование анода и, следовательно, уменьшают емкость сетка—анод.

11.5. Эквивалентные схемы электронных ламп

Как отмечалось в гл. 3, 4, 5 и 6, при расчете и разработке радиоэлектронных устройств электронные приборы представляют в виде различного рода моделей. Достаточно полно свойства приборов могут быть отражены на основе электрических (электротехнических) моделей, которые представляются в виде эквивалентных схем. В общем случае между электродами электронных ламп приложены постоянные напряжения, определяющие положение рабочих точек на характеристиках, и некоторые переменные напряжения, обусловленные преобразуемыми сигналами.

Эквивалентная схема двухэлектродной лампы. В случае малых сигналов нелинейностью анодной характеристики в окрестности рабочей точки можно пренебречь (считать ее в этой окрестности линейной) и представить диод в виде эквивалентного сопротивления R_i , равного дифференциальному (внутреннему) сопротивлению диода в данной точке анодной характеристики. Если частота преобразуемого сигнала достаточно велика, то необходимо учитывать влияние междуэлектродной емкости C , которая включается параллельно R_i . Емкость C учитывается, начиная с частот, когда емкостное сопротивление $X_C = 1/(\omega C)$ сравнимо с R_i или меньше. Диод в этом случае изображается в виде комплексного сопротивления \dot{Z} (рис. 11.7).

Эквивалентные схемы триодов и многоэлектродных ламп. В триоде, представляемом в виде активного линейного четырехполюсника, выходной ток I_a и входной ток I_c связаны с входными и выходными напряжениями через характеристические проводимости. Используя соотношения, подобные уравнениям, рассмотренным при анализе

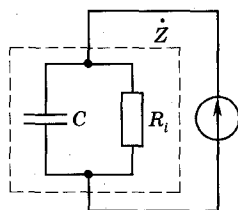


Рис. 11.7

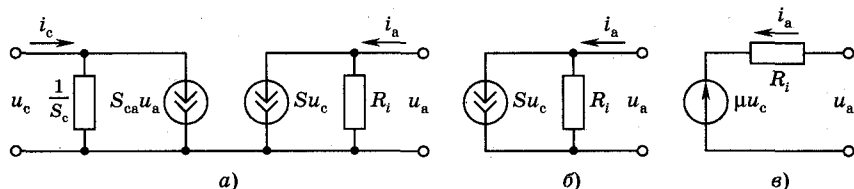


Рис. 11.8

транзисторов (4.4), и линейное приближение, когда в уравнениях приращения заменяются мгновенными малыми значениями переменных составляющих токов i_c , i_a и напряжений u_c , u_a , систему уравнений для линейного четырехполюсника запишем в следующем виде:

$$i_c = S_c u_c + S_{ca} u_a; \quad (11.4, a)$$

$$i_a = S u_c + u_a / R_i. \quad (11.4, б)$$

Эти уравнения можно представить в виде некоторой схемы замещения (рис. 11.8, а), в которой u_c и i_c являются входными, а u_a и i_a — соответственно выходными напряжениями и токами. Если $i_c = 0$, то схема замещения превращается в эквивалентную схему анодной цепи триода с генератором тока $S u_c$ (рис. 11.8, б). На этой схеме учтено и сопротивление нагрузки. Решая уравнение (11.4, б) относительно u_a и учитывая, что $S R_i = \mu$, получаем

$$u_a = R_i i_a - \mu u_c. \quad (11.5)$$

Этому выражению соответствует эквивалентная схема анодной цепи с генератором напряжения, представленная на рис. 11.8, в, где изображено и сопротивление нагрузки.

Проведенное рассмотрение справедливо не только для триода. Если тетрод или пентод используются с усилительной схемой, то цепи экранирующей и защитной сеток заземлены по переменной составляющей. Защитная сетка обычно соединена с катодом и через большую емкость, шунтирующую резистор смещения в цепи катода, подключена к земле. Экранирующая сетка также через большую емкость обычно соединяется с землей. Поэтому в цепях этих электродов не выделяются переменные напряжения. Входной цепью усилителя на тетраде или пентоде, как и усилителя на триоде, является цепь управляю-

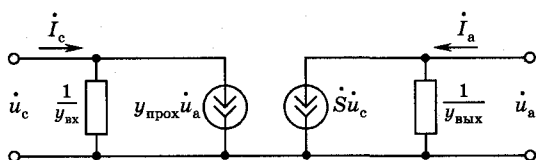


Рис. 11.9

щей сетки. Усиленное напряжение снимается с анодной нагрузки, и анодная цепь лампы оказывается выходной. Таким образом, схемы на рис. 11.8 могут служить еще и эквивалентными схемами для тетрода или пентода.

При рассмотрении работы лампы на высоких частотах характеристические проводимости лампы из-за наличия распределенных реактивностей (включая междуэлектродные емкости) становятся комплексными величинами. Эквивалентная схема лампы на высоких частотах для случая, когда I_a определяется только входной проводимостью, будет иметь вид, представленный на рис. 11.9. (Здесь \dot{I}_c , \dot{I}_a , \dot{u}_c , \dot{u}_a — комплексные токи и напряжения соответственно.)

В этом случае проводимости лампы могут быть определены согласно выражениям:

$$y_{\text{прох}} = \dot{I}_c / \dot{u}_c = j\omega C_{\text{ас1}}; \dot{y}_{\text{вых}} = \dot{I}_a / \dot{u}_a = 1/R_i + j\omega C_{\text{вых}};$$

$$C_{\text{вых}} = C_{\text{ак}} + C_{\text{ас3}} + C_{\text{ас2}}; S = \dot{I}_a / \dot{u}_c;$$

$$\dot{y}_{\text{вх}} = \dot{I}_c / \dot{u}_c = G_{\text{вх}} + j\omega C_{\text{вх}}; C_{\text{вх}} = C_{\text{ск}} (\text{триод});$$

$$C_{\text{вх}} = C_{\text{с1к}} + C_{\text{с1,с2}} (\text{пентод}); G_{\text{вх}} = 4/R_{\text{ic}}; R_{\text{ic}} — \text{сопротивление «сетка—катод»}.$$

11.6. Вакуумные интегральные схемы

В вакуумных интегральных схемах (ВИС) активными элементами являются электроракуумные лампы с размерами, близкими к размерам интегральных транзисторов. Электроракуумные микролампы по своим свойствам во многом подобны полевым транзисторам. Вакуумные интегральные триоды называют также *вакуумными полевыми транзисторами*, а их электроды — *катод, сетка, анод* называются соответственно *эмиттер, затвор, коллектор*.

В основе работы ВИС лежат те же физические явления, что и в основе работы рассмотренных электроракуумных триодов.

ВИС обладают рядом уникальных характеристик, в частности, скорость электронов в них может быть намного больше, чем в любом полупроводнике, их частотные свойства лучше частотных свойств кремниевых ИС и сравнимы со свойствами арсенид-галлиевых. Кроме того, ВИС обладают лучшей радиационной стойкостью. При изготовлении ВИС используется хорошо отработанная технология полупроводниковых ИС.

Одной из основных проблем при создании ВИС является разработка холодных (неподогреваемых) эмиттеров (катодов). В ВИС используется в основном электростатическая (автоэлектронная) эмиссия.

Наиболее исследованы интегральные триоды, устройство которых включает холодный эмиттер (катод), управляющий электрод (затвор — аналог управляющей сетки) и коллектор (анод).

Можно выделить четыре типа холодных катодов, используемых в ВИС:

- 1) полупроводниковые (кремниевые) решетки эмиттеров;
- 2) металлические катоды;
- 3) структуры металл—диэлектрик—полупроводник (МДП-структуры);
- 4) обратнсмещенные p — n -переходы, диоды Шоттки и p — i — n -диоды.

Структуры планарных приборов, создаваемых с использованием катодов 1-го и 2-го типов, во многом схожи (рис. 11.10).

При изготовлении металлических эмиттеров 2 (см. рис. 11.10) используется диэлектрическая изолирующая подложка 1 (например, стеклянная), на которой и формируется решетка эмиттеров. Расстояние между электродами (катод — 2, управляющий электрод — 3, анод — 4) колеблется от десятых долей микрона до нескольких микрон; радиус закругления кончика эмиттера — 20...30 нм; напряжение на коллекторе — от единиц до десятков, а иногда и до нескольких сотен В. В таких условиях в области эмиттера создается поле с напряженностью около 10^7 В/см, которое достаточно для создания силы, необходимой для выхода электронов из катода, т. е. для осуществления электростатической эмиссии. С помощью таких эмиттеров можно получить плотность тока $10^2...10^3$ А/см².

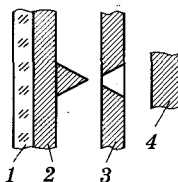


Рис. 11.10

Хорошими эмитирующими свойствами обладают и МДП-структуры (диоды) (рис. 11.11). В эмиттерах такого типа работа выхода полупроводника (обычно n -кремния) меньше работы выхода металла Al.

Металл и полупроводник разделены пленкой (15...20 мкм) диэлектрика (SiO_2) (см. рис. 11.11). При подаче прямого смещения между Al и n -Si электроны из n -Si через пленку SiO_2 туннелируют в металл и занимают энергетические уровни выше уровня Ферми металла $E_{\Phi M}$ (рис. 11.12, а), тем самым снижая его работу выхода и обеспечивая эмиссию электронов из структуры (кривая 1 на рис. 11.13).

Если же на рассматриваемую структуру катода подать обратное смещение («минус» источника подключается к металлу, рис. 11.12, б), то электроны будут туннелировать через SiO_2 из металла в полупроводник и ток потечет подобно тому, как это происходит для малых напряжений при прямом смещении. Однако при увеличении обратного напряжения обедненная область простирается все больше и больше в полупроводник, т. е. приложенное напряжение падает в основном на обедненном слое полупроводника. Это ограничивает число электронов, туннелирующих через оксидный слой, что вызывает насыщение тока (кривая 2 на рис. 11.13).

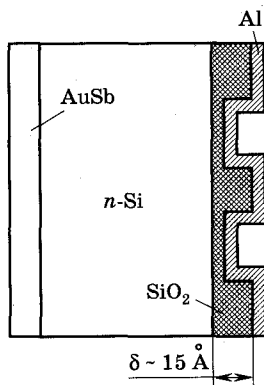


Рис. 11.11

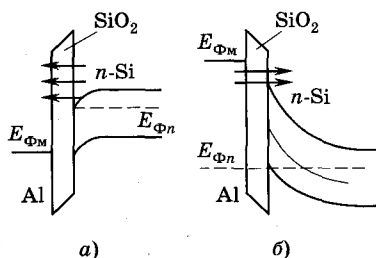


Рис. 11.12

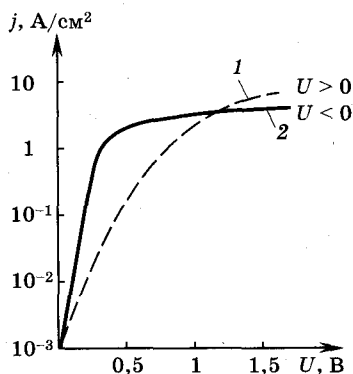


Рис. 11.13

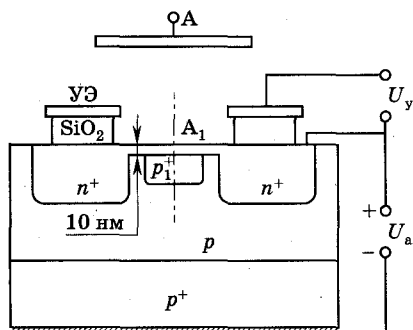


Рис. 11.14

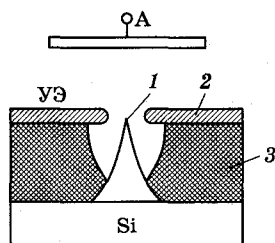


Рис. 11.15

Катоды на основе обратнсмещенных $p-n$ -переходов относятся к самым эффективным и хорошо изученным. Пример устройства таких катодов представлен на рис. 11.14.

Между тонкой n^+ -областью и p_1^+ , p^+ -областями создается обратное смещение. Рабочий (эмитирующий) $p-n$ -переход формируется между областями p_1^+ и n^+ . Этот переход смещается в область лавинного пробоя. Возникающая лавина локализуется практически на поверхности A_1 -структуры. Образующаяся в результате пробоя электронно-дырочная плазма и является источником эмиссии электронов с поверхности A_1 (см. рис. 11.14). Под действием поля, формируемого напряжениями U_y и U_a , приложенными к управляющему электроду $UЭ$ и аноду A (коллектору), электроны перемещаются к аноду (коллектору). Одним из существенных достоинств таких холодных кремниевых катодов является то, что форму активной эмитирующей электроны поверхности можно выбирать.

Эффективность эмиссии и электронная температура зависят от степени легирования областей структуры (n^+ и p_1^+), от пространственного распределения (профиля) легирующей примеси и очень сильно — от условий на поверхности A_1 . Качественно эти зависимости можно объяснить следующим образом.

Степень легирования и профиль легирования определяют величину и пространственное распределение электрического поля внутри обедненного слоя $p^+ - n^+$ -перехода и, следовательно, скорость и функцию распределения электронов в переходе. Часть этих электронов (наиболее быстрые, горячие) эмитируется в вакуум. Для уменьшения работы выхода электронов с поверхности и, следова-

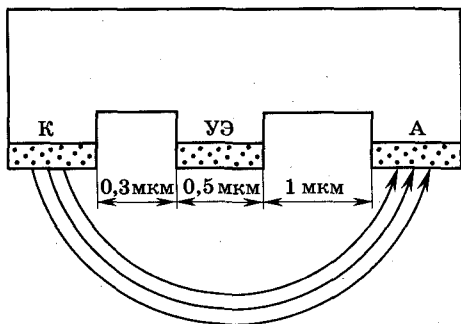


Рис. 11.16

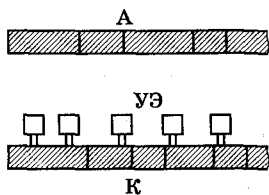


Рис. 11.17

тельно, увеличения тока эмиссии поверхность A_1 покрывают це-
зием. Такими методами получена плотность тока до 8000 А/см^2 .

Были разработаны и исследованы также катоды на основе переходов Шоттки, $p-i-n$ -диодов. Они имеют структуры, по-
добные рассмотренным для $p-n$ -перехода.

Примеры устройств вакуумных микротриодов плоской и вер-
тикальной конструкции приведены на рис. 11.15—11.17. На
рис. 11.15 катод 1 изготовлен в виде острия из кремния. Между
массивной частью катода и управляющим электродом (сеткой) 2,
выполненным из металлической пленки, располагается диэлек-
трический слой SiO_2 (3). Структура на рис. 11.15 выполнена на
плоскости (поверхности) диэлектрической подложки и накрыв-
вается диэлектрическим пустотелым колпачком. Аналогично
функционируют и структуры, показанные на рис. 11.16 и 11.17.
Однако в отличие от плоских микротриодов (см. рис. 11.16,
11.17) прибор на рис. 11.15 может быть сделан и в вертикальном
исполнении. В этом случае анод впаивается в диэлектрический
колпачок, через который делаются выводы других электродов.



Контрольные вопросы

1. Что такое работа выхода и какой физический смысл имеет это понятие?
2. Какие физические механизмы имеют место при различных видах эмиссии электронов?
3. Физические процессы в диодах. Анодные характеристики и параметры диодов.

4. Устройство и физические процессы в триодах.
5. Объяснить анодные и анодно-сеточные характеристики триодов. Параметры триодов.
6. Устройства, физические процессы, параметры и характеристики тетродов и пентодов.
7. Мощные модуляторные и генераторные лампы, их особенности, параметры и характеристики.
8. Эквивалентные схемы электровакуумных приборов.
9. Вакуумные интегральные схемы, их особенности, разновидности и устройство.
10. Каковы особенности катодов интегральных схем?

Глава 12

ЭЛЕКТРОННО-ЛУЧЕВЫЕ ПРИБОРЫ

12.1. Классификация, устройство и принцип действия электронно-лучевых приборов

Электронно-лучевыми приборами называются такие электровакуумные приборы, в которых используются сформированные в виде лучей потоки электронов. Различают одно-, двух- и многолучевые приборы. Приборы, у которых форма баллона выполнена наподобие трубки, называют *электронно-лучевыми трубками (ЭЛТ)*.

Основным классификационным признаком для электронно-лучевых приборов является назначение приборов. Различают следующие виды электронно-лучевых приборов: *приемные ЭЛТ, передающие трубки, запоминающие трубки, электронно-оптические преобразователи*.

Электронно-лучевые приборы являются приборами *отображения и преобразования информации*, представляемой в виде электрических или световых сигналов.

В приемных трубках, к которым относятся индикаторные трубки радиолокационных станций, осциллографические трубки, телевизионные трубки (кинескопы), трубки дисплеев, обеспечивающие вывод информации из ЭВМ, последовательности электрических сигналов преобразуются в определенную форму

видимого изображения. В передающих трубках, наоборот, оптическое изображение преобразуется в последовательность электрических сигналов.

В запоминающих трубках возможно как преобразование электрических сигналов в изображение и изображения в последовательность электрических сигналов, так и то и другое преобразования.

Для формирования и управления электронными потоками в ЭЛТ могут использоваться или только одни электрические поля (ЭЛТ с электростатическим управлением), или же электрические и магнитные поля (ЭЛТ с магнитным управлением).

В конструкциях большинства видов электронно-лучевых приборов имеются следующие основные элементы: **электронный прожектор 1**, **отклоняющая система 2**, **экран**, или **мишень 3** (рис. 12.1).

В трубках с магнитным управлением используются также магнитные фокусирующие катушки 4 (рис. 12.1, б).

Электронный прожектор, который часто называют **электронной пушкой**, создает и фокусирует электронный пучок (поток) в электронный луч. Он состоит из катода и, как правило, из нескольких электродов, формирующих электронный луч с заданными (необходимыми) энергетическими и геометрическими характеристиками. Отклоняющая система сканирует (перемещает) луч, сформированный прожектором, по заданному закону в пространстве. Отклонение осуществляется или с помощью взаимно перпендикулярных двух пар пластин определенной формы (**электростатическое отклонение**, см. рис. 12.1, а), или с помощью ортогонально расположенных двух пар магнитных катушек (**магнитное отклонение**, см. рис. 12.1, б). При магнитном управлении используются не только фокусирующие катушки, но и магнитные отклоняющие системы. На экране (мишени) происходит преобразование кинетической энергии электронов в оптическое излучение или в энергию электрического поля, формирующих соответствующий вид изображения.

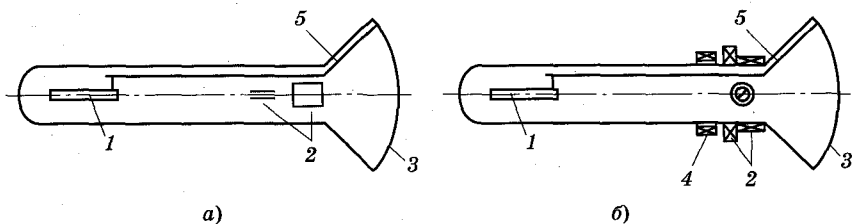


Рис. 12.1

Принцип работы большинства электронно-лучевых трубок сводится к следующему. Электроны луча, сфокусированные и ускоренные прожектором, поступают в отклоняющую систему, где под действием электрических и магнитных полей осуществляется пространственное перемещение луча.

После отклоняющей системы электронный луч попадает на экран, представляющий собой конструктивный элемент ЭЛТ со слоем люминофора, нанесенного на дно баллона трубки. При отсутствии отклонения электронный луч формирует в центре экрана маленькое яркое пятно. Для отвода вторичных электронов, выбиваемых лучом с поверхности экрана, на внутреннюю поверхность трубки наносится проводящий слой 5 (см. рис. 12.1, б), соединенный с прожектором. В трубках с магнитным управлением для фокусировки используется магнитная фокусирующая катушка.

12.2. Электронный прожектор с электростатической фокусировкой

В электронном прожекторе, состоящем обычно из катода и нескольких электродов, происходит управление величиной тока и плотностью электронного потока (фокусировка).

Управление плотностью (интенсивностью) электронного потока позволяет менять яркость пятна на экране ЭЛТ. Управление плотностью тока осуществляется за счет изменения характеристик электрического поля в основном в межэлектродных областях. Катод в ЭЛТ обычно выполняется в форме небольшого цилиндра 2, внутри которого помещен *подогреватель* 1 (рис. 12.2). Эмитирующей частью является дно цилиндра, покрытое оксидным слоем. Катод располагается внутри другого цилиндра с от-

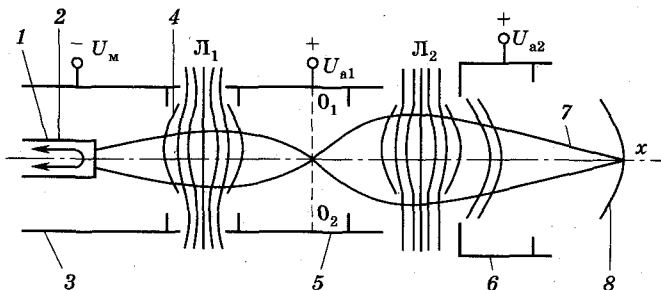


Рис. 12.2

вертием, являющегося *управляющим электродом (модулятором)* 3. Основное назначение модулятора — изменение тока электронного луча 7, и его действие подобно действию управляющей сетки в триоде. На модулятор подается небольшой отрицательный относительно катода потенциал U_m . В пространстве между катодом и модулятором формируется неоднородное электрическое поле 4, изменяющее объемный заряд около катода. Зависимость катодного тока I_k от потенциала модулятора, которая называется *модуляционной характеристикой* $I_k = f(U_m)$, в ЭЛТ несколько отличается от закона трех вторых и может быть выражена следующим соотношением:

$$I_k = b \left(\frac{U_m - U_{m0}}{U_{m0}} \right)^\gamma |U_{m0}|^{3/2}, \quad (12.1)$$

где U_m , U_{m0} — напряжение на модуляторе и запирающее напряжение в В; I_k — ток катода в мкА; b и γ — коэффициенты пропорциональности: $b = 2, 3 \dots 3$, $\gamma = 2, 5 \dots 3, 5$.

Отличие формулы (12.1) от закона трех вторых обусловлено тем, что ток катода зависит не только от потенциала модулятора, но и от величины поверхности катода, охваченной полем с положительным градиентом потенциала. Кроме того, в трубках с электростатическим управлением при малых отрицательных потенциалах на модуляторе, и, следовательно, при большой эмитирующей поверхности катода увеличивается диаметр луча и часть электронов не пропускается диафрагмой. В трубках с магнитной фокусировкой ограничивающие диафрагмы отсутствуют и ток луча близок к катодному току. По аналогии с триодом для напряжения запирающего модулятора U_{m0} можно записать:

$$U_{m0} \approx -D_m U_a, S_m = \frac{dI_m}{dU_m}, \quad (12.2)$$

где D_m , S_m — соответственно проницаемость модулятора и крутизна модуляционной характеристики $I_k = f(U_m)$ при постоянном напряжении 2-го анода $U_a = \text{const}$ и $U_a = U_{y\phi}$ или $U_a = U_{a1}$ (напряжение ускоряющего электрода или первого анода 5 (см. рис. 12.2), т. е. электрода, который расположен сразу за модулятором и имеет положительный потенциал). Как уже отмечалось, электронный прожектор содержит электроды, с помощью которых осуществляется не только управление плотностью электрон-

ного луча, но и его фокусировка. К таким электродам относятся ускоряющие электроды и аноды — обычно один или два. Принципы работы электронных прожекторов во многом аналогичны принципам действия оптических фокусирующих систем, поэтому раздел электроники, изучающий фокусировку электронных потоков и законы их движения, называется *электронной оптикой*. Неоднородные аксиально-симметричные электрические поля, формируемые системами электродов и используемые для изменения скорости и направления движения электронов, называются *электронными линзами*. Однородные и неоднородные магнитные поля могут быть также использованы в качестве линз.

Триодный прожектор. Большинство прожекторов современных электронно-лучевых приборов строят по двухлинзовой оптической схеме. Использование двухлинзового прожектора с первой сильной (короткофокусной) линзой и второй слабой (длиннофокусной) позволяет получать в плоскости экрана сечение электронного луча с диаметром порядка 0,1 мм при диаметре эмитирующей поверхности катода около 1 мм. Примерный вид распределений неоднородных электрических полей, образующих две линзы, и траектории электронов в триодном электронном прожекторе приведены на рис. 12.2.

Первая линза L_1 в этом прожекторе образована катодом 2, модулятором 3 и первым анодом 5, а вторая линза L_2 формируется между первым (5) и вторым (6) анодами. Следовательно, первая и вторая линзы прожектора, кроме фокусировки, еще и ускоряют электроны. В области L_1 электрическое поле с положительным градиентом доходит до поверхности катода и вытягивает электроны, испускаемые катодом. Поскольку первая линза вытягивает электроны из катодной области и ускоряет их, она должна быть только электростатической. Вторая линза прожектора может быть как электростатической, так и магнитной. В рассматриваемом случае обе линзы можно представить как совокупность собирающей (двоковыпуклой) и рассеивающей (двоковогнутой) элементарных линз. Собирающая линза образована неоднородным электрическим полем с эквипотенциальными поверхностями, обращенными выпуклостью к катоду, а рассеивающая — эквипотенциальными поверхностями, обращенными в сторону экрана. Преломляющее действие собирающей линзы больше, чем рассеивающей. Из-за большего потенциала в области рассеивающей линзы скорости движения электронов

в ее пространстве больше, чем в собирающей, поэтому фокусирующее (собирающее) действие последней преобладает. Рассеивающая линза увеличивает лишь фокусное расстояние всей системы.

Изменяя потенциал первого анода (5), в цепи которого протекает значительно меньший ток, чем в цепи второго (6), и поэтому изменение потенциала осуществляется с меньшими потерями, можно регулировать фокусное расстояние второй линзы, совмещая тем самым фокус второй линзы с поверхностью экрана.

Мы рассмотрели общие принципы работы *триодного прожектора* ЭЛТ, показанного на рис. 12.2. Теперь проиллюстрируем физический механизм изменения траектории электронов при движении в электрическом поле. Для этого из любой линзы выделим некую элементарную область и изобразим на ней несколько эквипотенциальных линий (рис. 12.3). Допустим, что электрон со скоростью \vec{v} пересекает эквипотенциаль ϕ_1 . Разложим вектор скорости \vec{v} на две составляющие, одна из которых \vec{v}_\perp направлена нормально к эквипотенциали, а другая \vec{v}_\parallel — по касательной, проведенной к эквипотенциали в точке пересечения ее электроном (точка A_1). Пусть потенциал второй эквипотенциали $\phi_2 > \phi_1$. Напряженность электрического поля $\vec{\epsilon}$ направлена по нормали к ϕ_2 в сторону ϕ_1 . Под воздействием электрического поля $\vec{\epsilon}$ составляющая скорости \vec{v}_\perp увеличивается, а \vec{v}_\parallel остается неизменной. В результате после прохождения электроном расстояния между эквипотенциалами ϕ_1 и ϕ_2 траектория электрона отклонится вниз. Аналогично рассуждая, приходим к заключению, что после прохождения между ϕ_2 и ϕ_3 ($\phi_2 < \phi_3$) электрон еще больше отклонится вниз. Если изменить направление градиента электрического поля, т. е. сделать $\phi_1 > \phi_2 > \phi_3$ и т. д., траектория электрона будет изгибаться вверх. Изменяя конфигурацию эквипотенциалей относительно вектора скорости или самонаправление вектора скорости \vec{v} , можно формировать траекторию движения электронов по необходимому закону.

Рассмотренный триодный прожектор обладает рядом недостатков. Основной из них — взаимное влияние процесса изменения плотности катодного тока, от которого зависит яркость свечения экрана, и качества фокусировки (диаметра пятна на экране).

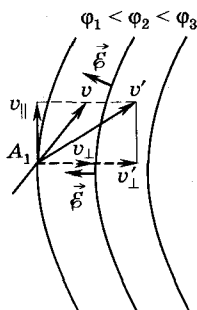
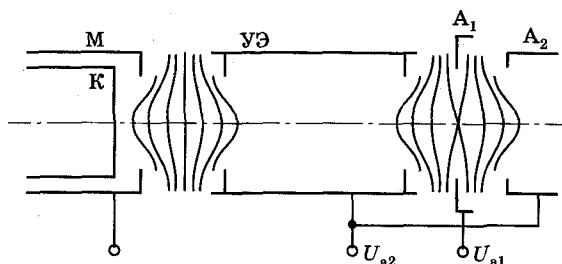


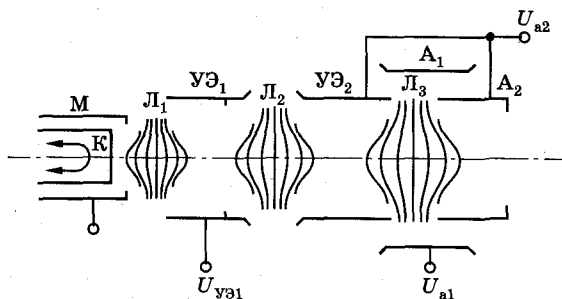
Рис. 12.3

Изменяя потенциал первого анода, чтобы сфокусировать электронный поток оптимальным образом, мы тем самым будем менять характеристики и конфигурацию электрического поля у поверхности катода и, следовательно, ток луча, т. е. яркость. С другой стороны, при изменении напряжения на модуляторе для регулировки яркости будет нарушаться фокусировка. Для устранения этих нежелательных явлений используют триодные прожекторы с нулевым током первого анода, которые применяются в большинстве современных приемных трубок.

В этом типе прожектора (рис. 12.4, а) между первым анодом A_1 и модулятором М располагается ускоряющий электрод УЭ в виде длинного цилиндра с диафрагмами, ограничивающими поперечные размеры электронного пучка. Первый анод выполнен в виде диафрагмы с отверстием в 2—3 раза больше диаметра пучка, что практически исключает попадание электронов на этот электрод. При такой конструкции A_1 и УЭ ток в цепи первого анода равен нулю, что позволяет питать электроды прожектора от общего делителя напряжения. Из-за отсутствия то-



а)



б)

Рис. 12.4

ка в цепи A_1 изменение напряжения на нем при фокусировке не влияет на распределение тока и напряжения на элементах делителя, т. е. не будет расстройки первой линзы L_1 ; аналогично и регулировка яркости не исказит фокусировку. Независимость действующего в плоскости модулятора потенциала от напряжения первого анода обусловлена малой проницаемостью УЭ с ограничивающими диафрагмами, которые являются надежным электростатическим экраном, препятствующим проникновению поля первого анода в прикатодную область.

Кроме перечисленных, триодным прожекторам присущи следующие недостатки: относительно большие углы расхождения электронов за плоскостью скрещения (плоскость O_1O_2 на рис. 12.2), что приводит к заметному ограничению тока пучка диафрагмой, установленной в ускоряющем электроде или первом аноде; большие значения запирающих напряжений на модуляторе при использовании высоких ускоряющих напряжений. Этих недостатков в значительной мере лишены электронные *прожекторы тетродного типа*. Схематическое устройство такого типа прожектора изображено на рис. 12.4, б.

Тетродный прожектор. Такой прожектор является трехлинзовой электронно-оптической системой: первая линза L_1 формируется катодом (К), модулятором (М) и первым ускоряющим электродом УЭ₁; вторая, слабая, линза L_2 заключена между двумя ускоряющими электродами УЭ₁ и УЭ₂; главная линза образована вторым ускоряющим электродом УЭ₂, первым A_1 и вторым A_2 анодами. Первый ускоряющий электрод обычно выполняется в виде короткого цилиндра с ограничивающей диафрагмой, которая уменьшает проницаемость этого электрода, а это, в свою очередь, уменьшает запирающее напряжение.

На первый УЭ₁ подают относительно невысокий потенциал (несколько сот вольт), а на второй — высокий положительный потенциал (до 15...20 кВ). Линза, формируемая между двумя ускоряющими электродами, уменьшает угол расходимости пучка после плоскости скрещения.

Применение той или иной конструкции электростатического электронного прожектора обуславливается функциональным назначением ЭЛТ, требованиями к качеству фокусировки, яркости свечения, допустимыми искажениями и т. п.

12.3. Электронный прожектор с магнитной фокусировкой

В ЭЛТ с магнитной фокусировкой в качестве второй (главной) проекционной линзы, отображающей скрещении траекторий электронов на экран, используется неоднородное магнитное поле (*магнитная линза*) короткой катушки, у которой диаметр соизмерим с ее толщиной (длиной). Катушка заключена в ферромагнитный панцирь и надевается на горловину трубки.

В трубках с магнитной фокусировкой используют как триодный, так и тетродный прожекторы. Тетродный прожектор применяют, когда необходимо получить высокую разрешающую способность и малые искажения при отклонении. В этом случае электроны после скрещенния имеют меньший угол расхождения по сравнению с триодным типом прожектора, что приводит к уменьшению сечения луча в области главной проекционной магнитной линзы и отклоняющей системы.

Неоднородные магнитные поля, используемые в электронно-лучевых приборах, создаются катушками, через которые протекает ток. Для получения нужной структуры магнитного поля используются полюсные наконечники и магнитные экраны, изготовляемые из материалов с большой магнитной проницаемостью. Применение таких материалов позволяет добиться нужной формы магнитного поля.

Магнитные поля, применяемые в большинстве ЭЛТ, являются аксиально-симметричными. Такие поля создаются круговыми катушками с ферромагнитными оболочками.

Принцип фокусировки электронного луча неоднородным магнитным полем короткой катушки можно проиллюстрировать с использованием рис. 12.5.

В общем случае вектор скорости электрона \vec{v} направлен под некоторым углом α к оси катушки. Если электрон находится в точке A (рис. 12.5, *a*), то вектор индукции магнитного поля \vec{B} и вектор скорости \vec{v} имеют как радиальные B_r и v_r , так и осевые (аксиальные) B_a и v_a составляющие. Как известно из курса физики, на электрон, движущийся под произвольным углом по отношению к направлению магнитного поля, действует сила Лоренца $\vec{F}_m = -e[\vec{v} \times \vec{B}]$. Составляющая силы Лоренца \vec{F}_{mt} (рис. 12.5, *b*), обусловленная скоростью электронов v_a и вектором B_r , направлена из-за плоскости чертежа на читателя. Векторное произведение $[\vec{v} \times \vec{B}]$ представляет собой правую тройку векторов, но при определении направления силы \vec{F} необходимо учитывать отрицательный знак заряда электрона и тогда $\vec{F}_m \sim -[\vec{v} \times \vec{B}]$. Сила

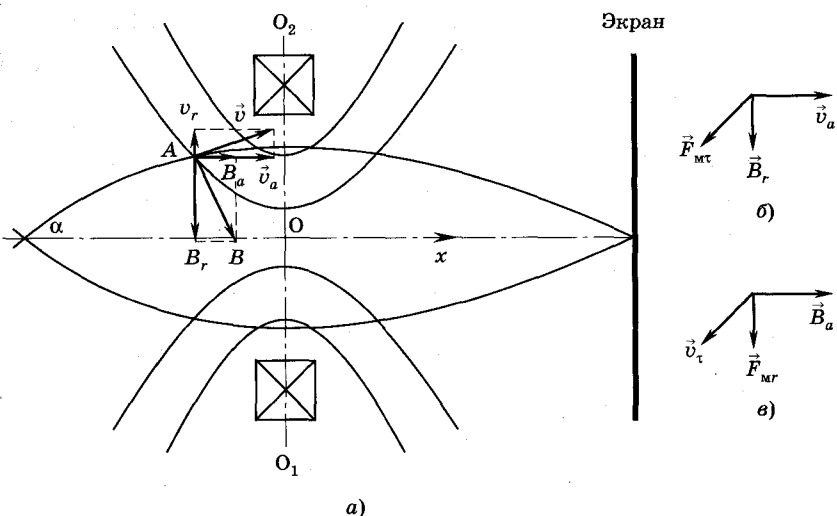


Рис. 12.5

$\vec{F}_{\text{мт}}$ вызывает вращение электронов около оси, т. е. появляется азимутальная составляющая скорости \vec{v}_t , которая совместно с B_a образует силу $\vec{F}_{\text{мг}}$, направленную к оси катушки (рис. 12.5, в).

Описанная ситуация сохраняется до плоскости, проходящей через центральную линию O_1O_2 катушки перпендикулярно ее оси. Таким образом, при движении электрона до этой плоскости, он перемещается по спирали с уменьшающимся радиусом, отклоняясь к оси катушки. После пересечения плоскости O_1O_2 направление радиальной составляющей индукции магнитного поля B_r меняется на противоположное, и, следовательно, сила $F'_{\text{мт}}$ будет направлена навстречу силе $\vec{F}_{\text{м}}$. Однако после пересечения этой плоскости электроны уже отклонились к оси, где значения B_r меньше, и поэтому поперечная сила $F_{\text{м}} \sim v_t B_a$ будет направлена в ту же сторону, что и ранее. Тангенциальная сила $F'_{\text{мт}}$ справа от плоскости O_1O_2 будет уменьшать скорость вращения электрона, который пересечет ось в некоторой точке C . Изменяя индукцию магнитного поля за счет регулирования тока в катушке, можно добиться пересечения траекторий электронов с осью трубки в плоскости экрана (см. рис. 12.5), обеспечивая тем самым фокусировку электронного потока.

В силу того что в магнитных линзах допустимы большие апертурные углы, разрешающая способность трубок с такой фокуси-

ровкой увеличивается, потери электронов внутри линзы отсутствуют, так как нет ограничивающих диафрагм и ток луча близок к току катода, т. е. крутизна модуляционной характеристики существенно больше; токи луча также намного больше (несколько мА), чем в трубках с электростатической фокусировкой (сотни мкА). Однако увеличение сечения пучка, как следствие больших апертурных углов в области главной проекционной линзы, приводит к большим, по сравнению с электростатической фокусировкой, диаметрам луча в области отклоняющей системы, что может вызывать сильные искажения при использовании электростатического отклонения. Поэтому магнитная фокусировка, как правило, требует и магнитных отклоняющих систем.

12.4. Отклоняющие системы

Основное назначение отклоняющих систем состоит в пространственном перемещении сфокусированного электронного луча. Есть два принципиально различных типа отклоняющих систем: электростатическая, в которой отклонение электронного луча осуществляется поперечным (по отношению к вектору скорости электронов) электрическим полем, и магнитная, использующая поперечное магнитное поле.

Отклоняющие системы должны обладать хорошей чувствительностью и малыми искажениями сигналов.

Если требуется перемещать луч последовательно по всей плоскости экрана, то простейшая электростатическая отклоняющая система состоит из двух пар попарно ортогональных пластин, расположенных последовательно вдоль оси трубки. Одна пара пластин отклоняет луч в вертикальном направлении, другая — в горизонтальном. Рассмотрим движение электронов между парой полубесконечных плоских пластин под действием поля \mathcal{E} , создаваемого разностью потенциалов, приложенной к пластинам.

Уравнения движения (см. приложение 2) нерелятивистских электронов в декартовой системе координат (рис. 12.6) при наличии только поперечного электрического поля $\mathcal{E} = -\mathcal{E}_z$ ($\mathcal{E}_x = \mathcal{E}_y = 0$) имеют вид

$$m \frac{d^2x}{dt^2} = 0; \quad (12.3)$$

$$m \frac{d^2y}{dt^2} = 0; \quad (12.4)$$

$$m \frac{d^2z}{dt^2} = e\mathcal{E}_z. \quad (12.5)$$

Допустим, что электрон влетает в пространство между пластинами ($x_0 = 0$ и $z_0 = 0$) в направлении оси x с начальной скоростью v_{x0} ($v_y = 0$, $v_{z0} = 0$). Интегрируя уравнения (12.3—12.5), получаем следующий результат:

$$x = v_{x0}t; v_z = \frac{e}{m} \mathcal{E}_z t; z = \frac{e \mathcal{E}_z}{2m} t^2. \quad (12.6)$$

Для получения уравнения траектории выражаем из первого уравнения (12.6) время t и подставляем этот результат в формулу для z :

$$z = e \mathcal{E}_z x^2 / (2m v_{x0}^2). \quad (12.7)$$

Таким образом, согласно (12.7) траекторией движения электрона в однородном электростатическом поле плоских пластин является парабола. Электроны на выходе из пластин длиной l_1 отклоняются на величину z_1 (см. рис. 12.6):

$$z_1 = e \mathcal{E} l_1^2 / (2m v_{x0}^2) = U_{\text{пл}} l_1^2 / (4d U_{a2}), \quad (12.8)$$

где $\mathcal{E} = \mathcal{E}_z = U_{\text{пл}}/d$; $U_{\text{пл}}$, d — соответственно разность потенциалов и расстояние между пластинами, $v_{x0} = \sqrt{\frac{2e}{m} U_{a2}}$, U_{a2} — напряжение на втором аноде, определяющее скорость электронов v_{x0} на входе в пластины.

Угол отклонения α направления движения электрона от первоначального на выходе из пластин определяется дифференцированием выражения (12.7) по x (при $x = l_1$):

$$\operatorname{tg} \alpha = \frac{dz}{dx} = \frac{e \mathcal{E}}{m v_{x0}^2} x \Big|_{x=l_1} = \frac{U_{\text{пл}} l_1}{2d U_a}.$$

Полное отклонение h_3 луча на экране (см. рис. 12.6) определяется формулами

$$z_1 = \frac{U_{\text{пл}} l_1^2}{4d U_{a2}}; z_2 = l_2 \operatorname{tg} \alpha = \frac{U_{\text{пл}} l_1 l_2}{2d U_{a2}};$$

$$h_3 = z_1 + z_2 = \frac{U_{\text{пл}} l_1}{2d U_{a2}} (l_1 + l_2) = \frac{U_{\text{пл}} l_1}{2d U_{a2}} \left(L + \frac{l_1}{2} \right).$$

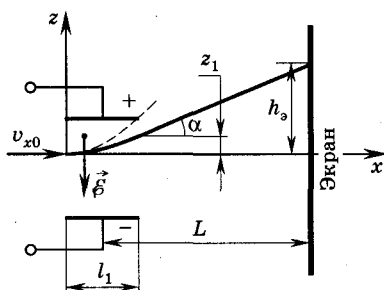


Рис. 12.6

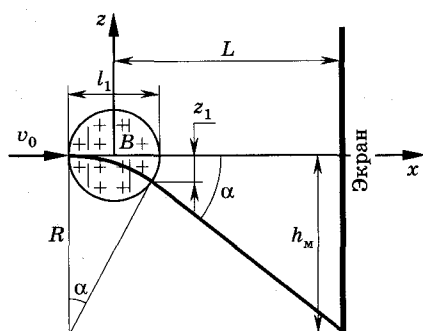


Рис. 12.7

которая обычно содержит две пары катушек, надеваемых на горловину трубки и образующих магнитные поля во взаимно перпендикулярных направлениях. Аналогично электростатической отклоняющей системе одна из пар катушек отклоняет луч в вертикальной, а другая — в горизонтальной плоскости. Рассмотрим отклонение электрона однородным магнитным полем одной пары катушек (рис. 12.7). В однородном магнитном поле ($B = B_y$, $B_x = B_z = 0$, при полном отсутствии электрического) уравнения движения электрона в декартовой системе координат имеют вид

$$m \frac{d^2x}{dt^2} = -e(v_z B_y); \quad m \frac{d^2y}{dt^2} = 0; \quad m \frac{d^2z}{dt^2} = -e(v_x B_y). \quad (12.10)$$

Система уравнений (12.10) справедлива для вычисления траектории электронов как в магнитных отклоняющих, так и в фокусирующих системах. Сила Лоренца, действующая на электрон, перпендикулярна вектору его скорости \vec{v} , поэтому величина полной скорости электрона $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ остается постоянной, а направление его движения непрерывно изменяется.

Если принять, что при $t = 0$ $x_0 = 0$, $z_0 = 0$, $v_{z0} = 0$, $v_{x0} = v_0$, то решение системы (12.10) можно записать в виде

$$z = \frac{mv_0}{eB} [1 - \cos(\omega_{ц} t)]; \quad x = \frac{mv_0}{eB} \sin(\omega_{ц} t), \quad (12.11)$$

где $\omega_{ц} = \frac{e}{m} B$ — электронная циклотронная частота.

Уравнения (12.11) описывают траекторию, которая представляет собой окружность радиуса $R = \frac{mv_0}{eB}$, следовательно, в одно-

С учетом того, что $L \gg l_1$, последнее соотношение можно упростить:

$$h_{\alpha} = \frac{U_{пл}}{2dU_{a2}} l_1 L. \quad (12.9)$$

Здесь, как и в дальнейшем, принято, что $U_{a2} = U_a$.

Перейдем теперь к вычислению траектории электронов при движении их в магнитной отклоняющей системе, кото-

родном магнитном поле электрон будет вращаться по окружности радиуса R с частотой ω_d . Поскольку диаметр l_1 отклоняющих катушек, как правило, существенно меньше R , на выходе из них отклонение электрона от оси трубки можно представить в виде приближенного соотношения $z_1 = l_1 \operatorname{tg} \alpha$ (см. рис. 12.7). Далее электрон будет двигаться по касательной к его криволинейной траектории в точке выхода из поля. При малых углах $\operatorname{tg} \alpha \approx \alpha$ и тогда $\alpha \approx \frac{z_1}{l_1}$, и если принять, что путь, пройденный внутри катушек, незначительно отличается от их диаметра, то

$$\operatorname{tg} \alpha \approx \alpha \approx \frac{l_1}{R} = \left(\frac{mv_0}{eB} \right)^{-1} l_1 = \sqrt{\frac{e}{2m}} \frac{Bl_1}{\sqrt{U_a}}, \quad (12.12)$$

где U_a — потенциал последнего электрода перед отклоняющей системой.

Угол α можно вычислить из формулы (12.12), если учесть, что индукция магнитного поля B пропорциональна числу ампер-витков катушки, т. е. $B = k_1 w I$ (k_1 — коэффициент пропорциональности, определяемый конструкцией катушки; w — число витков; I — ток, протекающий по виткам).

Полное смещение h электрона на плоском экране ЭЛТ, отстоящем на расстоянии L от центра отклоняющих систем (см. рис. 12.6 и 12.7), приближенно равно $h \approx L \operatorname{tg} \alpha$. Используя формулы (12.9 и 12.12), запишем смещения луча для электростатической $h_э$ и магнитной $h_м$ систем отклонения:

$$h_э = \frac{U_{пл} l_1 L}{2dU_a} \quad \text{и} \quad h_м = \sqrt{\frac{e}{2m}} \frac{L l_1}{\sqrt{U_a}} k_1 w I. \quad (12.13)$$

Отношения $\frac{h_э}{U_{пл}} = \varepsilon_э$ и $\frac{h_м}{I} = \varepsilon_м$ называются чувствительностью к отклонению соответственно для электростатической и магнитной систем. Из (12.13) получаем, что

$$\varepsilon_э = \frac{l_1 L}{2dU_a}; \quad \varepsilon_м = \sqrt{\frac{e}{2m}} \frac{l_1 L}{\sqrt{U_a}} k_1 w. \quad (12.14)$$

Параметры $\varepsilon_э$ и $\varepsilon_м$ соответственно характеризуют отклонение луча при разности потенциалов между отклоняющими пластинами в 1 В или при токе в 1 А, протекающем через катушки. Значе-

ние ϵ_0 измеряется в мм/В, а ϵ_m — в мм/А. Из (12.14) видно, что при магнитном отклонении изменение ускоряющего напряжения U_a существенно меньше влияет на чувствительность, чем при электростатическом, так как $\epsilon_m \sim \frac{1}{\sqrt{U_a}}$, а $\epsilon_0 \sim \frac{1}{U_a}$.

К достоинствам магнитного отклонения можно отнести меньшие aberrации, внешнее относительно ЭЛТ расположение катушек, что позволяет применять отклоняющие системы, вращающиеся вокруг оси трубки. Среди существенных недостатков магнитных отклоняющих систем — существенно большие потребляемые мощности, а также большая инерционность вследствие значительных собственных емкостей и индуктивностей. Электростатические отклоняющие системы из-за малых значений паразитных емкостей и малых времен пролета электронов могут работать на частотах до нескольких сотен МГц, а магнитные — только до нескольких десятков кГц.

Электростатические отклоняющие пластины располагаются в пространстве последовательно друг за другом. Совмещение вертикальных и горизонтальных отклоняющих систем нерационально из-за их сильного взаимного влияния, приводящего к большим искажениям при отклонении луча. Вертикальные и горизонтальные магнитные отклоняющие катушки обычно совмещают в пространстве, так как их взаимное влияние нетрудно исключить точной установкой и изготовлением.

В электростатических системах для увеличения максимального угла отклонения (чувствительности) и уменьшения искажений при отклонении луча применяют косо расставленные, изломанные и изогнутые пластины. Наилучшие результаты получаются в случае применения изогнутых пластин, при которых траектория электронов луча параллельна поверхности пластин отклоняющей системы.

12.5. Экраны электронно-лучевых трубок

После отклоняющей системы электроны попадают на экран ЭЛТ. Экран представляет тонкий слой люминофора, нанесенного на внутреннюю поверхность торцевой части баллона и способного интенсивно светиться при бомбардировке электронами.

В ряде случаев поверх слоя люминофора наносится проводящий тонкий слой алюминия. Свойства экрана определяются его

характеристиками и параметрами. К основным параметрам экранов относятся: *первый и второй критические потенциалы экрана, яркость свечения, световая отдача, длительность послесвечения.*

Потенциал экрана. При бомбардировке экрана потоком электронов с его поверхности возникает вторичная электронная эмиссия. Для отвода вторичных электронов стенки баллона трубки вблизи экрана покрываются проводящим графитовым слоем, который соединяется со вторым анодом. Если этого не делать, то вторичные электроны, возвращаясь на экран, вместе с первичными будут понижать его потенциал. В этом случае в пространстве между экраном и вторым анодом создается тормозящее электрическое поле, которое будет отражать электроны луча. Таким образом, для устранения тормозящего поля от поверхности непроводящего экрана необходимо отводить электрический заряд, переносимый электронным лучом. Практически единственным путем компенсации заряда является использование вторичной эмиссии. При падении электронов на экран их кинетическая энергия преобразуется в энергию свечения экрана, идет на его нагрев и вызывает вторичную эмиссию. Величина коэффициента вторичной эмиссии σ определяет потенциал экрана. Коэффициент вторичной эмиссии электронов $\sigma = I_{\text{в}}/I_{\text{л}}$ ($I_{\text{в}}$ — ток вторичных электронов, $I_{\text{л}}$ — ток луча, или ток первичных электронов) с поверхности экрана в широком диапазоне изменения энергии первичных электронов превышает единицу (рис. 12.8, $\sigma > 1$ на участке OA кривой при $U \leq U_{\text{кр}1}$ и при $U > U_{\text{кр}2}$).

При $U \leq U_{\text{кр}1}$ число уходящих от экрана вторичных электронов меньше числа первичных, что приводит к накоплению отрицательного заряда на экране, формированию тормозящего поля для электронов луча в пространстве между вторым анодом и экраном и их отражению; свечение экрана отсутствует. Потенциал $U_{\text{а}2} = U_{\text{кр}1}$, соответствующий точке A на рис. 12.8, называется *первым критическим потенциалом*. При $U_{\text{а}2} = U_{\text{кр}1}$ потенциал экрана близок к нулю.

Если энергия пучка становится больше $eU_{\text{кр}1}$, то $\sigma > 1$ и экран начинает заряжаться поло-

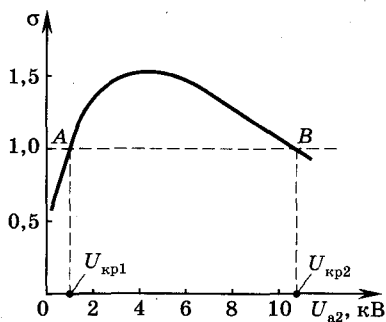


Рис. 12.8

жительно относительно последнего анода прожектора. Процесс продолжается до тех пор, пока потенциал экрана не станет приблизительно равным потенциалу второго анода. Это означает, что число уходящих с экрана электронов равно числу падающих. В диапазоне изменения энергии пучка от $eU_{кр1}$ до $U_{кр2}$ $\sigma \geq 1$ и потенциал экрана достаточно близок к потенциалу анода прожектора. При $U_{a2} \geq U_{кр2}$ коэффициент вторичной эмиссии $\sigma \leq 1$. Потенциал экрана вновь снижается, и у экрана начинает формироваться тормозящее для электронов луча поле. Потенциал $U_{кр2}$ (соответствует точке *B* на рис. 12.8) называют **вторым критическим потенциалом** или **предельным потенциалом**.

При энергиях электронного луча выше $eU_{кр2}$ яркость свечения экрана не увеличивается. Для различных экранов $U_{кр1} = 300 \dots 500$ В, $U_{кр2} = 5 \dots 40$ кВ.

При необходимости получения больших яркостей потенциал экрана с помощью проводящего покрытия принудительно поддерживают равным потенциалу последнего электрода прожектора. Проводящее покрытие электрически соединено с этим электродом.

Светоотдача. Это параметр, который определяет отношение силы света $J_{св}$, излучаемого люминофором нормально поверхности экрана, к мощности электронного луча $P_{эл}$, падающего на экран:

$$\eta = J_{св} / P_{эл}. \quad (12.15)$$

Светоотдача η определяет КПД люминофора. Не вся кинетическая энергия первичных электронов превращается в энергию видимого излучения, часть идет на нагревание экрана, вторичную эмиссию электронов и на излучение в инфракрасном и ультрафиолетовом диапазонах спектра. Светоотдачу измеряют в канделах на ватт: для различных экранов она изменяется в пределах 0,1...15 кд/Вт. При малых скоростях электронов свечение возникает в поверхностном слое и часть света поглощается люминофором. С увеличением энергии электронов светоотдача возрастает. Однако при очень больших скоростях многие электроны пробивают слой люминофора, не производя возбуждения, и происходит снижение светоотдачи.

Яркость свечения. Это параметр, который определяется силой света, излучаемого в направлении наблюдателя одним квадратным метром равномерно светящейся поверхности. Яр-

кость измеряют в кд/м². Она зависит от свойств люминофора (характеризуется коэффициентом A), плотности тока электронного луча j , разности потенциалов между катодом и экраном U и минимального потенциала экрана U_0 , при котором еще наблюдается люминесценция экрана. Яркость свечения подчиняется закону

$$B = Aj(U - U_0)^n. \quad (12.16)$$

Значения показателя степени n , потенциала U_0 для разных люминофоров изменяются в пределах соответственно 1...2,5 и 30...300 В. На практике линейный характер зависимости яркости от плотности тока j сохраняется примерно до 100 мкА/см². При больших плотностях тока люминофор начинает нагреваться и выгорать. Основным способом повышения яркости — увеличение U .

Разрешающая способность. Этот важный параметр определяется как свойство ЭЛТ воспроизводить детали изображения. Разрешающая способность оценивается числом отдельно различимых светящихся точек или линий (строк), приходящихся соответственно на 1 см² поверхности или на 1 см высоты экрана, либо на всю высоту рабочей поверхности экрана. Следовательно, для увеличения разрешающей способности необходимо уменьшать диаметр луча, т. е. требуется хорошо сфокусированный тонкий луч диаметром в десятые доли мм. Разрешающая способность тем выше, чем меньше ток луча и больше ускоряющее напряжение. В этом случае реализуется наилучшая фокусировка. Разрешающая способность также зависит от качества люминофора (крупные зерна люминофора рассеивают свет) и наличия ореолов, возникающих из-за полного внутреннего отражения в стеклянной части экрана.

Длительность послесвечения. Время, в течение которого яркость свечения уменьшается до 1% от максимального значения, называется временем послесвечения экрана. Все экраны разделяются на экраны с очень коротким (менее 10⁻⁵ с), коротким (10⁻⁵...10⁻² с), средним (10⁻²...10⁻¹ с), длительным (10⁻¹...16 с) и очень длительным (более 16 с) послесвечением. Трубки с коротким и очень коротким послесвечением широко применяются при осциллографировании, а со средним послесвечением — в телевидении. В радиолокационных индикаторах обычно используются трубки с длительным послесвечением.

В радиолокационных трубках часто применяют длительно светящиеся экраны, имеющие двухслойное покрытие. Первый слой люминофора — с коротким послесвечением синего цвета — возбуждается электронным лучом, а второй — с желтым цветом свечения и длительным послесвечением — возбуждается светом первого слоя. В таких экранах удается получить послесвечение до нескольких минут.

Типы экранов. Очень большое значение имеет цвет свечения люминофора. В осциллографической технике при визуальном наблюдении экрана используются ЭЛТ с зеленым свечением, наименее утомительным для глаза. Таким цветом свечения обладает ортосиликат цинка, активированный марганцем (виллемит). Для фотографирования предпочтительны экраны с синим цветом свечения, свойственным вольфрамату кальция. В приемных телевизионных трубках с черно-белым изображением стараются получить белый цвет, для чего применяются люминофоры из двух компонентов: синего и желтого.

Для изготовления покрытий экранов широко применяют также следующие люминофоры: сульфиды цинка и кадмия, силикаты цинка и магния, окислы и окисульфиды редкоземельных элементов. Люминофоры на основе редкоземельных элементов обладают целым рядом достоинств: они более стойки к различным воздействиям, чем сульфидные, достаточно эффективны, имеют более узкую спектральную полосу излучения, что особенно важно в производстве цветных кинескопов, где необходима высокая чистота цвета и т. д. В качестве примера можно привести сравнительно широко используемый люминофор на основе окисла иттрия, активированного европием $Y_2O_3:Eu$. Этот люминофор имеет узкую полосу излучения в красной области спектра. Хорошими характеристиками обладает также люминофор, состоящий из окисульфида иттрия с примесью европия $Y_2O_3S:Eu$, который имеет максимум интенсивности излучения в красно-оранжевой области видимого участка спектра и лучшую химическую стойкость, чем $Y_2O_3:Eu$ -люминофор.

Достаточно широкое распространение в электронно-лучевых приборах получили алюминированные экраны, у которых поверхность люминофора с внутренней стороны трубки нанесена тонкая (0,1...0,5 мкм) пленка алюминия, соединенная обычно со вторым анодом. К достоинствам таких экранов можно отнести

следующие: 1) потенциал экрана постоянен и равен U_{a2} , поскольку не зависит от коэффициента вторичной эмиссии, что позволяет работать при ускоряющих напряжениях, больших второго критического потенциала, и получать большую яркость свечения; 2) исключается паразитное засвечивание экрана с внутренней стороны трубки, что повышает контрастность изображения; 3) отражение пленкой излучаемого люминофором света повышает яркость свечения экрана и световую отдачу; 4) увеличивается стойкость экранов к бомбардировке их поверхности отрицательными ионами.

Алюминий химически инертен при взаимодействии с люминофорами экранов, легко наносится на поверхность испарением в вакууме и хорошо отражает свет. К недостаткам алюминированных экранов можно отнести то, что алюминиевая пленка поглощает и рассеивает электроны с энергией меньше 6 кэВ, поэтому в этих случаях светоотдача резко падает. Например, светоотдача алюминированного экрана при энергии электронов в 10 кэВ примерно на 60% больше, чем при 5 кэВ. Экраны трубок имеют прямоугольную или круглую форму.

12.6. Особенности электронно-лучевых приборов различного назначения

Электронно-лучевые приборы широко применяются для наблюдения, измерения и исследования различного рода электромагнитных сигналов, передачи и приема изображений в телевидении, записи и хранения сигналов в запоминающих устройствах и многих других областях.

Наиболее распространены *приемные ЭЛТ*, к которым относятся осциллографические, радиолокационные трубки, цветные и черно-белые кинескопы.

Осциллографические трубки. Это одно- или многолучевые электронные приборы, предназначенные для наблюдения или регистрации изменяющихся во времени электрических сигналов. Эти функции выполняются специальным измерительным прибором — *электронным осциллографом*, основным элементом которого является ЭЛТ. Эти приборы могут анализировать сигналы с амплитудой в несколько милливольт, что лежит за пределами чувствительности отклоняющих систем ЭЛТ, поэтому на входе осциллографа применяют усилители с большим коэффи-

циентом усиления. Обычно в осциллографах для наблюдения исследуемого сигнала в зависимости от времени на одну пару отклоняющих пластин или катушек (вертикально отклоняющих) подают напряжение сигнала, а на другую пару (горизонтально отклоняющих), расположенную ближе к экрану, подводится линейно изменяющееся во времени (пилообразное) напряжение, называемое *напряжением развертки*. Под действием линейного пилообразного напряжения развертки электронный луч пробегает с постоянной скоростью по экрану вдоль горизонтальной оси слева направо, затем быстро возвращается в исходную точку (обратный ход). Для удобства наблюдения регистрируемых сигналов след луча при обратном ходе устраняют путем его запираания на время обратного хода подачей отрицательного (гасящего) импульса на модулятор трубки. Осциллографические трубки должны иметь небольшое запирающее напряжение (несколько десятков В), чтобы не использовать гасящие импульсы большой амплитуды. Иначе усложняется схема генератора гасящих импульсов.

Если к вертикально отклоняющим пластинам (катушкам) подвести напряжение исследуемого сигнала, то электронный луч вычертит на экране временную зависимость исследуемого процесса, т. е. одновременно с линейным движением по горизонтали он будет отклоняться по вертикали за счет напряжения сигнала.

Современный электронный осциллограф должен позволять исследовать без искажений сигналы с частотой до ГГц, что можно обеспечить при малых емкостях и индуктивностях отклоняющих систем, малых временах пролета электронов через отклоняющие системы и высокой разрешающей способности трубки (несколько сот строк на рабочей поверхности экрана). При исследовании СВЧ-сигналов для осуществления высокоскоростной записи в осциллографических трубках применяют люминофоры с большой световой отдачей.

Обычно осциллографические трубки содержат двухлинзовый прожектор с электростатической фокусировкой, электростатическую отклоняющую систему и экран с люминофором, дающим зеленое или синее свечение. Для получения высокой яркости и разрешающей способности при относительно небольших ускоряющих напряжениях прожектор должен формировать луч с током до 100...500 мкА.

Для уменьшения взаимного влияния горизонтально отклоняющих и вертикально отклоняющих пластин между ними устанавливают электростатические экраны. Такие же экраны монтируют на входе и выходе отклоняющей системы.

Чтобы получить повышенные яркость свечения экрана и скорость записи, а также улучшить фокусировку часто используют трубки с послеускорением (рис. 12.9). В этих трубках сделаны

дополнительные ускоряющие аноды в виде проводящих колец на широкой части баллона с возрастающими в направлении экрана потенциалами. Поскольку ускорение электронов осуществляется после отклонения луча, увеличение яркости происходит без снижения чувствительности к отклонению. При получении очень большой яркости напряжение послеускорения достигает 20...25 кВ. Чтобы уменьшить искажения в пространстве послеускорения, дополнительные аноды выполняют в виде спирали с расстоянием между ними в несколько миллиметров.

Для одновременного исследования нескольких процессов используют многолучевые трубки с двумя, четырьмя, пятью или шестью независимыми прожекторами и отклоняющими системами. В СВЧ-осциллографах применяют специальные трубки, имеющие отклоняющую систему типа бегущей волны, в которой ряд коротких пластин укреплен по спирали, причем по мере отклонения луча вертикальное расстояние между пластинами увеличивается. Изменение потенциалов на пластинах и движение электронного луча синхронизированы. Такие трубки позволяют исследовать процессы до нескольких гигагерц.

Радиолокационные трубки. Электронно-лучевые трубки, используемые в радиолокационных установках, принципиальных отличий от осциллографических трубок не имеют, однако к ним предъявляется ряд специфических требований. В них необходимо получить на экране изображение импульса радиосигнала, отраженного от облучаемого объекта, в соответствии с положением объекта в пространстве, т. е. необходимо соотнести положение импульса на экране и положение объекта в пространстве. Сигнал на экране получают в полярных координатах, поэтому электронно-лучевая приемная радиолокационная трубка рабо-

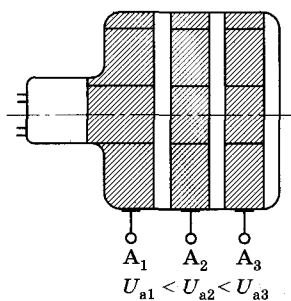


Рис. 12.9

тает в режиме яркостной отметки сигнала и имеет радиально-азимутальную развертку луча: в этом случае одновременно с радиальной разверткой луч равномерно отклоняется в азимутальном направлении. Для получения такой развертки используется магнитная отклоняющая система, состоящая из пары катушек, вращающихся около горловины трубки. Через катушки протекает ток линейной пилообразной формы, который создает переменное магнитное поле, отклоняющее луч от центра экрана по радиусу к периферии. За счет медленного вращения катушек луч прочерчивает на экране смещающуюся по азимуту радиальную линию. Вращения отклоняющих катушек и антенны РЛС синхронизованы. В исходном состоянии на модулятор трубки подано отрицательное напряжение, превышающее по модулю напряжение запираения. Трубка открывается отраженным от объекта радиоимпульсом. В результате на экране возникает светящееся пятно, удаление которого от центра экрана определяет расстояние до объекта, а смещение по азимуту от фиксированного значения соответствует азимуту лоцируемого объекта.

В РЛС используются также трубки с *круговой разверткой* и стержневым электродом, расположенным в центре экрана. При отсутствии отраженного сигнала луч описывает окружность на периферии экрана, центр которой совпадает с расположением стержневого центрального электрода. В момент появления сигнала, который подается на центральный электрод, описывающий окружность луч отклоняется к центру. Радиальное отклонение может быть также получено с помощью двух конусообразных электродов, расположенных после отклоняющей системы.

Точность определения положения цели зависит от размеров светящегося пятна и степени соответствия положения пятна на экране координатам объекта. Следовательно, трубка должна обладать высокой разрешающей способностью, большой яркостью свечения экрана, высоким контрастом и хорошей линейностью отклонения, а также длительным послесвечением (последнее требование связано с малой длительностью импульса, отпирающего электронный луч). Указанным требованиям удовлетворяют трубки с магнитной фокусировкой и магнитным отклонением, а также с большими ускоряющими напряжениями (6...25 кВ) и длительным послесвечением экрана.

В РЛС используют также ЭЛТ с записью темной трассой — *скиатроны*. Экран у них выполнен на основе хлористого калия, который темнеет при бомбардировке его быстрыми электрона-

ми. След электронного луча, перемещающегося по экрану, представляет собой темную фиолетовую линию и может сохраняться несколько дней и более. При необходимости обесцвечивание экрана достигается кратковременным прогревом путем пропуска импульса тока через прозрачный проводящий слой, который одновременно является подложкой люминофора.

Кинескопы черно-белого изображения. *Кинескоп* — это электронно-лучевая телевизионная трубка, предназначенная для приема изображения. Телевизионное изображение строится таким образом, что каждый кадр разбивается на определенное число строк. В нашей стране принят стандарт, при котором кадр содержит 625 строк, что определяет количество элементов разложения каждого кадра. Число элементов разложения в кадре определяет и требования к *разрешающей способности* кинескопа — для любого экрана она должна быть не менее 625 строк. Каждая строка изображения представляется сложным электрическим видеосигналом, в котором мгновенное значение напряжения в данный момент времени соответствует яркости изображения детали предмета (объекта). Следовательно, изменение напряжения внутри строчечного сигнала (видеоимпульса) будет определять и изменение яркости во времени по тому же самому закону, что и сигнал, т. е. кинескопы являются трубками с яркостной отметкой. Полное число строк, уместившихся на экране, воспроизведет один кадр изображения. Размеры экрана, а следовательно, и изображения, должны отвечать наилучшим условиям наблюдения. Для человеческого глаза оптимальный угол зрения — около 15° , и при расстоянии от экрана в несколько метров наиболее рациональные размеры экрана составляют 0,5...0,8 м, что и реализуется в выпускаемых промышленностью кинескопах.

Для кинескопа с диагональю экрана 0,6 м и высотой 0,36 м (стандартом установлено соотношение сторон изображения 4 : 3) размер пятна на экране при 625-строчной развертке составляет 0,57 мм, что легко обеспечивается использованием трубок с электростатической фокусировкой. Применение электростатических прожекторов в кинескопах приводит к менее жестким требованиям к стабильности источников питания, так как нестабильности питающих напряжений в рассматриваемом случае не нарушают фокусировки. Трубки с магнитной фокусировкой не допускают этого. Кроме того, прожекторы с электростатической фокусировкой более экономичны и удобны при массовом производстве кинескопов.

Электронный прожектор, используемый в кинескопах, строится по трехлинзовой схеме (см. рис. 12.4, б). Первый анод имеет больший диаметр, чем рядом расположенные ускоряющий электрод и второй анод. Благодаря такой конструкции ток первого анода близок к нулю, что не изменяет фокусировку электронного луча при регулировании напряжения на модуляторе. Применение промежуточного (первого) ускоряющего электрода с потенциалом 200...600 В позволяет иметь относительно небольшие отрицательные запирающие напряжения. Окончательные ускоряющие напряжения в кинескопах достигают 15...20 кВ, что наряду с использованием алюминированных экранов позволяет получать высокие яркости изображений. Наличие алюминиевой пленки надежно предохраняет люминофор от бомбардировки отрицательными ионами, образующимися внутри трубки.

В качестве покрытия экранов в кинескопах обычно используют механическую смесь желтого (цинк-кадмиевый сульфид с примесью серебра) и голубого (сульфид цинка, активированный серебром) люминофоров. Эти люминофоры позволяют получать световую отдачу не менее нескольких кд/Вт; яркость свечения не менее 30...40 кд/м² при ускоряющем напряжении 10...20 кВ; длительность свечения менее 0,02 с; разрешающую способность на кадр не менее 700...800 строк.

Отклоняющие системы кинескопов должны иметь высокую степень линейности отклонения (т. е. отсутствие искажений по всему экрану с предельным углом отклонения в 110°). Использование угла отклонения 110° уменьшает длину кинескопа в 1,5 раза по сравнению с углом отклонения в 70°. Эти требования реализуются за счет применения магнитной отклоняющей системы, которая содержит четыре катушки без ферромагнитных сердечников с последовательно складывающимися магнитными потоками. Вертикально-отклоняющие катушки (кадровые) и горизонтально-отклоняющие (строчные) совмещены в пространстве, что позволяет получить компактную конструкцию отклоняющей системы.

Колба кинескопа определяет многие эксплуатационные характеристики трубки. Давление окружающего воздуха на экран составляет ~ 10⁵ Па, поэтому для обеспечения высокой механической прочности и в целях безопасности экран выполняется выпуклой формы из стекла толщиной до 10 мм. Применение толстых экранов хотя и уменьшает яркость свечения, однако существенно снижает яркость ореолов, которые сильно влияют на качество изображения.

Для подачи высокого напряжения на второй анод прожектора внутреннюю поверхность колбы покрывают *аквадагом* (проводящим графитовым слоем). Наружная поверхность трубок в широкой части часто (при $U_a > 10$ кВ) тоже покрывается аквадагом. Внутреннее и внешнее покрытия электрически изолированы и образуют конденсатор фильтра высоковольтного выпрямителя.

Цветные кинескопы. Для воспроизведения цветного изображения в телевидении применяются системы, построенные на основе теории трехкомпонентного смешения синего, зеленого и красного цветов. В соответствии с этой системой ЭЛТ в цветном телевидении должны воспроизводить сигналы, несущие всю необходимую информацию о трех основных цветах, присутствующих в объекте, изображение которого должно быть передано. Каждому цвету должен соответствовать свой сигнал, который управляет яркостью свечения элементов экрана. Практическое распространение получили кинескопы с теньевыми масками и мозаичными трехкомпонентными экранами. Эти кинескопы имеют три прожектора с осями, расположенными или в вершинах равностороннего пространственного треугольника (дельтавидное расположение), или на одной линии в плоскости (планарное расположение) (рис. 12.10). В первом случае используется мозаичный экран, у которого зерна трех цветов люминофора, светящиеся красным, синим и зеленым цветом, образуют смежные дельтавидные триады *11*; во втором применяется линейчатый экран, у которого зерна люминофора в виде узких полос *12* красного, синего и зеленого цветов свечения располагаются последовательно и перпендикулярно строкам телевизионного изображения (см. рис. 12.10).

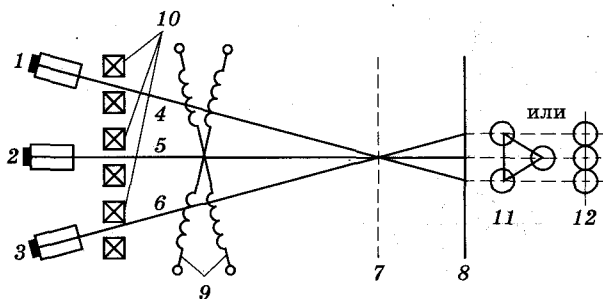


Рис. 12.10

Принцип получения цветного изображения заключается в следующем. Три прожектора (1, 2, 3) с электростатической фокусировкой создают три независимо управляемых электронных луча (4, 5, 6). За счет наклона и/или специальной сводящей системы 10 все три луча пересекаются в точках, расположенных в одной плоскости, в которой размещается теневая маска 7 (в виде тонкой металлической пластины толщиной $\sim 0,15$ мм, имеющей $(3...6)10^5$ отверстий диаметром 0,25 мм). За маской на расстоянии ~ 12 мм устанавливается либо мозаичный, либо линейчатый экран 8. Система сведения устроена так, что луч каждого прожектора будет возбуждать свечение только одного цвета при любом угле отклонения. Отклонение всех трех лучей по строкам осуществляется одновременно одной и той же магнитной отклоняющей системой 9. Вне зависимости от угла отклонения все три луча пересекаются в плоскости маски и проходят через одно отверстие, возбуждая свечение определенного цвета, что обеспечивает правильное воспроизведение цветов. Для сведения лучей в одну точку при больших углах отклонения в пределах маски используется электрическая коррекция пространственного положения лучей с помощью динамической сводящей системы, состоящей либо из электростатической линзы, либо из трех магнитов 10, располагаемых снаружи и смещающих лучи в радиальном направлении. На магнитопроводах сводящих магнитов располагают специальные катушки, через которые пропускают ток, пульсирующий с частотой строчной и кадровой разверток. Если для сведения используется электростатическая линза, то к ее электродам подводятся потенциалы, синхронизированные с кадровой и строчной развертками.

При планарном расположении прожекторов за счет использования определенным образом сформированных пространственных распределений горизонтально и вертикально отклоняющих полей удается получить хорошее сведение лучей по всему экрану, т. е. обеспечить самосведение электронных лучей. Для формирования нужных распределений магнитных полей используется специальная отклоняющая система. Такие кинескопы с планарным расположением прожекторов и самосведением лучей удобны и надежны в эксплуатации.

Более простая процедура сведения лучей является существенным достоинством кинескопов с планарным расположением прожекторов, однако они сложнее в производстве и требуют большей точности изготовления и сборки отдельных узлов трубки.

Чем сильнее различаются размеры зерен и лучей, тем лучше цветоделение. Однако с увеличением разности поперечных размеров зерен и лучей снижается яркость, увеличивается шаг маски и уменьшаются размеры отверстий, что приводит к потере части электронов, ухудшению разрешающей способности и т. д., поэтому при конструировании трубки приходится осуществлять многопараметрическую оптимизацию.

Запоминающие трубки. Предназначены для записи на диэлектрическом экране или специальной мишени сигналов с воспроизведением (считыванием) их через некоторый интервал времени в электрической, оптической или той и другой формах. Применяются запоминающие трубки в индикаторах РЛС, телевидении, автоматизированных системах управления, в устройствах отображения информации ЭВМ, для исследования однократных и медленно протекающих процессов и в ряде других областей.

Запоминающие трубки используют двойное преобразование информации: первое преобразование — *запись информации*, второе — *считывание*, или *воспроизведение информации*. При необходимости, кроме записи и считывания, в запоминающих трубках может использоваться и стирание информации. Время хранения информации — от долей секунды до нескольких дней.

Принцип действия этих приборов основан на использовании зависимости потенциала диэлектрического экрана (мишени) от энергии бомбардирующих его электронов. Для создания потенциального рельефа чаще всего используют вторичную электронную эмиссию. Запись и считывание осуществляется с помощью электронного луча. Энергия электронов, взаимодействующих с мишенью и определяющая коэффициент вторичной эмиссии, зависит от потенциала той точки мишени, куда попадет луч. Для диэлектрической поверхности этот потенциал может существенно отличаться от потенциала анода прожектора. В результате в различных точках мишени коэффициент вторичной эмиссии σ может быть как больше единицы, так и меньше. Это приводит к тому, что одни участки мишени будут заряжаться положительно, а другие отрицательно. На поверхности мишени формируется потенциальный рельеф. Существует несколько способов образования потенциального рельефа, реализуемых в различных типах запоминающих трубок.

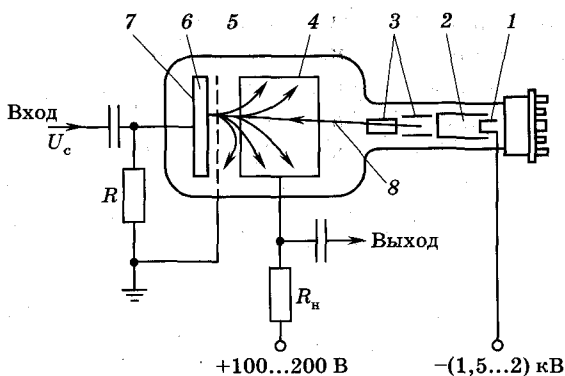


Рис. 12.11

Запоминающая трубка с барьерной сеткой является одним из наиболее распространенных типов электронно-лучевых приборов. Устройство трубки показано на рис. 12.11. Трубка содержит катод 1; электронный прожектор 2, создающий пучок быстрых электронов (энергия ~ 1 кэВ); отклоняющие пластины 3; коллектор вторичных электронов 4, имеющий большой положительный потенциал; тонкую металлическую барьерную сетку 5; диэлектрическую мишень 6, которая покрыта с одной стороны металлической пленкой (сигнальная пластина) 7. Барьерная сетка располагается на расстоянии $\sim 0,1$ мм от поверхности мишени.

Сигнальная пластина—диэлектрик—поверхность мишени образуют структуру, по своим свойствам подобную конденсатору, состоящему из множества элементарных конденсаторов.

При отсутствии сигнала электронный луч 8 будет выбивать с любой точки мишени одинаковое количество вторичных электронов, собираемых коллектором 4. Мишень приобретает одинаковый во всех точках равновесный потенциал, близкий к потенциалу коллектора.

При подаче на сигнальную пластину входной информации и одновременной развертке поверхности мишени электронным лучом с неизменным током элементы мишени приобретут новые равновесные потенциалы. Однако заряд, накапливаемый элементами мишени, зависит от величины входного сигнала. В результате на поверхности мишени относительно сигнальной пластины создается потенциальный рельеф. Рассмотрим этот процесс более подробно.

Если в некоторый момент времени на сигнальную пластину 7 подать положительное напряжение сигнала U_c , то потенциал барьерной сетки относительно мишени становится отрицательным, что будет препятствовать уходу некоторой части вторичных электронов на коллектор. Потенциал элемента мишени, куда попадает электронный луч, понижается по сравнению с равновесным значением. Этот элемент приобретает новый потенциал, величина которого отличается от равновесного на U_c , т. е. мгновенное напряжение сигнала U_c оказывается «записанным» на определенном элементе мишени. Если U_c изменяется во времени, то другим значениям напряжений сигнала будут соответствовать иные элементы мишени, где находится луч. Каждый элемент окажется заряженным до определенного значения, т. е. входной сигнал, изменяющийся во времени, наносится на поверхность мишени в виде пространственного потенциального рельефа. Одновременно с записью происходит повторение сигнала на резисторе R_n в цепи коллектора.

Считывание записанного сигнала производится при повторной развертке луча и отсутствии сигнала. Элементы мишени разряжаются по цепи «электронный луч (8) — земля — сигнальная пластина (7)» до значения равновесного потенциала, и ток в указанной цепи и цепи коллектора будет промодулирован по закону сигнала. Если при считывании на пластину поступает новый сигнал, элементы мишени заряжаются до напряжения, равного разности записанного и нового потенциалов, т. е. происходит вычитание сигналов. В силу сказанного такие трубки называют также *вычитающими*. В рассмотренном случае запись и считывание разнесены во времени. Записанная информация может сохраняться в виде потенциального рельефа на мишени, а затем по необходимости считываться.

Графekon. Такой тип запоминающих трубок с двумя электронными прожекторами 1 и 6 (рис. 12.12) применяется в радиолокации и телевидении. Прожектор 1 предназначен для записи, прожектор 6 — для считывания информации с диэлектриче-

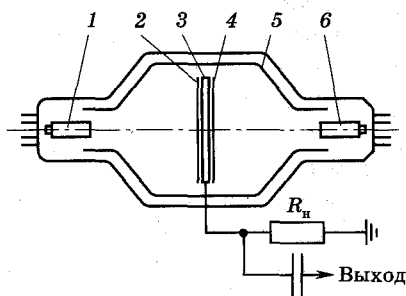


Рис. 12.12

ской мишени. Электронные прожекторы размещены по разные стороны мишени, причем прожектор 1 располагается со стороны сигнальной пластины 3 и барьерной сетки 2.

Напряжение сигнала подается на управляющий электрод записывающего прожектора и модулирует электронный луч по плотности. Электроны записывающего луча ускоряются высокими положительными потенциалами (~ 10 кВ). Они легко пронизывают сигнальную пластину и глубоко проникают в слой диэлектрика мишени 4, изменяя его проводимость. В месте нахождения пучка потенциал поверхности мишени 4 снижается и при больших токах луча приближается к потенциалу сигнальной пластины, на которую подается отрицательное относительно коллектора напряжение в несколько десятков В. Поскольку плотность электронного луча модулирована по закону сигнала, она приводит к различному изменению потенциала элементов мишени, т. е. на поверхности мишени создается потенциальный рельеф, максимальная глубина которого определяется разностью потенциалов между сигнальной пластиной и коллектором 5. Большая глубина потенциального рельефа позволяет осуществлять многократное считывание.

При считывании немодулированный луч с энергией ~ 1 кэВ движется по мишени и вызывает вторичную эмиссию с коэффициентом, пропорциональным записанному потенциалу. Вторичные электроны вытягиваются от поверхности мишени полем коллектора 5, разряжают элементарные конденсаторы, образованные сигнальной пластиной — элементом диэлектрика — поверхностью мишени, и в цепи сигнальной пластины протекают токи, которые создают на резисторе R_H напряжение, воспроизводящее сигнал.

Поскольку энергия считывающего луча много меньше, чем записывающего, характеристики потенциального рельефа изменяются незначительно и лишь на небольшой глубине, что позволяет считывать сигнал до нескольких сотен раз.

Запоминающие трубки с накоплением заряда, используемые в радиолокации, позволяют достаточно эффективно выделять сигнал на фоне помех. Сигнал многократно записывается на одних и тех же элементах мишени, и происходит накопление напряжения сигнала на фоне случайного сигнала помехи.

Существует разновидность запоминающих трубок для воспроизведения информации в виде изображения на люминесцентном экране. Это так называемые запоминающие трубки с видимым

изображением. По своим функциям они аналогичны осциллографическим радиолокационным трубкам и кинескопам, но в отличие от них могут длительное время сохранять информацию.

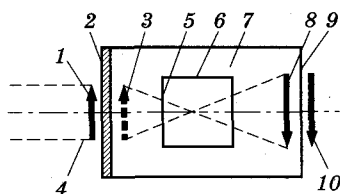


Рис. 12.13

Электронно-оптические преобразователи (ЭОП). Этот тип электровакуумных приборов предназначен для преобразования спектрального состава излучения и усиления яркости изображения. ЭОП применяются для преобразования изображений из инфракрасного диапазона в видимый (приборы ночного видения), для изучения быстропротекающих излучательных процессов, в качестве усилителей яркости: в телевидении, для скоростной киносъемки быстропротекающих процессов, для наблюдения слабоосвещенных и слабоизлучающих предметов, преобразования изображения из одной области спектра в другую (астрономия, медицина, радиоспектроскопия и спектроскопия и т. д.).

В ЭОП происходит двойное преобразование изображения объекта 4 (рис. 12.13). Сначала на фотокатод 2 оптическое изображение 1 трансформируется в электронное 3, последнее с помощью электронных потоков 5, управляемых электронными линзами 6, переносится через вакуумное пространство 7 на люминофорный экран 9, где из электронного изображения 8 преобразуется вновь в видимое изображение 10, но уже с другим спектральным составом.

Описанный однокамерный ЭОП усиливает яркость изображения в несколько десятков раз и имеет разрешение (разрешающую способность) 25 штрихов/мм. Разрешение — это параметр ЭОП, который оценивается наибольшим числом чередующихся светлых и темных линий (штрихов) на 1 мм изображения в центре фотокатода, видимых отдельно при оптимальной для наблюдателя яркости и достаточном увеличении. Единицей разрешения является пара линий/мм. Электронные линзы уменьшают электронное изображение, что позволяет увеличивать яркость.

Для увеличения яркости используются многокамерные ЭОП, которые представляют собой соединения нескольких однокамерных ЭОП в одном вакуумном объеме. На рис. 12.14 схематично представлен трехкамерный ЭОП, где камеры обозначены цифрами 1, 2, 3.

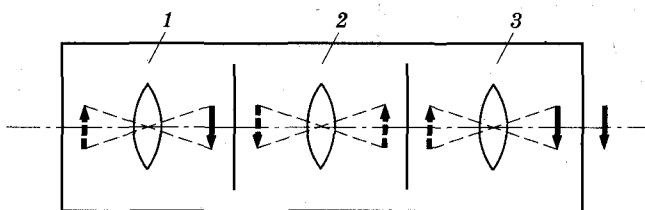


Рис. 12.14

Все шире применяются каналные ЭОП, в которых каскад усиления выполнен на микроканальной пластине. Микроканальная пластина толщиной 0,5...1,5 мм содержит большое количество параллельных отверстий (трубок, каналов), каждое из которых по существу является элементарным фотоэлектронным умножителем, усиливающим фототок элемента изображения. В рассматриваемых ЭОП электронное изображение с фотокатода с помощью электронной линзы переносится на пластину. Усиление в каждом канале происходит независимо от соседнего отверстия, т. е. изображение разбивается на элементы, число которых определяется количеством отверстий. После пластины за счет однородного электростатического поля электронное изображение переносится на люминесцентный экран.

Коэффициент усиления каждого канала $\sim 10^3 \dots 10^5$, разрешение 20...25 пар линий/мм при диаметре канала 20 мкм. Уменьшение диаметра каналов до 10 мкм увеличивает разрешение до 40 пар линий/мм.



Контрольные вопросы

1. Каковы классификация, устройство и принцип действия электронно-лучевых приборов?
2. Электронный прожектор. Электростатические фокусирующие системы триодного и тетродного типа: устройство, формирование электронных линз, принципы фокусирования, достоинства и недостатки.
3. Магнитная фокусировка: физические принципы фокусировки заряженных частиц в неоднородном магнитном поле, устройство фокусирующих систем, достоинства и недостатки, качество фокусировки.
4. Электростатические отклоняющие системы: принципы работы, параметры, характеристики, преимущества и недостатки.

5. Магнитные отклоняющие системы: принципы работы, параметры и характеристики, преимущества и недостатки.
6. Экраны электронно-лучевых трубок: устройство, разновидности, параметры и характеристики.
7. Каковы особенности радиолокационных ЭЛТ?
8. Каковы особенности осциллографических ЭЛТ?
9. Принципы работы, параметры и характеристики черно-белых и цветных кинескопов.
10. Электронно-оптические преобразователи: устройство, принципы работы, параметры и характеристики.

Глава 13

ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ СВЧ С ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ И ПРОДОЛЬНОМ ВЗАИМОДЕЙСТВИЕМ — ПРИБОРЫ ТИПА О

13.1. Общие сведения

В лампах, рассмотренных в предыдущих разделах, используется электростатическое управление электронным потоком, которое возможно на сравнительно низких частотах, где время пролета электронов между электродами лампы много меньше периода колебаний, поэтому можно считать, что каждый электрон будет перемещаться в статическом электрическом поле. В этих условиях переменное напряжение на управляющих электродах лампы будет вызывать синфазное изменение плотности электронного потока. В СВЧ-диапазоне (частоты от 300 МГц до 3000 ГГц, длины волн от 1 м до 0,1 мм) время пролета электронов между электродами оказывается порядка или даже много больше периода колебаний. Кроме того, индуктивности выводов, междуэлектродные емкости и СВЧ-потери в металле электродов, материале баллона и изоляторах также сильно проявляются в этом диапазоне частоты. Все это ухудшает характеристики приборов. Указанных недостатков лишены *электровакуумные приборы СВЧ с динамическим управлением электронным потоком*, при котором время пролета может быть порядка и больше периода СВЧ-колеба-

ний. Сущность такого управления состоит в том, что под действием переменного СВЧ-поля происходит изменение скорости электронов во времени по периодическому закону. В результате из-за разницы в скоростях электронов при их движении в пространстве дрейфа формируются разрежения и уплотнения электронов (образуются сгустки, т. е. области с повышенной плотностью электронов). Процесс образования сгустков в электронном потоке называют *модуляцией по плотности*, а процесс изменения скоростей электронов из-за СВЧ-поля — *модуляцией по скорости*. После формирования сгустков в приборах обеспечивается передача энергии от электронных сгустков СВЧ-полю колебательной системы. Поскольку в диапазоне СВЧ размеры прибора порядка и больше длины генерируемых и усиливаемых колебаний, то в качестве колебательных систем используются системы с распределенными постоянными, при этом узкополосные колебательные системы называются *резонаторами*, а широкополосные — *замедляющими системами*.

Электривакуумные приборы (ЭВП) по характеру энергообмена между электронным потоком и СВЧ-полем разделяются на *приборы О- и М-типа*. В приборах типа *О* в энергию СВЧ-поля преобразуется кинетическая энергия электронов. Магнитное поле в таких приборах либо отсутствует, либо играет вспомогательную роль. Эти приборы бывают двух видов — с кратковременным взаимодействием электронного потока и поля резонатора (пролетные и отражательные клистроны) и с длительным взаимодействием электронов и СВЧ-поля замедляющих систем (лампы бегущей волны типа *О* (ЛБВО), лампы обратной волны типа *О* (ЛОВО)).

В приборах типа *О* векторы напряженности постоянных электрических и магнитных полей (если последние присутствуют) не пересекаются, они параллельны. Кроме того, взаимодействие электронного потока происходит с составляющей электрического СВЧ-поля, вектор которой параллелен вектору скорости электронов потока. В этом плане приборы типа *О* можно назвать приборами с продольным взаимодействием, поскольку электрические и магнитные поля, обеспечивающие эффективность взаимодействия, направлены вдоль скорости электронного потока.

В приборах типа *М* происходит только длительное взаимодействие, и в энергию СВЧ-поля преобразуется потенциальная энергия электронов. Движение электронов в этих приборах осуществляется в скрещенных электрических и магнитных полях.

В настоящее время ЭВП СВЧ с динамическим управлением по сравнению с другими типами электронных приборов имеют наибольшие величины усиливаемой и генерируемой СВЧ-мощности, КПД, коэффициента усиления, полосы пропускаемых частот и ряда других параметров. В силу этого они занимают доминирующее положение в СВЧ-технике. Физические процессы в таких приборах описываются уравнениями, приведенными в приложении 2.

Специфическая особенность ЭВП СВЧ заключается в том, что они являются интегральными устройствами, включающими в себя вакуумный прибор, колебательную систему, магнитную систему, входные и выходные СВЧ-тракты и другие элементы. Элементы ЭВП СВЧ взаимосвязаны и взаимообусловлены, т. е. электронная часть создается с учетом свойств и особенностей всех других узлов прибора, особенно колебательной системы.

13.2. Клистроны

13.2.1. Пролетные клистроны.

В пролетных клистродах используются узкополосные колебательные системы (резонаторы). Как правило, количество резонаторов в этих приборах колеблется от двух до четырех-пяти. Основное назначение пролетных клистронов — генерация, усиление и умножение СВЧ-колебаний. Этот тип приборов может работать в диапазоне частот от сотен МГц до сотен ГГц, максимальные значения выходной мощности могут достигать сотен кВт в непрерывном режиме. Перестройка по частоте в этих приборах осуществляется механически.

Рассмотрим сначала двухрезонаторные клистроны (ДРК), которые наиболее ярко иллюстрируют принципы динамического управления электронным потоком, а затем перейдем к многорезонаторным клистродам, получившим наибольшее распространение. Схема устройства ДРК приведена на рис. 13.1. Катод 1 и ускоряющий электрод 2 формируют электрон-

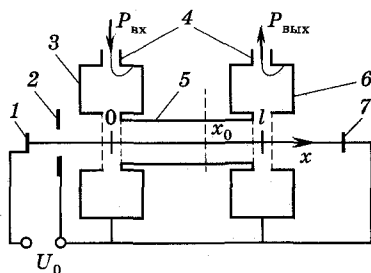


Рис. 13.1

ный поток с заданными характеристиками. Электроны подходят к сетке первого (входного) резонатора 3 с одинаковой скоростью

$$v_0 = \sqrt{\frac{2e}{m} U_0}, \quad (13.1)$$

e , m — заряд и масса электрона, U_0 — ускоряющее напряжение.

Входной резонатор 3 возбуждается через входное устройство 4 входным СВЧ-сигналом мощностью $P_{\text{вх}}$ и частотой ω . В результате между сетками этого резонатора существует переменное электрическое поле напряжением $u_1 = U_{m1} \sin(\omega t_1)$, которое в положительный полупериод дополнительно ускоряет электроны, а в отрицательный — тормозит (U_{m1} — амплитуда переменного напряжения на сетках резонатора 3). Следовательно, после первого резонатора электроны будут иметь разные скорости. Этот процесс называется *модуляцией электронов по скорости*. В пространстве дрейфа 5 электроны перемещаются без дополнительных воздействий. Из-за разной скорости электронов в электронном потоке будут образовываться уплотнения (сгустки) и разрежения, т. е. модуляция по скорости превращается в *модуляцию по плотности*. Частота следования сгустков равна частоте входного сигнала. Сгустки электронов, проходя между сетками выходного резонатора 6, настроенного на частоту входного сигнала, наводят в его цепи ток той же частоты, который через выходное устройство 4 поступает в нагрузку. Между сетками возникает СВЧ-напряжение. Сгустки пролетают выходной резонатор в тормозящем полупериоде и отдают СВЧ-полю часть энергии. Остальная часть энергии электронного потока рассеивается в виде тепла на коллекторе 7.

Процесс модуляции скорости электронов под действием синусоидального сигнала u_1 на входе можно описать формулой

$$u_1 = v_0(1 + M \sin(\omega t_1)), \quad (13.2)$$

где v_1 — скорость электронов в пространстве дрейфа; время t_1 определяет текущую фазу СВЧ-поля в момент прохождения электронами середины зазора входного резонатора 3 ($x = 0$);

$M = \frac{\beta U_{1m}}{2U_0}$ — коэффициент модуляции электронного потока, U_0 — ускоряющее напряжение (см. рис. 13.1); $\beta = \frac{\sin \Theta_s/2}{\Theta_s/2}$ — коэффици-

ент модуляции электронного потока, Θ_s — угол отклонения электронов в выходном резонаторе.

ент эффективности взаимодействия электронов с полем резонатора; $\Theta_3 = \Theta_2 - \Theta_1$ — фазовый угол пролета электрона в зазоре входного резонатора (Θ_1/Θ_2 — фазы СВЧ-поля соответственно в момент влета электрона в зазор и в момент вылета из зазора резонатора).

Эффективность группировки электронного потока характеризуется параметром группировки X :

$$X = M\theta_0, \quad (13.3)$$

где $\theta_0 = \omega \cdot t_0$ — абсолютный угол пролета электрона, отсчитываемый от момента пересечения середины зазора входного резонатора ($x = 0$, см. рис. 13.1) до рассматриваемой плоскости x в пространстве дрейфа; t_0 , соответственно, время пролета. Таким образом, фазовый угол θ_0 определяет, на сколько изменится фаза СВЧ-поля за время пролета t_0 .

Закон изменения переменного напряжения на сетках и график движения электронов в промежутке между резонаторами показан на пространственно-временной диаграмме (рис. 13.2, где $x_3 = l$ — расстояние между резонаторами).

Синусоида в нижней части рисунка изображает переменное напряжение между сетками входного резонатора. Цифрами на оси абсцисс отмечены моменты прохождения входного резонатора различными электронами.

Для простоты рассмотрения будем считать, что электроны мгновенно изменяют свою скорость в середине зазора первого резонатора ($x = 0$, см. рис. 13.1). На рис. 13.2 ось времени t проведена через точку $x = 0$ оси ординат, соответствующую середине зазора входного резонатора.

График движения электронов в пространстве за резонатором представлен прямыми линиями, тангенс угла наклона которых к оси абсцисс пропорционален скорости движения электронов. Электроны 2, 4, 6 и 8, пролетевшие сетки резонатора в моменты прохождения переменного напряжения через нуль, движутся со средней скоростью v_0 , т. е. они не изменяют скорости. Их называют невозмущенными. Прямые

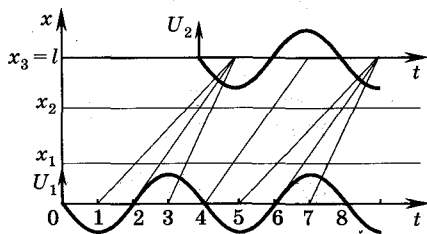


Рис. 13.2

движения для электронов 3, 7, прошедших зазор резонатора во время положительного полупериода, наклонены к оси абсцисс под большим углом, чем прямые для электронов 2, 4, 6 и 8. Графики для электронов 1, 5, прошедших зазор в тормозящий полупериод, составляют с осью абсцисс меньший угол. Электроны 1 и 5, вылетевшие из зазора раньше, чем электроны соответственно 2 и 6, но замедлившие свое движение, могут сблизиться в некоторый момент времени с этими электронами (плоскость $x_3 = l$). Электроны 3 и 7 ускоряются и в некоторый момент времени догонят соответственно электроны 2 и 6 в плоскости $x_3 = l$, соответствующей середине зазора выходного резонатора 6 (см. рис. 13.1 и 13.2). Таким образом, при переходе от тормозящего полупериода напряжения к ускоряющему происходит группирование электронов 1 и 3 около электрона 2, а электронов 5 и 8 — около электрона 6. Электроны 2 и 6 являются центрами группировки, а 4 и 8 — центрами разгруппировки.

Напряжения U_0 и U_{1m} можно подобрать таким образом, чтобы электроны встречались в середине зазора выходного резонатора, т. е. в точке $x = l$. Представленная картина является идеализированной, поскольку в действительности электроны 1, 2, 3 в одной точке не встретятся, а сблизятся на минимальное расстояние. Кроме того, с энергетической точки зрения существует оптимальная форма сгустка. В этом случае будет происходить наибольшая передача энергии от электронов полю. Большая часть электронов из потока сосредоточена в сгустках, которые будут тормозиться СВЧ-полем. Значит, число электронов, передающих энергию полю, больше, чем забирающих ее. Электроны 4 будут проходить второй резонатор в ускоряющий полупериод и будут отбирать энергию от СВЧ-поля.

Напряжение U_2 в верхней части рис. 13.2 является напряжением между сетками второго резонатора. Мощность колебаний, установившихся во втором резонаторе, превышает мощность колебаний, подводимых к входному резонатору, т. е. происходит усиление СВЧ-сигнала по мощности. Во входном резонаторе половина электронов ускоряется, а половина тормозится СВЧ-полем, поэтому расход энергии источника усиливаемых колебаний близок к нулю. В выходном резонаторе процесс отдачи (торможения) энергии электронами преобладает, поскольку моменты существования ускоряющей разности потенциалов соответствуют интервалам между сгустками, где электро-

нов мало. Исходя из рассмотренного механизма группирования электронов нетрудно заключить, что при неизменных напряжениях U_0 , U_{1m} и при фиксированной частоте ω параметр группировки X увеличивается с расстоянием x от входного резонатора. Параметр X определяет изменение во времени конвекционного тока i_{ex} в некотором сечении x пространства дрейфа. Как показывает теоретический анализ процессов в клистроне, закон изменения тока выглядит следующим образом:

$$i_{ex} = \sum_i I_0 / (1 - X \cos \omega t_i), \quad (13.4)$$

где I_0 — ток электронного потока в отсутствие модуляции, ωt_i — угол пролета центрального электрона i -го уплотнения внутри сгустка при $X > 1$. Если $X > 1$, то сгусток распадается на ряд (i) уплотнений, при $X \leq 1$ число $i = 1$.

Формула для i_{ex} показывает, что изменение тока во времени носит периодический, но несинусоидальный характер.

Если разложить эту периодическую функцию (13.4) в ряд Фурье, то ток можно записать в форме

$$i_{ex} = I_0 + \sum_{n=1}^{\infty} 2I_0 J_n(nX) \cos [n(\omega t_2 - \theta)], \quad (13.5)$$

где t_2 — время, определяющее текущую фазу СВЧ-поля в выходном резонаторе b , $J_n(nX)$ — функция Бесселя первого рода n -го порядка, θ — абсолютный угол пролета электронов от $x = 0$ до $x = l$.

В ДРК входной и выходной резонаторы настраиваются на одну и ту же частоту ω , поэтому наибольший интерес представляет первая гармоника электронного тока. При $n = 1$ и $\cos(\omega t_2 - \theta) = 1$ амплитуда первой гармоники тока $I_{em1} = 2I_0 J_1(X)$, где $J_1(X)$ — функция Бесселя 1-го рода 1-го порядка. Максимум амплитуды первой гармоники имеет место при $X = 1,84$, когда $J_1(X) = J_1(1,84) = 0,58$ (рис. 13.3).

Рассмотрим некоторые параметры и характеристики ДРК. Одним из важнейших параметров является **мощность колебаний в нагрузке вы-**

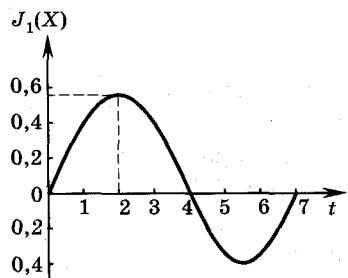


Рис. 13.3

ходного резонатора $P_{\text{вых}}$, которая равна мощности колебаний P_e , развиваемой электронным потоком:

$$P_e = \frac{1}{2} U_{2m} I_{2m} = P_{\text{вых}}, \quad (13.6)$$

$$I_{2m} = \beta_2 I_{em1} = 2I_0 \beta_2 J_1, \quad (13.7)$$

где β_2 — коэффициент взаимодействия электронов с полем выходного резонатора; U_{2m} , I_{2m} — соответственно амплитуды напряжения и тока в выходном резонаторе.

Как отмечалось выше, при $X = 1,84$ имеем $J_1(X) = 0,58$ и мощность

$$P_e = P_{e \text{ макс}} = 0,58 \beta_2 I_0 U_{2m}. \quad (13.8)$$

Для амплитуды напряжения на сетках второго резонатора можно записать

$$U_{2m} = I_{2m} / G_{\text{экр2}} = 2\beta_2 I_0 J_1(X) / G_{\text{экр2}}, \quad (13.9)$$

где $G_{\text{экр2}}$ — эквивалентная активная проводимость второго резонатора.

Таким образом, уровень выходной мощности во многом определяется режимом модуляции электронного потока во входном резонаторе.

Электронный коэффициент полезного действия $\eta_{\text{эл}}$ усилительного клистрона находится из соотношения

$$\eta_{\text{эл}} = \frac{P_{e \text{ макс}}}{U_0 I_0} = \frac{0,58 \beta_2 U_{2m}}{U_0}. \quad (13.10)$$

В реальных условиях $U_{2m} < U_0$, а $\beta_2 < 1$, но если взять предельные недостижимые значения $U_{2m} = U_0$ и $\beta_2 = 1$, то $\eta_e = \eta_{e \text{ макс}} = 0,58$. Таким образом, КПД ДРК не может превышать 58%.

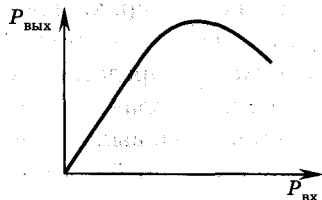


Рис. 13.4

Коэффициент усиления клистрона по мощности рассчитывается по формуле $K_P = 10 \lg P_{\text{вых}} / P_{\text{вх}}$ [дБ]. Коэффициент усиления, как и выходная мощность $P_{\text{вых}}$, зависит от режима работы клистрона. Амплитудная характеристика $P_{\text{вых}} = f(P_{\text{вх}})$ представлена на рис. 13.4.

Из приведенных выше формул следуют соотношения для входной и выходной мощности и функциональная связь между этими величинами:

$$P_{\text{вх}} = \frac{(U_{1m})^2}{2} G_{\text{экв1}} \quad (G_{\text{экв1}} \text{ — эквивалентная активная проводимость первого резонатора})$$

$$\text{и } X = M\theta_0 = \frac{\beta U_{1m}}{2U_0} \theta_0, \text{ то } P_{\text{вх}} \sim X^2, \text{ а}$$

$$P_{\text{вых}} \sim P_e = \beta_2 I_0 J_1(X) U_{2m} = \beta_2 I_0 J_1(X) \frac{I_{em1}}{G_{\text{экв2}}} = 2\beta_2^2 J_1^2(X) / G_{\text{экв2}}, \text{ т. е.}$$

$$P_{\text{вых}} \sim J_1^2(X) = f(X^2). \quad (13.11)$$

Таким образом, амплитудная характеристика $P_{\text{вых}} = f(P_{\text{вх}})$ (см. рис. 13.4) описывается зависимостью $J_1^2(X) = f(X^2)$, максимум которой, согласно приведенному выше анализу, соответствует значению $X = 1,84$. Увеличение $P_{\text{вых}}$ на восходящем участке характеристики при возрастании $P_{\text{вх}}$ обусловлено все более плотной группировкой электронов по мере роста амплитуды входного сигнала. Падение $P_{\text{вых}}$ при больших значениях $P_{\text{вх}}$ на нисходящем участке характеристики обусловлено перегруппировкой электронного потока, т. е. сгустки в этом случае формируются до прихода их во второй резонатор и при дальнейшем движении начинают расплываться из-за разных скоростей электронов, входящих в сгусток.

Форма волны конвекционного тока богата гармоническими составляющими, поэтому ДРК может эффективно работать в качестве умножителя частоты. В этом случае выходной резонатор настраивается на одну из высших гармоник частоты входного сигнала. ДРК обладают малыми КПД, коэффициентом усиления и выходной мощностью. Существенно лучшие характеристики имеют **многорезонаторные клистроны**, в которых реализуются условия многокаскадной группировки. Так, например, трехрезонаторный клистрон, имеющий входной резонатор, промежуточный ненагруженный и выходной резонаторы, можно представить как совокупность двух эквивалентных ДРК, сформированных соответственно входным и выходным и промежуточным и выходным резонаторами. При многокаскадной группировке удастся более оптимально сгруппировать электронный поток и поднять теоретический КПД в трехрезонаторном клистроне до 73%. Многорезона-

торные клистроны по сравнению с двухрезонаторными способны усиливать существенно более слабые сигналы. Уровень $P_{\text{вых}}$ в таких приборах может достигать сотен киловатт в непрерывном режиме, реальные значения КПД — до 40 и более процентов (по сравнению с 15...20% в ДРК), коэффициент усиления $\sim 30...50$ дБ.

13.2.2. Отражательные клистроны.

Эти приборы в основном используются в качестве маломощных генераторов с электронной перестройкой частоты. В них один и тот же резонатор осуществляет и модуляцию электронного потока по скорости, и отбор энергии. Это достигается тем, что модуляция по скорости происходит при прямом движении электронов, а отбор энергии от сгруппированного потока реализуется при обратном движении за счет отражения электронов постоянным тормозящим полем специального электрода (отражателя). Отражательные клистроны (ОК) работают вплоть до субмиллиметрового диапазона длин волн.

Устройство. Стационарные процессы движения электронов. Схематическое устройство ОК приведено на рис. 13.5, а, где цифрами обозначены следующие элементы: 1 — катод, 2 — ускоряющий электрод, 3 — резонатор, 4 — отражатель. На ускоряющий электрод и резонатор подается постоянный положительный потенциал (до нескольких сотен В), а на отражатель — отрицательный потенциал порядка $U_{\text{отр}} \approx -(100...200)$ В. Распределение постоянных потенциалов внутри ОК в направлении

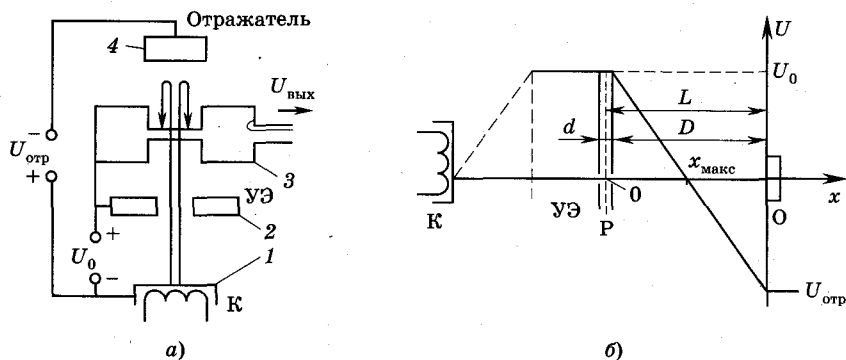


Рис. 13.5

от катода к отражателю дано на рис. 13.5, б. Если не учитывать СВЧ-поля в резонаторе, т. е. рассматривать условия, когда клистрон не возбужден, то до управляющего электрода (УЭ) электроны ускоряются, от УЭ до резонатора, включая его сетки, электроны летят с постоянной скоростью $v_0 = \sqrt{\frac{2e}{m}U_0}$, а в пространстве между резонатором (Р) и отражателем (О) они тормозятся и в плоскости $x = x_{\text{макс}}$ останавливаются и начинают свое движение в обратном направлении.

Принцип действия. Поток электронов, проходящий через сетки резонатора в силу теплового разброса скоростей электронов, вылетающих из катода, имеет флуктуирующую (шумовую) составляющую $\bar{v}_{\text{ш}}$ относительно v_0 , т. е. скорость электронов $v = v_0 + \bar{v}_{\text{ш}}$, где $\bar{v}_{\text{ш}}$ — усредненная шумовая (флуктуирующая) составляющая скорости. Вследствие флуктуаций электронного потока в резонаторе наводятся слабые колебания, т. е. на сетках резонатора из-за тепловых флуктуаций всегда существуют слабые колебания с частотой, близкой или равной собственной частоте резонатора. Резонатор из широкого спектра флуктуаций выделяет только те частоты, которые входят в его полосу пропускания. В результате процессов, которые будут описаны ниже, в клистроне устанавливаются при заданных U_0 и $U_{\text{отр}}$ стационарные СВЧ-колебания. Аналогично происходит возбуждение и других типов ЭВП СВЧ-генераторов: ЛОВО, магнетронов и др.

В установленном режиме поток электронов, пролетающих между сетками резонатора, подвергается воздействию переменного электрического поля.

Подобно ДРК скорость электронов в ОК изменяется по закону $v = v_0(1 + M \sin \omega t)$, т. е. как и в ДРК происходит модуляция электронов по скорости. Пролетев резонатор электроны движутся в тормозящем поле отражателя, теряя скорость до нуля в плоскости, где равна нулю мгновенная разность потенциалов $(U_0 + U_m \sin \omega t) + (-U_{\text{отр}}) = 0$ (U_m — амплитуда СВЧ-поля на сетках резонатора). После остановки около плоскости $x = x_{\text{макс}}$ электроны начнут свое движение назад к резонатору. Электроны, пролетевшие резонатор в течение той половины периода, когда переменное напряжение между сетками резонатора (середина зазора) изменялось от максимума к минимуму, образуют при возвращении сгусток. В отличие от ДРК в ОК центром сгу-

сгустка является электрон, пролетевший резонатор, когда мгновенное значение переменного напряжения равно нулю при изменении от положительного к отрицательному. Это связано с тем, что вектор скорости возвращающихся электронов противоположен по направлению вектору скорости электронов, движущихся от катода. Подбором U_0 и $U_{отр}$ можно добиться того, чтобы сгусток электронов пролетал сетки резонатора в максимально тормозящем СВЧ-поле, т. е. когда СВЧ-потенциал в середине зазора был бы максимально положительным (рис. 13.6). При изменении U_0 или $U_{отр}$ глубина проникновения электронов в пространство отражателя будет различной. Например, уменьшение $|U_{отр}|$ приводит к тому, что x_{\max} будет больше и плоскость остановки электронов приближается к отражателю и, следовательно, время движения электронов от резонатора к отражателю и назад будет возрастать. В результате электронный сгусток

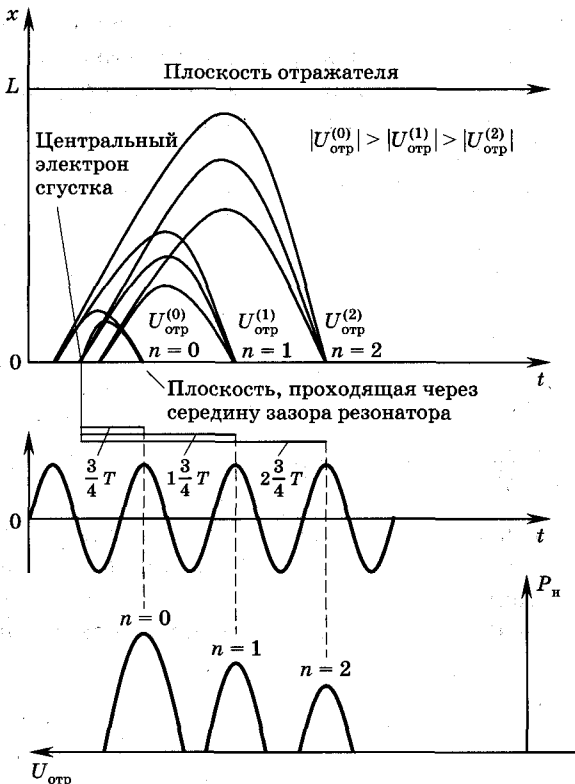


Рис. 13.6

может попадать в ближайший тормозящий период, когда $n = 0$, а время пролета невозмущенного электрона $t_{\text{пр}} = 3/4T$ при напряжении на отражателе $U_{\text{отр}} = U_{\text{отр}}^{(0)}$ или же в следующий тормозящий период, когда $t_{\text{пр}} = 3/4T + nT = (n + 3/4)T$, а $n = 1, 2, \dots$ (значение n представляет собой номер зоны генерации). На рис. 13.6 $U_{\text{отр}} = U_{\text{отр}}^{(0)}$ для $n = 0$, $U_{\text{отр}} = U_{\text{отр}}^{(1)}$ для $n = 1$ и $U_{\text{отр}} = U_{\text{отр}}^{(2)}$ для $n = 2$. Сгустки, попадая в тормозящее СВЧ-поле, передают ему энергию и поддерживают колебания в резонаторе при отводе ее в нагрузку. Зависимость мощности в нагрузке P_n для рассмотренных зон генерации приведена в нижней части рис. 13.6.

Ток в отражательном клистроне можно вычислить, пользуясь, как и в ДРК, законом сохранения заряда. Если i_e — конвекционный ток в плоскости, проходящей через середину зазора резонатора, а I_0 — ток электронного потока в отсутствие модуляции, то в результате расчетов можно получить, что при $i = 1$ $i_{ex} = I_0 / (1 + X \cos \theta)$, где $\theta = \omega t$ — фаза переменного напряжения в момент вылета электронов из резонатора. Это выражение отличается от формулы (13.4) для ДРК лишь знаком перед \cos , что объясняется сдвигом центра группирования в ОК на π по сравнению с ДРК. Следовательно, спектральный состав электронного тока в ОК не изменяется по сравнению с ДРК, и все соответствующие выводы, сделанные для ДРК, справедливы и для ОК.

Условия генерации. Эти условия, как и для любого автогенератора, определяются балансом фаз и балансом мощности. Физически баланс фаз в ОК выполняется при условии, что сгусток электронов приходит от отражателя в середину между сетками резонатора в максимуме тормозящего поля, т. е.

$$\theta_{\text{опт}} = \omega t_{\text{пр}} = \frac{2\pi}{T} \left(n + \frac{3}{4} \right) T = 2\pi \left(n + \frac{3}{4} \right).$$

Как видим из этого выражения, условие генерации существует для различных зон генерации (различные значения n).

Условие баланса фаз на основе анализа эквивалентной схемы (рис. 13.7) можно записать в виде равенства реактивных проводимостей

$$B_e + B_p = 0,$$

где B_e , B_p — соответственно реактивная проводимость электронного потока и резонатора.

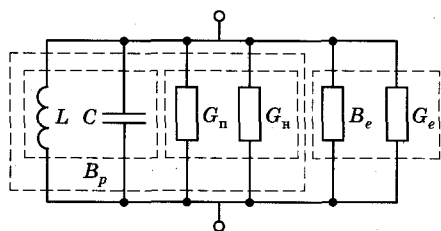


Рис. 13.7

Угол пролета $\theta_{\text{опт}}$ зависит от U_0 и $U_{\text{отр}}$, и поэтому, изменяя эти напряжения, можно получить различные углы пролета, соответствующие разным величинам $n = 0, 1, 2, 3, \dots$, поскольку меняется $t_{\text{пр}}$.

Сгруппированные электроны, возвращаясь к сеткам резонатора, тормозятся СВЧ-полем и тем самым передают энергию этому полю, поддерживая колебания в резонаторе. Сгусток электронов наводит в резонаторе ток, амплитуда которого, так же как и в ДРК (см. выражение (13.7)), равна $I_{2m} = 2\beta I_0 J_1(X)$. (Здесь индекс у β опущен, поскольку в ОК только один резонатор.) Мощность, передаваемая электронным потоком резонатору, равна $P_e = \beta I_0 U_m J_1(X)$. Часть этой мощности ($P_{\text{п}}$) расходуется на восполнение потерь в самом резонаторе, а остальная мощность ($P_{\text{н}}$) отводится в нагрузку:

$$P_e = P_{\text{п}} + P_{\text{н}}.$$

При проводимостях потерь $G_{\text{п}}$ и нагрузки $G_{\text{н}}$ мощности потерь и нагрузки представим соответственно в виде $P_{\text{п}} = (U_m^2 G_{\text{п}})/2$, $P_{\text{н}} = (U_m^2 G_{\text{н}})/2$ (см. рис. 13.6, внизу). При этом $P_e = \beta I_0 U_m J_1(X) = U_m^2 (G_{\text{п}} + G_{\text{н}})/2$. Последнее выражение является условием баланса мощностей в ОК. Если ввести эквивалентную активную проводимость электронного потока G_e (см. рис. 13.7), то условие баланса мощностей можно записать в виде равенства нулю суммы активных проводимостей $G_e + G_{\text{п}} + G_{\text{н}} = 0$, при этом G_e имеет отрицательное значение.

Электронная настройка частоты. В ОК существуют различные зоны генерации не при фиксированных U_0 и $U_{\text{отр}}$, а в определенном диапазоне значений U_0 или $U_{\text{отр}}$. Это связано с тем, что при изменении ускоряющего напряжения и напряжения на отражателе происходит также и изменение частоты колебаний. Такое явление называется *электронной настройкой частоты*. Обычно в ОК частоту изменяют за счет регулировки напряжения на отражателе, так как в его цепи ток не протекает, а значит, и нет потери мощности при электронной настройке. Электронная настройка

частоты присуща всем генераторам на основе ЭВП СВЧ с динамическим управлением: генераторы ЛОВО, магнетроны и другие генераторы на приборах типа *M*, и т. д.

Рассмотрим механизм этого явления в ОК. Формально, из условия баланса фаз $\theta_{\text{опт}} = \omega t_{\text{пр}} = 2\pi \left(n + \frac{3}{4} \right)$ видно, что правая часть выражения является константой при фиксированных *n*, следовательно, при изменении U_0 или $U_{\text{отр}}$, когда изменяется $t_{\text{пр}}$, для выполнения условия баланса фаз соответствующим образом должна изменяться частота ω , т. е. если $t_{\text{пр}}$ растет (при уменьшении $|U_{\text{отр}}|$), то величина ω должна возрастать. Эту ситуацию можно объяснить следующим образом. При уменьшении $|U_{\text{отр}}|$ относительно $|U_{\text{отр}}|_{\text{опт}}$, соответствующего центру зоны генерации, когда $P_{\text{н}} = P_{\text{н макс}}$ (рис. 13.8), электроны образуют сгусток, проникая ближе к отражателю, т. е. $x_{\text{макс}}$ возрастает, следовательно, электроны будут находиться в пространстве отражателя большее время, и сгусток сформируется, не долетев до резонатора. В этом случае напряжение на сетках резонатора будет опережать по фазе ток i_e на некоторый угол, и реактивное сопротивление резонатора будет иметь индуктивный характер. Однако поскольку сгусток формируется раньше прилета в резонатор, то ток (между сетками резонатора) в электронном потоке опережает

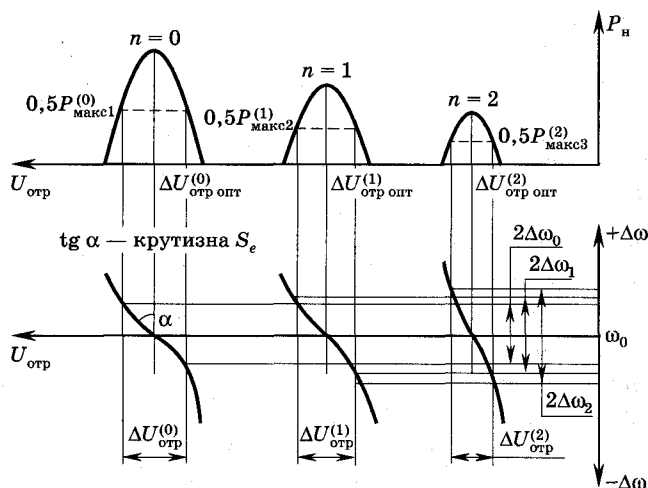


Рис. 13.8

напряжение, поскольку сгусток сформировался раньше прилета в резонатор, т. е. электронная проводимость носит емкостной характер. В результате изменение индуктивной проводимости гасится равным, но противоположным по знаку изменением емкостной проводимости, и сумма реактивных проводимостей опять оказывается равной нулю. Однако резонансная частота всей системы станет меньше, поскольку возросли индуктивность и емкость в эквивалентной схеме ОК. При увеличении $|U_{\text{отр}}|$ относительно $|U_{\text{отр}}|_{\text{опт}}$ будет наблюдаться аналогичная картина, только проводимость резонатора будет носить емкостной, а электронная проводимость — индуктивный характер, и частота возрастет.

Следует отметить, что как в том, так и в другом случаях амплитуда на сетках резонатора станет меньше, поскольку добротность резонатора упадет, так как работа происходит на частотах, отличающихся от резонансной частоты ω_0 . Оптимальный режим соответствует случаю, когда проводимость резонатора и электронного потока являются чисто активными и частота генерируемых колебаний $\omega = \omega_0$.

С увеличением номера колебаний (номера n) время пролета электронов возрастает на соответствующее n число периодов. Для того чтобы электроны окончательно сформировались в сгустки в моменты их прихода в резонатор, амплитуда колебаний на сетках резонатора будет меньше, поскольку сгустки должны формироваться на протяжении большего времени. Допустимые пределы изменения частоты, называемые *диапазоном электронной настройки частоты*, ограничиваются значениями напряжения на отражателе, при которых мощность колебаний уменьшается не более чем в два раза по сравнению с ее максимальным значением в центре области (см. рис. 13.8). Анализ показывает, что величина диапазона электронной настройки возрастает с увеличением I_0 , с уменьшением U_0 и при возрастании n . Такая зависимость связана с тем, что увеличение I_0 приводит к увеличению количества электронов в сгустках и росту мощности колебаний в зоне и ее расширению по оси $U_{\text{отр}}$ и, как результат, росту $\Delta\omega_p$. Это нетрудно понять из рис. 13.8, когда кривые P_n станут выше и шире для каждой зоны. Увеличение параметра n приводит к тому, что при одинаковых изменениях $\Delta|U_{\text{отр}}|$ в случае больших n фазовый сдвиг между токами и напряжениями будет значительно больше, чем при меньших n . Например, возьмем изменение напряжения на отражателе $\Delta|U_{\text{отр}}| = 10$ В для любого n . Пусть при

$n = 1$ $|U_{\text{отр}}|_{\text{опт1}} = 200$ В (центр зоны), тогда $\Delta|U_{\text{отр}}|/|U_{\text{отр}}|_{\text{опт1}} = 0,05$. Для $n = 3$ $|U_{\text{отр}}|_{\text{опт3}} = 50$ В и $\Delta|U_{\text{отр}}|/|U_{\text{отр}}|_{\text{опт3}} = 0,2$, т. е. изменение в четыре раза больше, поэтому и сдвиг по фазе между током и напряжением резонатора во втором случае будет пропорционально больше по сравнению с первым случаем. В результате диапазон и крутизна электронной настройки будут увеличиваться с ростом n .

Поскольку ширина диапазона электронной перестройки обычно мала ($\Delta\omega \leq 0,01\omega_0$), то для работы в широкой полосе частот применяют механическую перестройку частоты отражательного клистрона путем изменения размеров резонатора.

Параметры ОК. Наиболее типичные параметры современных отражательных клистронов лежат в следующих пределах: в среднем выходная мощность ОК находится в диапазоне величин 10...500 мВт, значения напряжения резонатора 200...300 В для ОК сантиметрового и 1000 В для ОК миллиметрового диапазонов длин волн, КПД $\leq 2...3\%$, а электронный КПД — несколько процентов.

Типичные значения крутизны электронной настройки в рабочих зонах генерации ОК: 0,3...0,5 МГц/В — десятисантиметровый; 1,0...2,5 МГц/В — трехсантиметровый; 5...8 МГц/В — восьмимиллиметровый диапазоны длин волн. Значения диапазона электронной перестройки лежат обычно в пределах до 1%. ОК имеют малую массу $\sim 0,05...0,5$ кг, большой срок службы до $(15...20) \cdot 10^3$ часов.

К недостаткам ОК относятся малый КПД и высокий уровень шумов. Дальнейшим развитием ОК являются **минитроны**, представляющие собой приборы с линейными размерами в несколько мм и массой ~ 1 г. Дальнейшая миниатюризация ЭВП СВЧ может привести к созданию интегральных устройств, подобных интегральным схемам на полупроводниковых приборах.

13.3. Лампы бегущей волны.

Устройство ЛБВО. Замедляющие системы

Энергия, передаваемая СВЧ-полю, пропорциональна напряженности поля, числу электронов и расстоянию, на котором происходит взаимодействие электронов и СВЧ-колебаний. Увеличение пути взаимодействия электронов с полем в клистронах

ограничено необходимостью иметь малый угол пролета электронов в зазоре резонатора (если $\theta_0 \rightarrow 0$, то $\beta \rightarrow 1$).

Повышение напряженности СВЧ-поля в зазоре можно осуществить либо за счет увеличения амплитуды колебаний, либо путем уменьшения ширины зазора резонаторов. В первом случае требуются контуры с очень высокой добротностью, а во втором при уменьшении зазора возрастает емкость резонатора и, следовательно, падает добротность; кроме того, амплитуда колебаний напряжения не может быть больше, чем э. д. с. (напряжение) U_0 источника питания. Усилители и генераторы на ЛБВО и ЛОВО в значительной мере лишены указанных недостатков. В этих приборах происходят те же процессы, что и в пролетных клистродах, т. е. под действием СВЧ-поля электроны группируются в сгустки, которые тормозятся электрическим полем волны, передавая последней часть своей кинетической энергии. Отличием ЛБВО и ЛОВО от клистронов является то, что модуляция скорости электронов осуществляется не в определенных, небольших по протяженности, участках пространства, а на достаточно большом отрезке пути в процессе почти синхронного движения волны и электронов у поверхности замедляющей системы.

Устройство. Устройство ЛБВО изображено на рис. 13.9. ЛБВО чаще всего представляет собой длинную стеклянную колбу небольшого диаметра, в одном конце которой размещены электроды 1, 2, 3, образующие электронную пушку, которая формирует электронный поток с заданными характеристиками. Колебательная система прибора состоит из спирали 5 и цилиндрического волновода, являющегося одновременно и каркасом фокусирующей катушки 6. Входной и выходной сигналы поступают соответственно через входной 4 и выходной 7 СВЧ-тракты. Электронный поток, выходящий из пушки, проходит по оси волновода, взаимодействуя с полем волны, распространяющейся по спиральной поверхности волновода. Выходной сигнал снимается с катушки 6. Электронный поток, выходящий из волновода, попадает на электроды 8, 9, 10, которые служат для его замедления и фокусировки.

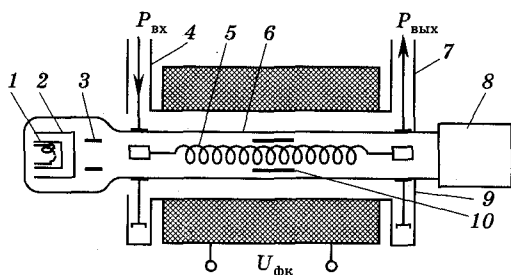


Рис. 13.9

Концы спирали выполнены в виде полых цилиндров, что позволяет обеспечить лучшую связь входа и выхода прибора с колебательной системой. Коллектор 8 собирает отработанные (провазимодействовавшие с СВЧ-полем) электроны. Основное назначение фокусирующей катушки состоит в предотвращении поперечной разгруппировки электронных сгустков под действием объемного заряда электронов. Соленоид формирует продольное, относительно оси прибора, магнитное поле, которое создает силу Лоренца, прижимающую электроны, имеющие поперечную составляющую скорости, к оси прибора. Серьезной проблемой в усилителях на ЛВВО является согласование колебательной системы с волновым сопротивлением входного фидера и сопротивлением нагрузки. При плохом согласовании часть энергии СВЧ отражается от выхода и движется в направлении ко входу, т. е. возникает обратная связь, что может вызвать автоколебания, для устранения которых и используются согласующие трансформаторы 9 и поглотитель 10.

Замедляющие системы. Для получения эффективного взаимодействия электронного потока с СВЧ-волной требуется примерное равенство фазовой скорости волны и скорости электронов (условие синхронизма), что достигается применением специальных волноведущих систем, которые называются замедляющими. Основное назначение замедляющей системы (ЗС) состоит в уменьшении фазовой скорости v_ϕ волны. При ускоряющих напряжениях в сотни и тысячи В скорость движения электронов $v_e \sim 6 \cdot 10^8 \dots 3 \cdot 10^9$ см/с, что составляет сотые и десятые доли скорости света c . Следовательно, для получения синхронизма коэффициент замедления $k_s = c/v_\phi$ должен быть в пределах 3...50. Уменьшить фазовую скорость v_ϕ можно или за счет использования в волноводах материалов с большими значениям диэлектрической и магнитной постоянных ϵ и μ , так как $v_\phi = c/\sqrt{\epsilon\mu}$, или применением специальных волноведущих (замедляющих) систем с периодическими условиями на границах, примеры устройства которых даны на рис. 13.10, 13.11, 13.12

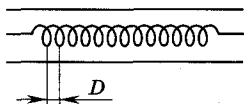


Рис. 13.10

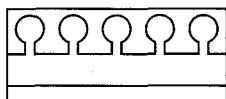


Рис. 13.11

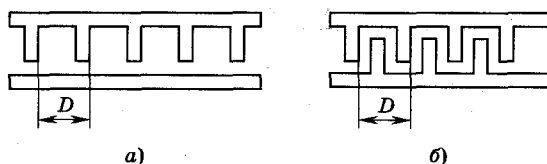


Рис. 13.12

(рис. 13.10 — спирально-коаксиальный волновод, рис. 13.11 — ЗС в виде цепочки резонаторов «щель—отверстие»; рис. 13.12, а — гребенчатая ЗС, рис. 13.12, б — встречно-штыревая ЗС).

Принцип замедления проиллюстрируем с помощью спирально-коаксиального волновода. При возбуждении такой ЗС от источника СВЧ-колебаний волна распространяется по проводнику спирали со скоростью, близкой к скорости света c . За время движения вдоль одного витка спирали длиной $2\pi r$ со скоростью c волна переместится по оси на шаг спирали D (см. рис. 13.10). Вдоль оси волна бежит с фазовой скоростью v_ϕ , следовательно, коэффициент замедления фазовой скорости в спирально-коак-

сиальном волноводе $k_3 = \frac{c}{v_\phi} = \frac{2\pi r}{D}$.

Принципы замедления в ЗС, изображенной на рис. 13.11, можно объяснить следующим образом. Любую ЗС можно представить в виде цепочки элементарных резонаторов с емкостью C_0 , сосредоточенной в основном между щелями резонатора, и с индуктивностью L_0 , которая образуется в основном отверстием. Электромагнитная волна частоты ω , поступающая в ЗС, возбуждает в резонаторах колебания с этой частотой. Амплитуда и фаза возбужденных колебаний зависят от отношения возбуждающей ω и собственной частоты резонатора $\omega_0 = 1/\sqrt{L_0 C_0}$. Амплитудная и фазовая характеристики элементарного резонатора аналогичны характеристикам резонансного контура, поэтому фазовый сдвиг вынужденных колебаний в зависимости от ω изменяется от $-\pi/2$ до $+\pi/2$. Вынужденные колебания резонаторов и возбуждающая волна, взаимодействуя друг с другом, образуют результирующую волну, которая сдвинута по фазе относительно исходной возбуждающей волны. При перемещении волны вдоль ЗС фазовый сдвиг накапливается, так как распространяющаяся волна постоянно подвергается воздействию вынужденных колебаний резонаторов, что увеличивает отставание по фазе резуль-

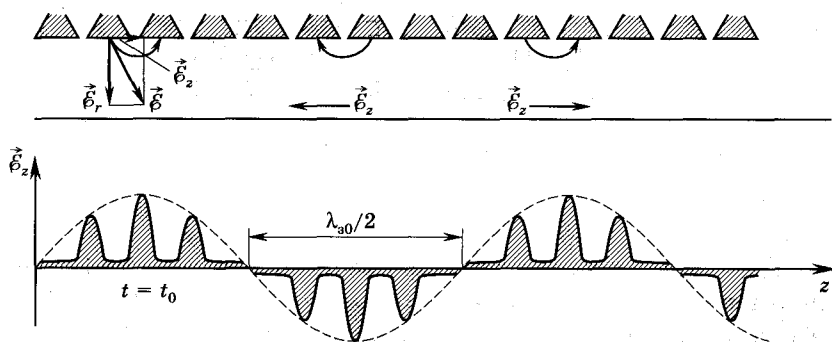


Рис. 13.13

тирующей волны от возбуждающей, а это означает, что фазовая скорость первой из них меньше фазовой скорости последней.

Пространственное распределение поля в ЗС является неоднородным. На рис. 13.13 дана зависимость одной из составляющих напряженности электрического поля, а именно продольной \vec{E}_z , от расстояния вдоль оси. Как видно из этого рисунка, распределение поля в пространстве носит периодический, но не гармонический характер, поэтому это распределение можно разложить в ряд Фурье; в результате закон изменения электрического поля волны в пространстве и во времени можно представить в виде

$$\vec{E}_z = \sum_{n=-\infty}^{\infty} a_n e^{i(\omega t - k_{zn} z)}.$$

Поле \vec{E}_z в такой форме представляет собой бесконечную последовательность (набор) бегущих волн (пространственных гармоник), с амплитудами a_n и постоянной распространения $k_{zn} = k_{z0} + \frac{2\pi n}{D}$, где k_{z0} — волновое число основной (нулевой) пространственной гармоники; λ_{z0} — длина замедленной основной волны, которая меньше длины волны λ в свободном пространстве в $\frac{c}{v_{\Phi}}$ раз. Длина волны λ_{z0} определяется огибающей периодического пространственного распределителя \vec{E}_z на рис. 13.13. Пространственные гармоники отличаются друг от друга как амплитудами, так и значениями фазовых скоростей $v_{\Phi n}$:

$$v_{\Phi n} = \frac{\omega}{k_{zn}} = \frac{\omega}{k_{z0} + 2\pi n/D}. \quad (13.12)$$

Фазовая скорость пространственных гармоник при положительных n имеет только положительные значения, а при отрицательных n — отрицательные. Положительным значениям n соответствуют прямые пространственные гармоники, а отрицательным — обратные.

Как известно, групповая скорость $v_{гр} = \frac{d\omega}{dk}$, тогда, дифференцируя выражение (13.12), можно получить соотношение для $v_{гр}$ следующего вида

$$\frac{1}{v_{гр}} = \frac{dk_{zn}}{d\omega} = \frac{d}{d\omega} \left(k_{z0} + \frac{2\pi n}{D} \right) = \frac{dk_{z0}}{d\omega}. \quad (13.13)$$

В формуле (13.13) взята обратная величина $v_{гр}$ для простоты дифференцирования. Из (13.13) видно, что вся совокупность пространственных гармоник имеет одну и ту же групповую скорость, а это означает, что по отдельности они существовать не могут. Они существуют как единая совокупность волн.

Одной из важнейших характеристик ЗС является *дисперсионная характеристика*, представляющая собой зависимость $v_{ф} = \varphi(\omega)$. Ее вид показан на рис. 13.14. Эта характеристика полностью аналогична характеристике, описывающей дисперсию света в кристаллическом твердом теле. Разница состоит только в том, что в оптическом диапазоне элементарные осцилляторы формируются кристаллической решеткой, а в ЗС элементарные осцилляторы — это резонаторы ЗС (ω_0 на рисунке — это собственная частота элементарного резонатора).

Если период замедляющей системы D много меньше длины волны λ_z , распространяющейся в ЗС, то такие ЗС называются *однородными*. В них волна как будто «скользит» по резонаторам, не замечая их; характеристики поля в двух соседних резонаторах отличаются незначительно, и фазовая скорость волны в таких системах слабо зависит от частоты (области I на рис. 13.14). Если $D \sim \lambda_z$, то такие системы называются *неоднородными*. В этом случае поля в соседних резонаторах отличаются значительно, и фазовая скорость претерпевает сильные изменения в зависимости от частоты (область II). Приборы с однородными ЗС обладают широкой полосой пропускания частот, в то время как устройства с неоднородным ЗС имеют большой диапазон электронной перестройки частоты.

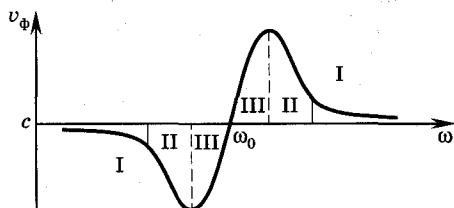


Рис. 13.14

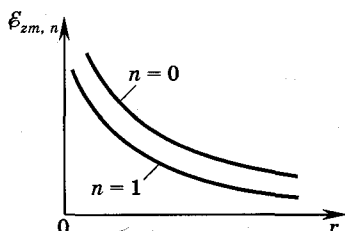


Рис. 13.15

Если фазовая скорость в функции частоты уменьшается, т. е. $\frac{dv_\phi}{d\omega} < 0$, то такая дисперсия называется нормальной (области I и II на рис. 13.14), а если возрастает, т. е. $\frac{dv_\phi}{d\omega} > 0$, то дисперсия аномальна (область III на рис. 13.14). В большинстве практических случаев необходимо, чтобы v_ϕ была меньше скорости света, поэтому в основном используется нижняя часть кривой рис. 13.14. Одним из важнейших параметров ЗС является сопротивление связи $R_{св}$, которое дает связь между напряженностью электрического поля и мощностью волны. Сопротивление $R_{св}$ определяет эффективность взаимодействия электронного потока с бегущей волной и вычисляется по формуле

$$R_{св} = \frac{\varepsilon_{zm, n}^2}{2k_{zn}P}, \quad (13.14)$$

где $\varepsilon_{zm, n}$ — амплитуда напряженности продольного поля пространственной n -й гармоники, P — мощность электромагнитной волны. Во всех ЭВП СВЧ наиболее важную роль играет продольная составляющая ε_{zn} электрического поля той пространственной гармоники, для которой обеспечивается условие синхронизма $v_e \sim v_{\phi n}$. Составляющая ε_{zn} при удалении от поверхности ЗС уменьшается (рис. 13.15, где r — текущее расстояние от поверхности ЗС). Следовательно, сопротивление связи $R_{св}$ будет иметь максимальные значения около поверхности и уменьшаться при удалении от нее, что важно при взаимодействии электронного потока и волны. Наиболее эффективно это взаимодействие будет при движении электронов на возможно близких расстояниях от поверхности ЗС.

13.4. Физические процессы в ЛБВО

Рассмотрим физические процессы в ЛБВО. Положим, что на вход ЛБВО подана СВЧ-волна с частотой ω , которая начинает распространяться вдоль ЗС, при этом длина волны в замедляющей системе λ_z меньше длины волны в свободном пространстве. Электронный поток бежит со скоростью v_e в том же направлении, что и волна. В той части ЗС, где векторы $\vec{\mathcal{E}}_z$ и \vec{v}_e имеют одно и то же направление, электроны тормозятся, а там, где эти векторы противоположно направлены, происходит ускорение электронов. В результате осуществляется модуляция электронного потока по скорости, которая приводит к образованию сгустков, причем в отличие от клистронов процесс формирования сгустков осуществляется при совместном движении электронов и СВЧ-волны и их непрерывном взаимодействии. Если первоначальная невозмущенная скорость электронов v_0 будет равна фазовой скорости волны v_ϕ , то сгустки, показанные на рис. 13.16 в виде темных прямоугольников, окончательно сформируются в области, когда потенциал СВЧ-поля растет и проходит через нуль, что отражено на пространственно-временной диаграмме рис. 13.16, а. Здесь z_1 определяет смещение электронов относительно волны в отсутствие модуляции по скорости, т. е. когда величина \mathcal{E}_z равна нулю. Поскольку движение электронов рассматривается в движущейся с фазовой скоростью $v_c = v_\phi$ системе координат, то при $v = 0$ траектории будут прямыми линиями (штриховые на рис. 13.16).

При наличии модуляции по скорости ($\mathcal{E}_z \neq 0$), электроны будут дополнительно ускоряться или замедляться, смещаясь на величину

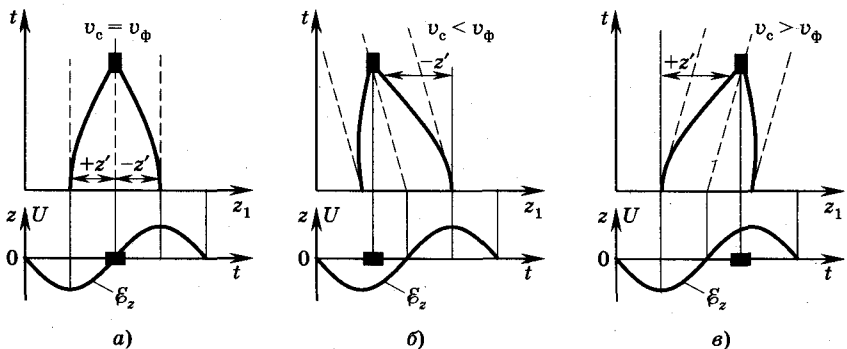


Рис. 13.16

ну z' по сравнению со случаем, когда $\mathcal{E}_z = 0$ (см. рис. 13.16, а, б, в). Если $v_e < v_\phi = v_k$, то сгустки сформируются в ускоряющем поле СВЧ-волны (рис. 13.16, б), а при $v_e > v_\phi$ — в тормозящем (рис. 13.16, в). Рабочим режимом ЛБВО является режим, когда $v_e > v_\phi$. В этом случае сгустки, двигаясь совместно с волной, все время тормозятся и передают часть своей кинетической энергии волне. Процесс передачи энергии от электронного потока будет продолжаться до тех пор, пока электроны не затормозятся волной до скорости, равной фазовой скорости СВЧ-волны.

Количественно процессы совместного движения электронов и волны характеризуются абсолютным и относительным углами пролета. **Абсолютный угол пролета** θ_e определяет движение электрона в неподвижной системе координат безотносительно к движению волны:

$$\theta_e = \frac{\omega Z}{v_e}, \quad (13.15)$$

где v_e — скорость электронов; z — расстояния от входа системы до рассматриваемой плоскости z в пространстве взаимодействия. При $z = l$, где l — расстояние между входом и выходом прибора, угол $\theta_e = 2\pi t/T$ определяет время движения электрона, выраженное в долях периода колебаний на всем протяжении пространства взаимодействия от входа до выхода.

В случае модуляции по скорости ($\mathcal{E}_z, v \neq 0$) электроны будут дополнительно ускоряться или замедляться, смещаясь на величину z' по сравнению со случаем, когда $\mathcal{E}_z, v = 0$ (см. рис. 13.16, а, б, в), т. е.

$$v_e = v_0 + \tilde{v}. \quad (13.16)$$

Здесь \tilde{v} — возмущенная СВЧ-полем часть скорости электронов,

$$v_0 = \sqrt{\frac{2e}{m} U_{\text{сп}}} — \text{невозмущенная скорость электронов,} \quad (13.17)$$

где $U_{\text{сп}}$ — постоянный потенциал либо спирали, либо последнего электрода электронной пушки, поскольку чаще всего спираль электрически соединена с ним. Следовательно, θ_e можно записать в виде

$$\theta_e = \theta_0 + \tilde{\theta}, \quad (13.18)$$

где $\theta_0 = \frac{\omega z}{v_0}$ — абсолютный угол пролета невозмущенного электрона, а $\tilde{\theta}$ — изменение угла пролета вследствие модуляции по скорости. В результате модуляции угла пролета θ_e и образуются сгустки. В отличие от клистронов процесс образования сгустков в ЛБВО происходит при непрерывном взаимодействии с волной, поэтому целесообразно вести рассмотрение в подвижной системе координат, которая перемещается со скоростью волны. При анализе работы ЛБВО удобно воспользоваться *относительным углом пролета* (φ_e) электронов, который определяет смещение электронов относительно волны в процессе их совместного движения, т. е.

$$\varphi_e = \omega z \left(\frac{1}{v_e} - \frac{1}{v_\phi} \right). \quad (13.19)$$

Относительный угол пролета, как и абсолютный, содержит постоянную и переменную составляющие, т. е. $\varphi_e = \varphi_0 + \tilde{\varphi}$. Наиболее эффективная передача энергии от электронов полю достигается в том случае, когда электронный сгусток во время взаимодействия остается в тормозящем поле волны. Формируясь в начале тормозящего полупериода, сгусток за время совместного движения с волной смещается к концу этого полупериода, т. е. $\varphi_0 = \pi$. Выходная мощность при $\varphi_0 = \pi$ оказывается максимальной.

13.5. Параметры и характеристики ЛБВО

В заключение рассмотрим параметры и характеристики ЛБВО. Одним из основных параметров является *коэффициент усиления* $K = \frac{P_{\text{вых}}}{P_{\text{вх}}}$.

Амплитуды $\mathcal{E}_{zm,0}$ продольной составляющей волны на входе и $\mathcal{E}_{zm,l}$ на выходе ЗС связаны между собой соотношением

$$\mathcal{E}_{zm,l} = \mathcal{E}_{zm,0} e^{\gamma l}, \quad (13.20)$$

где l — длина ЗС, γ — параметр усиления.

Выходная мощность однозначно определяется входной амплитудой

$$P_{\text{вх}} = \frac{\mathcal{E}_{zm0}^2 G_{\text{вх}}}{2}, \quad (13.21)$$

а выходная мощность

$$P_{\text{вых}} = \frac{\varepsilon_{zm}^2 G_{\text{вых}}}{2}, \quad (13.22)$$

где $G_{\text{вых}}$ и $G_{\text{вх}}$ — соответственно выходная и входная проводимости.

При полном согласовании ЗС с выходной и входной СВЧ-линиями $G_{\text{вых}} = G_{\text{вх}}$, поэтому коэффициент усиления по мощности $K = P_{\text{вых}}/P_{\text{вх}}$ с учетом соотношений (13.20)—(13.22) равен $K = \exp(2\gamma l)$ и в децибелах выражается соотношением

$$K = 10 \lg \frac{P_{\text{вых}}}{P_{\text{вх}}} = 10 \lg (e^{2\gamma l}) = 8,68\gamma l, \quad (13.23)$$

где параметры усиления $\gamma = f(C; N)$; $C = \sqrt[3]{\frac{R_{\text{св}} I_0}{4U_0}}$; $N = \frac{l}{\lambda_3}$; $R_{\text{св}}$ — сопротивление связи; I_0 — ток электронного пучка; U_0 — напряжение ускорения электронов; λ_3 — длина замедленной волны.

Если использовать параметры C и N , то, как показывает теоретический анализ, коэффициент усиления $K \sim CN$. Реальные значения $N \approx 10 \dots 30$, $C \approx 0,02 \dots 0,5$, а $K = 20 \dots 50$ дБ.

Электронный коэффициент полезного действия ЛБВО определяется соотношением $\eta_e = P_{\text{вых}}/(U_0 I_0)$. Мощность $P_{\text{вых}}$ пропорциональна уменьшению кинетической энергии электронов в результате взаимодействия с волной. Если принять, что конечная скорость электрона равна фазовой скорости волны, то изменение кинетической энергии электрона, обусловленное взаимодействием его с волной, будет равно $\Delta W_{\text{кин}} = (m v_0^2 / 2) - (m v_{\phi}^2 / 2)$. Оценки показывают, что $\Delta W_{\text{кин}} \approx C(m v_0^2 / 2)$, где C — параметр усиления, $m v_0^2 / 2 = W_0$ — энергия, которая затрачивается на ускорение электронного потока, т. е. $P_0 = U_0 I_0 \sim W_0$. Тогда электронный КПД $\eta_e = \Delta W_{\text{кин}} / W_0 \approx C$. Для увеличения КПД в ЛБВО применяют такие ЗС, в которых v_{ϕ} уменьшается по мере замедления электрона. В результате время пребывания сгустков в тормозящем поле волны увеличивается, что и позволяет получить дополнительное усиление и большую выходную мощность. ЗС, в которых v_{ϕ} уменьшается вдоль системы, называются *изохронными*.

Существует еще ряд способов увеличения КПД. В специально разработанных ЛБВО КПД может быть доведен до 50...60%,

а выходная мощность — до нескольких десятков кВт в непрерывном режиме.

Одним из достоинств усилителей на ЛБВО является возможность получения низкого уровня шума, в частности путем охлаждения спирали, снижения тока вторичных электронов из коллектора на спираль, использования малошумящих электронных пушек. В малошумящих ЛБВО коэффициент шума (см. гл. 22) принимает значения

$$K_{\text{ш}} = 2,5 \dots 20 \text{ (т. е. } 4 \dots 13 \text{ дБ)}.$$

Рассмотрим основные характеристики ЛБВО. На рис. 13.17 изображена зависимость выходной мощности $P_{\text{вых}}(I_{\text{ф.к}})$ от тока $I_{\text{ф.к}}$ в фокусирующей катушке $P_{\text{вх}} = \text{const}$, $\omega = \text{const}$ и $U_0 = U_{0\text{ опт}} = \text{const}$. Фокусирующая катушка в ЛБВО предназначена в основном для предотвращения поперечной расфокусировки электронных сгустков, происходящей под действием объемного заряда электронов в сгустках. (В современных ЛБВО вместо соленоида в основном применяются системы постоянных магнитов.) Следовательно, изменение тока $I_{\text{ф.к}}$ в фокусирующей катушке приводит к изменению напряженности продольного магнитного поля H . При малых токах, протекающих через катушку, напряженность магнитного поля, препятствующая уходу электронов на ЗС, недостаточна, и значительная часть электронов попадает на ЗС, т. е. не участвует в процессе преобразования энергии, поэтому выходная мощность мала. По мере увеличения тока $I_{\text{ф.к}}$ возрастает удерживающая со стороны магнитного поля сила, и все большее число электронов передает энергию бегущей СВЧ-волне. В результате выходная СВЧ-мощность возрастает. При больших напряженностях магнитного поля (большие токи $I_{\text{ф.к}}$) выходная мощность падает из-за уменьшения сопротивления связи (электроны «прижимаются» полем ближе к оси ЛБВО и взаимодействуют с продольным электрическим СВЧ-полем меньшей напряженности).

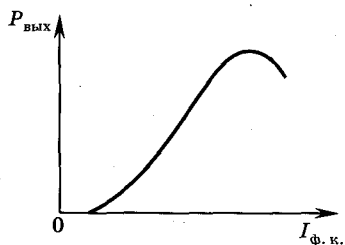


Рис. 13.17

Характеристика взаимодействия
 $P_{\text{вых}}(U_0)$ при $P_{\text{вх}} = \text{const}$, $I_{\text{ф.к.}} = I_{\text{ф.к. опт}} = \text{const}$, $\omega = \text{const}$ и при

постоянных напряжениях на остальных электродах показана на рис. 13.18.

При изменении напряжения на спирали изменяется скорость электронов, поступающих в ЗС, где распространяется электромагнитная волна. Доля кинетической энергии, передаваемой электронным потоком волне, зависит от разницы в скоростях v_ϕ и v_e , т. е. от относительного угла пролета $\varphi = \omega t \left(\frac{1}{v_\phi} - \frac{1}{v_e} \right)$. При $\varphi_0 = \pi$ усредненная за период энергия, передаваемая волне от электронов, максимальна. В этом случае электронный сгусток находится все время в тормозящем поле волны и отдает максимум энергии. При увеличении v_e (по сравнению со случаем $\varphi_0 = \pi$) угол пролета φ возрастает. В этом случае уменьшается коэффициент модуляции электронного потока по скорости. Электроны группируются в сгустки дальше от входа ЛБВО (ближе к выходу), и сгусток вблизи выхода оказывается в ускоряющем полупериоде и отбирает у волны часть энергии.

При уменьшении v_e (по сравнению со случаем $\varphi_0 = \pi$) уменьшается относительный угол пролета, а коэффициент модуляции по скорости возрастает, что может приводить к разгруппировке электронного потока, особенно на выходе ЛБВО (см. далее объяснение амплитудной характеристики на рис. 13.19). При некотором значении $v_e < v_\phi$ мощность падает до нуля, в этом случае электроны формируются в сгустки в области ускоряющего поля волны, т. е. они ускоряются волной и, следовательно, отбирают энергию у нее.

Амплитудная характеристика $P_{\text{вых}}(P_{\text{вх}})$ при $\omega = \text{const}$, $U_0 = U_{0 \text{ опт}} = \text{const}$, $I_{\text{ф. к. опт}} = \text{const}$ и постоянных напряжениях на остальных электродах изображена на рис. 13.19.

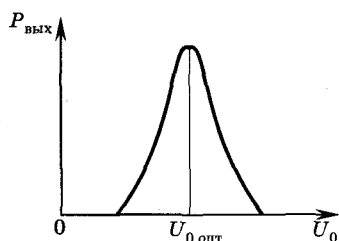


Рис. 13.18

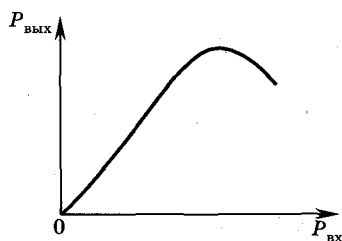


Рис. 13.19

При увеличении амплитуды входного СВЧ-сигнала возрастает коэффициент модуляции электронного потока по скорости, вследствие чего электронные сгустки формируются ближе ко входу замедляющей системы и, следовательно, возрастает расстояние, на котором сформированный сгусток взаимодействует с волной. Доля передаваемой кинетической энергии от сгустка волне увеличивается, что приводит к росту выходной мощности $P_{\text{вых}}$. При дальнейшем росте входной мощности $P_{\text{вх}}$ начинает уменьшаться, что может быть связано с перегруппировкой электронного сгустка из-за большой разности скоростей электронов, образующих сгустки. Электроны, образующие сгустки, имеют разные скорости, и при большой амплитуде входного сигнала это различие может быть значительным, что и вызывает продольную разгруппировку сгустков. Часть электронов сгустка при приближении к выходу ЛВВО может попадать в ускоряющую область СВЧ-волны и забирать у нее энергию, что и вызывает уменьшение $P_{\text{вых}}$.

13.6. Лампы обратной волны типа О

ЛОВО используются как для генерирования СВЧ-колебаний, так и в качестве регенеративных усилителей. Эти приборы наиболее широко применяются в диапазонах от дециметрового до субмиллиметрового включительно.

Генераторы на этих приборах отличаются большим диапазоном электронной настройки частоты. Схематическое устройство усилительной и генераторной ЛОВО показано на рис. 13.20, *а* и *б* соответственно. Катод *К* и управляющий электрод УЭ условно изображают электронную пушку. В ЛОВО обычно используются неоднородные замедляющие системы (на рисунке заштрихованы). Коллектор «*Кол*» выполняет те же функции, что и в ЛВВО. Поглотитель *П* обеспечивает необходимое распределение СВЧ-поля в ЗС.

Направление групповой скорости переноса энергии в ЛОВО в отличие от ЛВВО противоположно направлению скорости электронов v_e , а фазовая скорость обратной гармоники совпадает с вектором v_e . При выполнении условия синхронизма для обратной гармоники ($v_e \approx v_{\phi}$) можно обеспечить ее взаимодействие с электронным потоком таким же образом, как в ЛВВО.

В усилительной ЛОВО сигнал подается от внешнего генератора со стороны коллекторного конца. В генераторной лампе в

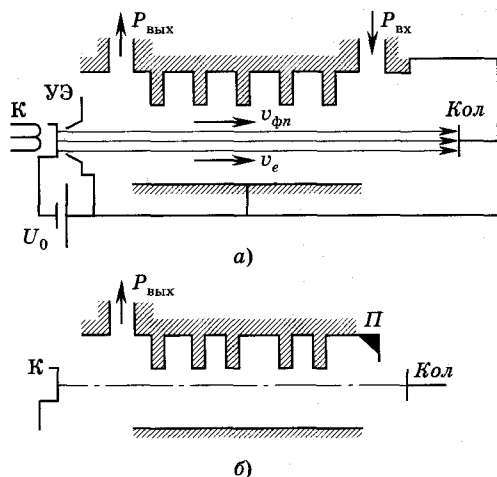


Рис. 13.20

результате флуктуации электронного потока наводятся слабые колебания. Распространяющиеся по ЗС волны представляют сумму пространственных гармоник. В ЛОВО условие синхронизма выполняется, как правило, для первой обратной гармоники, что обеспечивается конструкцией ЗС, подбором ее параметров и изменением электрических режимов прибора. Взаимодействие электронов с полем обратной гармоники, возбужденной или внешним сигналом в усилителе, или же за счет флуктуации в генераторах, приводит к их группировке. Сгруппированные электроны, как и в ЛБВО, передают часть своей кинетической энергии СВЧ-полю. Использование электронного потока, движущегося навстречу потоку электромагнитной энергии, обеспечивает положительную обратную связь, без которой невозможна работа автогенератора. Условным входом генератора на ЛОВО, так же как и в усилителе, можно считать коллекторный конец ЗС; так как энергия переносится волной от коллектора к выходу, расположенному со стороны электронной пушки. Для обеспечения возбуждения необходимо выполнить условие баланса фаз, которое для ЛОВО записывается в следующем виде:

$$\frac{\omega l}{v_{\phi(-1)}} - \frac{\omega l}{v_0} - \pi = 2\pi n, \quad (13.24)$$

где $n = 0, 1, 2, 3$, v_0 — скорость электронного потока в отсутствие модуляции.

Первое слагаемое в этом выражении определяет изменение фазы волны при прохождении расстояния l , равного длине ЗС, второе слагаемое — изменение фазы за время движения электронов в ЗС. (Знак «минус» в обозначении $U_{\phi(-)}$ в формуле (13.24) появляется из-за противоположного направления скоростей v_{ϕ} и v_0 .) Слагаемое π определяет оптимальные условия для передачи энергии от потока полю, когда образовавшийся сгусток электронов не выходит из тормозящего СВЧ-поля.

Первые два члена в выражении для баланса фаз, по сути дела, определяют относительный угол пролета ϕ_0 , и при $n = 0$ он будет равен π , т. е. это условие максимального отбора волной энергии от электронного потока, когда сгусток электронов на протяжении всего совместного движения находится в максимально тормозящем поле, усредненном по полупериоду волны.

При $n = 1$ относительный угол пролета равен 3π , т. е. электроны сгустка в течение двух полупериодов тормозятся и в течение одного — ускоряются или, иначе, при пролете одного пространственного периода ($\phi_0 = 2\pi$) усредненная за период энергия взаимодействия равна нулю, а в течение полупериода π электроны тормозятся. Для $n = 2$ относительный угол пролета равен $\phi_0 = 5\pi$ и обмен энергией электронов и волны происходит на протяжении $2^{1/2}$ периодов волны. В течение двух полных периодов баланс энергии обмена равен нулю. Результатом является то, что опять на протяжении оставшегося полупериода происходит преобладающая передача энергии от электронов волне и т. д.

Величину n называют порядком или номером зоны колебаний в ЛОВО. Наибольшая выходная мощность получается в основной зоне при $n = 0$. Условие баланса фаз можно использовать для определения частоты генерируемых колебаний в различных зонах, если известна дисперсионная характеристика обратной пространственной гармоники. Частота, как следует из условия баланса фаз, в выбранной зоне зависит от v_e , а следовательно,

и от U_0 , так как $v_e = \sqrt{\frac{2e}{m} U_0}$. Зависимости колебательной мощности $P_k = \phi_1(U_0)$ и частоты $\omega = \phi_2(U_0)$ генерируемых колебаний от ускоряющего напряжения при $I_{\phi.k.} = I_{\phi.k. \text{ опт}} = \text{const}$ представлены на рис. 13.21, а и б.

Скорость электронов v_e , взаимодействующих с обратной гармоникой электромагнитной СВЧ-волны, определяется ускоряю-

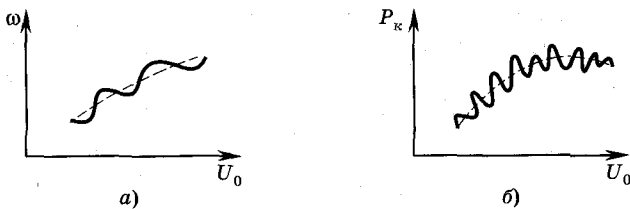


Рис. 13.21

щим напряжением U_0 , т. е. $v_e = \sqrt{\frac{2e}{m} U_0}$. Замедляющие системы в ЛОВО обладают сильной дисперсией (зависимостью фазовой скорости электромагнитной волны от частоты). За счет флуктуаций плотности электронного потока, вызванных тепловым разбросом скоростей электронов, которые эмитированы катодом, в замедляющей системе наводится широкий спектр электромагнитных флуктуаций. Изменяя скорость электронного потока, добиваются оптимальных условий взаимодействия с обратными гармониками разных частот ($v_e \approx v_{\Phi(-1)}$). Увеличение скорости электронов соответствует взаимодействию с обратной гармоникой, имеющей большую фазовую скорость и, соответственно, большую частоту. Таким образом, увеличение U_0 приводит к увеличению частоты генерируемых колебаний, и при снятии зависимости колебательной мощности от ускоряющего напряжения $P_k(U_0)$ происходит одновременное изменение частоты, т. е. $P_k = F(\omega)$.

При работе на низких частотах (малые значения U_0) длина замедленной волны λ_3 большая, и поэтому электронный поток разбивается на относительно малое число сгустков в пространстве взаимодействия (по длине спирали укладывается малое число длин волн, т. е. отношение $N = \frac{l}{\lambda_3}$ невелико). В силу этого из-за уменьшения расстояния, на котором сформированный сгусток взаимодействует с волной, ухудшаются условия передачи энергии от электронного потока волне, что усугубляется к тому же возрастанием объемного заряда в сгустках (из-за большого числа электронов в каждом сгустке). Последнее обстоятельство приводит к увеличению поперечной и продольной разгруппировок сгустков. Поэтому на низких частотах наблюдается «завал» характеристики $P_k(U_0)$.

При больших U_0 (большие частоты) уменьшается сопротивление связи электронного потока с полем СВЧ-волны. В этом случае поле концентрируется в основном ближе к поверхности ЗС, и на границе электронного потока на более высоких частотах существует продольное электрическое СВЧ-поле меньшей напряженности, что и является причиной уменьшения мощности.

Рассмотрим баланс мощностей. Получаемая волной от электронного потока энергия должна быть больше энергии потерь (затухание в ЗС, потери согласования и т. д.), т. е. ее величина должна быть достаточной для поддержания процесса формирования электронных сгустков. Расчет генератора на ЛОВО показывает, что в режиме самовозбуждения ток электронного пучка должен превышать некоторую величину, называемую пусковым током. Величина пускового тока должна быть разной для возбуждения колебаний различных порядков, так как с увеличением n требуются большие величины U_0 , т. е. большие затраты подводимой от источника постоянного тока энергии при мало изменяемой величине полезной (а в действительности при уменьшении), передаваемой от потока волне. Таким образом, КПД генератора уменьшается при возрастании n .

Для возбуждения колебаний при $n = 1$ величина пускового тока должна быть в несколько раз больше, чем при $n = 0$. Обычно в генераторах на ЛОВО реализуются режимы для $n = 0$. Для предотвращения самовозбуждения колебаний высших порядков длина ЗС выбирается равной $l = N\lambda_s$.



Контрольные вопросы

1. Каковы классификация и области применения пролетных и отражательных клистронов?
2. Устройство, характеристика элементов и принципы работы двухрезонаторных клистронов (ДРК). Объемные резонаторы и их параметры.
3. Пространственно-временные диаграммы, конвекционные и наведенные токи в двухрезонаторных пролетных клистронов.
4. Параметры и характеристики ДРК.
5. Особенности и принципы работы отражательных клистронов. Зоны генерации. Электронная настройка частоты.

6. Устройство ЛБВО. Характеристики и выполняемые функции элементов ЛБВО. Устройство, параметры и характеристики замедляющих систем.
7. Параметры и характеристики ЛБВО.
8. ЛОВО: устройство, принципы взаимодействия электронов с обратной волной, параметры и характеристики.

Глава 14

ЭЛЕКТРОННЫЕ ПРИБОРЫ СВЧ С ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ И СКРЕЩЕННЫМИ ПОЛЯМИ — ПРИБОРЫ ТИПА М

14.1. Общие сведения

Основные различия между механизмами взаимодействия в приборах типа О и М заключаются в следующем.

1. В приборах типа О передача энергии от луча к волне осуществляется только тогда, когда скорость электронов больше фазовой скорости возмущенной волны. В приборах типа М передача энергии осуществляется при равенстве средней скорости электронов и скорости волны.
2. В приборах типа О процесс передачи энергии от электронного потока СВЧ-волне продолжается до тех пор, пока электроны не замедлятся полем волны до скорости, равной фазовой скорости волны. Поскольку начальное превышение скорости электронов над скоростью волны обычно мало, то волне передается лишь небольшая часть кинетической энергии пучка, и КПД, как правило, невелик. В приборах типа М в энергию волны переходит потенциальная энергия, причем процесс взаимодействия осуществляется при неизменной средней скорости электронов. В энергию СВЧ-поля переходит вся потенциальная энергия электрона за исключением небольшой части, необходимой для создания синхронной скорости. С ростом анодного напряжения относительная доля энергии, необходимая для создания синхронной скорости, уменьшается, и КПД растет, достигая 80...90%.

3. В приборах типа М в отличие от приборов типа О переменная компонента объемного заряда ρ практически не играет роли в процессе преобразования энергии. В приборах типа О влияние объемного заряда может быть весьма значительно, что сильно снижает эффективность преобразования энергии.

14.2. Магнетроны. Движение электронов в скрещенных полях

В магнетронах процесс преобразования потенциальной энергии электронов происходит в скрещенных электрических и магнитных полях.

Магнетроны используются в качестве узкополосных генераторов. Эти приборы могут работать в диапазоне от метровых до миллиметровых длин волн и генерировать СВЧ-энергию до 100 кВт в непрерывном и до десятков и даже сотен мегаватт в импульсном режиме.

Основными достоинствами этих приборов являются следующие: высокие значения КПД (до 70% и более), сравнительно невысокий уровень питающих напряжений, достаточно большие мощности как в непрерывном, так и в импульсном режимах, высокая эффективность на весьма коротких длинах волн (миллиметровый диапазон длин волн), небольшие габариты и масса.

В настоящее время используются преимущественно многорезонаторные магнетроны. Устройство многорезонаторного магнетрона представлено на рис. 14.1.

Магнетрон состоит из катода 1, эмитирующего электроны; анодного блока 2, который объединяет функции электрода, собирающего отработанные электроны, и колебательной системы, состоящей из замкнутой цепочки резонаторов 3; магнитной системы 4, которая условно отмечена кружочком с точкой (вектор индукции магнитного поля направлен из плоскости чертежа на наблюдателя); устройства вывода энергии (магнитной петли) 5. Магнитная система выполняется на основе постоянных магнитов или магнитных катушек (электромагнитов).

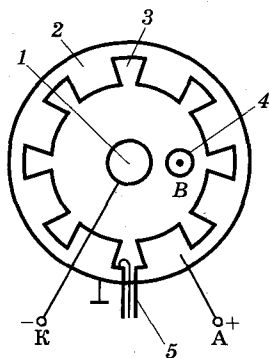


Рис. 14.1

Анодный блок выполняется в виде медного или латунного диска с прорезя-

ми для формирования резонаторов. Из-за больших размеров и трудности изоляции анод обычно заземляется, а на катод подается большой отрицательный потенциал через специальный элемент конструкции, позволяющий изолировать катод. Конструкция магнетрона, как правило, аксиально симметрична.

Движение электронов в пространстве взаимодействия магнетрона происходит в высокочастотных скрещенных постоянных электрических и магнитных полях по сложным траекториям. Структура и свойства СВЧ-полей в пространстве взаимодействия также весьма сложны, поэтому, прежде чем рассматривать принцип действия магнетрона, рассмотрим по отдельности движение электронов в скрещенных постоянных электрическом и магнитном полях при отсутствии СВЧ-полей, а потом свойства колебательной системы, т. е. структуру СВЧ-полей в пространстве взаимодействия. Если не учитывать переменные поля, то магнетрон можно рассматривать как вакуумный диод, помещенный в магнитное поле.

Хотя магнетрон выполнен в виде цилиндрической конструкции (рис. 14.2, а), движение электронов для простоты будем рассматривать для плоской конфигурации (модели) (рис. 14.2, б). Такое приближение достаточно корректно для магнетронов, у которых радиусы катода и анода не сильно отличаются.

Сделаем следующие допущения: существует постоянное электрическое поле напряженностью $\vec{\epsilon}_0 = \epsilon_y$, имеющее единственную составляющую, направленную по оси y , и постоянное магнитное поле, направленное вдоль оси z ($B_0 = B_z$, см. рис. 14.2, б); отсутствует влияние пространственного (объемного) заряда элект-

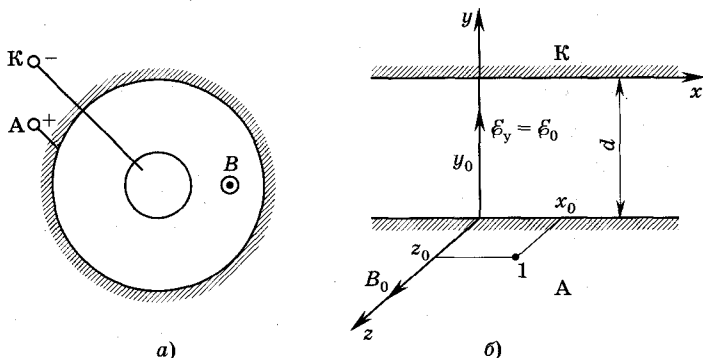


Рис. 14.2

ронов; начальные скорости электронов приняты равными нулю: $v_{x0} = v_{y0} = v_{z0} = 0$; релятивистские поправки не учитываются. В этих условиях необходимо определить (вычислить) траекторию электронов. Запишем уравнение движения электрического заряда в присутствии электрического и магнитного полей, т. е. под действием силы Лоренца:

$$m \frac{d\vec{v}}{dt} = -e\{\vec{\mathcal{E}}_0 + [\vec{v} \times \vec{B}_0]\}, \quad (14.1)$$

где m , e — соответственно масса и абсолютная величина заряда электрона, \vec{v} — вектор его скорости.

Раскладывая векторы \vec{v}_0 , $\vec{\mathcal{E}}_0$, \vec{B}_0 и $[\vec{v} \times \vec{B}_0]$ по ортам и учитывая, что $v_x = \frac{dx}{dt}$; $v_y = \frac{dy}{dt}$; $v_z = \frac{dz}{dt}$, получаем в декартовой системе координат три скалярных уравнения движения электрона, решение которых с соответствующими начальными условиями, сформулированными выше, позволяет вычислить траекторию движения электрона (при $x(t=0) = x_0$; $y(t=0) = y_0 = 0$ и $z(t=0) = z_0$).

При нахождении электрона в точке 1 (рис. 14.2), когда $v_{x0} = v_{y0} = v_{z0} = 0$, $y_0 = 0$ (x_0, y_0, z_0 — начальные координаты по осям x, y, z) и при сделанных ранее допущениях о направлении полей \mathcal{E} и \vec{B} , решение уравнения (14.1) имеет вид:

$$x = x_0 + \frac{\mathcal{E}_0}{B_0} t - \frac{m\mathcal{E}_0}{eB_0^2} \sin \omega_{\text{ц}} t, \quad (14.2)$$

$$y = -\frac{m\mathcal{E}_0}{eB_0^2} (1 - \cos \omega_{\text{ц}} t), \quad (14.3)$$

$$z = z_0, \quad (14.4)$$

где $\omega_{\text{ц}} = \frac{e}{m} B_0$ — циклотронная частота.

Соотношения (14.2), (14.3), (14.4) описывают траектории движения электронов. Эти траектории являются кривыми, лежащими в плоскости $z = z_0 = \text{const}$. Начальная координата x_0 только сдвигает траекторию вдоль оси z . Уравнения (14.2) и (14.3) описывают кривую, которая называется *циклоидой*. Каноническая форма уравнений, описывающих циклоиду, имеет вид

$$x = R(\varphi - \sin \varphi); \quad (14.5)$$

$$y = R(1 - \cos \varphi). \quad (14.6)$$

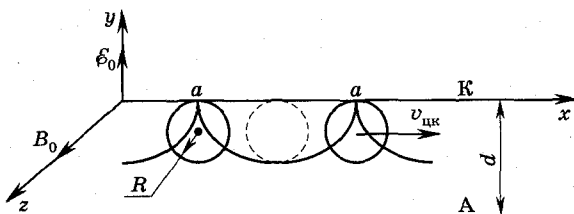


Рис. 14.3

Циклоида — это кривая, которую описывает точка, расположенная на ободу круга (например, на рис. 14.3 точка a) при качении последнего без скольжения по плоскости (на рис. 14.3 начало координат $x_0 = y_0 = 0$ перенесено на катод). В нашем случае радиус

окружности $R = \frac{m\varepsilon_0}{eB_0^2}$; $\varphi = \omega_{\text{ц}}t$ — угол поворота круга. Скорость перемещения центра круга $v_{\text{цк}} = \varepsilon_0/B_0$. Циклотронная частота $\omega_{\text{ц}}$ определяет угловую скорость качения круга и, следовательно, угловую скорость движения электрона по циклоидальной орбите.

Таким образом, электроны, вылетая с катода, будут двигаться в пространстве взаимодействия между катодом и анодом по циклоидальной кривой. Изменяя напряженность ε_0 электрического и индукцию B_0 магнитного полей, можно изменить радиус циклоиды и, следовательно, параметры траектории движения. При $R = \frac{d}{2}$ (d — расстояние между катодом и анодом, см. рис. 14.3) электроны будут касаться поверхности анода в вершине циклоиды; при $R > \frac{d}{2}$ электроны попадают на анод, и магнетрон в этом случае ведет себя как обыкновенный вакуумный анод.

Магнитное поле, при котором прекращается анодный ток, называется критическим. Индукцию критического поля $B_{\text{кр}}$ можно определить из условия $R = d/2$.

Используем полученное ранее выражение для радиуса окружности $R = \frac{m\varepsilon_0}{eB_0^2}$ и учтем, что поле ε между анодом и катодом

можно считать однородным. Тогда $\varepsilon_0 = \frac{U_0}{d}$, где U_a — потенциал анода, и уравнение для нахождения $B_{\text{кр}}$ имеет вид

$$\frac{mU_a}{edB_{\text{кр}}^2} = \frac{d}{2}, \quad (14.7)$$

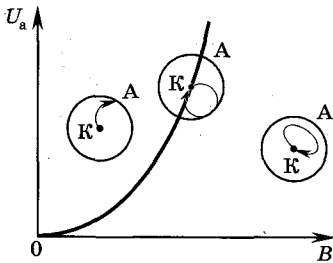


Рис. 14.4

откуда следует

$$B_{\text{кр}} = \frac{1}{d} \sqrt{\frac{2m}{e}} U_a.$$

Аналогично можно получить следующее выражение:

$$U_a = U_{\text{кр}} = eB^2 d^2 / (2m).$$

Это выражение описывает так называемую *параболу критического режима*, которая изображена на рис. 14.4, где также представлены траектории электронов для различных значений U_a и B . Левее параболы располагаются траектории электронов для значений U_a и B , соответствующих диодному режиму, а правее находятся траектории для U_a и B , соответствующих магнетронным режимам. Конструкция магнетронов является цилиндрической и рассмотренное приближение для плоского случая может оказаться достаточно грубым, тогда используют цилиндрические координаты. В цилиндрических координатах циклоида превращается в кривую, называемую эпициклоидой.

14.3. Колебательная система магнетрона

Анодный блок магнетрона как колебательная СВЧ-система представляет собой N связанных резонаторов, свернутых (замкнутых) в кольцо. В такой колебательной системе может возникнуть N видов колебаний. В колебательной системе из N резонаторов для возникновения колебаний необходим полный фазовый сдвиг СВЧ-колебаний при обходе всей развертки анода с возвращением в исходную точку, равный

$$\Phi = 2\pi n, \quad (14.8)$$

где $n = 0, 1, 2, 3, \dots, N/2$.

Целое число n определяет количество длин волн, укладываемых вдоль развертки анодного блока. В противном случае, т. е. при n нецелом, возникает интерференция волн и самовозбуждения магнетрона не происходит.

Разность фаз СВЧ-колебаний в соседних резонаторах (фазовый сдвиг) определяется очевидным соотношением:

$$\phi = 2\pi n / N. \quad (14.9)$$

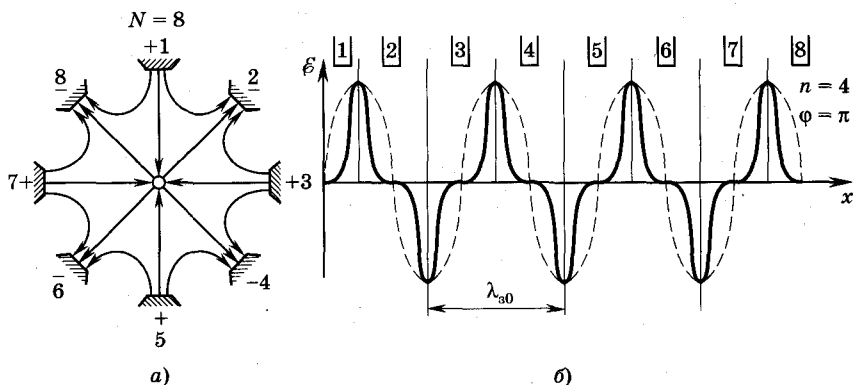


Рис. 14.5

На рис. 14.5, а представлена структура электрических СВЧ-полей для $N=8$ и $n=4$ ($\phi=\pi$).

Таким образом, различные виды колебаний в магнетроне можно характеризовать либо числом n , которое определяет номер вида колебаний, либо фазовым сдвигом ϕ .

Одним из наиболее эффективных видов колебаний является π -вид, для которого $\phi=\pi$. Из формулы (14.9) видно, что колебания вида π реализуются при $n = \frac{N}{2}$. Это возможно при четном числе резонаторов N .

Распределение напряженности переменного электрического поля вдоль анода для $\phi=\pi$ при $N=8$ дано на рис. 14.5, б. Электрическое поле в основном сосредоточено вблизи щелей резонаторов. Если построить огибающую распределения $\epsilon=f(x)$, то получим четыре периода, т. е. четыре длины волны укладывается вдоль анодного блока в рассматриваемых условиях.

Как видно из рис. 14.5, б, распределение поля в пространстве взаимодействия носит периодический, но не гармонический характер, хотя изменение напряженности поля во времени в каждой точке пространства носит синусоидальный (гармонический) характер. Такое пространственное распределение при разложении в ряд Фурье содержит бесконечную последовательность пространственных гармоник.

Как следует из изложенного в п. 13.3, фазовая скорость любой пространственной гармоники $v_\phi = \omega/k_{zp}$, k_{zp} — волновое число пространственной гармоники номера $p = 0, \pm 1, \pm 2, \dots$; $k_{zp} = k_{30} + k_{zp}'$; $p' = \pm 1, \pm 2, \dots$; $k_{30} = 2\pi/\lambda_{30}$; λ_{30} легко определяется из

рис. 14.5, б: $\lambda_{30} = 2\pi r_a/n$ (r_a — радиус анодного блока); $k_{3p} = 2\pi p'/D$, где D — период структуры, т. е. расстояние между серединами зазоров двух соседних резонаторов: $D = 2\pi r_a/N$. Тогда

$$v_{\Phi(n,p)} = \frac{\omega}{k_{3p}} = \frac{\omega}{\frac{n}{r_a} + \frac{pN}{r_a}} = \frac{\omega r_a}{n + pN}. \quad (14.10)$$

Таким образом, замкнутая колебательная система магнетрона в отличие от разомкнутой системы в ЛБВО обладает бесконечной последовательностью пространственных гармоник для каждого n -го номера вида колебаний.

Если рассматривать частотный спектр колебаний анодного блока, то оказывается, что соседние виды колебаний очень незначительно различаются по частоте. При работе магнетрона в результате различных флуктуаций в источнике питания может наблюдаться «перескок» в режиме работы магнетрона с одного вида колебаний на другой, сопровождаемый скачкообразным изменением частоты генерации колебаний. Существует два основных метода разделения колебаний: использование связей, применение разнорезонаторных анодных блоков. **Связки** — это проводники, соединяющие точки анодного блока, находящиеся под одним и тем же СВЧ-потенциалом. Для колебаний вида π связки имеют электрический контакт на сегментах анодного блока, расположенных через один резонатор. **Разнорезонаторные анодные блоки** выполняются, как правило, из двух групп резонаторов с различными размерами, причем эти резонаторы чередуются друг с другом.

Механизм возникновения незатухающих колебаний в магнетроне сходен с подобными явлениями в клистроне и других электровакуумных генераторных приборах СВЧ с динамическим управлением. Колебания в магнетроне возникают из-за флуктуаций плотности электронного потока. Колебания поддерживаются за счет энергии источника электрического питания. В магнетроне, как и в других ЭВП СВЧ, происходит модуляция по плотности и энергетическое взаимодействие сформировавшихся сгустков с электрическим полем.

Рассмотрим физический механизм взаимодействия электронного потока, перемещающегося в скрещенных $\mathcal{E} \times B$ -полях, с СВЧ-полем. Электрическое СВЧ-поле в пространстве взаимо-

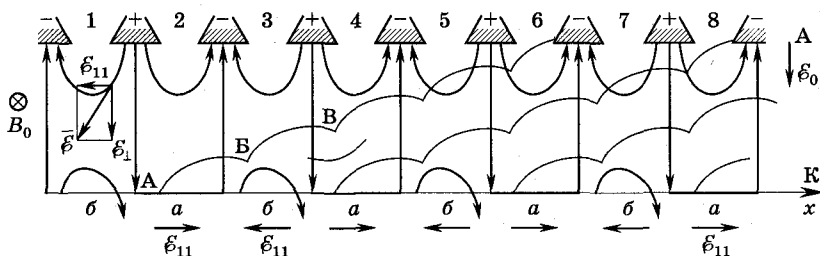


Рис. 14.6

действия имеет две составляющие: продольную \mathcal{E}_{\parallel} (вдоль поверхности катода или анода) и поперечную \mathcal{E}_{\perp} , направленную перпендикулярно к поверхности электродов.

Рассмотрим процессы взаимодействия с каждой составляющей. Начнем с продольной \mathcal{E}_{\parallel} . На рис. 14.6 показана развертка магнетрона, там же представлена структура СВЧ-поля в некоторый момент времени $t = t_0$. Все пространство взаимодействия разбито на две группы участков a и b , на каждом из которых поле \mathcal{E}_{\parallel} имеет свое направление. Сначала рассмотрим движение электронов, вылетающих с участков a катода. Если бы СВЧ-поле отсутствовало, то электроны, вылетев с катода, описали бы циклоиду и с той же энергией, с которой вылетели, вернулись бы на катод. Наличие же переменной составляющей электрического поля приводит к тому, что электроны на протяжении всего участка a будут перемещаться в тормозящем СВЧ-поле. В результате они потеряют часть энергии и при обратном движении к катоду не смогут достигнуть его и остановятся в точке Б. Для наглядности предположим, что время движения электрона от точки А до остановки в точке Б равно $T/2$ (T — период СВЧ-колебаний). Следующий цикл электрон начинает свое движение из точки Б в условиях, аналогичных моменту вылета электрона из катода (точка А), и остановится через полпериода в точке В и т. д. Через $T/2$ направления полей \mathcal{E}_{\parallel} на участках a и b изменяются на противоположные. Аналогичная картина будет наблюдаться на всех участках a пространства взаимодействия. Электроны, вылетающие с катода b , будут ускоряться СВЧ-полем, получать дополнительную энергию и будут возвращаться на катод. В результате весь электронный поток разобьется на ряд сгустков (спиц). Каждая спица образуется на одном простран-

венном периоде СВЧ-волны, поэтому число спиц в пространстве взаимодействия магнетрона будет определяться номером вида колебаний. Таким образом, продольная составляющая электрического СВЧ-поля ϵ_{\parallel} разбивает весь электронный поток на ряд спиц и формирует составляющую скорости электронов, направленную к аноду.

Время перемещения электрона от точки А до точки В не обязательно равно $T/2$. Оно может быть (и чаще всего бывает) меньше $T/2$, но это не изменяет общей картины взаимодействия электронов с СВЧ-полем. Итак, траекторию движения электрона можно приближенно уподобить траектории точки окружности, катящейся по прямой, уходящей от катода к аноду. Скорость качения этой окружности определяется отношением $v_{цк} = \epsilon_0/B_0$. Число петель в траектории электрона и крутизна прямой, по которой катится окружность, зависят от анодного потенциала и интенсивности высокочастотного поля. Подчеркнем еще раз, что на анод, описывая циклоидальные траектории, устремляются лишь электроны, вылетевшие с участков катода, против которого в данный момент времени существует тормозящее поле.

При взаимодействии с радиальной составляющей поля происходит уплотнение электронов в спице. Механизм этого уплотнения легко можно уяснить из рис. 14.7.

Рассмотрим три электрона, находящихся внутри спицы. Для электрона 1 суммарное радиальное СВЧ-поле $\epsilon = \epsilon_0 + \epsilon'_1$, для электрона 2 — ϵ_0 , а для электрона 3 — $\epsilon_0 - \epsilon''_1$ (см. рис. 14.7).

Переносная скорость электрона 1 $v_1 = \frac{\epsilon_0 + \epsilon'_1}{B_0}$, электрона 2 $v_2 = v_{цк} = \epsilon_0/B_0$ и электрона 3 $v_3 = \frac{\epsilon_0 - \epsilon''_1}{B_0}$, т. е. $v_1 > v_2 > v_3$. В результате такого соотношения скоростей электроны 1 и 3 будут

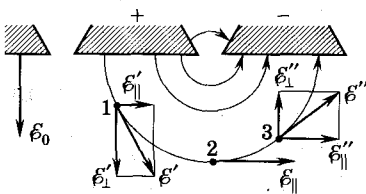


Рис. 14.7

приближаться к электрону 2, тем самым уплотняя спицу, т. е. происходит более плотная группировка электронов. Каждая спица находится в синхронизме с волной, бегущей вдоль анодного блока, т. е. спицы внутри пространства взаимодействия вращаются.

14.4. Характеристики и параметры магнетронов

Условием для отдачи электронами максимальной энергии высокочастотному полю является равенство фазовой скорости бегущей волны и средней скорости движения электронов в отсутствие колебаний: $v_{\Phi(n,p)} = v_{\text{цк}}$, т. е. в магнетроне реализуется условие синхронизма. Средняя скорость движения электронов неизменна для определенного режима работы магнетрона. Следовательно, и средняя кинетическая энергия электронов остается постоянной. Потенциальная же энергия электронов изменяется по мере подъема электронов к аноду. Она максимальна на катоде и равна eU_a , т. е. определяется полной разностью потенциалов между катодом и анодом, и минимальна (равна нулю) на аноде.

Пользуясь условием синхронизма, можно приближенно определить электрические и магнитные поля, соответствующие рабочим режимам магнетрона:

$$v_{\Phi(n,p)} = \frac{\omega r_{\text{cp}}}{n + pN}; v_{\text{цк}} = \frac{E_0}{B_0} = \frac{U_a}{dB_0}, \quad (14.11)$$

где ω — частота генерируемых колебаний; $r_{\text{cp}} = \frac{r_a + r_k}{2}$ — радиус окружности в середине пространства взаимодействия (r_a , r_k — соответственно радиусы анода и катода); $d = r_a - r_k$ — расстояние между катодом и анодом. Условие синхронизма меняется при переходе от одной точки к другой в направлении анода. Рассмотрим это условие для некоторой средней линии между катодом и анодом. Для пространства ближе к аноду от этой средней линии рассматриваемое условие будет выполняться с избытком, ближе к катоду с недостатком:

$$v_{\Phi(n,p)} = v_{\text{цк}}, \text{ т. е. } \frac{\omega(r_a + r_k)}{n + pN} = \frac{U_0}{(r_a - r_k)B_0}. \quad (14.12)$$

Отсюда

$$U_a = \frac{\omega(r_a^2 + r_k^2)B_0}{2(n + pN)}. \quad (14.13)$$

Из этого выражения видно, что для каждого вида колебаний существуют оптимальные соотношения между U_a и B_0 , при этом π -вид колебаний требует для своего возбуждения наименьшего

анодного напряжения. Одним из признаков возбуждения колебаний является протекание конвекционного тока в анодной цепи магнетрона. Постоянное анодное напряжение U_a , при котором возникают СВЧ-колебания, принято называть *пороговым напряжением* магнетрона. Анализ показывает, что простое напряжение U_n может быть определено по формуле

$$U_n = \frac{\omega(r_a^2 - r_k^2)B_0}{2(n + pN)} - \frac{m r_a}{2e} \left(\frac{\omega}{n + pN} \right)^2. \quad (14.14)$$

В соответствии с последней формулой построим графики (рис. 14.8). На этом рисунке изображена парабола критического режима и прямые линии, касательные к параболе. Каждая прямая определяет пороговое напряжение как функцию магнитного поля для каждого вида колебаний n (рассматривается восьмirezонаторный анодный блок, т. е. $N = 8$). Значения анодных напряжений для различных n в точках касания прямых называются *потенциалами синхронизации* U_{cn} . При $U_a = U_{cn}$ ($n = 1, 2, 3, 4$, см. рис. 14.8) электроны, двигаясь около анода параллельно его поверхности, оказываются в синхронизме с СВЧ-волной. Для того чтобы электроны попадали на анод, рабочее анодное напряжение магнетрона всегда должно удовлетворять условию $U_a > U_{cn}$, т. е. находиться выше линии порогового напряжения, при этом U_a не может сильно отличаться от U_n , иначе будет нарушаться условие синхронизма. Таким образом, при рабочих режимах магнетрона анодные напряжения имеют значения, близкие к величинам, соответствующим линиям порогового напряжения для различных видов колебаний. На рис. 14.8 штриховкой отмечены области значений U_a и B_0 , где вообще нет рабочих режимов магнетрона. При значениях U_a и B_0 , находящихся выше параболы критического режима, электроны, вылетая с катода и не совершая циклоидальных колебаний, попадают сразу на анод. Ниже линии $n = 4$ (π -вид) не существует видов колебаний, поскольку π -вид колебаний соответствует наибольшему n .

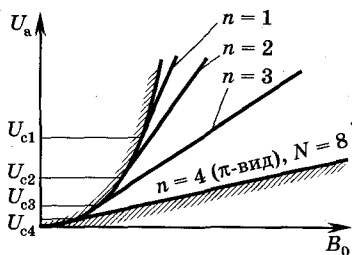


Рис. 14.8

В качестве *рабочих характеристик* магнетрона приняты зависимости $U_a = f(I_a)$ при постоянных значениях таких параметров прибора,

как индукция магнитного поля $B_0 = B_1, B_2, B_3, \dots$, колебательная мощность P_k , КПД η и частота генерируемых колебаний ω . Обычно рабочие характеристики линии постоянных значений B_0, P_k и η — изображаются на одном графике в координатах U_a, I_a (рис. 14.9).

Рассмотрим каждое семейство кривых $U_a = f(I_a)$ при $B_0 = \text{const}$, затем при $\eta = \text{const}$ и $P_k = \text{const}$.

КПД магнетрона η определяется электронными потерями η_e и потерями в колебательной системе η_k , т. е. $\eta = \eta_e \eta_k$. Электронные потери η_e определяются отношением энергии, получаемой СВЧ-полем, к полной энергии электронов, приобретенной от источника электрического питания. Основную долю потерь в η_e составляют потери на аноде, и, поскольку электроны попадают на анод со скоростью, не меньшей, чем фазовая скорость волны данного вида колебаний, минимальная скорость электронов определяется потенциалом синхронизации. Кинетическая энергия, соответствующая этой скорости, рассеивается в виде тепла. Вторую группу электронных потерь составляют электроны, которые попадают в неблагоприятную фазу СВЧ-волны и возвращаются на катод, разогревая его. Поэтому часто после воз-

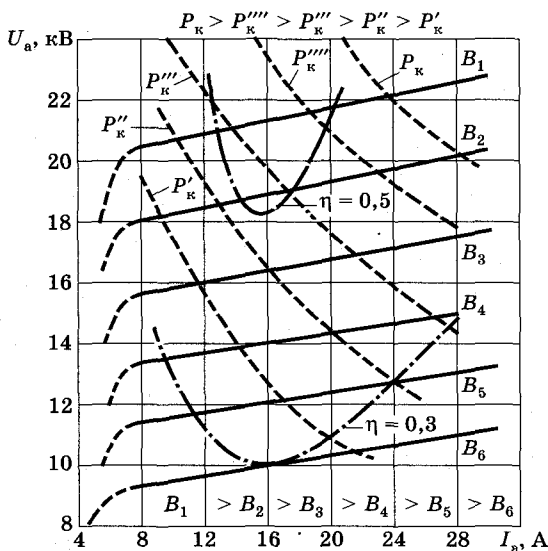


Рис. 14.9

никновения генерации напряжение накала (подогрева) катода уменьшают. КПД колебательной системы определяется активным сопротивлением резонатора.

При неизменной величине индукции магнитного поля $B_0 = \text{const}$ и увеличении анодного напряжения от нуля до $U_a < U_{\text{п}}$ ($U_{\text{п}}$ — пороговый потенциал) анодный ток очень мал, поскольку магнетрон работает в закритическом режиме, и СВЧ-колебания отсутствуют. При достижении пороговой величины $U_a = U_{\text{п}}$ магнетрон возбуждается, и анодный ток нарастает. При работе магнетрона необходимо выполнение условия синхронизма, т. е. $v_{\Phi(n,p)} = v_e$. Здесь

$$v_e = \frac{U_a}{(r_a - r_k)B}. \quad (14.15)$$

При постоянной индукции магнитного поля $B_0 = \text{const}$ возможно изменение $U_{\text{п}}$ в небольших пределах (не отражено на рис. 14.8). Увеличение анодного напряжения выше порогового вызывает увеличение скорости вращения спицы, в результате чего центр спицы смещается относительно максимума тангенциальной составляющей напряженности электрического СВЧ-поля. Наводимый в анодной цепи ток начинает опережать СВЧ-напряжение, и эквивалентная проводимость анодного блока приобретет емкостной характер. Частота генерируемых колебаний возрастает. Это явление носит название «электронного смещения частоты» (явление, аналогичное электронной настройке частоты). Изменение анодного напряжения, вызывающее пропорциональное изменение частоты, позволяет автоматически поддерживать условие синхронизма $v_{\Phi(n,p)} = v_e$.

При изменении B для сохранения условия синхронизма

$$v_{\Phi(n,p)} = v_e = \frac{U_a}{(r_a - r_k)B} = \text{const} \quad (14.16)$$

необходимо пропорциональное изменение U_a .

Чтобы объяснить семейство кривых $\eta = \text{const}$, рассмотрим зависимости $\eta = f(I_a)$ при $U_a = \text{const}$ и $\eta = f(U_a)$ при $I_a = \text{const}$, т. е. заменим параметр функцией, а функцию и аргументы по очереди будем делать параметром. Указанные две зависимости полностью объясняют линии постоянного КПД на рабочих характеристиках. Рассмотрим сначала зависимость $\eta_e = f(I_a)$ при

$U_a = \text{const}$. Как легко увидеть из семейства рабочих характеристик, КПД η сначала увеличивается с возрастанием тока I_a , а потом уменьшается.

При малых токах I_a мала и амплитуда переменного напряжения на анодном блоке магнетрона; фазовая группировка электронов в спице, обусловленная радиальной составляющей электрического СВЧ-поля, в этом случае проявляется слабо. Спица плохо сформирована. Много электронов выпадает из синхронизма с волной и возвращается на катод, что приводит к малому уровню мощности P_k и соответственно малому η_e . Увеличение I_a улучшает группировку электронов в спице. Все большее количество электронов в спице приближается к области максимального значения тангенциальной составляющей напряженности электрического СВЧ-поля, эффективность взаимодействия электронов с СВЧ-полем улучшается, растет его напряженность и выходная мощность увеличивается. При больших I_a генерируемая мощность P_k и η_e уменьшаются из-за разгруппировки спицы под действием большого объемного заряда электронов (большой ток — много электронов в спице — большой объемный заряд — большие кулоновские силы расталкивания).

Как видно из рабочих характеристик, функция $\eta = f(U_a)$ при $I_a = \text{const}$ является возрастающей, т. е. η становится больше при увеличении U_a . При объяснении этой зависимости учтем, что кинетическая энергия электронов, падающих на анод, не может быть меньше eU_{cn} . Если пренебречь потерями на катоде, то

$$\eta_e = \frac{I_a U_a - I_a U_{cn}}{I_a U_a} = 1 - \frac{U_{cn}}{U_a}, \quad (14.17)$$

где $P_0 = I_a U_a$ — мощность, затрачиваемая источником питания на формирование электронного потока, $U_c = U_{cn}$.

Выражение (14.17) показывает, что η_e растет при увеличении U_a . Однако реальные значения КПД меньше, чем вычисленные по последней формуле. Объясняется это тем, что кинетическая энергия электронов, падающих на анод, оказывается значительно больше, чем eU_{cn} , поскольку величина eU_{cn} определяет скорость $v_{цк}$ центра образующей окружности, а электроны могут попасть на анод в любой точке циклоиды. С увеличением U_a для соблюдения условия синхронизма пропорционально увеличи-

вается и индукция магнитного поля, поскольку $U_{\Phi(n,p)} = U_{\text{цк}} = \frac{U_a}{(r_a - r_k)B} = \text{const}$ для данного вида колебаний. Однако пропорциональное увеличение U_a и B приводит к уменьшению радиуса образующей окружности R , так как $R = \frac{mU_a}{(r_a - r_k)eB^2}$. Следовательно, высота и период циклоиды уменьшаются, что приводит к уменьшению средней кинетической энергии электронов, бомбардирующих анод. Уменьшение радиуса R эквивалентно уменьшению мгновенного значения кинетической энергии при постоянстве скорости перемещения центра круга $v_{\text{цк}}$.

Рассмотрим кривые $P_k = \text{const}$. Колебательную мощность можно записать в виде $P_k = \eta U_a I_a$, т. е. $U_a = \frac{P_k}{\eta I_a}$. Таким образом, в координатах U_a, I_a кривые $P_k = \text{const}$ близки к гиперболе. Отклонение от гиперболы обусловлено зависимостью η от U_a и I_a .

14.5. Приборы магнетронного типа

Приборы магнетронного типа (типа М) могут быть как генераторами, так и усилителями. Отметим свойства наиболее распространенных из них.

Митроны. Магнетроны, перестраиваемые напряжением (митроны), это широкополосные генераторы магнетронного типа.

Электрическая перестройка частоты в широких пределах в митроне осуществляется в основном либо путем ограничения эмиссионного тока катода (или тока анода), либо применением внешней колебательной системы малой добротности. В качестве колебательной системы в митроне используется встречно-штыревая структура 3 (рис. 14.10), свернутая в кольцо и связанная с внешней системой, имеющей низкую добротность (1,5...10). В магнетронах добротность, как правило, равна нескольким сотням и более.

Ограничение тока и управление им в митронах осуществляется, в частности, за счет того, что эмитирующий катод 1 вынесен за пределы пространства взаимодействия, т. е. электроны посту- пают в это про-

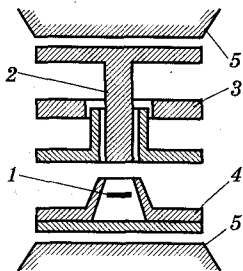


Рис. 14.10

странство извне. Для формирования необходимой структуры СВЧ-поля в центр прибора помещается холодный (не эмитирующий) катод 2. Между горячим катодом и анодной структурой располагается управляющий электрод 4. Так же как и в магнетроне, поперечное по отношению к электрическому магнитное поле формируется с помощью постоянного магнита 5.

Различают два основных типа митронов:

- 1) широкополосные митроны, характеризуемые отношением максимальной генерируемой частоты к минимальной, равным 2 : 1 или 3 : 1, при уровне выходной мощности до 1,5...3 Вт;
- 2) узкополосные митроны с перестройкой по частоте в пределах 5...20% при уровне выходной мощности от 3 до 150 Вт.

Митроны работают на частотах вплоть до 1 мм (300 ГГц).

При возрастании анодного напряжения U_a изменяется напряженность радиального электрического поля в пространстве взаимодействия и, следовательно, скорость вращения спиц, поскольку

$v_{цк} = \frac{U_a}{(r_a - r_k)B}$. Частота, с которой спицы будут проходить среднюю линию каждого резонатора, равна $f = v_{цк}/D$ (D — расстояние между соседними резонаторами, или период структуры). Если N — число элементарных резонаторов, то $f =$

$\frac{U_a N}{2\pi B(r_a^2 - r_k^2)}$, т. е. частота генерации является линейной функцией U_a , что подтверждается экспериментально. Зависимость выходной мощности от частоты $P_{вых} = F(f)$ подобна резонансной кривой анодной системы при ее низкой добротности. Это значит, что мощность на выходе пропорциональна импедансу, на который нагружен электронный поток. Зависимость $I_a = F_1(f)$ аналогична зависимости $P = F(f)$ мощности от частоты при $B = \text{const}$.

Лампа бегущей волны типа М (ЛБВМ). ЛБВМ используются как эффективные усилители мощных СВЧ-колебаний. Достоинствами этих магнетронных усилителей являются высокий КПД (до 70%), относительно низкие рабочие напряжения, например при $P_{вых} = 100$ кВт анодное напряжение составляет $U_a = 10...12$ кВ, а при $P_{вых} = 1$ МВт $U_a = 25...40$ кВ. Импульсные ЛБВМ имеют $P_{вых}$ до 10 МВт при $U_a = 100$ кВ с полосой пропускания $\Delta f \approx 17\%$.

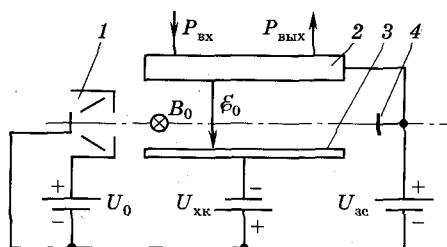


Рис. 14.11

Обычно ЛВВМ имеют цилиндрическую конструкцию, реже — линейную. Физические процессы удобнее рассматривать с помощью линейной ЛВВМ, устройство которой представлено на рис. 14.11, где 1 — электронная пушка (ЭП); 2 — замедляющая система (ЗС); 3 — холодный катод (ХК); 4 — коллектор; U_0 — напряжение на последнем ускоряющем электроде ЭП; $U_{хк}$ — напряжение на ХК; $U_{зс}$ — напряжение на коллекторе и ЗС.

ЭП формирует электронный поток с заданными параметрами. ЗС и ХК образуют пространство взаимодействия, где электроны взаимодействуют с СВЧ-полем ЗС в скрещенных постоянных электрическом \vec{E}_0 и магнитном \vec{B}_0 полях. Электронный поток вводится в пространство взаимодействия таким образом, чтобы электрическая составляющая силы $\vec{F}_e = -e\vec{E}_0$, действующая на электроны в направлении ЗС, была уравновешена магнитной составляющей $\vec{F}_m = -e[\vec{v} \times \vec{B}_0]$. В результате в отсутствие СВЧ-полей траектория электронов будет прямолинейной, т. е. циклоида вырождается в прямую линию. Скорость электронного потока в пространстве взаимодействия в этом случае будет $v_e = E_0/B_0$. В каждый момент времени все пространство взаимодействия, как и в магнетроне, можно разбить на ряд участков, где соседние участки имеют продольные составляющие СВЧ-поля, направленные противоположно. В результате на одних участках электроны будут замедляться, и электрическая составляющая силы F_e будет преобладать над магнитной. Электроны на этих участках будут подниматься в направлении ЗС, на других, наоборот, электроны ускоряются, тогда преобладает сила F_m . Электроны при этом двигаются в сторону холодного катода. Если $v_\phi = v_e$, то одни электроны будут подниматься, приближаясь к ЗС, а другие падать на катод, т. е. процессы взаимодействия здесь во многом аналогичны

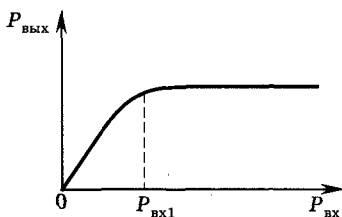


Рис. 14.12

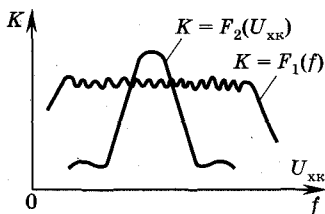


Рис. 14.13

процессам, протекающим в магнетроне. В электронном потоке образуются уплотнения (сгустки) и разрежения. В ЛБВМ потенциальная энергия преобразуется в энергию СВЧ-поля, и по мере приближения к ЗС потенциальная энергия электронов уменьшается, а энергия СВЧ-поля увеличивается. Поперечная составляющая СВЧ-поля, нормальная к поверхности ХК (иначе — отрицательного электрода), играет такую же роль, что и в магнетроне, т. е. она способствует лучшей группировке внутри сгустка.

Примеры основных характеристик ЛБВМ приведены на рис. 14.12 и 14.13. На рис. 14.12 изображена амплитудная характеристика. С увеличением входной мощности выходная $P_{\text{вых}}$ возрастет из-за того, что напряженность СВЧ-поля становится больше и электроны более интенсивно поднимаются к ЗС, отдавая все более значительную часть своей энергии. На пологом участке электроны при $P_{\text{вх}} > P_{\text{вх1}}$ начинают достигать ЗС, передав всю свою энергию СВЧ-полю. Незначительное увеличение $P_{\text{вых}}$ в этой части кривой вызвано тем, что по мере роста $P_{\text{вх}}$ все большее число электронов достигают ЗС.

Зависимости коэффициента усиления K от напряжения холодного катода $U_{\text{хк}}$ и частоты f (см. рис. 14.13) подобны соответственно характеристике взаимодействия и частотной характеристике ЛБВО.

Лампа обратной волны типа М (ЛОВМ). Наряду с ЛБВМ в СВЧ-технике достаточно широко применяются приборы, работающие на обратной пространственной гармонике — различные виды ЛОВМ, которые могут использоваться в качестве как усилителей, так и генераторов. Устройство генераторов на ЛОВМ (рис. 14.14) сходно с устройством маг-

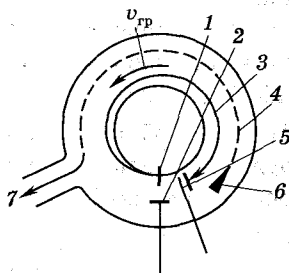


Рис. 14.14

нетронного усилителя типа ЛБВМ. Замедляющая система (ЗС) 4 обычно свернута в незамкнутое кольцо. Выходная мощность отводится в нагрузку по СВЧ-тракту 7, подключаемому вблизи катода 1, у ЛБВМ в этом месте находится вход прибора. У коллекторного конца ЗС ЛОВМ монтируется поглотитель 6, который поглощает электромагнитную волну, отраженную от выходной СВЧ-линии и катодного конца системы. В ЛБВМ у коллекторного конца располагается выход прибора. Пространство взаимодействия в ЛОВМ, как и в ЛБВМ, формируется холодным катодом 3 и ЗС 4. На рис. 14.14 показаны также анод электронной пушки 2 и коллектор 5; ЗС конструируется таким образом, чтобы в ней преобладающую роль играла первая обратная гармоника, вектор фазовой скорости $\vec{v}_{\phi(-1)}$ которой совпадает с направлением движения электронов. Групповая скорость СВЧ-волны (скорость переноса энергии $v_{гр}$ на рис. 14.14) направлена навстречу движению потока.

Под воздействием флуктуаций плотности электронного потока в ЗС возникают слабые электромагнитные колебания. Дальнейший механизм взаимодействия электронного потока с СВЧ-полем — формирование сгустков, отбор рабочих электронов, передача энергии полю волны — подобен механизму взаимодействия в ЛБВМ, за исключением того, что взаимодействие осуществляется не с прямой волной, а с обратной гармоникой. Однако фазовая скорость этой волны, как и в ЛБВМ, направлена в ту же сторону, что и направление движения электронного потока.

Одним из достоинств генератора на ЛОВМ является возможность (за счет изменения напряжения источников питания) электронного смещения частоты генерируемых колебаний в пределах нескольких процентов. Генераторы на ЛОВМ являются достаточно мощными приборами, у которых мощность в непрерывном режиме достигает нескольких кВт, КПД — 30%. Они работают вплоть до миллиметрового диапазона. ЛОВМ может работать как усилитель, в котором вход располагается около коллектора. Коэффициент усиления таких приборов 20...30 дБ.

Амплитроны. Это разновидность усилительных ЛОВМ, устройство которых сходно с устройством магнетронов. Отметим некоторые отличительные особенности устройства амплитрона (рис. 14.15, стрелками показаны направления движения СВЧ-энергии). Поскольку это усилительный прибор, то для устранения возбуждения на колебаниях вида λ , амплитрон,

в отличие от магнетрона, имеет нечетное число резонаторов. В магнетроне ЗС образована замкнутой цепочкой резонаторов, а в амплитроне ЗС 2 разомкнута за счет разрыва системы связок 3. Магнитное поле 4 формируется постоянным магнитом. Входной сигнал $P_{вх}$ возбуждает через связки поле в резонаторах. Электроны, эмитированные катодом 1, поступают в пространство взаимодействия и движутся по таким же траекториям, что и в магнетроне. Если скорость перемещения электронов равна фазовой скорости одной из пространственных гармоник СВЧ-поля, то в результате процессов группировки образуются спицы, в которых электроны передают свою потенциальную энергию СВЧ-полю. Амплитуда волны, бегущей вдоль связок, возрастает за счет потенциальной энергии электронов. В амплитроне взаимодействие осуществляется с обратной гармоникой, так же как и в ЛОВМ, т. е. амплитрон объединяет в себе свойства магнетрона и ЛОВМ. Амплитрон является очень мощным прибором, его мощность в непрерывном режиме достигает сотен кВт, а в импульсном — десятков и более МВт, усиление — до 20...30 дБ, рабочий диапазон частот — обычно единицы ГГц, КПД до 70...80%.

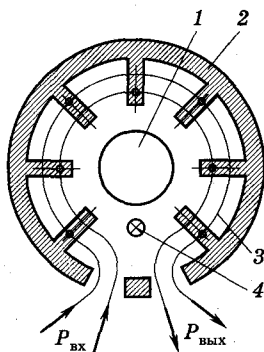



Рис. 14.15

Контрольные вопросы

1. Общие сведения о приборах типа «М», достоинства и недостатки. Области применения.
2. Каково устройство магнетрона? Движение электронов в скрещенных полях. Парабола критического режима.
3. Особенности колебательной системы магнетрона: виды колебаний, колебания π -вида, разделение видов колебаний.
4. Каковы рабочие характеристики и параметры магнетронов и амплитронов?
5. Каковы особенности, устройство, параметры и характеристики ЛВВМ, митронов, ЛОВМ?



ПРИБОРЫ ОТОБРАЖЕНИЯ ИНФОРМАЦИИ.

ОПТОЭЛЕКТРОННЫЕ ПРИБОРЫ

Глава 15

ГАЗОРАЗРЯДНЫЕ ПРИБОРЫ И ИНДИКАТОРЫ

15.1. Общие сведения

Газовый разряд. *Электрическим разрядом* называют совокупность процессов, связанных с прохождением электрического тока через газовую среду. *Ионные* или *газоразрядные приборы (ГРП)* — это приборы, использующие свойства электрического разряда в газе или парах металлов при давлении от 10^{-1} Па и выше.

Существует множество видов разрядов в газах, которые используются в приборах. Их можно разделить на две большие группы: *стационарные* и *импульсные*. Стационарные разряды характеризуются установлением равновесных значений всех параметров газовых разрядов. В *импульсных разрядах* ток протекает только в короткие интервалы времени, когда в разряде не успевают установиться равновесные процессы возбуждения, ионизации, рекомбинации, излучения и т. д. В интервалах между импульсами происходит установление равновесных распределений энергетических состояний атомов и молекул газа.

Наиболее распространенными параметрами газовых разрядов, используемых в приборах, являются следующие: плотность разрядного тока, напряженность электрического поля \mathcal{E} в разряде, давление газа p , диаметр канала разряда d , произведение давления на диаметр pd , отношение \mathcal{E}/p . Свойства разрядов

сильно зависят от граничных условий, рода газа, характера изменения напряжения, давления газа и т. д.

Приборы, построенные на основе газового разряда, в настоящее время находят применение в мощной силовой электротехнической и радиотехнической аппаратуре, в энергетических установках, в технологии, в том числе и в области микроэлектроники, в лазерной технике, в различных радиотехнических и электронных устройствах и т. д. Кроме того, ГРП широко применяются в индикаторной технике.

Индикаторные приборы. В гл. 12 уже были рассмотрены некоторые типы приборов, используемых как индикаторы — это электронно-лучевые приборы (трубки). В этой главе будут суммированы сведения об индикаторах и рассмотрены основные типы электронных приборов, используемых в качестве средств отображения информации.

Индикатором называется электронный прибор, предназначенный для преобразования электрических сигналов в изображение с определенным пространственным распределением яркости или контраста.

Все типы индикаторов могут быть разбиты на две группы: **активные** и **пассивные**. В **активных индикаторах** происходит преобразование электрической энергии в световую за счет использования следующих физических эффектов: **свечения накаливаемых тел в вакууме, низковольтной катодолюминесценции** (возбуждение твердого тела потоком быстрых частиц), **свечения газового разряда, электролюминесценции**. На основе этих явлений разработаны следующие типы активных индикаторов: **электронно-лучевые, вакуумные накаливаемые, вакуумные люминесцентные, газоразрядные, полупроводниковые** (светодиоды, электролюминесцентные конденсаторы).

В **пассивных индикаторах** под воздействием электрических полезных сигналов изменяются локальные оптические характеристики среды, что позволяет изменять в пространстве и времени амплитуду, фазу, длину световой волны, плоскость поляризации и направление распространения волн. Наибольшее применение в пассивных индикаторах получила модуляция интенсивности света. Для модуляции светового потока широко используются **электрооптические эффекты** в жидких кристаллах, **электрохромный эффект** (изменение цвета вещества под действием электрического поля), **электрофорез** (перемещение заряженных цветовых частиц под действием электрического поля).

Различные типы индикаторов могут быть разбиты на следующие группы: *знакомодулирующие*, в которых световой или электронный пучок профилируется по форме знака; *знакогенерирующие*, в которых знаки синтезируются по принципу фигур Лиссажу; *знакосинтезирующие*, в которых изображение формируется с помощью мозаики независимо управляемых элементов-преобразователей «сигнал—код». По характеру отображений информации индикаторы всех типов делятся на *единичные (точечные)*, *шкальные*, *цифровые* и *буквенно-цифровые* (одно- и многоразрядные), *матричные* и *мнемонические (мнемосхема — это условное изображение объектов, их состояния, процессов, явлений)*.

15.2. Элементарные процессы в газовых разрядах

Движение электронов и ионов в разрядном промежутке, формируемом двумя электродами (один из которых — анод, а другой — катод (рис. 15.1)), определяется как действием электрического поля, так и беспорядочным перемещением, обусловленным их взаимодействием с частицами газовой среды. Под взаимодействием двух частиц понимается любое вынужденное изменение характеристик их движения или внутреннего состояния при сближении. Все процессы взаимодействия между частицами в газовом разряде делятся на *упругие* и *неупругие*. При *упругих взаимодействиях (столкновениях)* суммарная кинетическая энергия взаимодействующих частиц остается постоянной. Для *неупругих взаимодействий* характерно изменение суммарной кинетической энергии частиц, которое обусловлено изменением их внутренней энергии. Если суммарная кинетическая энергия взаимодействующих частиц убывает, то внутренняя энергия по крайней мере одной из частиц возрастает (*неупругие столкновения первого рода*) и, наоборот, если кинетическая энергия растет, то внутренняя энергия хотя бы одной из частиц уменьшается (*неупругие столкновения второго рода*).

Упругое взаимодействие, или рассеяние, приводит к превращению направленного потока частиц в беспорядочное, хаотическое движение. В газоразрядных приборах наиболее существенны три типа упругих взаимодействий: 1) электрон-электронные; 2) атом-атомные, ион-атомные, ион-ионные; 3) электрон-атомные и электрон-ионные.

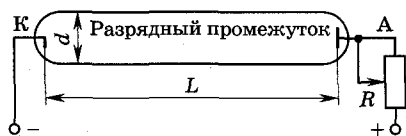


Рис. 15.1

В газовых разрядах, как и в твердом теле, вероятность столкновения частиц характеризуется эффективным и полным сечением взаимодействия. Под эффективным сечением взаимодействия понимают суммарную площадь, в пределах которой реализуется столкновение. Полное сечение процесса равно произведению эффективного сечения на концентрацию взаимодействующих частиц.

Сечения взаимодействий сильно зависят от относительных скоростей частиц. При малой степени ионизации газа, когда число ионов составляет $10^{-5} \dots 10^{-4}$ от числа атомов, наибольшую роль играют электрон-электронные взаимодействия, за счет которых функция распределения частиц по энергиям становится функцией Максвелла. При распределении Максвелла средняя кинетическая энергия хаотического движения частиц определяется их температурой, которая для отдельного сорта частиц, особенно электронов и ионов, сильно отличается от окружающей температуры. Это связано с тем, что в условиях разряда заряженные частицы приобретают энергию, ускоряясь в электрическом поле. Из-за упругих столкновений энергия от заряженных частиц передается нейтральным, вследствие чего направленное вдоль поля перемещение заряженных частиц (дрейф) превращается в хаотическое движение. Увеличение средней энергии хаотического движения различно для разных частиц. Наиболее интенсивно обмениваются энергией частицы с близкими массами, т. е. ионы с атомами и электроны между собой. При упругом соударении электрона с атомом практически не происходит обмена энергией из-за большой разницы их масс. Температура атомов T_a приближается к температуре ионов T_i только в дуговых разрядах очень большой плотности из-за высокой степени ионизации и большой частоты соударений. В разрядах, которые используются в приборах, температуры частиц сильно различаются (T_e — температура электронов): $T_e \gg T_i > T_a$. Температура частиц зависит от произведения pd , рода газа, наличия магнитных полей и т. д.

К неупругим столкновениям 1-го рода, играющим определяющую роль в приборах, относятся следующие процессы:

- 1) возбуждение атомов электронным ударом — $\vec{e} + A = e + A^*$
(\vec{e} , e — соответственно быстрый и медленный электроны, A , A^* — соответственно невозбужденный и возбужденный атомы);

- 2) ступенчатое электронное возбуждение — $\vec{e} + A^* = e + A^{**}$ (A^{**} — атом, перешедший на более высокий энергетический уровень с уровня, соответствующего A^*);
- 3) ионизация — $\vec{e} + A = 2e + A^+$ (A^+ — положительный ион). В результате этих неупругих взаимодействий часть кинетической энергии электронов переходит в энергию возбуждения или ионизации.

При взаимодействии медленного электрона e с возбужденным атомом A^* возможно столкновение 2-го рода $e + A^* = \vec{e} + A$. Этот процесс называется электронным девозбуждением и является основным при разрушении возбужденных состояний.

Обобщая сказанное, отметим, что упругие столкновения приводят к установлению определенной температуры частиц в газовом разряде. Электронная температура T_e определяет вероятность большинства неупругих взаимодействий (возбуждение, ионизация, девозбуждение, ступенчатое возбуждение).

Переход возбужденных атомов, ионов и молекул в более низкие энергетические состояния, включая основное, а также процесс рекомбинации электрона с ионом часто сопровождаются излучением фотона. Спектр оптического излучения в газовом разряде простирается от ИК до глубокой УФ области (вплоть до рентгеновского излучения).

Рассмотрим некоторые общие физические процессы, характерные для любого вида электрического разряда в газе. Первоначальные свободные электроны и ионы в газе могут появиться за счет внешних световых потоков, ионизирующих излучений различной природы, термоэмиссии, фотоэмиссии с поверхности катода и т. д. Если к электродам разрядного промежутка приложено внешнее постоянное напряжение плюсом к аноду, то электроны начинают перемещаться по направлению к аноду, а положительные ионы — к катоду.

Установившиеся разряды в постоянном электрическом поле можно разделить на несамостоятельные и самостоятельные. Несамостоятельные разряды при подаче напряжения на электроды существуют только при наличии внешних ионизирующих источников. В газоразрядных приборах используются самостоятельные разряды, которые существуют (возникают) в условиях, когда внешний ионизирующий источник, например фоновое ионизирующее излучение, оптическое излучение, тепловое излучение и т. д., необязательны или даже вообще отсутствуют. Основным процессом ионизации газа в разрядах такого типа является ударная ионизация (ионизация электронным ударом).

Разность потенциалов в разрядном промежутке, при которой электроны приобретают энергию, достаточную для возбуждения (ионизации) нейтральных атомов, называют **напряжением возбуждения (ионизации)**. Первоначальные электроны, набрав энергию в электрическом поле, будут ионизировать молекулы или атомы газа. В результате появляются добавочные свободные электроны, число которых нарастает при движении к аноду, так как они, в свою очередь, также участвуют в процессе ударной ионизации. Образовавшиеся в результате ионизации атомов положительные ионы перемещаются по направлению к катоду и выбивают с его поверхности новые электроны, которые, в свою очередь, также участвуют в ионизации. Таким образом, число электронов лавинообразно увеличивается до установления стационарного состояния, определяемого параметрами \mathcal{E}/d , pd , родом газа, конфигурацией электродов, внешними ионизирующими источниками, материалом и температурой катода, сопротивлением внешней цепи и рядом других факторов.

Чаще всего в разрядах реализуется ситуация, при которой суммарные объемные заряды электронов и ионов равны, а дебаевский радиус экранирования (дебаевская длина, см. п. 6.2) существенно меньше размеров приборов. Такое состояние ионизированного газа называют **газоразрядной плазмой**.

Типичные распределения параметров плазмы газового разряда (потенциала φ , напряженности электрического поля \mathcal{E} , объемного заряда ρ , плотности токов электронного j_e , ионного j_i и разрядного $j_p = j_e + j_i$, яркости свечения V) приведены на рис. 15.2.

Можно выделить три области разряда: прикатодную I, промежуточную II, прианодную III.

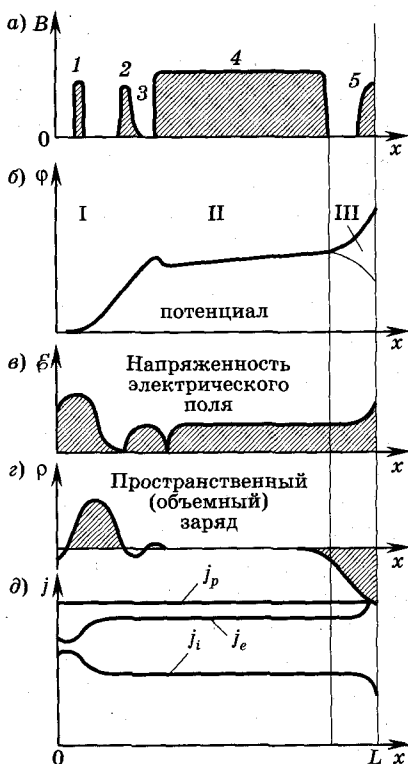


Рис. 15.2

Во многих приборах используется промежуточная область разряда, которая часто называется положительным столбом. Положительный столб характеризуется небольшим положительным падением потенциала. Прикатодная область называется областью прикатодного падения потенциала. В анодных областях разряда падение потенциала может быть как положительным, так и отрицательным, а в катодных — как правило, положительным.

15.3. Приборы тлеющего разряда. Плазменные панели

Тлеющий разряд используется в газоразрядных стабилизаторах напряжения (стабилитронах), в электропреобразовательных приборах — тиратронах, в приборах отображения информации (плазменных панелях, индикаторах), в газовых лазерах, в переключающих приборах (декатронах) и т. д.

Физические процессы в тлеющем разряде. Тлеющий разряд возникает в разрядном промежутке с холодным (неподогретым) катодом и характеризуется большим катодным падением потенциала, небольшой плотностью тока, наличием специфически выраженных катодных областей (см. рис. 15.2).

В непосредственной близости от катода электроны еще не набрали энергии, требующейся для возбуждения газа. Эта часть разряда не имеет свечения. Как только электроны начнут обладать энергией, достаточной для прямого или ступенчатого возбуждения частиц газа, возникает свечение, обусловленное девозбуждением атомов и молекул газа. Эта область разряда называется *катодной светящейся пленкой* (область 1, см. рис. 15.2, а). Так как энергия электронов по мере их продвижения к аноду возрастает, то вероятность возбуждения атома падает из-за уменьшения эффективного сечения взаимодействия при больших энергиях электронов, что приводит к уменьшению яркости свечения в этой области по сравнению с предыдущей (пространство между областями 1 и 2 на рис. 15.2, а). Однако здесь происходит интенсивное возбуждение большого числа различных атомных уровней и ионизация атомов. Число электронов лавины быстро возрастает.

Область тлеющего разряда, непосредственно примыкающая к катодному падению, называется областью отрицательного или *тлеющего свечения* (область 2 на рис. 15.2, а). В этой области еще имеется значительная доля электронов, сохранивших направленную скорость, приобретенную на катодном падении.

За область тлеющего свечения лежит *темная область*, где преобладает беспорядочное движение электронов (область 3, см. рис. 15.2, а, б). Положительные ионы здесь уравнивают пространственный заряд электронов (см. рис. 15.2, в), что приводит к уменьшению градиента потенциала. Абсолютное значение потенциала в этой области может несколько снижаться из-за проскакивания наиболее быстрых электронов, летящих от катода через область максимума потенциала (см. рис. 15.2, б). В рассматриваемой части разряда и в начале следующей (область 4) наблюдается наибольшая концентрация медленных электронов и положительных ионов,двигающихся в сторону катода, а также создаются благоприятные условия для излучательной рекомбинации электронов и ионов, как видно из рис. 15.2, а.

При дальнейшем продвижении к аноду вдоль области 4 беспорядочное тепловое движение электронов сохраняется, а потенциал $\varphi(x)$ незначительно возрастает (область положительно-го столба (ПС)) (см. рис. 15.2, б, область II). Увеличение потенциала обеспечивает медленный дрейф электронов к аноду и объясняется небольшим преобладанием концентрации положительных ионов в плазме столба. Что касается продольной (в направлении анода) напряженности электрического поля \mathcal{E} , то в этой области она остается постоянной (см. рис. 15.2, в). Электрический ток в ПС переносится параллельно оси разряда за счет дрейфа ионов по направлению осевого электрического поля, а электронов — против этого поля. В прикатодные области ПС поступает со стороны катода в единицу времени ровно столько электронов, сколько необходимо для поддержания заданной величины электрического тока в разряде. Аналогично этому столько ионов, сколько необходимо для поддержания ионной компоненты тока, поступает из прианодной области. Поскольку с поверхности анода никакой эмиссии положительных ионов нет, обеднение прианодной области III ионами должно непрерывно пополняться в ней за счет добавочной ионизации газа электронами, что реализуется благодаря большей напряженности поля в этой области по сравнению с продольной напряженностью в ПС (см. рис. 15.2, в).

Если с помощью внешнего переменного резистора, включенного в анодную цепь (см. рис. 15.1), изменять величину разрядного тока от нуля, то сначала возникает несамостоятельный разряд, который происходит только при наличии внешнего ионизатора. Начало самостоятельного тлеющего разряда характеризуется тем, что

при некотором напряжении между электродами (напряжения зажигания) и соответствующем токе разряда у поверхности одного из участков катода, где существует наибольшая напряженность поля и наибольшая эмиссия электронов, образуется область (столбик) повышенной ионизации газа, т. е. в этом объеме разряда процессы ионизации преобладают над рекомбинацией и диффузией заряженных частиц в окружающее пространство. После образования такого столбика плотность разрядного тока достигает нормальной плотности тока тлеющего разряда (см. рис. 15.2, *д*), определяемой сочетанием материала катода и природы газа и мало зависящей от давления газа. При дальнейшем увеличении тока разряда в некотором диапазоне его плотность не изменяется. Возрастание тока приводит лишь к расширению столбика, т. е. области катода, охваченной разрядом, при этом падение напряжения на разряде практически не меняется. Описанные процессы характерны для *нормального тлеющего разряда*, ВАХ которого (зависимость выходного напряжения от тока через прибор) представлена на рис. 15.3. Постоянство выходного напряжения при изменении тока через прибор (в области значений от $I_{ст. мин}$ до $I_{ст. макс}$, см. рис. 15.3) лежит в основе работы стабилитронов, тиратронов, газоразрядных панелей, некоторых типов лазеров и т. д.

Когда в разряде участвует вся поверхность катода, обращенная к аноду, то дальнейшее уменьшение сопротивления и увеличение разрядного тока приводит к существенному увеличению разрядного напряжения. В этом случае плотность тока у катода и катодное падение потенциала превышают нормальные значения. Такая форма разряда называется *аномальным тлеющим разрядом*.

После перехода в аномальный тлеющий разряд ВАХ прибора становится резко возрастающей (см. рис. 15.3).

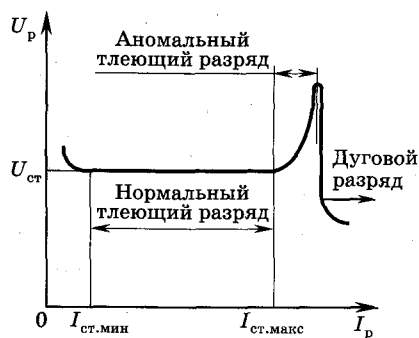


Рис. 15.3

При аномальном тлеющем разряде из-за высокого напряжения энергия ионов, бомбардирующих катод, существенно возрастает. Это приводит к дополнительному разогреву поверхности катода, что будет приводить к увеличению термоэмиссии с поверхности катода. Дополнительные электроны сталкиваются с атомами

и ионизируют их. В результате ток разряда резко увеличивается и аномальная форма тлеющего разряда переходит в *дуговую* (см. п. 15.4).

Многие черты чисто плазменных процессов, без учета приэлектродных явлений, характерных для тлеющего и дугового разрядов в постоянном электрическом поле, свойственны и разрядам в быстропеременных полях, в которых наличие электродов в разрядном промежутке, что имеет место в разряде на постоянном токе, вообще не является обязательным. Поэтому разряды часто классифицируют исключая признаки, связанные с электродными и приэлектродными эффектами. К таким признакам относятся: характер состояния ионизированного газа при действии внешнего поля и частотный диапазон поля. По первому признаку основные черты разряда в переменных полях совпадают с разрядом на постоянном токе. По признаку частоты f разряды разделяются на следующие типы:

- 1) *постоянные, низкочастотные* и не слишком кратковременные *импульсные* электрические поля, при которых свободные электроны за характерные времена изменения разряда успевают достичь электродов ($f \sim 10^2 \dots 10^5$ Гц);
- 2) *высокочастотные (ВЧ)* поля (*радиочастотный разряд* $f \sim 10^5 \dots 10^8$ Гц), в которых явления наиболее сложны и многообразны;
- 3) *сверхвысокочастотные (СВЧ)* поля, для разрядов которых характерна малость амплитуды пространственных колебаний электрона по сравнению с размерами разрядного объема.

В СВЧ-разрядах электроны и ионы при их движении под действием поля не достигают стенок, ограничивающих разряд. Эмиссионные процессы не играют роли, электронная лавина локализована. ВЧ-разряд занимает промежуточное состояние между разрядом на постоянном токе и СВЧ-разрядом.

По способу возбуждения разряды в переменных полях разделяются на *индукционные* и *емкостные*. В индукционных разрядах для зажигания используются индукционные катушки (соленоиды), изолированные от разрядного промежутка и формирующие переменное электромагнитное поле в разрядном промежутке.

В емкостных разрядах применяются внешние электроды, изолированные диэлектриком от рабочего плазменного объема.

Стабилитроны. Это двухэлектродные газоразрядные приборы, предназначенные для стабилизации напряжения. Катод стаби-

литрона выполнен в виде полого цилиндра с активированной внутренней поверхностью. Стержневой анод располагается по оси катода внутри него, т. е. поверхность анода существенно меньше поверхности катода. Оба электрода помещаются в стеклянный баллон, который наполняется смесью инертных газов (Ne—Ar, He—Ar и др.), давление которых колеблется от 10^{-1} до 10^4 Па. ВАХ стабилитрона совпадает с характеристикой нормального тлеющего разряда (см. рис. 15.3). Постоянное напряжение горения тлеющего разряда называется *напряжением стабилизации* ($U_{ст}$, см. рис. 15.3). Токи, соответствующие ВАХ тлеющего разряда, максимальный $I_{ст. макс}$ и минимальный $I_{ст. мин}$, называются *токами стабилизации*.

В настоящее время стабилитроны тлеющего разряда используются для *стабилизации напряжения* ($\Delta U_{ст}/U_{ст}$) = 0,02...0,06; $I_{ст} < 30...40$ мА, $U_{ст} \sim 70...150$ В или в *режиме опорного элемента* ($I_{ст} < 5$ мА, $U_{ст} \sim 80...90$ В). У опорных стабилитронов выделяют дополнительные параметры: нестабильность напряжения (< 0,1 В за час работы), температурный коэффициент напряжения стабилизации (несколько мВ/°С в диапазоне температур -60...+100 °С) и ряд других. Опорные стабилитроны отличаются высокой допустимой температурой (до 300 °С), долговечностью (тысячи часов), малыми габаритами, высокой механической прочностью.

Тиратроны. Тиратроны тлеющего разряда — это ионные трех- или четырехэлектродные приборы, которые используются как индикаторы и генераторы релаксационных колебаний и как другие активные элементы схем. Тиратроны имеют анод, холодный катод, одну (трехэлектродный прибор) или две (тетродный тиратрон) сетки. Все электроды помещены в миниатюрный стеклянный баллон, заполненный инертным газом при давлении порядка 10^3 Па. В зависимости от типа прибора конструкция отдельных электродов различна, но в любом из них катод, чаще всего молибденовый или никелевый, активированный материалом с малой работой выхода (цезий, барий и др.), имеет более развитую поверхность по сравнению с анодом. Сетки обычно выполняются в виде пластин с круглыми отверстиями (диафрагмами), а анод — из молибденовой проволоки. Трехэлектродные (триодные) тиратроны относятся к приборам с токовым управлением, а тетродные — с потенциальным управлением.

В *триодных тиратронах* при подаче напряжения между катодом и сеткой возникает вспомогательный разряд. Образовавшиеся в этом разряде электроны облегчают зажигание разряда между сеткой и анодом, а следовательно, и в основном разрядном промежутке катод—анод. Изменением тока I_c (напряжения U_c) на

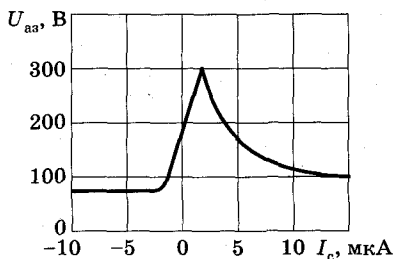


Рис. 15.4

сетке можно управлять анодным напряжением зажигания $U_{аз}$ тиратрона (рис. 15.4). Зависимости $U_{аз} = f(I_c)$ или $U_{аз} = f(U_c)$ называются *пусковыми характеристиками* тиратрона.

Отрицательные сеточные токи I_c на рис. 15.4 соответствуют подаче положительного напряжения U_c на сетку. В результате при определенных значениях U_c в пространстве между катодом и сеткой зажигается разряд, который при достаточных анодных напряжениях U_a охватывает и область между анодом и сеткой (левая часть характеристики с $U_{аз} = \text{const}$).

В окрестности, близкой к $I_c \approx 0$ (растущая часть пусковой характеристики), напряжения на сетке U_c недостаточно для зажигания вспомогательного разряда, и поэтому требуются большие анодные напряжения, чтобы проникающее в пространство сетка—катод электрическое поле смогло зажечь разряд. Следовательно, в этой области анодное напряжение зажигания $U_{аз}$ растет. Правая подающая часть характеристики связана с зажиганием разряда сначала между анодом и сеткой и перекидыванием его в область катод—сетка.

Характерной особенностью тиратронов любого типа является то, что после зажигания основного разряда сетка теряет свои управляющие свойства, т. е. изменение потенциала на сетке не влияет на разрядный (анодный) ток I_a и напряжение U_a . Это объясняется тем, что поле сетки экранируется заряженными частицами плазмы разряда на расстоянии порядка дебаевской длины (см. приложение 2 и гл. 2). После выключения анодного напряжения происходит деионизация плазмы в разрядном промежутке за время 10...100 мкс, и управляющее действие сетки восстанавливается.

В *четырёхэлектродном тиратроне* (тетроде) на первую сетку подается положительное напряжение 50...100 В, обеспечивающее

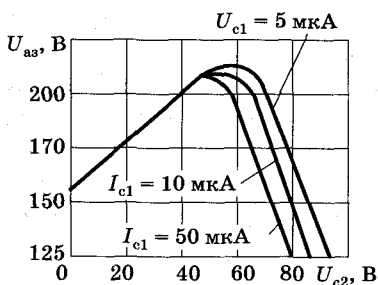


Рис. 15.5

вспомогательный разряд между катодом и первой сеткой. Напряжение на второй сетке тоже положительное, но меньше, чем на первой. С увеличением $U_{с2}$ (потенциальное управление) тормозящее поле между сетками уменьшается, электроны проходят к аноду и в пространстве между второй сеткой и анодом ионизируют газ, вызывая появление тлеющего

разряда. Из пусковых характеристик $U_{аз} = f(U_{с2})$ (рис. 15.5) видно, что зажиганием разряда можно управлять за счет изменения тока $I_{с1}$ первой сетки и напряжения $U_{с2}$ на второй.

Анодные характеристики тиратронов $U_a = f(I_a)$ представляют собой обычную характеристику нормального тлеющего разряда (см. рис. 15.3).

К основным параметрам тиратронов относятся: напряжения на электродах относительно катода при возникновении разряда, максимальное значение анодного тока I_a , время восстановления (время деионизации), характерные пусковые токи и напряжения; долговечность, диапазон рабочих температур, механическая прочность. Тиратроны тлеющего разряда имеют относительно небольшие размеры, высокую стабильность, долговечность, широкий диапазон рабочих температур ($-60...+100$ °С).

Декатроны. Декатроны — многоэлектродные приборы тлеющего разряда, используемые иногда в качестве счетчиков, цифровых индикаторов в десятичной системе, делителей частоты и пр.

Промышленностью выпускаются **одноимпульсные декатроны**, управляемые одиночными импульсами; **коммутирующие декатроны**, у которых каждый катод имеет отдельный вывод, что дает возможность распределить поступающие импульсы по десяти каналам; **многоанодные декатроны** — **полиатроны** для управления знаковыми индикаторами. Полиатроны можно также использовать в качестве счетных приборов.

Газоразрядные (плазменные) панели. Газоразрядные (плазменные) панели (ГРП) — многоэлектродные приборы, использующие оптическое излучение тлеющего разряда. ГРП применяются в качестве приборов отображения информации (различ-

ного типа индикаторы, дисплеи, многоцветные экраны, в том числе и для воспроизведения изображений).

Широкое внедрение персональных компьютеров и телевидения практически во все сферы человеческой деятельности стимулировало, помимо улучшения параметров ЭЛТ, развитие альтернативных технологических способов отображения информации, к которым относятся плоскочелпанельные дисплеи на основе газового разряда — плазменные панели (ПП) и на основе жидких кристаллов (ЖК).

По сравнению с ЭЛТ газоразрядные ПП обладают возможностью использовать большие размеры экрана при меньшей толщине прибора, лучшие эксплуатационные и эргономические характеристики и в перспективе не меньшую по сравнению с ЭЛТ информационную емкость при отображении информации. Большая яркость и контрастность, отсутствие искажений и дрожания изображения, отсутствие магнитных полей и, следовательно, проблем сведения электронных лучей, их фокусировки, а также стойкость ГРП к внешним электромагнитным полям и т. д. делают их более перспективными по сравнению с ЭЛТ.

ГРП имеют много конструктивных разновидностей. В настоящее время наилучшими характеристиками обладают следующие матричные газоразрядные индикаторные панели: ГРП постоянного тока, ГРП переменного тока (высокочастотные), ГРП постоянного тока с самосканированием. Однако общим конструктивным признаком для большинства матричных ГРП является наличие двух ортогональных систем электродов (катодных и анодных), нанесенных или уложенных на стеклянные плоские или гофрированные пластины, отделенные друг от друга диэлектрической пластиной с матричной системой отверстий, оси которых проходят через точку пересечения осей электродов (рис. 15.6). Полосковые электроды, как правило, прозрачны, чтобы не препятствовать выходу видимого свечения разряда. Пространство между электродами заполняется либо чистыми инертными газами, либо смесью газов при давлении несколько сотен Па. Свечение отдельных частей тлеющего разряда, как указывалось ранее, в основном обусловлено переходами возбужденных атомов с верхних энергетических уровней на более низкие. Состав смеси газов подбирается таким образом, чтобы увеличить интенсивность процессов ионизации, возбуждения и рекомбинации. Например, при добавлении к неону атомов ксенона велика вероятность реакции $Ne^* + Xe \rightarrow Xe^* + e + Ne$, т. е. в этом случае образование заряжен-

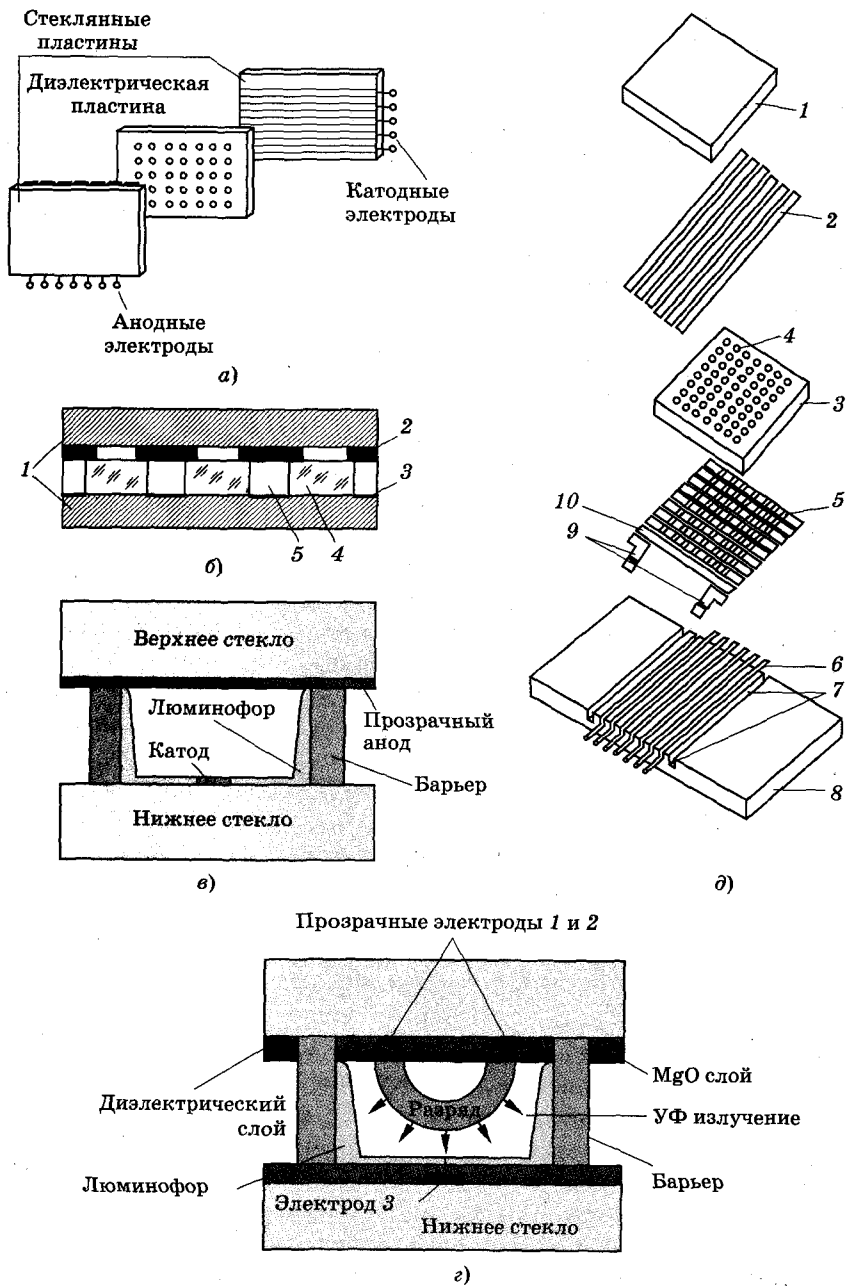


Рис. 15.6

ных частиц идет не только за счет столкновений электронов с нейтральными атомами, но и при взаимодействии возбужденных атомов основного газа Ne^* с атомами примеси Xe . Излучение тлеющего разряда в инертных газах обычно лежит в инфракрасной (ИК), красной или ультрафиолетовой (УФ) областях спектра. Невидимое, особенно УФ, излучение играет основную роль в цветных ГРП. Наиболее интенсивно излучающими областями тлеющего разряда, которые используются в ГРП, являются отрицательное свечение и положительный столб (области 2 и 4 на рис. 15.2, а). Если положительный столб ограничен металлическими или диэлектрическими стенками, что и реализуется в ГРП, то яркость может достигать $\sim 10^3$ кд/м² при светоотдаче в перспективе до 5 лм/Вт.

Устройство ГРП постоянного тока приведено на рис. 15.6, а, б, в, где отмечены все основные элементы конструкции: 1 — стеклянные пластины, 2 и 3 — электроды, 4 — диэлектрическая пластина с отверстиями, 5 — газоразрядная ячейка.

Принцип работы ГРП заключается в следующем: при подаче достаточного для зажигания разряда напряжения между каким-либо катодом и анодом (100...200 В) в ячейке, где эти электроды пересекаются, возникает тлеющий разряд. При подаче по определенному закону напряжения на несколько катодов и анодов можно с помощью точечного растра воспроизвести любую фигуру. Надежный и безошибочный запуск многочисленных разрядных ячеек в соответствии с входными сигналами индикации является сложной задачей в ГРП матричного типа. Для увеличения быстродействия и стабильности используют вспомогательный разряд в каждой ячейке ГРП. ГРП постоянного тока могут давать и многоцветное изображение. В этом случае на боковые или торцевые стенки ячеек наносятся покрытия из люминофоров, дающие свечение определенного цвета. Располагая рядом ячеек с покрытиями из люминофоров, дающих основные цвета (синий, зеленый, красный), можно получать цветные изображения. Зажигание в разряде ГРП осуществляется по принципу развертки.

В ГРП переменного тока используется более сложная конструкция; их принцип действия можно уяснить на примере одной ячейки (рис. 15.6, г), в которой электроды, в отличие от ГРП постоянного тока, отделены от газового пространства диэлектрическим слоем. В таких ГРП используется разряд на переменном токе емкостного типа, когда необходимое для зажигания разряда электрическое поле проникает в рабочее газовое простран-

во через емкости, формируемые диэлектрическими покрытиями электродов. Зажигание разряда в ячейке, показанной на рис. 15.6, *г*, осуществляется с помощью электродов 1 и 3, а поддержка горения планарного разряда реализуется благодаря электроду 2. Для повышения эффективности ГРП необходимо увеличивать поток УФ излучения из разряда на люминофор. Объем области отрицательного свечения, используемого в таких ГРП, обычно не превышает 10% от газового объема ячейки, следовательно, за счет использования более эффективных механизмов разряда и других светящихся областей разряда можно повысить их эффективность. В ГРП переменного тока для замены части изображения вводятся дополнительные электроды. В наиболее совершенных ГРП переменного и постоянного тока с автоматическим сдвигом строк благодаря возможности многострочной индикации можно резко сократить число элементов схемы управления.

На рис. 15.6, *д* изображено устройство электродной системы *матричного индикатора постоянного тока с самосканированием*: 1 — прозрачная стеклянная пластина; 2 — аноды индикации; 3 и 4 — соответственно диэлектрическая пластина и отверстия в ней; 5 — полосковые катоды сканирования с отверстиями диаметром ~ 0,05 мм; 6 — аноды сканирования, расположенные в канавках 7, вырезанных в опорной стеклянной пластине 8; 9 — парные электроды дежурного разряда; 10 — сплошной катод сброса (КС). Приведенное устройство условно можно разделить на две части: *сканирующую*, образованную анодами и катодами сканирования соответственно 6 и 5, и *индикаторную* с анодами индикации 2 и катодами сканирования 5.

При работе такого ГРП в сканирующей части реализуется направленное перемещение разряда из ячейки в ячейку по катодам сканирования 5 по мере переключения напряжения на них. Цикл работы начинается с разряда на катоде КС 10. Для уменьшения статического времени разброса возникновения разряда на КС используются дежурные электроды 9, обеспечивающие постоянно горящий разряд. После КС разряд переносится по катодам сканирования. Из-за малых размеров отверстий в катодах 5 свечение сканирующего разряда практически не воспринимается наблюдателем. Основное назначение сканирующего разряда состоит в облегчении зажигания разряда в индикаторных ячейках, образованных катодом и противоположным участком анода индикации. Активные заряженные частицы (электроны) проника-

ют в индикаторную ячейку из сканирующей ячейки через отверстия сканирования 5.

Для получения изображения синхронно с разверткой по катодам на индикаторный анод подаются импульсные напряжения. Если в сканирующей ячейке, связанной с индикаторной ячейкой, на анод которой подается импульсное напряжение, горит разряд, то в индикаторной ячейке тоже возникает разряд. После возникновения разряда и свечения в индикаторной ячейке напряжение на аноде индикации уменьшается до уровня, препятствующего возникновению разряда в других ячейках. Горящие, светящиеся индикаторные ячейки формируют изображение.

В заключение отметим, что основные тенденции в развитии ГРП связаны с уменьшением числа схем возбуждения, с возможностью интегрального исполнения части логики и изготовления высоковольтного блока возбуждения в одном керамическом корпусе с индикатором.

Основные параметры современных ГРП: яркость от 100...300 кд/м² до 650 кд/м²; контрастность до 3000 : 1; размер пикселя > 0,3 мм при размерах панели 400 × 300 мм и информационной емкости 1280 × 1024 пикселей; разброс размеров и взаимного расположения ячеек ~ 0,03...0,05 мм; долговечность не хуже 30 000 часов; угол зрения по горизонтали ~ 160°.

15.4. Приборы дугового разряда

В газоразрядных приборах, применяемых в технике, используется особая форма *дугового разряда*, возникающая в разрядном промежутке с накаливаемым катодом. Эта форма разряда называется *низковольтной дугой*.

Дуговой разряд характеризуется малым падением напряжения между электродами (катодом и анодом) и большими токами через прибор. Небольшие напряжения обусловлены тем, что катод разогревается не ионами, а посторонним источником напряжения. Низковольтная дуга может возникнуть при потенциале анода, меньше потенциала ионизации (*нормальная дуга*) или даже потенциала возбуждения (*аномальная дуга*).

Тиратрон дугового разряда — это ионный прибор, содержащий анод, накаливаемый катод, одну или несколько сеток. В качестве рабочей среды в тиратронах дугового разряда используются инертные газы, пары ртути, водород.

При наличии разряда проводимость тиратрона велика (сопротивление мало), а в отсутствие разряда сопротивление очень велико, поэтому тиратроны могут использоваться как управляемые переключатели, выпрямители и т. д. В настоящее время тиратроны применяются в основном в мощных энергетических и электротехнических установках, как и приборы самостоятельно-го дугового разряда — *газотроны* и *игнитроны*. В радиотехнике вместо тиратронов в основном используются тиристоры.

Помимо рассмотренных областей, низковольтная дуга находит широкое применение в таких приборах, как *плазменные ускорители*, которые в настоящее время широко внедряются в технологические процессы микроэлектроники, машиностроения и других отраслей промышленности. Плазменные ускорители используются также в установках термоядерного синтеза, в космических исследованиях, в качестве электроракетных двигателей и т. д.

Высоковольтная дуга (дуга с холодным катодом) широко используется в мощной коммутационной аппаратуре, предназначенной для работы в линиях электропередач, энергетических сетях, электростанциях, в мощных источниках света и т. д.

Высокочастотный (ВЧ) импульсный разряд применяется в разрядниках, которые используются для защиты чувствительных трактов приемных устройств.

ВЧ стационарный и импульсный разряды применяются в термоядерных установках для нагрева плазмы и в газовых и плазменных лазерах для формирования активной среды с нужными свойствами.

В технике и науке находят применение и другие формы газового разряда, такие, как коронный, искровой, факельный и т. д.

15.5. Электровакуумные и электролюминесцентные индикаторы

К электровакуумным индикаторам относятся *электронно-лучевые приборы (ЭЛП)* (гл. 12), *вакуумные накаливаемые индикаторы* и *вакуумные люминесцентные индикаторы*.

Электронно-лучевые приборы. Поскольку ЭЛП рассмотрены ранее в гл. 12, здесь отметим лишь некоторые их особенности как средств отображения информации (СОИ). Основными достоинствами ЭЛП как индикаторов являются следующие: высокая свето-

отдача, хорошая передача цвета и полутонов, простота управления (адресации), широкая полоса пропускания, высокое быстродействие, хорошая разрешающая способность и контрастность, стабильность изображения и т. д. В настоящее время ЭЛП с диагональю экрана в 51 см позволяют отображать до 6 тыс. знаков при ускоряющем напряжении ~ 15 кВ, яркости $100\dots 200$ кд/м². Помимо черно-белых, разработаны и используются ЭЛП с полицветной индикацией.

Возможность реализации полицветных устройств СОИ является несомненным достоинством ЭЛП. В качестве полицветных индикаторов могут использоваться обычные цветные кинескопы. Номенклатура выпускаемых в мире масочных цветных кинескопов очень велика. Цветные кинескопы с тремя прожекторами сложны в изготовлении и настройке. Этих недостатков в значительной мере лишен ЭЛП типа тринитрон, в котором реализован способ получения трех лучей с помощью одного прожектора. Три электронных потока, испускаемые тремя катодами, фокусируются первой общей линзой таким образом, что они пересекаются в одной точке (первое скрещение). Боковые лучи после точки пересечения смещаются отклоняющей системой так, что все три луча без дальнейшей фокусировки сходятся на люминофорном экране. Таким образом, первая большая апертурная линза осуществляет фокусировку лучей, а отклоняющее устройство — сведение лучей. В тринитроне можно получить цветное изображение примерно в 1,5 раза ярче, чем в обычном цветном кинескопе. Анодное напряжение при этом составляет $25\dots 27$ кВ, фокусирующее напряжение $5\dots 6$ кВ, ток луча $0,5$ мА.

В настоящее время разработаны и эксплуатируются различные варианты полицветных ЭЛП; в наиболее сложных из них используются одна пушка в комбинации с тремя источниками напряжения, специальная система сканирования и три люминофора. В таких приборах реализуются три режима работы: при напряжении 8 кВ возбуждается люминофор с зеленым свечением, при 11 кВ возбуждаются вместе красное и зеленое свечения, что позволяет получить желтый цвет, при напряжении 18 кВ возбуждается люминофор с оранжевым длительным послесвечением. Такой ЭЛП может работать в радиолокаторах.

В настоящее время продолжают работы по созданию плоских и совершенствованию проекционных ЭЛП.

Вакуумные накаливаемые индикаторы. В вакуумных накаливаемых индикаторах (ВНИ) происходит нагрев тела до температуры $(2...3) \cdot 10^3$ К под действием электрического тока, что вызывает его яркое свечение. Нагреваемое твердое тело выполняется либо из вольфрама, рения, гафния (в виде нитей накала), либо на полупроводниковой основе (SiC).

Для обеспечения большей долговечности ВНИ эксплуатируются при пониженном напряжении накала, обеспечивая яркость до $500...600$ кд/м². Такая яркость позволяет использовать ВНИ при высоком уровне окружающей освещенности. Номинальные напряжения ВНИ составляют $2...5$ В, время переходных процессов — $10^{-4}...10^{-2}$ с, потребляемый ток — $12...15$ мА/segment (для сегментных ВНИ), потребляемая мощность не превышает $50...60$ мВт/segment, угол обзора $> 90^\circ$, средний срок службы $\sim 10^5$ часов. Промышленностью выпускаются в основном два типа ВНИ: собственно лампы накаливания и сегментные вакуумные накаливаемые индикаторы. Сегментные ВНИ имеют от 4 до 10 сегментов, что позволяет отображать цифры от 0 до 9, буквы русского и некоторые буквы латинского алфавитов.

Вакуумные люминесцентные индикаторы (ВЛИ). ВЛИ в отличие от ЭЛП используют низковольтную катодолюминесценцию, вызываемую электронным потоком. Низковольтная катодолюминесценция возникает в люминофорах на проводящей основе ($ZnO : Zn$, $Zn : Cl$, $SnO_2 : Eu$), смешанных люминофорах ($ZnS : Ag + In_2O_2$) и некоторых других. Теоретически яркость низковольтной катодолюминесценции может достигать величины ~ 1600 кд/м² при плотности тока ~ 1 мА/м², энергии электронов ~ 20 эВ и энергетическом КПД $\sim 25...30\%$. При малой энергии электронов, бомбардирующих люминофор, свет излучается из поверхностного слоя толщиной в несколько атомных слоев. Из-за малой глубины проникновения первичных электронов плотность возбужденных атомов в поверхностных слоях люминофора велика, что приводит к насыщению яркости в зависимости от плотности тока, перегреву люминофора и температурному гашению люминесценции.

Конструктивно ВЛИ выполняются в виде цилиндрических и плоских баллонов. Цилиндрические ВЛИ бывают как одно-, так и многоразрядными, а плоские — только многоразрядными. Выпускаются также матричные, сегментные, аналоговые и другие типы ВЛИ. Большинство индикаторов выполняется вместе

со схемой управления и питания. Потребляемая мощность у 10-разрядного ВЛИ 1,33 Вт, у 40-разрядного — 2,66 Вт, среднее время наработки на отказ 10^5 ч, высота знаков от 5 до 15 мм, яркость свечения до 700 кд/м². На основе матричных ВЛИ разработаны индикаторные модули. Основой ВЛИ является стеклянная, либо керамическая пластина 1 (рис. 15.7), в сегментных углублениях которой выполняются

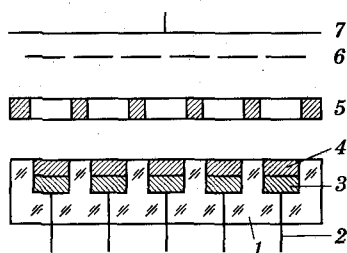


Рис. 15.7

проводящие слои 3 с выводами 2 от каждого сегмента. Поверх проводящих слоев наносится люминофор 4. Плата с проводящими покрытиями выполняет роль анода. Над платой устанавливается металлический экранирующий электрод с отверстиями 5, расположенными напротив соответствующих сегментов. На некотором расстоянии от экранирующего электрода монтируется сетка 6, а за ней прямонакальный оксидный катод 7. Стеклянный баллон изнутри покрывается проводящим слоем. Отрицательным по отношению к катоду напряжением на сетке можно полностью запирает электронный поток и прекращать свечение люминофора. С помощью матричных ВЛИ получают телевизионное изображение удовлетворительного качества.

Электролюминесцентные индикаторы. В этих приборах используется предпробойная электролюминесценция, которая возникает на границах зерен микроучастков порошковых и пленочных электролюминофоров при напряженностях электрических полей, близких или равных пробивным.

В настоящее время на практике получили распространение следующие виды электролюминесцентных активных индикаторов: порошковые, пленочные и органические. На рис. 15.8, а

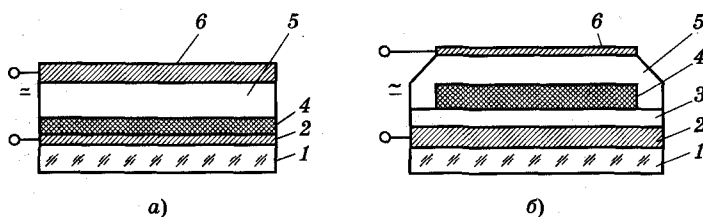


Рис. 15.8

схематически представлено возможное устройство порошкового, а на 15.8, б — пленочного индикаторов.

Порошковые и пленочные индикаторы изготавливаются посредством напыления прозрачного электрода 2 на стеклянную подложку — пластину 1, с последующим нанесением изоляционных слоев 3 (см. рис. 15.8, б) и 5 (см. рис. 15.8, а, б) люминофора 4 и непрозрачного электрода 6.

В качестве излучающего слоя в порошковых индикаторах используются смеси диэлектрической связки либо с порошковым люминофором ZnS при наличии примеси свинца, марганца, свинца с медью, либо с $ZnSe$, $ZnSiO_4$. В пленочных индикаторах люминофором служит пленка $ZnS : TbF_3$ толщиной около 0,2 мкм. Этот люминофор дает свечение зеленого цвета. Свечение красного цвета можно получить посредством введения фторида самария (SmF_3), синего цвета — фторида тербия (TbF_3). В порошковых и пленочных индикаторах происходят во многом аналогичные физические процессы. При приложении постоянного напряжения к порошковому или пленочному люминофору уровень Ферми и границы зон перехода металл—изолятор смещаются таким образом, что происходит туннелирование дырок и электронов в пленку или зерна люминофора. В результате этого возбуждаются ионы примеси Mn , Pb и т. д. или TbF_3 , SmF_3 и возникает излучательная рекомбинация. Рассмотренный механизм реализуется в приборах, где металлический электрод находится в контакте с люминофором и при питании индикатора постоянным током. В общем случае механизм свечения пленочных и порошковых электролюминесцентных слоев обусловлен рекомбинацией носителей заряда, инжектированных кристаллом люминофора и электродами или образованных в результате туннельного эффекта и ударной ионизации. Электролюминесцентные индикаторы зачастую имеют диэлектрические слои между электродами и люминофором (см. рис. 15.8, б). Поскольку пробивная напряженность диэлектрика превышает пробивную напряженность люминофора, то при увеличении приложенного напряжения ZnS пробивается раньше и «горячие» электроны возбуждают ионы примеси.

Яркость индикаторов, работающих на переменном токе, растет с увеличением частоты питающего напряжения (рис. 15.9). С увеличением частоты возрастает и напряженность электрического поля в люминофоре, поскольку уменьшается емкостное сопротивление между металлическими электродами и люмино-

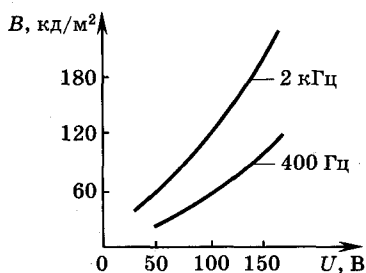


Рис. 15.9

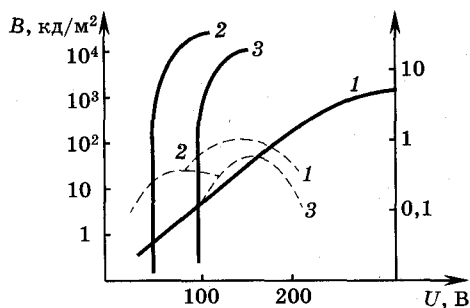


Рис. 15.10

фором, и, соответственно, возрастает яркость. Однако во многих материалах увеличение частоты питания (выше 10 кГц) приводит к значительным диэлектрическим потерям при незначительном увеличении яркости.

Основное преимущество индикаторов переменного тока по сравнению с индикаторами постоянного тока — существенно большая светоотдача. На рис. 15.10 приведены результаты сравнения яркости V (сплошные линии) и световой отдачи η (штриховые линии) (см. п. 12.5) для различных типов индикаторов.

Кривые 1 дают зависимости V и η от напряжения питания U для порошковых индикаторов переменного тока ($f = 5$ кГц), кривые 2 — для порошковых индикаторов постоянного тока; кривые 3 — для пленочных индикаторов переменного тока ($f = 5$ кГц).

Отечественная промышленность выпускает свыше 20 типов индикаторов с различными цветами свечения, которые позволяют отображать цифры, буквы, знаки, геометрические фигуры и т. д.

15.6. Пассивные индикаторы. Жидкокристаллические индикаторы

Жидкокристаллические индикаторы (ЖКИ). Жидкокристаллическое, или мезоморфное состояние — это состояние вещества, при котором оно обладает свойствами, присущими как твердым кристаллам, так и жидкостям. Достоинства ЖКИ по сравнению с другими индикаторами: малая потребляемая мощность (единицы мкВт/см²); низкие рабочие напряжения (для большинства ЖКИ это 1,5...5 В) и хорошая совместимость с КМОП интег-

ральными схемами¹ (см. гл. 7); простота конструктивного исполнения при малой толщине индикатора; большая долговечность.

К основным недостаткам можно отнести сравнительно низкое быстродействие, ограниченный угол обзора и необходимость внешнего освещения.

Жидкие кристаллы (ЖК), как и твердые, имеют анизотропию физических свойств, обладают упорядоченностью ориентации и в то же время им, как и жидкости, свойственна текучесть, поверхностное натяжение и вязкость. Выделяют следующие два основных типа ЖК: *смектические* (текучие кристаллы), *нематические* (собственно жидкие кристаллы). Смектические ЖК состоят из сильно вытянутых, с длиной до 3 мкм и шириной до 1 мкм, линейных молекул, которые образуют слои, расположенные один под другим на равном расстоянии и имеющие одинаковую толщину, близкую к длине молекул. В нематических ЖК не происходит образования отдельных слоев. Разновидностью смектических ЖК являются *холестерические вещества*, в которых ориентация молекул меняется от слоя к слою. Жидкие кристаллы имеют среднее удельное сопротивление $10^6 \dots 10^{10}$ Ом·см, зависящее от наличия примеси. Диэлектрическая проницаемость ϵ и показатель преломления n зависят от ориентации молекул и поэтому являются анизотропными величинами. В отличие от активных индикаторов ЖКИ не генерируют оптическое излучение, а модулируют его интенсивность за счет изменения таких его характеристик, как амплитуда, фаза, длина волны, плоскость поляризации и направление распространения. Так, например, при прохождении света через ЖК происходит поворот плоскости поляризации электромагнитной волны на угол $\sim 7 \cdot 10^4$ град/мм, в то время как у кварца эта величина порядка 15 град/мм. В ЖКИ запись информации происходит в результате целенаправленных изменений параметров жидкокристаллической среды под действием электрических, магнитных и акустических полей. Эти изменения приводят к вариации коэффициентов отражения, поглощения и рассеяния, изменению показателя преломления. Конструкция ЖКИ очень проста: между двумя стеклянными пластинами, на которые нанесены полупрозрачные электроды, заключено жидкокристаллическое вещество, толщина которого определяется прокладками между пластинами (10...20 мкм).

¹ Комплементарные металл—оксид—полупроводниковые интегральные схемы.

Принцип действия ЖКИ основан на изменении оптических свойств ЖК под действием электрического поля. Для этих изменений, как правило, требуются малые напряжения и низкая потребляемая мощность. На практике используются ЖКИ, работающие на просвет и отражение. В первом типе ЖКИ обе стеклянные пластины прозрачны. При работе ЖКИ на отражение второй электрод должен быть зеркально отражающим. Под индикатором, работающим в проходящем свете, размещается источник света и матово-черный экран. Между верхними фигурными электродами, формирующими определенный знак, и нижним общим электродом подается управляющее напряжение от единиц до десятков В, ток на один знак достигает ~ 1 мкА, т. е. затраты плотности мощности на один знак равны ~ 1 мкВт/см². Различные типы ЖКИ работают в диапазоне температур от -20 до 55° С. Долговечность ЖКИ составляет десятки тысяч часов.

Рассмотрим более подробно основные физические эффекты, используемые в различных типах ЖКИ. Наиболее широкое практическое применение нашли ЖКИ на основе нематических жидких кристаллов (НЖК), обладающих оптической и диэлектрической анизотропией. Обычно меру диэлектрической анизотропии определяют величиной $\Delta\epsilon = \epsilon_{\parallel} - \epsilon_{\perp}$, где ϵ_{\parallel} и ϵ_{\perp} — соответственно диэлектрическая проницаемость вдоль и поперек направления преимущественной ориентации молекул ЖК. Величина и знак $\Delta\epsilon$ определяются частотой электрического поля. Обычно величину $\Delta\epsilon$ определяют на частотах порядка десятков кГц. Если $\Delta\epsilon > 0$, то длинные оси молекул НЖК стремятся расположиться вдоль поля, а при $\Delta\epsilon < 0$ — поперек поля. В основном именно эти свойства НЖК, вызывающие электрооптические эффекты, связанные с движением вещества (*динамическое рассеивание*), с поворотом молекул и плоскости поляризации света в электрическом поле (*твист-эффект*), с анизотропией оптического поглощения растворов дихроичных красителей («гость») в НЖК («хозяин») (*эффект «гость—хозяин»*), используются в подавляющем большинстве ЖКИ.

Эффект динамического рассеивания реализуется в НЖК с $\Delta\epsilon < 0$, куда вводится легирующая примесь, приводящая к образованию свободных ионов. Под действием внешнего переменного электрического поля происходит ориентация большей доли молекул НЖК в определенном направлении, и, кроме того, возникают колебания объемного заряда. В результате при напряженностях внешнего поля, меньших некоторой пороговой величины, в ЖК

устанавливается течение молекул в виде упорядоченных цилиндрических вихрей. При пороговой напряженности заметные токи проводимости разрушают собранную полем упорядоченную структуру из-за возникновения турбулентности (турбулентного течения), что в конечном счете приводит к сильному рассеиванию света. Таким образом, при напряженностях, меньших пороговой, устанавливается гомогенная ориентация молекул, при которой оптическая ось образовавшейся структуры ЖК перпендикулярна лучу падающего света, а при напряженностях, больших пороговой, возникает их хаотическая ориентация, приводящая к деполяризации света.

Если задние электроды сделать зеркальными, а передние полупрозрачными, то при напряженности поля между электродами меньше пороговой ЖК будет прозрачным и зеркальные электроды будут хорошо видны в отраженном свете. При критических или еще больших напряженностях ЖК становится мутным, при этом зеркальные электроды будут не видны, т. е. если установить на обеих пластинах систему электродов и менять управляющее напряжение, можно получить то или иное видимое изображение. Устройство ячейки на эффекте динамического рассеивания представлено на рис. 15.11, а, где 1 — стеклянные пластины, 2 — прозрачный электрод, 3 — изоляционная прокладка, 4 — прозрачный электрод, 5 — слой ЖК, 6 — выводы.

ЖКИ на эффекте динамического рассеивания характеризуются уровнями потребляемой мощности $5 \dots 10 \text{ мВт/см}^2$ для плотности постоянного тока $j = 0,5 \dots 1 \text{ мкА/см}^2$ и $50 \dots 200 \text{ мВт/см}^2$

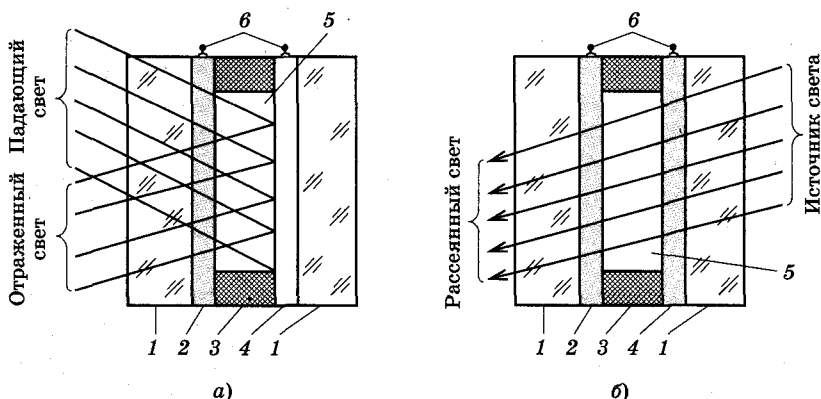


Рис. 15.11

для переменного тока $j = 2 \dots 10$ мкА/см². Рабочее напряжение для этих приборов обычно не превышает 20 В, срок службы (долговечность) на переменном токе — более 40 тыс. часов.

Твист-эффект основан на повороте плоскости поляризации света, возникающем при распространении света вдоль оси холестерической спирали или при скручивании нематической структуры. В нематических ЖК скручивания структуры добиваются специальной обработкой поверхности пластин, что позволяет ориентировать длинные оси приповерхностных молекул ЖК или перпендикулярно плоскости пластин (**гометронная ориентация**), или параллельно некоторому общему направлению в плоскости пластин (**гомогенная ориентация**). Если длинные оси молекул различных областей ЖК параллельны разным направлениям плоскости пластин, то ориентация называется **квазигомогенной**. В отсутствие внешних полей ориентация молекул в объеме жидкокристаллической ячейки определяется ориентацией молекул, находящихся вблизи поверхности стеклянных пластин, при этом укладка молекул вблизи обеих пластин гомогенная, но направление осей молекул вблизи одной пластины составляет угол 90° с этим направлением у другой пластины. Таким образом, в объеме ЖК возникает спиральная укладка нематических плоских структур, при этом поворот плоскостей составляет всего четверть витка спирали. Проходящий через такой ЖК плоскополяризованный свет превращается в слабо эллиптически-поляризованный, азимут которого повернут на 90° относительно поляризации падающего света. Если с обеих сторон от ячейки поставить скрещенные поляризатор и анализатор, то в отсутствие управляющего напряжения на электродах свет свободно проходит через анализатор и данный участок ЖК-ячейки будет светлым. При напряжениях на электродах больше порогового плоскость поляризации света не будет вращаться и тогда свет не будет пропускаться скрещенными поляризаторами, т. е. участок ячейки будет темным.

ЖКИ, использующий твист-эффект, представлен на рис. 15.11, б, где использованы те же обозначения, что и на рис. 15.11, а.

Эффект «гость—хозяин» использует анизотропию оптического поглощения дихроичных (от греч. «двухцветных») красителей в нематическом жидком кристалле. Введение молекул красителя («гостя») в НЖК в небольших концентрациях (1...2%) сохраняет температуры фазовых переходов и вязкоупругие свойства смеси

на уровне НЖК-хозяина, но при этом заметно изменяется электропроводность и оптические свойства красителя. Контраст изображения в ЖКИ на эффекте «гость—хозяин» зависит от оптических свойств красителя, толщины смеси, концентрации в ней красителя и условий наблюдения. Оптическая анизотропия в рассматриваемом эффекте обусловлена тем, что молекулы НЖК эффективно ориентируют молекулы красителя в определенном направлении.

Поскольку ЖКИ — пассивные устройства, то основным их оптическим параметром является контрастность, а не яркость, как в активных индикаторах. Контрастность знака по отношению к фону определяется коэффициентом K , который представляет отношение разности коэффициентов яркости фона (B_Φ) и знака индикатора (B_a) к коэффициенту яркости фона, выраженное в процентах, т. е. $K = \frac{B_\Phi - B_a}{B_\Phi} 100\%$. К другим параметрам

ЖКИ относятся ток потребления I_Π , напряжение управления $U_{упр}$, рабочая частота напряжения управления $f_{раб}$, минимальное напряжение управления $U_{упр. мин}$, максимально допустимое напряжение управления $U_{упр. макс}$, время реакции $t_{реак}$ и время релаксации $t_{рл}$. Одной из основных характеристик ЖКИ является вольт-контрастная.

ЖКИ обычно питают переменным напряжением, чтобы избежать направленных электрохимических процессов, которые сокращают срок службы индикаторов, поэтому вольт-контрастная характеристика ЖКИ представляет зависимость контрастности изображения K от действующего значения приложенного напряжения управления $U_{упр}$. С увеличением напряжения $U_{упр}$ до некоторого порогового значения $U_{пор}$ контраст сильно возрастает. При $U_{упр} > U_{пор}$ темп увеличения контраста с увеличением $U_{упр}$ становится много меньше. Значение $U_{упр. мин}$ выбирается на пологом участке кривой вблизи $U_{пор}$, т. е. $U_{упр. мин} \approx U_{пор}$.

Рассмотрим некоторые характеристики ЖКИ, связанные с потребляемым током. Вольт-амперная характеристика (ВАХ) ЖКИ представляет зависимость тока потребления $I_\Pi = f(U_{упр})$ при $T = \text{const}$ и $f_{раб} = \text{const}$ (рис. 15.12, а).

Ток потребления I_Π зависит от управляющего напряжения, рабочей частоты, температуры и конструктивно-технологичес-

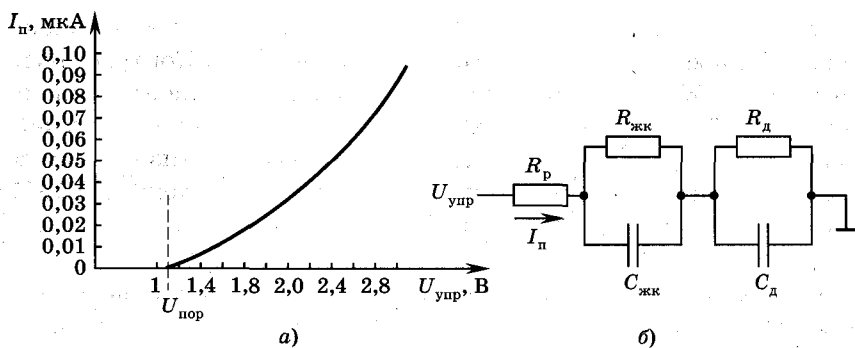


Рис. 15.12

ких особенностей ЖКИ. В самом общем случае $I_{п}$ включает активную (ионную) и реактивную (емкостную) составляющие, что достаточно хорошо иллюстрируется эквивалентной схемой на рис. 15.12, б, где R_p — сопротивление разводки, $R_{жк}$, $C_{жк}$ — соответственно сопротивление и емкость ЖК-слоя, R_d , C_d — сопротивление и емкость двойного электрического перехода.

Емкостная составляющая тока $I_{п}$ является определяющей на большей части ВАХ. На низких частотах ВАХ ЖКИ на твист-эффекте практически линейна. Нелинейность ВАХ при малых напряжениях (см. рис. 15.12, а), близких к $U_{пор} \sim U_{упр. мин}$, обусловлена ростом емкости ЖКИ, что связано с поворотом слоев длинных осей молекул НЖК в направлении электрического поля.

На низких и не очень высоких частотах управляющего напряжения из-за того, что $C_{жк} \ll C_d$, частотная зависимость $I_{п}$ будет определяться процессом заряда емкости $C_{жк}$ через малое сопротивление R_p и линейно зависеть от частоты (рис. 15.13, участок I). Начиная с частоты 15...20 кГц линейность частотной характеристики $I_{п} = F_1(f)$ может нарушаться (см. рис. 15.13, участок II) из-за изменения соотношения между емкостным сопротивлением ЖК и активным и реактивным сопротивлениями разводки.

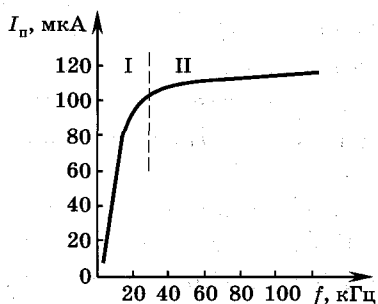


Рис. 15.13

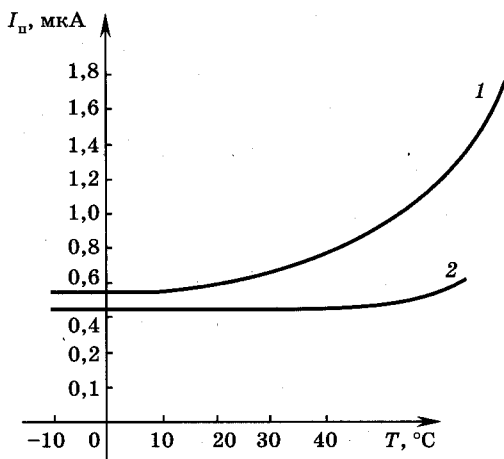


Рис. 15.14

Температурная зависимость $I_{\text{п}} = f(T)$ также определяется соотношением между емкостной и полной составляющими тока $I_{\text{п}}$ и может быть значительной, если доля ионной составляющей велика (рис. 15.14, кривая 1, доля ионного тока 50%). При малой доле ионного тока зависимость $I_{\text{п}}$ от температуры очень незначительна (рис. 15.14, кривая 2, доля ионного тока 20%).

Временные характеристики ЖКИ связаны со скоростью протекания электрооптических эффектов. Времена реакции $t_{\text{реак}}$ и релаксации $t_{\text{рл}}$ пропорциональны вязкости жидкости, квадрату толщины ЖК и приблизительно обратно пропорциональны квадрату разности напряжений между приложенным и пороговым значениями. Время реакции $t_{\text{реак}}$ определяется интервалом времени при включении индикатора, в течение которого ток потребления $I_{\text{п}}$ увеличивается до 0,8 от максимального значения; время релаксации — интервал времени при выключении, в течение которого ток потребления снижается до 0,2 от максимального значения. Поскольку скорость переключения определяется толщиной ЖК, его толщину делают возможно малой.

В настоящее время разработаны ЖКИ с цветными и полицветными изображениями. Цветное изображение возможно получить различными методами: например, изменять цвет индикатора, осуществлять циклическое изменение цветов в системе изображения, создавать фотографические объекты с различными характеристиками поглощения путем добавки дихроичных

красителей к ЖК. При ориентации молекул красителя за счет электрического поля для получения желаемого цвета применяют ЖК с отличающимися свойствами по сравнению с ЖК, используемыми для динамического рассеивания света. НЖК, в который вводится краситель, должен иметь большой дипольный момент вдоль длинной оси молекулы. При подаче напряжения молекулы красителя ориентируются в соответствии с направлением поля, поскольку дипольный момент НЖК совпадает с направлением поля, что и вынуждает молекулы красителя ориентироваться. Выключение напряжения заставляет молекулы НЖК (хозяин) вместе с молекулами красителя (гость) перестраиваться в исходную конфигурацию, и ячейка принимает свой первоначальный цвет. Для получения цветного изображения в ЖКИ применяются также оптические фильтры с электронной настройкой. Применение разноцветных слоев ЖК позволяет создать фильтры с электронным управлением цветом во всей видимой части спектра. За счет изменения концентрации красителя можно оптимизировать контрастность. Если поддерживать концентрацию красителя на уровне 1% или ниже, то можно получить максимальную контрастность. В этих условиях достигается наибольшее изменение адсорбирующих свойств. Это изменение убывает на половину своего максимального значения при концентрации красителя $\geq 5\%$.

Использование фотохромных (одноцветных) красителей, изменяющих свой цвет при различных длинах волн падающего света, позволяет также построить ЖК переключатели цвета. В фотохромных красителях можно изменять цвет как электронными, так и оптическими средствами. В одной и той же ячейке ЖКИ используются смеси красителей различного типа, например плеохроических (многоцветных) и фотохромных. Это позволяет изменять цвет в пределах принимаемого спектра (плеохроические красители) с использованием электрического поля или при облучении светом определенной длины волны (фотохромные красители), либо обоими способами одновременно. Плеохроизм, частный случай дихроизма, — различное поглощение анизотропным веществом проходящего через него света (а следовательно, и различное окрашивание) в зависимости от направления его распространения и поляризации.

Электрохромные индикаторы (ЭХИ). В ЭХИ используется явление изменения цвета под действием приложенного электрического поля. В настоящее время разработаны ЭХИ на основе ре-

акции восстановления металлических окислов WO_3 и реакции восстановления органических соединений — виологенов (компаундов и других соединений) на поверхности электродов. Конструктивное исполнение электрохромных индикаторов во многом подобно ЭЛИ.

В работе ЭХИ на основе тонких пленок WO_3 можно выделить три стадии: запись, запоминание, стирание. Запись и стирание реализуются при подаче на электроды ЭХИ постоянного напряжения (1,0...1,5 В) или тока, при этом в первом случае ток, протекающий между электродами, будет переменным, а во втором — напряжение между электродами будет зависеть от времени.

ЭХИ имеют время реакции ~ 50 мс при $T = 300$ К, срок службы $\sim 10^7$ циклов, работают при низких температурах (до -25 °С). Энергетические характеристики ЭХИ: минимальный заряд, требующийся для окрашивания ЭХИ на WO_3 , — 4 мКл/см², на виологенах — 2 мКл/см², потребляемая мощность при напряжении 1 В соответственно 4 и 2 мДж/см², рассеиваемая мощность для сигнала с частотой 1 Гц и скважностью 2 — от 2 до 8 мВт/см². Достоинства ЭХИ: малая потребляемая мощность, широкий диапазон рабочих температур, малые рабочие напряжения (~ 1 В), большая контрастность, не зависящая от угла зрения. Недостатки ЭХИ: небольшой выбор цветов, большие времена реакции и релаксации, относительно малый срок службы.

Электрофоретические индикаторы (ЭФИ). *Электрофорез* — направленное движение дисперсных частиц в жидкости (коллоидных и взвешенных твердых частиц, пузырьков газа, капель другой жидкости) под действием электрического поля.

В ЭФИ реализуется движение взвешенных заряженных пигментных частиц в окрашенной безводной жидкости под действием электрического поля. В результате электрофореза изменяется коэффициент отражения и цвет рабочего вещества.

ЭФИ состоит из слоя коллоидной жидкости толщиной 25...100 мкм, помещенной между двумя стеклянными пластинами с расположенными на их внутренних поверхностях прозрачными электродами.

Типичные параметры элемента ЭФИ толщиной 50 мкм: время перемещения пигментных частиц 100 мс при максимальном значении в импульсе напряжения 30 В и плотности тока ~ 1 мкА/см², потребляемая мощность < 12 мкВт/см² для импульсов с частотой повторения 1 Гц и скважностью 2, рекомендуемый

диапазон температур $-15...+50^{\circ}\text{C}$, ресурс работы $\sim 10^8$ циклов. Увеличение температуры приводит к возрастанию рассеиваемой мощности.

15.7. Сравнение индикаторов различного типа

Выбор количественных мер для оценки качества индикаторов является достаточно сложной задачей, так как нет общих методов и критериев оценки из-за весьма разнообразного назначения индикаторов, поскольку крайне сложно учитывать совместно набор параметров среды, оказывающих отрицательное влияние на оператора, и полезный эффект от работы индикатора. Реакция оператора на информацию, получаемую с помощью индикатора, часто является весьма субъективной, поэтому трудно разработать качественные методы моделирования, удовлетворяющие заданным требованиям. В силу этого часто используется простейший способ — сравнение достоинств и недостатков приборов.

Перечень параметров для оценки качества индикаторов включает следующие показатели: яркость, разрешающую способность, контрастность, число градаций яркости, цвет, размеры, угол обзора, времена реакции и релаксации, память, долговечность, надежность, экономичность, возможность мультиплексного управления и ряд других. Поскольку простейший способ оценки носит описательный характер, то выделим наиболее характерные достоинства индикаторов.

Полупроводниковые индикаторы имеют яркое и насыщенное свечение, легко управляются, хорошо сочетаются с интегральной техникой, экономичны, технологичны, долговечны.

Электронно-лучевые индикаторы обеспечивают хорошую информационную емкость, яркость в широкой цветовой гамме, совмещение изображений различных типов, высокое быстродействие.

Вакуумно-люминесцентные индикаторы отличаются хорошими эргономическими параметрами: большим углом обзора, экономичны. Они позволяют реализовать непосредственное сопряжение с МОП-интегральными схемами; находят широкое применение как многообразные индикаторы.

Газоразрядные индикаторные панели могут иметь большие размеры, высокую разрешающую способность, точность и надежность, обладают цифровой матричной адресацией высокой точности, обеспечивают совмещение функций индикации с адресацией и запоминанием информации.

Пассивные индикаторы имеют низкий уровень управляющих напряжений, малую потребляемую мощность, простую конструкцию; они технологичны, хорошо совмещаются с микросхемами.



Контрольные вопросы

1. Классификация газоразрядных приборов и индикаторов.
2. Каковы элементарные процессы в плазме газовых разрядов?
3. Процессы в тлеющем разряде и других типах разрядов, включая высокочастотные ВАХ газового разряда с холодным катодом.
4. Пространственное распределение потенциалов электрического поля, объемных зарядов и светимости областей газового разряда.
5. Индикаторные тиратроны.
6. Плазменные панели и их разновидности, устройство и принцип функционирования.
7. Плазменные панели на переменном и постоянном токе, физические процессы, параметры, характеристики.
8. Приборы дугового разряда.
9. Каковы особенности устройства, процессов и характеристик электровакуумных и электролюминесцентных индикаторов?
10. Электрохромные и электрофоретические индикаторы, принцип устройства и функционирования, достоинства и недостатки.
11. Жидкокристаллические индикаторы: классификация, принцип устройства и функционирования ЖКИ различных типов, их параметры и характеристики.

Глава 16

ОПТОЭЛЕКТРОННЫЕ ПРИБОРЫ

16.1. Общие сведения

Оптоэлектроника — область науки и техники, где процессы взаимодействия оптического излучения с веществом и оптическое излучение вещества используются для изучения фундаменталь-

ных свойств материи, а также для создания приборов, передачи, приема, переработки, хранения и отображения информации на основе взаимного преобразования электрических и оптических сигналов.

Оптический диапазон спектра включает электромагнитные волны длиной от 1 мм до 1 нм. Этот интервал, определенный Международной комиссией по освещению, достаточно условен. Часто длинноволновую область оптического диапазона ($\lambda \sim 1 \dots 0,1$ мм) относят к радиоволнам СВЧ-диапазона, а коротковолновую ($\lambda = 80 \dots 1$ нм) — к мягкому рентгеновскому излучению, а оптический диапазон разбивается на три поддиапазона: *инфракрасный* (ИК; $\lambda = 0,78 \dots 1000$ мкм), *видимый* ($\lambda = 0,38 \dots 0,78$ мкм) и *ультрафиолетовый* (УФ; $\lambda = 0,001 \dots 0,38$ мкм).

Широкое освоение оптического диапазона определяется рядом принципиальных преимуществ световых волн по сравнению с радиоволнами:

- 1) большей информационной емкостью оптической связи, что обусловлено значительно более высокой частотой световых волн (в видимом участке спектра $f \sim 10^{15}$ Гц; для передачи обычного телевизионного изображения необходима полоса частот $\Delta f = 6$ МГц, поэтому в УКВ и дециметровом диапазонах можно разместить до нескольких сотен телевизионных каналов; в оптическом диапазоне это число возрастает до сотен миллионов и более);
- 2) высокой направленностью излучения из-за малого отношения длины волны к размерам апертуры излучателя;
- 3) возможностью реализации идеальной гальванической связи входа и выхода, однонаправленностью потока информации, высокой помехозащищенностью, исключением взаимных наводок и паразитных связей между различными элементами схемы (все это достигается за счет того, что фотоны являются электрически нейтральными частицами, которые не взаимодействуют между собой и с внешним электрическим и магнитным полями);
- 4) высокой плотностью записи информации в оптических запоминающих устройствах, что открывает новые перспективы для построения ЭВМ последующих поколений.

Для реализации этих преимуществ необходимы прежде всего оптоэлектронные приборы, имеющие соответствующие характеристики. *Оптоэлектронные приборы* — это устройства, в которых основные процессы происходят с участием фотонов. В зависимос-

ти от особенностей протекающих процессов все оптоэлектронные приборы можно разделить на три группы:

- 1) **светоизлучатели**, преобразующие электрическую энергию в оптическое излучение (светодиоды, полупроводниковые лазеры, люминесцентные конденсаторы);
- 2) **фотоприемники** (фотодетекторы), которые преобразуют оптическое излучение в электрические информационные сигналы (фоторезисторы, фотодиоды, фототранзисторы и т. д.);
- 3) **солнечные преобразователи**, преобразующие оптическое излучение в электрическую энергию (солнечные батареи, фотовольтаические приборы).

Помимо указанных приборов, в различных областях науки и техники широко используются **оптоэлектронные пары** — полупроводниковые приборы, состоящие из светоизлучающего и фотоприемного элементов, между которыми существует оптическая связь по оптическому каналу, обеспечивающему электрическую изоляцию между входом и выходом. Светоизлучатель, фотоприемник и оптический канал, реализующий гальваническую развязку между входом и выходом, конструктивно объединены в одном корпусе.

16.2. Оптические явления в полупроводниках

В этом параграфе описаны физические явления, лежащие в основе работы оптоэлектронных полупроводниковых приборов. Работа как отдельных элементов, так и оптоэлектронных систем в целом построена на основе использования различных оптических явлений в твердых телах.

Поглощение света. Результат взаимодействия света с веществом часто оценивают с помощью коэффициентов отражения (R), поглощения (α) и пропускания (T):

$$R = I_{\text{отр}}/I_0; T = I_{\text{пр}}/I_0; \alpha l = I_{\text{погл}}/I_0,$$

где $I_{\text{отр}}$, $I_{\text{пр}}$, $I_{\text{погл}}$, I_0 — соответственно интенсивность отраженного, прошедшего, поглощенного и падающего на вещество света; l — расстояние, на котором интенсивность падающего света I_0 ослабляется веществом в e раз.

Коэффициент поглощения α (интегральный) обычно измеряется в см^{-1} . Он определяет долю поглощенной веществом энергии из светового пучка во всем спектральном диапазоне в слое

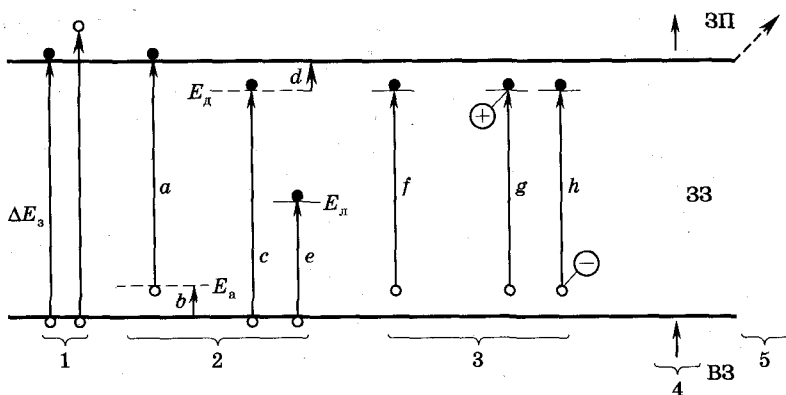


Рис. 16.1

толщиной (длиной) $l = 1$ см. Зависимость спектрального коэффициента поглощения от частоты падающего на вещество света $\alpha(\nu)$ или от длины волны $\alpha(\lambda)$ называют спектром поглощения.

Зависимость $R(\nu)$ или $R(\lambda)$ называют спектром отражения. На микроскопическом (квантовом) уровне процессы поглощения и излучения энергии веществом определяются переходами частиц (атомов, молекул) из одного энергетического состояния в другое, при этом происходит изменение энергетического состояния свободных и связанных с атомами электронов или изменение колебательной и вращательной энергии атомов в молекулах.

Различают пять основных типов оптического поглощения в полупроводниках: 1) *собственное*, 2) *примесное*, 3) *экситонное*, 4) *поглощение свободными носителями заряда*, 5) *решеточное*.

На рис. 16.1 схематически представлены основные оптические переходы в полупроводнике, связанные с указанными типами поглощений оптического излучения.

Рассмотрим кратко физические механизмы каждого вида поглощения.

Собственное, или фундаментальное, поглощение. Этот вид поглощения ($\alpha \sim 10^4 \dots 10^8 \text{ см}^{-1}$) связан с возбуждением под действием падающего излучения электронов и переходом их из валентной зоны в зону проводимости. Соответствующие переходы отмечены на рис. 16.1 позицией 1. Вероятность этих переходов (а следовательно, и обратных переходов, сопровождаемых оптическим излучением) зависит от строения энергетиче-

ческих зон полупроводников, которые можно разделить на две группы. В первой из них минимуму энергии в зоне проводимости соответствует импульс $p_{\text{мин}}$, а максимуму энергии в валентной зоне — импульс $\bar{p}_{\text{макс}}$, причем $\bar{p}_{\text{мин}} = \bar{p}_{\text{макс}}$, т. е. экстремумы зоны проводимости и валентной зоны в пространстве импульсов (или волновых чисел) совпадают. У второй группы полупроводников экстремумы зоны проводимости и валентной зоны находятся при различных \bar{p} (см. рис. 3.17), т. е. $\bar{p}_{\text{макс}} - \bar{p}_{\text{мин}} = \Delta p \neq 0$.

В качестве примера полупроводников первой группы можно привести антимонид индия и арсенид галлия, а второй — кремний и германий.

Для полупроводников первой группы переходы электронов через запрещенную зону с наибольшей вероятностью происходят между энергетическими состояниями, соответствующими значениям $\Delta \bar{p} \sim 0$; такие переходы называются *прямыми*. Для прямых переходов поглощение квантов с энергией, меньшей ширины запрещенной зоны, не происходит, поэтому спектр собственного поглощения со стороны длинных волн (или малых энергий) должен резко обрываться (пунктирные кривые на рис. 16.2). Это так называемая *красная (длинноволновая) граница внутреннего фотоэффекта*, соответствующая длине волны $\lambda = \lambda_{\text{кр}}$. Для полупроводников второй группы экстремумы зоны проводимости и валентной зоны находятся при различных p , т. е. $\Delta p \neq 0$. В этом случае реализуются *непрямые переходы*. Вероятность таких переходов много меньше вероятности прямых переходов. Непрямые переходы осуществляются с испусканием или поглощением фононов (акустических квантов колебаний кристаллической решетки). Эти переходы определяют поглощение в той части спектра, которая расположена у длинноволновой границы собственного поглощения, обусловленного прямыми переходами. Например, пороговое значение энергии фотонов в германии, соответствующее красной границе для непрямых переходов, составляет 0,64 эВ, а для прямых оптических переходов оно существенно больше и составляет 0,81 эВ. На рис. 16.2 представлен спектр собственного поглощения чистых по-

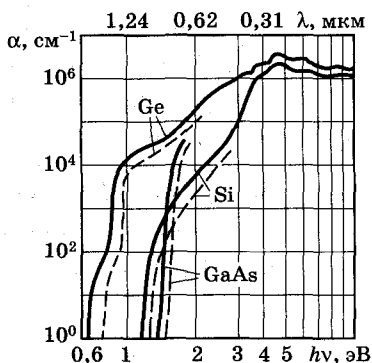


Рис. 16.2

тому спектра, которая расположена у длинноволновой границы собственного поглощения, обусловленного прямыми переходами. Например, пороговое значение энергии фотонов в германии, соответствующее красной границе для непрямых переходов, составляет 0,64 эВ, а для прямых оптических переходов оно существенно больше и составляет 0,81 эВ. На рис. 16.2 представлен спектр собственного поглощения чистых по-

лупроводников. Пунктирные кривые, как отмечено выше, соответствуют прямым переходам, а сплошные — непрямым. Для длинных волн, находящихся за пределами красной границы, поглощение крайне мало. Уменьшение α для коротких волн (коротковолновая граница фотоэффекта) обусловлено тем, что в этой области спектра коэффициент поглощения α очень велик ($> 10^5 \text{ см}^{-1}$) и излучение поглощается в тонком приповерхностном слое, где времена рекомбинации малы, т. е. фотоносители рекомбинируют раньше, чем они могут уйти из поверхностного слоя.

Примесное поглощение. Этот вид поглощения ($\alpha \sim 10 \dots 10^3 \text{ см}^{-1}$) связан с ионизацией или возбуждением примесных атомов под действием падающего света. В этом случае переходы электронов соответствуют позиции 2 на рис. 16.1. С увеличением концентрации донорных примесей в полупроводниках уровень Ферми поднимается вверх и в вырожденном полупроводнике n -типа располагается выше дна зоны проводимости. Поскольку состояния, расположенные ниже уровня Ферми, заполнены, поглощение, связанное с переходами в эти состояния, невозможно. Возможны только переходы на более высокие уровни, поэтому край спектра собственного поглощения в вырожденных полупроводниках смещается в сторону больших частот (энергий). Если же поглощение связано с переходом электронов с уровней акцепторов E_a в зону проводимости или из валентной зоны на уровни доноров E_d (переходы a и c на рис. 16.1), а также на уровни ловушек E_l (переход e), то это приводит к смещению спектра по сравнению с границей собственного поглощения в сторону меньших энергий или больших длин волн. Однако вероятность этих переходов на несколько порядков меньше, чем собственных ($\alpha \sim 10^3 \text{ см}^{-1}$). В примесном полупроводнике поглощение света может происходить за счет перехода носителей между границей валентной зоны или зоны проводимости и соответственно примесными уровнями E_a и E_d в запрещенной зоне (переходы b и d на рис. 16.1, обусловленные термической ионизацией атомов примесей¹). Для таких переходов коэффициент поглощения еще меньше ($\alpha \sim 10 \dots 10^2 \text{ см}^{-1}$).

¹ В связи с этим приемники ИК диапазона, использующие примесное поглощение, охлаждаются до криогенных температур ($\sim 77 \text{ К}$).

Экситонное поглощение. Экситон — это квазичастица, представляющая собой связанную электронно-дырочную пару, возникающую при возбуждении электрона валентной зоны энергией, меньшей энергии ширины запрещенной зоны. Экситон обладает нулевым спином, поскольку спины электрона и дырки антипараллельны. При размерах экситона (расстоянии между электроном и дыркой), заметно превышающих период решетки, взаимодействие электрона и дырки представляет собой кулоновское притяжение между положительным и отрицательным зарядом, ослабленное средой. В полупроводниках с прямыми переходами образование экситона сопровождается появлением очень узкого пика, а при непрямых переходах появляются ступеньки на краю спектра собственного поглощения (пики наблюдаются при $\lambda \geq \lambda_{кр}$ и не показаны на рис. 16.2). На рис. 16.1 переходы f , g , h (позиция 3) иллюстрируют экситонное поглощение ($\alpha \sim 10 \dots 10^2 \text{ см}^{-1}$). Переход f определяет свободный экситон, а g , h — связанные экситоны. Связанные экситоны возникают при поглощении света полупроводником, при котором возникает такое возбуждение электронов и дырок, когда кулоновское взаимодействие между ними приводит к объединению их с образованием ионоподобных или молекулярноподобных комплексов. Переходы g , h на рис. 16.1 иллюстрируют случай, когда свободная дырка объединилась с нейтральным донором (обозначен знаком «плюс», переход g) или свободный электрон — с нейтральным акцептором (обозначен знаком «минус», переход h). В результате образовались экситонные ионы.

Поглощение свободными носителями заряда. При поглощении фотонов свободные электроны переходят с одного уровня на другой в пределах той же зоны (см. рис. 16.1, переходы 4). Коэффициент поглощения свободными носителями ($\alpha \sim 10 \dots 10^3 \text{ см}^{-1}$) пропорционален их концентрации, квадрату длины волны падающего света и обратно пропорционален эффективной массе носителей и времени их релаксации. Спектр поглощения свободными носителями практически непрерывный и смещен в длинноволновую часть оптического диапазона.

Решеточное поглощение. Многие полупроводниковые материалы состоят из атомов различного типа, которые можно рассматривать как электрические диполи. В таких полупроводниках, как Si, Ge и др., диполи индуцируются световой волной. Наибо-

лее сильное поглощение наблюдается, когда частота собственных колебаний диполей близка к частоте падающего оптического излучения. Такое поглощение связано с возбуждением колебаний кристаллической решетки и называется решеточным ($\alpha \sim 1 \dots 10 \text{ см}^{-1}$, см. рис. 16.1, переходы 5). Решетка кристалла поглощает свет только при определенных значениях энергии фотона, поэтому спектр решеточного поглощения характеризуется рядом пиков поглощения, лежащих в далекой инфракрасной области спектра. При взаимодействии света с решеткой поглощаются только такие фотоны, импульс которых равен квазиимпульсу фонона, который мал по сравнению с импульсом фотона; поэтому, как следует из закона сохранения импульса, испускается, как правило, несколько фононов, что и определяет сложную структуру спектра поглощения колебаниями решетки.

Таким образом, при собственном и примесном поглощениях образуется избыточная (неравновесная) концентрация свободных носителей за счет квантовых электронных переходов между энергетическими уровнями. Экситонное, фононно-решеточное и поглощение на свободных электронах вызывают в конечном счете лишь разогрев кристаллической решетки.

Возникновение свободных носителей заряда под действием излучения приводит к увеличению проводимости материала (**фоторезистивный эффект**) и изменению контактной разности потенциалов в электрических переходах, что сопровождается появлением фото-ЭДС (**фотовольтаический эффект**). Эти два эффекта и используются в основном при создании фотоприемников, хотя имеется еще ряд фотоэлектрических эффектов, связанных с возникновением свободных носителей.

Излучение света полупроводниками. Излучающие приборы преобразуют электрическую энергию в энергию оптического излучения. В основе принципа действия полупроводниковых излучающих приборов лежит явление электролюминесценции. **Электролюминесценция** — излучение света телами под действием электрического поля — является частным случаем люминесценции. **Люминесценция** — это явление излучения света, избыточное над равновесным тепловым излучением тела при данной температуре, с длительностью, значительно большей периода световых волн. Эти особенности позволяют выделить люминесценцию среди других явлений вторичного свечения, в частности таких, как отражение и рассеяние света, тормозное излучение заряженных частиц и индуцированное излучение.

В излучающих приборах используются два наиболее характерных вида электролюминесценции твердых тел: *предпробойная*, происходящая в порошковых и пленочных электролюминофорах, и *инжекционная*, вызванная рекомбинацией электронов и дырок в переходах, включенных в прямом направлении. В рабочем диапазоне длин волн интенсивность светового потока излучающих приборов должна многократно превышать интенсивность фонового излучения окружающей среды.

16.3. Светодиоды

Светодиод — полупроводниковый излучающий прибор с одним или несколькими электрическими переходами, преобразующий электрическую энергию в энергию некогерентного оптического (светового) излучения. Используются светодиоды в оптических линиях связи, модуляторах, индикаторных устройствах, в оптопарах и т. д.

В основе принципа действия полупроводниковых излучающих приборов лежит явление электролюминесценции, связанное с самопроизвольной излучательной рекомбинацией носителей заряда, инжектируемых через электронно-дырочный переход. Излучение обусловлено рекомбинацией неравновесных носителей и происходит в $p-n$ -переходе и прилегающих к нему областях.

Одно из главных требований, предъявляемых к индикаторным светодиодам, — спектр излучения должен попадать в видимый диапазон. Поскольку в светодиодах основную роль играет межзонная излучательная рекомбинация, необходимая ширина запрещенной зоны полупроводников определяется граничными частотами фотонов $\nu_{кр}$ и $\nu_{ф}$ видимого диапазона: $h\nu_{кр} \leq \Delta E_z \leq h\nu_{ф}$, т. е. $1,8 \text{ эВ} < \Delta E_z = (h\nu) < 3,2 \text{ эВ}$. Из-за относительно большой ширины запрещенной зоны используемого полупроводника ток рекомбинации через $p-n$ -переходы оказывается большим по сравнению с током инжекции, особенно при малых прямых напряжениях (см. п. 2.3), т. е. процесс рекомбинации в этом случае реализуется в основном в $p-n$ -переходе. В светодиодах возможно получить излучательную рекомбинацию и при подаче обратных смещений, достаточных для ударной ионизации атомов в $p-n$ -переходе (см. п. 2.3). Образовавшиеся в результате ионизации неравновесные носители рекомбинируют в $p-n$ -переходе с излучением света, однако такие приборы менее эффективны.

В качестве основных полупроводниковых материалов для светодиодов применяют арсенид галлия GaAs, фосфид галлия GaP, нитрид галлия GaN, карбид кремния SiC, трехкомпонентный твердый раствор фосфида и арсенида галлия $\text{GaAs}_{1-x}\text{P}_x$, где $x \in 0 \dots 1$, и ряд других двойных и многокомпонентных полупроводниковых соединений. Использование этих материалов позволяет создать светодиоды, работающие в ИК, видимой и УФ областях спектра. КПД рассматриваемых приборов в основном зависит от **внутреннего квантового выхода** $\eta_{\text{ф}}$, который равен отношению количества излученных фотонов к числу рекомбинирующих пар носителей. Вероятность излучательной рекомбинации, определяющая внутренний квантовый выход, непосредственно связана с видом переходов в используемом полупроводнике. Внутренний квантовый выход в полупроводниках с прямыми переходами во много раз больше, чем с непрямыми.

Светодиоды на основе фосфида галлия. На рис. 16.3 (кривая $x = 1,0$) представлена энергетическая диаграмма чистого фосфида галлия, где минимумы энергии дна зоны проводимости при значении импульса $\bar{p} = \bar{p}_1$ соответствуют прямым переходам (ширина запрещенной зоны $\Delta E_g = 2,8$ эВ), а при $\bar{p} = \bar{p}_2$ — непрямым переходам ($\Delta E_g = 2,26$ эВ). Следовательно, чистый фосфид галлия GaP относится к непрямозонным полупроводникам. Квантовый выход для таких переходов незначителен, однако он широко используется для изготовления светодиодов, так как обеспечивает излучение в видимой области спектра, что необходимо для применения в индикаторных устройствах.

Для увеличения эффективности излучательных процессов в непрямозонных полупроводниках внедрением примесей создают рекомбинационные центры — ловушки. Например, при введении кадмия Cd, цинка Zn, меди Cu, азота N образуются ловушки акцепторного типа, уровни энергии которых относительно потолка валентной зоны соответственно равны: ($E_{\text{в}} + 0,097$) эВ для Cd, ($E_{\text{в}} + 0,064$) эВ для

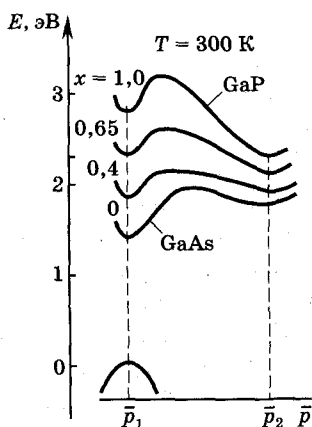


Рис. 16.3

Zn, ($E_{\text{в}} + 0,68$) эВ для Cu, ($E_{\text{п}} - 0,008$) эВ для N. Внедрение кислорода O и серы S приводит к появлению донорных уровней, которые располагаются ниже дна зоны проводимости: ($E_{\text{п}} - 0,896$) эВ для кислорода, ($E_{\text{п}} - 0,104$) эВ для серы. Излучательные переходы между донорными и акцепторными ловушками позволяют получить генерацию света на различных длинах воли (рис. 16.4, 16.5). При внедрении цинка, кадмия и кислорода реализуется красное излучение; кадмия, серы и азота — зеленое (см. рис. 16.4). Физические явления, происходящие в фосфиде галлия при наличии примесей, можно проиллюстрировать на примере легирования азотом. Азот замещает атомы фосфора в узлах кристаллической решетки. Азот и фосфор являются элементами одной группы периодической системы, имеют одинаковую внешнюю, но различную внутреннюю электронную структуру. Различие в строении приводит к возникновению энергетического уровня ловушки вблизи зоны проводимости. Инжектированные в *p*-область светодиода электроны попадают сначала на уровни ловушек, которые затем захватывают дырку из валентной зоны. В результате фотоны

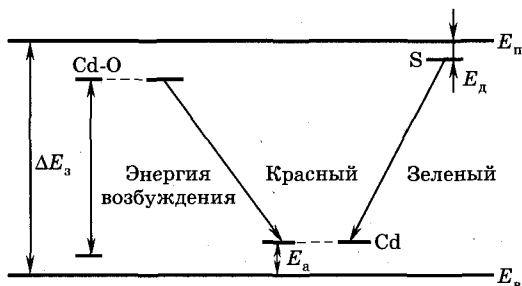


Рис. 16.4

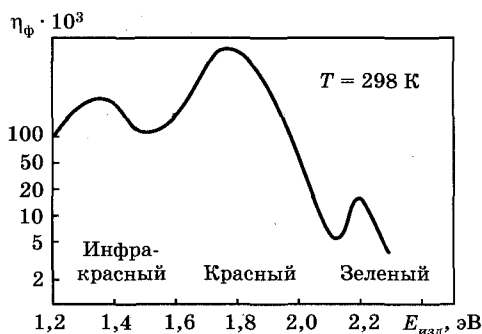


Рис. 16.5

рождаются с энергией, примерно равной разности между шириной запрещенной зоны и энергией связи атома ловушки.

Зависимость внутреннего квантового выхода $\eta_{\text{кр}}$ (в относительных единицах) от энергии излучаемых фотонов представлена на рис. 16.5. Полный КПД светодиода, излучающего зеленый свет, приблизительно равен 0,1%, а излучающего красный ~ 3%. Хотя КПД светодиодов с зеленым свечением мал, они применяются в индикаторной технике, поскольку чувствительность глаза к зеленому свету в 30 раз выше, чем к красному.

Светодиоды на основе трехкомпонентных твердых растворов. В светодиодах наиболее часто используются трехкомпонентные твердые растворы фосфида арсенида галлия $\text{GaAs}_{1-x}\text{P}_x$ (энергетические диаграммы соответствуют кривым при $0 < x < 1$, например $x = 0,4$ и $0,65$ на рис. 16.3). При увеличении фосфора в решетке арсенида галлия изменяется энергетическая диаграмма полупроводника (см. рис. 16.3). При $x = 0$ энергетическая диаграмма соответствует чистому арсениду галлия, а при $x = 1,0$ — чистому фосфиду галлия. При возрастании x от 0 до 0,45 ширина запрещенной зоны соединения ΔE_g увеличивается с 1,42 эВ до 1,98 эВ (см. рис. 16.3 и рис. 16.6). В светодиодах на основе таких материалов преобладают прямые переходы (кривая 1 на рис. 16.6). Дальнейшее увеличение содержания фосфора (увеличение x) приводит к непрямым переходам (кривая 2 на рис. 16.6), что вызывает уменьшение вероятности межзонной излучательной рекомбинации и, соответственно, внутреннего квантового выхода (кривая 1 на рис. 16.7).

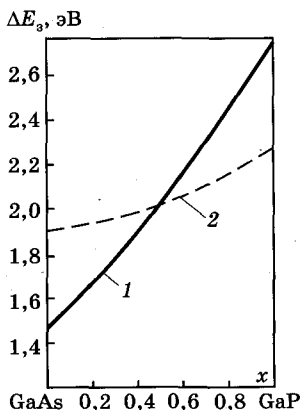


Рис. 16.6

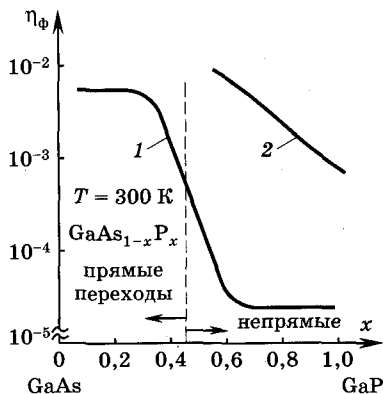


Рис. 16.7

Для увеличения эффективности, излучательной рекомбинации в фосфид арсенида галлия, как и в фосфид галлия, вводят примеси. На рис. 16.7 (кривая 2) проиллюстрировано влияние азота на величину квантового выхода $\eta_{\text{ф}}$.

Параметры светодиодов. К основным параметрам светодиодов относятся: яркость и мощность излучения, рабочее постоянное прямое напряжение, наибольшее постоянное или импульсное обратное напряжения, время нарастания и спада импульса излучения, длина волны излучаемого света (или его цвет), наибольший прямой постоянный или импульсный ток, КПД, долговечность и ряд других. Смысл большинства параметров понятен из их названия. Остановимся на специфических параметрах светодиодов.

Яркость B (кд/м²) характеризует свечение светодиода в избранном направлении. Для светодиодов яркость составляет несколько сот кд/м².

Минимальное прямое рабочее, или пороговое, напряжение $U_{\text{пор}}$ светодиода определяется энергией излучаемых фотонов; например, для зеленого света энергия фотона $\sim 2,2$ эВ, а пороговое напряжение $\sim 2,4$ эВ.

Максимальное рабочее напряжение ограничивается допустимой мощностью рассеяния светодиода. Оно в основном зависит от контактной разности потенциалов $p-n$ -перехода и сопротивления базы. Указанные напряжения определяют и соответствующие токи светодиода.

Постоянные времени нарастания и спада импульса излучения при импульсном возбуждении светодиодов характеризуют их инерционные свойства. Эти параметры измеряются между значениями яркости, составляющими 0,1 и 0,9 максимальной величины. Инерционность светодиодов определяется временем перезаряда емкости прибора. Для светодиодов значения постоянных времени составляют доли микросекунд.

КПД зависит от внутреннего квантового выхода и конструкции светодиодов. Потери энергии связаны с поглощением света в полупроводнике, контактах и элементах конструкции прибора.

Основные характеристики светодиодов. К ним относятся яркостная, спектральная и вольт-амперная (ВАХ) характеристики. **Яркостная характеристика** — это зависимость яркости B от тока через $p-n$ -переход, **спектральная характеристика** — зависимость интенсивности светового потока (или яркости, мощности, силы

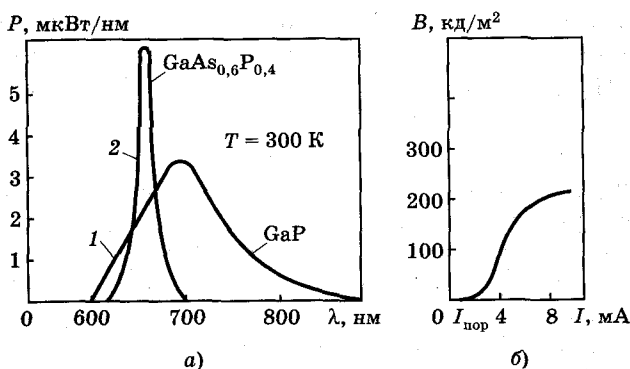


Рис. 16.8

света, энергии) от длины волны; ВАХ, как и в обычных диодах, есть зависимость $I = f(U)$.

На рис. 16.8, а представлены спектральные характеристики, дающие зависимость относительной мощности от длины волны излучения для светодиода из фосфида галлия (кривая 1) и фосфида арсенида галлия (кривая 2).

Вид яркостной характеристики зависит от структуры $p-n$ -перехода и области, в которой происходит преимущественная рекомбинация носителей заряда. При малых токах I и, соответственно, малых напряжениях излучение отсутствует. Излучение возникает при напряжениях, соответствующих энергии излучаемого фотона, приблизительно равной ширине запрещенной зоны, т. е. при $U = U_{\text{пор}}$ (напряжению $U_{\text{пор}}$ соответствует ток $I_{\text{пор}}$ на рис. 16.8, б). Рост напряжения (тока) увеличивает число рекомбинирующих носителей, следовательно, яркость возрастает. При больших токах начинает сильно проявляться безызлучательная рекомбинация из-за заполнения ловушек, в результате уменьшается квантовый выход и наклон характеристики к оси абсцисс становится меньше. Различия прямых ветвей ВАХ различных светодиодов объясняются разной шириной запрещенной зоны исходных материалов. Пробивное напряжение светодиодов при их обратном включении относительно невелико из-за малой толщины перехода.

Конструктивное исполнение светодиодов сильно влияет на величину внешнего квантового выхода, а следовательно, и на КПД прибора. Из-за высокого коэффициента преломления исходного материала светодиодов большая часть света испытывает полное

внутреннее отражение на границе раздела полупроводник—воздух. В результате из-за многократных переотражений от границ (рис. 16.9) происходит значительное поглощение света в полупроводнике, и только малая часть энергии излучения выходит из светодиодов простейшей плоской конструкции.

Светодиоды на основе фосфида арсенида галлия (см. рис. 16.9) получают наращиванием эпитаксиального слоя 2 на подложку из арсенида галлия 1. Излучаемый в области p — n -перехода 3 свет падает на подложку 1 и частично поглощается, что приводит к дополнительным потерям энергии. При более прогрессивной технологии эпитаксиальный слой 2 выращивают на прозрачной подложке из фосфида галлия с отражающим нижним покрытием 4. Это увеличивает выход полезного излучения. Внешний квантовый выход можно также увеличить за счет применения более сложных конструкций светодиодов. На рис. 16.10 показано устройство одного из типов таких светодиодов (3 — n -база, выполненная в виде полусферического монокристалла полупроводника, 1 и 2 — металлические контакты, 4 — эмиттер). Для повышения КПД светодиодов применяют прозрачные полусферические покрытия из стекла и пластмасс с высоким показателем преломления, просветляющие (прозрачные для излучаемых волн) покрытия внешней поверхности прибора и т. д.

Излучения различного цвета в индикаторах или индикаторных матрицах (см. гл. 15) реализуют в светодиодах с несколькими переходами. Пример двойной диодной структуры, которая излучает красный или зеленый свет, либо тот и другой одновременно, показан на рис. 16.11, где 1 и 2 — контакты к p -областям диода, генерирующим соответственно красный и зеленый свет; 6 — подложка; 4 и 5 — p — n -переходы соответственно для красного и зеленого диодов; 3 — n -GaP общий контакт.

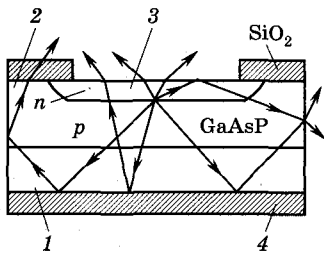


Рис. 16.9

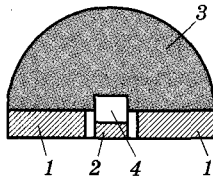


Рис. 16.10

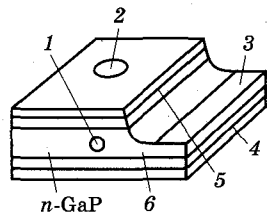


Рис. 16.11

На практике используются также приборы на основе чистого арсенида галлия и нитрида галлия, которые дают соответственно инфракрасное излучение ($\lambda \approx 900$ нм) и излучение голубого света, а также другие материалы.

Высокая надежность, большой срок службы (долговечность), малые рабочие напряжения и потребляемые мощности, небольшие масса и габариты светодиодов обусловили их широкое применение в устройствах самого различного назначения.

16.4. Полупроводниковые фотоприемники

Общая характеристика фотоприемников. По принципу действия все **фотоприемники** можно разделить на **тепловые** и **фотоэлектрические**; иногда среди последних выделяют еще фотоэлектронные приборы. В **тепловых фотоприемниках** используют изменение сопротивления чувствительного элемента в зависимости от его нагрева под действием поглощенного им излучения. В фотоэлектрических (фотоэлектронных) вакуумных приборах используют внешний фотоэффект, а в полупроводниковых фотоприемниках — внутренний.

Наибольшее применение находят полупроводниковые фотоэлектрические приемники, которые и будут основным предметом рассмотрения в этой главе. В общем случае во всех полупроводниковых фотоприемниках (**фотодетекторах**) происходят следующие процессы: генерация свободных носителей за счет воздействия излучения на полупроводник; перенос носителей и, возможно, их умножение в зависимости от типа прибора; взаимодействие с внешней цепью. В результате этих процессов формируется выходной сигнал, параметры которого зависят от падающего оптического излучения.

Можно выделить две основные группы оптоэлектронных полупроводниковых фотоприемников, предназначенных для приема, обработки и хранения информации: **дискретные фотоприемники** с малой апертурой — для приема оптических импульсов (сигналов); **многоэлементные фотоприемники** — для восприятия световых образов (изображений).

Среди фотоприемников первой группы лучшими с точки зрения высоких быстродействия и чувствительности в рабочем диапазоне длин волн и низкого уровня шумов являются фотодиоды (особенно с $p-i-n$ -структурой и лавинные), а среди второй — фоточувствительные приборы с зарядовой связью.

Фотоприемники являются крайне важной составной частью оптоэлектронных информационных систем. Потребности оптоэлектроники стимулировали создание фотоприемников, обладающих высокими быстродействием и чувствительностью, работающих в самых различных диапазонах длин волн, имеющих хорошие эксплуатационные характеристики и т. д. Перспективы развития оптоэлектроники в значительной мере определяются уровнем достижений в разработке фотоприемных устройств.

16.4.1. Фоторезисторы.

Фоторезистор представляет собой тонкую пластинку или пленку полупроводника 1 с омическими контактами 2 на двух противоположных концах (рис. 16.12), к которым подключается электрический источник питания с напряжением U . Полупроводник обычно наносится на стеклянную подложку 3. Наиболее распространены фоторезисторы на основе CdS и CdSe, работающие в длинноволновой части видимой области спектра. Фоторезисторы, изготовленные на основе PbS и PbSe, имеют наибольшую чувствительность в ИК диапазоне ($\lambda \sim 3 \dots 5$ мм). В ИК диапазоне работают также фоторезисторы на основе Ge, Si, GaAs, InSb и т. д.

Падающее на поверхность фоторезистора излучение генерирует в нем свободные носители за счет собственного или примесного поглощений (см. п. 16.2). В отсутствие светового потока проводимость фоторезистора называется темновой и описывается формулой (см. п. 1.3) $\sigma = q(\mu_n n + \mu_p p)$. Под действием оптического излучения за счет роста числа неравновесных свободных носителей происходит увеличение проводимости (и, следовательно, тока):

$$\sigma_{\phi} = q[\mu_n(n + \Delta n) + \mu_p(p + \Delta p)] = \sigma + \Delta\sigma,$$

где Δn , Δp — соответственно концентрации неравновесных электронов и дырок, $\Delta\sigma$ — фотопроводимость.

Для полупроводника с электронной проводимостью (в случае примесной (донорной) фотопроводимости или при условии, что дырки сразу же после рекомбинации захватываются центрами рекомбинации) фототок I_{ϕ} имеет только электронную составляющую и определяется соотношением

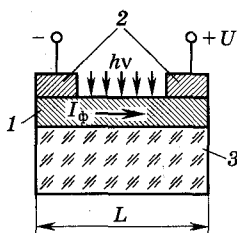


Рис. 16.12

$$I_{\phi} = qn_{\phi}v_{др}S = qn_{\phi}LS/t_{np} = qn_{\phi}V/t_{np}. \quad (16.1)$$

Здесь $n_{\phi} = \Delta n$ — концентрация фотоэлектронов, $v_{др} = L/t_{пр}$ — их скорость дрейфа, определяемая временем пролета $t_{пр}$ расстояния L между контактами фоторезистора, S — площадь поперечного сечения светочувствительного слоя (перпендикулярного направлению вектора плотности тока), $V = LS$ — объем полупроводника (см. рис. 16.12).

В стационарных условиях облучения и протекания постоянного фототока скорость генерации носителей в единице объема $G_{ген} = N_{\phi}A/V$ равна скорости их рекомбинации $G_{рек} = n_{\phi}/\tau$:

$$N_{\phi}A/V = n_{\phi}/\tau, \text{ или } N_{\phi}A = Vn_{\phi}/\tau, \quad (16.2)$$

где время жизни τ является количественной характеристикой интенсивности рекомбинации, а $N_{\phi}A$ — полное число фотонов, падающих на светочувствительную поверхность (площадью A) в единицу времени. (Предполагается, что толщина прибора больше глубины проникновения света.)

Очень важный параметр фоторезисторов — коэффициент внутреннего усиления фототока K_R , определяемый отношением числа фотоносителей, проходящих через прибор в единицу времени (фототок I_{ϕ}), к полному числу квантов оптического излучения, падающих на светочувствительную площадку в единицу времени (исходный ток $I_{\phiи}$), т. е.

$$K_R = I_{\phi}/I_{\phiи}. \quad (16.3)$$

Выражение для фототока $I_{\phiи} = qn_{\phi}V/\tau$ получается умножением формулы (16.2) на заряд носителя q .

После подстановки в формулу (16.3) этого выражения и соотношения (16.1) для фототока I_{ϕ} получим выражение для коэффициента внутреннего усиления в виде

$$K_R = \tau/t_{пр}, \quad (16.4)$$

где $(t_{пр})^{-1} = v_{обп}/L$, $t_{пр}$ — время пролета носителей между контактами фоторезистора (см. рис. 16.12).

Подчеркнем еще раз, что формула (16.4) получена для полупроводника с электронной проводимостью. Кроме того, считается, что генерация свободных носителей происходит только за счет облучения, т. е. не учитывается термогенерация носителей. В приборах с большим временем жизни носителей τ и малым расстоянием L между контактами коэффициент усиления может быть значительным, особенно в высокоомных полупроводниках ($K_R \sim 10^6 \dots 10^8$).

В полупроводниках, обладающих повышенной фотопроводимостью, K_R имеет меньшие значения из-за падения подвижности и насыщения дрейфовой скорости (см. п. 1.3).

Постоянная времени релаксации $\tau_{\text{рел}}$ фотопроводимости после прекращения оптического возбуждения пропорциональна времени жизни электронов τ . Время же фотосоответа (фотосоответки) определяется временем пролета свободных носителей между контактами прибора, т. е. величиной $t_{\text{пр}}$. Поскольку для фоторезисторов характерны большие расстояния между контактами и слабые электрические поля, их время фотосоответа обычно больше, чем у фотодиодов. Помимо коэффициента внутреннего усиления и постоянной времени релаксации, к основным параметрам фоторезистора (как и большинства фотоприемников) относятся:

монохроматическая чувствительность $S_{\text{ф}}(\lambda)$, которая определяется отношением фототока $I_{\text{ф}}$ к полной мощности монохроматического излучения $P_{\text{изл}}$ с длиной волны λ , падающей на чувствительную площадку фоторезистора, т. е.

$$S_{\text{ф}}(\lambda) = I_{\text{ф}}/P_{\text{изл}}(\lambda), \text{ А/Вт}; \quad (16.5)$$

интегральная чувствительность

$$S_{\text{ф. инт}} = \int S_{\text{ф}\lambda} d\lambda, \text{ А/Вт}, \quad (16.6)$$

где $S_{\text{ф}\lambda}$ — спектральная плотность чувствительности (измеряемая в А/(Вт · мкм), если длина волны λ выражена в мкм);

время нарастания (спада) $t_{\text{нр(сп)}}$ фототока, которое обычно определяется между уровнями 0,1 и 0,9 амплитуды импульса фототока при воздействии на фоторезистор идеально прямоугольного импульса излучения; при экспоненциальном нарастании и спаде фототока время нарастания (спада) связано со временем релаксации соотношением $t_{\text{нр(сп)}} = 2,2\tau_{\text{рел}}$;

границная частота $f_{\text{гр}}$, которая определяется частотой модуляции оптического излучения, соответствующей уменьшению чувствительности до уровня 0,707 от чувствительности для немодулированного излучения;

обнаружительная способность (см. [20])

$$D^* = S_{\text{ф}}(\lambda) \sqrt{A \Delta f / i_{\text{ш}}^2}, \quad (16.7)$$

где D^* измеряется в см · Вт⁻¹ · Гц^{1/2}; $\sqrt{i_{\text{ш}}^2}$ — среднеквадратическое значение шумового тока (см. гл. 22); Δf — рабочая полоса частот

фотоприемного устройства. Величина D^* позволяет сравнивать предельные возможности различных по устройству и условиям применения фотоприемников. Обнаружительная способность равна величине, обратной пороговой мощности фотоприемника, т. е. той мощности, которая еще различима на фоне шума для светочувствительной площадки в 1 см^2 и работе устройства в полосе частот 1 Гц.

Специфическими параметрами фоторезисторов являются сопротивления в темновом R_T , и засвеченном $R_{св}$ состояниях, их отношение $R_T/R_{св}$ и постоянная времени релаксации $\tau_{\text{рел}}$ фотопроводимости.

Для CdS- и CdSe-фоторезисторов типичные значения параметров: $R_T/R_{св} = 10^5 \dots 10^6$, $R_{св} = 10^2 \dots 10^8 \text{ Ом}$, $R_T = 10^7 \dots 10^{14} \text{ Ом}$. Инерционность переключения характеризуется временами около 100 мс. В миниатюрных планарных арсенид-галлиевых фоторезисторах удалось достигнуть $\tau_{\text{рел}} \sim 10^{-12} \text{ с}$. Однако такое быстрое действие реализуется лишь при возбуждении мощными лазерными импульсами; при малоинтенсивной засветке $\tau_{\text{рел}} \sim 10^{-9} \text{ с}$.

На рис. 16.13 представлены зависимости одного из важнейших параметров фотоприемников — обнаружительной способности — от длины волны. Отметим, что для приема излучения в ИК диапазоне с $\lambda > 2 \text{ мкм}$ фоторезисторы охлаждаются до температуры жидкого азота 77 К и паров жидкого гелия 4,2 К. При таких температурах уменьшаются тепловые эффекты, вызывающие термическую ионизацию и опустошение энергетических уровней, увеличиваются усиление и эффективность приема излучения. Фоторезисторы на CdS обладают наибольшей чувствительностью (см. рис. 16.13) на длинах волн $\lambda \sim 0,5 \text{ мкм}$; в диапазоне $\lambda \sim 10 \text{ мкм}$ используются фоторезисторы на основе HgCdTe. Для приема излучения в диапазоне длин волн $\lambda \sim 100 \dots 400 \text{ мкм}$ эффективно работают GaAs-приборы — особенно при приеме сигналов большой интенсивности. Фоторезисторы широко используются для детектирования в ИК области спектра при длинах волн больше нескольких микрометров. Для приема слабых сигналов на более коротких волнах в качестве высокочастот-

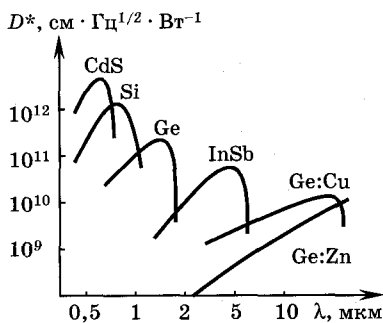


Рис. 16.13

ных оптических демодуляторов использование фоторезисторов ограничено. В этих случаях целесообразно применение фотодиодов.

16.4.2. Фотодиоды.

В семейство фотодиодов входят приборы с p — n -переходом, p — i — n -диоды, лавинные диоды, диоды с контактом металл—полупроводник и с гетеропереходом. Лучшими фотоприемниками являются кремниевые диоды с p — i — n -структурой и лавинные.

Структура любого фотодиода содержит монокристалл полупроводника, имеющего один или несколько электрических переходов, где присутствует обедненная область с сильным электрическим полем, в которой осуществляется разделение электронно-дырочных пар, рожденных оптическим излучением.

Конструкция диодов выполняется таким образом, чтобы их активная область была способна эффективно воспринимать оптическое излучение. Для этого в корпусе диода имеется прозрачное окно, за которым помещается светочувствительная область полупроводникового кристалла.

Фотодиоды обычно включаются в обратном направлении, при этом напряжение смещения не настолько велико, чтобы вызвать лавинный пробой. Исключение составляют лавинные фотодиоды, в которых внутреннее усиление реализуется именно за счет ударной ионизации при управляемом лавинном пробое. Большое обратное смещение позволяет уменьшить время пролета носителей через обедненную область, а также снизить емкость перехода и, соответственно, прибора, что в конечном счете улучшает высокочастотные и импульсные характеристики фотодиодов.

Наиболее важными параметрами фотодиодов как фотоприемников являются: квантовая эффективность, монохроматическая чувствительность, время фотоответа (фотоотклика) и шум, определяющие чувствительность приемников.

Квантовая эффективность η представляет собой отношение числа генерированных электронно-дырочных пар к общему числу падающих фотонов. Иначе, η — это количество фотоносителей, рожденных каждым фотоном, падающим на светочувствительный слой фотоприемника

$$\eta = (I_{\phi}/q)/(P_{\phi}/h\nu), \quad (16.8)$$

где I_{ϕ} — фототок, генерированный в полупроводнике за счет поглощения падающего оптического излучения, мощность которого P_{ϕ} , энергия фотона $h\nu$.

Одним из основных факторов, определяющих квантовую эффективность η , является коэффициент поглощения. На рис. 16.2 приведены зависимости коэффициентов поглощения Ge, Si и GaAs от длины волны. Для Ge, Si и полупроводниковых соединений типа $A^{III}B^V$ при увеличении температуры кривые сдвигаются в область больших длин волн. В УФ и видимой областях спектра хорошую квантовую эффективность имеют фотодиоды с контактом металл—полупроводник. Кремниевые диоды обладают высокой квантовой эффективностью в области длин волн $0,8 \dots 0,9$ мкм, германиевые — на $\lambda \sim 1 \dots 1,6$ мкм.

Инерционность фотодиодов (время фотоотклика) определяется временем диффузии носителей в обедненной области, временем дрейфа через обедненную область и емкостью обедненной области. Для уменьшения времени диффузии электрический переход формируют вблизи поверхности. Размер обедненной области перехода должен быть порядка $1/\alpha$, чтобы основная доля фотонов поглощалась в ней. Однако обедненная область перехода не должна быть и слишком широкой, иначе время пролета носителей через нее будет велико, а это ухудшает время фотоответа. В узких переходах велика емкость и инерционность растет за счет большой постоянной времени RC (R — сопротивление нагрузки). Оптимальная ширина (толщина) обедненного слоя реализуется, если время переноса носителей по порядку величины соответствует половине периода модуляции излучения, что при частоте модуляции $f \sim 10$ ГГц составляет величину ~ 5 мкм.

$p-i-n$ -фотодиоды. На рис. 16.14 представлены разновидности устройства $p-i-n$ -фотодиодов с освещением перпендикулярно переходу (рис. 16.14, а) и освещением параллельно переходу (рис. 16.14, б). Здесь 1 — просветляющее покрытие, позволяющее увеличивать квантовую эффективность; 2 — металлические контакты; 3 — слой окиси кремния SiO_2 ; 4 — отражающее по-

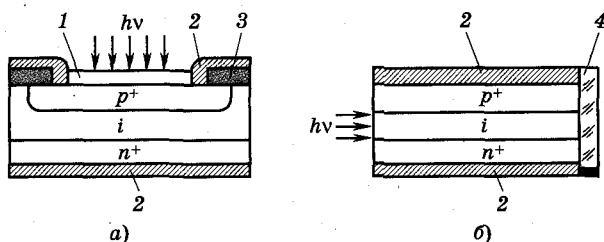


Рис. 16.14

крытие). На практике в $p-i-n$ -фотодиодах вместо области с собственной проводимостью (i -область) используется высокоомный слой n -типа (v -область) или высокоомный слой p -типа (π -область) (см. п. 5.4). К электродам $p-i-n$ -фотодиодов прикладывается обратное смещение.

Рассмотрим принцип действия фотодиода. При этом сначала будем полагать, что внешнее напряжение отсутствует и цепь разомкнута. Оптическое излучение, проникая с малыми потерями через прозрачное покрытие и p^+ -область, достигает i -базы, в которой в основном и поглощается (рис. 16.15). При этом происходит ионизация или собственных, или примесных атомов. Интенсивность проявления примесного поглощения существенно меньше собственного из-за малого количества примесных атомов по отношению к атомам собственного полупроводника, поэтому в фотодиодах в основном используется собственная фотопроводимость.

Образующиеся в i -базе электронно-дырочные пары за счет диффузии или дрейфа, если в базе есть электрическое поле, а чаще всего за счет того и другого начинают перемещаться. Дырки двигаются в направлении p^+ -области, а электроны — n^+ -области (рис. 16.15, б). За счет поля $p^+—i$ -перехода осуществляется разделение фотоносителей. Поскольку мы предположили, что цепь разомкнута, то в p^+ -области происходит накопление дырок, а в n^+ -области — электронов. В результате потенциальный барьер на переходе $p^+—i$ понижается, т. е. на контактах прибора появляется некоторая ЭДС. Это явление называется **фотовольтаическим эффектом**. (Оно будет более подробно рассмотрено при описании процессов в солнечных преобразователях, п. 16.6.)

Если теперь замкнуть цепь, приложив к диоду обратное смещение, то объемный заряд может занять всю i -область (см. энергетическую диаграмму на рис. 16.15, б) и во внешней цепи потечет ток I_{ϕ} , который будет складываться с обратным током I_T (темновым током) неосвещенного $p-i-n$ -диода. Таким образом, $I_{\phi} + I_T \approx I_{\phi}$. Увеличение мощности оптического излучения будет приводить и к возрастанию числа генерированных электронно-дырочных пар, а следовательно, и к увеличению фототока I_{ϕ} , что хорошо видно из ВАХ $p-i-n$ -фотодиода (рис. 16.16).

Большинство основных параметров $p-i-n$ -фотодиода было перечислено при рассмотрении фоторезисторов, а именно: монохроматическая чувствительность, время нарастания (спада), граничная частота, обнаружительная способность, темновой ток,

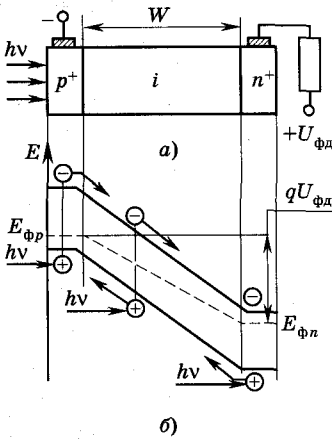


Рис. 16.15

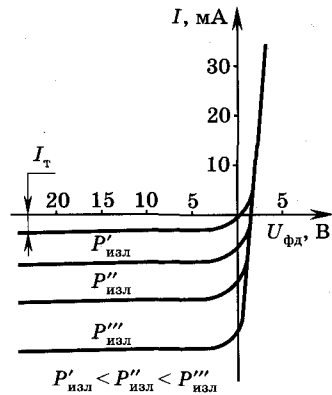


Рис. 16.16

площадь фоточувствительной площадки. Дополнительно у фотодиодов рассматриваются следующие параметры: максимально допустимое обратное напряжение $U_{\text{обр. макс.}}$, характеризующее предельные возможности фотодиода при включении в электрическую цепь; емкость фотодиода $C_{\text{фд}}$, которая равна сумме емкости корпуса и активной структуры. Во многих случаях именно $C_{\text{фд}}$ определяет инерционность диода.

На рис. 16.17 представлена спектральная зависимость чувствительности кремниевых $p-i-n$ -фотодиодов для различных значений толщины базы. Общий вид характеристики аналогичен спектральной характеристике материала, из которого выполнен прибор. Уменьшение чувствительности S_{ϕ} для более тонкой базы связано с уменьшением поглощения фотонов в базе и изменением условий собирания генерированных носителей. Уменьшение S_{ϕ} в области коротких волн обусловлено тем, что коэффициент поглощения α очень велик и излучение поглощается в поверхностном слое (область p^+), не проникая глубоко в базу. Времена рекомбинации в этом поверхностном слое малы. Вследствие этого фотоносители рекомбинируют раньше, чем уходят из поверхностного слоя.

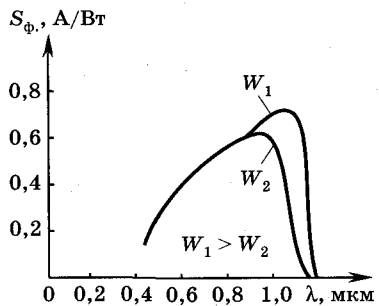


Рис. 16.17

Быстродействие рассматриваемых фотодиодов ограничивается временем пролета носителей через обедненный i -слой. Для увеличения быстродействия необходимо толщину i -слоя делать как можно меньше. Однако в этом случае возможно существенное снижение и квантовой эффективности. Как показывают расчеты и эксперимент, разумный компромисс между быстродействием и квантовой эффективностью достигается в том случае, если размер области поглощения изменяется от $1/\alpha$ до $2/\alpha$. Зависимость квантовой эффективности η кремниевого $p-i-n$ -фотодиода от ширины обедненной области для различных длин волн приведена на рис. 16.18. Изменение ширины обедненной области при увеличении обратного напряжения для различного удельного сопротивления i -области демонстрирует рис. 16.19. Если вся база диода представляет собой область пространственного заряда, то время пролета определяется только дрейфом носителей и в большинстве случаев очень мало. Так, например, при $U_{обр} = 10$ В и удельном сопротивлении базы $\rho = 3 \cdot 10^3$ Ом·см ширина обедненной области ~ 100 мкм (см. рис. 16.19) и при $v_{обр} = v_{др.нас} \approx 10^7$ см/с время пролета $t_{пр}$ через область базы будет порядка 1 нс. Напряжение, при котором наступает условие обеднения всей базовой области, называется напряжением отсечки. Приняв минимальную реальную толщину базы $W \approx 1$ мкм, для предельного быстродействия $p-i-n$ -фотодиодов получим $t_{пр} \approx 10 \dots 20$ пс, при этом значения граничной частоты $f_{гр} \approx 0,35/t_{пр} \approx 18 \dots 35$ ГГц. Для реальных диодов типичные значения $f_{гр}$ лежат в пределах 200...500 МГц. Параметры I_T , $C_{фд}$ и $U_{обр. макс}$ $p-i-n$ -диодов и обычных фотодиодов практически совпадают. Шумы $p-i-n$ -фотодиодов малы (см. гл. 22), но их чувствительность невелика.

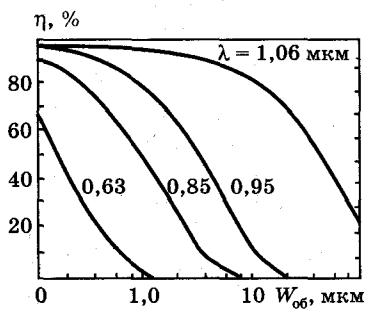


Рис. 16.18

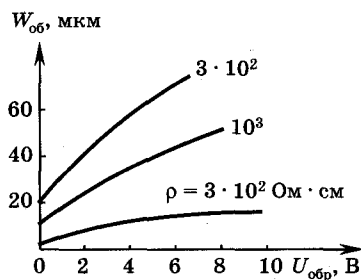


Рис. 16.19

Лавинные фотодиоды (ЛФД).

ЛФД работают при обратных смещениях, достаточных для размножения носителей. По сравнению с $p-i-n$ -фотодиодами они обладают внутренним усилением, что является преимуществом при приеме слабых сигналов, т. е. они имеют лучшую чувствительность, чем $p-i-n$ -фотодиоды, а это и обуславливает их широкое применение. ЛФД изготавливаются на основе Ge, Si, соединений группы $A^{III}B^V$ и других полупроводников. Определяющими параметрами при выборе материала фотодиода являются квантовая эффективность в заданном спектральном диапазоне, быстродействие и шумы. Германиевые ЛФД обеспечивают высокую квантовую эффективность в спектральном диапазоне 1...1,6 мкм, а кремниевые — особенно эффективны на длинах волн 0,6...1 мкм.

Существуют несколько конструктивных вариантов исполнения ЛФД. Для кремниевых ЛФД наиболее оптимальна $p-i-n$ -структура (рис. 16.20 и 16.21, а), распределение концентраций примесей для которой приведено на рис. 16.21, б. Эта структура является, по сути, совокупностью перекрывающихся между собой $p-i-n$ -фотодиода и лавинного $n-p$ -диода, образованного двумя верхними слоями структуры, изображенной на рис. 16.20; $p-i-n$ -диод образован n^+ -слоем, $p-i(\pi)$ -базой, которая выполняет роль i -слоя, и p^+ -подложкой. На рис. 16.20 обозначено: 1 — просветляющее покрытие; 2 — металлические контакты; 3 — окись кремния SiO_2 ; 4 — охранный кольцо; 5 — подложка.

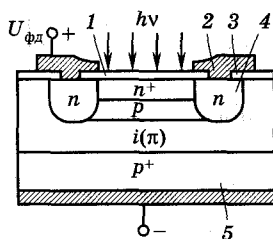


Рис. 16.20

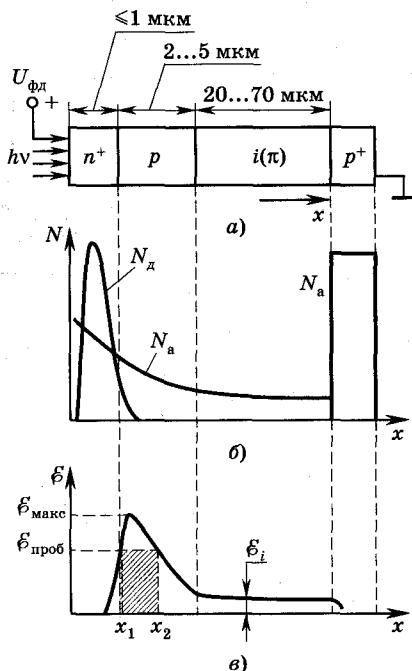


Рис. 16.21

Эта структура является, по сути, совокупностью перекрывающихся между собой $p-i-n$ -фотодиода и лавинного $n-p$ -диода, образованного двумя верхними слоями структуры, изображенной на рис. 16.20; $p-i-n$ -диод образован n^+ -слоем, $p-i(\pi)$ -базой, которая выполняет роль i -слоя, и p^+ -подложкой. На рис. 16.20 обозначено: 1 — просветляющее покрытие; 2 — металлические контакты; 3 — окись кремния SiO_2 ; 4 — охранный кольцо; 5 — подложка.

Применение охранного кольца в конструкции ЛФД обусловлено необходимостью устранения токов утечки и образования микроплазм по периферии перехода, вызываемых краевыми эффектами (более подробно эти явления рассмотрены ниже).

В ЛФД при номинальных режимах обедненная область занимает всю $p-i$ -базу. В высокоомной i -области напряженность электрического поля \mathcal{E}_i практически неизменна и существенно меньше, чем на границе n^+ - p -областей (рис. 16.21, в). В узкой p -области напряженность электрического поля максимальна ($\mathcal{E}_{\text{макс}} \sim (3...6)10^5$ В/см), и ее значения в некоторой области $x_1 \leq x \leq x_2$ (см. рис. 16.21, в) достаточны для возникновения и поддержания лавинного размножения. Область $x_1 \leq x \leq x_2$, где происходит лавинное размножение, на рис. 16.21, в заштрихована.

Механизм лавинного размножения при облучении диода светом выглядит следующим образом. При воздействии падающего излучения основная доля фотонов поглощается в i -области, что вызывает генерацию электронно-дырочных пар. Электроны под действием электрического поля $\mathcal{E}_i \approx 10^3...10^4$ В/см (см. рис. 16.21, в) перемещаются в направлении n^+ -области, достигнув области $x_1 < x < x_2$, где $\mathcal{E} > \mathcal{E}_{\text{проб}}$. Там они приобретают энергию, достаточную, чтобы при столкновении с атомами ионизовать их. В результате этих процессов будет происходить лавинное размножение фотоносителей.

В ЛФД необходимо обеспечить однородное размножение по всей fotocувствительной площади, т. е. должно быть исключено образование локализованных участков, в которых пробивное напряжение меньше, чем во всем переходе. Иначе будут возникать микроплазменные участки. Применение материалов с низкой плотностью дислокаций и выполнение активной области с размерами, необходимыми только для сбора светового пучка (диаметр от нескольких мкм до 100 мкм), а также использование охранного кольца позволяет уменьшить количество возможных микроплазм. Коэффициент лавинного усиления M , называемый также коэффициентом умножения, является основным параметром ЛФД. Он сложным образом зависит от приложенного напряжения. Для кремния эта зависимость представлена на рис. 16.22. При увеличении напряжения $U_{\text{фд}}$ на диоде от 0 до ~ 50 В обедненная область (область сосредоточения пространственного заряда) располагается в p -слое, подобно слою $x_1 < x < x_2$ на

рис. 16.21. Она полностью занимает p -слой при $U \sim 50 \dots 100$ В, т. е. координата x_2 будет на границе i -слоя.

На этом участке кривой лавинное размножение носителей происходит на границе $n-p$ -области (см. рис. 16.21). При изменении напряжения от ~ 50 до 300 В коэффициент размножения изменяется достаточно плавно. Здесь область пространственного заряда сначала занимает p -область, а потом и i -слой,

который по мере увеличения напряжения полностью заполняется объемным зарядом. Этот участок кривой на рис. 16.22 соответствует рабочим (номинальным) режимам ЛФД, и в рассматриваемом диапазоне напряжений наблюдается максимальный квантовый выход, т. е. все носители, генерированные в i -области, достигают внешних электродов. Резкое увеличение коэффициента M при больших напряжениях связано с лавинным размножением носителей в i -области. Стабильность параметров при этих напряжениях заметно ухудшается.

Коэффициент M имеет обычно неодинаковые значения в различных участках фоточувствительной поверхности; он максимален в центре и уменьшается в направлении границы. Температурная зависимость коэффициента M (см. рис. 16.22) объясняется так же, как и температурная зависимость пробивного напряжения (см. п. 2.5).

Отметим некоторые характерные свойства разновидностей ЛФД. Кремниевые лавинные фотодиоды со структурой металл—полупроводник эффективно работают в видимой и УФ областях спектра. ЛФД с контактом металл— n -Si особенно перспективны для применения в качестве высокочастотных детекторов ультрафиолетового излучения. Ультрафиолетовое излучение проходит через тонкие металлические электроды и поглощается в поверхностном слое кремния толщиной $\sim 10^{-6}$ см. В этом случае умножение носителей осуществляется в основном за счет электронов, что снижает уровень шума и повышает чувствительность.

ЛФД с гетеропереходами на базе таких соединений групп $A^{III}B^V$, как AlGaAs/GaAs, AlGaSb/GaSb, GaInAs/InP и GaInAs/InP, обладают рядом преимуществ по сравнению с приборами на основе германия и кремния. Эти соединения позволяют регули-

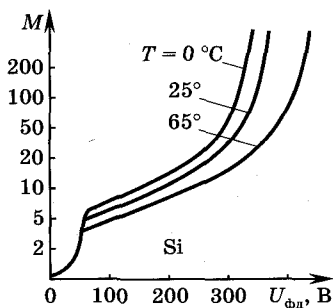


Рис. 16.22

ровать спектральные характеристики и чувствительность приборов, получать высокое быстродействие с высокой квантовой эффективностью за счет реализации прямозонных переходов. Кроме того, поверхностный слой гетероструктуры с более широкой запрещенной зоной формируется по технологии, которая обеспечивает хорошие высокочастотные характеристики и минимальные потери, связанные с рекомбинацией фотоносителей.

Лавинные фотодиоды характеризуются в основном той же системой параметров, что и другие фотоприемники. (Она была описана при рассмотрении фоторезисторов и $p-i-n$ -фотодиодов.) Различия, обусловленные спецификой работы и использования ЛФД, связаны с такими параметрами, как коэффициент усиления (умножения) M ; рабочее напряжение U_M , при котором достигается требуемое значение M ; произведение коэффициента усиления на полосу частот $Mf_{гр}$ (комбинированный параметр качества, позволяющий сопоставлять приборы с разными значениями M и $f_{гр}$). Как и у $p-i-n$ -фотодиода, у ЛФД используются такие параметры, как $f_{гр}$, I_T , $t_{пр}$, $C_{лфд}$, A и коэффициент шума.

Коэффициент усиления ЛФД был рассмотрен ранее, поэтому остановимся лишь на таком параметре, как рабочее напряжение. Диапазон возможных рабочих напряжений, типичных для $n-p-i-p$ -ЛФД, составляет 100...500 В, что несопоставимо больше по сравнению с $p-i-n$ -фотодиодом, который может работать без смещения.

Отметим ряд недостатков, присущих ЛФД-фотодиодам: сложность изготовления и высокая стоимость; большие рабочие напряжения и большая бесполезно расходуемая мощность; работа лишь в режиме усиления малых сигналов; жесткая стабилизация рабочего напряжения и температуры. Однако сочетание большого усиления и быстродействия и относительно малого уровня шумов обуславливает их широкое применение в оптоэлектронике.

16.4.3. Разновидности фотоприемников излучения.

Кроме рассмотренных фотоприемников (фоторезисторов, лавинных и $p-i-n$ -фотодиодов), широко используются такие фотоприемники, как фотодиоды с $p-n$ -переходом, с контактом металл—полупроводник, гетерофотодиоды, МДП-фотодиоды, фототранзисторы, фототиристоры и др.

Фотодиоды с $p-n$ -переходом ($p-n$ -фотодиоды). По своим свойствам, происходящим в них процессам, устройству, областям

применения, по используемым параметрам $p-n$ -фотодиоды во многом аналогичны $p-i-n$ -диодам. Устройство приборов этого типа понятно из рис. 16.23. В отличие от $p-i-n$ -фотодиода в относительно толстой n -базе $p-n$ -фотодиода доминируют диффузионные процессы, что приводит к большей инерционности и снижению fotocувствительности. В ИК диапазоне fotocувствительность $p-n$ -фотодиодов может быть на порядок хуже, чем у $p-i-n$ -приборов. Эти недостатки существенно ограничивают применение $p-n$ -фотодиодов в оптоэлектронике. К неоспоримым достоинствам этих приборов можно отнести простоту изготовления, повышенную однородность параметров изготавливаемых структур из-за использования низкоомной однородной полупроводниковой пластины, полную совместимость технологии изготовления кремниевых $p-n$ -фотодиодов с технологией микросхем. Последнее обстоятельство позволяет создавать интегральные фотоприемники, представляющие собой микросхему, где совместно с фотодиодом на том же кристалле изготавливаются высокочастотные транзисторы, операционные и ключевые усилители, пороговые схемы и т. д. Использование стандартной технологии обеспечивает низкую стоимость fotocувствительных микросхем, практически равную стоимости дискретных приборов, при высокой чувствительности, быстродействии и температурной стабильности. Такие схемы обладают очень широкими функциональными возможностями и перспективны в микроэлектронных оптических устройствах.

Фотодиоды с барьером Шоттки. Используются в качестве высокоэффективных фотодетекторов, обладающих малой инерционностью в видимой и УФ областях спектра. Основные физические процессы, протекающие при контакте металл—полупроводник, были рассмотрены в п. 2.6. Типичная структура фотодиода с барьером Шоттки изображена на рис. 16.24. Здесь 1 — металли-

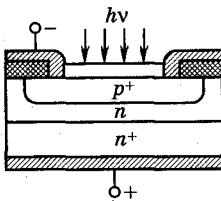


Рис. 16.23

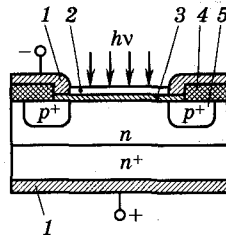


Рис. 16.24

ческие контакты; 2 — просветляющее покрытие из пленки ZnS, SiO и др.; 3 — тонкий полупрозрачный слой металла; 4 — окись кремния SiO₂; 5 — диффузионное охранное кольцо p^+ -типа.

По сравнению с $p-n$ - или $p-i-n$ -структурами металлическая пленка, обеспечивающая контакт с полупроводником n -типа, обладает значительно меньшим последовательным сопротивлением, чем в случае мелкого $n^+—n$ -перехода у перечисленных структур. Меньшее последовательное сопротивление дает и меньшие потери. Кроме того, паразитное поглощение коротковолнового (УФ) излучения в тонкой ($\sim 10^{-6}$ см) полупрозрачной пленке меньше, чем в более толстой p^+ -области. Все отмеченное, а также простота изготовления и широкий выбор возможных материалов (как металлов, так и полупроводников) несомненно являются достоинствами этих приборов. Из полупроводников в фотодиодах Шоттки используют Si, Ge, GaAs, GaAsP, InP, CdS, PbS, PbSe и др., а из металлов — Au, Ag, Pt, Cu, Mo, Ni, W, Cr, Sn, Zn, Cs и др. Приборы рассматриваемого типа хорошо совместимы технически и физически с интегральными структурами, что является существенным их достоинством.

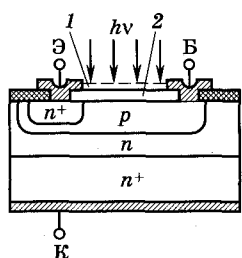
Для видимой и УФ областей спектра коэффициент поглощения в наиболее распространенных полупроводниках очень велик ($> 10^5$ см⁻¹), а эффективная глубина поглощения мала ($1/\alpha < 0,1$ мкм), поэтому соответствующим подбором металла и просветляющего покрытия добиваются, чтобы падающее излучение в основном поглощалось вблизи поверхности полупроводника. Пленка из золота толщиной $\sim 10^{-6}$ см пропускает более 95% падающего излучения, а при толщине $\sim 5 \cdot 10^{-6}$ см коэффициент пропускания снижается до 30%. Квантовая эффективность лучших приборов в видимом и УФ диапазонах колеблется в пределах 20...70%, время нарастания фотоотклика $\sim 10^{-10}$ с.

Фототранзисторы и фототиристоры. Биполярные и полевые транзисторы, а также тиристоры, рассмотренные соответственно в гл. 4—6, при соответствующей конструктивной модификации могут выполнять роль фотоприемников. Характерной особенностью этих приборов является наличие высокого коэффициента внутреннего усиления, что обеспечивает и высокую чувствительность. Однако эти фотоприемники отличаются по сравнению с фотодиодами большей инерционностью и конструктивной сложностью.

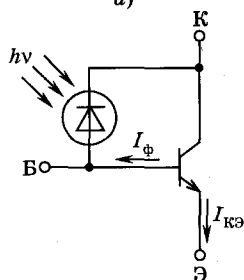
Устройство биполярного фототранзистора показано на рис. 16.25, а, а эквивалентная схема — на рис. 16.25, б. Основные области, электроды и их назначение — те же, что и в обычном транзисторе. Конструктивным отличием фототранзистора является наличие фотоприемного окна, состоящего из антиотражающего покрытия 1 и прозрачного защитного окисла 2. Через окно оптическое излучение проникает сначала в базовый слой (Б), а потом в коллекторную область (К). Переход база—коллектор выполняет роль фоточувствительного элемента и поэтому имеет большую площадь. На рис. 16.25, б он показан в виде фотодиода.

Рассмотрим основные процессы в биполярном фототранзисторе, который обычно включается в электрическую цепь по схеме с общим эмиттером (Э). Фототранзистор может работать в режиме как с плавающей базой p , когда ее электрическая цепь разорвана, так и при наличии базового тока подобно обычному биполярному транзистору. Специфическим является режим с плавающей базой, поэтому дальнейшее рассмотрение будет касаться именно его. При наличии напряжения между коллектором n и эмиттером n^+ и в отсутствие облучения в цепи коллектора будет протекать темновой ток. В этом случае эмиттерный переход смещен в прямом направлении, а коллекторный — в обратном. Напряжение коллектор—эмиттер распределено между этими двумя переходами, причем подавляющая часть падает на коллекторном переходе, а на эмиттерном падение напряжения незначительно. Поэтому и ток инжекции через этот открытый переход, и темновой ток I_T в цепи коллектора будут также малы.

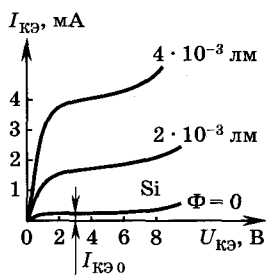
При освещении транзистора светом в его базовой и коллекторной областях генерируются электронно-дырочные пары, которые разделяются полем коллекторного перехода аналогично



а)



б)



в)

Рис. 16.25

тому, как это происходило в фотодиодах. Электроны движутся в область коллектора. Дырки, фотогенерированные в базе, и те, которые приходят в базу из коллектора, понижают потенциальный барьер на эмиттерном переходе, что вызывает дополнительную инжекцию электронов в базу. С увеличением интенсивности светового потока Φ растет число генерированных фотоносителей и, соответственно, увеличивается коллекторный ток, что хорошо видно из выходных ВАХ биполярного фототранзистора (рис. 16.25, в). Общий коллекторный ток

$$I_{КЭ} = I_{\Phi} + h_{21Э}I_{\Phi} + I_{КЭ0} = (1 + h_{21Э})I_{\Phi} + I_{КЭ0}, \quad (16.9)$$

где $I_{КЭ0}$ — ток в коллектор-эмиттерной цепи при $\Phi = 0$, $h_{21Э}$ — коэффициент передачи в схеме с ОЭ (см. п. 4.4).

Поскольку $h_{21Э} = \beta \gg 1$, эффективный квантовый выход фототранзистора в $(1 + h_{21Э})$ раз больше квантового выхода фотодиода (обычно на 1...2 порядка), роль которого в фототранзисторе выполняет переход база—коллектор. Если в цепи базы протекает ток от внешнего источника (на базу подано смещение), этот ток добавляется к току I_{Φ} и общий коллекторный ток возрастает. Биполярные фототранзисторы совмещаются с другими интегральными приборами. Это можно ярко проиллюстрировать на примере составного транзистора, устройство и схема которого представлены соответственно на рис. 16.26, а, б, где 1 — антиотражающее покрытие, 2 — защитный окисел, 3 — металлический контакт, 4 — SiO_2 . Транзистор VT_1 функционирует как обычный биполярный фототранзистор. Ток инжекции транзистора VT_1 является одновременно входным током транзистора VT_2 , что значительно повышает общий коэффициент

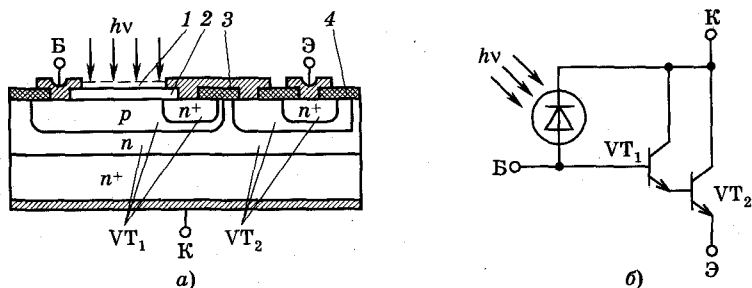


Рис. 16.26

усиления. В составном транзисторе коэффициент усиления достигает значений $10^3 \dots 10^4$. Большая емкость база—коллектор в биполярных фототранзисторах ограничивает их быстродействие, которое снижается при увеличении усиления за счет обратной связи. Типичные времена переключения: для фотодиодов $\sim 10^{-8}$ с, для биполярных фототранзисторов $5 \cdot 10^{-6}$ с, для составных фотоприемников $5 \cdot 10^{-5}$ с.

Фототиристоры. Устройство и принцип действия фототиристоров во многом аналогичны устройству и принципу действия обычных тиристоров и их разновидностей, описанных в гл. 5. Отличительная особенность устройства фототиристора — наличие полупрозрачного окна в области базы с управляющим электродом УЭ (рис. 16.27, а и б). Освещение фототиристора приводит к генерации электронно-дырочных пар в базе и в переходе Π_2 . Переход Π_2 разделяет электроны и дырки аналогично коллекторному переходу фототранзистора и переходу фотодиода. Избыточные носители усиливают механизм обратной связи тиристора, что приводит к переключению его из запертого состояния в открытое при меньших анодных напряжениях, аналогично тому, как это происходит при наличии управляющего тока в обычном тиристоре. ВАХ фототиристора полностью аналогичны характеристикам триодного фототиристора (см. рис. 5.1, е), где в качестве параметра вместо тока управляющего электрода выступает мощность светового потока или любой другой параметр, характеризующий интенсивность падающего излучения. Фототиристоры используются в основном в устройстве силовой автоматики и сильноточной электроники, так как позволяют переключать большие токи и напряжения при облучении их световыми сигналами малой интенсивности.

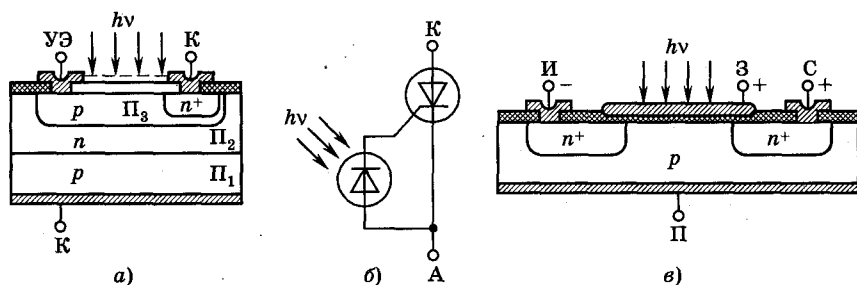


Рис. 16.27

Чувствительность фототиристоров не хуже, чем у составных биполярных фототранзисторов, а быстродействие фототиристоров ($10^{-5} \dots 3 \cdot 10^{-5}$ с) даже несколько лучше.

Полевые фототранзисторы. Структура кремниевого МДП-транзистора представлена на рис. 16.27, в, где И — исток; С — сток; П — подложка; З — затвор. Излучение через полупрозрачный электрод затвора проникает в область канала и подложки p -типа. Генерированные носители разделяются электрическим полем канала: электроны поступают в канал, увеличивая его проводимость, а дырки вытесняются в подложку. Увеличение проводимости канала под действием света приводит к изменению порогового напряжения на затворе и крутизны характеристики. Выходной сигнал, снимаемый с резистора, включенного в цепь стока, пропорционален фототоку в процессе облучения.

Если МДП-фототранзистор работает в ИК диапазоне, то обычно используется эффект не собственного, а примесного поглощения. В этом случае p -подложка легируется индием и бором. Бор дает мелкий акцепторный уровень, т. е. уровень, расположенный близко к границе валентной зоны, а индий образует глубокий акцепторный уровень, расположенный выше границы валентной зоны на 0,16 эВ. Если положительный потенциал на затворе (см. рис. 16.27, в) больше порогового, то происходит образование канала. Все примесные центры (уровни) захвачены дырками и нейтральны. Под действием ИК излучения с энергией фотонов, достаточной для ионизации атомов индия, дырки переходят в валентную зону и образуются отрицательные ионы акцепторов (в основном в поверхностном слое, т. е. недалеко от границы канала в обедненной области). Суммарная плотность отрицательного объемного заряда в обедненной области увеличивается, изменяется ширина перехода на границе канала и подложки, что, в свою очередь, модулирует проводимость канала и изменяет ток стока. Эти изменения пропорциональны величине светового потока и, соответственно, фототоку. При использовании вместо индия других примесных атомов можно изменять спектральную характеристику прибора, получая максимум чувствительности в требуемом спектральном диапазоне.

Достаточно распространен вариант полевого МДП-транзистора, состоящего из $p-n$ -перехода истоковой области и МДП-транзистора. Характерным конструктивным отличием прибора этой модификации от обычного МДП-транзистора является

наличие развитой площади истока, которая облучается светом, т. е. в этом случае переход исток—подложка выступает как $p-n$ -фотодиод, фототок которого усиливается МДП-транзистором. Кроме того, фото-ЭДС $p-n$ -перехода изменяет пороговый потенциал затвора и вся сток-затворная характеристика смещается, увеличивая ток стока. Такой интегральный приемник может работать в режиме накопления заряда. В этом режиме переход исток—подложка выступает в роли $p-n$ -фотодиода, и на емкости этого фотодиода, т. е. емкости перехода исток—подложка, накапливается заряд, а сам МДП-транзистор работает в ключевом режиме. В нормальном состоянии в МДП-транзисторе канал не сформирован и ключ на МДП-транзисторе разомкнут. Во время считывания на затвор подается короткий импульс, который формирует канал, ключ замыкается и снимаемый с сопротивления в цепи стока выходной сигнал будет пропорционален заряду, накопленному на емкости $p-n$ -перехода исток—подложка (емкости фотодиода). Такой режим позволяет значительно повысить чувствительность фотоприемника. Выигрыш в усилении увеличивается примерно в $t_n/t_{изм}$ раз, где t_n — время накопления заряда, которое равно времени освещения транзистора; $t_{изм}$ — время измерения, которое составляет величину порядка длительности импульса считывания, когда ключ замкнут.

МДП-транзисторы весьма перспективны за счет простоты технологии изготовления и возможности широкого использования в многоэлементных фотоприемниках большой степени интеграции.

Существует еще достаточно много различных видов фотоприемников, к которым можно отнести биполярные и полевые гетерофототранзисторы, гетерофотодиоды, однопереходные фототранзисторы, диоды Ганна, МДП-фотодиоды и т. д. Однако все эти приборы имеют в той или иной мере ограниченное применение и здесь не рассматриваются.

16.4.4. ПЗС-фотоприемники.

Эти фотоприемники относятся к классу многоэлементных интегральных приборов (интегральных схем), не имеющих аналогов среди дискретных полупроводниковых приборов. Наибольшее распространение получили матричные фото-ПЗС. В настоящее время приемники такого типа находят широкое применение в астрономии, в телевизионной технике, в цифровых фотоаппаратах и

других сферах, где используется видеотехника. Появление миниатюрных телекамер с применением ПЗС-матриц с размерами пиксела порядка нескольких мкм позволило использовать их в микрохирургии, микробиологии, микровидеооптике. ПЗС применяются в устройствах специального назначения, в космической технике и других отраслях науки и промышленности.

Основные физические процессы в фото-ПЗС, во многом их устройство и особенности переноса информации не сильно отличаются от обычных ПЗС, рассмотренных в гл. 10, поэтому здесь будут отмечены некоторые функциональные особенности и специфические параметры и характеристики фото-ПЗС.

Формирователи сигналов изображений на ПЗС по сравнению с ЭЛТ различного устройства характеризуются конструктивной и технологической простотой, малыми габаритными размерами и массой, большой долговечностью и надежностью и малой потребляемой мощностью. Конструктивно-технологическое интегрирование функций фоточувствительных и сканирующих элементов в одном приборе, обусловленное самосканированием, осуществляемым в ПЗС за счет передачи зарядовых пакетов на выход формирователя сигналов изображений, является весьма перспективным для создания полностью твердотельных формирователей сигналов изображения.

ПЗС-элементы в формирователе сигналов изображений, как и обычные ПЗС (гл. 10), работают в трех режимах: *восприятие (интегрирование) изображения*, когда происходит преобразование светового потока в зарядовые пакеты (в обычных ПЗС преобразуются электрические сигналы в заряды); *хранение зарядовых пакетов*; *передача (сканирование—считывание) зарядовых пакетов* на выход устройства.

В режиме восприятия изображений световой поток, отраженный от объекта, поступает на поверхность ПЗС и вызывает генерацию электронно-дырочных пар в полупроводниковой подложке. В потенциальных ямах ПЗС носители разделяются, в результате чего накапливается пространственно распределенная структура зарядовых пакетов, соответствующая воспринимаемому образу объекта.

Широкое распространение получили две разновидности формирователей сигналов изображений на ПЗС: *линейные (строчные)*, регистрирующие за период интегрирования одну линию изображения, и *матричные (плоскостные)*, в которых записывается весь образ сразу. При этом в обоих случаях используются ли-

бо временное, либо пространственное разделение режимов восприятия и сканирования.

При временном разделении функции восприятия и сканирования выполняются одними и теми же элементами за счет схем управления. При пространственном разделении ПЗС включает светочувствительную область, которая принимает световой поток и преобразует его в картину распределения зарядов, и область хранения, защищенную от света, в которую после интегрирования передается вся картина распределения зарядов и затем в режиме сканирования информация из этой области поступает на выход.

При временном методе все элементы ПЗС используются как светочувствительные ячейки, что позволяет получить максимальное разрешение. К недостаткам можно отнести: усложнение управляющих схем; влияние засветки, поскольку за время кадра изображение не только воспринимается, но и передается.

При пространственном разделении для получения той же разрешающей способности требуется удвоенное количество элементов, а следовательно, и увеличение площади кристалла. Достоинством метода является увеличение длительности интегрирования и защищенности от света области сканирования.

Для реализации указанных методов используются четыре основных способа организации матричных формирователей сигналов изображений на ПЗС: *кадровая, строчная, строчно-кадровая, адресная*.

При кадровой, строчной и строчно-кадровой организации формирователей сигналов изображения на ПЗС требуется отсутствие дефектов во всех элементах матрицы. Неисправность одного элемента вызывает потерю информации всего передающего столбца или строки. Адресная организация устраняет этот недостаток, поскольку в этом случае информация по координатно выбирается из отдельных элементов и реализуется в ПЗС с *инъекцией* заряда в подложку. В этой разновидности ПЗС зарядовая связь существует в парах МДП-конденсаторов, связанных между собой с помощью p -области. Затворы конденсаторов, подключенных соответственно к горизонтальной и вертикальной шинам ПЗС, представляют собой матрицу таких светочувствительных пар. Другим достоинством адресной организации является то, что для считывания зарядового пакета из любого элемента требуется только один перенос заряда, поэтому практически отсутствуют потери зарядового пакета и искажения передаваемой

информации. К недостаткам этого способа относятся: малый динамический диапазон из-за малости тока сигнала, сравнимого с паразитными токами; необходимость иметь чувствительные пороговые схемы на выходе; относительно большое время считывания (несколько мкс), что обусловлено временем рекомбинации и инжектированных из светочувствительного элемента дырок с электронами подложки; для управления по координатной выборке необходимость использования мощных сдвиговых регистров, занимающих большую площадь кристалла; меньшая, чем в обычных ПЗС, степень интеграции из-за большого количества пересекающихся шин и больших размеров кристалла.

Совершенствование технологии изготовления позволяет существенно улучшить параметры таких ПЗС.

Структура пиксела ПЗС-матриц, применяемых в цветном телевидении, отличается от структуры пиксела черно-белых матриц. Пиксел цветной матрицы состоит из комбинации трех пикселов, каждый из которых с помощью соответствующего фильтра регистрирует свет либо в красной, либо в зеленой, либо в голубой части оптического спектра. Таким образом, с каждого элемента цветной ПЗС-матрицы снимается три сигнала.

ПЗС в качестве приемников изображения изготавливают с полихромными электродами, получаемыми осаждением кремния из газовой фазы. После легирования бором или фосфором его можно использовать в качестве проводящего слоя. Термическое же окисление поликремния позволяет получить качественный межфазный диэлектрик, а его прозрачность облегчает использование ПЗС в качестве приемников изображения. Применение этой технологии позволяет осуществлять прием света не со стороны электродов, а с противоположной стороны, что приводит к ряду преимуществ, поскольку, в частности, уменьшаются потери из-за поглощения и отражения света электродами и межсоединениями.

Основными специфическими параметрами и характеристиками ПЗС как приемников изображения, помимо характеристик обычных ПЗС, являются:

- светочувствительность;
- пороговая светочувствительность;
- область спектральной чувствительности;
- разрешающая способность;
- время интегрирования;
- частотно-контрастная и светосигнальная характеристики;
- динамический диапазон.

Рассмотрим коротко некоторые из перечисленных параметров и характеристик.

Одним из основных факторов, определяющих **светочувствительность (фоточувствительность)**, является коэффициент поглощения α (см. п. 17.1). Коэффициент α резко уменьшается при увеличении длины волны света λ . Поэтому область длин волн, в которой осуществляется эффективное преобразование светового потока в информационные заряды, называемая **областью спектральной чувствительности**, ограничена. Длинноволновая граница определяется шириной запрещенной зоны полупроводника и для кремния она составляет 1,1 мкм. Коротковолновая граница ограничена величиной 0,3—0,4 мкм из-за сильного поглощения коротковолновых квантов света в узком приповерхностном слое, где происходит интенсивная рекомбинация фотогенерируемых носителей заряда. Светочувствительность (чувствительность) тесно связана с **квантовой эффективностью**, представляющей отношение числа зарегистрированных (принятых) зарядов к числу поступающих на светочувствительную область ПЗС фотонов. Квантовая эффективность — это квантовый выход светорегистрирующей части приемника, умноженный на коэффициент преобразования заряда фотоносителей в зарегистрированный полезный сигнал. Поскольку коэффициент преобразования заряда меньше единицы, то квантовая эффективность также меньше квантового выхода, являющегося отношением числа образовавшихся в полупроводнике фотоносителей к числу поступающих на полупроводник фотонов. Чувствительность и квантовая эффективность связаны между собой линейной зависимостью. Часто чувствительность делят на интегральную (суммарную во всем спектральном диапазоне приемника) и монохроматическую, которая измеряется в мА/Вт и характеризуется потоком излучения в спектральной полосе шириной 1 нм.

Пороговая светочувствительность часто определяется как минимальная (пороговая) освещенность объекта, при которой можно различить переход от черного к белому, или минимальная (пороговая) освещенность матрицы. Минимально возможное значение освещенности обусловлено темновыми токами в каждом элементе матрицы.

От числа элементов ПЗС-матрицы напрямую зависит один из основных параметров приборов изображения — **разрешающая способность**, на которую также влияют характеристики электронной схемы обработки сигналов и оптики. Разрешающая способность

определяется как максимальное количество черных и белых полос (т. е. количество переходов от черного к белому, формируемых *штриховой мирой* с различными пространственными периодами расположения черных и светлых штрихов), которые могут быть переданы камерой и различимы системой регистрации на предельно обнаруживаемом контрасте. За единицу измерения разрешающей способности принимается телевизионная линия (ТВЛ). Разрешение по вертикали у большинства камер ограничивается телевизионным стандартом — 625 строк телевизионной развертки, а по горизонтали камеры имеют разрешение 570 и более ТВЛ. Эффективное разрешение цветных камер хуже примерно на 30—40%, чем черно-белых, поскольку особенности формирования цветного изображения требуют дополнительных элементов, что приводит к дополнительным потерям.

Характеристики фотоприемных матриц отражают их функциональные возможности по преобразованию оптической информации в электрический видеосигнал. Светосигнальная и частотно-контрастная характеристики достаточно полно описывают свойства матричных ПЗС-фотоприемников. Светосигнальная характеристика (рис. 16.28, а) связывает выходное напряжение $U_{\text{вых}}$ с входной освещенностью E , где точка 1 соответствует темновому напряжению U_T на выходе матрицы при $E = 0$; точка 2 соответствует режиму насыщения, когда потенциальная яма заполнена; точка 3 является рабочей точкой.

Напряжение насыщения на выходе матрицы $U_{\text{нас}}$ зависит от глубины потенциальной ямы используемых МДП-структур и соответствует заполнению ямы зарядами. Дальнейшая фотогенерация носителей (увеличение $E > E_n$) не приводит к увеличению накапливаемого заряда, а следовательно, не приводит к возрастанию выходного видеосигнала.

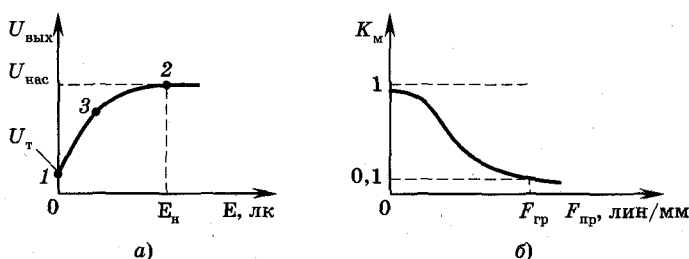


Рис. 16.28

Частотно-контрастная характеристика (ЧКХ) определяет разрешающую способность матрицы, т. е. возможность приемника различать мелкие детали изображения. ЧКХ (рис. 16.28, б) — это зависимость коэффициента модуляции входной освещенности $K_M = (E_{\text{макс}} - E_{\text{мин}})/(E_{\text{макс}} + E_{\text{мин}})$ ($E_{\text{макс}}$, $E_{\text{мин}}$ — соответственно освещенность в изображении светлого и темного штриха миры) от параметра $F_{\text{пр}}$ — пространственной частоты миры (величины, обратной пространственному периоду штриховой миры, т. е. расстоянию, равному сумме ширин соседних светлого и темного штрихов). Большие значения $F_{\text{пр}}$ соответствуют меньшему расстоянию между штрихами; $F_{\text{пр}}$ характеризуется числом линий на мм (лин/мм). Ход кривой ЧКХ зависит от тех же факторов, что и разрешающая способность прибора.

Важным параметром ПЗС является *динамический диапазон*, определяемый как отношение максимально возможного сигнала, сформированного светоприемником, к его собственному шуму. Для ПЗС этот параметр определяется как отношение наибольшего зарядового пакета, накапливаемого в пикселе, к шуму считывания. Динамический диапазон ПЗС достигает значения 50 000.

Источниками шумов в ПЗС являются: фотонный шум, шум темного сигнала, шум считывания и шум сброса. Фотонный шум является следствием дискретной природы света, представляющего поток дискретных частиц фотонов. Шум темного сигнала определяется термоэлектронной эмиссией.

Шум переноса связан с перемещением зарядового пакета по элементам ПЗС, когда часть частиц захватывается на дефектах и примесях, причем этот процесс носит стохастический характер. Шум считывания возникает при выводе сигнала из матрицы и его преобразовании в напряжение с последующим усилением.

Шум сброса связан с выводом имеющегося заряда из детектирующего узла перед вводом в него сигнального заряда. Для этой операции используется транзистор сброса. Поскольку электрический уровень сброса зависит только от температуры и емкости детектирующего узла, то это и является причиной шума.

16.5. Оптопары

Оптопарой называется прибор, содержащий в общем корпусе фотоприемник и излучатель, между которыми осуществляется электрическое и/или оптическое взаимодействие. Оптическое

взаимодействие реализуется посредством оптического канала. Оптический канал либо выполняется в виде световода, передающего энергию непосредственно от излучателя к приемнику, либо изготавливается из материала, оптические свойства которого могут изменяться при внешних воздействиях (управляемый оптический канал). Иногда между приемником и излучателем отсутствует какой-либо материал и излучение распространяется через воздушный или вакуумный промежуток.

Прибор, в котором помимо излучателя, фотоприемника и оптической среды в едином корпусе размещаются и дополнительные микронэлектронные или оптические элементы, в том числе и другие оптопары, называется *оптроном*. Различают три основные разновидности оптронов: оптопара с прямой оптической и разорванной электрической связью (рис. 16.29, а), оптрон с прямой электрической и разорванной оптической связью (рис. 16.29, б), оптрон с электрической и оптической связями (рис. 16.29, в). На рис. 16.29 обозначено: ФП — фотоприемник; И — излучатель; УС — устройство связи. Оптопары используются чаще всего в качестве элемента практически идеальной гальванической развязки и подробно будут рассмотрены ниже. Второй тип оптрона выполняет роль преобразователя параметров светового сигнала, т. е. осуществляет усиление, преобразования спектра и поляризации, а также конверсию некогерентного излучения в когерентное и т. д. При использовании многоэлементных фотоприемников можно получить и преобразование изображений. Третий тип оптронов, имея цепь обратной связи (ОС, см. рис. 16.29, в), мо-

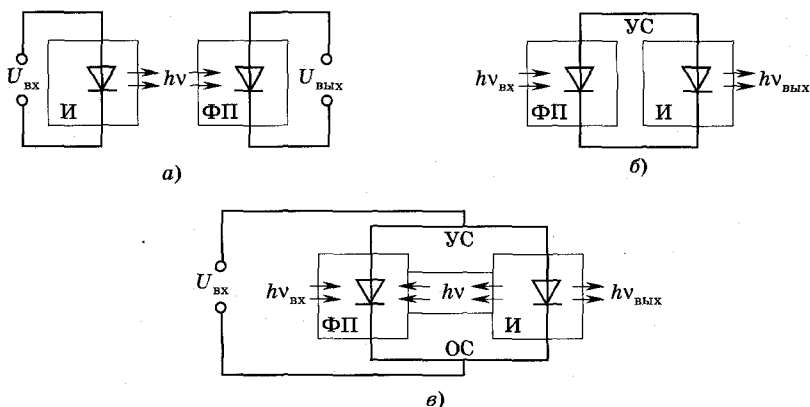


Рис. 16.29

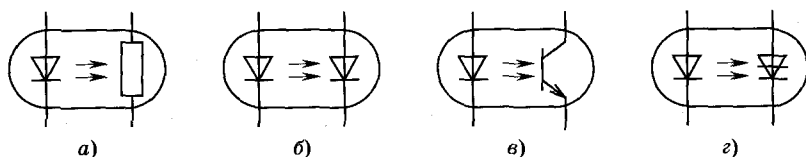


Рис. 16.30

жет выступать в качестве регенеративного устройства и производить усиление, частичное или полное восстановление входного сигнала. В таких оптронах возможны самые различные сочетания оптических и электрических входных и выходных сигналов.

Рассмотрим более подробно простейший вид оптронов — оптопары. Они обладают идеальной гальванической развязкой, большой широкополосностью, простотой конструкции, технологической, эксплуатационной и функциональной совместимостью с интегральными микросхемами.

Наибольшее распространение получили оптопары, в которых в качестве фотоприемников используются: фоторезисторы, фотодиоды, фототранзисторы и фототиристоры. Оптопары получили название по типу приемника: например, резисторная (рис. 16.30, а), диодная (рис. 16.30, б), транзисторная (рис. 16.30, в), тиристорная (рис. 16.30, г). Существует достаточно много конструктивных форм изготовления оптопар. На рис. 16.31, а приведено устройство так называемой «панельной» оптопары с расположением активных элементов в одной плоскости (1 — излучатель, 2 — фотоприемник, 3 — оптическая среда, 4 — микросхема). Эта конструкция позволяет использовать широко применяемое в микроэлектронике оборудование для монтажа кристаллов, припайки выводов. Наиболее перспективны монолитные оптопары в виде интегрированной твердотельной структуры. Пример устройства такой оптопары приведен на рис. 16.31, б (1 — светодиод; 2 — оптический канал из GaAlAs; 3 — фотодиод). Однако в настоящее время характеристики таких приборов имеют невы-

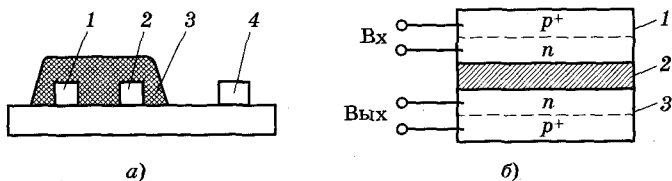


Рис. 16.31

сокие надежность и устойчивость к внешним воздействиям, плохие светопередачу и совместимость используемых материалов и т. д. Кроме того, в большинстве конструкций монолитных оптопар выводы располагаются с разных сторон кристалла и на различных уровнях, поэтому при их монтаже требуется специальное нестандартное оборудование.

В последние годы началось производство специфического вида оптронов — *волстронов* — приборов, в которых излучатель и фотоприемник жестко связаны друг с другом отрезком длинного (от десятков см до нескольких метров) волоконно-оптического кабеля. Такие оптроны очень эффективны для применения в высоковольтной и мощной аппаратуре, в установках с повышенной опасностью для работы обслуживающего персонала и т. д. Волстроны дешевы, обладают хорошей передаточной характеристикой, высокой надежностью и т. д.

Широкое распространение получили оптоэлектронные схемы, включающие оптопары и микросхемы. Например, оптоэлектронный переключатель состоит из диодной оптопары и ключевого усилителя. Такое сочетание позволяет добиться полного согласования по уровню сигналов со стандартными логическими микросхемами. Существует достаточное разнообразие оптоэлектронных микросхем, находящих применение в различных областях.

Принцип работы оптопары можно пояснить, используя прибор, представленный на рис. 16.31, б. Предположим, на вход оптопары «Вх» поступает импульс тока $I_{вх}$ или напряжения $U_{вх}$, который преобразуется светоизлучателем 1 в импульс светового потока. Излучение проходит через оптическую среду (канал) 2 и в фотоприемнике 3 преобразуется в электрический сигнал. Исходный сигнал, претерпевая двойное преобразование, испытывает некоторые искажения, которые должны быть минимальны.

Гальваническая развязка входной цепи «Вх» и выходной «Вых» достигается за счет оптически прозрачной диэлектрической среды между приемником и излучателем, которые должны быть оптически согласованы. Спектральное согласование светоизлучателя и фотоприемника реализуется соответствующим выбором их материалов. В качестве светоизлучателей в оптопарах используются преимущественно светодиоды. Применение оптических квантовых генераторов (лазеров) оправдано в быстродействующих устройствах.

Для оценки свойств оптронов используются следующие параметры, определяющие характеристики излучателей, фотоприемников, оптического сигнала и электронных устройств:

$S_{\text{и}}$ — эффективность излучателя, зависящая от типа прибора, режима возбуждения, температуры и т. д., Вт/А; $S_{\text{ф}}$ — чувствительность фотоприемника с учетом внутреннего усиления, А/Вт; $K_{\text{опт}}$ — передаточный коэффициент оптического канала; $K_{\text{у}}$ — передаточный коэффициент электронной схемы связи, в том числе и обратной, осуществляющей взаимодействие приемника и излучателя.

При расчетах оптронов в зависимости от режима работы применяются как дифференциальные, так и интегральные параметры.

Одним из основных параметров оптронов является коэффициент передачи по какому-либо параметру. Для оптопары это может быть коэффициент передачи по току, представляющий собой отношение выходного и входного токов:

$$K_i = i_{\text{вых}}/i_{\text{вх}} = S_U K_{\text{опт}} S_{\text{ф}}. \quad (16.10)$$

Используя значения S_U и $S_{\text{ф}}$, рассмотренные при описании излучателей и фотоприемников, и зная коэффициент передачи оптического канала (среды), нетрудно рассчитать общий коэффициент передачи оптопары.

Для регенеративного оптрона обычно рассматривают коэффициент регенерации оптрона $K_{\text{рег}}$. Этот параметр характеризует изменение величины сигнала при прохождении замкнутой системы приемник—излучатель—устройство связи:

$$K_{\text{рег}} = S_U K_{\text{опт}} S K_{\text{у}}, \quad (16.11)$$

где S — чувствительность устройства связи.

В регенеративном оптроне при $K_{\text{рег}} \gg 1$ может возникнуть генерация, т. е. вся система становится неустойчивой. Из-за двукратного преобразования энергии в этом типе оптронов (например, оптический сигнал сначала преобразуется в электрический и затем электрический в оптический) КПД мал.

Для диодных и транзисторных оптопар, применяемых в цифровых системах передачи информации, используют комбинированный параметр (параметр качества), представляющий собой отношение коэффициента передачи по току K_i к времени за-

держки распространения сигнала $t_{зд}$, т. е. $K_i/t_{зд} = K$ (c^{-1}). Диодные оптопары имеют лучший параметр качества $K \leq 10^7 c^{-1}$ по сравнению с транзисторными, но очень низкий коэффициент K_i ($\sim 0,02 \dots 0,03$), и поэтому после диодных оптопар требуется дополнительное усиление выходного сигнала, как правило, с использованием микросхем.

Транзисторные и тиристорные оптопары, применяемые для оптической коммутации силовоточных высоковольтных цепей, имеют в качестве критерия качества коэффициент передачи по мощности, определяемый как отношение мощности в выходной цепи к мощности на входе. Этот параметр принимает значения $10^6 \dots 10^7$. Времена переключения тиристорных и транзисторных пар имеют типичные значения от 1 до 50 мкс.

Резисторные оптопары характеризуются выходным сопротивлением, которое может изменяться на 7—8 порядков в зависимости от режима во входной цепи. Этот вид оптопар наиболее распространен в аналоговых устройствах. Недостаток резисторных оптопар — их низкое быстродействие и нестабильность параметров при изменении температуры.

При анализе цепей оптопара представляется обычно четырехполюсником. В оптроне с управляемым оптическим каналом между излучателем и фотоприемником помещается электрооптический или магнитооптический элемент, светопропускание которого регулируется электрически. Такой оптрон представляет собой шестиполюсник с двумя двухполюсными входами (по цепи излучателя и по цепи управления оптическим каналом) и одним двухполюсным выходом.

16.6. Солнечные преобразователи

Солнечный преобразователь (СП), или солнечный фотоэлемент, представляет собой полупроводниковый прибор, преобразующий солнечное излучение в электрическую энергию.

В основу работы СП положен фотовольтаический эффект в $p-n$ -переходе. Совокупность СП, объединенных в единой, как правило плоской, панельной конструкции, называется *солнечной батареей*. Солнечные батареи являются основным источником энергии на спутниках и космических кораблях. Они начинают успешно применяться и на наземных энергоустановках различной мощности.

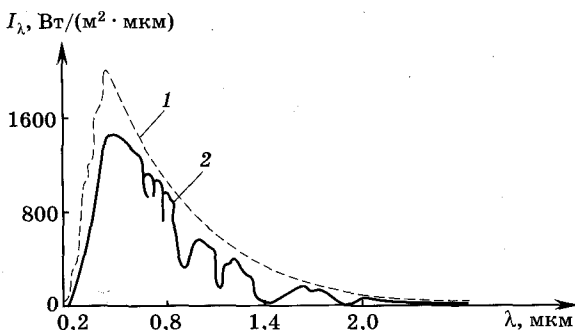


Рис. 16.32

Поскольку СП предназначены для преобразования солнечной энергии, необходимо стремиться идеально согласовать их спектральные характеристики со спектром излучения Солнца. На рис. 16.32 приведено распределение интенсивности солнечного излучения по спектру: кривая 1 соответствует солнечному спектру в свободном пространстве за пределами атмосферы, т. е. в условиях работы СП на спутниках и космических летательных аппаратах; 2 — солнечный спектр на поверхности Земли, когда Солнце находится в зените. Для первого случая (кривая 1) полная мощность излучения составляет 1353 Вт/м^2 , для второго (кривая 2) — 925 Вт/м^2 .

Существует достаточно много разнообразных конструкций солнечных элементов. Рассмотрим в качестве примера устройство кремниевого солнечного элемента с $p-n$ -переходами (рис. 16.33), поскольку оно является своеобразным эталоном при создании всех солнечных батарей. Солнечный элемент, изображенный на рис. 16.33, состоит из мелкого $p-n$ -перехода (1), сформированного у поверхности; лицевого (поверхностного) полоскового омического контакта 2; просветляющего покрытия на лицевой поверхности 3 тылового омического сплошного контакта 4. Рис. 16.33, б иллюстрирует структуру полоскового электрода 2 на виде сверху. Недостаток этой конструкции — наличие тени от металлических контактов, уменьшающее эффективность прибора, и большой коэффициент отражения света от поверхности элемента.

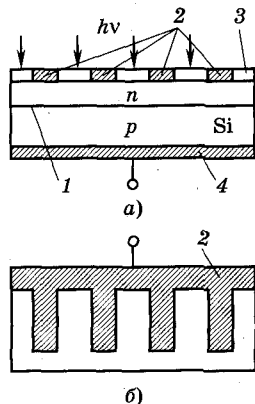


Рис. 16.33

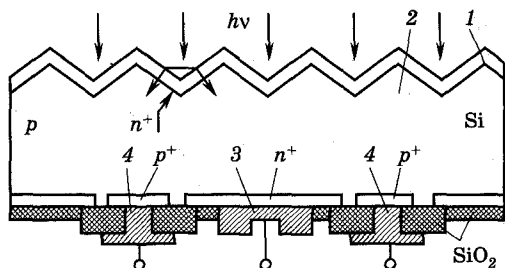


Рис. 16.34

Одним из лучших по своим характеристикам является солнечный элемент, показанный на рис. 16.34, где 1 — текстурированная поверхность с повышенным уровнем легирования донорами (n^+); 2 — p -область элемента; 3 — металлические контакты n^+ коллектора; 4 — металлические выводы базы. Сверху элемент имеет просветляющее покрытие (на рисунке не изображено).

Текстурированная поверхность создается с помощью анизотропного травления поверхности кремния, ориентированной в определенном кристаллографическом направлении. В рассматриваемом элементе отсутствует тень от металлических электродов и проще реализуются межсоединения элементов, поскольку здесь электроды располагаются на тыловой поверхности. Этот прибор функционирует подобно биполярному транзистору $n^+ - p - n^+$ -типа с изолированным n^+ -эмиттером, расположенным на текстурированной лицевой поверхности. Рассмотрим основные физические процессы, протекающие в солнечном преобразователе (элементе).

В основе устройства обычного солнечного элемента лежит освещаемый $p - n$ -переход, работающий без внешнего электрического смещения. В солнечных элементах используется собственное поглощение. Когда на элемент попадает солнечный свет, фотоны с энергией, большей энергии ширины запрещенной зоны, генерируют электронно-дырочные пары, при этом для ионизации атомов достаточно энергии фотона $h\nu = \Delta E_g$. Остальная энергия фотонов переходит преимущественно в тепло. Как неоднократно отмечалось, скорость генерации при удалении от поверхности уменьшается, при этом короткие волны поглощаются в более верхних слоях по сравнению с более длинными (см. п. 16.2). Скорости генерации электронно-дырочных пар как функции расстояния x от поверхности солнечного элемента, показанного на рис. 16.35, а

(заштрихованная часть соответствует $p-n$ -переходу), для длинноволнового и коротковолнового излучений приведены на рис. 16.35, б. Поскольку в верхних слоях образуется множество фотоносителей, за счет диффузии они начинают перемещаться в сторону обедненной области перехода, где разделяются внутренним полем $p-n$ -перехода. Электроны остаются в n -области, а дырки, для которых поле перехода является ускоряющим, перебрасываются в p -область. Электроны и дырки, появляющиеся в обедненном слое под действием света, выносятся из него сильным электрическим полем перехода, практически не рекомбинируя. Поэтому фототок обедненного слоя определяется числом фотонов, поглощаемых в этом слое в единицу времени.

Полный фототок, возникающий в фотоэлементе при поглощении света, равен сумме дырочного и электронного токов, протекающих через границы перехода, и дрейфового фототока, рожденного непосредственно в $p-n$ -переходе. Если цепь элемента разомкнута, фотоносители накапливаются в соответствующих областях и компенсируют часть объемного заряда ионов примеси в обедненной области $p-n$ -перехода. Потенциальный барьер на переходе уменьшается пропорционально величине компенсации этого заряда, и на выводах элемента возникает напряжение, которое при разомкнутой внешней цепи называется **напряжением холостого хода** $U_{\text{хх}}$. Это явление, как отмечалось в п. 16.4, называется **фотовольтаическим эффектом**. Если цепь замкнуть накоротко ($R_{\text{н}} = 0$), в ней потечет ток, обусловленный фотоносителями. Этот ток называется **током короткого замыкания** $I_{\text{кз}}$. При наличии нагрузки, т. е. когда $R_{\text{н}} \neq 0$, в цепи протекает ток меньше $I_{\text{кз}}$ и на $R_{\text{н}}$ создается падение напряжения $U_{\text{н}}$. Если на освещенный элемент подавать внешнее обратное смещение, к фототоку добавляется обратный ток $p-n$ -перехода, т. е. при обратном смещении полный ток равен $I_{\text{кз}} + I_{\text{обр}}$. Этот режим работы используется в фотоприемниках (см. п. 16.4). При прямых смещениях, больших $U_{\text{хх}}$, ток инжекции через переход начинает преобладать над фототоком и освещаемый переход рабо-

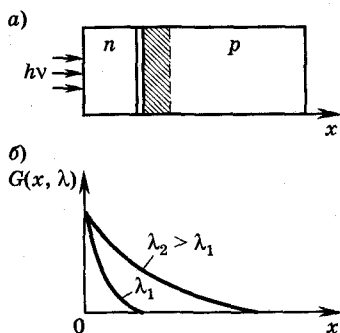


Рис. 16.35

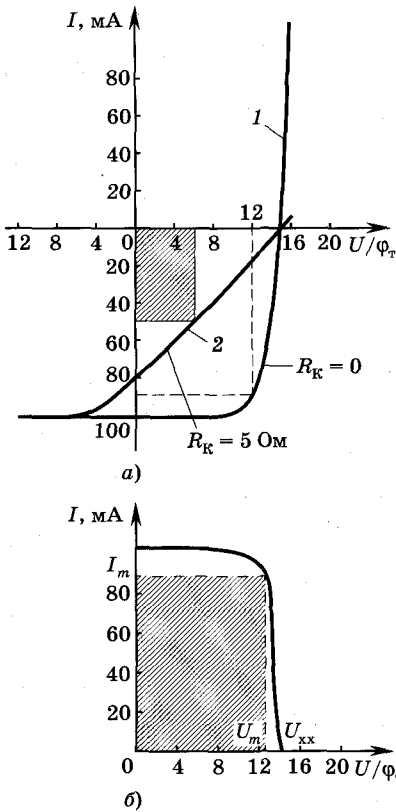


Рис. 16.36

ном на рис. 16.36, б. Если в (16.12) положить $I = 0$, можно получить выражение для напряжения холостого хода ($U = U_{xx}$):

$$U_{xx} = \varphi_T \ln [(I_\phi/I_0) + 1] \approx \varphi_T \ln (I_\phi/I_0). \quad (16.13)$$

Из формулы (16.13) хорошо видна роль тока I_0 . Величина U_{xx} возрастает с уменьшением обратного тока, когда он достигает насыщения при $I_\phi = \text{const}$, поэтому в солнечных фотоэлементах стремятся по возможности существенно понизить силу обратного тока. Формулы (16.12) и (16.13) позволяют рассчитывать значения тока I_m и напряжения U_m , при которых фотоэлемент выдает максимальную выходную мощность $P_m = I_m U_m$. Для этого, умножая левую и правую части (16.12) на U , получаем формулу для выходной мощности P и находим ее максимум из условия $dP/dU = 0$.

тает как обычный $p-n$ -переход. ВАХ такого прибора определяется выражением (см. п. 2.3)

$$I = I_0 [\exp (qU/kT) - 1] - I_\phi, \quad (16.12)$$

где I_0 — обратный ток идеализированного $p-n$ -перехода (см. п. 2.3); I_ϕ — ток, обусловленный фотоносителями.

Если взять за основу данные, соответствующие расчетной ВАХ идеализированного $p-n$ -перехода (п. 2.3, рис. 2.4), и учесть роль I_ϕ , то график ВАХ, определяемый формулой (16.12), имеет для $I_\phi = 100$ мА и $I_0 = 0,1$ мкА вид, изображенный на рис. 16.36, а, кривая 1. Значения токов и напряжений, соответствующие части ВАХ, расположенной в четвертом квадранте, отвечают случаю, когда прибор является источником энергии. Часто эту часть ВАХ солнечного элемента представляют в виде, показанном на рис. 16.36, б.

Соответствующие выражения для I_m и U_m имеют вид

$$I_m = [I_0 U_m \exp(U_m/\varphi_T)]/\varphi_T; \quad (16.14)$$

$$U_m = \varphi_T \ln \{[(I_\phi/I_0) + 1]/[1 + U_m/\varphi_T]\} = \\ = U_{xx} - \varphi_T \ln(1 + U_m/\varphi_T). \quad (16.15)$$

(Вторая формула соотношения (16.15) получена с учетом выражения (16.13) для U_{xx} .)

На рис. 16.36, б нанесены значения I_m и U_m и построен прямоугольник, площадь которого соответствует максимальной мощности P_m , т. е. для нахождения P_m необходимо в ВАХ фотоэлемента вписать прямоугольник с максимальной площадью. Таким образом, полученные выражения (16.14) и (16.15) определяют максимальную мощность, которая выделяется в нагрузку при поглощении излучения и при оптимальном согласовании внутреннего сопротивления элемента с сопротивлением внешней цепи. Максимальная эффективность преобразования η (максимальный КПД преобразования оптического излучения в электрическую энергию) реализуется при таком выборе и изготовлении материала, когда обратный ток будет минимальным; η_m определяется отношением максимальной выходной мощности P_m к мощности падающего излучения. Предельные значения КПД идеализированного преобразователя можно получить, если допустить, что потери в фотоэлементе обусловлены только излучательной рекомбинацией.

Казалось бы, для получения максимального КПД необходимо иметь минимальные значения I_0 и, следовательно, целесообразно использовать полупроводники с возможно более широкой запрещенной зоной. Однако между шириной запрещенной зоны ΔE_g и энергией фотона $h\nu$ существуют оптимальные соотношения для каждого реального кристалла. Если $h\nu > \Delta E_g$, то разность энергий $h\nu - \Delta E_g$ переходит в основном в тепло. Кроме того, как видно из рис. 16.32, в спектре излучения Солнца большая часть потока фотонов обладает энергией

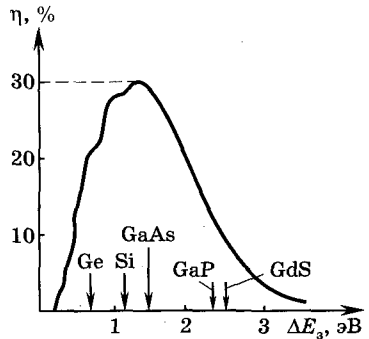


Рис. 16.37

порядка одного эВ. С учетом этого на основании проведенных расчетов было показано, что максимальный КПД достигается при $\Delta E_g = 1,35$ эВ. На рис. 16.37 приведена зависимость $\eta = f(\Delta E_g)$ для солнечного элемента, расположенного на Земле при $T = 300$ К и при освещении Солнцем, находящимся под углом 45° к горизонту. Если осуществить тысячекратную концентрацию солнечного света с помощью оптических систем, то максимальный КПД возрастет с 31 до 37%, что связано с увеличением U_{xx} .

В реальных солнечных преобразователях максимальный КПД заметно ниже идеализированного из-за влияния сопротивления высокоомной базовой области, различных видов рекомбинации, потерь на сопротивлении контактов и ряда других факторов, которые будут рассмотрены ниже. Определяющее влияние на КПД рассматриваемых приборов оказывает квантовая эффективность, или спектральный отклик, который представляет собой число генерируемых фотоэлектронов в полном токе прибора, приходящихся на один падающий фотон при облучении монохроматическим светом с длиной волны λ . Этот параметр связан с коэффициентом поглощения. В фотоэлементе, представленном на рис. 16.33, при поглощении фотонов с низкими энергиями основная доля носителей генерируется в базовой области, поскольку коэффициент поглощения в Si мал. Если энергия фотонов больше 2,5...3 эВ, то основная их доля поглощается в лицевом слое. Скорость поверхностной рекомбинации на лицевой поверхности достаточно велика, и это приводит к значительному уменьшению фотоотклика. Для анализа электрической цепи, в которую включен фотоэлемент, весьма полезно знать его эквивалентную схему, которая представлена на рис. 16.38. Генератор тока I_ϕ определяет генерацию неравновесных носителей за счет облучения светом, диод VD моделирует идеализированный $p-n$ -переход, ВАХ которого описывается выражением (2.20); параллельное сопротивление R_y обусловлено токами утечки; последовательное сопротивление R_k определяется в основном сопротивлением высокоомной базовой области

и для солнечного элемента зависит от глубины залегания перехода, концентрации примесей в n - и p -областях и от конструкции лицевого омического контакта (см. рис. 16.33). Теоретический анализ эквивалентной схемы

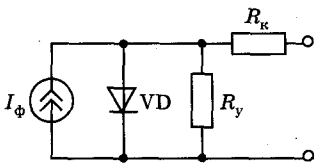


Рис. 16.38

показывает, что даже такое малое шунтирующее сопротивление, как $R_y = 100 \text{ Ом}$, практически не влияет на выходную мощность прибора. В действительности R_y много больше, поэтому его влиянием тем более можно пренебречь.

Последовательное же сопротивление R_k даже при значениях $R_k < 1 \text{ Ом}$ оказывает существенное влияние на выходную мощность. Так, при $R_k = 5 \text{ Ом}$ выходная мощность уменьшается более чем на 30% по сравнению с оптимальной величиной при $R_k = 0$. Таким образом, при учете последовательного сопротивления R_k , как и в обычном $p-n$ -переходе (см. п. 2.3), ток, напряжение и выходную мощность во внешней цепи можно определить соответствующими выражениями:

$$I = I_0 \{ \exp [q(U - IR_k)/kT] - 1 \} - I_\phi; \quad (16.16)$$

$$U = (kT/q) \ln \{ [(I + I_\phi)/I_0] + 1 \} + IR_k; P = |IU|.$$

Для иллюстрации влияния сопротивления R_k на выходную мощность можно привести следующий пример. Если при $R_k = 0$ выходная мощность $P = 1 \text{ Вт}$, то для $R_k = 2 \text{ Ом}$, $P = 0,57 \text{ Вт}$, а при $R_k = 5 \text{ Ом}$ выходная мощность, выделяемая во внешней цепи, равна лишь $0,27 \text{ Вт}$. Типичные значения R_k для кремниевого солнечного элемента, изображенного на рис. 16.33, составляют $0,4 \dots 0,7 \text{ Ом}$. На рис. 16.36, а (кривая 2) проиллюстрировано влияние R_k на ВАХ, там же показан заштрихованный прямоугольник, площадь которого определяет максимальную мощность, производимую фотоэлементом при $R_k = 5 \text{ Ом}$, и незаштрихованный прямоугольник — для максимальной мощности при $R_k = 0$.

В реальных солнечных элементах, в отличие от идеализированных, на величину прямого тока оказывает влияние рекомбинация носителей в $p-n$ -переходе, за счет чего КПД преобразования падает. Так, в кремниевых солнечных элементах при комнатной температуре рекомбинационный ток приводит к снижению эффективности преобразования на 25%. Повышение температуры также уменьшает эффективность преобразования (КПД) в основном из-за уменьшения напряжения U_{xx} , вызванного увеличением обратного тока насыщения $I_{обр}$ при высоких температурах, как это следует из формулы (16.13).

Отдельный солнечный кремниевый элемент площадью 2 см^2 имеет значения напряжения холостого хода $U_{xx} = 0,5 \dots 0,6 \text{ В}$ и

тока короткого замыкания $I_{кз} = 30...60$ мА. Большое количество таких элементов, соединенных последовательно-параллельно и собранных в единую солнечную батарею, позволяет получить источник электрической энергии с существенно большими токами и напряжениями. На рис. 16.39 показана ВАХ солнечной батареи, которая вырабатывает в наземных условиях при $T = 300$ К максимальную мощность $\sim 11,5$ Вт с КПД $\sim 13\%$. Батарея освещается Солнцем, находящимся в зените. На этом же рисунке приведены и линии постоянной мощности ($P = \text{const}$).

Выше были рассмотрены плоские кремниевые $p-n$ -фотоэлементы. Существует достаточное разнообразие конструктивных решений приборов этого вида, в которых за счет тех или иных структурных и конструктивных модификаций удается повысить эффективность. Одной из таких модификаций является элемент, устройство которого дано на рис. 16.34. Как уже отмечалось, этот прибор работает подобно биполярному транзистору с изолированным n^+ -эмиттером. Неравновесные электронно-дырочные пары, рожденные светом в n^+ -эмиттере (1) или p -базе (2), движутся (как в обычном транзисторе) к n^+ -коллектору и разделяются коллекторным переходом. В n^+ -коллектор поступают электроны, а дырки остаются в p -базе, выводы которой (4) осуществляются через p^+ -области. В этом фотоэлементе вблизи тыловой (нижней на рис. 16.34) поверхности перед металлическими омическими контактами созданы сильнолегированные полупроводниковые p^+ -слои. Между двумя базовыми областями p^+ и p возникает потенциальный барьер $q\phi_p$, препятствующий выходу электронов p -области.

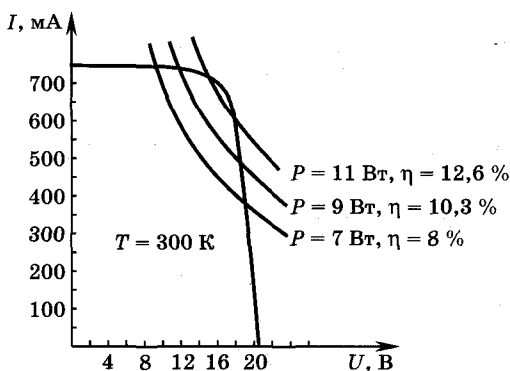


Рис. 16.39

В результате резко снижается скорость поверхностной рекомбинации электронов вблизи тыловой поверхности, т. е. уменьшаются потери фотоносителей, что вызывает увеличение спектрального отклика, особенно для длинноволнового излучения. Плотность тока короткого замыкания возрастает, повышается и напряжение холостого хода за счет роста тока короткого замыкания $I_{кз}$, уменьшения рекомбинационного тока тылового контакта и из-за дополнительного потенциального барьера между p - и p^+ -областями. Таким образом, наличие текстурированной поверхности, уменьшающей потери на отражение, и высоколегированных областей на тыловой поверхности позволяет получить КПД порядка 20% и более.

Рассмотрим теперь основные свойства других фотоэлементов, обладающих целым рядом достоинств. К таким приборам можно отнести солнечные элементы с гетеропереходами, поверхностные и тонкопленочные солнечные элементы.

Гетеропереходы представляют собой переходы, которые образуются при контакте двух полупроводников с различной шириной запрещенной зоны (см. п. 2.7). Если у фотоэлемента верхний слой, на который падает свет, сделать из полупроводника с широкой запрещенной зоной ΔE_{31} , а нижний — с узкой ΔE_{32} , то при облучении квантами света с энергией $\Delta E_{31} < h\nu < \Delta E_{32}$ фотоны проходят через слой первого полупроводника и поглощаются во втором. Первый слой с широкой запрещенной зоной ΔE_{32} играет роль оптического окна; остальные процессы аналогичны процессам в солнечных элементах с p - n -гомопереходами. Основные преимущества солнечных элементов с гетеропереходами перед приборами с p - n -гомопереходами следующие:

- 1) повышение спектрального отклика на коротких длинах волн при условии, что $h\nu < \Delta E_{31}$ и фотоны поглощаются в обедненном слое второго полупроводника;
- 2) возможность получения низкого последовательного сопротивления за счет сильного легирования верхнего слоя;
- 3) высокая радиационная стойкость при достаточно толстом и широкозонном верхнем слое полупроводника.

К достоинствам *солнечных элементов с барьерами Шоттки* можно отнести:

- 1) большой выходной ток и хороший спектральный отклик из-за близкого расположения обедненного слоя к поверхности, что ослабляет негативное влияние малого времени жизни и высокой скорости поверхностной рекомбинации;

- 2) высокую радиационную стойкость;
- 3) отсутствие необходимости проведения высокотемпературной диффузии, что позволяет использовать низкотемпературную технологию, применяемую при изготовлении поликристаллических тонкопленочных солнечных элементов.

При изготовлении солнечных элементов на барьерах Шоттки слой металла должен быть настолько тонким ($\sim 10^{-6}$ см), чтобы основная доля света достигала полупроводника. В рассматриваемых приборах фототок, как правило, имеет три компонента, которые связаны с поглощением: фотонов с энергией порядка энергии барьера на границе металл—полупроводник; фотонов коротковолновой части оптического спектра в обедненной области перехода; длинных волн в нейтральной области полупроводника, где генерируются электронно-дырочные пары, диффундирующие в сторону обедненного слоя.

Первая компонента обычно составляет $\sim 0,01$ от полного тока и, как правило, не учитывается. Таким образом, две основные компоненты спектрального отклика и фототока связаны с генерацией носителей в обедненном слое и в объеме базовой области. Разделение носителей в обедненном слое происходит так же, как и в обычном $p-n$ -переходе. Переход выполняется таким образом, чтобы сильное поле обедненного слоя выносило из него носители за время, меньшее времени рекомбинации. Для увеличения фототока следует увеличивать коэффициент пропускания структуры и диффузионную длину фотоносителей. Спектральный отклик элемента с барьером Шоттки несколько меньше, чем у элемента с $p-n$ -переходом, из-за отражения и поглощения света металлической пленкой. Однако последний вид потерь может быть сведен к минимуму. Коэффициент пропускания света золотыми пленками толщиной $\sim 10^{-7} \dots 10^{-6}$ см при наличии просветляющего покрытия может достигать 90...95%.

Расчеты показывают, что КПД идеализированного элемента, не имеющего потерь на отражение и резистивных потерь, возрастает с увеличением высоты потенциального барьера на контакте металл—полупроводник. Его максимум достигается при энергии барьера, равной ширине запрещенной зоны и составляет: $\sim 25\%$ для GaAs; $\sim 22\%$ для Si; менее 13% для Ge. Эти значения КПД близки к предельным КПД элементов на $p-n$ -гомопереходах.

Для получения большой высоты потенциального барьера при контакте с полупроводником n -типа используют металлы с вы-

сокой работой выхода, а для полупроводников p -типа — металды с низкой работой выхода.

Солнечные элементы на МДП-структурах имеют между металлом и полупроводником тонкий ($\sim 10^{-5}$ см) слой диэлектрика. Преимущества этих элементов: повышенная эффективность при облучении коротковолновым светом из-за сильного поля в поверхностном слое, отсутствие кристаллических дефектов, возникающих при диффузии примесей во время создания элементов с p — n -переходами. Такие элементы имеют КПД $\sim 18\%$ в условиях облучения на Земле при освещении Солнцем, находящимся в зените. При создании таких структур не надо проводить высокотемпературную диффузию, их можно изготавливать на полукристаллических и аморфных подложках. Поэтому солнечные элементы на МДП-структурах как источники энергии наиболее выгодны с экономической точки зрения по сравнению с другими типами СП.

Тонкопленочные солнечные элементы выполняются из поликристаллических или неупорядоченных полупроводниковых пленок, нанесенных или выращенных на электрически активных или пассивных подложках (керамика, металл, стекло, пластмасса, кремний и т. д.). Они находят широкое применение из-за низкой стоимости, обусловленной простотой технологии и дешевизной используемых материалов. Однако эти приборы имеют низкий КПД и деградацию характеристик во времени, обусловленные тем, что пока не удается изготовить высококачественные пленки, слабо реагирующие с окружающей средой.

В последнее время в качестве одного из основных материалов для изготовления тонкопленочных солнечных элементов используют аморфный кремний, который называется α -Si. Кристаллический и аморфный кремний сильно различаются. В отличие от кристаллического, аморфный кремний близок по своим свойствам к прямозонному полупроводнику с шириной запрещенной зоны $\sim 1,6$ эВ, что хорошо видно из сравнения зависимостей коэффициента поглощения от длины волны для этих материалов (рис. 16.40). Пленки α -Si толщиной 1...3 мкм обычно выращивают на стеклянных подложках, покрытых слоем металла. На них были созданы солнечные элементы с p — n -перехо-

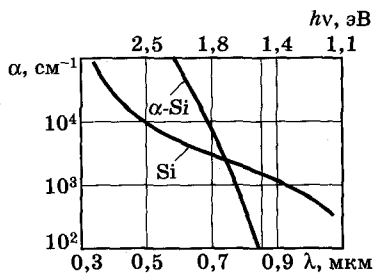


Рис. 16.40

дами и барьерами Шоттки. Для α -Si коэффициент поглощения излучения в видимой части спектра составляет $10^4 \dots 10^5 \text{ см}^{-1}$, поэтому большинство фотоносителей генерируется на расстоянии меньше одного мкм от поверхности, диффузионная длина имеет тот же порядок величины, последовательное сопротивление велико ($\sim 10^4 \text{ Ом} \cdot \text{см}$), КПД соответственно мал. Для элемента с барьером Шоттки и прозрачным металлическим слоем наилучший КПД $\sim 6\%$.

16.7. Электровакуумные фотоприемники

16.7.1. Фотоэлементы. Устройство. Принцип работы.

Электровакуумные фотоэлектронные приборы — это приборы, которые преобразуют энергию электромагнитного излучения в электрические сигналы. Принципы работы электровакуумных фотоэлектронных приборов основаны на использовании фотоэлектронной эмиссии (см. п. 11.2). В этой главе рассматриваются только основные свойства электровакуумных фотоэлектрических приборов, к которым относятся **фотоэлементы** и **фотоэлектронные умножители**.

Фотоэлементом называют **электровакуумный прибор**, использующий при своей работе явление **внешнего фотоэффекта**. Различают **электровакуумные** и **газонаполненные фотоэлементы**, которые отличаются друг от друга степенью разреженности газа в рабочем пространстве. В настоящее время наиболее широко применяются электровакуумные фотоэлементы, которые имеют два электрода: **фотокатод**, служащий источником **фотоэлектронов**, и собирающий их **анод**. Анод изготавливают в виде плоской сетки, кольца, диска и т. д. (рис. 16.41); конструктивно анод вы-

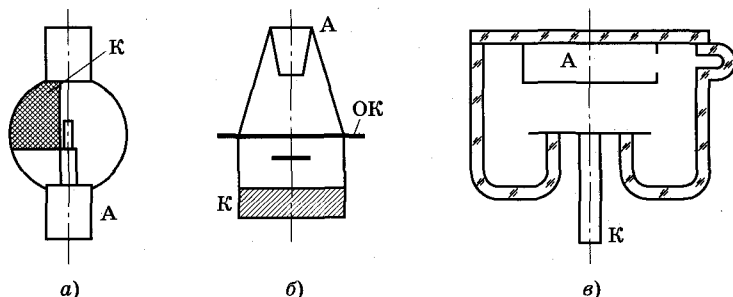


Рис. 16.41

полняется так, чтобы он не создавал больших потерь для света, падающего на фотокатод.

Основными параметрами фотокатодов являются: интегральная и монохроматическая токовая чувствительности, квантовый выход, удельное (поверхностное) сопротивление на 1 см^2 площади фоточувствительного слоя, плотность тока термоэмиссии при комнатной температуре. По области рабочего спектра фотокатоды разделяются на фотокатоды, работающие в УФ, видимой и ИК областях облучения. При этом наибольшее применение находят фотокатоды на основе полупроводниковых материалов. Фотоэлементы классифицируются на основе конструктивных признаков фотокатодов. В соответствии с этим все фотоэлементы можно разбить на три группы (см. рис. 16.41):

- 1) фотоэлементы с массивным непрозрачным фотокатодом (К), нанесенным на часть стеклянной колбы (баллона) (рис. 16.41, а) заштрихованная часть баллона); используются как приемники постоянных или модулированных низкой частотой ($f < 20 \text{ кГц}$) световых потоков в контрольно-измерительной аппаратуре, автоматике, аппаратуре звуковоспроизведения;
- 2) фотоэлементы с массивным небольшим катодом, нанесенным или на дно баллона, или на специальную металлическую пластину (рис. 16.41, б); предназначены для приема сфокусированного излучения малой интенсивности; для уменьшения токов утечки служит охранное кольцо (ОК), а выводы катода (К) и анода (А) располагаются с противоположных сторон баллона;
- 3) импульсные сильноточные элементы с фотокатодом на металлической подложке (рис. 16.41, в) имеют малое продольное сопротивление, малую инерционность ($\sim 10^{-3} \text{ с}$), линейную энергетическую характеристику и большую эмиссионную способность, достигающую 100 А при длительности импульса до 10^{-9} с .

Фотоэлементы первой, второй и третьей групп имеют рабочее напряжение менее 240 В , 100 В и 2 кВ соответственно. Для уменьшения токов утечки в баллон часто вваривают специальные охранные кольца ОК (см. рис. 16.41, б). При конструировании импульсных фотоэлементов стремятся уменьшить влияние времени пролета фотоэлектронов между катодом и анодом, междуэлектродные емкости, сопротивления и индуктивности выводов, поэтому выводы катодов делают короткими, а выводы анодов

часто выполняют в виде кольца или применяют коаксиальные выводы. Импульсные элементы, предназначенные для измерений коротких световых вспышек, имеют междуэлектродную емкость 3...4 пФ, а временное разрешение $\tau \sim 10^{-8} \dots 10^{-11}$ с.

16.7.2. Параметры и характеристики фотоэлементов.

В качестве параметров фотоэлементов, как и для полупроводниковых фотоприемников, используются чувствительность, квантовый выход (квантовая эффективность), шумы, минимально регистрируемая мощность излучения (пороговый поток), обнаружительная способность, темновой ток, постоянная времени, сопротивление. Эксплуатационные и конструктивные параметры: максимально допустимая рассеиваемая мощность, нестабильность чувствительности и темнового тока во времени, температурный коэффициент чувствительности и др.

К основным характеристикам фотоэлементов относятся: спектральные, вольт-амперные, энергетические, частотные, температурные.

Вольт-амперная характеристика (ВАХ) $I_{\phi} = f(U_a)$. ВАХ фотоэлементов называют зависимость фототока I_{ϕ} от напряжения анода U_a при неизменном световом потоке $\Phi = \text{const}$. Типичная ВАХ фотоэлемента представлена на рис. 16.42. При малых значениях U_a (участок I на рис. 16.42) в приборе реализуется режим объемного заряда (см. гл. 11), т. е. у поверхности катода за счет фотоэмиссии образуется область отрицательного объемного заряда и не все электроны попадают на анод (ток ограничен полем объемного отрицательного заряда и определяется законом трех вторых). По мере

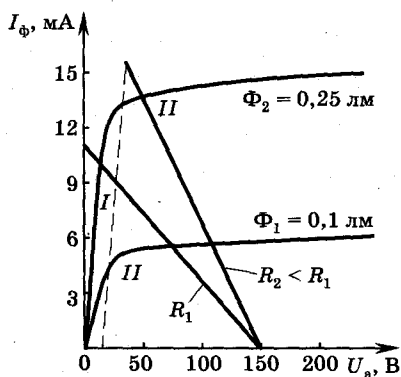


Рис. 16.42

увеличения U_a все большее количество электронов из области объемного заряда имеет возможность уйти на анод, ток растет, а плотность объемного заряда уменьшается. Часть электронов попадает на поверхность баллона, вызывая вторичную электронную эмиссию с коэффициентом $\sigma < 1$.

Участок II на рис. 16.42 соответствует режиму, при котором объемный заряд около катода отсутствует и все эмитиро-

важные электроны попадают на анод (режим насыщения). Согласно закону Столетова увеличение интенсивности светового потока будет вызывать нарастание тока эмиссии и величина тока насыщения будет больше. Значение напряжения, соответствующее началу участка насыщения, определяется конструкцией прибора и возрастает при увеличении интенсивности светового потока из-за возрастания плотности объемного заряда у поверхности фотокатода.

Энергетические (световые) характеристики $I_{\Phi} = f(\Phi)$. Энергетическими (или световыми) характеристиками называются зависимости фототока I_{Φ} от интенсивности светового потока Φ при неизменном анодном напряжении U_a . Эти характеристики линейны в широком диапазоне изменения Φ (рис. 16.43), что определяется законом Столетова, согласно которому фототок I_{Φ} пропорционален интенсивности светового потока. Отклонение от линейности при больших значениях Φ обусловлено влиянием объемного заряда (поскольку $U_a = \text{const}$) и утомлением фотокатода, которое связано с физико-химическими процессами, происходящими в фотокатоде под действием ионной бомбардировки и при взаимодействии с остаточными газами и т. д.

Величину утомления определяют как отношение

$$(S_{I_n} - S_I) / S_I = \Delta S_I / S_I,$$

где S_{I_n} — начальная (при отсутствии утомления) токовая чувствительность; S_I — токовая чувствительность при наличии утомления, когда наступает стабилизация параметров, связанных с током.

Относительное уменьшение токовой чувствительности до наступления стабилизации определяют при неизменных анодном потенциале и световом потоке.

При $\Phi = 0$ фототок I_{Φ} несколько отличается от нуля. Существует темновой ток, обусловленный термоэлектронной эмиссией электронов с фотокатода при комнатной температуре и токами утечки (проводимости) по стеклу баллона.

При наличии нагрузки (сопротивления R) в цепи фотоэлемента световая характеристика также может существенно отклоняться от линейной, осо-

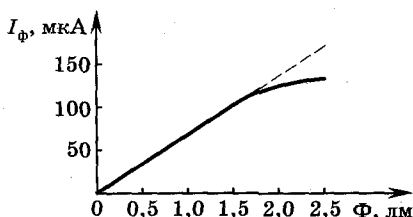


Рис. 16.43

бенно если нагрузочная прямая будет пересекать ВАХ при больших Φ в области, соответствующей режиму ограничения тока объемным зарядом (участок I , например, при $R = R_1$ на рис. 16.42).

Частотные характеристики $S_f = \psi(f)$. Частотными характеристиками фотоэлементов называют зависимость их чувствительности от частоты изменения (модуляции) интенсивности светового потока, воздействующего на фотокатод.

Частотные характеристики определяются инерционностью фотоэлементов. На рис. 16.44 приведены частотные характеристики фотоэлемента с массивным фотокатодом (кривая 1) и импульсного фотоэлемента (кривая 2). Уменьшение чувствительности при высоких частотах модуляции интенсивности света связано с временем пролета электронов от катода до анода и объясняется переходными процессами (временем релаксации) в электрической цепи, в которую включен фотоэлемент. Для обычных (неимпульсных) промышленных электровакуумных фотоэлементов время пролета $\tau_{пр}$ составляет $\sim 10^{-8} \dots 10^{-9}$ с, а междуэлектродная емкость $C \sim 10 \dots 50$ пФ, для импульсных фотоэлементов $\tau_{пр} \sim 10^{-10} \dots 10^{-11}$ с, $C \sim 3 \dots 4$ пФ. Серийные обычные электровакуумные фотоэлементы имеют частотный диапазон до 10^6 Гц, а импульсные — до 10^9 Гц.

Спектральные характеристики $S_\phi(\lambda) = f(\lambda)$. Спектральными характеристиками фотоэлементов называют зависимости монохроматической чувствительности S_λ от длины волны λ (или частоты) электромагнитного излучения, воздействующего на фотокатод электровакуумного фотоэлемента.

Спектральные характеристики для сурьмяно-цезиевого (1) и кислородно-цезиевого (2) фотокатодов показаны на рис. 16.45. Эти

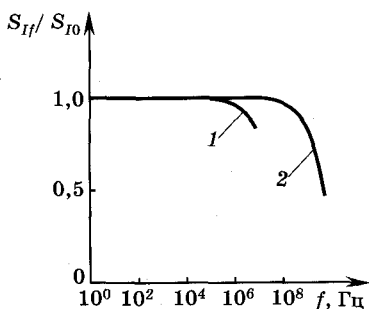


Рис. 16.44

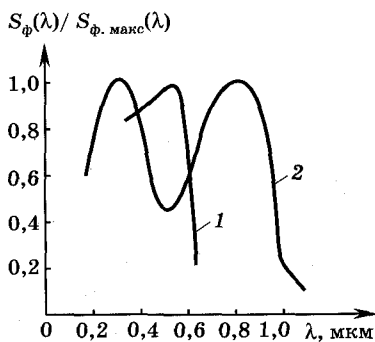


Рис. 16.45

характеристики определяются в основном электрофизическими параметрами полупроводниковых материалов, из которых изготовлены фотокатоды. Физические процессы, определяющие ход рассматриваемых характеристик, полностью аналогичны явлениям, происходящим в полупроводниковых фотоэлектрических приборах (см. п. 16.4). Конструктивные свойства также влияют на характер зависимости $S_\lambda = f(\lambda)$; к ним можно в первую очередь отнести толщину фотокатода, материал подложки, физические свойства окна (стекла) фотоприемника.

Кроме описанных сурьмяно-цезиевых и кислородно-цезиевых фотокатодов, в фотоэлементах применяют многощелочные фотокатоды, образованные соединениями сурьмы с атомами калия, натрия и цезия $\text{Na}_2\text{KSb}(\text{Cs})$. Максимальная чувствительность фотокатода достигается, когда отношение натрия к калию приблизительно равно двум, а содержание цезия много меньше, чем калия.

Рассмотренные параметры и характеристики фотоэлементов подвержены изменениям под влиянием температуры, что обычно отражают в виде зависимости параметров фотокатодов от температуры или температурных коэффициентов.

16.7.3. Фотоэлектронные умножители. Устройство. Принцип работы.

Фотоэлектронные умножители (ФЭУ) — это электровакуумные приборы, в которых ток фотоэлектронной эмиссии усиливается посредством вторичной электронной эмиссии.

ФЭУ представляет собой электровакуумный фотоэлемент, объединенный с электронной усилительной системой в едином корпусе (баллоне). Действие усилительной системы основано на явлении вторичной электронной эмиссии. Устройство фотоэлектронного умножителя одного из типов показано на рис. 16.46.

Эмитируемые фотокатодом ФК фотоэлектроны n_k , ускоряемые и фокусируемые электродами Э (в ФЭУ, изображенном на рис. 16.46, вместо этих электродов использована только сетка Э), попадают на первый **динод** D_1 . Динод (анод) — это электрод, обладающий большим коэффициентом вторичной эмиссии ($\sigma \sim 6 \dots 8$). Часть электронов теряется в фокусирующей и ускоряющей системе (сетка на рис. 16.46, а), что при расчетах обычно учитывается с помощью коэффициента эффективности γ (γ_1 — отношение числа фотоэлектронов, достигающих в единицу времени первого динода D_1 , к числу ежесекундно эмитируемых катодом фотоэлектронов n_k). Таким образом, на первый динод по-

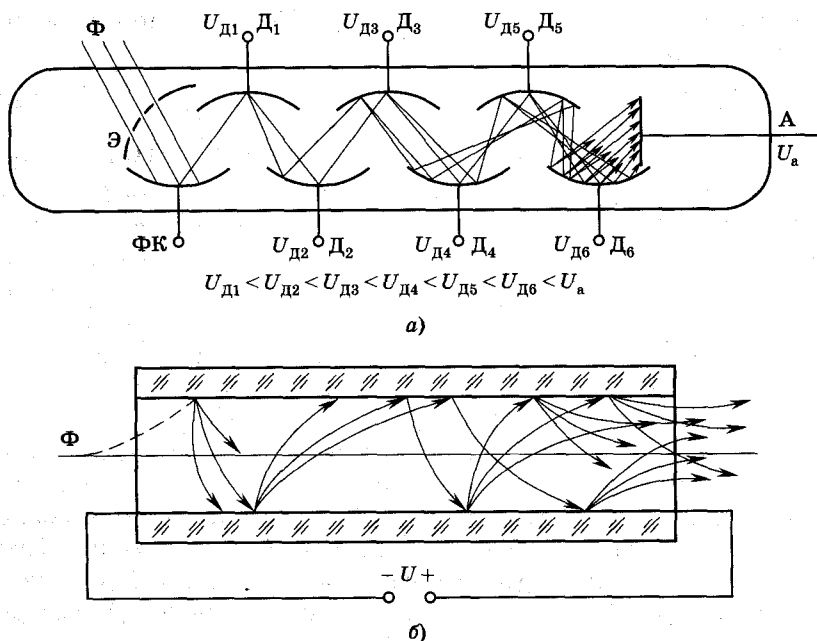


Рис. 16.46

ступит $n_k \gamma_k$ электронов. С поверхности первого динода выйдет в σ_1 раз большее число электронов, чем на него упадет. Появившиеся после бомбардировки первого динода вторичные электроны ускорятся полем второго динода D_2 , выбивают из него вторичные электроны, т. е. со второго динода при $\gamma_2 = 1$ уйдет в $\sigma_1 \sigma_2$ раз большее число электронов и т. д. К аноду придет поток электронов в σ^m раз больший (при $\gamma_k = \gamma_1 = \gamma_2 = \dots = \gamma_m = 1$; $\sigma_1 = \sigma_2 = \dots = \sigma_m$), чем было испущено катодом (m — число динодов). В общем случае с учетом эффективности каскадов количество электронов, попадающих в единицу времени на анод, можно вычислить по следующей формуле:

$$n_A = n_k \gamma_1 \sigma_1 \gamma_2 \dots \sigma_m \gamma_m = n_k \gamma_k \prod_{i=1}^m \sigma_i \gamma_i = \gamma_k M n_k, \quad (16.17)$$

где γ_i — эффективность i -го каскада усиления ($\gamma \sim 0,7 \dots 0,95$), равная отношению числа электронов, достигающих $(i + 1)$ -го динода, к числу электронов, эмитированных i -м динодом; σ_i — коэффициент вторичной эмиссии i -го динода; $\sigma_i \gamma_i$ — коэффициент усиления i -го каскада; m — число динодов.

С учетом (16.17) запишем выражение для тока, протекающего в цепи анода:

$$I_a = \gamma_k M I_k, \quad (16.18)$$

где I_k — ток фотоэмиссии с катода, величина $M = \prod \sigma_i \gamma_i$ называется коэффициентом усиления фотоэлектронного умножителя по току.

Коэффициент усиления M и эффективности γ_i сбора электронов диодами зависят как от эмитирующей способности диодов, так и от конструкции входных камер и диодных систем. В настоящее время существует достаточно большое количество типов фотоумножителей, предназначенных для различных применений (телевидение, оптоэлектроника, лазерная техника, дозиметрия, астрономия и т. д.).

Входная камера, как правило, состоит из фотокатода и электронно-оптической системы, обеспечивающей фокусирование потока фотоэлектронов в направлении первого диода. Фотокатод в зависимости от конструкции и назначения ФЭУ может быть как полупрозрачным, расположенным в торце прибора, так и массивным при боковом входе оптического сигнала. Диаметры катодов достигают нескольких десятков см, однако наиболее часто встречаются ФЭУ с размерами фотокатодов от 1,0 до 5 см.

Одним из основных требований, предъявляемых к электронно-оптическим системам ФЭУ, особенно быстродействующим, является требование минимального разброса времен пролета электронов от поверхности фотокатода до первого диода (изохронность траекторий электронов). Неодинаковость времени пролета обусловлена разбросом начальных скоростей и длин пробега электронов, вылетевших с различных участков фотокатода, неоднородностью электрических полей и рядом других факторов. Наилучшие результаты получаются во входных камерах со сферической формой электродов (см. рис. 16.46, а), где входная камера ФЭУ повторяет конструкцию каскада умножительной системы.

Диодные системы весьма разнообразны по конструкции. К ним предъявляются следующие требования: большое усиление и быстродействие, линейность энергетических (световых) характеристик, высокая эффективность, простота изготовления и эксплуатации. Наибольшее распространение получили диодные системы с электростатическими полями, обладающие наилучшими эксплуатационными характеристиками. Они могут быть разделены на системы с дискретными диодами и системы с распределенными диодами.

В первой группе наибольшее распространение получили системы, использующие фокусировку электронов, и системы «сквозного» типа (жалюзи, сетки, пленки на «прострел»), в которых умножительные каскады сконструированы таким образом, что они не требуют специальной фокусировки электронных пучков в пространстве между вторичными эмиттерами (см. рис. 16.46, б). Системы на дискретных диодах сквозного типа малочувствительны к воздействию внешних магнитных полей и из-за развитой рабочей поверхности диодов обеспечивают работу при больших токах нагрузки. К их недостаткам следует отнести более низкую эффективность диодов и временное разрешение по сравнению с ФЭУ, имеющими электростатическую фокусировку.

Системы на распределенных диодах бывают трех типов: пластинчатые, целевые и каналные. Последние из них в простейшем случае представляют собой трубку определенного калибра (отношение длины к диаметру), внутренняя поверхность которой обладает нужным электрическим сопротивлением и хорошим коэффициентом вторичной электронной эмиссии. Если на концы трубки с калибром 50...100 подать высокий потенциал ($> 2...2,5$ кВ), то в канале трубки сформируется однородное электрическое поле.

Фотоэлектроны выбивают с внутренней поверхности трубки вторичные электроны, которые под действием электростатического поля ускоряются и бомбардируют стенки канала, находящиеся под большим потенциалом (см. рис. 16.46, б). Коэффициент усиления M трубки зависит от ее калибра, поверхности канала и приложенного к его концам напряжения. Величина M достигает значений $10^5...10^6$. Канальные системы не требуют внешнего делителя напряжения, необходимого для систем на дискретных диодах, имеют простую конструкцию и малые размеры.

В последние годы были разработаны гибридные ФЭУ, в которых в качестве умножающих элементов используются полупроводниковые диодные или транзисторные структуры. Их принцип действия основан на образовании свободных носителей в полупроводнике с $p-n$ -переходом при бомбардировке его электронами с энергией ~ 10 кэВ. На $p-n$ -переход подано обратное смещение. Образовавшиеся в результате бомбардировки электронно-дырочные пары разделяются полем $p-n$ -перехода, образуя ток в цепи анода. Коэффициент усиления пропорционален коэффициенту умножения носителей в полупроводнике и

достигает величины 10^3 для диодных и 10^6 для транзисторных структур.

Гибридные ФЭУ имеют большие выходные токи ($\sim 0,5$ А в стационарном режиме и до 20 А в импульсном), малые габариты, высокое быстродействие ($\sim 10^{-10}$ с). Они не чувствительны к внешним магнитным полям.

16.7.4. Параметры и характеристики фотоэлектронных умножителей (ФЭУ).

В зависимости от вида регистрируемого сигнала в ФЭУ различают: статические параметры при измерении световых немодулированных сигналов, частотные параметры при работе с модулированными оптическими сигналами и импульсные параметры для импульсных сигналов.

Большинство характеристик и параметров ФЭУ соответствуют аналогичным параметрам и характеристикам фотоэлементов, имеющих фотокатоды таких же типов. Здесь мы рассмотрим только или специфические параметры и характеристики ФЭУ, или их особенности.

Коэффициент усиления диодной системы M и анодная токовая чувствительность S_{Ia} показаны на рис. 16.47, где кривая 1 относится к диодам с сурьмяно-цезиевым покрытием, а кривая 2 — к сплавным вторично-электронным эмиттерам. Увеличение коэффициента усиления M с ростом напряжения питания вызвано возрастанием коэффициентов вторичной эмиссии σ_i диодов. При большем числе диодов и прочих равных условиях наклон кривых увеличивается, то же самое будет наблюдаться и при больших коэффициентах вторичной эмиссии диодов.

Анодную токовую чувствительность $S_{Ia} = dI_a/d\Phi$ [А/лм] измеряют при полном освещении поверхности катода и значениях световых потоков, соответствующих линейной части световой характеристики, при которых анодные токи I_a не превышают предельных значений.

Обычно величины световых потоков выбирают в диапазоне 10^{-4} ...
... 10^{-8} лм.

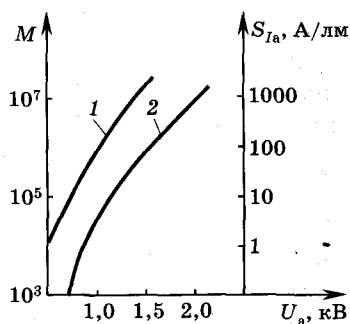


Рис. 16.47

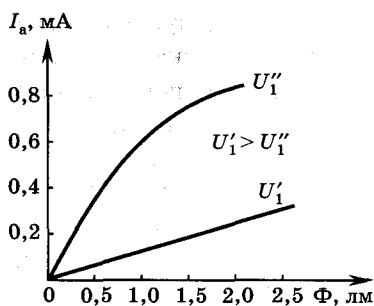


Рис. 16.48

Энергетические (световые) характеристики (рис. 16.48). Отклонение от линейности при больших Φ обусловлено влиянием пространственного заряда и утомлением фотокатода. При интенсивных световых потоках возникают большие плотности токов с фотокатода, что и приводит к формированию около его поверхности области объемного заряда, ограничивающего число электронов, летящих в на-

правлении первого динода. При импульсных сигналах диапазон линейных режимов значительно шире. Влияние на ток анода напряжения на первом диноде определяется его ВАХ. На рис. 16.48 в режиме объемного заряда это влияние иллюстрируется кривыми при двух значениях напряжения на первом диноде ($U_1'' > U_1'$). При регистрации излучения, модулированного гармоническими сигналами, инерционность оценивается по частотной характеристике. Умножители различного типа работают в частотном диапазоне от 15 до 150 МГц.

ФЭУ, предназначенные для регистрации коротких вспышек слабой интенсивности, должны иметь хорошие быстродействие ($\sim 10^{-3} \dots 10^{-10}$ с) и чувствительность. Способность ФЭУ выделять световые импульсы, незначительно различающиеся по интенсивности, характеризуется **собственным амплитудным разрешением** ФЭУ, которое тем выше, чем ниже уровень собственных шумов прибора. Минимальные значения (нижний порог) интенсивности измеряемых потоков излучения определяются темновым током и шумами. Темновой ток ФЭУ обусловлен термо- и автоэлектронной эмиссией с поверхности фотокатода и динодов, токами утечки между электродами прибора.

Темновой ток, являющийся параметром ФЭУ, дается в справочниках для определенного режима его работы. Флуктуации этого тока служат источниками собственных шумов ФЭУ, к которым добавляются флуктуации (шумы) коэффициентов усиления σ динодов.

Пороговые потоки (пороговая чувствительность) в единичной полосе частот¹ лучших ФЭУ достигают 10^{-13} лм \cdot Гц^{-1/2} [21].

¹ Этот параметр определяется световым энергетическим потоком в люменах, падающим на фотокатод ФЭУ и вызывающим анодный фототок, равный эффективному значению темнового тока, приведенному к полосе частот в 1 Гц.

Быстродействие ФЭУ оценивается по форме импульса на выходе прибора при освещении всей поверхности фотокатода световым импульсом с длительностью $< 3 \cdot 10^{-9}$ с. На основе анализа формы выходного импульса определяют следующие импульсные параметры:

τ_{ϕ} — длительность фронта анодного импульса тока, которая измеряется между уровнями 0,1 и 0,9 от его амплитудного значения I_{am} ; $\tau_{0,5}$, $\tau_{0,1}$ — соответственно длительности выходных импульсов по уровням $0,5I_{am}$ и $0,1I_{am}$;

$S = 0,8(I_{am}/\Phi)$ — крутизна фронта импульса (в линейной области изменения тока, см. рис. 16.48).

В зависимости от рабочей частоты все ФЭУ можно разделить на две большие группы:

- 1) низкочастотные ФЭУ для измерения предельно малой интенсивности световых потоков;
- 2) импульсные и высокочастотные ФЭУ для регистрации кратковременных и быстро изменяющихся малых потоков излучения.

Характерные параметры ФЭУ первой группы: темновые токи менее $2 \cdot 10^{-8}$ А, $S_{Ia} \sim 1000$ А/лм; пороговый поток в единичной полосе частот $\sim (2,5 \dots 3) \cdot 10^{-13}$ лм \cdot Гц $^{-1/2}$, диаметр фотокатодов $\sim 5 \dots 6$ мм. Малые диаметры фотокатодов обусловлены необходимостью работы при малых темновых токах и требованиями, предъявляемыми к аппаратуре, используемой в космической технике.

Быстродействующие ФЭУ конструируются таким образом, чтобы получить возможно более высокую изохронность траекторий электронов, вылетающих с различных участков электродов. Параметры таких ФЭУ: длительность фронта анодного импульса $\tau_{\phi} < 2,5$ нс; $S \sim 200$ мА/нс, $I_{am} \sim 0,5$ А. Для расширения полосы рабочих частот выводы ФЭУ выполняют коаксиальными, что обеспечивает хорошее согласование с внешними высокочастотными и сверхвысокочастотными цепями.

ФЭУ, предназначенные для приема излучения лазеров, имеют монохроматические катоды с чувствительностью ~ 4 мА/Вт, с длительностью фронта и выходного импульса $\tau_{\phi} \sim 2 \dots 3$ нс, $\tau_{0,1} \sim 0,7$ нс, с пороговым потоком в единичной полосе частот $\sim 2 \cdot 10^{-13}$ лм \cdot Гц $^{-1/2}$.

**Контрольные вопросы**

1. Классификация оптоэлектронных приборов и краткая характеристика каждого типа.
2. Какие преимущества имеет использование световых волн в оптоэлектронике по сравнению с радиоволнами?
3. Каковы основные физические механизмы поглощения света полупроводниками?
4. Излучение света полупроводниками. Светодиоды: типы, характеристики и параметры.
5. Общая характеристика фотоприемников.
6. Фоторезистивный эффект и фоторезисторы. Параметры и характеристики фоторезисторов.
7. Фотовольтаический эффект, $p-n$ -фотодиоды: ВАХ, параметры.
8. Устройство, параметры, характеристики $p-i-n$ -фотодиодов и лавинных фотодиодов.
9. Фототранзисторы, фоторезисторы, составные фототранзисторы: особенности устройства и параметры.
10. Типы оптронов; схемы и принципы работы.
11. Параметры и характеристики оптронов.
12. Солнечные преобразователи: устройство, ВАХ, параметры, эквивалентная схема, солнечные батареи и их характеристики.
13. Солнечные элементы на гетеропереходах, с барьером Шоттки, тонкопленочные солнечные элементы: достоинства и недостатки, особенности параметров.
14. Фотоэлементы: устройство, принцип работы, характеристики и параметры.
15. Фотоэлектронные умножители: устройство, принцип работы, параметры и характеристики.

ПРИБОРЫ КВАНТОВОЙ ЭЛЕКТРОНИКИ

Глава 17

ОСНОВЫ КВАНТОВОГО УСИЛЕНИЯ

17.1. Индуцированные и спонтанные переходы

Внутренняя энергия атомов, молекул, ионов, различных соединений и сред, образованных указанными частицами, квантована. Каждая молекула (атом, ион) может взаимодействовать с электромагнитным излучением, совершая переход с одного энергетического уровня на другой. При этом происходит изменение внутренней энергии от одного значения, соответствующего определенному движению и ориентации электронов и ядер, к другому значению, соответствующему другим движениям и ориентациям.

Энергия поля излучения также квантована, так что обмен энергией между полем и взаимодействующими с ним частицами может происходить только дискретными порциями.

Частота излучения, связанного с переходом атома (молекулы, иона) между энергетическими состояниями, определяется частотным постулатом Бора

$$E_2 - E_1 = h\nu,$$

где E_1 , E_2 — соответственно энергия частицы (атом, молекула, ион) в верхнем и нижнем энергетических состояниях, h — постоянная Планка, ν — частота.

Не все переходы между энергетическими состояниями являются возможными. Если частица находится в верхнем состоянии, то имеется определенная вероятность, что через некоторый период времени она перейдет в нижнее состояние и произойдет изменение энергии. Этот переход может быть как излучательным, так и безызлучательным, как под влиянием внешнего воздействия, так и без него. В среде, обладающей дискретными уровнями энергии, существуют три вида переходов: *индуцированные, спонтанные и релаксационные*.

При индуцированных переходах квантовая система может переводиться из одного энергетического состояния в другое как с поглощением квантов энергии внешнего поля, так и с излучением кванта электромагнитной энергии. Индуцированное, или вынужденное, излучение стимулируется внешним электромагнитным полем. Вероятность индуцированных переходов (как излучательных, так и безызлучательных) отлична от нуля только для внешнего поля резонансной частоты, энергия кванта которого совпадает с разностью энергий двух рассматриваемых состояний. Индуцированное излучение полностью тождественно излучению, вызывающему его. Это означает, что электромагнитная волна, созданная при индуцированных переходах, имеет ту же частоту, фазу, поляризацию и направление распространения, что и внешнее излучение, вызвавшее индуцированный переход.

Если рассматриваемая квантовая система обладает двумя уровнями энергии $E_2 > E_1$ (рис. 17.1), при переходах между которыми излучается или поглощается квант энергии $h\nu$, то частицы рассматриваемой системы находятся в поле их собственного излучения, спектральная объемная плотность энергии которого на частоте перехода равна ρ_ν . Это поле вызывает переходы как из нижнего состояния в верхнее, так и из верхнего в нижнее (рис. 17.1, а). Вероятности этих индуцированных

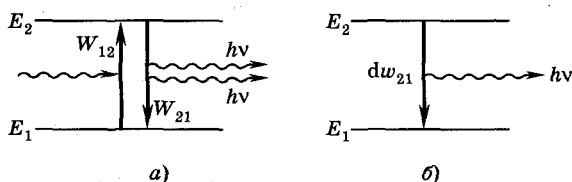


Рис. 17.1

переходов для поглощения и излучения W_{12} и W_{21} в единицу времени соответственно пропорциональны ρ_ν :

$$\begin{aligned} W_{12} &= B_{12}\rho_\nu, \\ W_{21} &= B_{21}\rho_\nu, \end{aligned} \quad (17.1)$$

где B_{12} , B_{21} — коэффициенты Эйнштейна соответственно для индуцированного поглощения и излучения.

Спонтанные переходы (рис. 17.1, б) происходят из верхнего энергетического состояния E_2 в нижнее E_1 самопроизвольно — без внешнего воздействия — с излучением кванта $h\nu$, т. е. они являются излучательными. Вероятность dw_{21} таких переходов не зависит от внешнего электромагнитного поля и пропорциональна времени. За время dt

$$dw_{21} = A_{21} dt, \quad (17.2)$$

где A_{21} — коэффициент Эйнштейна для спонтанного излучения.

Полное число переходов в единицу времени из энергетического состояния E_2 («верхнего») в «нижнее» состояние E_1 (переход $2 \rightarrow 1$) равно произведению числа частиц n_2 в состоянии 2 на вероятность перехода $2 \rightarrow 1$ в единицу времени для одной частицы.

При термодинамическом равновесии ансамбль частиц не теряет и не приобретает энергии, т. е. число излученных квантов (число переходов из верхнего энергетического состояния E_2 в нижнее E_1 состояние) должно быть равно числу поглощенных квантов (числу переходов из состояния E_1 в E_2).

При тепловом равновесии распределение населенности частиц по уровням энергии подчиняется закону Больцмана

$$n_2/n_1 = g_2 \exp(-E_2/kT)/g_1 \exp(-E_1/kT), \quad (17.3)$$

где n_1 , n_2 — соответственно число частиц, находящихся в состояниях E_1 и E_2 ; g_1 , g_2 — статистические веса (кратности вырождения) уровней 2 и 1. Пропорциональность населенностей уровней их статистическим весам обусловлена тем, что вероятность пребывания частицы в некотором квантовом состоянии определяется только энергией этого состояния, а различные квантовые состояния, целиком определяемые полным набором квантовых чисел, могут иметь одинаковые энергии.

При термодинамическом равновесии число излучательных переходов из верхнего состояния в нижнее (N_2) равно числу переходов из нижнего состояния в верхнее (N_1), происходящих с поглощением излучения. Число переходов N_2 определяется вероятностью одного перехода, умноженной на населенность уровня с энергией E_2 , т. е.

$$N_2 = n_2(W_{21} + w_{21}) = g_2(\rho_\nu B_{21} + A_{21}) \exp[-E_2/(kT)]. \quad (17.4)$$

Аналогично число индуцированных переходов из нижнего состояния в верхнее, определяющих поглощение энергии, равно

$$N_1 = n_1 W_{12} = g_1 \rho_\nu B_{12} \exp[-E_1/(kT)]. \quad (17.5)$$

Соотношение между коэффициентами A_{21} , B_{21} , B_{12} находится из условия термодинамического равновесия, при котором $N_1 = N_2$. Приравнявая выражения (17.4) и (17.5), можно определить спектральную плотность поля собственного (равновесного) излучения рассматриваемой равновесной системы

$$\rho_\nu = (A_{21}/B_{21}) / \{ [g_1 B_{12}/(g_2 B_{21})] \exp[-(E_1 - E_2)/(kT)] - 1 \}. \quad (17.6)$$

Если приравнять выражение (17.6) формуле Планка

$$\rho_\nu = (8\pi\nu^2/c^3)h\nu / \{ \exp[h\nu/(kT)] - 1 \} \quad (17.6, a)$$

(что справедливо для равновесной системы) и использовать частотное условие Бора $h\nu = E_2 - E_1$, то, сделав предположение о равенстве вероятностей индуцированного поглощения и излучения, т. е. $g_1 B_{12} = g_2 B_{21}$, получим соотношение для коэффициентов Эйнштейна для спонтанного и вынужденного излучения:

$$A_{21} = (8\pi\nu^2/c^3)h\nu B_{21}. \quad (17.7)$$

Вероятность излучательных переходов в единицу времени (с испусканием квантов спонтанного и вынужденного излучения) равна

$$W_{\text{изл}} = \{ [(8\pi\nu^2/c^3)h\nu] + \rho_\nu \} B_{21}. \quad (17.8)$$

Оценки показывают, что для СВЧ и оптического диапазонов $A_{21} \ll B_{21}$, т. е. вероятность спонтанного излучения много меньше, чем индуцированного, а поскольку спонтанное излучение определяет шумы, то в квантовых приборах роль шумов незначительна.

Необходимо отметить, что равновесное излучение всей системы частиц по отношению к каждой из частиц является внешним электромагнитным полем, стимулирующим поглощение или излучение частицей энергии в зависимости от ее состояния. Величина $8\pi\nu^2/c^3$, входящая в выражения (17.7) и (17.8), определяет число типов волн или колебаний в единичном объеме и в единичном интервале частот для области, размеры которой велики по сравнению с длиной волны $\lambda = c/\nu$.

Кроме индуцированных и спонтанных переходов в квантовых системах существенное значение имеют безызлучательные релаксационные переходы. Безызлучательные релаксационные переходы играют двойную роль: они приводят к дополнительному уширению спектральных линий (см. п. 17.3) и осуществляют установление термодинамического равновесия квантовой системы с ее окружением.

Релаксационные переходы происходят, как правило, вследствие теплового движения частиц. Поглощение тепла сопровождается переходами частиц на более высокий уровень и, наоборот, превращение энергии частицы в тепло происходит при переходе ее на более низкий уровень энергии. Таким образом, релаксационные переходы приводят к установлению вполне определенного для данной температуры равновесного распределения частиц по энергиям.

В реальных системах влиянием спонтанного излучения на естественную ширину спектральных линий можно пренебречь по сравнению с релаксационными процессами, которые более эффективно сокращают времена жизни возбужденных состояний, что и приводит к уширению спектральных линий (как это следует из соотношения неопределенностей для энергии—времени). Механизм этих процессов релаксации сильно зависит от конкретной системы. Например, для парамагнитных кристаллов, в частности в случае электронного парамагнитного резонанса, существенный вклад в уширение линий излучения вносят *спин-спиновые* и *спин-решеточные* взаимодействия и связанные с ними процессы релаксации с характерными временами соответственно порядка $10^{-1} \dots 10^{-3}$ с и $10^{-7} \dots 10^{-8}$ с.

Таким образом, релаксационные процессы, способствующие установлению теплового равновесия в среде, обеспечивают непрерывность процесса поглощения энергии внешнего электромагнитного излучения.

17.2. Усиление в квантовых системах

Если квантовая система находится в условиях термодинамического равновесия, то она поглощает энергию внешнего излучения, так как число переходов с нижних уровней на верхние (см. формулу (17.5)) в этих условиях превосходит число обратных переходов (см. формулу (17.4)). Для получения усиления света (электромагнитных волн) в среде необходимо, чтобы число переходов с излучением энергии превышало число переходов с поглощением энергии, а для этого необходимо нарушить термодинамическое равновесие.

Действительно, в соответствии с формулами (17.4) и (17.5) можно записать следующее соотношение для излучаемой и поглощаемой мощностей $P_{\text{изл}}$ и $P_{\text{погл}}$ на частоте ν :

$$\begin{aligned} P_{\text{изл}} - P_{\text{погл}} &= (n_2 W_{21} - n_1 W_{12}) h\nu = \\ &= \rho_\nu (B_{21} n_2 - n_1 B_{12}) h\nu = \rho_\nu B h\nu (n_2 - n_1). \end{aligned} \quad (17.9)$$

Здесь принято, что $B_{12} = B_{21} = B$ и $g_1 = g_2$, а также не учтена мощность спонтанного излучения, которое считается пренебрежимо малым. Тогда, как это следует из соотношений (17.9), для того чтобы излучаемая при вынужденных переходах мощность превосходила поглощаемую, т. е. для получения усиления света в квантовой системе необходимо, чтобы выполнялось условие $n_2 > n_1$, т. е. населенность верхнего уровня должна превышать населенность нижнего. Состояние системы, при котором населенность верхнего энергетического состояния превышает населенность нижнего, называется *состоянием с инверсной населенностью уровней* или *инверсией населенностей*. Среда (вещество), в которой реализуется инверсия населенностей и, соответственно, возможно усиление света, называется *активной средой* (активным веществом).

17.3. Основные элементы устройства квантовых генераторов

Активное вещество. До этого рассматривалась двухуровневая идеальная квантовая система, которая с достаточной степенью приближения реализуется в полупроводниковых лазерах. Большинство же активных веществ, используемых в приборах квантовой электроники, являются многоуровневыми системами. Однако, как правило, невзирая на очень большое количество уровней, участвующих в процессе генерации лазерного излучения,

большинство активных сред можно условно разделить на системы с тремя или четырьмя рабочими уровнями, так как в процессе создания инверсии населенностей и генерации излучения реально участвует три или четыре совокупности уровней энергии. Такие совокупности уровней могут образовывать энергетические полосы (например, уровни 3 на рис. 17.2, а, б). Таким образом, лишь условно можно называть активную среду трех- или четырехуровневой. Сначала рассмотрим условия получения инверсии населенности в трехуровневой квантовой системе (см. рис. 17.2, а). Вспомогательное излучение (излучение накачки) действует в диапазоне частот перехода ν_{13} и заставляет атомы или молекулы активного вещества переходить с уровня 1 на широкий уровень 3. Через некоторое время частицы из состояния 3 безызлучательно переходят (релаксационный переход) на метастабильный уровень 2, время жизни на котором больше времени жизни на третьем уровне, как правило, на несколько порядков. В результате на втором уровне происходит накопление частиц и реализуется инверсия населенностей ($n_2 > n_1$). Из решения кинетических уравнений для n_2 и n_1 следует, что для получения инверсии населенностей при минимальной величине накачки необходимо выполнение условий:

$$\tau_{31} \gg \tau_{32}; \tau_{21} \gg \tau_{32},$$

где τ_{31} , τ_{21} — времена спонтанных переходов и τ_{32} — время релаксации между соответствующими уровнями.

Первое неравенство означает, что частицы, перешедшие под воздействием накачки в состояние 3, должны в основном переходить на уровень 2, а не возвращаться на уровень 1. Согласно

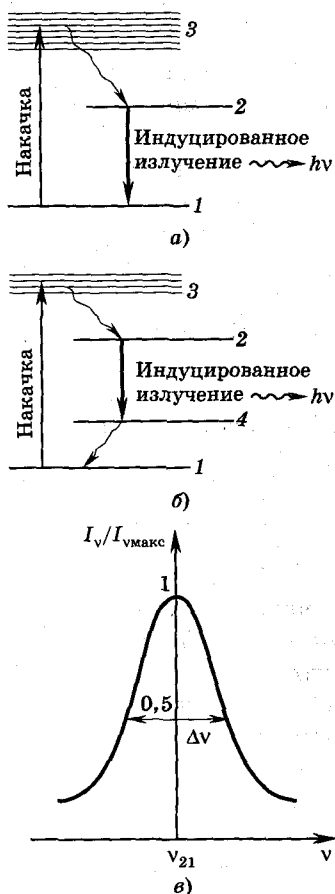


Рис. 17.2

второму соотношению, частицы на уровне 2 должны находиться много большее время, чем на уровне 3, что позволяет им накапливаться на верхнем уровне 2 (*метастабильном уровне*) излучательного перехода $2 \rightarrow 1$. Трехуровневая система обладает целым рядом недостатков, важнейшим из которых является большая мощность накачки из-за наличия кроме индуцированных большого числа спонтанных переходов $2 \rightarrow 1$.

Существенными преимуществами по сравнению с трехуровневой схемой обладают схемы с четырьмя энергетическими уровнями (см. рис. 17.2, б). Дополнительный четвертый уровень энергии по отношению к трехуровневой схеме располагается между вторым и первым уровнями. В четырехуровневой схеме для получения инверсии населенности возбужденные частицы с уровня 3 должны преимущественно переходить на уровень 2, т. е. выполняется условие $\tau_{32} \ll \tau_{31}, \tau_{34}$. Помимо этого, необходимо выполнение неравенства $\tau_{21} \gg \tau_{24}$. Это означает, что переход $2 \rightarrow 4$ является более вероятным по сравнению с переходом $2 \rightarrow 1$ и время жизни частиц на метастабильном уровне 2 значительно больше, чем на уровне 4, т. е. $\tau_{24} \gg \tau_{41}$.

При анализе процесса взаимодействия квантовой системы с внешним излучением предполагалось, что ширина энергетических уровней E_1 и E_2 равна нулю и усиление происходит только на одной частоте ν_{21} . Однако в действительности излучение реализуется в некоторой частотной полосе, что связано с уширением энергетических уровней. Излучаемая квантовой системой мощность максимальна на некоторой частоте $\nu_{21} = (E_2 - E_1)/h$ (соответствующей энергиям $E_{1,2}$, для которых плотность распределения частиц по энергиям максимальна) и убывает по обе стороны от этой частоты. Зависимость спектральной интенсивности I_ν излучения (поглощения) от частоты называют спектральной линией излучения (поглощения). Типичная зависимость $I_\nu = f(\nu)$ дана на рис. 17.2, в. Частотный интервал, на границах которого мощность излучения уменьшается до половины максимальной величины, называется шириной спектральной линии ($\Delta\nu$).

Очевидно, что величина $\Delta\nu$ определяется шириной энергетических уровней: $\Delta\nu = (\Delta E_2 + \Delta E_1)/h$, где $\Delta E_2, \Delta E_1$ — соответственно энергетическая ширина уровней 2 и 1. Наименьшая ширина спектральной линии называется *естественной шириной*. Она непосредственно определяется вероятностью спонтанного излучения и составляет величину от единиц Гц до единиц МГц. Сущест-

венное влияние на уширение спектральных линий оказывает эффект Доплера, поскольку излучающие атомы, молекулы, ионы движутся в газе или колеблются в кристаллической решетке твердого тела в различных направлениях со случайными скоростями. В результате пропорционально скорости перемещения указанных частиц относительно наблюдателя будет происходить смещение частоты. Поэтому наблюдаемая спектральная линия будет представлять собой огибающую суммы спектральных линий всех частиц рассматриваемой системы, причем для определенной частоты вклад в эту сумму пропорционален количеству частиц, обладающих данной резонансной частотой. Уширение, связанное с несовпадением резонансных частот различных микрочастиц, называют неоднородным. Неоднородное уширение наиболее существенно, если средний разброс резонансных частот микрочастиц больше ширины спектральной линии, определяемой однородным уширением, при котором линии каждого отдельного атома и системы в целом уширяются одинаково.

Сильное уширение спектральных линий вызывается также столкновениями частиц, насыщением населенности энергетических уровней и некоторыми другими факторами. Общая ширина спектральной линии в оптическом диапазоне может превышать несколько сотен МГц. В результате уширения энергия индуцированного излучения, получаемая за счет инверсии населенности, распределяется по всей ширине спектральной линии. Инверсия населенности в разных типах лазеров, как будет показано дальше, создается различными методами. Процесс создания инверсной населенности в квантовых системах называется *накачкой*. Квантовая система, способная усиливать электромагнитное излучение, принципиально должна содержать два основных элемента — активную среду и источник накачки, обеспечивающий инверсию населенностей. Если в активной среде с инверсной населенностью распространяется в направлении x плоская электромагнитная волна (рис. 17.3), то, как показывают расчеты,

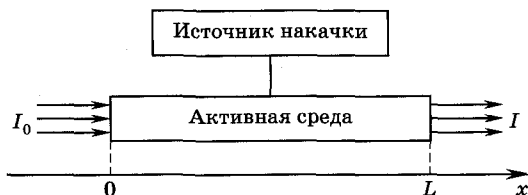


Рис. 17.3

ее интенсивность (плотность потока энергии) $I = \rho v$ (v — скорость электромагнитной волны) в любом сечении $x = \text{const}$ в пределах активной среды ($0 \leq x \leq L$) определяется соотношением

$$I = I_0 \exp \{ [B'_{21}(h\nu/\nu)(n_2 - g_2 n_1/g_1) - \beta]x \}, \quad (17.10)$$

где I_0 — интенсивность волны на входе активной среды (в сечении $x = 0$), β — постоянная, характеризующая величину потерь на единицу длины в активной среде, например за счет рассеяния на неоднородностях, дефектов кристаллов и т. д. Коэффициент $B'_{21} = B_{21}/(2\pi \Delta\nu)$, где $\Delta\nu$ — ширина контура спектральной линии, пропорционален коэффициенту Эйнштейна B_{21} . Параметр

$$\alpha = B'_{21}(h\nu/\nu)(n_2 - g_2 n_1/g_1) \quad (17.11)$$

определяет усиление волны на единицу длины. При наличии инверсной населенности, когда $n_2 > g_2 n_1/g_1$, интенсивность волны при распространении в среде экспоненциально возрастает. Из выражения (17.10) видно также, что при $x = L$, т. е. на выходе активной среды, интенсивность равна $I_0 e^{(\alpha - \beta)L}$, т. е. с увеличением инверсии населенностей усиление возрастает экспоненциально. Приведенные формулы (17.10) и (17.11) учитывают влияние на усиление факторов, связанных с шириной контура спектральной линии $\Delta\nu$, неодинаковой кратностью вырождения рабочих уровней активной среды g_1 и g_2 , а также потерями в системе.

Оптические резонаторы. Любой лазер представляет собой генератор электромагнитных колебаний. Колебания в любом генераторе поддерживаются за счет наличия положительной обратной связи. Отличительной особенностью квантовых генераторов является принципиальная возможность генерации в отсутствие специально введенной системы обратной связи. Дело в том, что положительная обратная связь содержится в самом явлении индуцированного излучения: электромагнитное поле, воздействующее на активную среду, вызывает индуцированное излучение, которое, в свою очередь, ведет к возрастанию поля, и, следовательно, ведет к дальнейшему росту интенсивности индуцированного излучения, и т. д.

Однако при индуцированном излучении инверсная населенность, вообще говоря, уменьшается, поэтому колебания в активной среде будут нарастать до предельной интенсивности, определяемой нелинейным процессом насыщения усиления.

Практически во всех квантовых приборах возбуждение колебаний происходит в резонансной системе, внутри которой размещается активная среда. В этом случае необходимая для генерирования колебаний положительная обратная связь обеспечивается за счет взаимодействия электромагнитного поля внутри резонатора с активной средой. Оптический резонатор обеспечивает многократное отражение волн, что увеличивает эффективную плотность электромагнитной энергии в активной среде, а следовательно, приводит к увеличению вероятности индуцированных переходов.

В оптическом диапазоне в качестве резонаторов используют отражатели (зеркала), между которыми располагается активный элемент. Основная особенность таких открытых оптических резонаторов заключается в том, что размеры их значительно больше длины волны, в результате чего оказывается возможным одновременное возбуждение большого числа собственных колебаний, что существенно ухудшает выходные параметры оптических генераторов.

Общие свойства оптических резонаторов наиболее наглядно можно оценить на примере простейшего резонатора — интерферометра Фабри—Перо, представляющего собой два плоских зеркала, расположенных на расстоянии L друг от друга.

При бесконечно протяженных зеркалах каждое собственное колебание резонатора будет образовываться в результате сложения плоских волн, движущихся в противоположных направлениях между зеркалами. В результате формируются волны, распространяющиеся вдоль оси резонатора, которые называются *продольными (аксиальными) модами* (типами колебаний), резонансные частоты которых определяются формулой

$$\omega_q = (q\pi - \Delta\phi)L^{-1}(c/n); q = 1, 2, 3, \dots, \text{ т. е. } \lambda_q = 2L/(q - \Delta\phi/\pi), \quad (17.12)$$

где c — скорость света в вакууме, n — показатель преломления активной среды, заполняющей резонатор, c/n — скорость распространения волны в активной среде, λ_q — длина волны q -го колебания, $\omega = 2\pi\nu$, q — номер типа колебаний, $\Delta\phi$ — изменение фазы при отражении волны. Собственные колебания отличаются друг от друга числом q полуволн, укладывающихся на длине резонатора. Частотный интервал между соседними типами продольных колебаний, т. е. теми, у которых номер q отличается на единицу, как следует из формулы (17.12), равен

$$\Delta\omega = \omega_q - \omega_{q-1} = \pi c/(Ln), \quad (17.13)$$

или в длинах волн

$$\Delta\lambda = \lambda/(q - \Delta\phi/\pi) \approx \lambda/q. \quad (17.14)$$

Поскольку длина резонатора L в лазерах имеет порядок нескольких см (5...50 см), легко подсчитать, что в оптическом диапазоне волн ($\lambda \approx 5 \cdot 10^{-5}$ см) число q будет очень большим ($\sim 10^5 - 10^6$), поэтому соседние типы колебаний располагаются близко друг к другу. Помимо рассмотренных продольных мод, существуют также типы колебаний, образованных плоскими волнами, распространяющимися под некоторым углом к оси резонатора. Однако для этих типов колебаний добротность оптического резонатора оказывается существенно меньшей.

Для случая конечных размеров зеркал, например для квадратных с размерами сторон, равными D , длина волны собственных колебаний (мод) в такой системе определяется выражением

$$1/\lambda = [(m_1/2D)^2 + (m_2/2D)^2 + (q/2L)^2]^{1/2}, \quad (17.15)$$

где m_1, m_2, q — целые числа, характеризующие структуру поперечных и продольных мод резонатора. Число q , относящееся к продольным модам, определяет число полуволн, которые укладываются на длине резонатора L (вдоль оси x). Числа m_1, m_2 относятся к поперечным модам и определяют число изменений направления поля вдоль поверхности зеркал, т. е. по осям y и z .

Одно из зеркал резонатора, а иногда и оба делаются частично пропускающими для вывода энергии из оптического генератора.

Если принять, что коэффициент отражения зеркал $r = 1$, то с учетом дифракционных явлений, обусловленных конечными размерами зеркал, их непараллельностью, шероховатостью поверхностей и т. д., добротность Q плоскопараллельного резонатора с достаточной точностью может быть записана в виде

$$Q = (1/2\beta_1)(2\pi L/\lambda), \quad (17.16)$$

где β_1 — затухание, определяющее суммарные потери для продольных видов колебаний в резонаторе.

Последняя формула показывает, что добротность для всех продольных типов колебаний в таких резонаторах практически одинакова. Относительная ширина резонансной линии для любого продольного вида колебаний номера q на уровне половины мощности определяется известной формулой

$$(\Delta\nu/\nu)_q = (\Delta\omega/\omega)_q = (\Delta\lambda/\lambda)_q = 1/Q = 2\beta(\lambda/2\pi L). \quad (17.17)$$

Помимо плоскопараллельных резонаторов (типа Фабри—Перо), в лазерах достаточно широко используются конфокальные, сферические резонаторы, а также полусферические резонаторы, образованные плоским и сферическим зеркалами. Каждый тип резонатора имеет свои достоинства и недостатки.

Резонаторы Фабри—Перо требуют повышенной точности установки зеркал относительно друг друга (параллельность зеркал, перпендикулярность их к оси, определенное расстояние между зеркалами и строгая выдержка их поперечных размеров, а также ряд других требований).

Конфокальным называют резонатор, образованный двумя одинаковыми сферическими отражателями, расположенными друг от друга на расстоянии, равном радиусу кривизны зеркала. Поскольку фокусное расстояние сферического зеркала составляет половину радиуса кривизны, то фокусы двух зеркал совпадают. Расчеты показывают, что в конфокальном резонаторе потери существенно ниже потерь в случае плоских зеркал. Это объясняется тем, что поле в конфокальном резонаторе сильно сконцентрировано у оси резонатора и поэтому быстрее спадает к краям зеркал, что уменьшает выход излучения во внешнее пространство. Кроме того, потери энергии для различных типов колебаний, особенно поперечных, могут сильно различаться между собой. Так, в конфокальном резонаторе, в отличие от плоскопараллельного, имеется принципиальная возможность подавления высших типов колебаний.

Сферические резонаторы имеют, как правило, расстояние между зеркалами большее по сравнению с конфокальными. Характеристики сферических резонаторов хуже, чем конфокальных. Однако такая конфигурация зеркал представляет большой интерес с точки зрения селекции видов колебаний и, кроме того, они не требуют столь высокой точности юстировки и технологически более просты.

На практике очень часто используются полусферические резонаторы, в которых, как упоминалось выше, одно зеркало является плоским, а второе — сферическим. Добротность полусферического резонатора, как показывают расчеты, определяется главным образом сферическим зеркалом, поскольку размер пятна на нем больше, чем на плоском зеркале. Эта система характеризуется низкими потерями и малой критичностью к изменению расстояния между зеркалами.

В незаполненных резонаторах спектр собственных колебаний полностью определяется конечными размерами зеркал и их конфигурацией. При наличии в резонаторе активного вещества, особенно твердотельного диэлектрика, образование колебаний следует рассматривать как результат многократных отражений от зеркал различных типов поверхностных волн, распространяющихся вдоль диэлектрического волновода. При этом надо учитывать и дифракционные явления на зеркалах, которые приводят к изменению структуры поверхностных волн при каждом отражении.

17.4. Условия баланса мощности и фаз в лазерах (оптических квантовых генераторах)

Поле в резонаторе можно представить как сумму волн, распространяющихся в противоположных направлениях и отражающихся от зеркал. Если помимо коэффициентов отражения от зеркал r_1 и r_2 учесть создаваемые ими дифракционные потери ξ , потери на единицу длины активной среды β , то условие равновесия в квантовом генераторе можно записать в виде

$$r_1 r_2 (1 - \xi)^2 e^{-2\beta L} e^{2\alpha L} = 1, \quad (17.18)$$

где L — расстояние между зеркалами резонатора, α — коэффициент усиления на единицу длины в активной среде.

Последнее выражение является условием баланса амплитуд (мощностей) для существования стационарного режима генерации, при котором индуцированное излучение компенсирует суммарные потери в системе. Условие баланса фаз представляется в следующем виде:

$$\Delta\varphi_1 + \Delta\varphi_2 + (2\pi/\lambda)2L = 2\pi q \quad (q = 0, 1, 2, 3, \dots), \quad (17.19)$$

где $\Delta\varphi_1$, $\Delta\varphi_2$ — изменения фазы при отражении от зеркал резонатора.

Условие баланса фаз в формуле (17.19) означает, что в установленном режиме генерации волна, пройдя путь $2L$ в резонаторе и дважды отразившись от зеркал, получает фазовый сдвиг $2\pi q$, кратный целому числу периодов. Тем самым в генераторе обеспечивается положительная обратная связь, благодаря которой непрерывно воспроизводятся колебания неизменной частоты (образуются стоячие волны). Выражение (17.19) по смыслу совпадает с (17.12).

В формуле (17.18) (баланс мощности) экспоненциальный член $e^{2\alpha L}$ описывает индуцированное усиление, а остальные множители в левой части равенства характеризуют потери. Отсюда следует, что стационарный режим генерации лазера установится лишь в том случае, если индуцированное усиление в активной среде компенсирует все потери в резонаторе. Таким образом, из формулы (17.18) следует условие стационарного режима генерации, сводящееся к тому, чтобы коэффициент усиления α имел определенное стационарное значение $\alpha_{\text{ст}}$

$$\begin{aligned}\alpha_{\text{ст}} &= \beta - (1/L) \ln(1 - \xi) - (1/2L) \ln r_1 r_2 = \\ &= \beta + (1/L) \ln \frac{1}{(1 - \xi)\sqrt{r_1 r_2}}.\end{aligned}\quad (17.20)$$

Величина $\alpha_{\text{ст}}$, как следует из формулы (17.11), реализуется при некоторой пороговой населенности $n_{2\text{пор}}$ верхнего уровня рабочего лазерного перехода.

17.5. Спектр и характеристики излучения квантовых генераторов

Замечательные свойства лазерного излучения обусловлены как свойствами вынужденного излучения, так и использованием оптических резонаторов, что позволяет получить когерентное оптическое излучение, характеристиками которого можно управлять.

Спектр излучения. Количество типов колебаний, которые могут возникнуть в резонаторе, определяется спектральным контуром излучения активного вещества (рис. 17.4, а) и резонансными свойствами интерферометра (резонатора) (рис. 17.4, б). На рис. 17.4, б показаны только аксиальные моды колебаний; интенсивность излучения представлена в относительных единицах и обозначена как $I_{\text{отн}}$.

В случае когда инверсия населенностей очень мала и ни на одной из частот резонатора не выполняется условие самовозбуждения, в системе возможно лишь спонтанное излучение, спектр которого определяется в основном формой контура спектральной линии активной среды. С увеличением инверсии населенностей на рабочем переходе усиление в активной среде возрастает и условие самовозбуждения (баланса амплитуд) начинает сначала выполняться для одной или нескольких аксиальных мод колебаний,

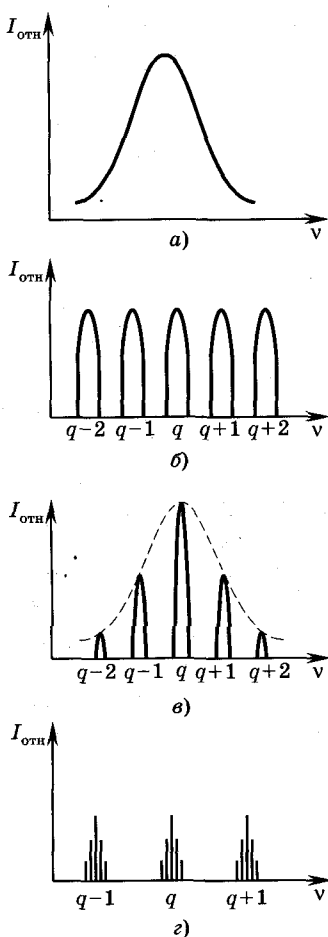


Рис. 17.4

имеющих, как правило, частоты, ближайшие к максимуму спектрального контура излучения активной среды (см. рис. 17.4, а). В этом случае на фоне спонтанного излучения будут выделяться интенсивные колебания на отдельных частотах.

При дальнейшем росте инверсии населенностей усиление в активной среде на всех частотах резонатора увеличивается и условие самовозбуждения начинает выполняться для все большего числа мод колебаний резонатора, т. е. спектр генерации лазера расширяется (рис. 17.4, б). В диапазоне СВЧ ширина спектральной линии излучения достаточно мала и условие баланса амплитуд выполняется, как правило, лишь для одной частоты. В оптическом же диапазоне ширина спектральной линии превышает несколько сотен МГц, и в ее пределах размещаются частоты значительного числа как аксиальных (несколько десятков), так и поперечных типов колебаний (формула (17.15)), частоты которых группируются около аксиальных мод (рис. 17.4, в). Ширина диапазона частот, в котором распределяются поперечные типы колебаний, составляет обычно единицы МГц.

Анализ спектра излучения лазера показывает, что на структуру спектра существенное влияние оказывает тип активной среды, характер уширения спектральной линии и пространственная модуляция инверсной населенности.

Наличие многих частот в спектре лазерного излучения часто является существенным недостатком. Например, в связанной линии оптического диапазона это может привести к появлению в приемнике ряда комбинационных частот, которые представляют помеху для сигнала. Сужение спектра излучения принципиально может быть достигнуто различными путями: снижением интенсивности накачки до величины, близкой к порогу воз-

буждения генератора (ситуация, подобная изображенной на рис. 17.4, в); путем увеличения потерь для нежелательных типов колебаний до величин, при которых условие порога самовозбуждения (баланса амплитуд) не выполняется; использованием фокусирующих линз и диафрагм в резонаторе (селекция частот резонатора); применением полусферического резонатора, состоящего из плоского и сферического зеркал, которые располагаются на расстоянии, приблизительно равном радиусу кривизны сферического зеркала (этим методом селективируются поперечные типы колебаний). Применение неустойчивых резонаторов, в которых поле вне их внутренней полости зеркалами не концентрируется, а рассеивается, расположение диспергирующих призм между плоскими зеркалами и т. д. также позволяют устранить нежелательные типы колебаний.

Направленность и поляризация излучения. Характерной чертой лазеров является высокая *направленность пучка (малая угловая расходимость) излучения*, что обусловлено как самим механизмом лазерной генерации, так и большими характерными размерами излучающей системы по сравнению с длиной волны излучения λ . Предельно достижимая расходимость пучка определяется дифракционными явлениями на зеркалах или пятне, через которое выводится излучение, в этом случае угловая ширина диаграммы направленности по уровню половинной мощности ($\theta_{0,5}$) определяется формулой $\theta_{0,5} \approx 1,22\lambda/D$, где D — апертура пучка на выходном зеркале.

Типичные значения величины $\theta_{0,5}$ для многих типов лазеров заключены в диапазоне от единиц до нескольких десятков угловых минут, что превышает предельные теоретические значения в 20...50 раз. Таким образом, для лазеров направленность излучения является весьма высокой, что позволяет получить высокое угловое разрешение в оптических локаляторах и высокую угловую точность в системах наведения по лучу, обеспечить хорошую скрытность при передаче информации, повысить дальность действия систем связи, локации и управления без увеличения излучаемой мощности.

Увеличение расходимости луча реального лазера по сравнению с предельными теоретическими значениями обусловлено следующими причинами: сферичностью волны на выходе ОКГ, многомодовостью излучения, наличием неоднородностей в активной среде и несовершенством элементов резонатора. Все эти

причины в основном вызывают искажения распределения поля на зеркалах резонатора. Искажения диаграммы направленности излучения сводятся к появлению нескольких максимумов, их смещению относительно центральной оси всей диаграммы. Влияние оптических неоднородностей наиболее сильно проявляется в полупроводниковых, в меньшей мере — в твердотельных, и почти отсутствует в газовых лазерах.

Уменьшение расходимости излучения на выходе лазеров достигается в первую очередь путем устранения высших поперечных типов колебаний. Наибольшую направленность имеют одномодовые лазеры, в которых возбуждается только один поперечный тип колебаний TEM_{00} . Проблема обеспечения одномодового режима в газовых лазерах небольшой мощности решена. Для твердотельных лазеров, активная среда которых обладает большим усилением, обеспечение одномодового режима достаточно сложно, так как трудно создать значительные потери для высших типов колебаний, не ухудшая энергетических характеристик лазера.

Важным параметром лазеров является *поляризация излучения*, под которой понимается преимущественное направление вектора электрического поля. В лазерах поляризация излучения определяется величиной усиления в резонаторе для волн различной поляризации, что обуславливается или самой активной средой, или введенными в резонатор поляризующими элементами. Когда поляризующие элементы (устройства) отсутствуют и активная среда изотропна, условия возникновения колебаний для всех плоскостей поляризации одинаковы, т. е. излучение оказывается неполяризованным. В большинстве случаев излучение лазеров оказывается плоскополяризованным, что часто связано с мерами по уменьшению внутренних паразитных потерь на отражение. В резонаторе для уменьшения паразитных отражений граничные поверхности активной среды выполняются не перпендикулярными к оси резонатора, а наклонными.

Зависимость коэффициента отражения плоскопараллельной стеклянной пластинки от угла падения луча для различной ориентации вектора напряженности поля показывает, что отражение для излучения, поляризованного в плоскости падения при некотором угле φ_0 (угле Брюстера), близко к нулю. Это свойство широко используется в лазерах. В газовых лазерах окна газоразрядной кюветы выполняются под углом Брюстера. То же самое можно сказать и о твердотельных лазерах, где грани стержней активного вещества также часто скашивают под углом Брюстера. Такое вы-

полнение элементов активной среды определяет поляризацию излучения. Для излучения, поляризованного в плоскости падения, потери в резонаторе минимальны, что обычно и реализуется в большинстве лазеров.

Поляризация излучения может определяться и самой активной средой, особенно кристаллической. Как правило, усиление в кристалле зависит от угла между его оптической осью и плоскостью поляризации света. Так, например, в рубине наибольшая вероятность индуцированных переходов и наибольшее усиление наблюдается для излучения, плоскость поляризации которого перпендикулярна оптической оси, т. е. когда оптическая ось кристалла перпендикулярна оси резонатора. Излучение лазера в этом случае будет поляризованным в плоскости перпендикулярной оптической оси кристалла. Для излучения, поляризованного в этой плоскости, условие самовозбуждения выполняется в первую очередь. Именно это излучение и возникает в лазере. В оптических линиях связи использование поляризационной модуляции позволяет увеличить, по сравнению с амплитудной, дальность действия системы за счет уменьшения потерь света при модуляции.



Контрольные вопросы

1. Что такое инверсия населенности и каковы методы ее получения?
2. Каковы свойства индуцированного и спонтанного излучений?
3. Каковы особенности создания инверсной населенности в трех- и четырехуровневых квантовых системах?
4. Назвать и объяснить причины уширения контура спектральных линий.
5. Каковы механизмы усиления волн в квантовых системах?
6. Активные среды и их характеристики.
7. Как формируется спектр собственных колебаний резонатора?
8. Типы резонаторов и их параметры.
9. Спектр излучения лазера и свойства излучения.
10. Условия генерации в лазерах.
11. Чем определяется направленность и поляризация излучения в лазерах?
12. Виды переходов в квантовых системах и их свойства.

Глава 18

КВАНТОВЫЕ ПРИБОРЫ СВЧ-ДИАПАЗОНА (МАЗЕРЫ)

**18.1. Активное вещество твердотельных мазеров.
Парамагнитные уровни энергии**

В 1954—1955 гг. Н. Г. Басову и А. М. Прохорову в СССР и Ч. Таунсу, Дж. Гордону и Х. Цайгеру в США удалось независимо друг от друга осуществить усиление и генерацию электромагнитных волн на частоте 23 870 МГц (СВЧ-диапазон) за счет использования инвертированных состояний колебательных уровней пучка молекул аммиака. Американцы назвали это устройство «мазер», что является аббревиатурой английского словосочетания: «*microwave amplification by stimulated emission of radiation*»¹. Мазер на пучке молекул аммиака, отличающийся строгим постоянством частоты, вскоре нашел важное применение в качестве сверхстабильного стандарта частоты (атомные часы). В 1956 г. Н. Бломбергенем (США) был предложен работающий в непрерывном режиме трехуровневый *твердотельный мазер*, в котором используется электронный *парамагнитный резонанс*. В 1957 г. впервые был создан действующий СВЧ-усилитель с рекордно малым уровнем шумов. Кроме того, созданный мазер имел все достоинства перестраиваемого СВЧ-усилителя, работающего в непрерывном режиме и усиливающего сигналы в широкой полосе частот.

Основной особенностью парамагнитных материалов является наличие атомов или ионов, обладающих постоянным магнитным моментом. Это свойство присуще только некоторым типам атомов, в основном атомам переходной группы периодической системы элементов, которые по своей природе являются парамагнитными, что обусловлено преимущественно нескомпенсированным спиновым магнитным моментом электронов.

Магнитные свойства атомов. Объяснение магнитных свойств веществ с классических позиций базируется на планетарной модели атомов, в соответствии с которой электрон вращается

¹ Английская аббревиатура соответствующего словосочетания, в котором «*microwave*» (микроволновое) заменяется на «*light*» (свет), есть «*laser*» (лазер).

вокруг ядра по орбите радиуса r и обладает **орбитальным моментом импульса** (количества движения)

$$\vec{l} = mr^2\vec{\omega}_e, \quad (18.1)$$

где m — масса электрона, $\vec{\omega}_e$ — угловая скорость его движения по орбите (псевдовектор, направлен вдоль оси вращения).

Движущийся по круговой орбите электрон образует замкнутый круговой ток, который представляет собой магнитный диполь с моментом

$$\vec{\mu}_l = -\frac{er^2\vec{\omega}_e}{2}. \quad (18.2)$$

Минус в (18.2) обусловлен отрицательным зарядом электрона и указывает на то, что магнитный момент противоположен по направлению вектору момента импульса. Из выражений (18.1) и (18.2) следует связь между орбитальным механическим и магнитным моментами электрона:

$$\vec{\mu}_l = -\frac{e}{2m}\vec{l}.$$

В рамках классических представлений взаимодействием магнитных моментов электронов атома с внешним магнитным полем можно объяснить парамагнитные свойства вещества. Количественное описание этих эффектов дает квантовая механика, в соответствии с основными законами которой механический момент электрона квантуется, т. е. может принимать лишь дискретный ряд значений, пропорциональных *постоянной Планка* $\hbar = h/2\pi$:

$$|\vec{l}| = \sqrt{l(l+1)}\hbar.$$

Здесь l — *орбитальное квантовое число*, которое может принимать только целые значения (в том числе и ноль). При достаточно больших значениях $l \gg 1$ приближенная *формула квантования момента импульса* имеет вид

$$|\vec{l}| \approx l\hbar.$$

Таким образом, физический смысл орбитального квантового числа l состоит в том, что оно приближенно определяет модуль момента импульса орбитального движения электрона в единицах \hbar .

Соответственно квантуется и орбитальный магнитный момент:

$$|\vec{\mu}_L| = \frac{e\hbar}{2m} \sqrt{l(l+1)} = \beta \sqrt{l(l+1)}, \quad (18.3)$$

где $\beta = \frac{e\hbar}{2m} = 9,27 \cdot 10^{-24} \text{ А} \cdot \text{м}^2$ — магнетон Бора.

Для объяснения целого ряда экспериментальных фактов квантовая физика учитывает наряду с орбитальным также собственный, или спиновый момент электрона (спин).

Спиновый магнитный момент электрона, определяемый его спином, в соответствии с квантовой теорией равен

$$|\vec{\mu}_S| = 2\beta s, \quad (18.4)$$

где $s = 1/2$ — спиновое квантовое число для электрона.

Для атома в целом справедливы соотношения между суммарными орбитальными и спиновыми моментами импульса \vec{L} и \vec{S} и соответствующими магнитными моментами $\vec{\mu}_L$ и $\vec{\mu}_S$, как и для электрона¹, т. е.

$$\vec{\mu}_L = -\beta\vec{L}, \quad \vec{\mu}_S = -2\beta\vec{S}. \quad (18.5)$$

Ориентацию векторов \vec{L} и \vec{S} для атома нельзя рассматривать независимо, поскольку в сумме они должны всегда давать полный момент импульса атома $\vec{J} = \vec{L} + \vec{S}$, который квантуется определенным образом. Аналогично векторы $\vec{\mu}_L$ и $\vec{\mu}_S$ связаны друг с другом таким образом, что в сумме получается полный магнитный момент атома $\vec{\mu}_J$. Эта связь приводит к следующему соотношению между полным магнитным моментом атома $\vec{\mu}_J$ и полным моментом импульса \vec{J} :

$$\vec{\mu}_J = -g_J\beta\vec{J}, \quad (18.6)$$

где g_J (g -фактор Ланде) принимает значения от нуля до двух.

Если преобладает орбитальный момент, т. е. $\mu_J \approx \mu_L$, то $g_J \approx 1$, в противоположном случае, т. е. при $\mu_J \approx \mu_S$, фактор $g_J \approx 2$.

Зеемановское расщепление уровней. Хорошо известно, что если магнитный диполь, в рассматриваемом случае это атом, поместить в постоянное магнитное поле, то энергия диполя будет

¹ Сложение орбитальных и спиновых моментов электрона осуществляется в соответствии с правилами квантовой механики и зависит от типа связи (см., например, [37], § 8).

зависеть от его ориентации во внешнем магнитном поле. Назовем эту энергию ориентационной. Ориентационная энергия магнитного диполя в магнитном поле с индукцией \vec{B} равна

$$E_M = |\vec{\mu}_J \cdot \vec{B}| = \mu_J B |\cos \theta| = \mu_z B, \quad (18.7)$$

где θ — угол между векторами $\vec{\mu}_J$ и \vec{B} .

Вектор \vec{J} , а значит, и вектор $\vec{\mu}_J$ в соответствии с квантовой теорией могут иметь только некоторые разрешенные (по отношению к вектору \vec{B}) ориентации, определяемые *магнитным квантовым числом* M_J , принимающим $2J + 1$ значений, где J — *квантовое число полного момента импульса атома*. В силу этого ориентационная энергия E_M может принимать тоже только соответствующие дискретные значения (всего их $2J + 1$). Если принять, что вектор \vec{B} направлен вдоль оси Z и учесть выражение (18.6), то *ориентационную энергию* можно записать следующим образом:

$$E_M = \mu_z B = M_J g_J \beta B. \quad (18.8)$$

Соотношение (18.8) определяет разрешенные уровни энергии атома в магнитном поле. В отсутствие магнитного поля ($B = 0$) из (18.8) следует, что $E_M = 0$, т. е. все квантовые состояния, определяемые различными ориентациями вектора $\vec{\mu}_J$, *вырождены*, т. е. любому разрешенному значению μ_z соответствует одна и та же энергия.

Наложение магнитного поля приводит к *снятию вырождения* этих квантовых состояний. Это означает, что в соответствии с (18.8) каждому значению квантового числа M_J , определяющему разрешенную ориентацию магнитного момента атома, соотносится свое определенное значение энергии E_M , т. е. состояния с разными ориентациями вектора $\vec{\mu}_J$ имеют разные ориентационные энергии. Следовательно, в магнитном поле происходит расщепление уровней парамагнитных атомов. Такое расщепление уровня, соответствующего квантовому числу J , на $2J + 1$ уровней называется *зеemanовским*, а получающиеся уровни *зеemanовскими*, или *парамагнитными энергетическими уровнями*, разность энергий между которыми прямо пропорциональна магнитной индукции (см. формулу (18.8) и рис. 18.1, а).

Энергетические уровни парамагнитного кристалла. В электрическом поле кристаллической решетки твердого тела или в сильных внешних электрических полях происходит также рас-

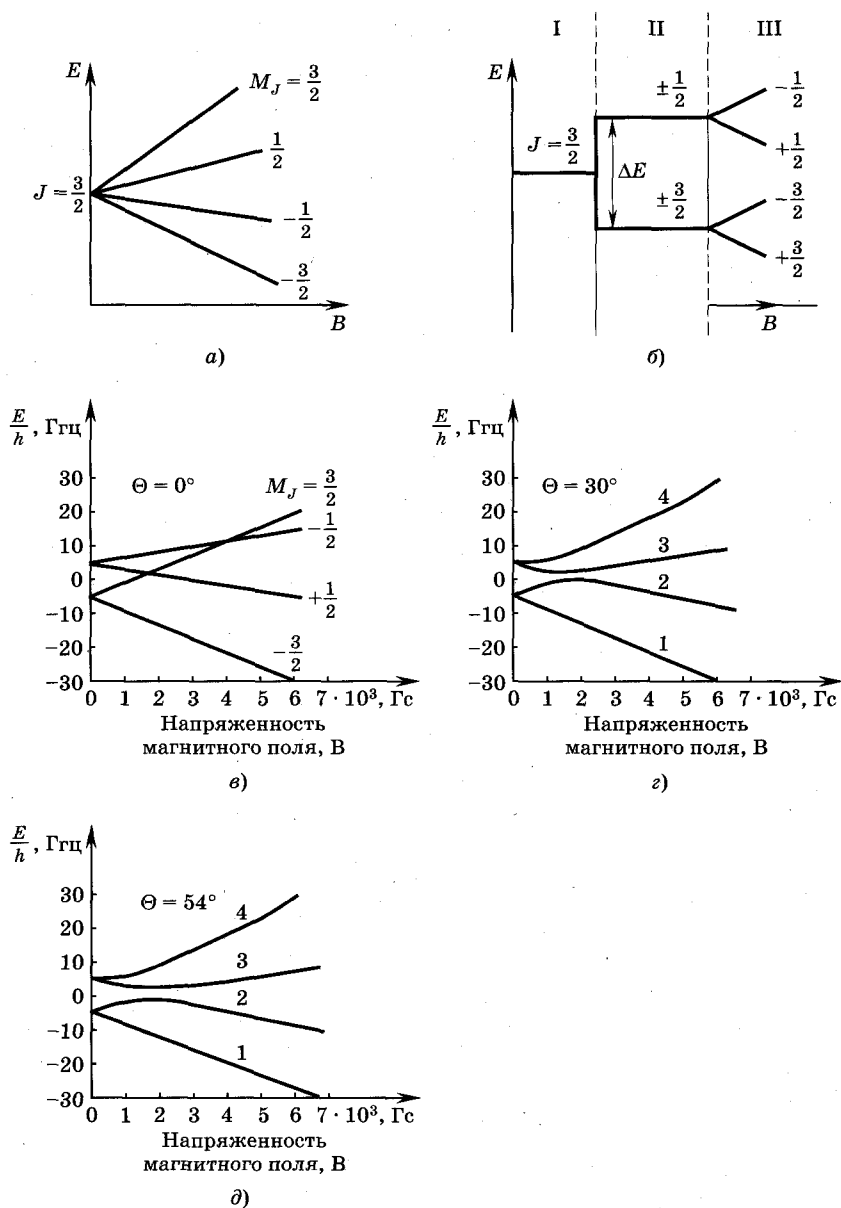


Рис. 18.1

щепление энергетических уровней атомов, которое называется **штарковским**.

В твердотельных лазерах в качестве активного вещества широко используется рубин, который представляет окись алюминия Al_2O_3 с небольшой примесью парамагнитных ионов трижды ионизированного хрома Cr^{3+} , замещающих в кристаллической решетке часть ионов алюминия. Ион Cr^{3+} имеет электронную конфигурацию $3d^3$, которой соответствует основной уровень ${}^4F_{3/2}$, имеющий квантовые числа $S = 3/2$, $L = 3$, $J = 3/2$ свободного иона¹. Свободный ион и ион в кристаллической решетке имеют $2J + 1 = 4$ зеемановских уровней. Однако особенностью иона хрома в решетке является то, что этим четырем уровням соответствуют $2S + 1$ спиновых состояний, поскольку орбитальный момент импульса \vec{L} в решетке сильно подавлен. Это хорошо иллюстрируется тем фактом, что в решетке рубина $g \approx 2$, т. е. фактор Ланде изотропен, в то время как для свободного иона $g \approx 0,4$. Свободный ион хрома обладает системой зеемановских уровней, подобной показанной на рис. 18.1, а. Однако ион Cr^{3+} в рубине имеет более сложную систему зеемановских уровней, характеризующуюся наличием нулевого расщепления ($B = 0$) за счет электрического поля кристаллической решетки и сильной анизотропией, определяемой свойствами решетки рубина.

Расщепленные в электрическом поле решетки уровни иона хрома отличаются друг от друга только абсолютной величиной проекции магнитного момента на ось кристалла и являются двукратно вырожденными, что иллюстрирует рис. 18.1, б, где участок I соответствует свободному иону, II — иону в кристаллической решетке, на участке III показаны энергетические зеемановские уровни при наличии магнитного поля с индукцией B .

Таким образом, при $B = 0$ уровни Cr^{3+} в рубине разделены на два подуровня, у которых $M_J = \pm 3/2$ и $M_J = \pm 1/2$, разность энергий этих состояний равна $\Delta E = h \times 11,4 \text{ ГГц} \approx 2,6 \cdot 10^{-6} \text{ эВ}$ ($1 \text{ Гц} \Rightarrow \Rightarrow 2,24 \cdot 10^{-16} \text{ эВ}$, $1 \text{ ГГц} \Rightarrow 2,24 \cdot 10^{-7} \text{ эВ}$).

На рис. 18.1, в, г и д показано изменение энергетической структуры иона хрома при различной ориентации вектора магнитного поля относительно кристаллографической оси рубина. Если угол θ между направлением оси и вектором \vec{B} больше нуля,

¹ Здесь и в дальнейшем для описания квантовых состояний атомов (ионов) используются данные, приведенные в [37].

то энергетические уровни не пересекаются и все запрещенные (квантовыми «правилами отбора») переходы становятся разрешенными.

При $\theta = 54^\circ 44'$ энергетические уровни должны быть симметрично расположенными относительно проведенной между ними средней линии (см. рис. 18.1, ∂), что используется в некоторых схемах двойной накачки мазеров, поэтому угол $\theta = 54^\circ 44'$ называется *углом двойной накачки* (см. п. 18.2).

Разность энергий двух соседних зеемановских уровней можно оценить, воспользовавшись формулой Планка и соотношением (18.8):

$$h\nu = hf = \Delta E_{M+1, M} = E_{M+1} - E_M = g_J \beta B, \quad (18.9)$$

откуда следует, что частота перехода f дается формулой

$$f = \frac{g_J \beta}{h} B = \gamma B, \quad (18.10)$$

где $\gamma = \frac{g_J \beta}{h}$.

Поскольку β и h — константы, то при $g_J = 2$ соотношение (18.10) можно записать в виде

$$\frac{f}{B} = \frac{2\beta}{h} = 2,8 \text{ МГц/Гс.}$$

При $B \geq 10^2$ Гс частота перехода лежит в СВЧ-диапазоне (на СВЧ принято частоту обозначать буквой « f »). Рабочий диапазон мазеров требует напряженности магнитного поля более 1000 Гс.

Явление перехода ионов хрома с нижних парамагнитных уровней на верхние, происходящее под влиянием внешнего СВЧ электрического поля и сопровождающееся поглощением его энергии, называется *электронным парамагнитным резонансом*.

18.2. Квантовые парамагнитные усилители

Усиление электромагнитной волны в квантовых приборах СВЧ-диапазона реализуется при взаимодействии ее с частицами активного парамагнитного вещества, в котором в результате накачки создана инверсная населенность. Для эффективного взаимодействия волны со средой необходимо увеличивать плотность потока индуцирующего излучения, поскольку вероятность вынужден-

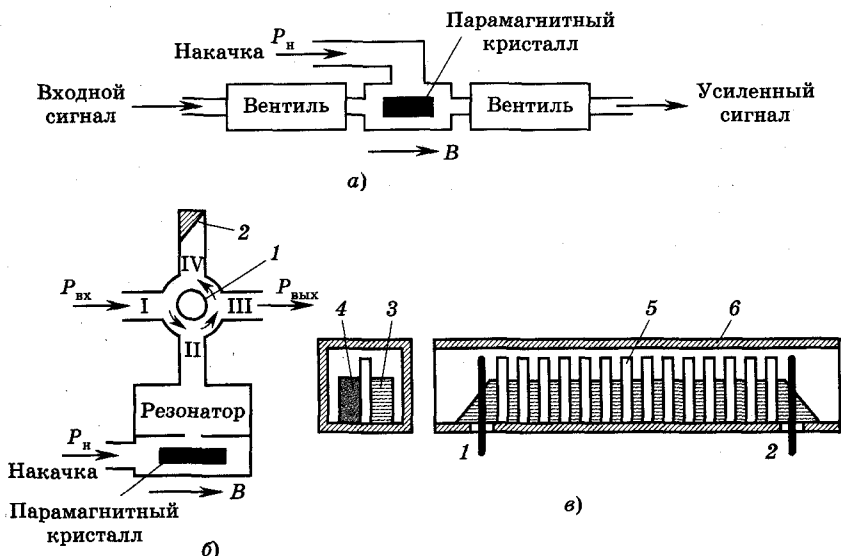


Рис. 18.2

ных переходов прямо пропорционально спектральной плотности потока излучения (см. п. 17.1). В квантовых парамагнитных усилителях (КПУ) плотность потока индуцированного излучения увеличивается или за счет многократного прохождения волн через среду (резонаторные усилители, рис. 18.2, а, б), или вследствие замедления волны (усилители бегущей волны, рис. 18.2, в).

18.2.1. Резонаторные квантовые парамагнитные усилители (РКПУ).

Различают несколько типов РКПУ: усилители проходного типа (см. рис. 18.2, а), отражательного (циркуляторного) типа и много-резонаторные, как правило, циркуляторного типа (рис. 18.2, б). Основные элементы перечисленных усилителей практически одинаковы. Устройство проходных и отражательных усилителей различается способами разделения входных и выходных трактов.

Для устранения самовозбуждения в проходном КПУ на входе и выходе включаются вентильные элементы. В КПУ циркуляторного (отражательного) типа входной и выходной сигналы проходят по одному и тому же волноводу (или коаксиальной линии). Для развязки входной и выходной цепей в этом случае используется циркулятор 1 (см. рис. 18.2, б). Входное плечо I циркулятора подключается к антенне; сигнал, пройдя первое плечо

чо, поступает во второе, а затем в резонатор. Усиленный за счет индуцированного излучения в парамагнитном кристалле сигнал возвращается через тот же волновод в циркулятор и направляется в плечо III и далее к нагрузке. Такие направления входных и выходных сигналов обеспечиваются за счет **невзаимного элемента** циркулятора, представляющего чаще всего феррит, находящийся во внешнем магнитном поле. Это устройство с малыми потерями пропускает только сигналы определенной поляризации и направления, а поскольку входные и выходные сигналы в плече II имеют противоположные направления и поляризацию, то они распространяются указанным образом. Шумы, отраженная неоднородностями часть мощности сигналов и другие паразитные колебания, попадающие с выхода по плечу III в циркулятор, поглощаются в согласованной нагрузке 2.

Основным элементом КПУ является резонатор с активным веществом (парамагнитным кристаллом), одна из мод которого настроена на частоту сигнала, а другая — на частоту накачки.

Инверсия населенностей создается за счет облучения кристалла электромагнитной волной накачки, имеющей частоту, большую частоты сигнала.

На рис. 18.2 не указаны криогенные системы охлаждения. Без них КПУ не может работать из-за наличия спин-спиновой и спин-решеточной релаксации, которые непрерывно стремятся восстановить тепловое равновесие спиновой системы с кристаллической решеткой (спиновой системой называется в этих условиях ансамбль из возбужденных парамагнитных ионов хрома). Магнитная система, необходимая для получения зеемановских уровней, обозначена на рисунке вектором \vec{B} .

Рассмотрим более подробно процессы в парамагнитном кристалле, связанные с получением инверсии населенностей и усилением сигнала.

На рис. 18.3, *а*, *б*, которые аналогичны по смыслу схемам, показанным на рис. 18.1, изображены энергетические диаграммы иона хрома в решетке рубина. В отсутствие накачки и сигнала распределение населенности всех уровней будет равновесным и подчиняется закону Больцмана (штриховая экспоненциальная (падающая) кривая на рис. 18.3, *б*), согласно которому

$$n_i/n_j = e^{-(E_i - E_j)/(kT)}$$
. В отличие от «орбитальных» энергетических уровней разность энергий ($E_i - E_j$) между парамагнитными уровнями обычно меньше характерной величины тепловой энергии kT .

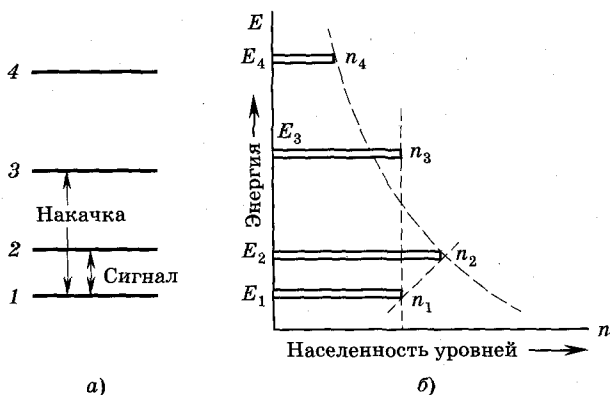


Рис. 18.3

Будем считать, что квант электромагнитной волны накачки, облучающей кристалл рубина, имеет энергию $hf = hf_{13}$, где f_{13} — частота перехода с уровня 1 на уровень 3.

Под действием энергии накачки ионы хрома возбуждаются и переходят с уровня 1 на уровень 3. Будут и обратные переходы, однако на начальном этапе населенность уровня 3 растет, а уровня 1 уменьшается из-за его большей заселенности в равновесном состоянии. Этот процесс продолжается до тех пор, пока населенности этих уровней не сравняются ($n_1 = n_3$), как показано на рис. 18.3, б вертикальной штриховой прямой. При этом в обоих направлениях будет происходить примерно одинаковое число переходов, так что, пока на систему действует достаточно сильное излучение накачки, она будет удерживаться в состоянии, далеком от теплового равновесия, при этом говорят, что переход $1 \rightarrow 3$ насыщен.

В процессе насыщения перехода $1 \rightarrow 3$ населенности уровней 2 и 4 могут несколько измениться, так как изменение населенностей уровней 1 и 3 влияет на релаксационные процессы, в которых участвуют уровни 2 и 4. Однако можно считать, что эти изменения населенностей уровней 2 и 4 незначительны и ими можно пренебречь. После насыщения перехода $1 \rightarrow 3$ устанавливается распределение населенностей между четырьмя уровнями, показанное на рис. 18.3, б, из которого видно, что при $n_1 = n_3$ населенность второго уровня больше первого, т. е. $n_2 > n_1$. Следовательно, переход $2 \rightarrow 1$ характеризуется «отрицательной температурой», или инверсией населенностей (наклонная штриховая прямая между уровнями 1 и 2). Если теперь на парамагнитный

кристалл (кристалл рубина) поступает слабый сигнал с частотой f_{12} такой, что $hf_{12} = E_2 - E_1$, то вследствие индуцированного перехода с уровня 2 на уровень 1 произойдет усиление этого сигнала. В реальном кристалле число возбужденных ионов хрома в спиновой системе достаточно велико, поэтому среднее число переходов в единицу времени, вызываемое приходящим сигналом, практически постоянно, а процесс усиления происходит непрерывно, несмотря на то что переход любого отдельного спина, индуцированный приходящим сигналом, представляет собой случайный процесс. Можно говорить о том, что отдельный спин совершит переход с некоторой вероятностью под воздействием приходящего сигнала. Поскольку вероятность индуцированного излучения прямо пропорциональна спектральной плотности энергии излучения (см. п. 17.1), то число излучательных переходов в единицу времени прямо пропорционально начальному уровню мощности усиливаемого сигнала. В определенном диапазоне мощностей входного сигнала процесс квантового усиления можно считать линейным.

Существует несколько методов, позволяющих повысить инверсию населенности по сравнению с рассмотренным случаем использования трехуровневой системы. Прежде всего при высоких частотах сигнала инверсию желательно получать путем накачки на более низкой частоте, поскольку генераторы накачки на более высоких частотах менее выгодны с различных точек зрения.

В парамагнитной спиновой системе квантового усилителя с четырьмя или большим числом рабочих уровней инверсию населенностей можно повысить за счет одновременной накачки на двух или более переходах, причем иногда для этих целей можно использовать один и тот же источник накачки. Наибольшее распространение получили методы с многократной накачкой, поясняемые рис. 18.4, особенно первый из них на рис. 18.4, а (метод *пушпульной накачки*) и разновидность его на рис. 18.4, з. Под воздействием накачки происходит одновременный переход спиновых систем (ионов хрома) с уровня 1 на верхний сигнальный (рабочий) уровень 3 и с нижнего сигнального уровня 2 на уровень 4, т. е. в этом случае на рабочем переходе происходит заселение верхнего рабочего уровня с одновременным освобождением нижнего, что и приводит к увеличению инверсии.

Расчеты показывают, что коэффициент инверсии при такой накачке пропорционален $(f_n/f_c) - 1$, а в трехуровневой системе этот коэффициент пропорционален $(f_n/2f_c) - 1$, где f_n , f_c — частоты накачки и сигнала соответственно. Для рубина симмет-

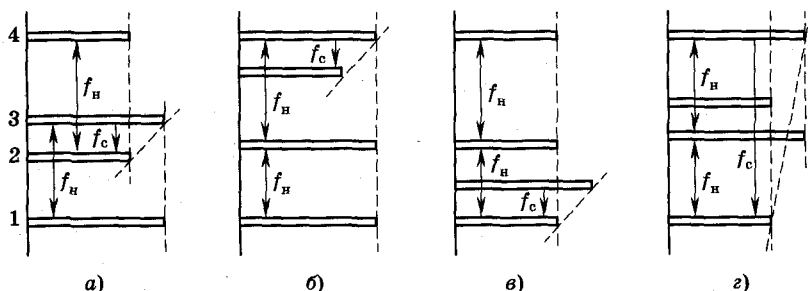


Рис. 18.4

ричная схема энергетических уровней (см. рис. 18.4), необходимая для пушпульной накачки одним источником, автоматически реализуется при угле θ между вектором \vec{B} и осью кристалла, равным $\theta = 54^\circ 44'$.

На рис. 18.4, б и 18.4, в показаны две схемы пушпульной накачки. В отдельных частных случаях коэффициент инверсии для этого случая совпадает с соответствующим выражением для мазера с пушпульной накачкой, но в любом случае он больше, чем для трехуровневого мазера.

К другим методам повышения инверсии относятся: использование явления кросс-релаксации; сокращение времени релаксации; накачка на более низких частотах (например, $f_H = f_{12} < f_{32}$ (см. рис. 18.4, з)); гармоническая накачка, связанная с одновременным насыщением двух переходов (см. рис. 18.4, г).

Кросс-релаксация представляет собой процесс обмена энергией между двумя магнитными переходами, которые принадлежат к многоуровневой системе спинов.

Для усиления широкополосных сигналов эффективность однорезонаторных усилителей оказывается недостаточной из-за малой полосы пропускания. Многорезонаторные усилители и усилители бегущей волны позволяют существенно увеличить частотную полосу усиливаемых сигналов. Использование таких усилителей дает увеличение эффективности в несколько раз, а в последовательно-параллельной схеме — в десятки раз. На рис. 18.2, б представлена блок-схема усилителя с двумя связанными резонаторами (контурами). Активное вещество может располагаться как в верхнем, так и нижнем резонаторе, но предпочтительным является последнее расположение. На практике используется также сложная схема из четырех последовательно-параллельно соединенных резонаторов.

18.2.2. Мазеры бегущей волны (МБВ).

МБВ имеют лучшие параметры, чем многорезонаторные усилители. Усиление в МБВ получается за счет длительного взаимодействия усиливаемого сигнала с активным веществом. Полоса пропускания усилителя бегущей волны ограничена полосой пропускания замедляющей системы и шириной контура спектральной линии рабочего вещества. В качестве замедляющей системы (см. п. 13.3) часто применяется гребенчатая структура, при использовании которой однонаправленное излучение достигается без заметного усложнения конструкции прибора. Чтобы получить однонаправленное усиление, т. е. устранить отражения, вызванные рассогласованием на входе и выходе, вводят специальный поглотитель обратной волны и определенным образом располагают активное вещество по отношению к замедляющей системе (см. рис. 18.2, в). На рис. 18.2, в введены следующие обозначения: 1 и 2 — входной и выходной штыри связи соответственно с источником сигнала и нагрузкой; 3 — активное вещество (рубин); 4 — невзаимный элемент (обычно феррит в магнитном поле); 5 — гребенчатая замедляющая система; 6 — волновод.

Замедляющая система, активное вещество и невзаимный элемент размещаются в отрезке волновода 6, который одновременно является низкодобротным резонатором.

Накачка может быть осуществлена с использованием волноводного переходника или специального штыря, дополнительно введенного в секцию волновода. На рис. 18.2 опущены такие элементы конструкции МБВ, как магнит (электромагнит) и криогенное устройство для охлаждения активного вещества.

18.3. Охлаждение мазеров и их параметры

Усиление в твердотельных мазерах можно получить при глубоком охлаждении кристалла до температуры жидкого гелия (1,2...4 К). Причины, по которым приходится это делать, связаны со спецификой больцмановского распределения применительно к СВЧ-диапазону частот, когда энергия кванта электромагнитной волны на частоте сигнала и накачки много меньше тепловой энергии колебаний атомов и ионов вплоть до криогенных температур, т. е. $hf_{\text{СВЧ}} \ll kT$.

При больцмановском распределении отношение населенностей нижнего и верхнего уровней равно

$$n_i/n_j = e^{\frac{-(E_i - E_j)}{kT}} = e^{\frac{hf}{kT}}.$$

При $hf \ll kT$ разность населенностей крайне мала; так, при $T \approx 300$ К $(n_i - n_j)/n_i$ составляет всего лишь 10^{-4} , поэтому для получения заметной разности населенностей необходимо охладить спиновую систему (парамагнитный возбужденный кристалл) до гелиевых температур, хотя возможность усиления существует и при температуре жидкого азота (77 К) и даже сухого льда (195 К). На практике весь резонатор или всю волноводную систему погружают в жидкий хладагент, содержащийся обычно в хорошо изолированных сосудах Дьюара с двойными стенками. Для жидкого гелия применяются двойные сосуды, в которых внутренний сосуд с гелием помещают во внешний с азотом. Необходимость охлаждения мазеров является их существенным недостатком.

Однако наличие низкой температуры имеет и положительную сторону, связанную с крайне низким уровнем шумов, который для мазеров обычно оценивают не коэффициентом шумов, а эквивалентной избыточной входной шумовой температурой усилителя $T_{\text{экв}}$ (см. п. 22.3). Как показывают расчеты, шумовая температура связана с коэффициентом шума F соотношением, вытекающим из формулы (22.8):

$$F = 1 + \frac{T_{\text{экв}}}{T_0},$$

где T_0 — шумовая температура генератора (источника сигнала).

Обычно T_0 принимается равной температуре окружающей среды при нормальных условиях, т. е. $T_0 \approx 290$ К.

Эквивалентная шумовая температура усилителя $T_{\text{экв}}$ имеет тот же порядок величины, что и его рабочая температура, т. е. составляет около 4 К, тогда коэффициент шума составляет $F \approx 1,01$, т. е. шум-фактор равен нескольким сотым дБ. В реальных системах полные шумовые характеристики усилителя всегда определяются не шумами усилителя, а омическими потерями и шумами во входных линиях, потерями в антенне и т. д.

Во всех реальных случаях шумы самого усилителя оказываются пренебрежимо малыми. Следует отметить, что тепловые шумы квантовой системы тесно связаны со спонтанным излучением, которое для спиновых систем в СВЧ-диапазоне играет существенно меньшую роль, чем для переходов, соответствующих оптическому диапазону (см. гл. 19, 20). Однако это основной фи-

зический процесс, который в конечном счете обуславливает собственные тепловые шумы квантовых усилителей.

Мазеры как усилители со сверхнизким уровнем шума целесообразно использовать только в случаях, когда остальные шумы в системе, и в том числе шумовой фон источника излучения, достаточно малы. Поэтому реализация очень малых потерь в волноводах, циркуляторах, вращающихся сочленениях и других элементах становится определяющей при создании радиосистем. Помимо этого, при конструировании квантового приемника приходится учитывать не только ширину диаграммы и коэффициент направленности излучения, но также боковые и задние лепестки, поскольку через них может поступать на вход тепловое (шумовое) излучение окружающих предметов, хотя главный луч и будет направлен на источник излучения холодного неба.

Современные КПУ работают начиная с дециметрового и вплоть до миллиметрового диапазона длин волн, имеют усиление 20...40 дБ, полосу пропускания до 20 МГц (многорезонаторные усилители) и до 40 МГц (усилители с бегущей волной). КПУ имеют малый динамический диапазон: насыщение в резонаторных усилителях наступает при мощности входных сигналов 10^{-9} ... 10^{-10} Вт, а в усилителях бегущей волны — при мощности 10^{-7} Вт, время восстановления порядка 0,01...0,1 с. Последние два фактора можно отнести к существенным недостаткам рассматриваемых усилителей.

КПУ нашли применение в первую очередь в системах исследования космического пространства, для управления и связи с искусственными спутниками Земли и другими летательными аппаратами, в системах ПРО для обнаружения баллистических ракет, в РЛС дальнего обнаружения и других областях науки и техники.

18.4. Квантовые генераторы СВЧ-диапазона

Основным достоинством квантовых генераторов СВЧ-диапазона является чрезвычайно высокая стабильность частоты, много лучшая, чем у других типов генераторов. Однако уровень мощности излучения этих генераторов крайне мал, поэтому они могут применяться как опорные источники высокостабильных колебаний, корректирующих частоту кварцевых генераторов, а также использоваться в *квантовых стандартах частоты и времени*, в навигации, при исследовании космоса и в других областях науки и техники. В качестве эталона времени принят период колебаний, соот-

ветствующий центру контура спектральной линии атома цезия Cs^{133} на частоте 9 192 631 770 Гц. Эталон длины 1 м сравнивается с длиной волны λ спектральной линии криптона Kr^{86} , а именно на отрезке в 1 м укладывается 1 650 763,77 длин волн этой линии.

Стандарты частоты разделяются на *пассивные* и *активные*. В *пассивных приборах* используется внешний СВЧ-генератор, частота которого подстраивается под частоту одного из переходов атомов рубидия или цезия, колба с которыми помещается в резонатор и облучается газосветной лампой, заполненной также парами рубидия или цезия. Излучение от лампы имеет частоты, соответствующие частотам переходов атомов в колбе, размещенной в резонаторе, поскольку она заполнена теми же атомами, что и газосветная лампа. Излучение от газосветной лампы, проходящее через колбу, вызывает вынужденные переходы между оптическими уровнями (частота переходов этих уровней лежит в оптическом диапазоне), причем из-за большей населенности нижних уровней число переходов с поглощением больше, чем с излучением, что приводит к ослаблению интенсивности света, проходящего через колбу. Если в резонатор, где расположена колба, подать СВЧ-мощность с частотой, соответствующей частоте перехода между нижними уровнями атомов рубидия или цезия (частота этих переходов лежит уже в СВЧ-диапазоне и верхний уровень СВЧ перехода является нижним для указанного оптического перехода), то произойдет перераспределение населенности между оптическими уровнями. В результате поглощение света колбой возрастет, поскольку увеличилась населенность нижнего уровня оптического перехода за счет энергии СВЧ-колебаний, что вызовет уменьшение интенсивности светового потока, регистрируемого фотодетектором. Для увеличения точности настройки СВЧ-генератора на частоту перехода атомов в колбе применяется низкочастотная фазовая модуляция СВЧ-сигнала специальным генератором. Далее промодулированный низкочастотный сигнал используется для управления частотой кварцевого генератора.

В службе времени и других областях науки и техники, где существует потребность в сигналах, составляющих сетку стандартных высокостабильных частот (1 кГц, 10 кГц, 100 кГц, 1 МГц и др.), для этих целей наиболее целесообразно использовать современные *активные квантовые стандарты частоты*, среди которых следует выделить часто применяемые автогенераторы на атомах водорода. Схема одного из них приведена

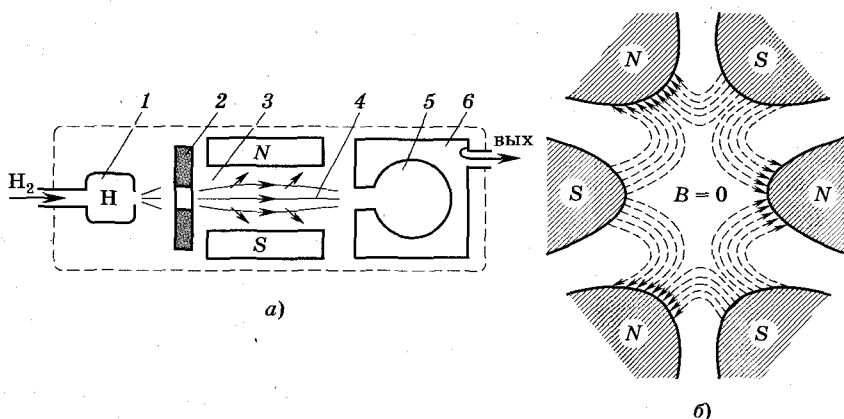


Рис. 18.5

на рис. 18.5. В генераторе используется переход с частотой 1420,405 МГц между энергетическими уровнями сверхтонкой структуры основного состояния атома водорода.

Принцип действия и устройство водородного генератора, схематично показанного на рис. 18.5, *а*, состоит в следующем. В источнике 1 под действием высокочастотного разряда происходит разложение молекулярного водорода на атомы. Разряд осуществляется при давлении 10...50 Па. Диафрагма 2, представляющая собой систему параллельных каналов с диаметром $\sim 1,4$ мм, формирует параллельный атомарный пучок 4, проходящий магнитное сортирующее устройство 3. Около резонатора 6 с накопительной ячейкой 5 поддерживается давление 10^{-5} Па. Разность давлений в источнике 1 и на входе в резонатор заставляет атомы водорода двигаться от источника через каналы в камеру.

Перед входом в магнитное сортирующее устройство 3 в пучке находятся атомы водорода как на нижнем E_1 , так и на верхнем E_2 энергетических уровнях рабочего перехода. Сортировка атомов по указанным энергетическим состояниям осуществляется неоднородным магнитным полем, создаваемым шестиполюсным постоянным магнитом, формирующим структуру поля, показанную на рис. 18.5, *б*. Атомы, находящиеся в верхнем энергетическом состоянии, фокусируются вблизи оси шестиполюсного магнита, где магнитная индукция равна нулю ($B = 0$), и попадают в размещенную в резонаторе кварцевую колбу (накопительная ячейка 5), стенки которой покрыты тефлоном. Суть метода сортировки состоит в том, что на частицы, в данном случае на атомы водорода,

обладающие магнитным моментом, в неоднородном магнитном поле действует сила, направление которой определяется знаком производной $\frac{\partial B}{\partial r}$ (r — радиус-вектор, направленный от оси сортирующей системы к периферии). В селекторе Z напряженность магнитного поля возрастает от оси к периферии, и производная $\frac{\partial B}{\partial r}$ имеет один и тот же знак. С увеличением магнитного поля энергия нижнего уровня (E_1) уменьшается, а верхнего (E_2) возрастает, поэтому и силы, действующие на атомы, находящиеся в состояниях с неравными энергиями E_1 и E_2 , будут различны. В результате частицы с энергией E_2 отклоняются к оси и попадают в резонатор, а с энергией E_1 — уходят от оси и удаляются насосом, благодаря чему и создается инверсная населенность уровней.

Полный поток атомов водорода, поступающих в кварцевую накопительную ячейку, покрытую изнутри тефлоном, равен $10^{12} \dots 10^{13}$ частиц в секунду¹. Атомы в колбе находятся в течение нескольких секунд, испытывая в среднем 10^5 соударений со стенками колбы, и вылетают через то же входное отверстие. Особенность тефлона состоит в том, что при соударении с ним возбужденных атомов происходит сохранение возбужденного состояния. В результате атомы водорода сохраняют возбужденное состояние в течение времени гораздо большего, чем время свободного пробега в резонаторе, что увеличивает время жизни на уровне и уменьшает ширину спектральной линии генерации. Отражаясь от стенок накопителя, атомы движутся по различным направлениям, причем усредненная за время их взаимодействия с полем излучения скорость практически равна нулю, чем резко ослабляется влияние эффекта Доплера на уширение контура спектральной линии. В целом ширина спектральной линии излучения атомов водорода в этой системе имеет чрезвычайно малую величину $\sim 0,0057$ Гц (в молекулярном генераторе на аммиаке ширина линии составляет единицы кГц). Стабильность частоты водородного генератора характеризуется величиной не хуже $\sim 10^{-12}$, что является лучшим показателем для всех генераторов такого

¹ Накопительная ячейка применяется для увеличения времени пребывания атомов в резонаторе, что приводит к росту числа излучательных переходов и обеспечивает генерацию.

типа. Для получения столь высокой стабильности необходима тщательная магнитная экранировка. Мощность водородного генератора $\sim 5 \cdot 10^{-12}$ Вт, что достаточно для стабилизации частоты более мощных генераторов. Он работает совместно со специальными радиотехническими устройствами, служащими для формирования сетки частот и управления атомными часами.



Контрольные вопросы

1. Охарактеризуйте особенности активного парамагнитного вещества. Как образуются парамагнитные уровни энергии и каковы их характеристики?
2. С чем связаны зеемановское расщепление уровней и явление электронного парамагнитного резонанса?
3. Резонаторные квантовые парамагнитные усилители: устройство, способы создания инверсии населенностей, особенности накачки в трех- и четырехуровневых системах.
4. Квантовые парамагнитные усилители (КПУ) бегущей волны: устройство, физические механизмы усиления сигнала, параметры и характеристики.
5. Необходимость охлаждения для работы мазеров. Влияние криогенных температур на характеристики мазеров.
6. Квантовые генераторы СВЧ-диапазона, особенности создания инверсной населенности за счет сортировки частиц в неоднородном магнитном поле. Параметры современных квантовых стандартов частоты.

Глава 19

ГАЗОВЫЕ ЛАЗЕРЫ

19.1. Общие сведения

Газовые лазеры — наиболее широко используемый вид лазеров. Среди них можно найти лазеры такого типа, которые будут удовлетворять почти любому требованию.

Особенности газовых лазеров большей частью обусловлены тем, что они, как правило, являются источниками излучения атомных или молекулярных спектров. Поэтому длины волн переходов точно известны. Они определяются атомной структурой и практически не зависят от условий окружающей среды. Генерация может быть осуществлена в любой части оптического диапазона — от УФ (~ 200 нм) до далекой ИК области ($\sim 0,4$ мм), частично захватывая микроволновую область.

Разреженность рабочего газа обеспечивает оптическую однородность среды с низким коэффициентом преломления, что позволяет применять простую математическую теорию для описания структуры поперечных мод резонатора. Необходимо также отметить, что линии спонтанного излучения в газовом лазере наиболее узки ($10^{-8} \dots 10^{-2}$ Å). Ширина линии генерации газовых лазеров минимальна среди всех видов лазеров и может составлять десятки герц. Благодаря высокой однородности активной среды в газовых лазерах можно получать наименьший угол расходимости лазерного луча по сравнению с другими типами лазеров (порядка минуты без всяких дополнительных коллимирующих устройств).

Газовые лазеры можно разделить на три вида: *атомарные, ионные и молекулярные*.

Лазеры на нейтральных атомах являются главным образом генераторами ближнего ИК и красного излучения, поскольку для нейтральных атомов энергия, соответствующая лазерным переходам, составляет обычно $1 \dots 2$ эВ.

В ионных газовых лазерах энергия излучения квантов значительно больше — $2 \dots 5$ эВ и более, поэтому ионные газовые лазеры генерируют видимое и УФ-излучение.

В молекулярных лазерах используются колебательные и вращательные уровни молекул, для которых энергия переходов составляет сотые и десятые доли эВ, что соответствует излучению в далеком ИК и субмиллиметровом диапазонах длин волн.

19.2. Процессы создания инверсии населенностей в газовых лазерах

Свойства газовых лазеров зависят от величины давления и природы газа, от взаимного расположения и времени жизни рабочих уровней, от энергии и плотности свободных электронов, от размеров объема, где расположен газ, от способа подведения энергии и ряда других факторов (см. п. 15.2). Можно указать

следующие механизмы создания инверсии населенности в газовых лазерах: неупругие столкновения частиц 1-го и 2-го рода в газовом разряде, оптическая накачка, диссоциация молекул, фотодиссоциация. В большинстве случаев в газовых лазерах накачка осуществляется за счет газового разряда, создаваемого в активной среде, где располагаются электроды. Основные физические процессы в газовом разряде рассмотрены в гл. 15. Здесь будут перечислены и просуммированы те положения, которые доминируют при получении инверсии населенностей.

Из всего большого разнообразия видов разряда в газе (см. п. 15.2) в технике газовых лазеров наиболее часто используются тлеющие и дуговые разряды, в которых можно выделить три характерные области: прикатодную, прианодную и область положительного столба (см. п. 15.2). Напомним, что положительный столб отличается от приэлектродных областей небольшим градиентом потенциала и отсутствием объемного заряда. В плазме положительного столба в основном и создается инверсия населенностей, поскольку в этой области присутствуют как электроны, так и нейтральные возбужденные и невозбужденные атомы, молекулы и ионы, как положительные, так и отрицательные.

Все частицы в плазме положительного столба находятся в непрерывном хаотическом движении, а заряженные частицы еще участвуют и в непрерывном дрейфовом движении. В процессе своего движения частицы взаимодействуют (соударяются) друг с другом, обмениваются энергией, в результате чего и происходит возбуждение и ионизация атомов и молекул. Процессы взаимодействия между частицами в плазме разряда принято делить на две группы: упругие и неупругие.

За счет упругих взаимодействий энергия от заряженных частиц передается нейтральным, а направленное движение частиц превращается в хаотическое. Поэтому средняя энергия хаотического движения частиц в плазме возрастает и, следовательно, их температура оказывается выше, чем температура окружающей среды, однако увеличение средней энергии различно для разных частиц.

Среди неупругих процессов взаимодействия 1-го рода наиболее существенными являются следующие: прямое электронное возбуждение, ступенчатое электронное возбуждение, ионизация атома (см. п. 15.2). При прямом электронном возбуждении возбужденные состояния атомов заселяются при малой концентрации электронов в разряде. При ступенчатом возбуждении засе-

ляются возбужденные состояния атомов при достаточно большой концентрации электронов в газовом разряде. Ионизация приведет к увеличению количества заряженных частиц в плазме. Все эти три неупругих процесса являются процессами 1-го рода, так как в результате взаимодействия суммарная кинетическая энергия частиц убывает (см. п. 15.2).

При взаимодействии медленного электрона с возбужденным атомом возможно соударение 2-го рода, при этом атом отдаст свою внутреннюю энергию и перейдет в невозбужденное состояние. Результатом этого процесса является обеднение возбужденных состояний. Другим эффективным процессом неупругих соударений 2-го рода являются соударения нейтральных и возбужденных атомов в смеси различных газов, когда невозбужденный атом одного газа сталкивается с возбужденным атомом другого газа и в результате этого происходит передача возбуждения. Результатом такого процесса является дополнительное заселение возбужденных состояний за счет взаимодействия с атомами буферного газа. Вероятность этого процесса тем выше, чем ближе друг к другу расположены энергетические уровни, между которыми осуществляется обмен возбуждениями, т. е. в идеальном случае эти уровни должны иметь одно и то же значение энергии.

Перечисленные процессы далеко не исчерпывают все многообразие явлений в плазме положительного столба. Однако они позволяют в первом приближении трактовать механизмы создания инверсии населенностей в наиболее распространенных видах газовых лазеров. В заключение подчеркнем роль упругих и неупругих процессов в создании инверсии населенностей. Инверсия населенностей непосредственно обязана неупругим соударениям. Упругие же столкновения определяют среднюю энергию хаотического перемещения электронов (электронную температуру) в газовом разряде, которая определяет вероятность и значение сечения возбуждения или ионизации. Сечения процессов немонотонно зависят от электронной температуры.

19.3. Атомарные газовые лазеры.

Гелий-неоновый лазер

Первым газовым лазером непрерывного действия был лазер, в котором в качестве активной среды использовалась смесь двух газов — гелия и неона. Схематическое устройство He—Ne лазера

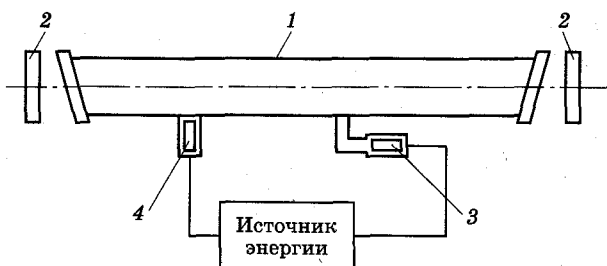


Рис. 19.1

приведено на рис. 19.1, где 1 — газоразрядная трубка со смесью газов гелия и неона, 2 — зеркала резонатора, 3 — накаливаемый катод, 4 — анод, 5 — источник электрического питания.

Газоразрядная трубка (ГРТ) 1 изготавливается из кварца или высококачественного стекла, подвергается термической обработке, откачивается и заполняется рабочей смесью. Отношение парциального давления гелия (He) и неона (Ne) обычно составляет величину $P_{\text{He}}/P_{\text{Ne}} = 5 \dots 15$. Выходные окна ГРТ устанавливаются под углом Брюстера к оптической оси резонатора. Наличие брюстеровских окон обуславливает поляризацию генерируемого излучения и минимальное отражение его от торцов ГРТ. Наибольшее усиление имеет место для волны, поляризованной в плоскости, проходящей через ось резонатора и нормаль к плоскости выходного окна (на рис. 19.1 плоскость чертежа, являющаяся плоскостью падения для луча лазера). В этой плоскости поляризовано выходное излучение газового лазера.

Возбуждение газовой среды может быть обеспечено как посредством безэлектродного высокочастотного индукционного разряда, так и за счет разряда на постоянном токе (см. рис. 19.1).

В газовых лазерах на нейтральных атомах общее давление газовой смеси составляет величину порядка 1 мм рт. ст.

Индукцированное излучение создается атомами неона, а атомы гелия облегчают получение инверсной населенности в неоне за счет неупругих процессов соударения 2-го рода.

Схема нижних возбужденных состояний атомов гелия и неона приведена на рис. 19.2. Нижним возбужденным состояниям атома гелия соответствуют значения энергии 20,61 и 19,82 эВ. Переходы $2^1 S_0 \rightarrow 1^1 S_0$ и $2^3 S_1 \rightarrow 1^1 S_0$ запрещены, поэтому состояния $2^1 S_0$ и $2^3 S_1$ являются метастабильными с большим временем жизни $\tau \sim 10^{-3}$ с. Энергетические состояния неона 2s,

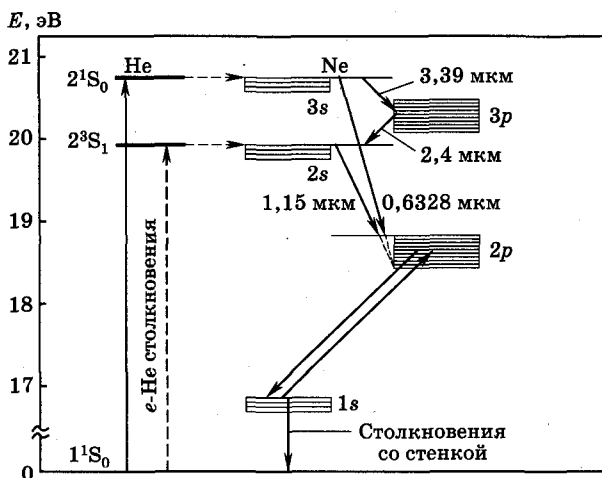


Рис. 19.2

$3s$, $2p$, $3p$, как это показано на рис. 19.2, представляют собой в каждом случае несколько близко расположенных энергетических уровней. Уровни $3s$ и $2s$ неона энергетически близки к метастабильным уровням гелия. Состояния $1s$ неона являются метастабильными (с временем жизни $\tau \sim 10^{-3}$ с); релаксационные переходы из этих состояний происходят преимущественно за счет столкновений атомов со стенками трубки. Накопление атомов в этих состояниях является нежелательным, поскольку это приводило бы к дополнительному заселению уровней $2p$ (за счет захвата фотонов, возникающих при релаксационном излучательном переходе $2p \rightarrow 1s$), а следовательно, к уменьшению инверсной населенности на рабочих лазерных уровнях.

Возбуждение атомов гелия и неона в разряде происходит за счет прямого электронного соударения. Населенность состояния в установившемся режиме определяется произведением скорости заселения на время жизни данного состояния. Следует отметить, что времена жизни s -состояний неона несколько больше времен жизни p -состояний, поэтому даже при равных скоростях возбуждения возможна инверсия населенностей в стационарном режиме на переходах $s-p$.

Присутствие в разряде метастабильных атомов гелия приводит к передаче возбуждения от метастабилей гелия 2^1S_0 и 2^3S_1 к атомам неона (состояния $3s$ и $2s$). Поскольку эффект передачи возбуждения является резонансным, то он приводит к селек-

тивному заселению тех состояний неона, энергия которых наиболее близка к энергии исходных состояний гелия. Таким образом, возникает дополнительный механизм заселения состояний $2s$ и $3s$, и их населенность при определенных условиях разряда может превысить населенность состояний группы $2p$ и $3p$.

При рассмотрении условий получения инверсии на переходах $2s \rightarrow 2p$ и $3s \rightarrow 3p$ следует учитывать, что состояния $1s$ неона хорошо заселяются в разряде и при больших токах разряда процесс ступенчатого возбуждения атомов из состояний $1s$ в состояния $2p$ и $3p$ может стать существенным. Этот процесс ведет к дополнительному подзаселению нижних рабочих состояний и, следовательно, к уменьшению, инверсии и срыву генерации. Генерация на переходах с $\lambda \sim 1,15$ мкм, $3,39$ мкм и $0,63$ мкм наблюдается примерно в одинаковых условиях разряда и характеризуется аналогичными зависимостями мощности генерации от параметров разряда.

Выделение генерации на определенном переходе с определенной длиной волны осуществляется подбором интерференционных зеркал, введением в резонатор диспергирующего или селективно поглощающего элемента. Наибольшее усиление соответствует переходу $3s \rightarrow 3p$ с $\lambda \sim 3,39$ мкм. Мощность генерации P в первом приближении пропорциональна превышению инверсии ΔN над пороговым значением $\Delta N_{\text{пор}}$, обеспечивающим компенсацию потерь в резонаторе, т. е. $P \sim (\Delta N - \Delta N_{\text{пор}})$. Значение инверсии и мощность генерации в значительной степени зависят от тока разряда, общего давления смеси, соотношения компонент смеси, диаметра разрядной трубки, формы сечения разрядной трубки.

Зависимость мощности генерации от тока разряда для генерации на длине волны $\lambda \sim 0,63$ мкм представлена на рис. 19.3. С увеличением разрядного тока возрастает плотность электронов в плазме положительного столба разряда, все процессы электронного возбуждения интенсифицируются, населенность всех возбужденных, в том числе и рабочих состояний, возрастает. При небольших токах разряда I_p , когда процессами ступенчатого возбуждения из состояний $1s$ можно пренебречь, возрастание населенности рабочих состояний происходит пропорционально току, соответственно увеличивается и инверсия населенностей. Когда величина инверсии превысит пороговое значение ($I_p \sim 2 \dots 20$ мА), возникает генерация. Дальнейшее увеличение тока приводит к возрастанию мощности излучения.

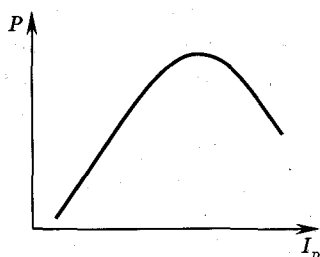


Рис. 19.3

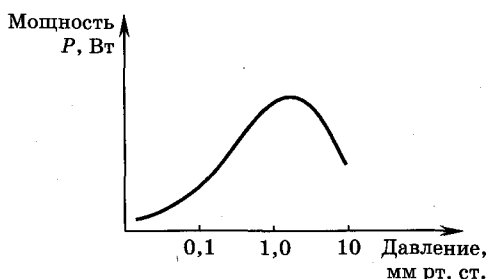


Рис. 19.4

Как отмечалось выше, при больших разрядных токах, когда концентрация электронов в плазме велика, начинают сказываться процессы ступенчатого возбуждения нижних рабочих состояний $2p$ и $3p$ с долгоживущих уровней $1s$. Скорость этого процесса приблизительно пропорциональна квадрату концентрации электронов, в то время как скорость прямого возбуждения приблизительно линейно связана с концентрацией электронов. В результате дополнительного заселения нижних рабочих состояний инверсия снижается, мощность генерации падает и при токах I_p порядка нескольких сотен мА генерация срывается.

Увеличение общего давления смеси при малых давлениях до $1...2$ мм рт. ст. (примерно $100...200$ Па) приводит к возрастанию мощности генерации. Это связано с увеличением концентрации атомов неона и гелия и с общим ростом населенностей возбужденных состояний. При большом давлении существенным оказывается снижение электронной температуры, приводящее к резкому уменьшению числа электронов, способных возбуждать атомы смеси. Характерная кривая зависимости мощности излучения от общего давления смеси дана на рис. 19.4.

Инверсия на рабочих переходах сильно зависит от передачи возбуждения с метастабилей гелия к атомам неона. В разряде с равной вероятностью происходит и обратный процесс, связанный с передачей возбуждения от атомов неона в состояниях $2s$ и $3s$ к невозбужденным атомам гелия. Вероятности двух встречных процессов практически равны, их скорости пропорциональны населенностям исходных состояний. Для того чтобы направить процесс в смеси от гелия к неону, необходимо создать избыточную концентрацию гелия, поэтому в рабочей смеси гелия всегда больше. Оптимальное соотношение парциальных давлений неона и гелия находится в диапазоне от $1 : 5$ до $1 : 15$.

На величину выходной мощности сильное влияние оказывает диаметр газоразрядной трубки. Увеличение диаметра приводит к увеличению объема активной среды, в результате чего мощность должна была бы нарастать. Однако с ростом диаметра трубки снижается электронная температура, что снижает инверсию. Кроме того, одним из основных процессов, приводящих к уменьшению населенности состояния $1s$, является диффузия возбужденных атомов неона к стенкам и их столкновение со стенками, что приводит к освобождению состояния $1s$. С увеличением диаметра трубки время диффузии атомов неона к стенкам увеличивается, следовательно, эффективное время жизни в состоянии $1s$ также возрастает, вследствие чего инверсия населенностей снижается. Таким образом, населенность нижних рабочих состояний лазерного перехода растет с увеличением диаметра. В результате существует оптимальный диаметр ГРТ, который для трубки метровой длины равен 7...9 мм. Повышения мощности генерации можно добиться за счет применения трубок с эллиптическим сечением. Делая сечение трубки эллиптическим, можно, сохраняя минимальный размер сечения неизменным, увеличить объем активной смеси за счет другого размера сечения.

19.4. Ионные лазеры

В ионных лазерах используются не атомные переходы, а переходы между возбужденными состояниями ионов. Ионные лазеры способны генерировать гораздо большую непрерывную мощность по сравнению с лазерами на атомных переходах. Это объясняется тем, что рабочие состояния ионных переходов располагаются значительно выше на энергетической диаграмме. Поэтому выравнивание заселенностей инвертированных уровней с ростом тока разряда (например, заселение уровней $1s$ и $2p$ в неоне) за счет ступенчатых процессов, наблюдавшееся на атомных состояниях, в ионных лазерах не проявляется. Вероятности ионных переходов, как правило, больше, чем вероятности атомных. Разность энергий между уровнями рабочего перехода больше и, следовательно, излучение происходит в более коротковолновой области по сравнению с атомарными лазерами. Ионные лазеры генерируют свет в видимом и УФ диапазонах длин волн. Чтобы получить значительную долю ионов в разряде, плазма должна быть высокоионизованной, что можно обеспечить использованием сильноточного дугового разряда. Рабочий ток в ионных

лазерах достигает нескольких десятков А, а плотность тока — 1000 А/см^2 . Это вызывает необходимость использования охлаждаемых электродов. Охлаждается также и газоразрядная трубка, которая для получения высокой плотности газа делается малого диаметра, в виде капилляра. Рабочее давление газа, как правило благородного, составляет десятые доли мм рт. ст. Одним из наиболее распространенных ионных лазеров непрерывного действия является аргоновый лазер, работающий в диапазоне длин волн $454...514 \text{ нм}$. В аргоновом лазере генерация осуществляется на переходе $4p-4s$. При создании инверсии населенностей заселение возбужденных состояний иона аргона происходит в основном за счет ступенчатого электронного возбуждения из основного ионного состояния. Измеренные значения времен жизни для уровней $4p$ в несколько раз превышают времена жизни уровней $4s$, что позволяет реализовать инверсию населенностей на этом переходе с некоторой разностью населенностей, которую обозначим ΔN . При увеличении давления газа концентрация электронов возрастает, а электронная температура уменьшается. Это приводит к существованию некоторого оптимального давления, при котором инверсия ΔN и мощность генерации максимальны. Оптимальное давление для различных линий несколько различается, причем для переходов, верхние уровни которых располагаются выше, оптимальное давление меньше. Область оптимальных давлений — десятые доли мм рт. ст.

При возрастании тока разряда увеличивается концентрация заряженных частиц, которая зависит от плотности тока более сильно, чем при линейном законе. При токах в несколько десятков А электронная температура от тока почти не зависит, поэтому с увеличением тока величина ΔN и мощность генерации возрастают (рис. 19.5) примерно по квадратичному закону $P \sim I^2$. Фактором, который в конечном счете должен привести к насыщению инверсии ΔN при увеличении тока, является поглощение на переходе между нижними рабочими и основным ионным состояниями. Этот процесс при большой концентрации ионов в основном состоянии приводит к тому, что любой излучаемый квант тут же вновь поглощается и, таким образом, место

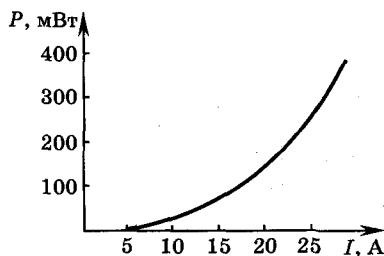


Рис. 19.5

выбывшего из состояния $4s$ иона занимает новый. Происходит увеличение эффективного времени жизни $4s$ -состояния и уменьшение инверсной населенности ΔN . Суммарная мощность генерации ионных лазеров достигает десятков Вт на трубках длиной до 2 м, коэффициент полезного действия составляет сотые—десятые доли процента.

19.5. Молекулярные газовые лазеры

Атомарные и ионные лазеры имеют весьма малый коэффициент полезного действия. Это связано с тем, что верхний рабочий уровень высоко расположен над исходным состоянием, с которого идет возбуждение, поэтому в процессе создания инверсии населенностей принимает участие лишь малая доля общего числа электронов.

Для гелий-неонового лазера значение энергии электрона, необходимое для возбуждения гелия, около 20 эВ (см. рис. 19.2). Средняя энергия хаотического движения электронов (электронная температура) в разряде порядка $7 \dots 8$ эВ, электронов с энергией выше 20 эВ в разряде не более $5 \dots 6\%$. Такого же порядка и эффективность использования энергии накачки. Подобная ситуация имеет место и в ионных лазерах. Кроме того, в указанных типах лазеров низка и эффективность использования энергетических уровней, которая определяется отношением энергии излучения ($E_{\text{изл}}$) на рабочем переходе к энергии возбуждения ($E_{\text{возб}}$). Для гелий-неонового лазера отношение $E_{\text{изл}}/E_{\text{возб}} \approx 2/20 \approx 0,1$, т. е. в этих лазерах КПД не может быть более 10%. Сказанное означает, что с точки зрения повышения КПД газовых лазеров перспективным является использование низко расположенных энергетических уровней, например вращательных и колебательных молекулярных уровней. В молекуле, состоящей из нескольких атомов, внутренняя энергия определяется не только энергией электронов каждого атома, но и энергией колебательного и вращательного движения атомов в целом около некоторого положения равновесия.

Рассмотрим молекулярный газовый лазер, работающий на смеси азота и углекислого газа. Схема нижних колебательных состояний молекул CO_2 и N_2 представлена на рис. 19.6. Генерация возникает на переходах $4 \rightarrow 3$ и $4 \rightarrow 2$ в молекуле CO_2 . Отметим, что верхний рабочий уровень в молекуле CO_2 отстоит от основного состояния меньше чем на 0,35 эВ (для сравнения у неона примерно 20 эВ). Нижнее колебательное возбужденное состояние

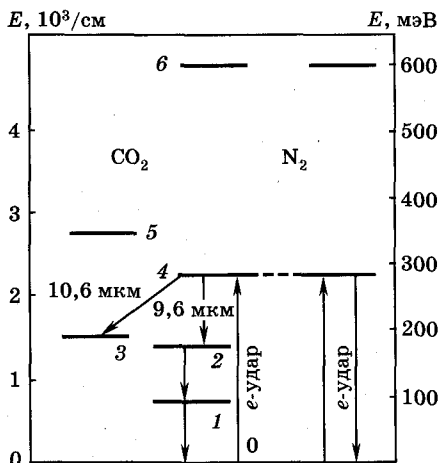


Рис. 19.6

молекулы азота заселяется в разряде за счет рекомбинации молекулярного иона азота. Переход между этим и основным состояниями запрещен. Поэтому возбужденное состояние N_2 характеризуется большой заселенностью. При взаимодействии возбужденных молекул азота с молекулами CO_2 , находящимися в основном состоянии, возможна передача возбуждения, ведущая к преимущественному заселению состояния 4. В результате возникает инверсия на переходах $4 \rightarrow 3$ и $4 \rightarrow 2$. Рабочему переходу $4 \rightarrow 3$ соответствует длина волны 10,6 мкм. Опустошение нижнего лазерного уровня 2 или 3 происходит очень быстро вследствие колебательного и вращательного обмена, так что молекула попадает в нижнее колебательное состояние 1, причем выравнивание заселенностей лазерных уровней 2 и 3 происходит в течение $10^{-8} \dots 10^{-9}$ с. Опустошение колебательного состояния 1 происходит на несколько порядков медленнее, так что необходимы дополнительные компоненты в активной среде для достижения быстрого опустошения уровня 1 и тем самым поддержания инверсной населенности на необходимом уровне. Оптимальными компонентами являются гелий, водород и вода. Соответственно этому используется газовый разряд со смесью CO_2 , N_2 и He (отношение парциальных давлений $P_{CO_2} : P_{N_2} : P_{He} = 1 : 1 : 8$). В лазерах же на чистой смеси CO_2-N_2 рабочий объем заполняется смесью углекислого газа и азота в соотношении $(1 : 1) \dots (1 : 5)$ при общем давлении

нии порядка 1 мм рт. ст. Добавка гелия значительно увеличивает мощность генерации. К уменьшению инверсии населенностей приводит диссоциация молекул CO_2 (50...80% молекул распадается в течение 0,1...1 с) с образованием молекул CO и повышением температуры газа, что приводит к более высокому тепловому заселению нижнего лазерного уровня и ускорению столкновительной релаксации верхнего лазерного уровня 4.

CO_2 — N_2 лазер является одним из важнейших типов лазеров. С помощью CO_2 -лазера достигается максимальная непрерывная мощность лазерного излучения. Его высокий КПД (больше 20%) превышает КПД почти всех других лазеров, за исключением полупроводниковых, существует несколько разновидностей CO_2 -лазеров. К ним относятся: CO_2 -лазеры с непрерывным возбуждением, импульсные CO_2 -лазеры, ТЕ — лазер на CO_2 высокого давления, волноводный CO_2 -лазер, газодинамический CO_2 -лазер и ряд других. Наиболее мощным из перечисленных типов CO_2 -лазеров являются *газодинамические*, с помощью которых можно получать непрерывную генерацию с мощностью более 100 кВт. В этих лазерах инверсия населенностей образуется благодаря тому, что смесь газов (CO_2 , N_2 , H_2O) с высокой температурой (~ 1500 К) и большим давлением (~ 2 МПа) вытекает через сверхзвуковое сопло с высокой скоростью. Из-за того, что при истечении газодинамической струи происходит ее расширение и охлаждение, нижний лазерный уровень CO_2 релаксирует к состоянию с более низкой температурой существенно быстрее, чем верхний, поэтому в направлении потока образуется инверсия населенностей.

В дальней ИК области вплоть до миллиметрового диапазона длин волн используются колебательно-вращательные уровни таких молекул, как CH_3Br , NH_3 , H_2O , H_2S , SO_2 , HCN , CH_3F , CH_3OH и т. д. Когерентное излучение в УФ диапазоне длин волн можно получить в лазерах на H_2 , а также на атомах Ne , Ar , Kr , Xe , Cl , F , C , O , N и т. д., находящихся в различных ионизационных состояниях.

19.6. Разновидности газовых лазеров

19.6.1. Эксимерные лазеры.

В качестве активной среды в лазерах этого типа выступают молекулы, которые могут существовать только в возбужденном состоянии. Такие молекулы называются эксимерами. В экси-

мерных молекулах для лазерной генерации используются электронно-колебательные переходы между устойчивым возбужденным (E_1) и химически неустойчивым основным состоянием (E_0). В качестве примера эксимерных молекул можно привести возбужденные молекулы благородных газов и их соединения: Xe_2^* , Kr_2^* , Ar_2^* , XeF^* , XeCl^* , KrF^* , ArF^* и т. д. (звездочка означает возбужденную молекулу).

Если в рабочем объеме создать большую концентрацию эксимерных молекул, например, пучком электронов с энергией до 1 МэВ и плотностью тока до 10^{11} А/см², то может быть получена инверсная населенность возбужденного уровня E_1 . В результате возникает инверсия населенностей на переходах между верхним связанным состоянием E_1 и нижним неустойчивым состоянием E_0 . Из-за того, что верхнее состояние E_1 представляет собой полосу электронно-колебательных уровней, а в нижнем состоянии отсутствует дискретная вращательно-колебательная структура уровней энергии, излучение эксимерного лазера происходит в сравнительно широком спектральном диапазоне, что позволяет перестраивать частоту генерации в пределах этого перехода.

Помимо широкополосности излучения, можно отметить также то, что эксимерные переходы из возбужденного в неустойчивое основное состояние из-за быстрой диссоциации эксимерной молекулы сопровождаются практически мгновенным опустошением нижнего лазерного уровня. Время жизни эксимеров мало ($\sim 10^{-9} \dots 10^{-8}$ с), поэтому эксимерные лазеры работают в основном в импульсном режиме, генерируя короткие импульсы излучения с энергией до 10^5 Дж и КПД до 10%. Эксимерные лазеры излучают энергию в УФ диапазоне длин волн ($\sim 1 \dots 3,5$) 10^2 нм.

19.6.2. Химические лазеры.

Инверсия населенностей в химических лазерах достигается за счет энергии химических реакций между отдельными составляющими активной среды, при этом используются только экзотермические реакции. В результате этих реакций в газовых смесях, например таких, как HF, происходит выделение энергии, большая часть которой переходит в колебательную энергию молекул.

Таким образом, основным достоинством химических лазеров является прямое преобразование химической энергии в энер-

гию электромагнитного лазерного излучения с высоким КПД ($\sim 10\%$). К другим достоинствам химических лазеров можно отнести большие мощности излучения в непрерывном режиме (≥ 10 кВт), высокое значение удельной энергии (~ 10 Дж/л), отсутствие громоздких источников питания. Химические лазеры работают на колебательно-вращательных уровнях в спектральном диапазоне 3...10 мкм.

19.6.3. Рентгеновские лазеры ($\lambda < 100$ нм).

Разработка этих лазеров обусловлена требованиями развития нанотехнологий, особенно в области микроэлектроники, применениями в оптической микроскопии, голографии, в перспективных системах противоракетной обороны и т. д. Энергия кванта излучения в рентгеновских областях спектра составляет десятки и даже сотни эВ, что создает большие трудности при разработке подходящих активных сред с нужным спектром энергетических уровней. Кроме того, в рентгеновском диапазоне возрастает роль спонтанных переходов, поскольку отношение их вероятности к вероятности индуцированного излучения пропорционально кубу частоты $A_{mn}/B_{mn} \sim \omega^3$, что затрудняет генерацию индуцированного излучения (см. п. 17.1).

В качестве активных сред в рентгеновских лазерах используются плазменные среды с многократно ионизированными атомами. Энергетические состояния ионов по своей структуре аналогичны энергетической структуре соответствующих атомов с той лишь разницей, что значения разности энергий уровней ионизированных атомов намного больше, чем для нейтральных. Например, спектры разрешенных значений энергии восьмикратно ионизированного атома аргона Ar^{8+} или иона селена Se^{24+} аналогичны спектру атома Ne, но значения разности энергий уровней у этих многократно ионизированных атомов Ar и Se больше, чем у Ne, в 50 и 500 раз соответственно.

Получение плазмы с такими многократно ионизированными (многозарядными) ионами возможно только в установках типа тех, которые используются для термоядерного синтеза, поэтому это ограничивает их применение в настоящее время. Однако в силу важности отмеченных выше возможных применений, работы по созданию этих лазеров в развитых странах проводятся достаточно интенсивно. В настоящее время уже получена лазерная генерация в режиме сверхкоротких импульсов на Se^{24+} ,

Ar^{8+} , Ag^{37+} и ряде других элементов в диапазоне длин волн от 3 до 50 нм.

19.6.4. Лазеры на свободных электронах.

В лазерах этого типа используются основные преимущества электронных вакуумных приборов с динамическим управлением, например ЛВВО и ЛВВМ (см. гл. 13, 14). В лазерах на свободных электронах активной средой являются электронные потоки, ускоренные до релятивистских скоростей и движущиеся через ондулятор. Ондулятор представляет устройство с периодически изменяющимся в пространстве электрическим или магнитным полем. В магнитном ондуляторе поле формируется набором расположенных друг за другом магнитов с чередующейся полярностью. Кроме поступательного движения, электроны под действием такого магнитного поля совершают периодические колебания (осцилляции), которые сопровождаются электромагнитным излучением с частотой $\omega_0 \approx v/\lambda$, где v — скорость переносного (продольного) движения электронов, $D = \lambda$ — пространственный период изменения магнитного поля.

Расчеты показывают, что для получения излучения электронов в видимом диапазоне спектра необходимо разогнать их до энергии в 50 МэВ (при $D = 1$ см). Современные ускорители заряженных частиц позволяют разогнать электроны до энергий, превышающих 500 МэВ, что будет соответствовать излучению на красном краю рентгеновского диапазона.



Контрольные вопросы

1. Каковы особенности газовой среды как активного вещества лазеров?
2. Какие процессы определяют получение инверсии населенностей посредством газового разряда? Охарактеризовать роль каждого из процессов.
3. Роль гелия в He—Ne-лазерах. Рабочие переходы и селекция частот генерации.
4. Объяснить физические процессы, определяющие зависимость мощности излучения от тока разряда в He—Ne-лазерах.
5. Особенности энергетической структуры рабочей (активной) среды в молекулярных лазерах; специфика их параметров и характеристик.

6. Объяснить факторы, определяющие КПД газовых лазеров.
7. Особенности устройства и физические механизмы генерации эксимерных и химических лазеров. Применение и параметры эксимерных лазеров.
8. Особенности устройства и физические механизмы генерации рентгеновских лазеров и лазеров на свободных электронах. Применение указанных типов лазеров и их параметры.

Глава 20

ТВЕРДОТЕЛЬНЫЕ И ЖИДКОСТНЫЕ ЛАЗЕРЫ

20.1. Особенности активных сред

В твердотельных лазерах активным веществом является кристаллический или аморфный диэлектрик, в котором генерирующими центрами являются специально введенные активные атомы. По принципу действия к лазерам на твердом теле очень близки жидкостные лазеры, в которых активным веществом служат жидкие диэлектрики с растворенными в них примесями.

Структура энергетических уровней примесных атомов в ряде растворов оказывается весьма близкой к той, которая наблюдается в твердых диэлектриках. В силу сказанного, процессы в твердотельных и жидкостных лазерах имеют много общего и, следовательно, оба типа лазеров целесообразно рассматривать вместе. Концентрация активных атомов обычно составляет единицы или доли процента от полного числа атомов в среде, т. е. активное вещество в рассматриваемых типах лазеров представляет собой раствор или аналогичную среду слабой концентрации, подобную идеальному газу. Однако по сравнению с газовыми лазерами концентрация активных частиц в твердотельных и жидкостных лазерах все же намного больше, что позволяет в конденсированных средах достигнуть большей плотности инверсии населенностей и большего удельного энергосъема. Существенным недостатком таких лазеров, из-за особенностей активного вещества, является ограничение способов накачки практически единственным, а именно облучением вещества светом (оптическая накачка).

Отдельные атомы в твердых (и жидких) телах находятся в поле, создаваемом всеми остальными, близко расположенными частицами вещества, что приводит к существенному отличию энергетической структуры этих атомов (ионов) в конденсированной среде от соответствующих уровней энергии изолированных частиц.

Это вызывает расщепление отдельных дискретных уровней и превращение их в энергетические зоны или полосы. Особенно это характерно для уровней энергии, соответствующих внешним электронам в атоме. На рис. 20.1 показаны уровни энергии в изолированном атоме (рис. 20.1, а) и соответствующие им энергетические зоны в твердом теле (рис. 20.1, б).

При такой энергетической структуре спектр поглощения света активным веществом должен состоять из отдельных полос и линий, ширина которых зависит от ширины соответствующих энергетических уровней (зон). Основная причина различия спектров изолированных атомов и твердого или жидкого тела связана с электрическим взаимодействием отдельных атомов с внутренним полем вещества.

Перечислим основные требования, предъявляемые, исходя из энергетических соображений, к активным средам твердотельных лазеров.

1. Активное вещество должно обладать метастабильными уровнями энергии, на которых из-за большого времени их жизни можно создать значительную населенность, т. е. получить достаточное число активных атомов. Наличие метастабильных уровней с высоким квантовым выходом проявляется в существовании интенсивных линий люминесценции, достаточно узких, что является необходимым условием для увеличения коэффициента усиления в среде с инверсной населенностью.
2. Активное вещество должно обладать широкими полосами поглощения. При широких полосах поглощения эффективно используется энергия излучения немонахроматического источника накачки. В этом случае большинство атомов, поглотивших фотоны накачки, переходит затем в метастабильное состояние, что увеличивает КПД лазера, при этом необходимо, чтобы частота возбуждающего излучения не очень сильно превышала частоту лазерного излучения.

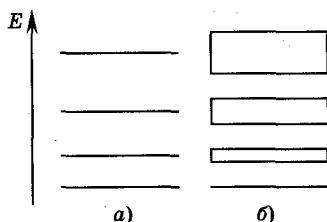


Рис. 20.1

3. Активное вещество должно обладать по возможности малыми оптическими потерями за счет рассеяния на неоднородностях на частоте генерации. Эти потери уменьшают добротность резонатора и, следовательно, КПД лазера.

Наиболее полно указанным требованиям удовлетворяют активные диэлектрики, представляющие собой твердые растворы элементов с недостроенными внутренними электронными оболочками (парамагнитные ионы) в различных кристаллических и аморфных матрицах. Из-за хорошей экранировки недостроенных внутренних электронных оболочек переходных элементов периодической системы от внешнего поля валентными электронами при введении таких элементов в конденсированную среду не происходит при взаимодействии с решеткой *коренной перестройки* их энергетического спектра. Однако активный парамагнитный ион в диэлектрике (матрице) находится под воздействием электростатического поля окружающих его частиц среды, в результате чего изменяются энергетическая структура иона, а также вероятности излучательных и безызлучательных переходов. Лазерные рабочие переходы осуществляются между уровнями энергии активных ионов. Поэтому нахождение оптимальной матрицы является важной задачей.

Сформулируем основные требования, которым должны удовлетворять матрицы лазерных диэлектриков. Матрица должна обладать:

- хорошими оптическими свойствами — оптической однородностью, прозрачностью для лазерного излучения;
- хорошим согласованием с активатором (активным веществом), чтобы можно было вводить заданный активатор в регулируемых количествах без нарушения ее механических и оптических свойств;
- хорошей теплопроводностью, чтобы исключить термические деформации и термооптические искажения;
- высокой стойкостью по отношению к лазерному излучению и мощному излучению накачки.

В настоящее время не создано материала, удовлетворяющего в полной мере перечисленным требованиям. Наилучшими свойствами обладают: рубин ($\text{Al}_2\text{O}_3 : \text{Cr}^{3+}$), алюмоиттриевый гранат, активированный ионами неодима ($\text{Y}_3\text{Al}_5\text{O}_{12} : \text{Nd}^{3+}$, или YAG : Nd^{3+}),

стекло с неодимом, иттербий — эрбиевое стекло, алюминат иттрия с неодимом ($YAl_2O_3 : Nd^{3+}$) и ряд других соединений.

При ламповой оптической накачке наибольший КПД получен в лазере на стекле с примесью неодима (8%).

20.2. Системы накачки

В лазерах рассматриваемого типа в основном используется оптическая накачка. Системы накачки включают в себя источник света и оптическую систему, которая с минимальными потерями транспортирует световую энергию в активную среду.

20.2.1. Источники света.

Наиболее интенсивные полосы поглощения большинства активных сред находятся в видимой части спектра и соседних участках ИК и УФ диапазонов. Спектр излучения источника всегда находится в области частот, больших частоты излучения лазера, т. е. излучение лампы накачки смещено в более коротковолновую часть спектра. В силу этого для генерации лазерного излучения в видимом диапазоне необходимо, чтобы лампа накачки излучала свет в области фиолетовых и УФ длин волн. Для этого требуются источники с высокой температурой поверхности, поскольку согласно закону Стефана—Больцмана полная мощность E_T (интегральная по спектру) с единицы поверхности абсолютно черного тела пропорциональна четвертой степени его температуры:

$$E_T \approx \sigma T^4,$$

где $\sigma \approx 5,67 \cdot 10^{-12}$ Вт \cdot см $^{-2}$ \cdot К $^{-4}$.

Длина волны, соответствующая максимуму излучения, определяется законом смещения Вина и обратно пропорциональна температуре абсолютно черного тела. Под абсолютно черным телом понимается идеальное тело, которое поглощает все падающее на тело электромагнитное излучение. Излучения реальных тел всегда несколько отличаются от закона излучения абсолютно черного тела, но с достаточной для практических целей точностью излучение реальных источников света может рассматриваться как излучение черного тела с некоторой эффективной температурой.

Максимум спектральной плотности потока излучения черного тела, в соответствии с формулой Планка, при температуре ~ 3000 К приходится на ИК диапазон. Подобные источники могут быть использованы для возбуждения таких кристаллов, как $\text{CaF}_2 : \text{Dy}^{2+}$ и $\text{YAG} : \text{Nd}^{3+}$, имеющих полосы поглощения в ближней ИК области. При увеличении эффективной температуры источника согласно закону Вина максимум в спектре излучения смещается в коротковолновую область. Расчеты и эксперименты показывают, что спектральный максимум на длине волны 560 нм наблюдается при температуре около 6500 К, а на длине волны 410 нм — при 10 000 К. Указанные полосы характерны для спектра поглощения рубина, широко используемого в твердотельных лазерах. Суммарное значение эффективности (КПД) для двух указанных полос составляет в максимуме излучения около 30%. Эффективные температуры излучения порядка 5000...15 000 К реализуются в газоразрядной плазме, в силу чего газоразрядные лампы получили широкое распространение среди источников накачки твердотельных лазеров.

Наиболее важными параметрами таких источников являются КПД, спектральный состав излучения, предельно допустимая мощность (энергия). КПД лазера $\eta_{\text{л}}$ определяется как отношение излучаемой энергии $W_{\text{л}}$ к потребляемой электрической энергии $W_{\text{е}}$, т. е. $\eta_{\text{л}} = W_{\text{л}}/W_{\text{е}}$.

Как правило, для накачки твердотельных лазеров используются импульсные газоразрядные лампы, заполненные ксеноном или парами ртути, либо галогенные лампы накаливания, в которых лампа заполняется парами иода или фтора, либо смесью различных галогенов. Наиболее распространены лампы с иодным циклом. Галогенные лампы в основном используются для накачки в непрерывном режиме. Галогены восстанавливают накальную нить и предотвращают ее распыление. В процессе горения лампы при высокой температуре происходит распыление вольфрамовой нити. В результате происходит затемнение поверхности стеклянного баллона за счет оседания распыленного вольфрама. Галогены вступают в реакцию с парами

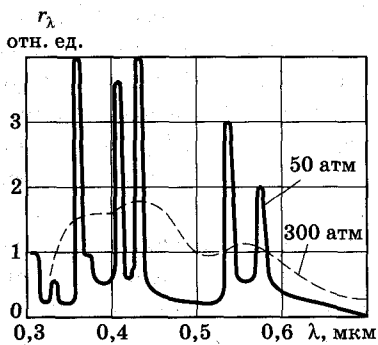


Рис. 20.2

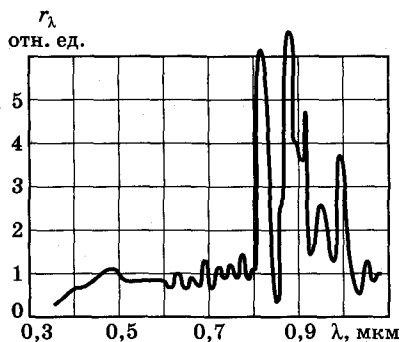


Рис. 20.3

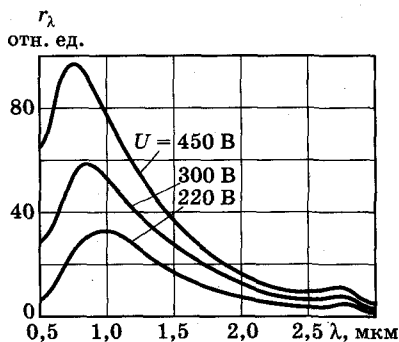


Рис. 20.4

вольфрама, и образующееся соединение возвращается на нить, где происходит восстановление вольфрама.

На рис. 20.2, 20.3, 20.4 приведены характерные спектры излучения газоразрядных ламп: ртутной (при различных давлениях паров ртути), ксеноновой и лампы накаливания с иодным циклом (при различных напряжениях накала U).

20.2.2. Системы оптической накачки.

Создание инверсии населенностей в активной среде, достаточной для лазерной генерации, требует интенсивного ее облучения. Пороговые интенсивности достигают десятков Вт/см². Соответствующие энергии проще получать в импульсном режиме. Система оптической накачки, обеспечивающая создание лазерной генерации, содержит в общем случае лампу накачки и светооптические элементы (например, эллиптические отражатели), концентрирующие лучистый поток, испускаемый лампой накачки, на активном элементе. Эффективность системы накачки можно определить как отношение поглощаемого в активном материале лучистого потока к мощности, потребляемой всеми лампами накачки. Несколько видов систем оптической накачки представлено на рис. 20.5, где 1 — активная среда, 2 — лампа накачки, 3 — зеркальная поверхность, 4 — оптический фильтр (см. рис. 20.5, *б*).

Одна из первых систем накачки, которая была использована создателем рубинового лазера Т. Г. Мейманом в 1960 г., изображена на рис. 20.5, *а*. Активный материал в этой системе размещен таким образом, что спиральная газоразрядная лампа обвивает его. Лампа вместе с активным стержнем помещается в цилиндриче-

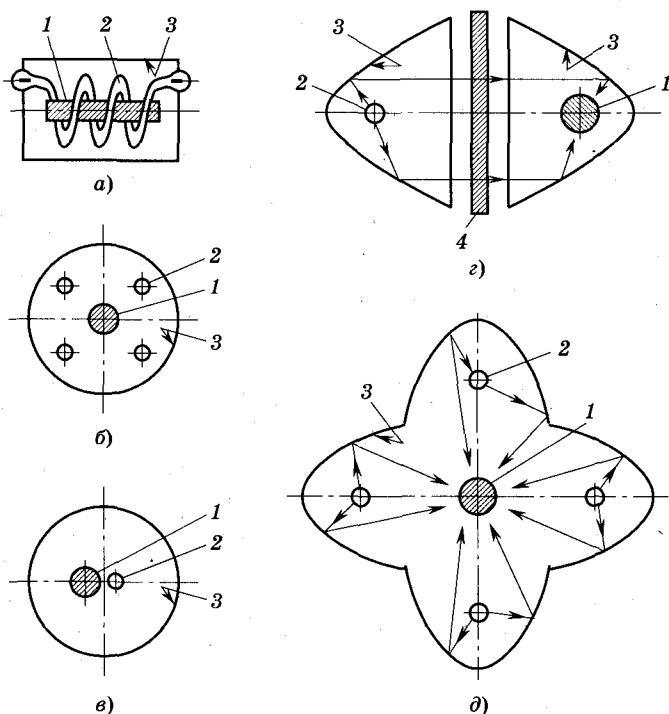


Рис. 20.5

ский отражатель, внутренние стенки которого покрыты зеркальным слоем. Облучение активного кристалла происходит как прямым (от лампы) лучистым потоком, так и отраженным от поверхности зеркального покрытия цилиндрического рефлектора.

Очевидно, что от эффективности выбранной системы накачки в прямой зависимости находится КПД всего генератора в целом, поэтому вопросам рационального конструирования их уделяется большое внимание. Эффективность первой рассмотренной системы относительно невысока, при этом большая часть излученной энергии рассеивается вне образца. Для сравнительной оценки эффективности различных систем накачки удобно пользоваться значением пороговой мощности или энергии возбуждения, при которой возникает генерация в определенном активном образце.

В рассмотренной системе накачки пороговая энергия для рубинового кристалла имеет порядок 1000 Дж. На один—два порядка большей эффективностью обладают системы, в которых используется фокусировка излучения на активном образце.

Если использовать другое расположение источника излучения и активного образца в цилиндрическом рефлекторе, то можно добиться существенно лучших результатов. В качестве примера можно привести систему накачки, состоящую из цилиндрического рефлектора кругового сечения, в котором лампа и активный образец расположены симметрично относительно продольной оси цилиндра и вдоль нее на одинаковом расстоянии (рис. 20.5, б). Эффективность передачи энергии в активный образец в рассматриваемой системе зависит от осевого расстояния активного образца и лампы относительно продольной оси цилиндра. При близком расположении образца и лампы КПД может быть высоким (см. рис. 20.5, в) и приближаться к КПД более совершенных систем с эллиптическими отражателями. Однако с увеличением осевого расстояния в системах накачки с круговыми цилиндрическими отражателями КПД уменьшается значительно быстрее. Очевидно, что с учетом потерь на отражение от поверхности активного образца и при неполном отражении от рефлектора, максимальная эффективность передачи энергии в активный элемент для рассматриваемой системы и системы с эллиптическими отражателями не может превосходить величины $\eta_{\max} = r_p(1 - r_{\text{обр}})$, где r_p и $r_{\text{обр}}$ — коэффициенты отражения от поверхности рефлектора и активного образца при нормальном падении.

Система накачки с эллиптическими отражателями (рис. 20.5, г) обладает фокусирующими свойствами. Эллипс представляет собой геометрическое место точек, сумма расстояний от которых до двух точек внутри эллипса, называемых фокусами эллипса, есть величина постоянная и равная большей оси эллипса, поэтому лучи, исходящие из одного фокуса, будут собираться во втором. Таким образом, световая энергия, излученная лампой, расположенной вдоль одной из фокальных линий, концентрируется эллиптическим отражателем в месте расположения образца (вдоль другой фокальной линии). Из-за конечных поперечных размеров лампы через вторую фокальную линию будут проходить только лучи, проекции которых на поперечное сечение отражателя перпендикулярны поверхности лампы. Поскольку другие лучи фокусируются неточно, то изображение лампы занимает некоторую область вблизи фокальной линии. Следовательно, КПД данной системы зависит от поперечных размеров источника накачки, активного образца и отражателя, а также от эксцентриситета эллипса. Чем меньше эксцентриситет, тем лучше фокусировка.

В ряде случаев с целью увеличения энергии накачки применяются отражатели, образованные несколькими, чаще всего четырьмя, «эллиптическими цилиндрами» (рис. 20.5, *д*), имеющими одну общую фокальную линию, где и располагается активный образец. Это дает возможность фокусировать на активном образце излучение от нескольких источников, т. е. увеличивать плотность потока энергии на поверхность активного кристалла. Однако эффективность использования энергии в такой системе (КПД) оказывается более низкой, чем в случае одноэллипсного цилиндра, что связано с уменьшением фокусирующей поверхности каждого из эллипсов. Максимальный КПД эллиптических систем, так же как и цилиндрических, как уже указывалось, определяется коэффициентами отражения от поверхности рефлектора и образца при нормальном падении.

Поглощение энергии активным элементом может быть значительно увеличено при многократном прохождении его излучением накачки, что реализуется в системах с плотным заполнением, в которых поперечные размеры отражателя и активного элемента отличаются незначительно (рис. 20.6, где 1 — активный элемент, 2 — отражатель, 3 — лампа накачки), так как в таких системах нивелируется неточность фокусировки излучения лампы накачки. Эффективность системы с плотным заполнением оказывается примерно такой же, как и в случае эллиптического рефлектора с размерами излучателей (ламп), по возможности близкими к точечным излучателям. Однако системы с плотным заполнением проще в изготовлении. Достоинство же эллиптических отражателей заключается в возможности расположения активного образца на значительном расстоянии от лампы, что важно для мощных лазеров.

Высокая эффективность использования энергии источника накачки достигается в сфероидальном эллипсоиде вращения (рис. 20.7), где 1 — лампа, 2 — активный образец, 3 — отража-

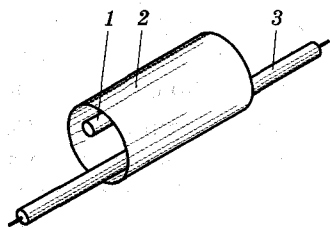


Рис. 20.6

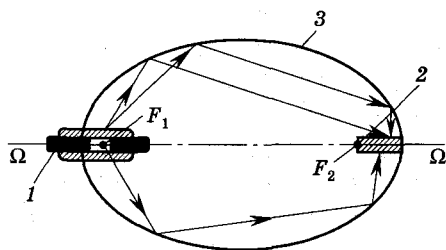


Рис. 20.7

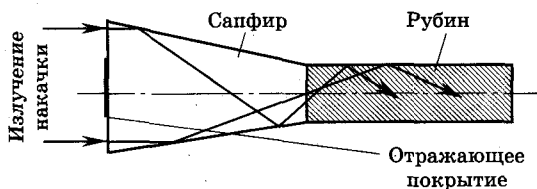


Рис. 20.8

тель, Ω — большая ось вращения эллипсоида. Лампа и активный кристалл располагаются вдоль оси Ω между соответствующим фокусом и поверхностью эллипсоида. Лучи, испускаемые лампой левее фокуса F_1 , будут фокусироваться в области правее фокальной точки F_2 (см. рис. 20.7). В этом случае практически весь световой поток лампы попадает на активный образец, и распределение интенсивности излучения на нем получается приблизительно таким же, как и на лампе.

В целях увеличения поглощения энергии накачки иногда используется торцевая накачка, при которой возбуждающее излучение поступает в активный кристалл через торцевую поверхность и распространяется вдоль оси активного элемента, представляющего собой рубиновый стержень с нарастанным на него сапфировым конусом (рис. 20.8). За счет полного внутреннего отражения света от поверхности сапфирового конуса осуществляется концентрация световой энергии, падающей на торец активного элемента, в кристалле рубина.

20.3. Рубиновый лазер

Первый лазер был создан в 1960 г. на основе искусственного кристалла рубина. Из твердотельных лазеров рубиновые в настоящее время являются одними из наиболее распространенных. Рубин отличается высокой химической стойкостью, механической прочностью и высокой теплопроводностью. Эти качества и обусловили широкое использование рубина в лазерных устройствах.

Рубин — это кристаллический минерал, основу которого составляет решетка Al_2O_3 с внедренными в нее трехзарядными ионами хрома Cr^{3+} , которые замещают часть ионов алюминия Al^{3+} . Содержание хрома, как правило, колеблется в пределах от 0,05 до 0,5%. Чем больше хрома, тем более красный оттенок имеет рубин. Рубин может выращиваться искусственно. К кристаллам рубина, используемым в лазерах, предъявляются жесткие тре-

бования по оптической однородности, размерам, составу, поэтому технология выращивания кристаллов рубина является весьма прецизионной. В рубиновом лазере кристаллическая решетка Al_2O_3 является матрицей, а ионы хрома — активными элементами, т. е. генерирующими центрами. Вследствие взаимодействия между ионами кристаллической решетки основной уровень иона хрома расщеплен на ряд подуровней. Схема энергетических уровней иона Cr^{3+} в матрице Al_2O_3 приведена на рис. 20.9, где энергия указана в см^{-1} (1 см^{-1} соответствует энергии $1,24 \cdot 10^{-4}$ эВ). Эта достаточно сложная система уровней практически может быть приведена к классической трехуровневой квантовой системе.

Длины волн для середин полос поглощения равны $0,56 \text{ мкм}$ (560 нм) для нижнего уровня 3 и $0,41 \text{ мкм}$ (410 нм) для верхнего уровня 3, а ширина каждой из них составляет около $0,1 \text{ мкм}$ (100 нм). Поглощение в рубине зависит от направления распространения света относительно оптической оси кристалла. Ионы хрома, поглощая энергию накачки в зеленой ($\lambda \sim 0,56 \text{ мкм}$) и синей ($\lambda \sim 0,41 \text{ мкм}$) полосах оптического спектра, переходят в возбужденное состояние. Спектр поглощения рубина представлен на рис. 20.10, где две кривые соответствуют двум направлениям падающего излучения относительно оптической оси кристалла (\parallel — падающий свет параллелен оси кристалла, \perp — падающий свет перпендикулярен оси кристалла). Время жизни ионов хрома Cr^{3+} в состояниях 3 (см. рис. 20.9) составляет $5 \cdot 10^{-8}$ с. Основная часть

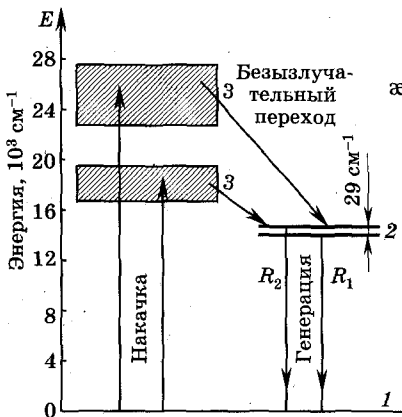


Рис. 20.9

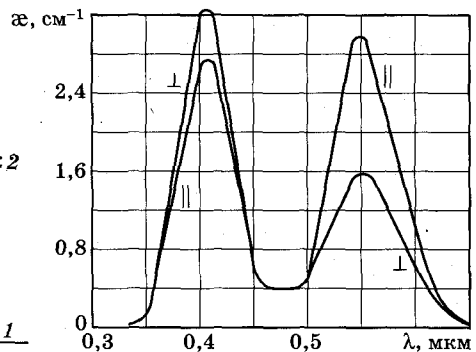


Рис. 20.10

возбужденных ионов хрома ($\sim 75\%$) с уровней 3 за счет безызлучательной релаксации переходит на метастабильный уровень 2. Остальные 25% возбужденных ионов переходят в исходное состояние 1. Уровень 2 состоит из двух близко расположенных подуровней, разность энергий между которыми приблизительно составляет $3,6 \cdot 10^{-3}$ эВ (29 см^{-1}). Оба подуровня находятся в тепловом равновесии, и, в соответствии с распределением Больцмана, населенность нижнего подуровня несколько больше, чем верхнего. Перераспределение частиц между подуровнями и восстановление равновесного состояния происходит с постоянной времени $\sim 10^{-7}$ с. Верхний подуровень подпитывает нижний, непрерывно восполняя уход возбужденных частиц из нижнего состояния.

Накопление возбужденных частиц в состоянии 2 и, следовательно, получение инверсии населенностей для рабочего лазерного перехода $2 \rightarrow 1$ возможно благодаря большому ($3 \cdot 10^{-3}$ с) времени жизни частиц на метастабильном уровне 2. Процесс накопления активных частиц будет продолжаться до тех пор, пока инверсная населенность не превысит пороговую. Для усиления волны в кристалле необходимо выполнение условия $N_2 - (g_2/g_1) \cdot N_1 \geq 0$. Расчеты показывают, что в режиме генерации колебаний в лазере на рубине на втором уровне должно находиться не менее трети от общего числа всех возбужденных частиц. В рубиновом лазере из-за наличия двух подуровней 2 генерация осуществляется на двух линиях, которые обозначаются R_1 и R_2 (см. рис. 20.9). Длины волн этих линий зависят от температуры кристалла, так как температура изменяет внутривещеточное расщепление энергетических уровней основного состояния. Зависимость длины волны генерации от температуры кристалла является специфической особенностью твердотельных лазеров. Длина волны генерации лазера на рубине при комнатной температуре на линии R_1 составляет 694,3 нм, а на линии R_2 — 692,9 нм.

Максимальный показатель поглощения α в полосах уровней 3 составляет $2 \dots 4 \text{ см}^{-1}$ при оптимальном содержании ионов хрома. При этом диаметр активного рубинового элемента не превышает 2 см. Обычно применяют искусственные стержни диаметром около 1,0 см, оптическая ось которых составляет угол 90° или 60° с осью стержня. Лазерное излучение в этих случаях линейно поляризовано с вектором электрического поля, расположенным перпендикулярно плоскости, проходящей через оптическую ось кристалла рубина и ось стержня.

Характерное пороговое значение интенсивности накачки в рубиновом лазере на полосе поглощения $\lambda \sim 0,56$ мкм составляет величину порядка ~ 3 Дж/см². Удельный съём энергии в режиме свободной генерации (см. п. 20.6.1) составляет $0,2 \dots 0,25$ Дж/см², т. е. КПД не превышает $1 \dots 2\%$. Низкий КПД рубинового лазера обусловлен его работой по трехуровневой схеме.

20.4. Лазеры на средах, активированных неодимом

В средах, активированных примесью редкоземельных элементов в качестве генерирующих центров, реализуется работа по четырехуровневой схеме, а не по трехуровневой, как в лазере на рубине, что повышает КПД и улучшает некоторые параметры таких твердотельных лазеров.

Материалы на основе иттрий-алюминиевого граната $Y_3Al_5O_{12}$, вольфрамата кальция $CaWO_4$ и стекло с примесями иона неодима Nd^{3+} , а также фторицевого кальция CaF_2 с двухвалентным ионом диспрозия Dy^{2+} или трехвалентным ионом урана U^{3+} требуют сравнительно малой энергии возбуждения, поскольку работают по четырехуровневой схеме, которая, как отмечалось ранее, обладает большей эффективностью по сравнению с трехуровневой. Следует подчеркнуть, что из всех 14 редкоземельных элементов наилучшие результаты получены с использованием трехвалентных ионов неодима. Механизм генерирования оптических колебаний по четырехуровневой схеме можно проиллюстрировать на примере лазера на вольфрамате кальция ($CaWO_4$) с примесью

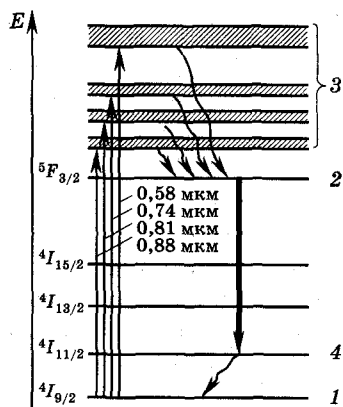


Рис. 20.11

ионов Nd^{3+} . Для матрицы YAG — иттрий-алюминиевого граната (ИАГ) происходящие при накачке процессы принципиально те же, что и для матрицы $CaWO_4$.

Диаграмма энергетических уровней Nd^{3+} в решетке $CaWO_4$ приведена на рис. 20.11, где показаны сравнительно узкие полосы поглощения в видимой ($\lambda \sim 0,58$ мкм) и ближней инфракрасной ($\lambda \sim 0,74; 0,81; 0,88$ мкм) областях спектра. Эти полосы поглощения выполняют роль третьего уровня (см. рис. 17.2, б).

Наиболее интенсивная полоса поглощения с максимумом $\lambda = 0,58$ мкм имеет ширину около 0,002 мкм. Полосы с $\lambda \sim 0,81$ мкм и 0,88 мкм очень удобны для накачки полупроводниковыми лазерами и светодиодами на основе арсенида галлия. Возбужденные за счет энергии источника накачки ионы неодима на уровнях 3 быстро переходят на метастабильный уровень 2, время жизни которого велико ($\tau_{24} = 0,23$ мс), и накапливаются на нем, создавая инверсию населенностей относительно уровня 4.

Рабочие лазерные переходы с уровня 2 возможны в состоянии ${}^4I_{15/2}$, ${}^4I_{13/2}$, ${}^4I_{11/2}$, ${}^4I_{9/2}$, но наиболее сильным является переход ${}^4F_{3/2} \rightarrow {}^4I_{11/2}$ (на уровень 4) с длиной волны $\lambda = 1,06$ мкм. Нижний рабочий лазерный уровень 4 расположен примерно на 0,25 эВ выше основного уровня 1. Быстрый безызлучательный переход с уровня 4 в основное состояние 1 обеспечивает эффективное опустошение нижнего рабочего уровня и повышает инверсию населенностей между уровнями 2 и 4. Такая схема переходов и определяет четырехуровневый характер генерации излучения неодимового лазера.

Энергетический зазор в 0,25 эВ между уровнями 1 и 4 существенно больше энергии kT ($kT = 0,026$ эВ при $T = 300$ К), поэтому уровень 4 при комнатных температурах в равновесном состоянии (в отсутствие накачки) практически не населен и для достижения инверсной населенности на переходе $2 \rightarrow 4$ требуется сравнительно небольшая энергия в отличие от трехуровневой схемы рубинового лазера. Пороговая инверсная населенность второго уровня, при которой в четырехуровневом лазере возможна генерация, определяется в основном потерями в резонаторе. В неодимовом лазере верхний рабочий уровень 2 расщеплен на два подуровня, а нижний уровень 4 — на шесть подуровней (не показаны на рис. 20.11). В результате излучение на длине волны $\lambda = 1,06$ мкм имеет сложную (тонкую) структуру. Наиболее сильная линия имеет полосу около 195 ГГц при комнатной температуре.

Процесс развития генерации излучения в неодимовом лазере реализуется так же, как и в других типах лазеров. Когда населенность уровня 2 превысит пороговую величину, в резонаторе лазера усиливаются волны на частоте рабочего перехода. Под действием возникшего первоначального поля излучения ионы неодима совершают индуцированные излучательные переходы $2 \rightarrow 4$, а затем с уровня 4 безызлучательно переходят в основное состояние. На-

личие промежуточных уровней (уровни между 2 и 4) не особенно сказывается на характере описанных процессов, поскольку вероятности переходов на эти уровни относительно невелики.

Недостатком лазеров на основе вольфрамата кальция является необходимость компенсации с помощью одновалентных ионов щелочных металлов неравновесного заряда в решетке, который возникает вследствие того, что трехвалентные ионы неодима замещают в узлах решетки двухвалентные ионы кальция. Возникающие вследствие этого локальные электрические поля приводят к неконтролируемым сдвигам энергетических уровней ионов неодима, что вызывает увеличение пороговой энергии. Компенсация заряда с помощью ионов щелочных металлов снижает увеличение пороговой мощности почти на порядок.

Активная среда на основе YAG с примесью неодима ($\text{YAG} : \text{Nd}^{3+}$) в значительной мере лишена недостатков, присущих вольфрамату кальция. В $\text{YAG} : \text{Nd}^{3+}$ трехвалентные ионы неодима заменяют трехвалентные же ионы иттрия, что исключает необходимость компенсации неравновесного заряда и поэтому снижает пороговую энергию возбуждения лазера. Кроме того, кристаллы граната обладают высокой теплопроводностью и малыми оптическими потерями. Благодаря указанным достоинствам лазеры на гранате работают как в непрерывном, так и импульсном режиме генерации с большой частотой повторения импульсов и средней мощностью до нескольких сотен Вт.

К недостаткам всех кристаллов, активированных редкими землями, относится отсутствие широких полос поглощения. Уровни 3 имеют незначительную ширину, поэтому использование ламповой накачки оказывается малоэффективным. Для увеличения эффективности накачки при использовании газоразрядных ламп накачивания в кристаллическую матрицу граната наряду с активными ионами вводят ионы другого вида, называемые сенсibilизаторами, которые имеют широкие полосы поглощения. К таким ионам относятся ионы Cr^{3+} , которые являются сенсibilизаторами для лазера $\text{YAG} : \text{Nd}^{3+}$. Добавление сенсibilизатора приводит к расширению эффективной полосы поглощения излучения накачки кристаллами граната с примесями, что повышает КПД лазера.

Помимо кристаллических матриц, широко используются активные среды на основе стекол с примесью редкоземельных элементов, чаще всего Nd^{3+} . Стекла являются аморфными средами и по сравнению с рассмотренными ранее обладают более высо-

кой оптической однородностью, технологичностью, возможностью изготовления активных элементов больших размеров и возможностью введения примеси в необходимых для получения большой мощности концентрациях с равномерным распределением по объему. К наиболее заметным недостаткам стекол как активных элементов относятся: низкая теплопроводность, высокий температурный коэффициент линейного расширения, ограниченная область оптической прозрачности ($0,33 \dots 2,5$ мкм), которая при введении примеси еще более сужается. Кристаллические и аморфные матрицы хорошо дополняют друг друга.

В стеклах из-за местных неоднородных электростатических полей, вызванных примесью, происходит сильное неоднородное уширение спектральных линий (см. п. 17.2) как при излучении, так и поглощении. Длина активных элементов на стеклах может достигать 1 м с поперечным сечением более 500 см^2 (диаметр около 25 см). С использованием таких элементов при большой концентрации активной примеси можно получать большую энергию импульсов (до нескольких тысяч Дж). Однако величина снимаемой средней мощности с активных стеклянных сред ограничена низкой теплопроводностью стекол. В импульсном режиме возможна генерация сверхкоротких световых импульсов длительностью до $5 \cdot 10^{-13}$ с мощностью в импульсе до 10^{13} Вт и более. КПД лазеров на стеклах выше, чем на гранатах. Он достигает значений порядка $8 \dots 10\%$.

20.5. Жидкостные лазеры

Жидкости, используемые в лазерах, как и твердые вещества, являются диэлектриками в конденсированной фазе, т. е. по своим физико-химическим свойствам жидкости и аморфные твердые среды близки друг к другу; так, например, стекло по существу представляет собой переохлажденную жидкость.

Использование жидкостей в качестве активной лазерной среды достаточно перспективно, поскольку в этом случае реализуется целый ряд достоинств по сравнению с другими средами. К таким преимуществам можно отнести легкость получения активной среды с требуемыми характеристиками, простоту изменения ее состава, возможность прокачки среды для поддержания ее в высокоэффективном состоянии, легкость охлаждения при большой мощности генератора, оптическую однородность и т. д.

Однако подобрать жидкость в качестве основы активной среды, удовлетворяющей таким требованиям, как интенсивная флуоресценция с большим квантовым выходом на рабочем переходе, широкие полосы поглощения энергии накачки, отсутствие потерь на частоте рабочего перехода, достаточно сложно. Многочисленные исследования различных жидкостей позволили реализовать лазерную генерацию лишь в относительно небольшом их количестве, при этом лазеры на этих средах, как правило, уступают твердотельным лазерам по своим рабочим характеристикам, за исключением жидкостных лазеров на органических красителях.

Эти лазеры не имеют аналогов с точки зрения возможности плавной перестройки длины волны генерации в очень широком диапазоне — во всей видимой области спектра, включая ближний УФ и ИК диапазоны, при этом реализуется излучение с очень высокой монохроматичностью, достигающей 1...1,5 МГц. Лазеры на органических красителях, работающие в импульсном режиме, имеют энергию в импульсе до нескольких сотен Дж и мощность в десятки Вт в непрерывном режиме генерации. КПД лазеров на органических красителях достигает десятков процентов при лазерной накачке.

В лазерах на красителях активное вещество создается на основе молекул органических красителей в органических растворителях или воде. Органические красители обладают интенсивными полосами поглощения в видимой области спектра, что и определяет их насыщенную окраску. Эти свойства краситель проявляет как в жидких, так и в твердых растворах. В кристаллической фазе указанные свойства сильно ослабевают. Структура молекулы красителя содержит бензольные (C_6H_6), пиридиновые (C_5H_5N), азотные ($C_4H_4N_2$) и другие кольцевые структуры.

Наибольшее распространение в лазерах на органических красителях получил родамин 6G, структура которого представлена на рис. 20.12. Энергетическая структура молекул органических

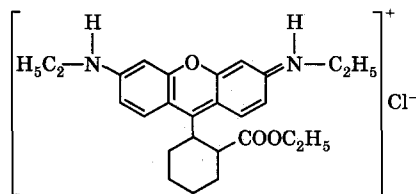


Рис. 20.12

красителей включает разрешенные уровни энергии электронных, колебательных и вращательных состояний. Каждому электронному состоянию соответствует набор колебательных уровней энергии, а каждому электронно-колебательному со-

стоянию полоса вращательных уровней. Полная энергия молекулы с достаточной степенью точности может быть определена как $E = E_{\text{эл}} + E_{\text{кол}} + E_{\text{вр}}$, где $E_{\text{эл}}$, $E_{\text{кол}}$, $E_{\text{вр}}$ — энергии электронных состояний и соответственно колебательного и вращательного движений атомов в молекуле.

Эти энергии находятся в следующем соотношении:

$$E_{\text{эл}} : E_{\text{кол}} : E_{\text{вр}} = 1 : \sqrt{m/M} : (m/M),$$

где m — масса электрона, M — масса молекулы.

Отношение m/M лежит в пределах 10^{-3} — 10^{-5} . Энергетические уровни электронных состояний по порядку величины составляют 1...3 эВ, колебательных — 10^{-2} ... 10^{-1} эВ и вращательных — 10^{-4} ... 10^{-3} эВ.

Колебательные и вращательные уровни перекрываются друг с другом, образуя энергетические полосы для определенных электронных состояний, которые являются мультиплетными. Мультиплетность — это степень вырождения квантового состояния атома или молекулы, характеризуемого заданным значением полного спинового момента. В лазерах на красителях используются две группы мультиплетных состояний: синглетные (S) и триплетные (T) состояния. Синглетные состояния имеют антипараллельную ориентацию спинов при полном спиновом квантовом числе $S = 0$, а триплетные — параллельную ориентацию спинов при $S = 1$. Структура энергетических мультиплетных уровней органического красителя представлена на рис. 20.13, где показаны излучательный переход — прямой стрелкой (флуоресценция) и безызлучательные переходы — волнистыми линиями; колебательные уровни энергии в синглетных S_0 , S_1 , S_2 и триплетных T_1 и T_2 электронных состояниях выделены жирными линиями. Каждое электронное состояние S_1 , S_2 , T_1 и T_2 расщеплено на серию колебательных подуровней, между

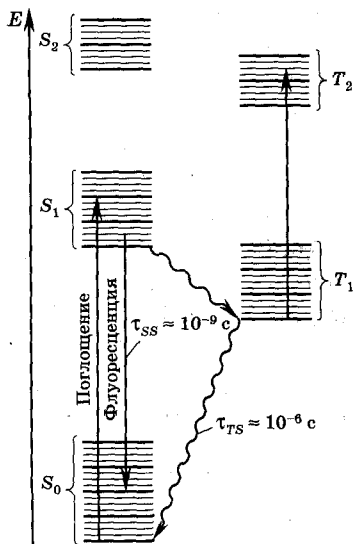


Рис. 20.13

которыми располагаются уровни вращательных энергетических состояний. В соответствии с квантово-механическими правилами отбора по спину оптические переходы разрешены между энергетическими состояниями с одинаковой мультиплетностью, т. е. в рассматриваемом случае синглет-синглетные ($S \rightarrow S$) (время жизни $\tau_{SS} \approx 10^{-9}$ с, см. рис. 20.13) и триплет-триплетные ($T \rightarrow T$). Синглет-триплетные и триплет-синглетные переходы (время жизни $\tau_{TS} \approx 10^{-6}$ с, см. рис. 20.13) запрещены: их вероятности на три порядка меньше, чем для разрешенных $S \rightarrow S$ и $T \rightarrow T$ переходов.

В равновесных условиях (при температуре $T = \text{const}$) молекулы красителя находятся в основном электронном состоянии S_0 с больцмановским распределением по колебательно-вращательным подуровням (в пределах энергии $\sim kT$). При облучении красителя светом от источника накачки происходит поглощение оптического излучения и молекулы переходят с уровней основного состояния на колебательно-вращательные уровни возбужденного состояния S_1 . Поскольку ширины энергетических полос состояний S_0 и S_1 велики, то спектр поглощения для этих переходов будет представлять достаточно широкую полосу. Для красителя родамин 6G этот спектр соответствует кривой 1 на рис. 20.14, где интенсивность поглощения характеризуется сечением поглощения фотона σ . С верхних энергетических подуровней состояния S_1 молекула очень быстро (за время $\sim 10^{-12}$ с) в результате релаксационных безызлучательных процессов переходит на нижние подуровни S_1 и затем из состояния S_1 в S_0 как излучательно, так и безызлучательно. Спектр спонтанной люминесценции (флуоресценции) красителя изображен кривой 2 на рис. 20.14, где Φ_ω — спектральная плотность светового потока,

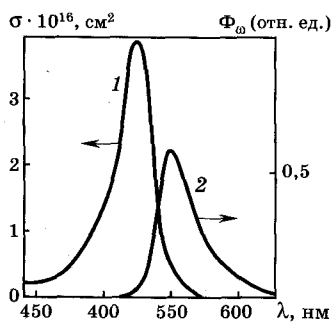


Рис. 20.14

вызванного флуоресценцией. Время жизни при излучательных переходах $S_1 \rightarrow S_0$ порядка 10^{-9} с. Спектр люминесценции представляет широкую полосу, смещенную в длинноволновую область относительно спектра поглощения.

Для получения инверсной населенности в лазерах на красителях используют оптическую накачку, создаваемую либо излучением импульсных газоразрядных ламп, либо дру-

гих лазеров. Так как основным достоинством лазеров на красителях является перестройка частоты в широком диапазоне, то для их накачки наиболее часто применяют излучение газовых лазеров различного типа, второй гармоники неодимового лазера, а также импульсных газоразрядных ламп и т. д.

Генерация в лазерах на органических красителях осуществляется по четырехуровневой схеме между уровнями S_1 и S_0 . Под воздействием интенсивной накачки достигается инверсия населенности между нижними подуровнями полосы S_1 и верхними подуровнями S_0 . Высоко расположенные подуровни S_0 не заселены: из-за тепловой релаксации их опустошение осуществляется весьма эффективно. Триpletные состояния T_1 и T_2 играют негативную роль в получении инверсии и, соответственно, лазерной генерации. С верхнего состояния S_1 молекулы красителя могут безызлучательно перейти в T_1 и не участвовать в индуцированном излучении, хотя вероятность этого перехода меньше вероятности лазерного. Состояние T_1 является метастабильным, т. е. переход $T_1 \rightarrow S_0$ запрещен, и поэтому молекулы накапливаются в этом состоянии. В большинстве красителей разрешенные переходы $T_1 \rightarrow T_2$ и $S_1 \rightarrow S_0$ близки по энергии, поэтому молекулы из состояния T_1 легко переходят в T_2 , поглощая энергию излучения лазерного перехода. Это паразитное поглощение может приводить к срыву генерации.

Спектр излучения жидкостного (как и любого другого) лазера определяется не только активными молекулами красителя, но и характеристиками резонатора. Для осуществления перестройки частоты лазера в широком диапазоне внутрь резонаторов вводят частотно-селективные элементы.

Вероятность оптических переходов $S_1 \rightarrow S_0$ велика, поэтому усиление для лазеров на красителях весьма значительно, примерно на два порядка выше, чем в твердотельных неодимовых и рубиновых лазерах, т. е. объем активной среды в жидкостных лазерах на красителях значительно меньше, чем в твердотельных.

Из-за интенсивной накачки и большой плотности лазерного излучения происходит быстрое разложение красителя. Чтобы избежать этого, необходимо осуществлять быструю прокачку активной жидкости, реализуя ее полную смену в активной области за время в несколько микросекунд.

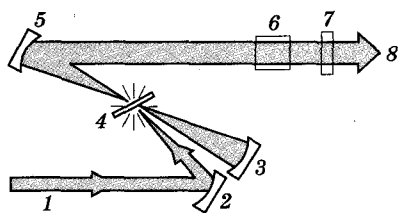


Рис. 20.15

Перестройку частоты излучения можно осуществлять за счет введения внутрь резонатора призм, дифракционных решеток, насыщающихся фильтров и других специальных частотно-селективных элементов. В качестве иллюстрации на рис. 20.15 приведена схема лазера

на красителе с ламинарным потоком. В этом лазере резонатор образован сферическим зеркалом 3 с высоким коэффициентом отражения ($\sim 98\%$) и плоским выходным зеркалом 7 с коэффициентом отражения порядка 95% . Вспомогательное зеркало 5 с коэффициентом отражения 100% служит для уменьшения габаритов лазера и обеспечивает сильную направленность выходного излучения 8. Излучение накачки 1 направляется на активную среду поворотным зеркалом 2. Поток красителя 4 направлен перпендикулярно к плоскости рисунка и для уменьшения потерь на отражение окно кюветы расположено под углом Брюстера к направлению излучения накачки. Селективный элемент 6 осуществляет перестройку длины волны лазера.

КПД лазеров на органических красителях при накачке лазерным излучением достигает $25\ldots 30\%$ и 1% при использовании в качестве накачки импульсных ламп. Мощность излучения в непрерывном режиме порядка нескольких Вт, а в импульсном — несколько МВт при длительности импульса порядка 20 нс и частоте повторения до 200 Гц, общий диапазон перекрытия длин волн — от $0,34$ до $1,17$ мкм при подборе различных красителей. Расходимость выходного лазерного излучения составляет $2\ldots 5$ мрад. Диапазон перестройки для каждого из красителей равен $40\ldots 80$ нм. Ширина спектра лазерного излучения ~ 10 нм при отсутствии диспергирующих элементов внутри резонатора.

20.6. Режимы работы и основные характеристики твердотельных лазеров

20.6.1. Режим свободной генерации.

При работе твердотельных лазеров реализуется два основных режима работы: *режим свободной генерации* и *режим модулированной добротности*, для реализации которого в лазер вводятся управляющие элементы.

Лазерное излучение в режиме свободной генерации часто носит нерегулярный характер с наличием серии релаксационных импульсов (пичков), положение которых во времени меняется случайным образом (пичковый режим). При нестационарном импульсном режиме генерации характерные времена пичков короче времени релаксации рабочих уровней активной среды. Рассмотрим качественную картину процесса нестационарной свободной генерации в твердотельном лазере. При использовании импульсной

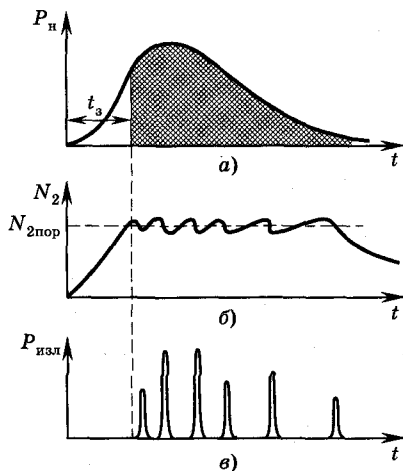


Рис. 20.16

газоразрядной лампы накачки энергия излучения накачки (P_n) изменяется во времени, как показано на рис. 20.16, а, и длительность импульса излучения составляет 0,1...1 мс.

После включения накачки под воздействием ее излучения населенность верхнего рабочего уровня возрастает и в некоторый момент достигает пороговой величины $N_{2\text{ пор}}$, при которой возникает генерация (рис. 20.16, б). Время задержки t_0 лазерной генерации относительно момента включения лампы накачки составляет величину порядка $10^{-5} \dots 10^{-4}$ с. Дальнейшее возрастание N_2 вызывает быстрый рост интенсивности электромагнитного поля на частоте генерации, что приводит к увеличению вероятности индуцированных переходов с верхнего рабочего лазерного уровня, которая прямо пропорциональна плотности излучения. В результате уменьшается населенность этого уровня N_2 и соответственно падает инверсия населенностей. Более подробно этот процесс выглядит следующим образом. В начальной стадии генерации, когда сразу же после времени запаздывания энергия индуцированного излучения мала, незначительно и число индуцированных переходов. Под действием энергии накачки продолжается накопление частиц на верхнем рабочем уровне. Поскольку интенсивность индуцированных переходов непрерывно растет, начиная с некоторого момента времени, поступление частиц на верхний уровень за счет накачки оказывается недостаточным из-за интенсивного индуцированного излучения.

Поэтому с этого момента N_2 начинает уменьшаться до тех пор, пока не снизится до $N_{2\text{ пор}}$ (см. рис. 20.16, б). При $N_2 < N_{2\text{ пор}}$ условие баланса амплитуд не будет выполняться и генерация соцвется. После этого вновь происходит накопление возбужденных частиц на верхнем рабочем уровне, и процесс повторяется. Возникают следующие пички, которые повторяются пока действует импульс излучения накачки. Как уже отмечалось, излучение ($P_{\text{изл}}$) твердотельных лазеров в свободном режиме генерации состоит из последовательности нерегулярных по амплитуде и временному положению импульсов — *пичков* (рис. 20.16, в). Рассмотренный процесс генерации часто наблюдается и во многих генераторах непрерывного режима. Причины возникновения незатухающих случайных колебаний могут быть самыми разнообразными. Они могут быть обусловлены нестабильностью параметров лазера, взаимодействием различных мод при многомодовой генерации, пространственной неоднородностью накачки и ее пульсациями, наличием неоднородностей в кристаллах и другими факторами. Длительность пичков составляет 0,1...10 мкс.

20.6.2. Режим модулированной добротности.

Режим модулированной добротности реализуется при импульсной работе лазера, когда за сравнительно большой промежуток времени за счет энергии накачки происходит значительное накопление частиц на верхнем рабочем уровне активного вещества, а излучение происходит в виде очень коротких импульсов с высокой пиковой мощностью («гигантских» лазерных импульсов).

Для предотвращения лазерной генерации в процессе накачки в резонатор вводятся элементы (*затворы*), управляющие его добротностью. Затворы по сути дела управляют моментами времени, когда начинают выполняться условия самовозбуждения лазера (баланса амплитуд и фаз). В качестве элементов, управляющих добротностью резонатора, используются механические, электрооптические и пассивные затворы. После того как населенность верхнего рабочего уровня достигнет достаточно больших значений, затвор выключается, добротность резонатора резко возрастает и реализуются условия, необходимые для развития мощного короткого импульса излучения. Лазеры с управляемой добротностью позволяют получить импульсы света с длительностью 10^{-7} ... 10^{-9} с и менее при пиковой мощности излучения до 10^{12} Вт.

Диаграммы, поясняющие процессы в лазере с модулированной добротностью, представлены на рис. 20.17, где показаны временные зависимости мощности источника накачки P_H (рис. 20.17, а), изменения добротности Q (рис. 20.17, б), населенности верхнего уровня рабочего перехода N_2 (рис. 20.17, в) и мощности излучения лазера $P_{изл}$ (рис. 20.17, г). За счет введения затвора к моменту включения источника накачки существенно увеличиваются потери в резонаторе, т. е. его добротность Q оказывается низкой, а поэтому пороговая населенность верхнего рабочего уровня $N_{2\text{ макс}}$ будет существенно превышать пороговую населенность $N_{2\text{ пор}}$ в отсутствие затвора. Таким образом, генерация не может возникнуть

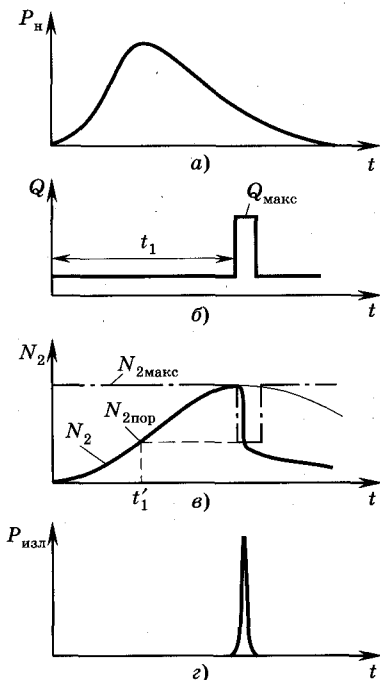


Рис. 20.17

при $t < t_1$ при данной мощности накачки P_H . В результате под действием источника накачки будет продолжаться накопление активных частиц на верхнем уровне. Через некоторое время t_1 (см. рис. 20.17, б) населенность N_2 достигнет максимального значения ($N_{2\text{ макс}}$) для данной мощности накачки P_H .

Приближенно можно считать, что вся энергия, которая в режиме свободной генерации излучалась пичками, к моменту t_1 будет накоплена в активном элементе.

В момент t_1 , когда $N_2 = N_{2\text{ макс}}$ затвор выключается, добротность Q резко возрастает до величины $Q_{\text{ макс}}$, причем при этой возросшей добротности $N_{2\text{ макс}} \gg N_{2\text{ пор}}$, где $N_{2\text{ пор}}$ — пороговая населенность верхнего лазерного уровня для возросшей добротности оптического резонатора. В результате с момента t_1 при выключении затвора происходит лавинообразное нарастание мощности лазерного излучения. Вся запасенная на рабочем лазерном переходе энергия излучается в одном коротком импульсе. Исследования показывают, что длительность импульса ла-

зерного излучения в режиме модулированной добротности примерно на два порядка меньше длительности каждого пичка при свободной генерации.

Таким образом, в режиме свободной генерации, когда в течение всего импульса накачки остается постоянной высокая добротность резонатора $Q = Q_{\text{макс}}$ (затвор отсутствует), лазерная генерация начинается как только инверсная населенность достигает порогового значения $N_{2\text{пор}}$ (например, в момент времени $t'_1 \ll t_1$, см. рис. 2.17, в). Поэтому для формирования каждого пичка свободной генерации используется существенно меньшая энергия накачки, чем в режиме модуляции добротности, вследствие чего эти импульсы имеют гораздо меньшую мощность и большую длительность. Режим модуляции добротности применяется не только при импульсной накачке, его можно использовать и при непрерывной накачке для получения последовательности коротких лазерных импульсов. Физические процессы будут такими же, как и при импульсной накачке. При включении затвора происходит накопление частиц на верхнем лазерном рабочем уровне (аналогично режиму модулированной добротности при импульсной накачке); после выключения затвора формируется мощный короткий импульс излучения, затем процессы повторяются.

В режиме модулированной добротности на характеристики излучения лазера существенное влияние оказывает скорость изменения добротности резонатора посредством управляющего элемента при переходе его из закрытого в открытое состояние. Экспериментальные исследования показывают, что при низкой скорости изменения добротности вместо одного импульса может излучаться пачка импульсов с убывающей во времени амплитудой. Такая структура лазерного излучения связана со свойствами индуцированных излучательных переходов. Многоимпульсная генерация в режиме модулированной добротности является нежелательным явлением из-за ухудшения характеристик лазерной системы. Для получения одноимпульсного излучения необходимо изменять добротность резонатора таким образом, чтобы она принимала установившиеся значения после окончания формирования первого импульса, т. е. управляющий элемент должен полностью изменять добротность резонатора за время, соизмеримое с длительностью генерируемого импульса ($\tau \approx 10^{-7} \dots 10^{-9}$ с).

Охарактеризуем наиболее типичные методы модуляции добротности лазеров.

Механические методы. Для реализации режима модуляции добротности используется метод, основанный на механическом вращении одного из зеркал резонатора. В механических затворах обычно используется не плоское или сферическое зеркало, а треугольная призма с полным внутренним отражением. Возбуждение колебаний в таком резонаторе возможно лишь в интервалы времени, когда угол между плоскостями зеркал не превышает нескольких угловых минут. Время переключения добротности 10^{-7} с можно реализовать при скоростях вращения призмы $\sim 30\,000$ об/мин. Включение лампы накачки синхронизируется с вращением зеркала. Разряд в лампе начинается за время t_1 до того, как зеркало придет в положение, соответствующее максимальной добротности резонатора. Использование призмы полного внутреннего отражения вместо плоского зеркала вызвано тем, что небольшие случайные колебания призмы вокруг горизонтальной оси не сказываются на работе лазера, что очень важно при эксплуатации такого затвора.

Электрооптические методы. Электрооптические затворы (модуляторы, Q-затворы) работают на основе использования эффекта двойного лучепреломления, возникающего под воздействием электрического поля в некоторых изотропных жидкостях (эффект Керра) и твердых кристаллах (эффект Погкельса). Существует несколько схем устройства такого типа Q-затворов. Одной из наиболее распространенных является схема, в которой внутри резонатора вдоль его оси между активной средой и зеркалом помещают электрооптический элемент Керра или элемент Погкельса, а за ним в непосредственной близости от зеркала — поляризатор. При подаче на электрооптический элемент постоянного напряжения определенной величины с вектором напряженности поля, направленным под углом в 45° по отношению к плоскости поляризации лазерного излучения, за счет эффекта двойного лучепреломления происходит суммарный поворот плоскости поляризации этого излучения на 90° . Поляризатор обычно ориентируют таким образом, чтобы он пропускал излучение свободной генерации (напряжение на электрооптическом элементе отсутствует), а излучение, поляризованное в перпендикулярной плоскости, задерживал, т. е. затвор закрыт при наличии соответствующего напряжения на электрооптической ячейке. За время, пока затвор закрыт, источник накачки создает наибольшую степень инверсной населенности в активном материале. После этого напряжение с электрооптического элемен-

та практически мгновенно снимается, добротность резонатора резко возрастает и возникает сверхмощный импульс излучения.

В качестве электрооптической среды в затворах широко применяются нитробензол (ячейка Керра) и кристаллы дигидрофосфата калия (KDP) и рубидия (RDP) (ячейка Погкельса). К достоинствам электрооптических затворов следует отнести высокое быстродействие ($\sim 10^{-10}$ с), высокую стойкость к воздействию мощного оптического излучения, отсутствие подвижных частей и возможность электрического управления.

Фототропные управляющие затворы. Такие затворы представляют собой нелинейную оптическую среду, для которой коэффициент поглощения существенно уменьшается при возрастании проходящего через нее излучения. Наиболее часто в этом случае применяют просветляющиеся красители. Эффект просветления определяется переходом поглощающих молекул красителя в возбужденное состояние и связанным с этим уменьшением показателя поглощения. Этот тип модуляторов является пассивным в отличие от первых двух.

20.6.3. Энергетические характеристики лазеров.

Энергия, излучаемая твердотельными лазерами, определяется энергией накачки и потерями в резонаторе, прежде всего коэффициентом пропускания полупрозрачного зеркала. Зависимости энергии излучения лазера на рубине и мощности лазера на алюмо-иттриевом гранате соответственно от энергии и мощности накачки представлены на рис. 20.18 (а — импульсный режим, б —

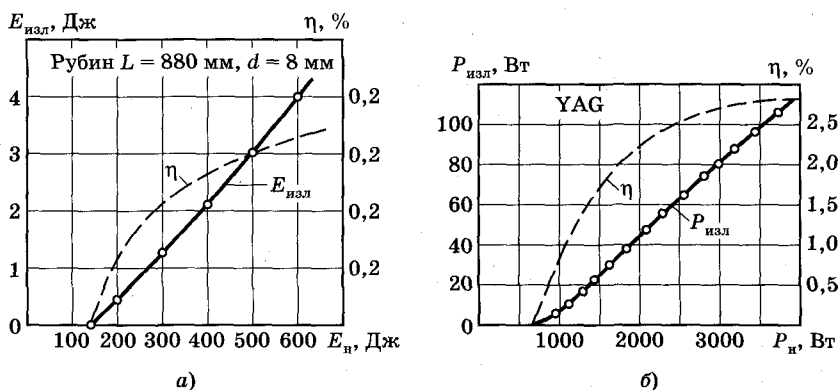


Рис. 20.18

непрерывный режим). Энергия излучения $E_{\text{изл}}$ для импульсной генерации (или мощность $P_{\text{изл}}$ для непрерывного режима) пропорциональна превышению энергии (мощности) накачки над пороговым значением. Практически линейная зависимость $E_{\text{изл}}$ (или $P_{\text{изл}}$) от энергии $E_{\text{н}}$ (или мощности $P_{\text{н}}$) источника накачки определяется свойствами индуцированных переходов, вероятность которых прямо пропорциональна плотности энергии электромагнитного поля.

Как следует из энергетических диаграмм для рубиновых и неодимовых лазеров, инверсная населенность на рабочих уровнях определяется заселенностью энергетических полос Z (см. рис. 20.9, 20.11). Энергетические полосы заселяются за счет индуцированного поглощения энергии накачки и, следовательно, их населенность прямо пропорциональна плотности энергии излучения лампы накачки. Таким образом, инверсная населенность на рабочем переходе и, следовательно, мощность (энергия) излучения лазера линейно связаны с мощностью (энергией) накачки. С использованием приведенных выше рассуждений и зависимости $P_{\text{изл}} = f(P_{\text{н}})$ КПД лазера, работающего в непрерывном режиме (см. рис. 20.18, б, штриховая кривая), равен

$$\eta = P_{\text{изл}}/P_{\text{н}} = C(P_{\text{н}} - P_{\text{н.пор}})/P_{\text{н}},$$

где $P_{\text{н.пор}}$ — пороговая мощность накачки, C — коэффициент, определяемый эффективностью источника накачки, отражателя, потерями в резонаторе и пропусканием полупрозрачного зеркала, через которое выходит лазерное излучение. Для лазеров импульсного режима поведение КПД определяется теми же факторами и полностью аналогично (см. рис. 20.18, а, штриховая кривая), только вместо значений мощностей в выражение для η необходимо поставить соответствующие энергии.

Рассмотрим характеристику $E_{\text{изл}} = f(1 - r_2)$, отражающую влияние коэффициента пропускания полупрозрачного зеркала $(1 - r_2)$ на энергию (или мощность) излучения $E_{\text{изл}}$ (рис. 20.19). С увеличением коэффициента пропускания зер-

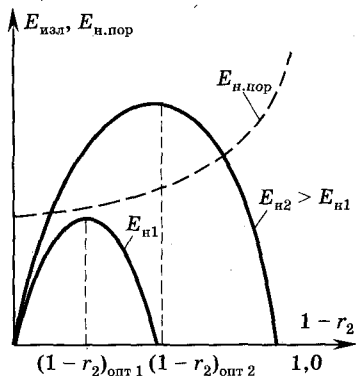


Рис. 20.19

кала $(1 - r_2)$ (r_2 — коэффициент отражения полупрозрачного зеркала) возрастают потери на излучение и пороговая энергия накачки $E_{н. пор}$, так как суммарные потери увеличиваются.

При $(1 - r_2) = 0$ энергия во внешнее пространство не выводится, она полностью остается в резонаторе. С ростом коэффициента пропускания до значения $(1 - r_2)_{опт}$, оптимального (для данной мощности накачки), энергия излучения $E_{изл}$ увеличивается за счет возрастания доли выводимой энергии. Снижение $E_{изл}$ при дальнейшем увеличении $(1 - r_2)$ обусловлено уменьшением плотности потока излучения в активной среде, а следовательно, и уменьшением вероятности индуцированных переходов, что приводит к уменьшению числа фотонов, генерируемых в активной среде. Очевидно, что с ростом пропускания $(1 - r_2)$, т. е. уменьшением добротности резонатора, увеличивается и пороговая энергия $E_{н. пор}$. Оптимальный коэффициент пропускания $(1 - r_2)_{опт}$ для большинства лазеров колеблется в пределах 20...60%, но для очень мощных лазеров он может достигать значений, близких к 90%. В последнем случае полупрозрачное зеркало вообще может отсутствовать. Его роль выполняет торцевая поверхность активного кристалла, отражающая 5...8% падающего на нее излучения.

На энергетические и спектральные характеристики твердотельных лазеров сильное влияние оказывает температурный режим активной среды. Изменение температуры активной среды влияет на пространственное распределение ионов в активном кристалле, что приводит к изменению внутреннего электрического поля, которое участвует в формировании энергетического спектра активных веществ. Можно отметить несколько эффектов, связанных с нагревом активного кристалла, которые существенно влияют на параметры лазеров.

1. Нарушается положение энергетических уровней, что вызывает изменение длины волны излучения.
2. Существенно ухудшается монохроматичность излучения из-за расширения контура спектральной линии люминесценции.
3. Максимальная частота повторения импульсов лазерного излучения ограничена вследствие нагрева активного элемента, что приводит к уменьшению средней мощности лазеров, работающих в режиме непрерывной накачки и модуляции добротности резонатора.

Для уменьшения влияния указанных факторов на параметры лазеров применяют специальные системы охлаждения активной среды, особенно в мощных лазерах.



Контрольные вопросы

1. Каковы особенности активной среды твердотельных и жидкостных лазеров?
2. Каким требованиям должна удовлетворять матрица активной среды твердотельных лазеров?
3. Какие системы накачки твердотельных лазеров являются наиболее эффективными и почему?
4. Механизм получения инверсии населенностей в рубиновом лазере и роль каждого из рабочих и вспомогательных энергетических уровней.
5. Каковы отличия и преимущества энергетической структуры неодимового лазера по сравнению с рубиновым?
6. Каковы процессы, определяющие режимы свободной генерации и модулированной добротности?
7. Основные характеристики твердотельных лазеров.
8. Характеристика и энергетическая структура активной среды жидкостных лазеров на органических красителях.
9. Устройство, особенности физических процессов, параметры и характеристики лазеров на органических красителях.

Глава 21

ПОЛУПРОВОДНИКОВЫЕ ИНЖЕКЦИОННЫЕ ЛАЗЕРЫ

21.1. Полупроводниковые материалы, используемые в источниках излучения

Полупроводниковые вещества как излучающие среды в приложении к светодиодам были частично рассмотрены в гл. 16. В этой главе эти вопросы применительно к полупроводниковым лазерам освещаются более подробно. Как уже отмечалось, все

полупроводники подразделяются на прямозонные и непрямозонные, в которых соответственно реализуются прямые и непрямые переходы (см. также п. 16.2). Физическая причина их различия определяется законом дисперсии $E = f(p)$, т. е. зависимостью энергии состояния E от квазиимпульса p .

В прямозонных полупроводниках (GaAs, InP, InSb и др.) максимум функции $E(p)$ в валентной зоне (ВЗ) и минимум $E(p)$ в зоне проводимости (ЗП) соответствуют одинаковому p . Достоинство прямозонных полупроводников — большая вероятность излучательного межзонного перехода.

В полупроводниках типа Ge, Si, SiC, GaP, AlAs (непрямозонные полупроводники) экстремумы зон смещены, поэтому переходы между ними сопровождаются большим изменением квазиимпульсов, причем последние превышают квазиимпульсы фотонов. По закону сохранения квазиимпульса излучательный переход без участия третьих тел в таких полупроводниках запрещен. В этом случае необходимо участие фононов, воспринимающих изменение квазиимпульса частиц при их межзонном переходе. При «непрямых» переходах излучательная рекомбинация свободных носителей тока, находящихся в энергетических состояниях на краях соответствующих зон, значительно меньше безызлучательной рекомбинации через примеси. Поэтому такие материалы редко используются для создания излучающих приборов.

Помимо упомянутых бинарных соединений в излучающих приборах большую роль играют взаимные растворы родственных соединений, включая и непрямозонные. Если смешиваемые соединения имеют одинаковый тип решетки, то они могут образовывать непрерывный ряд твердых растворов с постоянным изменением ширины запрещенной зоны ΔE_g . Тогда открывается возможность разработки приборов, которые в совокупности позволяют непрерывно перекрыть широкий диапазон спектра излучения. Простейший тип твердых растворов — трехкомпонентные соединения типа $\text{InP}_{1-x}\text{As}_x$, $\text{In}_x\text{Ga}_{1-x}\text{As}$ (группа АIII—BV) и большинство твердых растворов группы АIII—BVI ($\text{CdS}_x\text{Se}_{1-x}$, $\text{Zn}_x\text{Cd}_{1-x}\text{S}$, $\text{ZnSe}_x\text{Te}_{1-x}$ и др.). Электрические характеристики таких растворов слабо зависят от степени неупорядоченности кристалла.

Бинарные соединения обычно изготавливаются в виде крупных слитков, большинство твердых многокомпонентных растворов выращивается в виде эпитаксиальных слоев. Для лазе-

ров на основе гетеропереходов, работающих при повышенных температурах, важную роль играют соединения с одинаковыми периодами решетки (изопериодические пары). К ним относятся растворы, образуемые взаимным замещением галлия и алюминия, так как оба этих элемента имеют близкие ковалентные радиусы. В таких растворах, как AlGaAs, AlGaPAs, AlGaSb, период решетки остается почти постоянным при изменении соотношения алюминия и галлия, тогда как ширина запрещенной зоны E_g значительно изменяется. Многослойные гетероэпитаксиальные структуры на основе изопериодических составов AlGaAs позволили создать эффективные инжекционные лазеры (см. п. 21.2) в диапазоне длин волн 0,62...0,9 мкм.

Новые возможности создания изопериодических пар связаны с четырехкомпонентными твердыми растворами, например GaInPAs. Изопериодичность решетки в них достигается дозированным добавлением к InP примесей Ga и As, причем влияние обеих примесей на период решетки взаимно компенсируется.

Одно из основных требований, предъявляемых к материалам для светоизлучающих приборов, — излучение света в видимом участке спектра. Чувствительность человеческого глаза S_λ к различным длинам волн λ при постоянной энергии по спектру иллюстрируется кривой (рис. 21.1). Там же нанесены точки, соответствующие длине волны излучения некоторых полупроводниковых соединений.

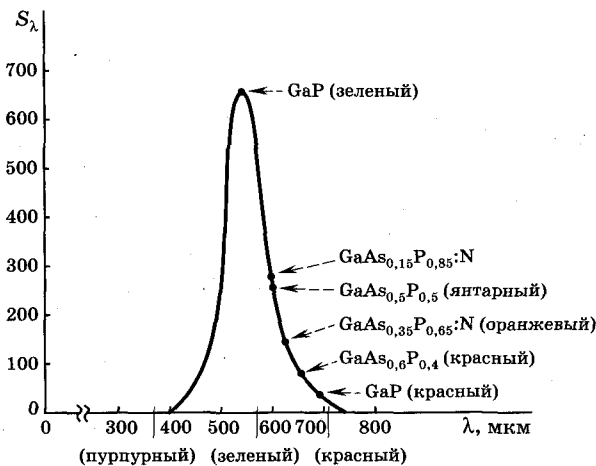


Рис. 21.1

21.2. Инжекционные полупроводниковые лазеры на основе гомопереходов

К достоинствам *инжекционных полупроводниковых лазеров* обычно относят малые размеры, возможность легкой интеграции с микроэлектронными устройствами, простоту реализации модуляции излучения, высокий КПД, а к недостаткам — широкий контур спектральной линии излучения, большой угол расходимости лазерного луча, относительно низкую временную когерентность, зависимость параметров от температуры.

В инжекционном полупроводниковом гомолазере активным элементом служит монокристалл арсенида галлия GaAs, в котором создан электронно-дырочный переход. В монокристалле арсенида галлия монокристалл GaAs, имеющий форму параллелепипеда с размером сторон в несколько десятых долей мм, укрепляется в массивном медном корпусе, служащем для отвода тепла. Две противоположные грани полупроводниковой пластины, перпендикулярные плоскости перехода, образуют оптический резонатор. Отражающие поверхности резонатора чаще всего получают путем скола кристалла вдоль кристаллографической плоскости, что обеспечивает получение идеально ровных и параллельных поверхностей. Ввиду значительного коэффициента преломления света на границе воздух—арсенид галлия коэффициент отражения от поверхностей резонатора составляет величину $\sim 0,3$, что достаточно для реализации условий самовозбуждения. Остальные четыре грани полупроводникового кристалла, чтобы исключить возникновение паразитных колебаний, делают шероховатыми.

В полупроводниковых лазерах используются монокристаллы GaAs с концентрацией примесей, превышающей концентрацию $n_{\text{выр}}$, при которой наступает вырождение полупроводника. Таким образом, в кристалле арсенида галлия электронно-дырочный переход образуется вырожденными n - и p -полупроводниками. Как известно, вследствие расщепления локальных уровней примесей в примесные зоны уровень Ферми в вырожденном n -полупроводнике располагается выше дна зоны проводимости, а в p -полупроводнике — ниже потолка валентной зоны.

Энергетическая диаграмма электронно-дырочного перехода на границе вырожденных полупроводников показана на рис. 21.2.

Энергетические уровни в зоне проводимости n -полупроводника вблизи (выше) уровня $E_{\text{д}}$ заняты электронами, концентрация ко-

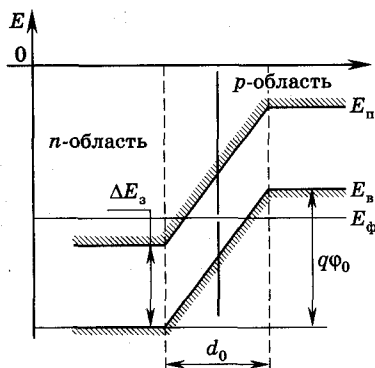


Рис. 21.2

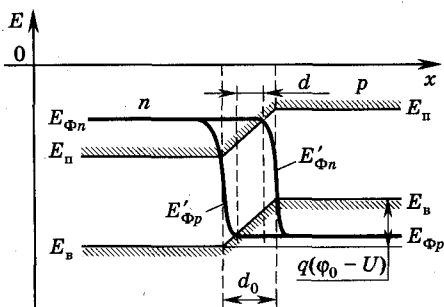


Рис. 21.3

торых убывает по мере удаления от уровня E_n . Для простоты дальнейших рассуждений будем считать, что уровни с энергией $E_n \leq E \leq E_\phi$ заняты электронами, а уровни с энергией $E > E_\phi$ свободны. Аналогично будем полагать, что в валентной зоне p -полупроводника уровни с энергией $E_v \geq E \geq E_\phi$ свободны, а энергетические состояния с энергией $E < E_\phi$ заняты электронами.

При подключении к p — n -переходу прямого внешнего напряжения U его энергетическая диаграмма изменяется (рис. 21.3). Запирающий слой (шириной d_0 при $U = 0$, см. рис. 21.2) сужается до ширины d ; потенциальный барьер уменьшается до величины $q(\phi_0 - U)$, а уровень Ферми в области запирающего слоя расщепляется на квазиуровни E'_{fn} и E'_{fp} . Вычисления показывают, что квазиуровни Ферми располагаются на энергетической диаграмме так, как это показано на рис. 21.3. Условия межзонной рекомбинации выполняются только в пределах узкой полоски внутри запирающего слоя, где уровни в зоне проводимости между состояниями E'_{fn} и E_n заняты электронами, а уровни в валентной зоне между состояниями E'_{fp} и E_v свободны. Слева от этой полоски все состояния в валентной зоне несвободны, а справа от этой полоски отсутствуют свободные электроны в зоне проводимости. Таким образом, акты межзонной рекомбинации могут происходить только в пределах слоя d' .

Как отмечалось выше, для энергетической диаграммы GaAs $E = f(p)$ характерно расположение минимума энергии в зоне проводимости и максимума энергии в валентной зоне при одном

и том же значении импульса электронов p . В таких веществах межзонная рекомбинация происходит в виде прямых излучательных переходов (при неизменном значении импульса частицы). Большинство таких переходов сопровождается излучением фотона с частотой, соответствующей разности энергий в зоне проводимости и в валентной зоне.

Допустим, что такой переход произошел спонтанно. Излученный при этом фотон при неоднократных отражениях от параллельных плоскостей кристалла, образующих оптический резонатор, может индуцировать последующие акты межзонных переходов частиц. Среди индуцированных переходов будут преобладать переходы вниз, сопровождающиеся излучением, если в пределах слоя d' будет существовать инверсная населенность. Состояние инверсной населенности достигается путем энергетической накачки: при подаче положительного смещения через переход течет ток инжекции и большинство энергетических уровней в зоне проводимости заселяются электронами, а в валентной зоне большая часть уровней оказывается свободной.

Вновь излученные в результате электронных переходов фотоны индуцируют новые излучательные переходы, и если энергия индуцированного излучения превзойдет по величине энергию потерь, то установится режим генерации. Энергия индуцированного излучения естественно зависит от числа активных частиц, т. е. от концентрации электронов в зоне проводимости, которая, в свою очередь, определяется величиной напряжения смещения или, иначе говоря, от тока инжекции. Таким образом, режим генерации устанавливается при определенном значении тока инжекции, называемого пороговым током.

Часто *полупроводниковый квантовый генератор (ПКГ)* работает в импульсном режиме. Энергетическая накачка для создания инверсной населенности осуществляется импульсами тока, поступающими к ПКГ от источника питания. Рассмотрим некоторые характеристики полупроводниковых инжекционных лазеров на гомопереходах.

Зависимость выходной мощности излучения от мощности накачки. При относительно небольших значениях мощности накачки P_n (тока инжекции) число активных частиц, а следовательно, и мощность излучения, растет практически линейно в зависимости от величины P_n . С дальнейшим ростом мощности накачки наблюдается быстрое уменьшение мощности излучения, что объясняет

ся разогревом диода. При увеличении температуры плотность пологового тока, при котором начинается генерация, возрастает. Это связано с изменением функций распределения электронов и дырок в энергетических зонах. При повышении температуры из-за смещения уровня Ферми к границам зоны проводимости и валентной зоны уменьшается вероятность заполнения электронами верхних рабочих уровней в окрестности зоны проводимости, с которых происходят индуцированные переходы в валентную зону. Одновременно увеличивается заполнение электронами нижних рабочих энергетических уровней вблизи потолка валентной зоны. В результате этих процессов снижается инверсия населенностей и выходная мощность падает.

На величину выходной мощности на средней длине волны излучения лазера влияет также изменение размеров резонатора, происходящее вследствие линейного расширения кристалла полупроводника при повышении температуры.

Рассмотрим распределение энергии излучения по поперечному сечению луча и плоский угол расхождения луча в горизонтальной и вертикальной плоскостях. В инжекционных лазерах расходимость луча зависит от материала, конструкции диодной структуры и ряда других факторов. Однако основной причиной, определяющей диаграмму направленности излучения, является дифракционная расходимость. Она определяется толщиной активного слоя инжекционного лазера. Оптическую модель полупроводникового лазера можно представить в виде активного тонкого плоского слоя толщиной d , заключенного между двумя пассивными областями (рис. 21.4, а).

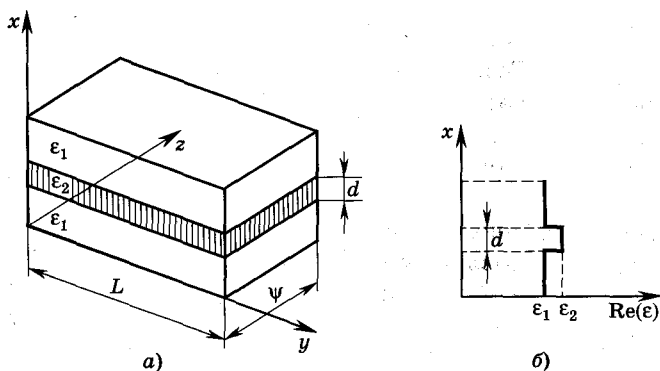


Рис. 21.4

Комплексная диэлектрическая проницаемость активной и пассивной областей различна и на границе указанных областей имеется скачок диэлектрической проницаемости (рис. 21.4, б). Действительная часть диэлектрической проницаемости активного слоя, как правило, больше, чем в окружающих ее пассивных областях. В результате происходит полное внутреннее отражение электромагнитных волн на границах активного слоя и канализация потока излучения в нем. Вне активного слоя имеются только волны поверхностного типа. Таким образом, диод подобен открытому диэлектрическому волноводу, характеристики которого и определяют диаграмму направленности излучения. При относительно большой ширине d активного слоя возбуждаются не только продольные, но и поперечные типы колебаний высоких порядков вдоль оси x , диаграмма направленности которых имеет много лепестков сравнимой интенсивности. Поперечные типы колебаний по оси x формируют структуру луча ПКГ в вертикальной плоскости, т. е. в поперечном плоскости перехода сечения (рис. 21.5). При малой толщине диэлектрического волновода (активной области) происходит подавление поперечных типов колебаний по оси x .

Если в активной области существует единственный низший тип колебаний, то угол расходимости α в вертикальной плоскости стремится к дифракционному пределу. Распределение интенсивности излучения в пучке зависит от угла α , который с хорошим приближением описывается выражением

$$\alpha \approx 2 \arctg [(2/kb)^2 + (b/R_0)^2]^{1/2}, \quad (21.1)$$

где $b \approx d$ — ширина излучающего пятна; R_0 — радиус кривизны волнового фронта на зеркале; $k = 2\pi/\lambda$ — волновой вектор. Как правило, кривизной волнового фронта можно пренебречь, и тогда ось главного лепестка диаграммы направленности совпадает с осью резонатора, а угол расхождения луча можно определить по формуле

$$\alpha = 2 \arctg (2/kd). \quad (21.2)$$

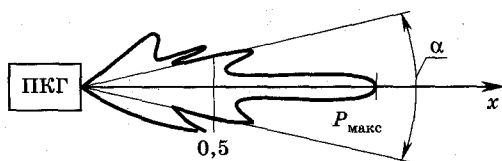


Рис. 21.5

21.3. Инжекционные лазеры на основе гетеропереходов

В лазерных диодах на гетеропереходах инжектированные носители, собранные в узкой активной области, могут создать инверсию населенностей при низких плотностях тока. Гетеропереходы позволяют получить лазерную генерацию в полупроводниках с непрямыми переходами. Наибольшее распространение в инжекционных лазерах на гетеропереходах получили соединения на основе арсенида галлия GaAs. Чтобы создать гетеропереходы на основе этого материала, необходим полупроводник с широкой запрещенной зоной, с потенциальными барьерами и коэффициентами отражений, ограничивающими как потоки носителей, так и фотонов. Таким условиям удовлетворяет $Al_xGa_{1-x}As$. Поскольку постоянные решетки AlAs и GaAs очень хорошо согласуются, смежная область гетероперехода имеет низкую плотность дефектов, что уменьшает скорость рекомбинации. Коэффициент отражения также изменяется незначительно. При содержании алюминия на уровне $x = 0,2$ коэффициент преломления составляет 3,27 по сравнению с 3,43 для чистого GaAs, что обеспечивает хорошие волноводные свойства этого материала. Кристаллы Ga и Al имеют одинаковый тип решетки, практически одинаковые периоды решетки и ковалентные радиусы, равные 1,26 Å. Замещение одного из этих элементов другим в гомополярных кристаллах происходит практически без изменения периода решетки. Вследствие различия температурного коэффициента расширения GaAs и AlAs полное совпадение их решеток имеет место при высокой температуре. Следовательно, выращивание гетероперехода при высоких температурах наиболее благоприятно и осуществляется практически без образования дефектов роста кристалла.

В лазерных диодах на гетеропереходах толщину области, в которой создается инверсия населенностей, удается легко ограничить с помощью создания широкозонного p -гетероперехода рядом с $p-n$ -переходом на основе более узкозонного материала. Если в гетеропереходе имеется барьер для электронов в зоне проводимости с высотой, заметно большей kT , то диффузии электронов в глубь p -области не происходит. Электроны отражаются барьером назад в активный слой, а не рекомбинируют безызлучательно, как это обычно имеет место на поверхности полупроводникового кристалла.

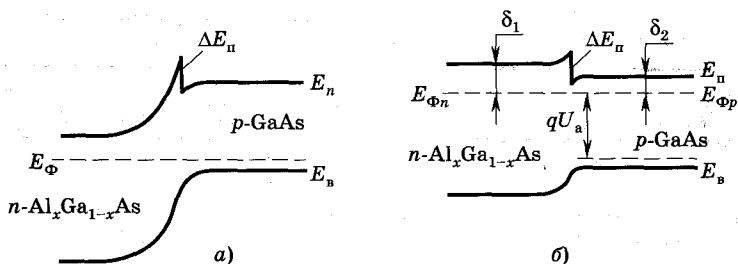


Рис. 21.6

Скачок ширины запрещенной зоны в гетеропереходах, образуемых GaAs и $\text{Al}_x\text{Ga}_{1-x}\text{As}$, происходит главным образом за счет разрыва края зоны проводимости. Так, например, для гетероперехода $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}-\text{GaAs}$ скачок разрыва зоны проводимости ΔE_n составляет 0,4 эВ (рис. 21.6, а) при ширине запрещенной зоны $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$, равной 1,85 эВ, а скачок разрыва валентной зоны ΔE_v пренебрежимо мал и не отражен на рис. 21.6, а.

В лазерах, изготовленных на основе односторонних гетероструктур (ОГС) типа $n\text{-GaAs}-p\text{-GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$, невозможно использовать тонкие активные слои (толщиной $d \sim 1 \dots 1,5$ мкм). В таких гетероструктурах с уменьшением d резко возрастает пороговый ток и генерация становится недостижимой. Критическая толщина активного слоя зависит от температуры так, что с ростом температуры происходит срыв эффективной генерации. Двусторонние гетероструктуры (ДГС) свободны от этого недостатка. Они имеют узкозонный активный слой, с обеих сторон заключенный между широкозонными слоями (например, структура $n\text{-Al}_{x_1}\text{Ga}_{1-x_1}\text{As}/p\text{-GaAs}/p^+\text{-Al}_{x_2}\text{Ga}_{1-x_2}\text{As}$, рис. 21.7, а'). В этих структурах диффузия инжектированных носителей тока ограничивается с помощью $p-p^+$ -гетероперехода и, кроме того, имеется инжектирующий $p-n$ -гетеропереход. Разрыв края зоны проводимости в гетеропереходах приводит к однонаправленному характеру инжекции носителей. Если эмиттер является широкозонным с проводимостью n -типа, то при прямом смещении уменьшается потенциальный барьер для электронов, инжектируемых в узкозонную p -область. В этом случае инжекция электронов в узкозонную активную p -область происходит независимо от концентрации доноров в эмиттерной области. В инжектирующих гетеропереходах в области инжекции может быть получено динамическое вырождение при невырожденной области

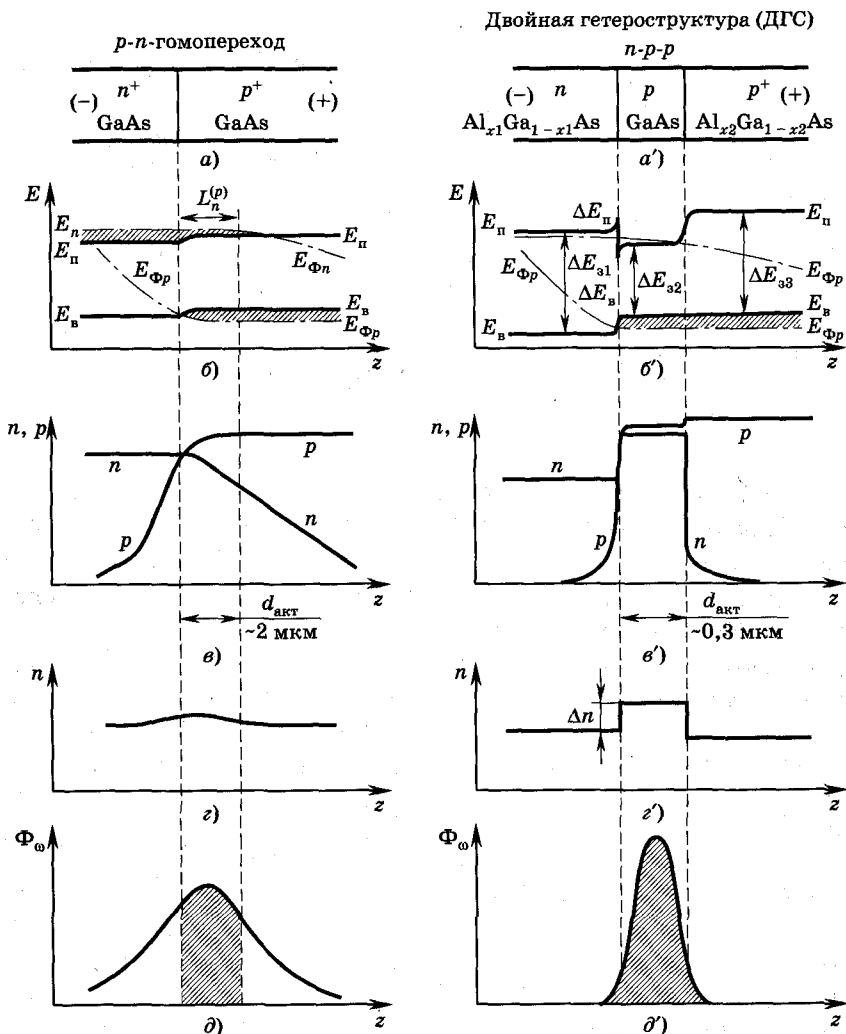


Рис. 21.7

эмиттера, что связано со значительным превышением числа неравновесных носителей тока над равновесными во время инжекции. Такой процесс называется суперинжекцией. Рис. 21.6, б иллюстрирует этот процесс. Расстояние уровня Ферми в области $n\text{-Al}_x\text{Ga}_{1-x}\text{As}$ от дна зоны проводимости δ_1 больше, чем δ_2 в области $p\text{-GaAs}$ (см. рис. 21.6, б). Это означает, что плотность электронов, инжектированных в GaAs, будет больше плотности

электронов, поступающих в $n\text{-Al}_x\text{Ga}_{1-x}\text{As}$ из $p\text{-GaAs}$. Энергетическая диаграмма для двойной гетероструктуры (ДГС) типа $n\text{-Al}_x\text{Ga}_{1-x}\text{As}/p\text{-GaAs}/p\text{-Al}_x\text{Ga}_{1-x}\text{As}$ может быть получена зеркальным отражением рисунка 21.6, б относительно вертикальной линии. Сравнительные характеристики, иллюстрирующие процессы получения инверсной населенности, распределения носителей n и p вдоль структур, изменение коэффициента преломления n и интенсивности света Φ_0 в области переходов и соседних областях приведены на рис. 21.7.

Из-за различия диэлектрических проницаемостей активного слоя и окружающих его слоев возникает скачок диэлектрической проницаемости, что приводит к образованию волновода (световода). В ДГС диэлектрический волновод является симметричным, т. е. локализованные волны существуют при любой сколь угодно малой толщине. При малых толщинах диэлектрического волновода усиление, развиваемое в активном слое, используется наиболее эффективно. В ДГС реализуются также наиболее низкие значения пороговой плотности тока при комнатной температуре, что связано с отсутствием межзонного поглощения в пассивных областях.

В гомоструктуре лазера ограничивающие (задерживающие) барьеры крайне малы. Распространение света по обе стороны активной области происходит достаточно интенсивно, что вызывает значительные потери энергии излучения.

В лазере на ОГС существует потенциальный барьер величиной $\sim 0,4$ эВ, ограничивающий движение электронов. Кроме того, коэффициент преломления уменьшается примерно на 5% при прохождении света от $p\text{-GaAs}$ к $p\text{-Al}_x\text{Ga}_{1-x}\text{As}$ (рис. 21.8, б). В результате происходит ограничение распространения света на этой межфазной границе перехода. Однако на межфазном $n\text{-}p\text{-GaAs}$ -переходе наблюдается только незначительное изменение коэффициента преломления и существенное число фотонов теряется в $n\text{-GaAs}$ -области. При большом смещении $p\text{-GaAs}$ -область становится тонкой, имеет место инжекция дырок в $n\text{-GaAs}$ -область и эффективность работы лазера возрастает.

В рассмотренной ДГС происходит пространственное ограничение распространения как носителей тока, так и фотонов. Сравнительные характеристики лазерных структур на гомопереходах, с одним гетеропереходом и с двумя гетеропереходами (с указанием характерных размеров) приведены на рис. 21.8,

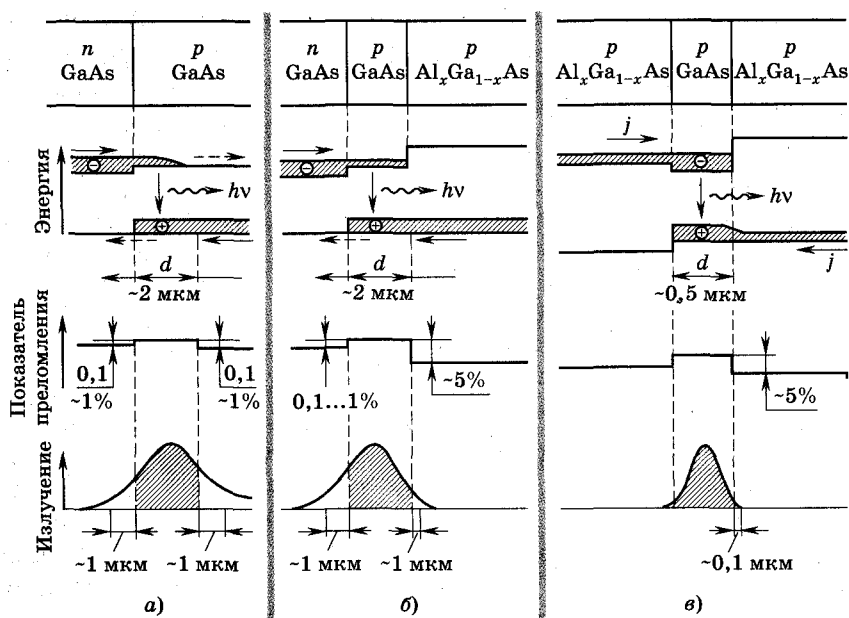


Рис. 21.8

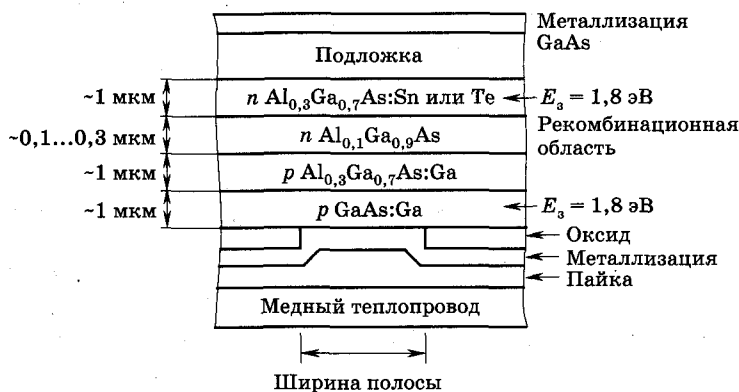


Рис. 21.9

где *a* — гомоструктура; *б* — ОГС; *в* — ДГС. Структура лазера на ДГС показана на рис. 21.9. Резонатор типа Фабри—Перо формируется за счет скола противоположащих граней.

Для получения оптического усиления в полупроводниковых лазерах, как и в лазерах на других активных средах, энергия индуцированного излучения должна превышать энергию потерь.

Поэтому существует некоторая пороговая величина тока через переход. Зависимость плотности порогового тока от потерь может быть определена следующим соотношением:

$$j_{\text{пор}} = [1/(L\alpha)][\beta L + \ln(1/R)], \quad (21.3)$$

где α — коэффициент усиления на единицу длины и единицу плотности тока, β — внутренние потери на единицу длины, R — коэффициент отражения. На рис. 21.10 даны зависимости плотности порогового тока от обратной длины резонатора L^{-1} для инжекционных гомо- (кривые 1, 2) и ДГС-лазеров (кривые 3, 4).

Обычно размеры лазеров лежат в пределах 200...500 мкм без учета выводов и корпуса. Спектр излучения гетеролазеров формируется большим количеством как продольных, так и поперечных мод. Диаграмма направленности излучения в дальней зоне зависит от числа мод резонатора и дифракционных ограничений для прямоугольной активной области. Для инжекционных лазеров на основе ДГС угол расходимости луча достигает 40° , вместо $10...15^\circ$ для гомолазеров и лазеров на ОГС. Поэтому задача коллимации лазерного луча для лазеров на ДГС более трудна. Лазерный луч от ДГС в основном поляризован таким образом, что вектор электрического поля излучения параллелен плоскости гетероперехода. Это соответствует ТЕ-модам внутри лазерного резонатора. В лазере на односторонней гетероструктуре излучение формируется главным образом ТЕМ-модами и не имеет хорошо выраженной поляризации. В инжекционных лазерах наблюдается резкое увеличение мощности излучения при увеличении тока, начиная с порогового значения, величина которого может снизиться на порядок при понижении температуры до 77 К.

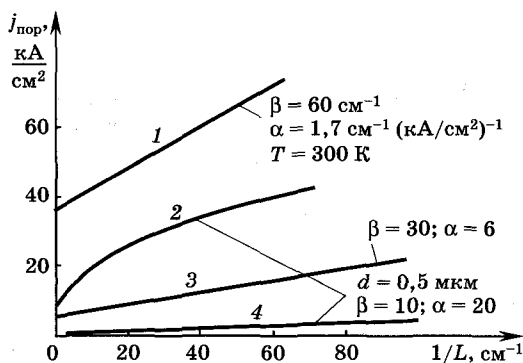


Рис. 21.10

21.4. Разновидности полупроводниковых лазеров

Полосковые лазеры. Было исследовано большое число конструкций лазеров, выполненных на основе полосковой геометрии. Одной из наиболее эффективных оказалась мезаполосковая структура, представленная на рис. 21.11. Приборы этого типа позволяют реализовать одномодовый режим работы при пороговом токе менее 50 мА. Размеры структуры: ширина — 6 мкм, длина — 100 мкм, толщина активной области — 0,5 мкм. Однако пороговая плотность тока — 6 кА/см² больше по сравнению с 1 кА/см² для структуры с большей шириной.

Гетеролазеры с распределенной обратной связью. Наиболее эффективным способом формирования требуемой оптической обратной связи является использование обратного рассеяния Брэгга при периодическом изменении коэффициента преломления и (или) усиления в лазерном волноводе. Такие изменения (возмущения) могут быть получены за счет травления периодической рифленой структуры на границе между активным *p*-GaAs-слоем и *p*-GaAlAs-слоем в односторонней гетероструктуре (рис. 21.12). Решетчатая структура получается за счет интерферометрической экспозиции фоторезиста на поверхности *p*-GaAs с последующим полным травлением через полученный фоторезист и наращиванием слоя AlGaAs. Период травления λ выбирается из условия

$$\lambda = m\pi/\beta \approx m\lambda_0/2n, \quad (21.4)$$

где λ_0 — длина волны излучения лазера в свободном пространстве; β — постоянная распространения в волноводе; $m = 1, 2, 3, \dots$ — целое число; n — коэффициент преломления волноводного материала.

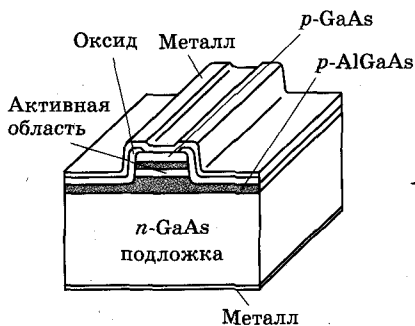


Рис. 21.11

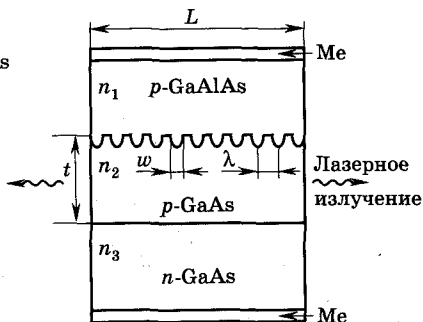


Рис. 21.12

Длина волны излучения GaAs/AlGaAs-лазеров при наличии распределенной обратной связи в решетчатой структуре изменяется приблизительно на $0,5 \text{ \AA}/\text{K}$ из-за изменения коэффициента преломления в зависимости от температуры. Для гомолазеров с резонаторами Фабри—Перо изменения длины волны от температуры составляют величину $\sim 4 \text{ \AA}/\text{K}$ и вызваны изменением ширины запрещенной зоны полупроводника. Из других особенностей лазеров с распределенной обратной связью следует отметить волновую селективность, возможность контроля продольных мод и хорошо сколлимированные выходные лучи.

В лазерах с распределенной обратной связью можно осуществить согласование оптических эффектов, обусловленных периодической структурой и резонаторами типа Фабри—Перо. Физический механизм такого согласования поясняется на рис. 21.13. По законам геометрической оптики лучи от выступов будут рассеиваться в фазе, если дополнительное расстояние $\lambda + b$ для лучей 1...4 при рассеивании от соседних выступов будет кратно длине волны света в материале (λ_0/n), т. е.

$$b + \lambda = m(\lambda_0/n), \quad (21.5)$$

причем b отсчитывается от плоского фронта волны (ПФВ), а λ — период решетки Брэгга. Поскольку $b = \lambda \sin \theta$, то

$$\sin \theta = (m\lambda_0/n\lambda) - 1. \quad (21.6)$$

Если $\lambda_0/n = \lambda$, тогда $\theta = 0$ при $m = 1$, волновой фронт будет параллелен p — n -переходу лазера и лучи будут ортогональны. При

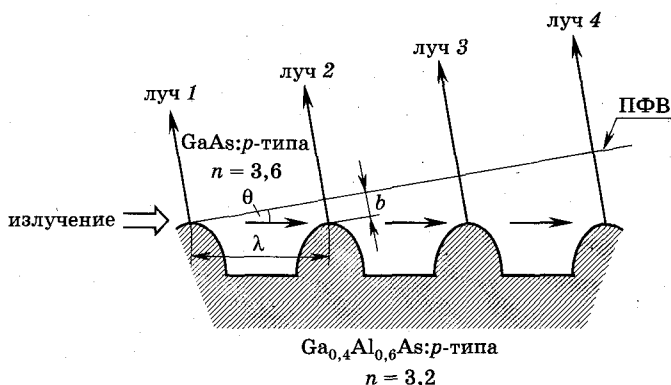


Рис. 21.13

$m = 0$ решение (21.6) будет описывать рассеяние света в прямом направлении, а при $m = 2$ — в обратном направлении. Как прямая, так и обратные волны рассеиваются, распространяясь как выше, так и ниже рифленой поверхности. Возможные углы лучей в воздухе после отражений определяются из соотношения

$$\sin \Phi = m\lambda_0/\lambda - n, \quad (21.7)$$

где Φ — угол волнового фронта, измеренный относительно нормали. Поскольку $\sin \Phi \leq 1$ и $n = 3,6$ для GaAs и AlGaAs, во внешнем пространстве будут видны только лучи внутри конуса с углом 16° . Оптимальное расположение поверхности рифления при распределенной обратной связи ортогонально переходу и определяется соотношением

$$\lambda/(\lambda_0/n) = p, \text{ где } p = 1, 2, 3, \dots \quad (21.8)$$

Формула (21.8) определяет брэгговское рассеяние порядка $2p$ в структуре с распределенной обратной связью. При $p = 1$ существует только излучательная мода, нормальная к плоскости рифления. Величина расходимости выходного луча вдоль направления рифления $\sim 0,35^\circ$ и может быть вычислена по формуле (21.7). Дифференцируя это выражение, получаем

$$d\Phi = [(m\lambda_0/(n\lambda) - (\lambda_0 dn/(n d\lambda))(\Delta\lambda n/\lambda_0)]/\cos \Phi.$$

При $\Delta\lambda = 6 \text{ \AA}$ угол расходимости составляет $0,21^\circ$. Расчеты показывают, что расходимость в ортогональном плоскости рифления направлении составляет 10° . Это хорошо согласуется с измеряемой диаграммой направленности в дальней зоне. Выходное излучение полностью поляризовано (вектор электрического поля волны параллелен плоскости рифления). Уровень безызлучательной рекомбинации в переходной области достаточно высок даже при получении рифления травлением. Поэтому часто используют структуру с разделением активной и волноводной областей (рис. 21.14).

Для уменьшения рабочего тока в ДГС-лазерах необходимо уменьшать толщину активной области d (см. рис. 21.8). Однако, если размер активной области становится меньше длины волны, в ней резко возрастают оптические поте-

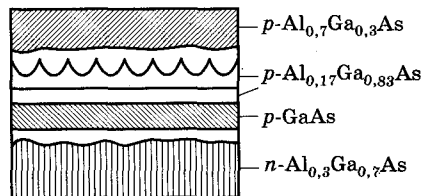


Рис. 21.14

ри, обусловленные просачиванием света в соседние с активной низкоомные области и ухудшением волноводных свойств активной среды. Чтобы снизить влияние этих эффектов, необходимо разделить области электронного и оптического ограничения. На рис. 21.14 сверхтонкий активный слой p -GaAs толщиной $d \ll \lambda$ размещается внутри более толстых слоев.

При подаче прямого напряжения электроны и дырки из широкозонных n - и p -областей инжектируются в прилегающие к ним области, так же, как это происходит в обычной ДГС-структуре. Профили изменения концентрации в p -областях подбираются таким образом, чтобы основная доля инжектированных носителей, создающих инверсию населенностей, была сосредоточена внутри слоя p -GaAs, а оптический волновод формируется на границах p -GaAs и p -GaAlAs.

На рис. 21.14 области преимущественного распространения света обозначены светлыми полосками рядом с p -GaAs.

В современных лазерах с отдельным оптическим и электронным ограничением активная область, где происходят накопление и излучательная рекомбинация неравновесных электронов и дырок, представляет собой **квантоворазмерные структуры** (см. гл. 10), в которых функция плотности энергетических состояний имеет ступенчатый вид, что уменьшает тепловое размытие в распределении носителей по энергии и облегчает достижение инверсной населенности. Максимальный коэффициент усиления возрастает, а температурная зависимость пороговой плотности ослабевает.

Для формирования квантоворазмерных структур слои должны быть тонкими, причем каждая из областей n - и p -гетероструктуры на рис. 21.14 является слоистой с различной концентрацией и различными соотношениями примесей Al и Ga.

Необходимость изготовления сверхтонких слоев ($\sim 0,01 \dots 0,1 \dots 1,0$ мкм) для создания активных структур в лазерах с отдельным ограничением тока накачки и излучаемого света предъявляет очень высокие требования к технологии изготовления и к контрольно-измерительной технике. По сути дела, необходимо контролировать структуры с точностью до одного атомного слоя. Технологии, позволяющие реализовать такие параметры, называются **нанотехнологиями** (см. гл. 10).

В заключение отметим, что использование квантоворазмерных структур в активной области ДГС-лазеров с отдельным ограничением дает возможность: уменьшить пороговую плотность

тока накачки, ослабить влияние температуры, увеличить коэффициент усиления активной среды на единицу длины, улучшить спектральные характеристики.



Контрольные вопросы

1. Какие преимущества имеют полупроводниковые лазеры по сравнению с лазерами другого типа?
2. Какие материалы используются в светоизлучающих полупроводниковых приборах и каковы характеристики этих материалов?
3. Почему в полупроводниковых лазерах используются вырожденные полупроводники и каков механизм получения инверсной населенности в них?
4. Параметры и характеристики полупроводниковых инжекционных гомолазеров.
5. Охарактеризуйте преимущества лазеров на основе ОГС и ДГС, механизмы получения инверсии населенностей в этих типах лазеров.
6. Энергетические диаграммы, характеристики и параметры лазеров на ОГС и ДГС.
7. Каковы особенности и характеристики ДГС-лазеров с распределенной обратной связью?
8. Разновидности полупроводниковых гетеролазеров, их особенности и параметры.

ВОПРОСЫ ПРИМЕНЕНИЯ
И ЭКСПЛУАТАЦИИ
ЭЛЕКТРОННЫХ ПРИБОРОВ

Глава 22

ШУМЫ ЭЛЕКТРОННЫХ ПРИБОРОВ

22.1. Общие сведения

Физические процессы, которые были рассмотрены при изучении электронных приборов, считались детерминированными, хотя на самом деле на эти процессы оказывают влияние различные случайные факторы, вызывающие флуктуации (случайные изменения) токов и напряжений в цепях приборов. Флуктуации напряжения (или тока) на выводах электронных приборов и устройств обычно называются *шумами*. Шумы внутренне присущи всем электронным приборам, и они обусловлены случайным характером движения (на микроскопическом уровне) носителей заряда внутри электронных приборов. Шумы часто являются фактором, накладывающим ограничения на характеристики электронных устройств, в частности, на их чувствительность. Иногда шум можно использовать как средство для исследования электрических характеристик самой системы.

Обычно шумы в электронных приборах рассматриваются как стационарные случайные процессы во временной или частотной области, поэтому основными характеристиками шумов являются: среднее значение реализации этого случайного процесса (математическое ожидание); спектральная плотность, дающая среднюю спектральную составляющую флуктуирующего сигнала;

дисперсия, характеризующая энергию шума с математическим ожиданием (средним значением), равным нулю; автокорреляционная функция, определяющая статистическую связь между двумя значениями случайной функции $x(t)$, разделенными временным интервалом τ , и ряд других характеристик, смысл и значение которых будут даны по мере их использования.

22.2. Источники шумов

К основным видам шумов, которые проявляются в электронных приборах, относятся: тепловой, диффузионный, дробовой, генерационно-рекомбинационный, $1/f$ -шум, шум токораспределения, шум вторичной эмиссии, лавинного умножения, взрывной, ионный и некоторые другие.

Тепловой шум возникает из-за случайных флуктуаций скорости носителей заряда в среде (в приборе). Флуктуации скорости носителей, в свою очередь, вызывают случайные изменения плотности носителей и, как следствие этих процессов, появляются флуктуации токов и напряжений. Тепловые флуктуации не нарушают теплового равновесия системы, так как за случайным (микроскопическим) отклонением какого-либо электрофизического параметра от равновесного состояния следует (в среднем) возвращение к нему. Большое число микроскопических событий такого рода ведет к заметному изменению тока или напряжения на клеммах прибора, резистора и т. д. (рис. 22.1).

Дробовой шум связан с прохождением свободных носителей заряда через потенциальный барьер. Впервые он был обнаружен в электронных лампах. Название этого вида шума связано с тем, что звук от динамика, установленного на выходе лампового усилителя, при большом коэффициенте усиления в отсутствие полезного сигнала на входе напоминал звук падающей на стальной лист дроби. Простой дробовой шум представляет собой флуктуации тока, вызываемые электронами или дырками, которые эмитируются случайно и независимо друг от друга, не взаимодействуя между собой. В твердотельных приборах, например в диодах и транзисторах, носители из эмиттера выходят случайным образом и перемещаются через потенциальный барьер обедненного слоя

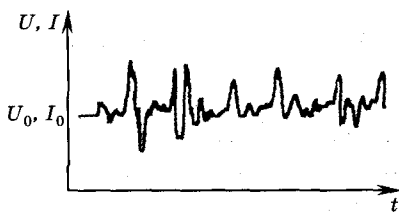


Рис. 22.1

p — n -перехода, вызывая дробовой шум. В вакуумных и газоразрядных приборах электроны эмитируются из катода случайным образом и затем перемещаются к аноду под действием электрического поля. Эти электроны случайным образом преодолевают потенциальный барьер, образующийся на границе катода с вакуумом или рабочей средой (см. п. 11.2). Как в твердотельных, так и в вакуумных приборах ток, создаваемый потоками носителей, флуктуирует около среднего уровня, что связано с дискретной природой эмиссии (см. рис. 22.1). Хотя физическая природа теплового и дробового шумов различна, но вид шумовых сигналов в обоих случаях аналогичен. Как в том, так и в другом случае его можно представить в виде большой последовательности случайных импульсов, похожих по форме и случайно распределенных во времени.

Однако подчеркнем, что дробовой шум в лампах, в отличие от теплового шума в проводниках, возникает при существенно неравновесном процессе, при котором движение носителей осуществляется благодаря внешнему полю и полностью отсутствует взаимодействие носителей как со средой, так и друг с другом.

Диффузионный шум обусловлен вероятностным характером диффузии носителей заряда. Для полупроводниковых материалов этот вид шума является причиной теплового шума, а для приборов с p — n -переходами — основной составляющей дробового шума.

Токовый шум, или $1/f$ -шум, фликкер-шум, шум контактов, избыточный шум — это все различные наименования $1/f$ -шума. Это название связано с тем, что спектральная плотность мощности этого шума изменяется в зависимости от частоты по закону $f^{-\alpha}$, где $\alpha \sim 0,8 \dots 1,2$.

Со стороны сверхнизких частот такая зависимость наблюдается до значений $f \sim 10^{-6}$ Гц. На высоких частотах $1/f$ -шум маскируется тепловым, дробовым или каким-либо другим видом шума. В данный момент нет единой теории этого явления, хотя отдельные частные модели достаточно хорошо могут описывать такие шумы в некоторых типах приборов. Судя по всему, $1/f$ -шум — это совокупность достаточно большого количества различных физических явлений, подчиняющихся одной и той же закономерности. На основе имеющихся данных можно заключить, что причины возникновения $1/f$ -шума в различных случаях разные. $1/f$ -шум можно представить как случайную последовательность импульсов определенной формы $u(t) \sim t^{-1/2}U(t)$, где $U(t)$ — единичная ступенчатая функция: $U(t) = 1$, при $t > 0$, $U(t) = 0$ при $t < 0$, для которой спектральная плотность изменяется как f^{-1} в широ-

ком частотном диапазоне. Такое формальное представление пока не дает представления о физическом механизме, порождающем импульсы указанной формы. Шум, подчиняющийся закону, когда спектральная плотность обратно пропорциональна частоте, проявляется у всех материалов и элементов, используемых в электронике: собственных полупроводников, резисторов, приборов на $p-n$ -переходах, пленок, жидких металлов и растворов электролитов, ламп с термокатодами, сверхпроводников и переходов Джозефсона и т. д. Несмотря на очень большое число опытных фактов, до сих пор нельзя сказать со всей определенностью, принадлежит ли $1/f$ -шум к объемным или поверхностным явлениям. У некоторых типов приборов, например у МДП-транзисторов, это поверхностный эффект. В этом случае в возникновении шума важную роль играет поверхность раздела полупроводник — окисел. У других типов приборов, например у однородных резисторов, — это объемный эффект, обусловленный флуктуациями числа или подвижности носителей заряда. Иногда полученные опытные данные говорят в пользу флуктуации одновременно и числа, и подвижности носителей. Наиболее достоверны данные о спектральной плотности $1/f$ -шума в однородных материалах. Существует эмпирический закон, согласно которому спектральная плотность $1/f$ -шума обратно пропорциональна общему числу носителей заряда в образце. Хотя теоретического обоснования этого закона нет и нет связи с каким-либо физическим механизмом возникновения $1/f$ -шума, он, по-видимому, отражает характерную особенность многих явлений, сопровождающих $1/f$ -шум. При обосновании этого эмпирического закона наиболее широко обсуждался механизм, связанный с наличием ловушек на поверхности и обменом энергией прибора или материала с окружающей средой при тепловом равновесии.

Как уже отмечалось, физический механизм, обуславливающий $1/f$ -шум, в каждом конкретном случае может проявиться в характерных особенностях формы шумового сигнала, иначе, предполагается, что анализ математического описания процесса, обладающего характеристиками $1/f$ -шума, может привести к пониманию физики, лежащей в основе этого явления. Наибольшее распространение в настоящее время получили две модели процессов, имеющих форму спектра $1/f$:

- 1) модель случайного цуга импульсов;
- 2) модель, основанная на суперпозиции большого числа релаксационных процессов с широкой вариацией характерных постоянных времени.

Вторая модель используется более широко, поскольку она имеет непосредственное отношение к $1/f$ -шуму в МДП-транзисторах, связанному с поверхностными явлениями. В соответствии с этим механизмом носители туннелируют между энергетическими состояниями полупроводника и уровнями ловушек, локализованных в слое диэлектрика (окисла), что подтверждается практическим отсутствием $1/f$ -шума в полевых транзисторах (ПТ) с $p-n$ -переходами, где поверхностные явления проявляются слабо. В то же время у МОП ПТ с большой поверхностью раздела полупроводник — окисел составляющая, обусловленная этим видом шума, является доминирующей в общем шумовом спектре. Как показывает эксперимент, и в этом случае интенсивность шума крайне чувствительна к качеству поверхности, а спектральная зависимость, как правило, остается той же, т. е. $\sim 1/f$, независимо от состояния поверхности. Следовательно, каждый новый, последующий слой окисла вызывает появление собственного $1/f$ -спектра, а не сдвигает ВЧ-границу уже существующего спектра в сторону более низких частот за счет больших постоянных времени у более глубоких ловушек. Получившая в настоящее время широкое признание теоретическая модель Мак-Уортера, основанная на захвате носителей поверхностными ловушками, описывает только весьма частный механизм шума в полупроводниках. Наличие $1/f$ -шума в металлах и других материалах, сильная температурная зависимость $1/f$ -шума у пленок металлов и ряд других процессов со спектральной зависимостью $1/f$ пока, по сути дела, полностью не объяснен.

Генерационно-рекомбинационный шум возникает в процессе генерации и рекомбинации как равновесных, так и неравновесных носителей (см. гл. 1). При этих процессах электроны и дырки появляются и исчезают случайным образом. Случайный характер появления и исчезновения свободных носителей приводит к флуктуациям величины сопротивления полупроводника или отдельных областей прибора, где процессы рекомбинации и генерации свободных носителей реализуются. Если через образец такого полупроводника пропустить постоянный ток, то на его концах, помимо постоянной, возникает еще и флуктуационная составляющая ЭДС.

В большинстве полупроводниковых приборов присутствуют обедненные носителями области, например различного рода электрические переходы. При диффузии носителей из той или другой объемной области в обедненный слой носители могут пересечь и покинуть его, или отразиться от него, или же, наконец, часть носителей может быть захвачена рекомбинационными центрами (ло-

вухками) внутри рассматриваемого слоя. В последнем случае во внешней цепи будут возникать импульсы тока, которые образуют рекомбинационный ток (см. п. 1.3, 2.5). Этот ток состоит из стационарной составляющей, на которую накладываются флуктуации, обусловленные случайным характером рекомбинации. Если носители образуются в обедненном слое, то электрическое поле в нем разделяет носители и заставляет их перемещаться в направлении тех объемных областей, где они являются основными.

Процессы генерации сопровождаются появлением во внешней цепи токовых импульсов, имеющих противоположный знак по сравнению с импульсами, возникающими при рекомбинации. Ток генерации также состоит из стационарной составляющей и наложенных на нее распределенных по случайному закону генерационных флуктуаций. Таким образом, процесс генерации и рекомбинации носителей связан с их случайным появлением и исчезновением, и эти явления могут рассматриваться как последовательность независимых случайных событий, подобных дробовому шуму. Поэтому для описания генерационно-рекомбинационных шумов часто можно использовать формулы, полученные для анализа дробовых шумов с введением соответствующих поправок, учитывающих специфику процесса.

Взрывной шум проявляется, как правило, в таких приборах, как диоды, транзисторы на $p-n$ -переходах, туннельные диоды, композиционные резисторы. В простейшем виде взрывной шум проявляется как бистабильный сигнал (с двумя уровнями) ступенчатой формы, с малыми изменениями амплитуды и со случайно распределенными интервалами времени между ступенями (подобно случайному телеграфному сигналу). Однако иногда встречаются сигналы с несколькими уровнями амплитуд (ступеней). Предполагается, что взрывной шум у обратносмещенных $p-n$ -переходов обусловлен нерегулярным включением и выключением поверхностных каналов проводимости. При прямом включении причиной этого шума являются дефекты кристалла в области перехода, причем основную роль играют не металлические примеси, а дефекты, подобные дислокациям. Характерный вид реализации бистабильного взрывного шума приведен на рис. 22.2. Он состоит из случайных ступенчатых выбросов, на которые наложен **белый шум** (обладающий равно-

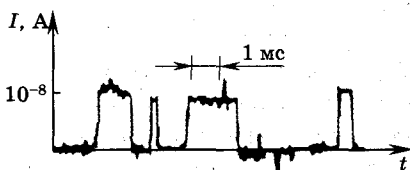


Рис. 22.2

мерным спектром в широком диапазоне частот от 0 до 10^{13} Гц). Считается, что сигнал симметричен, если среднее время длительности каждого из двух уровней ступени одинаково, и асимметричен, если имеет место значительное отклонение от такого условия.

Лавинный и микроплазменный шум связаны с лавинным умножением, имеющим место в областях сильного электрического поля, где кинетическая энергия свободных носителей заряда достаточна для ионизации атомов кристаллической решетки (см. п. 2.5). Процессы ионизации, лежащие в основе умножения носителей, происходят случайно, создавая таким образом шум в токе лавинной ионизации. Лавинное умножение носителей в конце концов приводит к пробоем $p-n$ -перехода (см. п. 2.5). Генерация электронно-дырочных пар во время умножения образует в конечном счете электронно-дырочную плазму. Достаточно часто процессы пробоя реализуются не по всей площади перехода, а в отдельных, порой микроскопически малых областях, где в силу различного рода дефектов существует наибольшая напряженность электрического поля и в результате пробоя образуется микроплазма. Реализация шума микроплазмы часто подобна реализации взрывного шума, но с амплитудой значительно большей ($\sim 10^{-5}$ А), в то время как амплитуда сигнала взрывного шума по порядку величины равна 10^{-8} А.

Сама микроплазма локализуется внутри перехода в областях сильного электрического поля с характерными размерами в несколько сотен ангстрем, в которых трещины и другие дефекты кристаллической решетки содержат ловушки. Эти ловушки, захватывая свободные носители, образуют повышенную плотность заряда, что вызывает возникновение сильного поля. Образование и последующее разрушение микроплазмы является случайным процессом, что и приводит к наблюдаемым ступенчатым изменениям тока перехода.

Шумы токораспределения возникают, когда ток разделяется между электродами прибора, например транзистора или пентода. Характер распределения тока носит случайный характер в основном из-за флуктуации поперечной составляющей скорости движения электрона. На эти шумы оказывают также влияние флуктуации катодного тока или тока эмиттера, флуктуации потенциалов на электродах, различные неоднородности полей, нестабильности электрофизических параметров электродов и т. д.

Все указанные причины приводят к случайному изменению направления движения электронов и, как следствие этого, перехват электронов сеткой в электровакуумной лампе или иным электродом в другом типе прибора является случайным, что и вызывает флуктуации тока в цепях электродов.

Шумы вторичной эмиссии обусловлены тем, что коэффициент вторичной эмиссии с электродов электронных ламп является непостоянным во времени из-за нестабильности первичного тока электронов, неидеальности поверхности, наличия различных примесей, неоднородностей, инородных включений в материале вторичного катода. В обычных усилительных лампах шумы, вызванные вторичной эмиссией, как правило, значительно меньше дробового шума и шума токораспределения.

Радиационный (фотонный) шум определяется флуктуациями параметров светового сигнала (излучения), падающего на приемник, т. е. флуктуациями числа падающих на светочувствительный слой фотоприемника фотонов, которые приходят от излучателя и элементов самого приемника. Кроме рассмотренных, в электронных приборах имеются некоторые другие виды и источники шумов.

22.3. Методы описания шумов

Одним из наиболее эффективных методов теоретического анализа шумов является метод Фурье. В этом методе флуктуирующая величина $x(t)$ может быть описана ее спектральной плотностью $\overline{S_x(f)}$. Например, шумовая ЭДС $U(t)$ в небольшом интервале (полосе) частот представляется через ее спектральную плотность $\overline{S_U(f)}$, а источник флуктуационного тока $i(t)$ в полосе частот Δf замещается генератором шумового тока со средним квадратом $\overline{i_{ш}^2} = \overline{S_i(f)}\Delta f$, где $\overline{S_i(f)}$ — спектральная плотность тока $i(t)$. Достоинство метода Фурье состоит в том, что средние квадраты величин, определяющие интенсивность (мощность) шума, можно рассчитывать при помощи теории цепей переменного тока.

Шумящий электронный прибор обычно представляется в виде шумящего двух- или четырехполюсника (многополюсника) (рис. 22.3). Для упрощения анализа такой системы с шумом многополюсник принимается «идеальным», т. е. свободным от шума (рис. 22.3, а). Он характеризуется обычно матрицей проводимости $\|Y\|$, составленной на основе системы уравнений (4.26). Шум моделируется либо путем включения параллельно входу и выходу шумовых генераторов тока $i_{ш1}(t)$ и $i_{ш2}(t)$ (рис. 22.3, б),

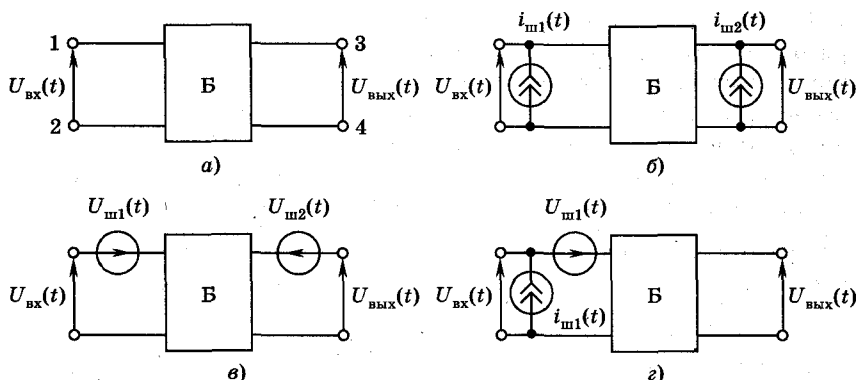


Рис. 22.3

либо последовательным включением шумовых ЭДС $U_{ш1}(t)$ и $U_{ш2}(t)$ (рис. 22.3, в), которые могут коррелировать между собой.

Часто при определении роли шумов в конкретных устройствах и приборах, помимо указанных представлений, используют включение на входе нешумящего четырехполюсника одновременно генераторов входного шумового тока и входного шумового напряжения, причем полагают, что шумы этих генераторов некоррелированы (рис. 22.3, г). Выходные шумы легко вычисляются через известные параметры нешумящего четырехполюсника (см. п. 22.4, шумы биполярных транзисторов).

Параметры шумовых генераторов тока и ЭДС для каждого прибора определяются присущими ему источниками шума. Если анализировать каждый прибор отдельно, то определение роли шумов сильно усложняется из-за большого разнообразия приборов. Ситуация существенно упрощается в случае, если применяются такие шумовые генераторы (источники) тока и напряжения, которые являются эталонными и относительно универсальными, что используется при анализе шумов большинства приборов. Кроме того, в случае большого многообразия источников шума используются более сложные эквивалентные схемы (см. п. 22.4, шумы полевых транзисторов).

В качестве эталонных источников шума обычно берут дробовый шум диода в режиме насыщения и тепловой шум сопротивления R при температуре T .

Если рассматривать любой двухполюсник или прибор, который можно представить как двухполюсник (диоды, сопротивления и т. д.), то источник шумового тока включается параллельно комплексной входной проводимости \dot{Y} , а источник шумовой ЭДС $\sqrt{\overline{U_{ш}^2}} = \sqrt{\overline{i_{ш0}^2} |Z|^2}$ — последовательно с комплексным сопротивлением \dot{Z} . Средний квадрат шумового тока определяется соотношением

$$\overline{i_{ш0}^2} = 2qI_{эKB} \Delta f, \quad (22.1)$$

где $I_{\text{экр}}$ — эквивалентный ток насыщения диода для данной цепи или прибора, q — заряд подвижного носителя (электрона или дырки). Смысл тока $I_{\text{экр}}$ состоит в том, что если измерена величина шума цепи (прибора) и на ее вход подключается параллельно диод в режиме насыщения с током $I_{\text{экр}}$, то мощность шума на выходе удваивается.

В соответствии с теоремой Найквиста тепловой шум сопротивления R при температуре T в частотной полосе Δf может быть представлен либо генератором напряжения с шумовой ЭДС $\sqrt{U_{\text{ш}}^2} = \sqrt{4kTR\Delta f}$, включенной последовательно с сопротивлением R , либо источником шумового тока с $\sqrt{i_{\text{ш}}^2} = \sqrt{4kTR\Delta f/R}$.

Необходимо отметить, что формула Найквиста справедлива, если энергия теплового движения kT больше энергии кванта теплового излучения hf или на частоте шума, или в рассматриваемом спектральном диапазоне, т. е. $hf \ll kT$, где h — постоянная Планка, а k — постоянная Больцмана. Во всем радиодиапазоне, вплоть до оптических частот, это соотношение справедливо. Действительно, лишь при частоте $f = 10^{13}$ Гц ($\lambda = 30$ мкм) энергия кванта hf оказывается по порядку величины равной kT (для $T = 300$ К).

Более общим выражением для мощности шума $P_{\text{ш}}$, справедливым для любого соотношения между hf и kT , является следующее:

$$dP_{\text{ш}} = \frac{hf}{\exp(hf/kT) - 1} df.$$

Учитывая, что мощность P , отдаваемая генератором в согласованную нагрузку R , равна $P = U^2/4R$, получаем значение усредненного квадрата шумовой ЭДС в полосе Δf в виде

$$\overline{U_{\text{ш}}^2} = 4hf \Delta f R / [\exp(hf/kT) - 1]. \quad (22.2)$$

Эта формула позволяет оценивать шумы и в оптическом диапазоне. При $hf \ll kT$ последнее выражение превращается в формулу Найквиста.

Если эти результаты перенести на электронный двухполюсный прибор, то при активной шумовой проводимости прибора $G_{\text{ш}}$ шумовой ток равен

$$\overline{i_{\text{шэ}}^2} = 4kTG_{\text{ш}} \Delta f. \quad (22.3)$$

Шумовая выходная ЭДС вычисляется согласно соотношению

$$\overline{U_{\text{ш}}^2} = \overline{i_{\text{шэ}}^2} |Z|^2 = 4kTR_{\text{ш}} \Delta f, \quad (22.4)$$

где $R_{\text{ш}} = \frac{1}{G_{\text{ш}}}$ — эквивалентное шумовое действительное (активное) сопротивление прибора при температуре T .

Можно преобразовать тепловой шумовой ток в эквивалентный ток диода в режиме насыщения. Для этого необходимо подставить выражение (22.1) для $\overline{i_{ш}^2}$ в формулу (22.3), в результате получим $2qI_{\text{экв}} = 4kTG_{\text{п}}$, откуда следует соотношение

$$I_{\text{экв}} = 2kT/(qR_{\text{п}}). \quad (22.5)$$

Часто для описания шумов, помимо тока диода в режиме насыщения и шумового сопротивления, вводят понятие эквивалентной шумовой температуры $T_{\text{экв}}$. Введение $T_{\text{экв}}$ особенно полезно, если в приборе доминирующим является тепловой шум.

При описании шумов прибора вместо действительной температуры T используют некую «опорную» температуру T_0 . Это связано с тем, что температуру прибора часто бывает трудно определить, так как для различных элементов прибора она может различаться. Так, например, в вакуумных приборах температуры всех электродов, как правило, различны. Поэтому для удобства выбирают фиксированную, стандартную комнатную температуру $T_0 = 290$ К.

При анализе шумов в многополюсных приборах (транзисторах, многоэлектродных лампах и т. д.) для эквивалентных схем применяют включение нескольких источников шума в виде генераторов тока и напряжения или их комбинаций, например генератора шумового тока, включаемого параллельно входной проводимости, и коррелированного с ним генератора шумового напряжения, включаемого последовательно (см. рис. 22.3, з).

Для оценки и сравнения радиоприемных устройств при выделении слабых сигналов на фоне шума широко используется понятие коэффициента шума F .

Коэффициент шума показывает, во сколько раз отношение мощности источника сигнала $P_{\text{ис}}$ к мощности источника шума $P_{\text{иш}}$ на входе больше аналогичного соотношения на выходе, т. е.

$$F = \frac{(P_{\text{ис}}/P_{\text{иш}})_{\text{вх}}}{(P_{\text{ис}}/P_{\text{иш}})_{\text{вых}}}. \quad (22.6)$$

Коэффициент шума часто измеряется в децибеллах (дБ):

$$F \text{ дБ} = 10 \lg F = 10 \lg \frac{(P_{\text{ис}}/P_{\text{иш}})_{\text{вх}}}{(P_{\text{ис}}/P_{\text{иш}})_{\text{вых}}}. \quad (22.7)$$

Эквивалентная шумовая температура системы $T_{\text{экв}}$ связана с коэффициентом шума F соотношением

$$T_{\text{экв}} = T_0(F - 1). \quad (22.8)$$

Из формулы (22.7) видно, что минимальное значение коэффициента шума равно единице ($F_{\text{мин}} = 1$).

22.4. Шумы электронных приборов различного типа

Шумы полупроводниковых диодов. Эти шумы можно оценивать с помощью эквивалентных схем шумящих двухполюсников, у которых интенсивность шумов определяется либо генератором шумового тока, подключенным параллельно проводимости, либо генератором шумового напряжения, включенным последовательно с комплексным сопротивлением диода. Величина и физический механизм шумов у диодов сильно зависят от схемы включения. При прямом включении диода напряженность поля внутри перехода мала, преобладающим оказывается диффузионный ток. В случае малых токов расчет эквивалентного шумового тока можно проводить по формуле Найквиста. При больших токах формула Найквиста дает неверные результаты. Это связано с тем, что при наличии диффузии принципиально важно пользоваться оценками для диффузионных, а не тепловых шумов. Диффузионный и тепловой шум оцениваются одинаково только тогда, когда справедливы соотношения Эйнштейна (1.21) и носители имеют равновесное распределение Максвелла по энергии.

При обратном включении диода напряженность электрического поля внутри перехода велика. Ток через прибор определяется дрейфом неосновных носителей, время перемещения которых меньше их времени жизни, поэтому шумами рекомбинации и флуктуациями скорости перемещения, возникающими из-за столкновений во время пролета перехода, можно пренебречь. В результате доминировать будет дробовой шум, величину тока которого можно оценивать по формуле для $\overline{i_{шз}^2} = 2qI_p \Delta f$, где в качестве $I_{эkv}$ принимается обратный ток $p-n$ -перехода $I_{обp}$. Если обратное напряжение на переходе близко к напряжению пробоя, то необходимо учитывать шумы лавины; эквивалентный ток диода по сравнению с $I_{обp}$ возрастает примерно в M^n раз (M — коэффициент лавинного умножения, $n \sim 3$).

Шумы биполярных транзисторов. Основными источниками шумов в биполярных транзисторах являются: дробовые шумы, обусловленные пролетом носителей через потенциальные барьеры эмиттерного и коллекторного переходов, т. е. дробовые шумы тока базы и тока коллектора; тепловые шумы объемного сопротивления базы и последовательно с ним включенного диффе-

ренциального сопротивления эмиттерного перехода $r_{\text{Э}}$ (см. п. 4.4); $1/f$ -шумы, которые на низких частотах в основном обусловлены поверхностными явлениями (поверхностной рекомбинацией) в области базы.

Шумы в биполярном транзисторе можно проанализировать, используя результаты анализа шумов в $p-n$ -переходах (диодах). Однако следует учитывать при этом рекомбинацию носителей в обедненном слое перехода эмиттер—база. В отсутствие такой рекомбинации транзистор можно считать «идеальным», т. е. его характеристики описываются моделью Эберса—Молла (см. п. 4.3). Составляющая от рекомбинации носителей будет проявляться на контактах прибора как независимая добавка к «идеальным» шумовым токам эмиттера и базы.

Спектральные плотности флуктуации в эмиттерном $\bar{S}_{i_{\text{Э}}}$ и коллекторном $\bar{S}_{i_{\text{К}}}$ токах можно записать в следующем виде:

$$\bar{S}_{i_{\text{Э}}} = 4q(I_{\text{Э}} - I_{\text{Бр}})(G_{\text{Э}}/G_{\text{Э0}} - 1/2) + 2qI_{\text{Бр}}; \quad (22.9)$$

$$\overline{i_{\text{шЭ}}^2} = 4q(I_{\text{Э}} - I_{\text{Бр}})(G_{\text{Э}}/G_{\text{Э0}} - 1/2) \Delta f + 2qI_{\text{Бр}}; \quad (22.10)$$

$$\bar{S}_{i_{\text{К}}} = 2qI_{\text{К}}, \quad \overline{i_{\text{шК}}^2} = 2qI_{\text{К}} \Delta f, \quad (22.11)$$

где $G_{\text{Э}}$ — «идеальная» проводимость перехода эмиттер—база, вычисленная на основе модели Эберса—Молла и формулы Шоттки (2.20); $G_{\text{Э0}}$ — высокочастотная проводимость; $I_{\text{Э}}$, $I_{\text{К}}$ — токи эмиттера и коллектора; $I_{\text{Бр}}$ — компонента базового (и эмиттерного) тока, обусловленная рекомбинацией носителей в обедненном слое. Разность $I_{\text{Э}} - I_{\text{Бр}}$ есть постоянная составляющая тока эмиттера идеального транзистора. В выражении (22.10) первый член определяется тепловыми флуктуациями и процессами рекомбинации—генерации в области базы (ток $I_{\text{Бр}}$), а второй член — это дробовый шум из-за рекомбинационных процессов в обедненном слое.

На низких частотах $G_{\text{Э}} = G_{\text{Э0}}$, и первое слагаемое сводится к выражению $I_{\text{Э}} - I_{\text{Бр}}$, описывающему обычный дробовый шум в токе, т. е. на низких частотах весь шум тока эмиттера определяется полным дробовым шумом. Ток коллектора (выражение (22.11)) для всех частот имеет только дробовую шумовую составляющую. Это выражение получено при условии, что через коллекторный переход проходят независимые (некоррелиро-

ванные) потоки носителей. Если допустить, что вклады от тепловых флуктуаций процессов рекомбинации—генерации в объемной области базы и обедненном слое эмиттерного перехода независимы друг от друга, то может быть получена спектральная плотность шумовых токов эмиттера и коллектора с учетом их взаимной корреляции $\bar{S}_{\text{ЭК}}$. Величина $\bar{S}_{\text{ЭК}}$ может быть вычислена на основе следующей формулы:

$$\bar{S}_{\text{ЭК}}(f) = -2qI_{\text{К}} \frac{\alpha_{\text{К}}}{\alpha_0} \frac{Y_{\text{Э}}}{Y_{\text{Э0}}}, \quad (22.12)$$

где $Y_{\text{Э}}$ — полная проводимость идеализированного перехода эмиттер—база; $\alpha_{\text{К}}$ — отношение переменного тока коллектора к переменному току идеального эмиттера; α_0 — статический коэффициент передачи по току в схеме с ОБ (см. п. 4.4); $Y_{\text{Э0}}$ — статическая проводимость.

На низких частотах $\alpha_{\text{К}} \approx \alpha_0$ и $Y_{\text{Э}} \approx G_{\text{Э0}}$, тогда $\overline{S_{\text{КЭ}}}(f) = -2qI_{\text{К}}$, т. е. флуктуации коллекторного и эмиттерного токов сильно коррелированы. Зная $\bar{S}_{\text{ИК}}$, $\bar{S}_{\text{ИЭ}}$ и $\bar{S}_{\text{КЭ}}$, можно вычислить спектральную плотность флуктуации тока базы $\overline{S_{\text{ИБ}}}(f)$ с учетом взаимной корреляции шумовых токов эмиттера и коллектора, а также спектральную плотность токовых флуктуаций коллектора и базы $\overline{S_{\text{КБ}}}(f)$ с учетом их взаимной корреляции (см. [16]), т. е.

$$\overline{S_{\text{ИБ}}}(f) = \overline{S_{\text{ИЭ}}}(f) + \overline{S_{\text{ИК}}}(f) + \text{Re } \overline{S_{\text{КЭ}}}(f); \quad (22.13)$$

$$\overline{S_{\text{КБ}}}(f) = -\overline{S_{\text{КЭ}}}(f) - \overline{S_{\text{ИК}}}(f). \quad (22.14)$$

Чтобы определить шумовые свойства транзисторного усилителя, оценим коэффициент шума биполярного транзистора, используя эквивалентную схему на рис. 22.3, *г* для включения транзистора с ОЭ. Включение двух генераторов на входе обусловлено тем, что на выходе усилителя шум будет как при коротком замыкании, так и при холостом ходе входной цепи (разомкнутой).

Таким образом, источники шумов в транзисторе представлены в виде внешних генераторов шума, подсоединенных к выводам «бесшумной» эквивалентной схемы. Параметры этих генераторов могут быть вычислены на основе выражений (22.10)—(22.14). Зная параметры генераторов, можно вычислить коэф-

коэффициент шума биполярного транзистора, наименьшее значение которого определяется формулой (см. [16])

$$F \approx 1 + (1/\beta_0 + 2\omega^2 t_{\text{пр}}^2 / 9)^{1/2}, \quad (22.15)$$

где β_0 — статический коэффициент передачи по току в схеме с ОЭ, т. е. $\beta_0 = h_{21Э}$; $t_{\text{пр}} = \frac{W_B^2}{2D}$ — время пролета области базы, ширина которой W_B (см. п. 4.5), D — коэффициент диффузии, $\omega = 2\pi f$.

На низких частотах $F \approx 1 + 1/\sqrt{\beta_0}$, что при $\beta_0 = 100$ дает значение $F \approx 0,4$ дБ. Если частота меньше величины $1/(t_{\text{пр}}\sqrt{\beta_0})$ и шумы генераторов являются некоррелированными, то

$$\overline{U_{\text{ш. вх}}^2} \approx (2qI_{\text{КЭ}}/G_{Э0}^2) \Delta f \text{ и } \overline{I_{\text{ш. вх}}^2} = 2qI_B \Delta f. \quad (22.16)$$

При учете сопротивления базы для $\beta_0 \gg 1$ и $r_{Э} \gg r_B$ коэффициент шума на низких частотах может быть вычислен по формуле (см. [16]):

$$F \approx [1 + (r_B + r_{Э}/2)/R_c + R_c/(2r_{Э}\beta_0)],$$

где $r_{Э} = 1/G_{Э0}$ (см. п. 4.5); r_B — объемное сопротивление базы (см. п. 4.4); R_c — активное сопротивление источника сигналов.

Отметим, что влияние сопротивления базы r_B на коэффициент шума учитывается добавлением r_B к шумовому сопротивлению транзистора. Изменяя R_c , можно добиться минимального коэффициента шума при $R_c = [2r_{Э}\beta_0(r_B + r_{Э}/2)]^{1/2}$:

$$F_{\text{мин}} \approx 1 + 2[(r_B + r_{Э}/2)/(2r_{Э}\beta_0)]^{1/2}. \quad (22.17)$$

Больше всего $1/f$ -шум сказывается на низких частотах, и его влияние можно учесть, если вместо $\beta_0 = h_{21Э}$ подставить

$$\beta'_0 = \beta_0/(1 + f_F/f), \quad (22.18)$$

где f_F — характерная частота $1/f$ -шума, которая для малошумящего кремниевого планарного транзистора не превышает 10^3 Гц, т. е. при частотах f сигнала в несколько кГц $1/f$ -шум можно не учитывать.

Шумы полевых транзисторов. Источники шума в полевых транзисторах (ПТ) с управляющим p — n -переходом и МДП(МОП)-транзисторах во многом схожи. Наиболее сильные различия наблю-

даются только для $1/f$ -шума, интенсивность которого мала для ПТ с $p-n$ -переходом и значительна на низких частотах в МОП-транзисторах.

Наиболее важным источником шума в ПТ являются тепловые флуктуации проводимости в канале транзистора. Такие флуктуации обуславливают тепловой шум тока стока и, за счет емкости затвор—канал $C_{зк}$, тока затвора. В кремниевых ПТ с $p-n$ -переходом генерация носителей в области пространственного заряда переходов канал—затвор формирует основную часть тока утечки затвора, а так как $p-n$ -переходы находятся при обратном смещении, то генерируемые носители движутся в области сильного электрического поля. Поэтому составляющая, обусловленная дробовым шумом в токе затвора, доминирует на низких частотах по сравнению с тепловой составляющей шума. Эти же процессы генерации носителей в обедненном слое вызывают появление шума в стоковом токе из-за того, что примесные генерационные центры все время меняют свое зарядовое состояние, вызывая местное изменение ширины обедненного слоя, а следовательно, ширины канала и его проводимости, что и обуславливает шумовой ток в стоковой цепи.

При комнатных температурах флуктуации концентрации носителей в канале транзистора из-за рекомбинационно-генерационных процессов являются еще одним источником шума. Однако наибольший уровень все же имеют генерационные шумы, вызванные примесными центрами (ловушками) в областях пространственного заряда ($p-n$ -переходах).

Таким образом, в ПТ областями, где сосредоточены источники шума, являются: канал, $p-n$ -переходы, подзатворные поверхностные области.

В зависимости от длины канала в нем могут проявляться либо тепловые шумы (длинный канал), либо дробовые шумы (короткий канал). В длинных каналах напряженность электрического поля существенно меньше, чем в коротких каналах или в коллекторном переходе биполярных транзисторов, поэтому в первом приближении можно полагать, что роль дрейфового движения в формировании интенсивности флуктуации при перемещении носителей от истока к стоку мала. Носители в длинном канале находятся почти в тепловом равновесии с кристаллической решеткой, тогда средний квадрат шумового тока стока $\overline{i_{шс}^2}$ на выходе для схемы с общим истоком и на линейном

участке стоковой характеристики можно представить в следующем виде:

$$\overline{i_{шC}^2} = 4kTG_{\text{вых}} \Delta f, \quad (22.19)$$

где T — температура канала; $G_{\text{вых}}$ — проводимость канала ПТ на линейном участке стоковой ВАХ.

При работе полевого транзистора в области насыщения выражение для среднего квадрата шумового тока приближенно описывается формулой

$$\overline{i_{шC}^2} = 8kTS \Delta f/3,$$

где S — крутизна на участке насыщения стоковой ВАХ.

Пересчитанная во входную цепь величина $\overline{i_{шC}^2}$ при работе транзистора на участке насыщения стоковой ВАХ определит параметры входного шумового генератора напряжения, который для схемы с общим истоком МДП-транзистора будет включен в цепь затвора и параметры которого можно вычислить по формуле

$$\overline{U_{ш1}^2} = \overline{i_{шC}^2}/S^2 = 8kT\Delta f/(3S^2), \quad (22.20)$$

где S — крутизна стоковой характеристики ПТ.

В случае преобладания дробовых шумов в канале (короткий канал) шум на выходе может быть оценен согласно выражению

$$\overline{i_{шC}^2} = 2qI_c \Delta f.$$

Шумовые токи во входной цепи (в цепи затвор—исток), кроме пересчитанных тепловых шумов канала, обусловлены еще и дробовыми флуктуациями тока утечки затвора в МДП-транзисторах и обратного тока p — n -перехода в полевых транзисторах с управляющим переходом, т. е.

$$\overline{i_{ш3}^2} = 2qI_3 \Delta f. \quad (22.21, а)$$

Помимо этого, флуктуации потенциала вдоль канала в МДП-транзисторах через емкость затвор—канал $C_{3К}$ вызывают (наводят) шумовую составляющую в токе затвора, которая на участке насыщения стоковой ВАХ оценивается по формуле

$$\overline{i_{ш3 \text{ нас}}^2} = kT(\omega^2 C_{3К}^2 / G_{\text{нас}}) \Delta f. \quad (22.21, б)$$

Из выражения (22.21, б) видно, что наведенная шумовая составляющая в цепи затвора носит тепловой характер.

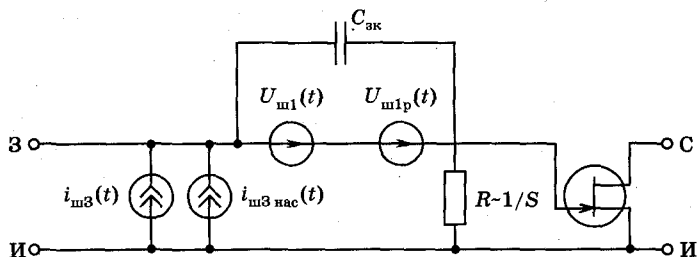


Рис. 22.4

Как показывают исследования, несмотря на то что тепловые шумовые токи затвора и канала обусловлены одинаковыми физическими процессами, взаимокорреляционная связь между ними весьма незначительна. Таким образом, полный шумовой ток во входной цепи образуется суммой наведенного (22.21, б) и дробового (22.21, а) шумов (рис. 22.4). Принято считать, что эти шумы некоррелированы. Во входной цепи полевых транзисторов существует также составляющая $1/f$ -шума из-за случайных рекомбинационных процессов на поверхности в МДП-транзисторах и в обедненном слое в транзисторах с управляющим $p-n$ -переходом. Этот шум учитывается введением в эквивалентную шумовую схему входной цепи ПТ (см. рис. 22.4) последовательно с шумовым генератором $\overline{U_{ш1}^2}$ (22.20) еще одного генератора шумового напряжения с ЭДС, определяемой формулой

$$\overline{U_{ш1p}^2} = (8kTS_{нас}/3)(f_0/f) \Delta f,$$

где f_0 — частота, на которой интенсивность $1/f$ -шума становится соизмеримой с интенсивностью белого шума.

Итак, в полевых транзисторах основными источниками шумов являются тепловые или дробовые шумы в канале, наведенные, дробовые и $1/f$ -шумы во входной цепи (цепи затвора).

Сравнение спектральных плотностей $S_U(f)$ шумовых ЭДС различного типа транзисторов представлено на рис. 22.5, из которого видно, что в области высоких частот МОП-транзисторы являются наиболее шумящими.

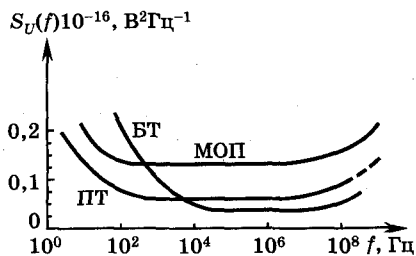


Рис. 22.5

Шумы фотоприемников. Фото-резисторы. Основные источники шума — тепловой, генерационно-рекомбинационный и дробовой. Эквивалентная схема фоторезистора с учетом генераторов шума представлена на рис. 22.6. На этом рисунке:

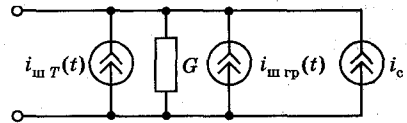


Рис. 22.6

- 1) генератор тока $\sqrt{i_{ш T}^2}$ моделирует тепловые шумы $\overline{i_{ш T}^2} = 4kTG \Delta f$, где G — проводимость, обусловленная тепловым током, фототоком и фоновым током;
- 2) шумовой генератор $\sqrt{i_{ш гр}^2}$ моделирует генерационно-рекомбинационный шум

$$\sqrt{i_{ш гр}^2} = (\tau/t_{пр})[4qI_0 \Delta f / (1 + \omega^2 \tau^2)], \quad (22.22)$$

где τ — время жизни фотоносителей; I_0 — стационарный выходной фототок; $t_{пр}$ — время пролета носителей между контактами; $\omega = 2\pi f$ — частота модуляции оптического сигнала;

3) генератор i_c определяет среднеквадратичный токовый полезный сигнал.

Качество фотодетекторов часто оценивают эквивалентной мощностью шума (*noise equivalent power*), которая определяется как среднеквадратичная мощность падающего излучения, необходимая для получения отношения сигнал/шум, равного единице в полосе 1 Гц.

Фотодиоды. Шумы фотодиодов всех типов (*p-i-n*-фотодиоды, с барьером Шоттки, с гетеропереходами) в основном имеют одинаковые источники, к которым относятся дробовые, генерационно-рекомбинационные и тепловые флуктуации. Эти источники, как и при рассмотрении шумовых свойств других приборов, моделируются генераторами шумовых токов.

При поглощении оптического сигнала и фонового излучения в фотодиодах генерируются электронно-дырочные пары, которые разделяются электрическим полем и дрейфуют в противоположные от *p-n*-перехода стороны. При этом фототок, обусловленный оптическим сигналом, равен $I_{ф}$, фоновым излучением — $I_{фон}$, а ток термической генерации электронно-дырочных пар в обедненной области — I_T . Вследствие случайного характера процессов генерации носителей (определяющих эти токи) и прохожде-

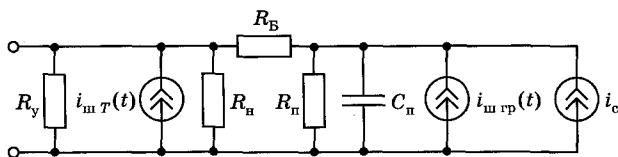


Рис. 22.7

нии их через сильное поле обедненной области возникают дробовые шумы с интенсивностью

$$\overline{i_{шд}^2} = 2q(I_{\phi} + I_{\text{фон}} + I_T) \Delta f. \quad (22.23)$$

Помимо этих шумов, наличие объемных сопротивлений областей диода вносит дополнительные тепловые шумы. На эквивалентной схеме фотодиода, приведенной на рис. 22.7, изображены шумовые генераторы токов и другие элементы схемы, определяемые параметрами фотодиодов. На схеме C_n и R_n — емкость и сопротивление перехода; R_B — последовательное объемное сопротивление базы диода; R_n — сопротивление нагрузки; R_y — входное сопротивление последующего усилителя полезного сигнала. Тепловые шумы определяются суммарной проводимостью фотодиода, так как

$$\overline{i_{шT}^2} = 4kTG_{\Sigma} \Delta f, \quad (22.24)$$

где

$$G_{\Sigma} = 1/R_{\Sigma} = (1/R_y) + (1/R_n) + (1/R_n).$$

Лавинные фотодиоды. В лавинном фотодиоде (ЛФД) происходит умножение сигнального I_{ϕ} , фонового $I_{\text{фон}}$ и темнового I_T токов. В результате для ЛФД выражение (22.23) принимает следующий вид:

$$\overline{i_{шд}^2} = 2q(I_{\phi} + I_{\text{фон}} + I_T)M^2F(M) \Delta f, \quad (22.25)$$

где M — коэффициент умножения (см. п. 16.4); $F(M) = \overline{M^2}/M^2$ — шум-фактор; $\overline{M^2}$ — средний квадрат внутреннего коэффициента усиления (умножения). Шум-фактор $F(M)$ определяет меру увеличения дробового шума по отношению к идеальному нешумящему усилителю. Эквивалентная схема ЛФД подобна обобщенной схеме фотодиодов (см. рис. 22.7), поэтому тепловой шум

в ЛФД вычисляется по формуле (22.24). Существенной особенностью ЛФД является то, что лавинное умножение может привести к увеличению отношения сигнал/шум, причем существует оптимальное значение коэффициента умножения M , которое обеспечивает максимальное отношение сигнал/шум при заданной оптической мощности. Это связано с тем, что лавинный процесс является статистическим, вероятностным процессом из-за неодинакового умножения, производимого отдельно электронами и дырками, поскольку коэффициенты ионизации электронов и дырок различны. В результате при приеме оптического сигнала может сложиться ситуация, при которой первичная флуктуация числа носителей в области умножения усиливается относительно меньше, чем принимаемый сигнал, т. е. отношение сигнал/шум улучшится. В формуле (22.25) шум из-за лавины характеризуется шум-фактором $F(M)$, который зависит от коэффициентов ионизации для электронов α_n и дырок α_p . Поэтому флуктуация числа первичных носителей вызывает после умножения существенное относительное изменение общего числа флуктуирующих носителей и шум-фактор возрастает. Если же один из коэффициентов ударной ионизации мал, например $\alpha_p = 0$, т. е. рожденные дырки вообще не производят ионизацию, то в этом случае после каждого последующего акта ионизации в области умножения будут появляться носители, образованные только электронами. В первом же случае рождались носители, образованные как электронами, так и дырками. Поэтому во втором случае флуктуация, соответствующая изменению числа носителей после умножения, будет заметно меньше. Таким образом, при существенном различии α_n и α_p значение шум-фактора может быть мало и, следовательно, шумы ЛФД будут меньше.

Таким образом, за счет особенностей образования лавины шумы в ЛФД могут усиливаться меньше, чем полезный сигнал.

Шумы электровакуумных приборов. Как ясно из изложенного, диод, работающий в режиме насыщения, является эталонным источником шума и средний квадрат эквивалентного шумового тока этого прибора вычисляется по формуле $\overline{i_{шз}^2} = 2qI_{э\text{кв}} \Delta f$. Однако если в диоде реализуется режим объемного заряда (см. п. 11.3), то пространственный заряд, формируя своего рода потенциальный барьер у катода, уменьшает интенсивность шумов на величину, которая характеризуется коэффициентом депрес-

сии γ и тогда в обобщенном виде шумы электровакуумного диода можно оценивать по формуле

$$\overline{i_{шд}^2} = 2qI_a \Delta f \gamma^2, \quad (22.26)$$

где I_a — ток анода; γ — коэффициент депрессии, который для режима насыщения равен единице. В общем случае $\gamma^2 = C_1 k T_k / (q U_a)$, где $C_1 \approx 7$ — константа; U_a — анодное напряжение; T_k — температура катода.

Шумы вакуумного триода вычисляются по формуле (22.26), только с иным значением γ . В выражении для коэффициента депрессии вместо U_a необходимо поставить действующее напряжение U_d (см. п. 11.3). Если величину U_d выразить через крутизну анодно-сеточной характеристики S , выражение для среднего квадрата шумового тока для триода примет вид

$$\overline{i_{штп}^2} = 4kT_k(0,644S) \Delta f. \quad (22.27)$$

Шумовое напряжение между катодом и управляющей сеткой лампы можно вычислить, если учесть, что для малых сигналов $S = i_a/u_c$, где i_a и u_c — соответственно мгновенные значения переменных составляющих анодного тока и сеточного напряжения (см. п. 11.5). Тогда, принимая, что $u_c = u_{вх}$, $i_{вх} = i_a$, можно получить

$$\overline{U_{штвх}^2} = \overline{i_{штп}^2} / S^2 = 4kT_k(0,644/S) \Delta f. \quad (22.28)$$

Выражение (22.28), по сути дела, описывает тепловые шумы, возникающие в активном сопротивлении $R_{шт}$ при температуре T_0 . Значение $R_{шт}$ вычисляется по формуле

$$R_{шт} = 0,644T_k / (ST_0), \quad (22.29)$$

где T_0 — стандартная температура окружающей среды, значение которой принимается равным 300 К.

Дополнительно в триоде существуют шумы, обусловленные токами утечки в цепи сетки и наведенными токами. Интенсивность последнего источника шума растет с увеличением частоты. Эти источники создают дополнительный дробовый шум. Кроме этого, на низких частотах может сильно сказаться взрывной $1/f$ -шум, особенно в приборах с оксидным катодом. $1/f$ -шум обусловлен флуктуациями работы выхода электронов с различ-

ных участков катода. В режиме работы триода с токами сетки необходимо учитывать еще и шумы токораспределения. Помимо этого, при наличии вторичной эмиссии с электродов появляются дополнительные шумы.



Контрольные вопросы

1. Каковы физический смысл понятия «шум» и причины появления шумов в электронных приборах?
2. Каково происхождение основных источников шумов?
3. Физические явления, определяющие $1/f$ -шум.
4. Объяснить природу генерационно-рекомбинационного шума, лавинных и микроплазменных шумов.
5. Методы описания шумов.
6. Объяснить природу шумов полупроводниковых приборов и методы расчета их параметров.
7. Физические механизмы возникновения шумов фотоприемников.
8. Объяснить происхождение и способы оценки шумов электровакуумных приборов.

Глава 23

ЭКСПЛУАТАЦИОННЫЕ УСЛОВИЯ РАБОТЫ, РЕЖИМЫ И НАДЕЖНОСТЬ ЭЛЕКТРОННЫХ ПРИБОРОВ

23.1. Эксплуатационные условия, параметры и режимы работы электронных приборов

Номинальные значения параметров электронных приборов могут быть реализованы только при определенных условиях их эксплуатации. Эксплуатация электронных приборов должна осуществляться в соответствии с требованиями ТУ и стандартами-руководствами по применению электронных приборов (общие положения) и руководством для конкретного класса приборов.

В процессе эксплуатации окружающая среда может оказывать сильные воздействия на электронные приборы, изменяя их параметры и характеристики. Влияние окружающей среды определяется механическими (вибрации, удары, ускорения и т. п.) и климатическими (температура, давление, влага, атмосферный воздух, радиация, пары химически активных веществ и т. п.) воздействиями. Свойства приборов при наличии указанных воздействий характеризуются следующими основными параметрами: механической устойчивостью, прочностью, климатической устойчивостью, радиационной стойкостью и рядом других.

Механическая устойчивость определяется виброустойчивостью, ударной устойчивостью и устойчивостью к воздействиям постоянных ускорений. **Виброустойчивостью** называется способность электронного прибора сохранять электрические параметры под воздействием вибраций в пределах норм, установленных техническими условиями. Вибрации характеризуются частотой и ускорением. Наибольшим **вибрационным воздействиям** подвергается радиоэлектронная аппаратура, устанавливаемая на борту летательных аппаратов.

Частотный диапазон наиболее интенсивных вибраций на современном самолете простирается от единиц Гц до 5 кГц, а полный диапазон вибраций превышает 10 кГц. При наибольших амплитудах вибраций ускорения достигают 15...20g. Вибрации могут носить резонансный характер.

Ударная устойчивость (ударостойкость) — это способность приборов противостоять механическим ударным воздействиям, при которых отклонения параметров прибора не превышают установленной нормы. Этот параметр характеризуется значением ударного ускорения и числом ударов, выдерживаемых прибором.

Устойчивость к воздействию постоянных ускорений оценивается величиной постоянных ускорений, реализуемых обычно посредством центрифуги. Ускорения при оценке механической устойчивости выражаются в единицах *g*.

Прочность — это способность прибора (изделия) противостоять разрушающему воздействию вибрации и ударов. Соответственно различают **вибропрочность** и **ударную прочность**. Эти параметры определяются предельными ускорениями (в единицах *g*) и длительностью воздействий, после которых прибор выполняет свои функции, регламентируемые ТУ. Значения ускорений, устанавливаемые ТУ для аппаратуры, работающей в наиболее тяжелых условиях, достигают 40 и более *g*.

Климатическая устойчивость — это способность прибора выполнять свои функции под воздействием таких факторов, как температура, влага, пыль, газ, давление окружающей среды и т. д. Для характеристики климатической устойчивости используют такие параметры, как термостабильность, влагостойкость, водостойкость и др.

Термостабильность приборов характеризуется температурным коэффициентом (ТК) того или иного параметра. ТК определяет обратимое изменение параметров прибора. Необратимые изменения параметров прибора определяются температурным коэффициентом нестабильности (ТКН).

Влагостойкость и **водостойкость** определяют способность приборов выдерживать соответственно воздействие влажности воздуха (окружающей среды) и воды. Для защиты полупроводниковых приборов от влияния окружающей среды используют герметизацию корпусов приборов. При использовании приборов в аппаратуре, работающей в условиях повышенной влажности, платы с полупроводниковыми диодами и транзисторами подвергаются многослойному покрытию специальными лаками или заливаются компаундами. Микросхемы и микросборки с бескорпусными приборами подвергаются общей герметизации. При применении заливки плат компаундами и покрытии их лаками необходимо учитывать ухудшение теплоотвода.

Радиационная стойкость характеризует способность электронных приборов работать под воздействием проникающего ионизирующего излучения с сохранением изменений основных параметров в пределах норм, регламентированных ТУ. Наибольшее влияние оказывает γ -излучение, при этом различают воздействия, носящие как обратимый, так и необратимый характер (см. п. 23.2).

В справочниках обычно приводятся значения параметров приборов, которые гарантируются техническими условиями для оптимальных (номинальных) или предельных режимов эксплуатации.

Электронные приборы являются, как правило, устройствами универсального применения. Они могут быть использованы не только по прямому назначению, но и в других случаях. Однако набор параметров и характеристик, приводимых в справочнике, соответствует в первую очередь прямому назначению прибора. Значения большинства параметров зависят от рабочего режима и температуры, поэтому обычно в литературе даются усреднен-

ные величины параметров и, кроме того, устанавливается интервал их изменения. Этот интервал соответствует или минимальным, или максимальным значениям разброса параметров.

На практике изменения токов и напряжений могут происходить в большом диапазоне значений. Ограничением возможности применения приборов служат значения параметров предельно допустимых режимов. Даже кратковременные превышения параметров предельно допустимых режимов могут приводить к выходу из строя приборов, особенно полупроводниковых, поэтому при работе необходимо принимать во внимание нестабильности источников питания, значение и характер нагрузки, амплитуды и длительности сигналов на выходе и т. д. Превышение предельной температуры, особенно в мощных приборах, может вызвать тепловой пробой $p-n$ -перехода, поэтому необходимо стремиться улучшать теплоотвод от прибора. Хороший и правильный тепловой режим работы приборов снижает интенсивность отказов транзисторов, улучшает стабильность параметров аппаратуры. В аппаратуре теплоотвод от приборов обеспечивается радиаторами или соответствующей конструкцией элементов и узлов.

23.2. Радиационная стойкость электронных приборов

Основные сведения об ионизирующих излучениях. Под *радиационной стойкостью* понимается возможность электронных приборов работать под воздействием проникающего ионизирующего излучения (ИИ) при изменении основных параметров в пределах норм, регламентированных техническими условиями.

По физической природе ИИ подразделяются на следующие основные виды: α - и γ -излучения, нейтронное (n), электронное (e), протонное (p) излучения.

Тяжелые частицы (α -частицы и осколки деления ядер) из-за их малой проникающей способности оказывают незначительное влияние на изделия электронной техники. Источником α -частиц может быть естественная радиация элементов конструкции ИС. α -частицы испускаются незначительным количеством радиоактивных примесей, присутствующих в материалах конструкции ИС.

Источниками ИИ являются ядерные энергетические установки, ядерные взрывы, естественные и искусственные радиационные пояса Земли, космические лучи и др. *Космические лучи* образуются за счет галактического излучения и излучения Солнца. *Галактические лучи* представляют собой ядра различных химических элементов (~ 85% протоны, ~ 15% ядра гелия и очень небольшая доля тяжелых элементов).

Искусственные радиационные пояса возникают в результате ядерных взрывов в верхних слоях атмосферы и имеют определенный период существования. **Естественные радиационные пояса Земли** образуются в результате захвата магнитным полем Земли потоков заряженных частиц, входящих в состав космических лучей. Эти пояса делятся на внутренний, расположенный на высоте 600...1500 км, и внешний радиационный пояс, расположенный на расстоянии $\sim 1,5$ радиуса Земли.

Основными характеристиками ИИ являются: **энергия частиц**, выражаемая в электронвольтах (эВ); **плотность потока частиц** N , определяемая числом частиц, проходящих через единичную площадку в 1 см^2 за 1 с ; **интегральный поток** Φ , т. е. усредненный по всему энергетическому распределению частиц со сложным энергетическим спектром (Φ определяется плотностью потока частиц, умноженной на время облучения, $\text{част}/\text{см}^2$); **доза (поглощенная доза)** D и **мощность дозы** (мощность поглощенной дозы) P — для излучения со сложным энергетическим спектром (D — это энергия, переданная веществу ионизирующим излучением в расчете на единицу массы; мощность дозы P — значение дозы, отнесенное к единице времени). Доза излучения D измеряется в Дж/кг и рад ($\text{Дж}/\text{кг} = 1 \text{ Гр} = 10^2 \text{ рад}$), а мощность дозы P — в Вт/кг.

Для количественного описания рентгеновского и γ -излучений с энергией квантов не выше 2 МэВ, ионизирующих вещество, используют такую характеристику, как **экспозиционная доза** D_γ . Единица измерения D_γ — Кл/кг — это экспозиционная доза, при которой в результате ионизации в единице массы (1 кг) возникает такое количество заряженных частиц, которое переносит заряд каждого знака в один кулон. Широко применяется внесистемная единица, определяющая экспозиционную дозу, — **рентген**, равный $2,58 \cdot 10^{-4}$ Кл/кг.

В зависимости от продолжительности воздействия на материалы различают **импульсное и непрерывное ИИ**. Импульсное ИИ реализуется в условиях ядерного взрыва, а непрерывное — при наличии ядерных энергетических установок, космических лучей и т. п. Для оценки импульсных ИИ вводят такой параметр, как экспозиционная доза мгновенного излучения, которая определяется соотношением

$$D_{\gamma \text{ мгн}} = P_\gamma t, \text{ Кл/кг},$$

где t — длительность импульса излучения; P_γ — мощность дозы в импульсе мгновенного γ -излучения, которая определяется в А/кг. Типичные значения t колеблются в пределах $10^{-7} \dots 10^{-3}$ с.

Как правило, эффекты, порождаемые импульсной радиацией, вызывают обратимые изменения электрических параметров изделий. Отметим некоторые характеристики наиболее мощных источников ионизирующих излучений. На расстоянии 1 км от центра ядерного взрыва интегральный поток нейтронов $\Phi \sim (3 \dots 5) \cdot 10^{16}$ нейтр/см². Излучение в активной зоне мощного реактора (~ 1000 МВт) характеризуется следующими величинами: плотность потока нейтронов $N \sim 10^{14} \dots 10^{15}$ нейтр/(см² · с);

мощность дозы γ -излучения $P_\gamma \sim 10^3$ А/кг. При мощности реактора 100 кВт $P_\gamma \sim 10^{-1}$ А/кг, $N \sim 10^{10}$ нейтр/(см²·с). Частицы галактических лучей обладают очень высокой энергией ($10^8 \dots 10^{20}$ эВ), но плотность потока их очень мала. Доза облучения на летательном объекте за год составляет в этом случае $D \sim 10^{-2} \dots 10^{-1}$ Дж/кг ($1 \dots 10$ рад).

Максимально возможные годовые дозы для протонов солнечных космических лучей достигают величин порядка $10 \dots 10^2$ Дж/кг ($10^3 \dots 10^4$ рад) на поверхности аппарата и $1 \dots 10$ Дж/кг ($10^2 \dots 10^3$ рад) за слоем защиты¹ 1 г/см². Годовые экспозиционные дозы излучения на поверхности космического летательного аппарата от внутреннего и внешнего радиационных поясов Земли при облучении электронами с энергией 0,02...5 МэВ соответственно составляют 10^{12} и $10^{11} \dots 10^{13}$ рад в год, а за слоем защиты 10^5 рад в год. Протоны внутреннего пояса дают экспозиционную дозу на поверхности аппарата 10^{10} рад в год, а за слоем защиты 10^5 рад в год.

Излучение искусственных радиационных поясов характеризуется следующими величинами: $D \sim 10^5$ Гр (10^7 рад), $P_\gamma \sim 10^{-1} \dots 10^{-4}$ А/кг. Максимально возможная годовая доза облучения радиоэлектронной аппаратуры, обусловленная потоками электронов из искусственных радиационных поясов, может составлять на поверхности аппарата $10^6 \dots 10^7$ Дж/кг ($10^8 \dots 10^9$ рад), а с учетом защиты аппаратуры корпусом летательного аппарата — $10^3 \dots 10^4$ Дж/кг ($10^5 \dots 10^8$ рад). Таким образом, радиационная обстановка на борту околоземных космических объектов определяется в основном протонами внутреннего радиационного пояса и электронами искусственных радиационных поясов Земли.

Проблема обеспечения радиоэлектронной аппаратуры материалами и приборами, устойчивыми к воздействию проникающей радиации, является весьма сложной из-за большого числа используемых в них материалов. Уровень воздействия ИИ на радиоэлектронные изделия, в том числе на электронные приборы, зависит от комплекса физических, химических, технологических и конструктивных факторов.

Основные типы радиационных дефектов в твердых телах. В результате взаимодействия ядерных излучений с атомами твердого тела происходят следующие процессы: смещение атомов из устойчивых положений в решетке, ионизация, иногда появление в решетке примесей за счет ядерного деления, а также задержки в решетке (внедрения в нее) тех атомов, которыми производилось облучение.

Характер радиационных повреждений зависит от типа связей в облучаемом твердом теле, от вида и условий облучения.

¹ Слой защиты δ в 1 г/см² определяется как произведение плотности вещества d (г/см³) на толщину слоя l (см), например для железа $l = \delta/d = 1/7,8 \approx 0,125$ см.

Все дефекты, возникающие в кристаллической решетке твердого тела под действием ионизирующего излучения, разделяются на простые (точечные) и сложные (комплексы), представляющие совокупности простых.

Электро- и теплопроводность, механические, оптические и магнитные свойства, коэффициенты диффузии свободных носителей и другие характеристики реальных кристаллов связаны с точечными дефектами.

Дефекты чаще всего проявляются в виде смещения атомов с образованием вакантных узлов и внедренных атомов. Существенное изменение электрофизических параметров твердого тела наблюдается при количествах радиационных дефектов и внедренных атомов, сравнимых с исходным числом дефектов и примесей, присущих исходному кристаллу.

Существует несколько теоретических моделей описания радиационных нарушений в твердом теле. Наиболее распространенные из них используют предположение об образовании первичных точечных дефектов в кристаллической решетке в результате упругого столкновения движущейся частицы с ядром или атомом вещества. При этом предполагается, что атом смещается из своего нормального положения в решетке, если приобретенная им энергия E_a больше некоторой пороговой энергии E_d , которая зависит от энергии связи атомов в кристаллической решетке. В узле, где атом находился до столкновения, образуется вакансия. В твердых телах с сильными связями пороговая энергия имеет порядок величины $E_d \approx 25$ эВ.

Первичный смещенный атом, если он обладает достаточной энергией, перемещаясь по решетке, вызывает возбуждение, ионизацию и образование вторичных смещенных атомов. Потеряв запас энергии, он остановится в междуузлии. Таким образом, при $E_a > E_d$ полное число смещений атомов превышает число актов их взаимодействия с ионизирующими частицами (ИИ) из-за развития каскадных процессов. Теория каскадного процесса образования сложных дефектов наряду со смещениями атомов в междуузлии учитывает их переход в другие узлы решетки за счет смещения атомов решетки движущимися атомами (частицами).

Процессы упругого смещения атомов определяются структурными особенностями твердого тела. В кристаллической решетке можно выделить направления с более плотной упаковкой атомов, где имеются условия для облегченной передачи энергии ИИ атомам. В этих направлениях при облучении кристалла образуются области уплотнения, а вдоль трека (направления движения) взаимодействующей частицы формируется область разрыхления (сложный дефект). В реальных кристаллах всегда существуют различного типа дислокации с поверхностной плотностью до 10^8 см⁻². При облучении вещества наиболее вероятны размножения дислокаций, которые также относятся к сложным дефектам. Наличие химических примесей в кристаллах может привести к возможности взаимодействия с ними простых радиационных дефектов и образования стабильных комплексных дефектов. Большинство устойчивых радиаци-

онных дефектов в германии, кремнии, арсениде галлия и других материалах появляются в результате взаимодействия точечных дефектов между собой и с атомами примесей, имеющих в веществе до облучения. Различные виды ИИ образуют дефекты, отличающиеся своими свойствами. Коротко рассмотрим воздействие на материалы каждого вида ИИ.

Быстрые нейтроны (кинетическая энергия > 100 кэВ), проникая в твердое тело, в основном образуют структурные радиационные дефекты в результате упругого взаимодействия с ядрами. При этом быстрый нейтрон передает ядру часть своей кинетической энергии, что может вызвать ионизацию атома. Энергетический спектр быстрых нейтронов ядерного реактора занимает диапазон $0,1 \dots 14$ МэВ при средней энергии $1,5 \dots 2$ МэВ.

При расчетах полного числа смещений обычно полагают, что все нейтроны обладают одной энергией, равной средней. Вклад нейтронов в диапазоне более 2 МэВ незначителен, и им пренебрегают. Из-за высокой проникающей способности быстрых нейтронов в веществах со средним атомным весом распределение дефектов по объему кристалла является равномерным.

Тепловые нейтроны (энергия 0,5 эВ и менее), имеющиеся в спектре реакторного излучения, хотя и приводят к ядерным превращениям, которые по своему действию подобны своеобразным легирующим примесям, вызывают меньшее число структурных дефектов по сравнению с быстрыми.

Протоны при взаимодействии с веществом теряют кинетическую энергию за счет упругих и неупругих столкновений с атомами. Упругое рассеяние преобладает при энергии протонов менее 50 МэВ. Если энергия протонов более 50 МэВ, то необходимо учитывать как упругое, так и неупругое рассеяние.

Расчеты показывают, что протоны с энергией $E_{пр} = 2,5$ МэВ передают при однократном столкновении с атомами кремния и германия относительно небольшую энергию ($E_{а.ср} \approx 130$ эВ) по сравнению с нейтронным облучением ($E_{а.ср} = 100$ кэВ для Si и $E_{а.ср} = 40$ кэВ для Ge при энергии нейтронов $E_n = 1,5$ МэВ), что свидетельствует о возможности создания большого числа первично смещенных атомов при передаче быстрыми протонами небольших порций энергии атомам решетки. По сравнению с нейтронами, которые рассеиваются на ядрах, протоны рассеиваются в результате кулоновских взаимодействий как с электронными оболочками атомов, так и с ядрами, что приводит к формированию большого числа дефектов.

Электроны при взаимодействии с веществом чаще всего испытывают неупругое рассеяние, вызывающее ионизацию атомов. Однако наряду с этим определяющим процессом некоторую роль может играть упругое рассеяние, приводящее к смещению атомов в междуузлия. Сказанное о характере взаимодействия нейтронов и заряженных частиц с атомами и ядрами хорошо иллюстрируется данными, приведенными в табл. 23.1.

Таблица 23.1

Зависимость числа первично смещенных атомов в 1 см^3 вещества от вида и энергии излучения (в расчете на одну быструю частицу)

Материал	Нейтроны $E_n = 2 \text{ МэВ}$	Протоны		Электроны $E_e = 2 \text{ МэВ}$
		$E_{пр} = 2 \text{ МэВ}$	$E_{пр} = 5 \text{ МэВ}$	
Кремний $E_d = 12,9 \text{ эВ}$	42	341	148	2,8
Германий $E_d = 14,5 \text{ эВ}$	62	482	213	1,2

Гамма-кванты (γ -кванты) при прохождении через вещество вызывают в основном ионизацию атомов и смещение атомов при упругих столкновениях комптоновских электронов (см. ниже) с атомами материала.

ИИ при ядерных и термоядерных взрывах, от импульсных ускорителей и различных изотопных источников содержит γ -кванты с кинетической энергией до 10 МэВ, для которых характерны три основных процесса взаимодействия с веществом; это фотоэлектрическое поглощение (фотоэффект), комптоновское рассеяние (Комптон-эффект) и эффект образования электронно-дырочных пар (соответственно области А, С и В на рис. 23.1). Каждый из упомянутых процессов проявляется при определенных значениях энергии γ -квантов и атомного номера элемента Z (см.

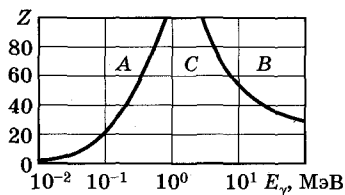


Рис. 23.1

рис. 23.1). В каждом из процессов возникают электроны с энергиями, сравнимыми с энергией падающих квантов. Следовательно, γ -облучение сопровождается внутренней бомбардировкой вещества быстрыми электронами, что может также вызвать смещение атомов. Энергия комптоновских электронов зависит от энергии падающего γ -кванта и атомного веса мишени.

При больших значениях мощности дозы γ -излучения происходит сильная ионизация вещества и окружающей его газовой среды, что может вызывать значительное изменение электрофизических свойств облучаемого образца.

Влияние ионизирующих излучений на параметры полупроводников. Характер дефектов кристаллической решетки полупроводниковых материалов, обусловленных воздействием радиации, не отличается от рассмотренных. Возникающие дефекты в объеме полупроводников приводят к изменению концентрации свободных носителей заряда n , их подвижности μ , времени жизни τ и диффузионной длины L .

Время жизни неосновных носителей заряда в объеме полупроводников чаще всего определяется не прямой рекомбинацией зона—зона,

а через энергетические уровни ловушек, расположенные в запрещенной зоне. Возникновение радиационных дефектов приводит к формированию глубоких энергетических уровней в запрещенной зоне полупроводника, что увеличивает скорость объемной рекомбинации и, соответственно, уменьшает время жизни пар неравновесных носителей заряда. Несмотря на появление в запрещенной зоне полупроводника большого числа уровней, многие из которых являются рекомбинационными, время жизни в облученном материале, как показывают эксперименты, определяется лишь одним-двумя доминирующими рекомбинационными центрами. Если концентрация рекомбинационных центров достаточно мала по сравнению с концентрацией равновесных носителей заряда, то, как показывают теоретические расчеты, в случае одного доминирующего уровня время жизни носителей τ обратно пропорционально концентрации рекомбинационных центров N_r , которая, в свою очередь, прямо пропорциональна интегральному потоку облучения. Следовательно, изменение обратного времени жизни носителей $\Delta\left(\frac{1}{\tau}\right) = \frac{1}{\tau} - \frac{1}{\tau_0}$ (τ_0 , τ — соответственно времена жизни неосновных носителей заряда до и после облучения) при облучении пропорционально концентрации рекомбинационных центров N_r , или интегральной плотности потока ионизирующих частиц Φ :

$$\Delta\left(\frac{1}{\tau}\right) = K_r \Phi, \quad (23.1)$$

где K_r — коэффициент радиационного изменения времени жизни неосновных носителей заряда, зависящий от концентрации основных носителей заряда, скорости образования радиационных центров и от степени заполнения их электронами.

Таким образом, при увеличении Φ время жизни уменьшается, а изменение обратного времени жизни $\Delta\left(\frac{1}{\tau}\right)$ увеличивается, что проиллюстрировано в качестве примера для n -Si с удельным сопротивлением $\rho \sim 100 \text{ Ом} \cdot \text{см}$ на рис. 23.2, где кривая 1 определяет влияние облучения на время жизни γ -квантами, 2 — электронами с $E_e = 2,5 \text{ МэВ}$, 3 — электронами с $E_e = 30 \text{ МэВ}$, 4 — быстрыми нейтронами.

Анализ экспериментальных результатов исследования влияния радиации на подвижность носителей показывает, что наблюдаемые изменения подвижности обусловлены не только рассеянием на ионизованных атомах, возникших в результате облучения (падающие правые вет-

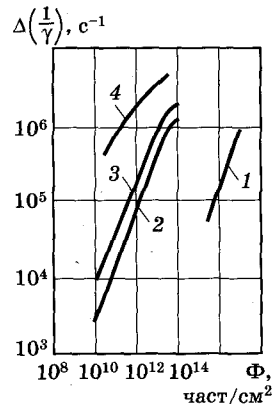


Рис. 23.2

ви кривых 4, 5, 6 на рис. 23.3), но и на фонах, появляющихся из-за эффектов смещения атомов (левые нарастающие ветви кривых 4, 5, 6 и кривые 2, 3). Кривые 2 и 4 получены при облучении потоками нейтронов соответственно с интенсивностью $\Phi_n = 1,6 \cdot 10^{14} \text{ см}^{-2}$ и $\Phi_n = 3,3 \cdot 10^{15} \text{ см}^{-2}$, кривая 3 — электронами с $\Phi_e = 10^{15} \text{ см}^{-2}$; 5, 6 — протонами с $\Phi_{пр} = 10^{14} \text{ см}^{-2}$ и $\Phi_{пр} = 6 \cdot 10^{14} \text{ см}^{-2}$ и энергией $E_{пр} = 660 \text{ МэВ}$. Значительное уменьшение подвижности при низких температурах (большие значения $\frac{1}{T}$) во время облучения по сравнению с исходным состоянием обус-

ловлено появлением ионизованных атомов и комплексов атомов. Радиация увеличивает число заряженных неподвижных частиц, что приводит к возрастанию вероятности кулоновских столкновений свободных носителей с ионами и, соответственно, к уменьшению подвижности. Поскольку физические механизмы взаимодействия с веществом полупроводников потоков облучения различной природы количественно и качественно могут существенно различаться (см. табл. 23.1), изменение подвижности в Si для каждого типа потоков будет разное, что хорошо видно на примере рис. 23.4 (кривая 1 — γ -кванты, 2 — электроны с энергией 10 МэВ, 3 — быстрые нейтроны).

Изменение концентрации свободных носителей в Si наиболее сильно сказывается при небольших уровнях легирования и облучении потоком нейтронов с $\Phi_n > 10^{14} \text{ нейтр/см}^2$.

Радиационная стойкость полупроводниковых диодов. К настоящему времени наиболее изучены радиационные эффекты в выпрямительных диодах. В кремниевых диодах при воздействии ионизирующего излучения любого вида с достаточно высокой энергией происходит деградация его вентильных свойств, связанная с увеличением обратной и уменьшением прямой проводимости. Скорость деградации в основном определяется шириной базы W , ее удельным сопротивлением ρ_B и значением прямого тока $I_{пр}$. Помимо этого, заметное влияние оказывают энергия излучения, концентрация примесей в эмиттерной области, градиенты концентрации и время жизни дырок в базе τ_p (для базы с электронной проводимостью). В большинстве случаев в кремниевых диодах наибольшие изменения при облучении претерпевает прямая ветвь ВАХ, поэтому радиационную стойкость диодов часто оценивают по двукратному увеличению прямого напряжения $U_{пр}$ при заданном прямом токе $I_{пр}$.

Основными электрофизическими параметрами полупроводников, определяющими изменение ВАХ при облучении, являются τ_p , N_d , ρ_B и W . Изменение времени жизни дырок τ_p в базе диода, которое является функцией упомянутых параметров N_d , ρ_B и W , в широком диапазоне значений $I_{пр}$ при различных видах облучения можно оценить по формуле (23.1).

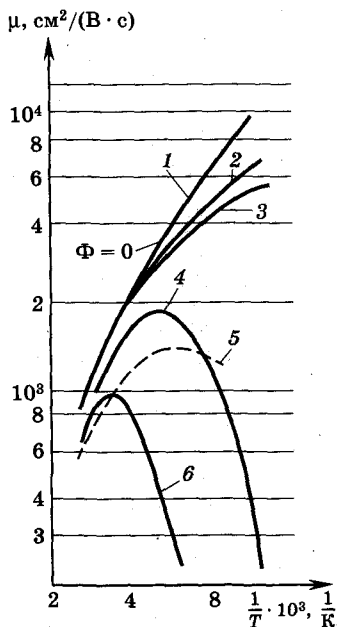


Рис. 23.3

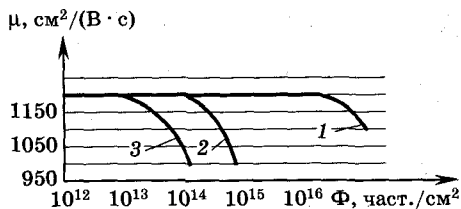


Рис. 23.4

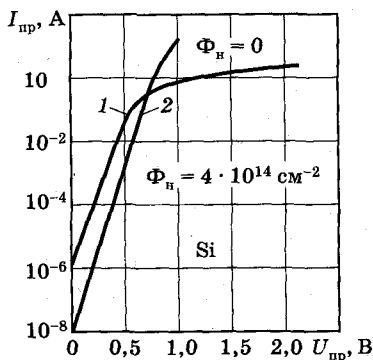


Рис. 23.5

ВАХ в широком диапазоне значений прямых токов и изменение прямого напряжения $U_{пр}$ кремниевых диодов при облучении нейтронами представлены на рис. 23.5 и 23.6 соответственно. На рис. 23.5 кривая 1 получена при $\Phi_n = 0$, т.е. в отсутствие облучения, а кривая 2 — при $\Phi_n = 4 \cdot 10^{14}$ нейтр/см², на рис. 23.6 $U_{пр0}$ — прямое напряжение при $\Phi_n = 0$, и заштрихованные области — разброс значений $U_{пр}/U_{пр0}$ для различных примесей (например, В и Al). Диоды, где в качестве примеси использован бор (В), имеют большую радиационную стойкость, чем при использовании алюминия (Al).

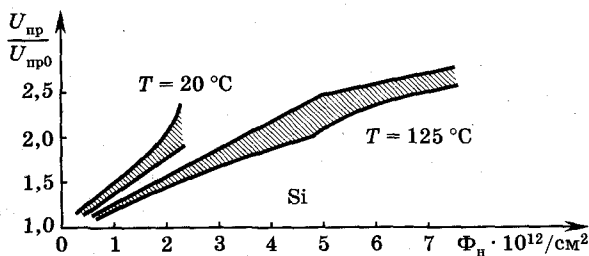


Рис. 23.6

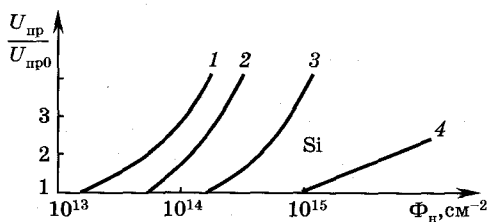


Рис. 23.7

Облучение диодов при повышенных температурах вызывает существенно меньшие изменения их свойств за счет частичного отжига радиационных дефектов. Отжиг — это ускоренное восстановление заряда внутри материала (см. далее *воздействие ИИ на МДП-транзисторы*).

Влияние облучения на различные типы диодов иллюстрирует рис. 23.7, где кривые 1, 2 получены при облучении различных марок выпрямительных диодов, а 3, 4 — импульсных; $U_{пр0}$ — прямое напряжение на диоде при $\Phi_n = 0$.

Влияние радиации на обратные ветви ВАХ кремниевых диодов в основном определяется процессами, происходящими в $p-n$ -переходе. Облучение изменяет эффективное распределение нескомпенсированной примеси вблизи $p-n$ -перехода, т. е. влияет на толщину перехода. Это связано с тем, что при облучении происходит изменение эффективной концентрации носителей в базовой области. Изменением же концентрации основных носителей в эмиттере можно пренебречь, так как для $p-n$ -перехода $N_a \gg N_d$ (см. п. 2.1), поэтому увеличение толщины перехода при облучении происходит в основном в сторону базы.

Кроме того, наличие в переходе значительных градиентов концентрации легирующей примеси приводит к тому, что при облучении может наблюдаться существенное изменение градиентов концентрации свободных носителей и, следовательно, токов, текущих через переход. Зависимость плотности обратного тока $j_{обр}$ от потока нейтронов для кремниевых диодов с $\rho_B = 2,5 \text{ Ом} \cdot \text{см}$ при $U_{обр} = 300 \text{ В}$ представлена на рис. 23.8 (кривая 1 получена при $T = 18^\circ \text{C}$, 2 — при $T = 100^\circ \text{C}$). Изменение ширины перехода при облучении естественно будет вызывать и изменение барьерной емкости.

Германиевые диоды имеют очень низкую радиационную стойкость, т. е. при облучении в них резко возрастает обратный ток $I_{обр}$, поскольку в отличие от диодов на основе кремния основной вклад в $I_{обр}$ вносит диффузионная, а не генерационная составляющая тока (см. табл. 23.1).

$p-i-n$ -диоды. В отличие от диодов с низкоомной базой при описании радиационных эффектов в этих диодах вместо традиционных параметров используют только одну радиационную константу K_r (23.1). Расчеты и эксперименты показывают, что радиационная чувствительность

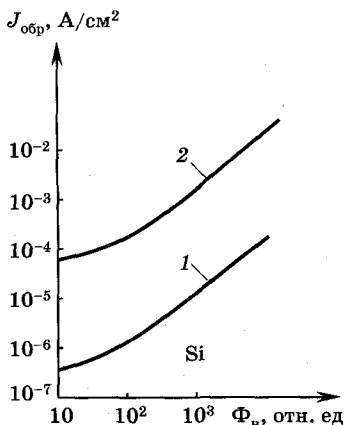


Рис. 23.8

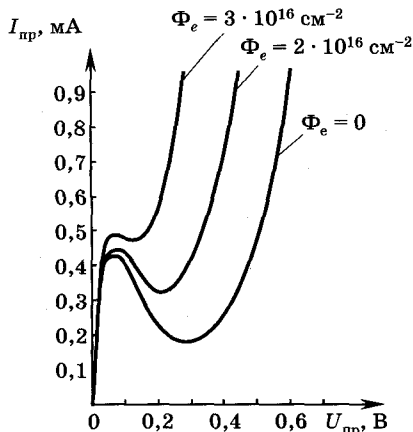


Рис. 23.9

$p-i-n$ -диодов, определяемая коэффициентом K_r , прямо пропорциональна квадрату толщины базовой области.

Туннельные диоды. Радиационные эффекты в туннельных структурах сложны и разнообразны из-за того, что процессы переноса носителей могут происходить одновременно как непосредственно между зонами, так и с участием уровней в запрещенной зоне, образуемых за счет ИИ. Особенно сильно изменяется ток впадины I_v (избыточный ток), который как раз и определяется туннелированием носителей на уровни в запрещенной зоне. Пиковый ток $I_{п}$ также заметно растет при облучении (рис. 23.9). Среди диодов различных типов туннельный является наиболее радиационно стойким.

В момент воздействия импульсного ИИ в $p-n$ -переходе диода образуются избыточные неосновные носители, которые под действием поля перехода разделяются, как при фотовольтаическом эффекте (см. п. 16.4). В результате во внешней цепи диода возникает импульс тока с амплитудой, пропорциональной площади перехода и мощности дозы излучения P , т. е.

$$I_{\Phi} = qG_nPS(L_{об} + L_n + L_p),$$

где L_n, L_p — диффузионные длины носителей; $L_{об}$ — ширина обедненной области перехода; G_n — фактор ионизации ($4,3 \cdot 10^{13} \text{ см}^{-3} \cdot \text{Гр}^{-1}$ для кремния).

Наиболее критичными параметрами к импульсному ИИ у диодов являются обратный ток и прямое падение напряжения. Мощность дозы γ -излучения для диодов $P_{\gamma} = 10^5 \dots 10^6 \text{ Гр/с}$. При этом изменение $U_{пр}$ обусловлено явлениями, подобными вентильному фотоэффекту, как в фотодиодах.

Радиационная стойкость транзисторов. При облучении большинство параметров биполярных транзисторов (БТ) изменяется. Однако, как правило, из всех параметров рассматривают один основной — статический коэффициент передачи тока. Изменение этого коэффициента, связанное с объемными и поверхностными явлениями, определяет радиационную стойкость большинства самых разнообразных схем, где используются БТ. Объемные изменения под действием ионизирующего излучения обусловлены в основном упругим рассеянием ИИ с образованием радиационных дефектов, приводящих к изменению времени жизни, концентрации и подвижности носителей заряда.

Неупругое рассеяние излучения ионизирует газ в корпусе прибора и генерирует свободные носители в веществе, которые захватываются поверхностными уровнями. В результате ионизации газа может происходить осаждение ионов на поверхность полупроводника. Как захват свободных носителей поверхностными ловушками, так и осаждение ионов изменяют поверхностные свойства полупроводника, в частности скорость поверхностной рекомбинации. При наличии объемных и поверхностных явлений в полупроводнике изменение статического коэффициента передачи тока в схеме с общим эмиттером ($h_{21Э}$) может быть оценено по формуле

$$\Delta(1/h_{21Э}) = t_{Эк} K_{эфф} \Phi, \quad (23.2)$$

где $t_{Эк}$ — время пролета носителей между эмиттером и коллектором; $K_{эфф}$ — эффективный коэффициент радиационных повреждений, связанный прежде всего с изменением времени жизни неосновных носителей заряда и скорости поверхностной рекомбинации, т. е. $K_{эфф}$ по сравнению с K_t учитывает большее число факторов, изменяющих время жизни в конкретных приборах. Например, $K_{эфф}$ в германиевых транзисторах определяется: радиационным изменением времени жизни неосновных носителей заряда в активной и пассивной областях базы; конфигурацией и геометрическими размерами как этих областей, так и эмиттера; величиной внутреннего поля в пассивной базе. Для уменьшения влияния пассивной базы на величину $h_{21Э}$ при облучении необходимо обрабатывать приэмиттерную поверхность базы транзисторов ускоренными ионами, а германиевые планарные приборы необходимо изготавливать с тонкими эмиттером и пассивной областью базы.

Величина $K_{эфф}$ для кремниевых транзисторов определяется радиационными процессами в активной базе и в области пространственного заряда эмиттерного перехода.

На рис. 23.10 представлены экспериментальные зависимости относительных значений коэффициента передачи тока кремниевых транзисторов малой и средней мощности в зависимости от величины интегрального потока нейтронного Φ_n , электронного Φ_e , протонного $\Phi_{пр}$ и

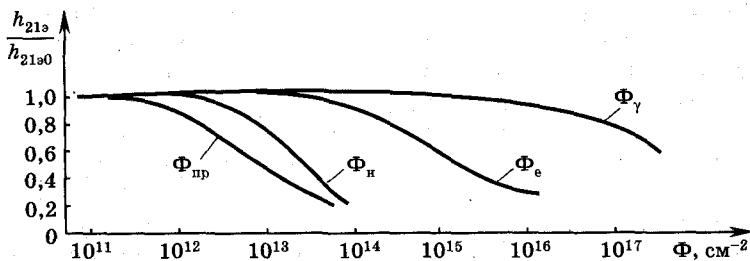


Рис. 23.10

γ -излучений Φ_{γ} . Приведенные зависимости позволяют проводить ориентировочную оценку стойкости транзисторов к любому виду излучений.

В ряде приборов, таких, как германиевые непассивированные (с не протравленной поверхностью кристалла) транзисторы, кремниевые транзисторы, работающие в микрорежимах, МДП-транзисторы, в основном именно поверхностные изменения при облучении определяют радиационную стойкость приборов и стабильность их параметров после облучения. Основные процессы на поверхности полупроводников при облучении можно классифицировать следующим образом:

- образование новых быстрых поверхностных состояний, что увеличивает скорость поверхностной рекомбинации;
- возбуждение поверхностных состояний или захват генерируемых облучением носителей ловушками, что изменяет заряд поверхностных состояний и величину поверхностного потенциала;
- ионизация газа в корпусе прибора при облучении и осаждение заряженных ионов на поверхности кристалла, что изменяет также поверхностный потенциал.

Для современных полупроводниковых приборов с пассивированной поверхностью, например для планарных транзисторов с пленкой SiO_2 , ионизация газа в корпусе при облучении практически не сказывается на изменении их параметров.

Под действием ионизирующего излучения в пленке SiO_2 формируется дополнительный положительный заряд, что в совокупности с ростом плотности быстрых поверхностных состояний приводит к изменению параметров приборов. Так, в МОП-транзисторах с p -каналом образование дополнительного положительного заряда смещает пороговое напряжение затвор—исток $U_{зи\text{ пор}}$ в сторону больших отрицательных значений, а увеличение плотности быстрых поверхностных состояний изменяет крутизну вольт-амперных и вольт-фарадных характеристик. Оба эти эффекта не исчезают после снятия ИИ.

Возникновение положительного заряда ионов можно объяснить следующим образом. Ионизирующее излучение вызывает образование электронно-дырочных пар в слое диэлектрика. Если к затвору приложено по-

ложительное относительно подложки напряжение, то, в первую очередь, из-за большой подвижности, электроны будут вытягиваться полем затвора. Дырки захватываются дырочными ловушками или рекомбинируют с электронами до выхода из окисла. В результате в диэлектрике SiO_2 формируется избыточный положительный заряд. Образующийся заряд Q при фиксированном потенциале затвора уменьшает напряженность поля в окисле, что приводит к насыщению Q при росте поглощенной дозы D . Объемный заряд расположен внутри окисла на расстоянии нескольких десятков нанометров от границы раздела кремний—диэлектрик. Величина объемного заряда определяется только поглощенной дозой облучения. Величина и локализация пространственного заряда определяется величиной и полярностью напряжения смещения на затворе в процессе облучения. При этом наблюдается линейная зависимость наведенного заряда от напряжения затвор—исток $U_{\text{зи пор}}$. Это связано с тем, что все приложенное к затвору напряжение падает на слой объемного заряда, создающегося в окисле во время облучения, а не на весь слой окисла.

Положительный пространственный заряд в слое диэлектрика модулирует проводимость канала полевого транзистора и изменяет все его характеристики.

Наиболее чувствительны к облучению ИИ n -канальные МОП-транзисторы, у которых наименьшие значения $U_{\text{зи пор}}$. Зависимость $U_{\text{зи пор}}$ для p -канального транзистора от D дана на рис. 23.11. При увеличении $t_{\text{зд. вкл}}$ соответственно возрастает и время задержки $t_{\text{зд. вкл}}$ сигнала при работе транзистора в импульсном режиме (см. рис. 23.11).

Облученные электронами образцы были стабильны при комнатной температуре в различных средах и при разных электрических режимах. Восстановление заряда до величины $Q \sim 10^{-11} \text{ см}^{-2}$ протекает в течение нескольких месяцев. Процесс ускоренного восстановления пространственного заряда в окисле, который называется отжигом, происходит за счет нейтрализации его электронами, введенными в окисел. Обычно нейтрализующие электроны получают за счет или повышения температуры прибора (материала) — термического отжига, или облучения светом — УФ отжига. Восстановление заряда при термическом отжиге

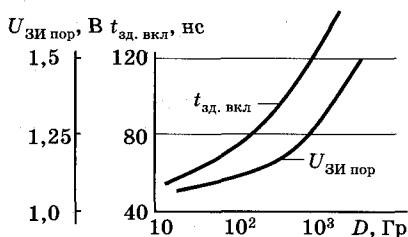


Рис. 23.11

обусловлено инжекцией электронов из кремния в окисел. Одновременно при температурах отжига (150...300 °С) происходит восстановление скорости поверхностной рекомбинации. УФ освещение вызывает заброс электронов в окисел из кремния. Этот процесс происходит с наибольшей интенсивностью при энергии кванта света $h\nu > 5 \text{ эВ}$, а начинается он при $h\nu \sim 3...5 \text{ эВ}$.

В настоящее время ведутся поиски новых диэлектриков для затвора, в которых скорости накопления заряда при облучении были бы минимальны. К таким диэлектрикам относятся структуры $\text{SiO}_2\text{—Si}_3\text{N}_4$ (нитрид кремния), а также Al_2O_3 . Одной из причин уменьшения скорости накопления заряда в этих диэлектриках при облучении ИИ является более высокая, чем у SiO_2 , проводимость, что приводит к частичной компенсации положительного объемного заряда дырок электронами проводимости.

Радиационная стойкость транзисторов к непрерывному ИИ в существенной мере зависит от их частотных свойств, технологии изготовления, конструкции. Для биполярных мощных НЧ-транзисторов предельные плотности потоков частиц (нейтронов) составляют $\Phi_n = 3 \cdot 10^{11} \dots 3 \cdot 10^{13} \text{ см}^{-2}$, для ВЧ-транзисторов — $\Phi_n = 5 \cdot 10^{15} \dots 5 \cdot 10^{17} \text{ см}^{-2}$, для МОП-транзисторов — $\Phi_n = 3 \cdot 10^{12} \dots 2 \cdot 10^{14} \text{ см}^{-2}$.

Наиболее критичными параметрами к импульсному ИИ являются токи через обратносмещенные переходы. Радиационная стойкость к импульсному ИИ составляет $10^3 \dots 10^6$ Гр/с для кремниевых биполярных транзисторов. В схеме с общим эмиттером (ОЭ) из-за ИИ первичный ток коллекторного $p\text{—}n$ -перехода может вызвать вторичный ток эмиттерного перехода. Первичный ток, вызванный ИИ, является результатом диффузии неосновных носителей из коллектора в базу транзистора. В базе эти носители становятся основными и при малых токах утечки начинают снижать потенциальный барьер эмиттерного перехода. Для схемы транзистора с ОЭ этот эффект накопления носителей в базе приводит к открыванию транзистора и усилению коллекторного тока.

Для тириستоров импульсные токи, возникающие под действием импульсного ИИ, могут вызвать его переключение из закрытого в открытое состояние.

Следует также обратить внимание на возможность ионизации газа в герметичном корпусе транзистора, что может приводить к изменениям параметров транзисторов.

Радиационные эффекты в арсенид-галлиевых полупроводниковых приборах и интегральных схемах. Высокая подвижность свободных носителей ($\mu = 5000 \text{ см}^2/(\text{В} \cdot \text{с})$), большая ширина запрещенной зоны ($\Delta E_g = 1,43 \text{ эВ}$), малые времена жизни носителей ($\tau \sim 10^{-8} \text{ с}$) арсенида галлия делают этот материал одним из наиболее перспективных полупроводников для изготовления приборов и интегральных схем, работающих при повышенных температурах, в широком диапазоне частот, включая миллиметровый диапазон.

Основными причинами изменения характеристик GaAs при воздействии ИИ, так же как у Ge и Si, являются процессы возникновения радиационных дефектов в кристаллической структуре и генерации неравновесных носителей заряда.

Сравнительный анализ радиационного изменения электрофизических параметров арсенида галлия и кремния показывает, что с точки зрения влияния объемных эффектов смещения GaAs не имеет преимуществ перед Si в широком диапазоне концентрации легирующей примеси. Арсенид галлия отличает от кремния существенно меньшее время жизни носителей заряда, которое в значительной мере определяется процессами излучательной и Оже-рекомбинации, особенно при повышенной концентрации неравновесных носителей. В арсениде галлия указанные процессы по сравнению с кремнием более вероятны. Наряду с объемными эффектами смещения, на характеристики GaAs-приборов могут существенно влиять изменение плотности поверхностных состояний и накопление заряда на границе раздела между областями прибора.

Рассмотрим остаточные эффекты в GaAs-приборах. В настоящее время наибольшее применение в микросхемах на арсениде галлия находят диоды с барьером Шоттки (выпрямляющий контакт металл—полупроводник) и полевые транзисторы с длиной затвора от единиц до десятых долей мкм. Эти приборы используются в СВЧ-устройствах и сверхбыстродействующих логических элементах и ячейках памяти, являющихся основой цифровых интегральных микросхем (ИМС).

Диоды с барьером Шоттки имеют высокую граничную частоту, малое время переключения $t_{пер}$ и ВАХ, близкую к идеальной. Коэффициент неидеальности (см. гл. 2) для GaAs-диода равен $m = 1,01...1,1$, а для кремниевого $p-n$ -перехода $m = 1,5$. Предельно достигнутые значения времени переключения $t_{пер}$ составляют единицы пс.

Детальное исследование радиационного изменения характеристик диодов Шоттки (ДШ) на основе контактов Au- n -GaAs и Al- n -GaAs при облучении быстрыми нейтронами (рис. 23.12) позволило выявить доминирующие механизмы деградации этих приборов. ДШ были выполнены на основе эпитаксиальной пленки n -GaAs с концентрацией легирующих

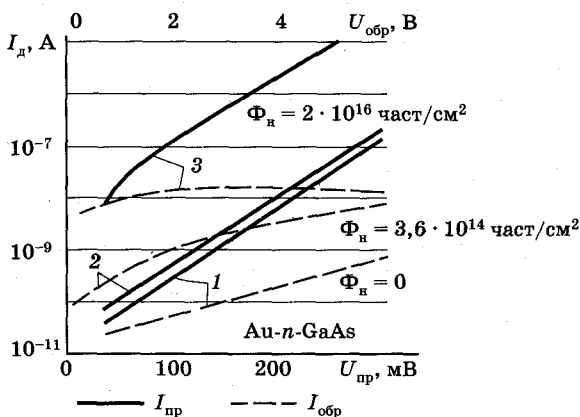


Рис. 23.12

примесей $10^{15} \dots 10^{17} \text{ см}^{-3}$. Значения потока нейтронов выбирались из условий малого изменения концентрации основных носителей. Причина увеличения обратного тока, а соответственно и прямого (рис. 23.13), в ДШ при облучении потоками нейтронов связана с увеличением туннельного тока, обусловленного переходами носителей из металла на радиационные уровни в запрещенной зоне обедненного слоя ДШ.

Ионное легирование базы (ионная имплантация) ДШ очень сильно изменяет радиационные характеристики приборов. Так, например, при облучении потоком нейтронов $\Phi_n = 3 \cdot 10^{15} \text{ нейтр/см}^2$ базы диода (контакт никель— n —GaAs), легированной ионами бора (концентрация 10^{13} см^{-3}), увеличивается коэффициент неидеальности ВАХ для прямого напряжения $U_{пр} = 0,2 \dots 0,3 \text{ В}$ с $m = 1,01$ ($\Phi_n = 0$) до $m \approx 4,25$, а при $U_{пр} > 0,6 \text{ В}$ — до $m \approx 2$. Столь сильное изменение m может быть обусловлено следующими физическими механизмами: туннелированием носителей через уровни ловушек в запрещенной зоне полупроводника, изменением высоты потенциального барьера, уменьшением концентрации свободных носителей, процессами генерации—рекомбинации в обедненном слое перехода.

Перейдем к рассмотрению остаточных эффектов при ИИ полевых транзисторов с коротким каналом, длина которого обычно $0,5 \dots 2 \text{ мкм}$. В этих транзисторах реализуются высокие значения электрического поля в активной области канала и наблюдается насыщение скорости дрейфа электронов под затвором транзистора. В результате начало и наклон пологого участка выходных ВАХ при росте напряжения сток—исток определяют не только перекрытием канала, но и насыщением дрейфовой скорости. На рис. 23.14 (взят из [27]) показано влияние облучения полевых тран-

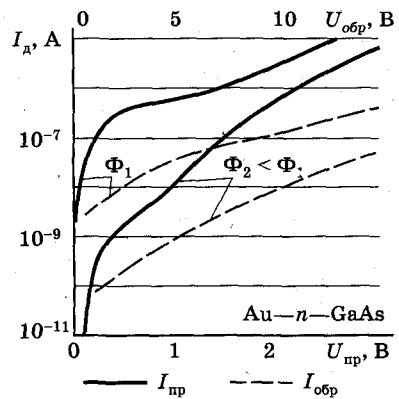


Рис. 23.13

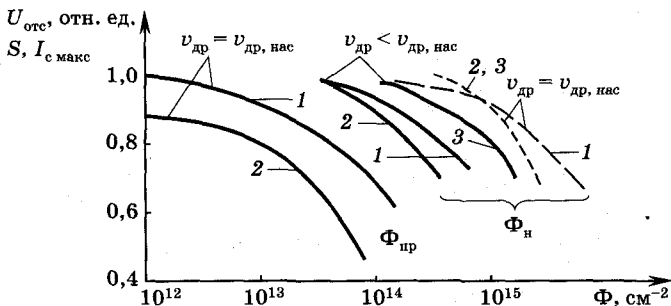


Рис. 23.14

зисторов нейтронами и протонами на изменение крутизны S (кривые 1), максимального тока стока $I_{C_{\max}}$ (кривые 2) и напряжения отсечки $U_{\text{отс}}$ (кривые 3). Как видно из рисунка, наиболее существенная деградация характеристик полевых транзисторов наблюдается при протонном облучении. Если в канале транзистора происходит насыщение скорости дрейфа носителей $v_{\text{др}} = v_{\text{др. нас}}$, то влияние радиации на указанные параметры заметно меньше, чем при $v_{\text{др}} < v_{\text{др. нас}}$. Это объясняется тем, что при $v_{\text{др}} = v_{\text{др. нас}}$ скорость дрейфа определяется в первую очередь процессами взаимодействия носителей с кристаллической решеткой и в меньшей степени изменением концентрации, процессами рассеяния на объемных дефектах смещения. К наиболее важным характеристикам ПТ относятся коэффициенты усиления K_y и шума $K_{\text{ш}}$. Повышенное изменение K_y и $K_{\text{ш}}$ при малых токах стока для напряжений затвора $U_3 \sim U_{\text{отс}}$ объясняется существенным уменьшением крутизны характеристики прибора с ростом потока нейтронов.

Радиационная стойкость кремниевых интегральных схем. При анализе радиационной стойкости изделий электронной техники все эффекты разделяются на остаточные (долговременные), связанные со структурными дефектами, и перемежающиеся, или переходные (кратковременные), обусловленные неравновесными носителями заряда.

Остаточные эффекты возникают, как правило, при воздействии непрерывного ИИ и могут быть обусловлены сложными дефектами кристаллической решетки, вызванными смещением атомов, появлением объемных зарядов при переносе носителей, а также локализованными неподвижными зарядами.

Переходные (кратковременные) эффекты существуют при воздействии импульсного ИИ и в течение относительно короткого времени после его прекращения и сопровождаются возникновением переходных процессов в схемах. Это проявляется в возникновении фототока, модуляции проводимости в полупроводниках, появлении элементарных дефектов, увеличении токов утечки. Интенсивность переходных эффектов может зависеть как от мощности дозы, так и от поглощенной дозы, а остаточные эффекты определяются в основном величиной поглощенной дозы ИИ.

Как известно, интегральные схемы (ИС) делятся на два класса: цифровые (логические, триггерные и т. д.) и аналоговые (многокаскадные усилители, стабилизаторы, преобразователи частоты и т. д.). В основе цифровых схем лежат простейшие транзисторные ключи (см. гл. 9), а в основе аналоговых (гл. 8) — простейшие усилительные каскады. Аналоговые ИС более чувствительны к ИИ по сравнению с цифровыми, что можно пояснить, используя передаточные характеристики (рис. 23.15). В транзисторном ключе, являющемся основой цифровых ИС, замкнутое состояние (точка B) характеризуется малым падением напряжения, а ра-

замкнутое (точка A) — большим падением напряжения на нем. Форма (вид) передаточной характеристики между точками A и B не играет заметной роли. Различные воздействия (температурные, радиационные, световые, электромагнитные и т. д.) изменяют передаточную характеристику в соответствии с пунктирными кривыми. Небольшие колебания около точки B из-за радиации практически не меняют значения выходного сигнала, а значит, не влияют на работу ключа. В усилительном каскаде (аналоговая ИС) используется также и участок между точками A и B . Любая «деформация» характеристики будет непосредственно отражаться на работе схемы. Отсюда видно, что аналоговые схемы более чувствительны к различного рода воздействиям, в том числе и радиационным, по сравнению с цифровыми. В литературе очень мало сведений о радиационной стойкости аналоговых интегральных микросхем, поэтому в дальнейшем рассматриваются только цифровые микросхемы.

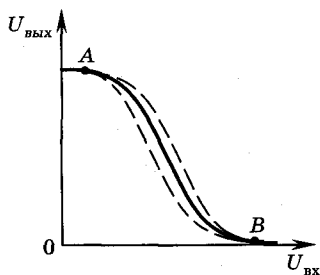


Рис. 23.15

Интегральная микросхема, в отличие от дискретных приборов, является более сложным неоднородным устройством, включающим самые различные материалы — металлы, диэлектрики, полупроводники. На основе данных по воздействию ИИ на перечисленные материалы (см. п. 23.1, 23.2), а также на дискретные электронные приборы можно ожидать при исследовании радиационной стойкости элементов микросхем следующих результатов:

- 1) изменения концентрации и подвижности носителей заряда при воздействии облучения, особенно в полупроводниковых слоях с низкой степенью легирования;
- 2) резкого увеличения сквозных токов через диэлектрик из-за возникновения дефектных локальных состояний и накопления объемных зарядов ионов в диэлектрических слоях;
- 3) из-за возникновения центров рекомбинации и захвата как в объеме, так и на поверхности переходных областей металл—диэлектрик, металл—полупроводник, полупроводник—диэлектрик и $p-n$ -переходов будут сильно изменяться свойства этих областей, в частности, возникать локальные поверхностные состояния, деформироваться профили распределения доноров и акцепторов и появляться внутренние электрические поля;
- 4) уменьшения времени жизни неосновных носителей.

Все эти эффекты будут в основном проявляться во всех элементах ИС. Стойкость микросхем хуже, чем транзисторов, из-за меньших размеров элементов и большого их числа, наличия изолирующих диэлектрических слоев, более сильного влияния поверхностных эффектов, их сложности и ряда других факторов конструктивно-технологического

плана. При этом физические эффекты, возникающие при облучении полупроводниковых устройств, можно разделить на объемные и поверхностные.

До этого рассматривались объемные эффекты. Остановимся кратко на поверхностных и ионизационных эффектах. Поверхностные эффекты изменяют величину поверхностного потенциала и увеличивают скорость поверхностной рекомбинации. Появление же ионизационных токов связано с генерацией неравновесных электронно-дырочных пар. При воздействии импульсов ИИ большой мощности указанные эффекты могут вызвать закорачивание некоторых типов ИС, т. е. переход их в режим работы, подобный режиму работы открытого диода. Такое явление часто называют *эффектом «защелкивания»*. В этом случае схема перестает функционировать до тех пор, пока не будет снято напряжение питания. Обычным типом отказа ИС при «защелкивании» являются короткие замыкания, вызывающие перегорание шины питания и выход схемы из строя. Режим «защелкивания» чаще всего наблюдается в КМОП и ТТЛ ИС (см. гл. 8, 9) с изолирующими $p-n$ -переходами; БИС КМОП (см. гл. 7—9) переходят в этот режим при меньших мощностях дозы, чем БИС ТТЛ.

В бортовой космической аппаратуре источниками ложных однократных срабатываний могут являться заряженные частицы высокой энергии (чаще всего протоны), которые, проходя через некоторую область микросхемы, формируют электрический сигнал, изменяющий логическое состояние цифровой схемы.

Многие параметры биполярных ИС функционально зависят от коэффициента передачи тока транзисторов в схеме с ОЭ $h_{21э}$, поэтому деградация этого коэффициента определяет радиационную стойкость. Изменение напряжения насыщения $U_{КЭ\text{нас}}$ транзисторов ИС при облучении также связано с изменениями других параметров. Таким образом, $h_{21э}$ и $U_{КЭ\text{нас}}$ являются теми параметрами, в которых интегрально проявляется воздействие ионизирующих излучений, поэтому радиационную стойкость биполярных ИС часто оценивают с помощью этих параметров. Для оценки влияния облучения на МДП ИС используют величины изменений пороговых напряжений $U_{\text{пор}}$, крутизны, коэффициента усиления, емкостей $C_{\text{зи}}$ и $C_{\text{зп}}$, напряжения отсечки $U_{\text{отс}}$, максимального тока стока и ряд других.

Радиационная стойкость любого типа ИС зависит от схемотехнических, конструктивных и технологических факторов.

Радиационная стойкость цифровых МДП-интегральных микросхем. Суммарная доза ионизирующего излучения более резко сказывается на МДП ИС, чем на биполярных. Она приводит к стабильному изменению порогового напряжения затвора $U_{\text{пор}}$. Это связано главным образом образованием поверхностных состояний на границе кремний—диэлект-

рик и накоплением положительных зарядов в подзатворном оксидном диэлектрике, захватываемых ловушками. Плотность поверхностных состояний резко возрастает при уровне накопленной дозы выше 10^6 рад. Все это изменяет крутизну МДП-транзисторов, сдвигает $U_{\text{пор}}$ как n -, так и p -канальных транзисторов, приводя в конечном счете к отказу схемы. КМДП ИС на кремниевых монокристаллических подложках более чувствительны к импульсному ионизирующему излучению, чем КМДП ИС на изолирующих подложках (примерно на два порядка). Это объясняется тем, что в МДП ИС вероятность возникновения радиационных дефектов и образования пар носителей в подложке и p - n -переходах, которые имеют больший объем, заметно больше, чем в КМДП ИС.

В современных БИС и СБИС используют транзисторы с характерными размерами областей порядка мкм и менее. Исследования влияния облучения на микросхемы показывают, что по мере уменьшения длины канала $L_{\text{эфф}}$ МДП-транзисторов влияние γ -излучения на изменение порогового напряжения возрастает (рис. 23.16). В приборах с коротким каналом наблюдается повышенная степень загрязнения окисла после проведения сток-истоковой диффузии, что приводит к увеличенному числу дефектов и поверхностных ловушек.

КМДП БИС, сформированные на изолирующей сапфировой подложке, обладают лучшими характеристиками по сравнению с микросхемами на полупроводниковой подложке. Практическое применение таких ИС выявило ряд проблем, связанных с их радиационной стойкостью и не характерных для КМДП-схем на полупроводниковой подложке. Одной из причин, снижающих радиационную стойкость таких ИС, является образование положительного заряда в изолирующей сапфировой подложке во время облучения, что вызывает инверсию проводимости эпитаксиального слоя кремния, в котором формируются области транзисторной структуры. В результате токи утечки КМДП-структур на сапфировой подложке при воздействии ионизирующего излучения выше, чем в схемах на полупроводниковой подложке. Накопление во время облучения положительного заряда в окисле приводит к смещению $U_{\text{з пор}}$ в сторону больших по модулю отрицательных значений для p -канальных и

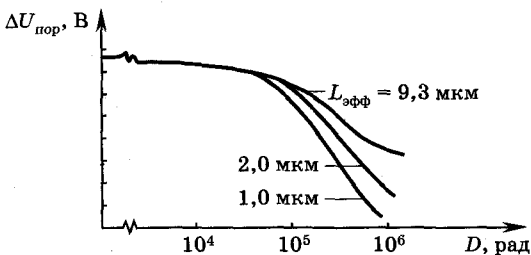


Рис. 23.16

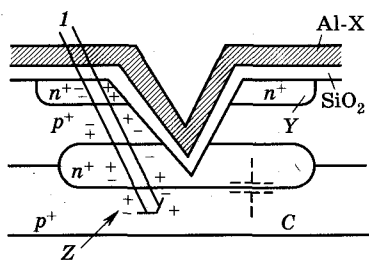


Рис. 23.17

меньших значений для *n*-канальных транзисторов. Однако при суммарной дозе радиации, превышающей 10^6 рад, пороговое напряжение для *n*-канальных транзисторов начинает возрастать, что объясняется образованием на границе Si—SiO₂ поверхностных состояний, которые компенсируют положительный заряд окисла. КМДП-схемы высокой интеграции на полупроводниковой и сапфировой подложках обладают примерно одинаковой радиационной стойкостью, первые выдерживают дозу $10^3 \dots 10^5$ рад, а вторые — $10^3 \dots 10^6$ рад. Возможности схем на сапфировой подложке еще не полностью исчерпаны.

В настоящее время наиболее изучены радиационные процессы в динамических и статических запоминающих устройствах (ЗУ), поэтому рассмотрим поведение этих схем при облучении.

Основным элементом МДП ЗУ является ячейка памяти. Распределенная емкость *C* элемента запоминающей ячейки (показана штрихами на рис. 23.17) в динамическом ЗУ хранит информацию (см. гл. 9). При прохождении ионизирующей заряженной частицы *I*, например α -частицы (α -частицы — двукратно заряженные ядра гелия, испускаемые элементами конструкции ИС), через области интегральной схемы, где существует электрическое поле, происходит разделение образовавшихся электронно-дырочных пар и накопление или электронов, или дырок в отдельных областях.

Если траектория α -частицы проходит через обедненный слой *n*-канального VМОП-транзистора (транзистора с *V*-образными канавками), то образовавшиеся электроны будут накапливаться под действием поля затвора в ячейке (рис. 23.18), а дырки — инжектироваться в подложку.

Носители заряда, образовавшиеся в подложке вне зоны обеднения, диффундируют через нее. При этом они либо исчезнут за счет рекомбинации, либо, достигнув обедненной области, будут разделены полем, т. е. электроны втянутся в обедненную область, а дырки останутся в подложке. В результате такого разделения зарядов и собирания электронов в ячейке памяти может произойти переключение ее из состояния «единицы» в состояние «ноль». Является ли заряд, генерированный α -частицами, достаточным для переключения ячейки, зависит от величины полного заряда, хранящегося в ячейке, энергии α -частицы, эффективности собирания заряда и т. д.

В рассмотренном случае только состояние логической единицы чувствительно к α -излучению, а состояние логического нуля нечувствительно. Генерированные α -излучением заряды могут распределиться по нескольким ячейкам, если расстояние между ними меньше диффузионной длины. В этом случае чувствительными к наведенному заряду могут ока-

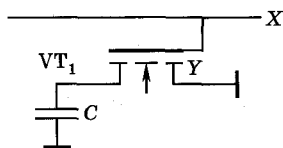


Рис. 23.18

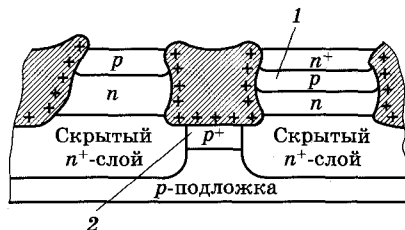


Рис. 23.19

заться другие элементы схемы, что может привести к потере как «ноля», так и «единицы».

Помимо перемежающихся (временных) отказов (сбоев) при больших дозах облучения в ИС возникают и необратимые повреждения. Необратимые повреждения в МДП ИС, обусловленные как γ -излучением, так и действием токов быстрых частиц, имеют по сути дела одну и ту же природу. Различие состоит в том, что γ -излучение приводит к отказу практически всех активных элементов ИС, а быстрые частицы вызывают отказы только отдельных приборов.

Влияние радиации на кремниевые интегральные схемы на биполярных транзисторах. Одной из основных отличительных черт активных элементов БИС, изготовленных по биполярной технологии, является повышенная стойкость их параметров к поглощаемой дозе ионизирующего излучения по сравнению с МДП ИС (особенно ИС с диэлектрической изоляцией элементов). Уровень устойчивости таких биполярных ИС к дозовым эффектам определяется параметрами окисла и границы $Si-SiO_2$. На рис. 23.19 дана изопланарная структура транзистора, являющаяся одной из наиболее распространенных (на рис. 23.19: 1 — индуцированный канал между эмиттером и коллектором, 2 — канал между скрытыми слоями). Захват дырок в окисле на границе $Si-SiO_2$ и генерация поверхностных состояний при облучении приводит к образованию канала между скрытыми n^+ -слоями из-за инверсии слоя под дном диэлектрической изоляции и канала между эмиттером и коллектором $n-p-p$ -транзистора (за счет инверсии области p -базы вдоль боковой стенки диэлектрической изоляции). Помимо этого, увеличиваются токи рекомбинации у боковой стенки диэлектрической изоляции в области p -базы. Это обусловлено возрастанием скорости поверхностной рекомбинации за счет образования быстрых поверхностных состояний на границе $Si-SiO_2$. В результате этих процессов изменяется коэффициент передачи тока.

Инверсия типа проводимости p -области базы в результате индуцирования положительного заряда в слое окисла (см. рис. 23.19, область 2) и изменение плотности поверхностных состояний вызывают рост токов утечки между коллектором и эмиттером транзисторной структуры. Пре-

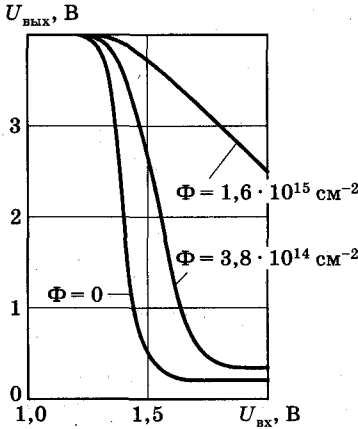


Рис. 23.20

имущественное проявление влияния канала между скрытыми областями или канала между эмиттером и коллектором $n-p-n$ -транзистора МС зависит от концентраций примесей области, предотвращающей образование канала под дном изолирующего окисла (p^+ -область), и области базы около боковой стенки. Чем выше уровень легирования p^+ -области, тем меньше утечка между скрытыми слоями.

В МС с изоляцией транзисторных структур за счет $p-n$ -переходов основным эффектом, возникающим в результате импульсной радиации, является генерация фототока в многочисленных изолирующих переходах. Значение фототока пропорционально площади перехода, поэтому основными источниками являются изолирующие переходы большого размера. Эти паразитные токи обуславливают связь между различными элементами и изменяют характеристики активных элементов микросхемы, что приводит к изменению ее параметров. В частности, ионизационные токи изменяют (уменьшают) нормальный и инверсный коэффициенты передачи тока α_N и α_I транзистора, а в режиме насыщения соответственно и напряжение $U_{КЭ\text{ нрс}} \sim U_{\text{вых}}$, которое возрастает.

В качестве примера такого влияния на рис. 23.20 показано изменение передаточной характеристики ИС ДТЛ (диодно-транзисторная логика) под действием быстрых нейтронов с энергией $E > 2,9$ МэВ.

Указанные изменения связаны с уменьшением коэффициента передачи тока в схеме с ОЭ $h_{21Э} = \beta = \frac{\alpha}{1 - \alpha}$ в зависимости от интенсивности радиационного облучения (рис. 23.21).

Радиационная стойкость оптоэлектронных приборов. Проблема радиационной стойкости достаточно остра для солнечных элементов, используемых на борту космических летательных аппаратов. Под влиянием облучения частицами высоких энергий мощность солнечных элементов понижается из-за образования дефектов в полупроводниках и изменения свойств других материалов, входящих в конструкцию данных приборов.

Продолжительность периода работоспособности бортовых солнечных батарей оценивается временем, в течение которого мощность солнечных батарей не выходит за пределы минимума, необходимого для функционирования основной аппаратуры летательного аппарата.

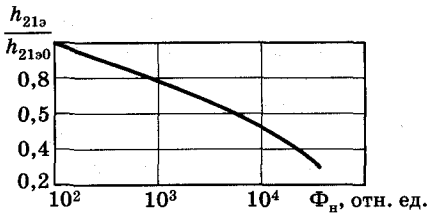


Рис. 23.21

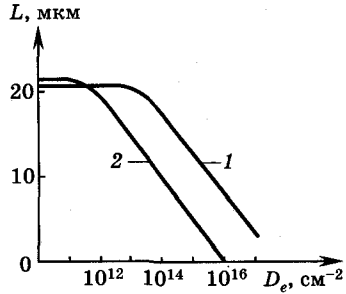


Рис. 23.22

Возникновение дополнительных дефектов в полупроводнике под воздействием ионизирующего облучения приводит к уменьшению времени жизни частиц τ и, соответственно, диффузионной длины L (поскольку $L = \sqrt{D\tau}$, см. п. 1.4) по закону

$$\frac{1}{\tau} = \frac{1}{\tau_0} + K'D,$$

или

$$\frac{1}{L^2} = \frac{1}{L_0^2} + KD, \quad (23.3)$$

где τ_0 , L_0 — соответственно время жизни и диффузионная длина до облучения; K' — коэффициент пропорциональности; $K = K'/D_{\text{диф}}$; $D_{\text{диф}}$ — коэффициент диффузии; D — доза радиации; $L = \sqrt{D\tau}$; $L_0 = \sqrt{D_{\text{диф}0}\tau_0}$. Второе выражение (23.3) следует из первого, если считать, что коэффициент диффузии не изменяется под действием радиации, т. е. $D_{\text{диф}} = D_{\text{диф}0}$, что на практике достаточно хорошо выполняется. На рис. 23.22 приведены зависимости диффузионной длины L от дозы облучения электронами D_e с энергией 1 МэВ для кремниевого солнечного элемента с р-базой (кривая 1) и n-базой (кривая 2). Из рисунка видно, что радиационная стойкость элементов с р-базой существенно выше, чем элементов с n-базой, что объясняется разной подвижностью и, соответственно, различными временами жизни электронов и дырок, а это, в свою очередь, определяет и диффузионную длину. Для повышения радиационной стойкости кремниевых солнечных элементов в них вводят литий, который нейтрализует дефекты и способствует стабилизации времени жизни при воздействии ИИ. Для защиты от ИИ перед лицевой поверхностью элемента делается тонкое прозрачное защитное покрытие, содержащее, например, церий. Для солнечных батарей предельные значения $\Phi_n = 10^{11} \dots 10^{13} \text{ см}^{-2}$, $D_\gamma = 10^3 \dots 10^4 \text{ Гр}$.

Воздействие облучения на фотоприемники сопровождается снижением коэффициента усиления, обусловленным эффектами, аналогичными тем, которые наблюдаются в диодах, транзисторах, солнечных элементах и других обычных полупроводниковых приборах.

Возникновение в полупроводнике при облучении уровней, формирующих центры безызлучательной рекомбинации, приводит к уменьшению мощности излучения светодиодов. Использование материалов с малыми временами жизни повышает радиационную стойкость этих приборов.

Радиационная стойкость оптронов. Изменение характеристик фотоприемников и светоизлучателей под действием ИИ уменьшает коэффициент передачи и ухудшает импульсные свойства *оптронов*. Допустимые потоки нейтронов Φ_n для современных оптопар составляют $10^{12} \dots 10^{14} \text{ см}^{-2}$, а дозы γ -излучения $D_\gamma = 10^5 \text{ Гр}$; при таких облучениях время задержки в оптопаре увеличивается в 2...3 раза, а потери интенсивности света в оптическом канале из-за изменения коэффициента пропускания возрастают на 10...20 дБ.

Радиационная стойкость электровакуумных приборов. Изменение характеристик этих приборов при облучении определяется в первую очередь ухудшением эмиссионных свойств катодов из-за деградации структуры поверхности, что вызывает изменение тока анода и крутизны характеристики.

Под воздействием ИИ может произойти растрескивание стекла, выделение электродами и конструкционными материалами газов, что нарушает вакуум внутри приборов. В результате этого снижаются предельно допустимые напряжения между электродами и ухудшаются другие характеристики. Импульсное ИИ вызывает в электровакуумных лампах увеличение токов утечки между электродами. Предельно допустимые дозы при этом составляют $10^5 \dots 10^7 \text{ Гр}$. Приемно-усилительные лампы в стеклянном баллоне выдерживают $\Phi_n = 10^{14} \dots 10^{15} \text{ см}^{-2}$ и $D_\gamma = 10^5 \dots 10^8 \text{ Гр}$, в металлокерамическом корпусе — $\Phi_n = 10^{18} \text{ см}^{-2}$.

Уровень радиационной стойкости изделий электронной техники зависит от конструкции, типа приборов, применяемых в них материалов и технологии их изготовления. Сравнение радиационной стойкости различных электронных приборов и ИМС позволяет произвести рис. 23.23, где введены следующие обозначения:

Д — диоды; ЭВП — электровакуумные приборы; $p-i-n$ -диоды; МБПТ — мощные биполярные транзисторы; ПТ — полевые транзисторы; МДПТ — МДП-транзисторы; ТД — туннельные диоды; ВЧТ — высокочастотные транзисторы; СБ — солнечные батареи; ИСДИ — интегральные схемы с диэлектрической изоляцией; МК — магнитные ком-

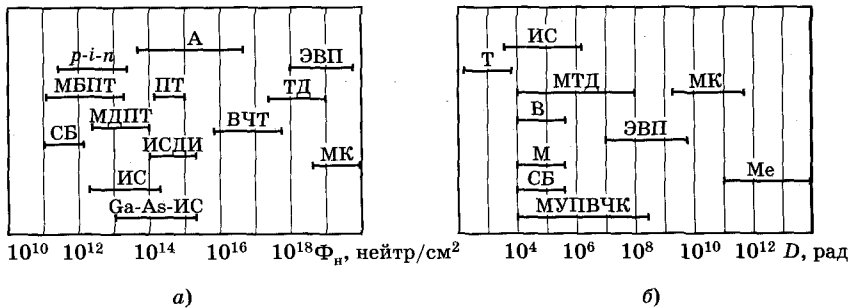


Рис. 23.23

поненты; ИС — интегральные схемы; GaAs-ИС — арсенид-галлиевые ИС; Т — транзисторы; МТД — мощные ТД; В — варисторы; Me — металл; МУПВЧК — мощные униполярные полевые ВЧ-ключи.

23.3. Надежность электронных приборов

Надежность — это свойство объекта выполнять заданные функции, сохраняя во времени значения установленных показателей в заданных пределах, соответствующих заданным режимам и условиям использования, технического обслуживания, ремонта, хранения и транспортирования.

Радиоэлектронное и электронное оборудование современных космических аппаратов включает в себя до десятка миллионов сложных устройств, поэтому проблема повышения надежности имеет крайне важное значение.

В качестве одного из основных количественных критериев надежности используется **интенсивность отказов**. При этом под отказом понимается событие, заключающееся в нарушении работоспособности изделия. Свойство электронного прибора сохранять работоспособность в течение определенного промежутка времени t называется **безотказностью**. Это свойство характеризуется вероятностью безотказной работы $p(t)$ в течение промежутка времени t . Она может быть вычислена по формуле: $p(t) = n(t)/N$, где $n(t)$ — число исправных приборов в любой рассматриваемый момент времени; N — начальное общее число испытываемых приборов; $n(t)/N$ — в любой момент времени t определяет вероятность безотказной работы отдельно взятого прибора.

Интенсивность отказов $\lambda(t)$ характеризует интенсивность выхода приборов из строя и определяется соотношением $\lambda(t) =$

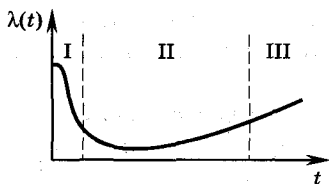


Рис. 23.24

$= \frac{\Delta n}{(N-d)\Delta t}$, где Δn — число отказов приборов за малый промежуток времени Δt ; N — общее число испытуемых приборов; d — число приборов, отказавших до начала интервала Δt .

Вероятность безотказной работы и интенсивность отказов связаны между собой соотношением

$$p(t) = \exp \left[- \int_0^t \lambda(t) dt \right].$$

Интенсивность отказов $\lambda(t)$ для электронных приборов имеет три характерных временных интервала (рис. 23.24). Начальный интервал времени (обл. I) соответствует ранним отказам (периоду приработки приборов), причины которых заключаются в быстро проявляющихся скрытых производственных дефектах. Второй участок времени (обл. II) соответствует периоду *нормальной эксплуатации*, когда значения $\lambda(t)$ мало изменяются во времени и существенно меньше по сравнению с периодом ранних отказов. На третьем временном участке (обл. III) интенсивность отказов вновь возрастает из-за износа и старения приборов. На этом участке преобладают отказы из-за изменения параметров приборов. Средняя наработка времени на один отказ t_{cp} определяется через $\lambda(t)$ по формуле $t_{cp} = 1/\lambda$.

При $\lambda = 5 \cdot 10^{-6} \text{ ч}^{-1}$ и $t = 5000 \text{ ч}$ $\lambda t = 2,5 \cdot 10^{-2}$, тогда $p(t) = p(5000) = 0,975$; $t_{cp} = 2 \cdot 10^5 \text{ ч}$, т. е. средняя наработка до отказа в данном примере составляет 20 000 ч.

Показатели надежности определяют такой важный эксплуатационный параметр приборов, как его *долговечность*, которая устанавливает гарантированную продолжительность работоспособного состояния электронных приборов в процессе их эксплуатации.

В соответствии с характером изменения параметров электронных приборов отказы могут быть разбиты на две большие группы:

- 1) *катастрофические (внезапные)*, характеризующиеся импульсным (скачкообразным) изменением параметров прибора;
- 2) *условные (постепенные)*, возникающие в результате плавного изменения параметров.

Первый вид отказов часто приводит к полной потере работоспособности приборов. Во втором случае прибор может сохра-

нять полную или частичную работоспособность, но его параметры с течением времени заметно отклоняются от первоначальных или комплексных значений.

Катастрофические отказы, относящиеся к **безусловным отказам**, как правило, обусловлены недостатками конструкции или технологии, а также условиями эксплуатации.

Конструктивные недостатки, приводящие к катастрофическим отказам, могут проявляться из-за различия коэффициентов теплового расширения материалов сочленяющихся элементов приборов, что приводит к механическим нарушениям различного вида контактов, растрескиванию кристаллов полупроводников, баллонов ламп. Это наиболее ярко проявляется в мощных приборах. Чтобы не допустить дефектов такого рода, применяют термокомпенсирующие прокладки между отдельными элементами конструкции.

В высокочастотных приборах размеры электродов составляют часто единицы и десятки мкм, особенно это относится к полупроводниковым приборам. Выводы у таких приборов выполняются тонкими проводниками (~ 8 мкм), которые невозможно сделать строго однородными по длине. В результате могут возникнуть обрывы или перегорание внутренних выводов. Перегорание может быть вызвано также значительным перепадом температур при использовании длинных внутренних выводов и технологическими нарушениями процесса пайки контактов. Кроме обрывов и перегораний выводов, в электронных приборах возможны пробой переходов или окисных слоев, короткие замыкания. Пробой возникают во всех типах электронных приборов. Они приводят к коротким замыканиям между различными электродами. Причины пробоев обусловлены нарушением и несовершенством технологии, неправильной эксплуатацией приборов, связанной с несоблюдением режимов работы приборов. Примером несовершенства технологии в планарных транзисторах является наличие в окисном слое микроскопических отверстий, являющихся одной из причин возникновения пробоев. Наличие пылинок на поверхности фоторезистора после процесса травления приводит к возникновению отверстий в месте их расположения.

Рассмотрим основные причины катастрофических отказов при неправильной эксплуатации, поскольку надежность работы приборов в схеме (устройстве) определяется не только его собственной надежностью, но и условиями его эксплуатации. До 70% всех отказов полупроводниковых приборов обусловлено непра-

вильной эксплуатацией. Наиболее часто встречающимся примером неправильной эксплуатации является использование приборов в режимах, превышающих максимально допустимые или близких к ним.

Одним из важнейших параметров, определяющих надежность прибора, является максимально допустимая температура перехода, которая определяется мощностью, рассеиваемой в приборе, температурой окружающей среды и условиями теплообмена между переходом и окружающей средой.

Повышение мощности, рассеиваемой прибором, будет приводить к разогреву элементов прибора. Увеличение температуры приводит к росту доли собственной электропроводности в примесном полупроводнике. В результате при некоторой температуре электронно-дырочный переход вообще перестает существовать. За максимально допустимую температуру принимается более низкая температура, при которой основные свойства прибора сохраняются. Температура перехода определяется материалом (германий, кремний, арсенид галлия и т. д.) и степенью легирования полупроводника. Полупроводниковые приборы с высокой степенью легирования (например, туннельные диоды) имеют наиболее высокую максимально допустимую температуру перехода ($T_{\text{макс. п}}$), а приборы на основе слаболегированных материалов (например, высоковольтные диоды) имеют самую низкую температуру $T_{\text{мин. п}}$. На практике для повышения надежности температуру окружающей среды ограничивают 30...40 °С, а рабочая мощность устанавливается менее половины от $P_{\text{макс}}$.

Максимально допустимые напряжения между электродами приборов часто устанавливаются значительно меньше напряжения пробоя перехода: из-за возможности пробоя как в объеме, так и на поверхности полупроводника; из-за требований отсутствия влияния размножения носителей в коллекторном переходе на коэффициент передачи тока; из-за технологических дефектов полупроводникового материала (неоднородности, посторонние включения и т. п.) и ряда других причин.

Максимально допустимые значения токов определяются из условий разогрева перехода с учетом неоднородностей материала. Катастрофические отказы составляют 15...20% от общего числа отказов.

Большинство отказов обусловлены постепенным изменением параметров приборов сверх допустимых отклонений. Условные отказы чаще всего вызваны физическими и химическими про-

цессами на поверхности и в объеме полупроводниковых приборов, и процессами на катодах в электронных лампах. Объемные процессы, приводящие к недопустимому изменению параметров, в полупроводниковых приборах играют незначительную роль. Постепенные изменения параметров полупроводниковых приборов связаны в основном с явлениями на поверхности (см. § 1.5). Изменение поверхностных состояний приводит к заметному изменению обратных токов, коэффициентов передачи и пробивных напряжений полупроводниковых приборов.

Одним из основных факторов, определяющих нестабильность электрофизических свойств полупроводников, является наличие влаги. Для предотвращения попадания влаги на поверхность полупроводника корпуса приборов тщательно герметизируют и, кроме того, внутри корпуса помещают *влагопоглотитель (сорбент): силикагель или алюмогель*. Другая причина нестабильности свойств поверхности связана с движением ионов щелочных металлов (натрия, кальция, калия) в слое окисла. Поверхностная плотность этих ионов может достигать 10^{14} ионов/см². Они образуют в пленке двуокиси кремния подвижные объемные заряды, влияющие на свойства переходов и параметры приборов, особенно при повышенных температурах.

Таким образом, влияние условий работы сказывается на интенсивности как катастрофических, так и условных отказов. Для транзисторов такое влияние можно проиллюстрировать на основе зависимостей, представленных на рис. 23.25, где отражено воздействие нормализованной температуры T_n и рассеиваемой мощности P_K на величину λ ($P_{K \text{ макс}}$ — максимальная допустимая мощность на коллекторе). Нормализованная температура определяется следующим образом:

$$T_n = \frac{T_{\text{раб}} - T_{\text{сн}}}{T_{\text{макс}} - T_{\text{сн}}},$$

где $T_{\text{раб}}$ — рабочая температура; $T_{\text{макс}}$ — максимально допустимая температура; $T_{\text{сн}}$ — температура начала снижения электрического режима.

Таким образом, при длительной и безотказной (надежной) работе при-

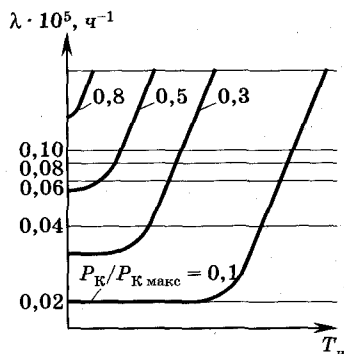


Рис. 23.25

боров целесообразно эксплуатационные режимы выбирать ниже максимально допустимых. Для определения надежности приборов проводятся механические, климатические испытания и испытания на долговечность.

В заключение можно привести конкретные значения интенсивности отказов электронных приборов. Период приработки (участок I зависимости $\lambda(t)$ на рис. 23.24) для различных электронных приборов колеблется от единиц до сотен часов. Заметно выраженного увеличения интенсивности отказов, соответствующих периоду старения и износа (участок III, рис. 23.24) для полупроводниковых приборов не наблюдается, что и отражает большую долговечность этих приборов. Интенсивность отказов $\lambda(t)$, соответствующая периоду нормальной эксплуатации при номинальных режимах (участок II, рис. 23.24), для диодов и транзисторов лежит в диапазоне $10^{-7} \dots 10^{-3} \text{ ч}^{-1}$, а для электронных ламп — $10^{-4} \dots 10^{-5} \text{ ч}^{-1}$.



Контрольные вопросы

1. Рассказать об эксплуатационных условиях, параметрах и режимах электронных приборов.
2. Характеристики и параметры ионизирующих излучений.
3. Каковы основные типы радиационных дефектов в твердых телах?
4. Объяснить механизмы влияния ионизирующих излучений на параметры полупроводников.
5. Рассказать о радиационной стойкости транзисторов и интегральных схем.
6. Радиационная стойкость цифровых МДП-интегральных микросхем.
7. Каково влияние радиации на характеристики кремниевых ИС на биполярных транзисторах?
8. Радиационная стойкость оптоэлектронных приборов.
9. Радиационная стойкость электровакуумных приборов.
10. Надежность электронных приборов.

Приложения

Приложение 1

Основные параметры Si, GaAs и Ge

Параметр (при $T = 300$ К)	Кремний Si	Арсенид галлия GaAs	Германий Ge
Собственное удельное сопротивление ρ , Ом · см	$2,3 \cdot 10^5$	10^8	47
Ширина запрещенной зоны ΔE_g , эВ	1,12	1,42	0,66
Эффективная масса электронов по отношению к массе свободного электрона m_n/m_0	0,33	0,07	0,22
То же для дырок m_p/m_0	0,55	0,5	0,39
Эффективная плотность состояний, см ⁻³ : в зоне проводимости N_c в валентной зоне N_v	$2,8 \cdot 10^{19}$ 10^{19}	$4,7 \cdot 10^{17}$ $7 \cdot 10^{17}$	10^{19} $6 \cdot 10^{18}$
Собственная концентрация n_i , см ⁻³	$1,45 \cdot 10^{10}$	$1,8 \cdot 10^6$	$2,4 \cdot 10^{13}$
Подвижность, см ² /(В · с): электронов μ_n дырок μ_p	1500 450	8500 400	3900 1900
Коэффициент диффузии, см ² /с: электронов D_n дырок D_p	36 13	290 12	100 45

Окончание прил. 1

Параметр (при $T = 300 \text{ K}$)	Кремний Si	Арсенид галлия GaAs	Германий Ge
Дрейфовая скорость насыщения $v_{\text{нас}}$, см/с: электронов дырок	10^7 $8 \cdot 10^6$	$6 \cdot 10^6$ —	$6 \cdot 10^6$ $6 \cdot 10^6$
Электрическое поле пробоя $E_{\text{пр}}$, В/см	$3 \cdot 10^5$	$4 \cdot 10^5$	10^5
Диэлектрическая проницаемость $\epsilon_0 \epsilon$, пФ/см	1,05	1,15	1,42
Температурный коэффициент энергии запрещенной зоны ТК ΔE_g , эВ/К	$-2,4 \cdot 10^{-4}$	$-4,3 \cdot 10^{-4}$	$-3,9 \cdot 10^{-4}$

Приложение 2

Основные уравнения, используемые для анализа работы электронных приборов

Физические явления, определяющие процессы преобразования сигналов и энергии в большинстве типов электронных приборов, связаны с переносом носителей заряда. Явления переноса включают такие процессы, как диффузия, электропроводность, теплопроводность, термоэлектрические явления и т. д. Все указанные процессы определяют величины токов и напряжений в электронных приборах. Поэтому одной из кардинальных задач при изучении приборов является определение и вычисление таких физических параметров, как ток, напряжение, напряженности электрических и магнитных полей, а также установление функциональных зависимостей между ними.

В приборах важны не индивидуальные свойства носителей, а их коллективное поведение. Когда в электронном газе имеется электрическое поле, вызванное нарушением квазинейтральности или приложенное извне, электроны перераспределяются таким образом, чтобы экранировать поле на расстоянии порядка дебаевской длины λ_D . Таким образом, дебаевская длина определяет предельное расстояние x , на котором начинается коллективное поведение, т. е. λ_D отделяет область коллективного описания ($\lambda_D > x$) от области описания индивидуальных частиц ($\lambda_D \leq x$).

Большинство наблюдаемых в приборах величин определяются средним поведением ансамбля частиц. Поэтому подходящим общим теоретическим описанием является статистическое описание. Задача микроскопического статистического подхода состоит в нахождении *кинетических коэффициентов* — подвижности, дифференциальной и статической проводимостей, коэффициента диффузии и т. д. в зависимости от параметров, определяющих свойства вещества (эффективные массы носителей, температура решетки, энергия ионизации и т. д.), и условий (напряжение, граничные условия и т. д.), реализуемых в конкретных приборах.

Кинетические коэффициенты получаются в результате усреднения по статистическому ансамблю, который характеризуется функцией распределения $f(\vec{r}, \vec{v}, t)$, зависящей от скорости частиц \vec{v} , координаты \vec{r} и времени t . Смысл функции $f(\vec{r}, \vec{v}, t)$ состоит в том, что произведение $f(\vec{r}, \vec{v}, t) d\vec{r} d\vec{v}$ представляет собой вероятное число заряженных частиц в элементе объема $d\vec{r} d\vec{v}$

где $L_{in}[f_{e0}]$ — интеграл неупругих соударений, учитывающий процессы возбуждения электронных и колебательных (для молекулярного газа) состояний и ионизации, N — концентрация нейтральных атомов и молекул массы M , $\sigma_T(u)$ — транспортное сечение рассеяния, определяемое энергией (скоростью) частиц.

При записи (П2.5) считается, что электрическое поле \mathcal{E} и его градиент направлены вдоль оси x , и, следовательно, состояние электронного газа в направлениях, перпендикулярных оси x , является пространственно-однородным, т. е. $f_e = f_e(x, u, \theta, t)$. (Здесь θ — угол между направлением электрического поля и вектором скорости.) Возможны также дальнейшие упрощения уравнения Больцмана, такие, как, например, предположение об однородности и стационарности электрического и магнитного полей.

При низких давлениях, когда длина свободного пробега электронов и ионов сравнима с размерами приэлектродных слоев и, тем более, межэлектродного промежутка, двучленное приближение для решения уравнения Больцмана становится неприменимым. Наиболее адекватной процедурой получения картины разряда является при этом метод Монте-Карло решения кинетического уравнения. В расчетах методом Монте-Карло разыгрывается вероятный сценарий движения индивидуальных заряженных частиц в поле с учетом упругих и неупругих столкновений, включая ионизирующие.

Обычно при вычислениях следят за траекториями каждой из частиц, что даже при оптимальном алгоритме приводит к огромным вычислительным затратам. Самосогласованные расчеты, включающие решение уравнения для поля, требуют огромных затрат даже в случае тлеющего разряда постоянного тока и применяются лишь для расчета катодного слоя. Поэтому в последнее время значительное развитие получили так называемые PIC-MCC алгоритмы, позволяющие значительно, по сравнению с обычным методом Монте-Карло, сократить объем вычислений и ресурсов оперативной памяти. PIC (*particle-in-cell*) методы основываются на различных вариантах метода крупных частиц, успешно используемого для решения широкого круга задач. Основная идея метода крупных частиц — замена реальной «лабораторной» плазмы некоторой модельной плазмой псевдочастиц с большими массой и зарядом, каждая из которых состоит из очень многих реальных частиц.

Квазичастицы движутся в соответствии с законами классической механики в самосогласованном электрическом поле и участвуют в столкновениях с нейтральными молекулами, как и обычные электроны и ионы. Поскольку плотность квазичастиц в модельной плазме ($\sim 10^3 \dots 10^5 \text{ см}^{-3}$) значительно меньше плотности зарядов в «лабораторной» плазме ($\sim 10^8 \dots 10^{11} \text{ см}^{-3}$), существенно легче реализовать на практике полностью самосогласованный алгоритм, в котором траектории квазичастиц и столкновительные процессы с их участием рассчитываются с использованием метода Монте-Карло. Поэтому PIC-метод, комбинированный с методом Монте-Карло (PIC-МСС метод), активно применяется для самосогласованного моделирования разрядов низкого давления.

На основе микроскопического уравнения Больцмана можно вывести более простые макроскопические гидродинамические уравнения. Использование гидродинамической модели означает, что мы заменяем электронный газ некоторой сплошной заряженной жидкостью, характеризуемой такими параметрами, как средняя плотность, средняя скорость, давление и коэффициент вязкости, который определяется частотой столкновений. Макроскопические уравнения гидродинамической модели получают из уравнения Больцмана путем вычисления моментов функции распределения. Вычисление моментов функции f означает умножение f на $Q = 1$, \vec{v} , $\vec{v}\vec{v}$ и т. д. и интегрирование по пространству скоростей. Если зависимость $c(f) = \left(\frac{\partial f}{\partial t}\right)_{\text{т}}$ описывает полное влияние на функцию распределения упругих столкновений, возбуждения, ионизации, прилипания и рекомбинации, то получаем уравнение

$$\int Q \left(\frac{\partial f}{\partial t}\right) d\vec{v} + \int Q \vec{v} \vec{\nabla}_{\vec{r}} f d\vec{v} + \int Q (F/m) \vec{\nabla}_{\vec{v}} f d\vec{v} = \int Q c(f) d\vec{v}. \quad (\text{П}2.6)$$

Нулевой момент функции распределения (при $Q = 1$) определяет среднюю плотность (концентрацию) электронов

$$n(\vec{r}, t) = \int f(\vec{r}, \vec{v}, t) d\vec{v}. \quad (\text{П}2.7)$$

Используя нулевые моменты уравнения Больцмана (при $Q = 1$ третий интеграл в левой части (П2.6) равен нулю [40]) и функции распределения, получаем из (П2.6) уравнение непрерывности

$$\frac{\partial n}{\partial t} + \vec{\nabla}_{\vec{r}}(n\vec{v}) = n(\vec{v}_i - \vec{v}_a - \vec{v}_r) \equiv n\bar{v}, \quad (\text{П}2.8)$$

где величины \bar{v}_i , \bar{v}_a , \bar{v}_r определяют усредненное по скорости изменение числа электронов в единице объема за единицу времени вследствие ионизации, прилипания и рекомбинации. Поскольку ионизация (\bar{v}_i) увеличивает число электронов в объеме, а прилипание (\bar{v}_a) и рекомбинация (\bar{v}_r) его уменьшают, то в правую часть (П2.8) эти величины входят с разными знаками.

Если умножить левую и правую части (П2.8) на заряд частицы, то уравнение непрерывности можно записать в виде

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_{\vec{r}} \vec{j} = \rho \bar{v}, \quad (\text{П2.9})$$

где $\rho = en$ и $\vec{j} = en\vec{v}$ — объемная плотность заряда и плотность тока соответственно.

Первый момент функции распределения относительно скорости определяет среднюю скорость частиц

$$\vec{v}_0(\vec{r}, t) = \left[\frac{1}{n(\vec{r}, t)} \right] \int \vec{v} f(\vec{r}, \vec{v}, t) d\vec{v}. \quad (\text{П2.10})$$

Используя первый момент ($Q = \vec{v}$) функции распределения (П2.10) и записывая первый момент уравнения Больцмана (П2.6), получаем уравнение движения (изменения импульса)

$$\frac{d\vec{v}_0}{dt} = -\frac{e}{m} (\vec{E} + \vec{v}_0 \times \vec{B}) - \nu_m \vec{v}_0 - \frac{1}{n(\vec{r}, t)} \vec{\nabla}_{\vec{r}} \int (\vec{v}_0 - \vec{v})(\vec{v}_0 - \vec{v}) f d\vec{v}, \quad (\text{П2.11})$$

где ν_m — эффективная частота столкновений с изменением импульса.

Для точного нахождения последнего члена в правой части уравнения (П2.11), который определяет внутреннее давление электронов в рассматриваемой системе, необходимо знать более высокий (второй) момент функции распределения, задающий вектор теплового потока, выходящего из рассматриваемой системы. В случае локального равновесия, когда ведущую роль играют столкновения частиц (см. пояснение к формуле (П2.1)), можно связать внутреннее давление ионизованной среды с температурой, и тогда последний член в уравнении (П2.11) примет

вид $\left(\frac{kT_e}{m} \right) \left(\frac{1}{\rho} \right) \vec{\nabla}_{\vec{r}} \rho$, где $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана.

Комбинируя уравнения движения заряженной жидкости с уравнениями Максвелла и граничными условиями, можно рассмотреть поведение следующих взаимодействующих подсистем: электрического поля, подсистемы носителей тока, подсистемы тепловых колебаний атомов кристаллической решетки. Такая система уравнений позволяет анализировать процессы в электронных приборах, в которых реализуются явления переноса (электровакуумные приборы, включая СВЧ-приборы с динамическим управлением, полупроводниковые, газоразрядные, плазменные, оптоэлектрические приборы).

В конечном итоге основные уравнения, используемые при анализе процессов и построении ВАХ приборов, имеют вид

$$\vec{\nabla} \times \vec{\epsilon} = \frac{\partial \vec{B}}{\partial t}; \quad \vec{\nabla} \times \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t}; \quad (\text{П2.12})$$

$$\vec{\nabla} \cdot \vec{B} = 0; \quad \vec{\nabla} \cdot \vec{D} = \rho; \quad (\text{П2.13})$$

$$\frac{d\vec{v}_0}{dt} = -\frac{e}{m} (\vec{\epsilon} + \vec{v}_0 \times \vec{B}) - \nu_m \vec{v}_0 - \left(\frac{kT_e}{m} \right) \left(\frac{1}{\rho} \right) \vec{\nabla}_r \rho; \quad (\text{П2.14})$$

$$\rho = en, \quad \vec{j} = \rho \vec{v}, \quad \vec{B} = \mu \mu_0 \vec{H}, \quad \vec{D} = \epsilon \epsilon_0 \vec{\epsilon}. \quad (\text{П2.15})$$

Плотность тока \vec{j} и объемная плотность заряда ρ связаны уравнением непрерывности (П2.9).

В отсутствие столкновений уравнение непрерывности может быть получено из уравнений Максвелла (П2.12) и (П2.13). Система уравнений (П2.12—П2.15) является наиболее общей для корректного описания физических процессов в большинстве электронных приборов, включая электровакуумные приборы СВЧ с динамическим управлением.

Уравнения магнитной гидродинамики (МГД) весьма сложны для решения даже в своем самом простом одномерном нестационарном случае. Поэтому аналитические методы их исследования могут быть применены только для ряда частных случаев с дополнительными упрощающими предположениями. Однако в последнее время благодаря бурному развитию вычислительной техники и последовавшему за этим широкому внедрению численных методов, ориентированных на высокопроизводительные компьютеры, удалось достичь значительных успехов при решении задач МГД. Разработанные численные методы обладают большей универсальностью по сравнению с аналитическими методами и

позволяют находить решение с любой наперед заданной точностью. Тем не менее, следует отметить, что если способы расчета одномерных и двумерных задач МГД достаточно хорошо проработаны, то трехмерные расчеты практически важных задач чрезвычайно редки.

Одними из возможных численных методов решения уравнений МГД являются вариационные методы и методы конечных разностей. Суть первых состоит в замене в рассматриваемой области непрерывной среды, поведение которой описывается функциями непрерывного аргумента, ее разностным аналогом. Эта модель среды представляется при помощи дискретных функций, определенных в конечном числе точек. Такое множество точек называется разностной сеткой. Дифференциальные уравнения исходной модели при этом переходят в конечно-разностные соотношения. В результате непрерывная модель поведения системы заменяется или, как принято говорить, аппроксимируется набором разностных уравнений — разностной схемой.

Разностная схема должна отражать основные свойства непрерывной среды. Поэтому необходимо требовать, чтобы в первую очередь для схемы были справедливы разностные аналоги исходных основных законов сохранения. Разностные схемы, обладающие этим важным свойством, называются консервативными. Однако на реальных сетках в задачах, решением которых являются быстро меняющиеся во времени и пространстве функции, такие разностные схемы могут приводить к результатам, значительно отличающимся от истинных, вследствие наличия в этих схемах фиктивных источников энергии. Поэтому помимо требования о выполнении основных законов сохранения обычно дополнительно из физических соображений вводят ряд сеточных соотношений, позволяющих избежать различных дисбалансов. Такие схемы называются полностью консервативными.

Особенность полностью консервативных разностных схем состоит в том, что такие схемы одновременно аппроксимируют различные виды записи системы дифференциальных уравнений, каждый из которых отражает определенный физический аспект явления. Благодаря этому такие схемы правильно передают, например, соотношения между кинетической и внутренней энергией, в то время как схемы других типов порождают фиктивные источники энергии, которые на грубых сетках могут заметно исказить решение.

Вариационные методы в своей наиболее общей формулировке заменяют задачу минимизации некоего функционала, заданного на бесконечномерном линейном пространстве, задачами его минимизации на последовательности конечномерных подпространств. При этом система МГД уравнений трактуется как операторное уравнение в гильбертовом пространстве с линейным, самосопряженным и положительно определенным оператором. Современные реализации вариационных методов чаще всего являются сеточными, для которых функции отличны от нуля лишь на конечном числе ячеек выбранной сетки. В результате матрица системы уравнений является разреженной, что значительно облегчает отыскание решений исходной системы уравнений. Указанные модификации вариационных методов также носят название методов конечных элементов [41, 42].

1. *Аваев Н. А., Шишкин Г. Г.* Электронные приборы / под ред. Г. Г. Шишкина. — М.: МАИ, 1996.
2. *Электронные приборы / Дулин В. Н., Аваев Н. А., Демин В. П. и др.* / под ред. Г. Г. Шишкина. — М.: Энергоатомиздат, 1989.
3. *Зи С.* Физика полупроводниковых приборов / пер. с англ. под ред. Р. А. Суриса. — М.: Мир, 1984.
4. *Пасынков В. В., Чиркин Л. К.* Полупроводниковые приборы. — М.: Высшая школа, 1987.
5. *Шалимова К. В.* Физика полупроводников. — М.: Энергоатомиздат, 1985.
6. *Аваев Н. А., Наумов Ю. Е., Фролкин В. Т.* Основы микроэлектроники. — М.: Радио и связь, 1991.
7. *Блихер А.* Физика силовых биполярных и полевых транзисторов / пер. с англ. под ред. И. В. Грехова — Л.: Энергоатомиздат, 1986.
8. *Полупроводниковые приборы. Диоды, тиристоры, оптоэлектронные приборы. Справочник / под ред. Н. Н. Горюнова* — М.: Энергоатомиздат, 1984.
9. *Полупроводниковые приборы. Транзисторы. Справочник / под ред. Н. Н. Горюнова.* — М.: Энергоатомиздат, 1985.
10. *Мозговой Г. П., Силин В. Д., Чахмахсазян Е. А.* Математическое моделирование и макромоделирование биполярных элементов электронных схем. — М.: Радио и связь, 1985.
11. *Бубенников А. Н.* Моделирование интегральных микротехнологий, приборов, схем. — М.: Высшая школа, 1989.
12. *Жеребцов И. П.* Основы электроники. — Л.: Энергия, 1985.
13. *Тугов Н. М., Шарунич Л. С.* Оптоэлектроника. — М.: Энергоатомиздат, 1984.
14. *Блихер А.* Физика тиристоров. — Л.: Энергоиздат, 1981.
15. *Быстров Ю. А., Литвак И. И., Персианов Г. М.* Электронные приборы для отображения информации. — М.: Радио и связь, 1985.
16. *Букингем М.* Шумы в электронных приборах и системах. — М.: Мир, 1986.
17. *Окснер Э. С.* Мощные полевые транзисторы и их применение. — М.: Радио и связь, 1985.
18. *Гоноровский И. С.* Радиотехнические цепи и сигналы. — М.: Дрофа, 2006.
19. *Milnes A. G.* Semiconductor devices and integrated electronics. — Van Nostr and Reinhold company, 1980.
20. *Носов Ю. Р.* Оптоэлектроника. — М.: Радио и связь, 1990.

21. Жигарев А. А., Шамаева Г. Г. Электронно-лучевые и фотоэлектронные приборы. — М.: Высшая школа, 1982.
22. Proceedings of the 2-nd Int. Conf. on Vac. Microelectronics — Bath: Inst. phys., Conf., Ser. № 99, Section 1, 1989.
23. Андрушко Л. М., Федоров Н. Д. Электронные и квантовые приборы СВЧ. — М.: Радио и связь, 1981.
24. Березин В. М., Буряк В. С., Гутцайт Э. М., Марин В. П. Электронные приборы СВЧ. — М.: Высшая школа, 1985.
25. Шишкин Г. Г. Электрорадиотехнические приборы. — М.: Изд-во МАИ, 1992.
26. Агазаян Т. М., Аствацатурьян Е. Р., Скоробогатов П. К. Радиационные эффекты в интегральных микросхемах. — М.: Энергоатомиздат, 1989.
27. Кулакова В. М., Ладыгина Е. А., Шеховцова В. Н. Действие проникающей радиации на изделия электронной техники / под ред. Е. А. Ладыгина. — М.: Радио и связь, 1980.
28. Бобровский Ю. Л., Корнилов С. А., Кратилев И. А. и др. Электронные, квантовые приборы и микроэлектроника. — М.: Радио и связь, 1998.
29. Шишкин Г. Г. Приборы квантовой электроники. — М.: Сайнс-Пресс, 2004.
30. Шукла А. А. / под ред. А. С. Сигова. Электроника, 2005.
31. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: МФТИ, 1999.
32. Пихтин А. Н. Оптическая и квантовая электроника. — М.: Высшая школа, 2001.
33. Федоров Н. Д., Федоров Д. Н. Толковый словарь по электронике. — М.: Радио и связь, 2001.
34. Степаненко И. П. Основы микроэлектроники. — М.; СПб.: Лаб. баз. знаний «Невский Диалект», физматлит, 2001.
35. Прянишников В. А. Электроника: Курс лекций. — СПб.: Корона принт, 2000.
36. Усанов Д. А., Скрипаль А. В. Физика полупроводников. Явления переноса в структурах с туннельно-тонкими полупроводниковыми слоями. — Изд-во Саратовского университета, 1996.
37. Сигмен А. Мазеры. — М.: Мир, 1966.
38. Соклоф С. Аналоговые интегральные схемы. — М.: Мир, 1988.
39. Стил М., Вюраль Б. Взаимодействие волн в плазме твердого тела. — М.: Атомиздат, 1973.
40. Шкаровский И., Джонстон Т., Бачинский М. Кинетика частиц плазмы. — М.: Атомиздат, 1969.
41. Самарский А. А. Теория разностных схем. — М.: Наука, 1983.
42. Галлагер Р. Метод конечных элементов. Основы. — М.: Мир, 1984.
43. Драгунов В. П., Неизвестный И. Т., Гридчин В. А. Основы наноэлектроники. — М., 2006.

Список основных использованных обозначений

- σ — удельная электрическая проводимость
 ρ — удельное сопротивление
 ΔE_z — ширина запрещенной зоны полупроводника
 n_i, p_i — концентрация электронов и дырок в собственном полупроводнике
 p_{p0}, n_{n0} — равновесная концентрация основных носителей (дырок и электронов) в акцепторном и донорном примесных полупроводниках
 p_{n0}, n_{p0} — равновесные концентрации неосновных носителей заряда
 n^+, p^+ — области полупроводников с повышенной концентрацией электронов и дырок
 E_d, E_a — энергетические уровни доноров и акцепторов
 $\Delta E_d, \Delta E_a$ — энергия ионизации доноров и акцепторов
 E_n, E_v — энергетические уровни дна зоны проводимости и потолка валентной зоны
 N_d, N_a — концентрация атомов доноров и акцепторов
 N_n, N_v — эффективная плотность энергетических состояний в зоне проводимости и валентной зоне
 E_F, E_{Fn}, E_{Fp} — энергии уровня Ферми для собственного, донорного и акцепторного полупроводников
 L_D — дебаевская длина экранирования
 τ_n, τ_p — время жизни электронов и дырок
 G — скорость генерации
 v — скорость носителей
 $v_{др}$
 $v_{др.нас}$ — скорость дрейфа и дрейфовая скорость насыщения
 μ — подвижность носителей
 μ_n, μ_p — подвижность электронов и дырок
 \mathcal{E} — напряженность электрического поля
 $\mathcal{E}_{кр}, \mathcal{E}_{пор}$ — критическая и пороговая напряженности электрического поля
 T — абсолютная температура
 j — плотность тока
 j_n, j_p — плотность электронного и дырочного токов
 $j_{др}, j_{диф}$ — плотность дрейфового и диффузионного токов

D	— коэффициент диффузии
Φ_T	— тепловой потенциал
L_n, L_p	— диффузионная длина электронов и дырок
Φ_0	— контактная резкость потенциалов
l_0	— полная ширина (толщина) обедненной области ($p-n$ -перехода)
$r_{\text{диф}}$	— дифференциальное сопротивление
r_6	— объемное сопротивление базы
$I_{\text{обр}}$	— обратный ток
I_0	— тепловой ток
$U_{\text{проб}}$	— напряжение пробоя
$C_{\text{бар}}$	
$C_{\text{диф}}$	— барьерная и диффузионная емкость $p-n$ -перехода
$R_{\text{обр}}$	— сопротивление обратносмещенного $p-n$ -перехода, обусловленное током термогенерации
$U_{\text{п}}$	— падение напряжения на переходе
$I_{\text{обр. макс}}$	— максимальный обратный ток
$q\Phi_m$	— работа выхода электронов из металла
$\Delta E_{\text{п}}$	
$\Delta E_{\text{в}}$	— величины энергетических скачков при разрыве зон проводимости и валентной зоны при гетеропереходах
Φ	— полная контактная разность потенциалов гетероперехода
$U_{\text{н}}$	— напряжение на нагрузке
$I_{\text{пр. макс}}$	
$I_{\text{обр. макс}}$	— максимально допустимый постоянный прямой и обратный токи
$U_{\text{обр}}$	— падение напряжения на диоде (электрическом переходе) при обратном смещении
$U_{\text{пр. ср}}$	
$U_{\text{обр. ср}}$	— среднее значение прямого и обратного напряжения
$C_{\text{корп}}$	— емкость корпуса прибора
$C_{\text{выв}}$	— емкость выводов
λ	— рабочая длина волны
β_I	— чувствительность по току
$T_{\text{ш}}$	— шумовая температура
$P_{\text{ш}}$	— мощность шумов
Δf	— полоса пропускаемых частот
$t_{\text{вос}}$	— время восстановления обратного сопротивления
$I_{\text{вос. макс}}$	— максимальный ток восстановления
$U_{\text{ст}}$	— напряжение стабилизации
$I_{\text{ст. мин}}$	
$I_{\text{ст. макс}}$	— минимальный и максимальный токи стабилизации
f_0	— резонансная частота
θ	— фаза, фазовый сдвиг

- j_e — плотность тока эмиссии
 I_a, I_{c1} — ток анода и первой (управляющей) сетки
 U_a, U_{c1} — потенциал анода и первой (управляющей) сетки
 U_{c2}, U_{c3} — потенциалы экранирующей и защитной сеток
 R_0 — сопротивление лампы при постоянном токе
 S, S_c — крутизна анодно-сеточной и сеточной характеристик триода
 D — проницаемость лампы
 μ — статический коэффициент усиления
 $C_{ск}, C_{ак},$
 C_{ac} — входная (сетка — катод), выходная (анод — катод), проходная (анод — сетка) емкости электронных ламп
 F_m — сила Лоренца
 B — индукция магнитного поля
 U_{a2} — напряжение на втором аноде
 $U_{пл}$ — потенциал отклоняющих пластин
 $\varepsilon_э, \varepsilon_m$ — чувствительность к отклонению соответственно диэлектростатической и магнитной систем
 $\omega_{ц}$ — электронная циклотронная частота
 X — параметр группировки
 θ_3 — угол пролета зазора между сетками резонатора
 θ_0, θ_x — абсолютный угол пролета электрона
 i_{ex} — конвекционный ток
 I_0 — ток электронного луча, потока
 J_n — функция Бесселя первого рода n -го порядка
 P_e — мощность, развиваемая электронным потоком
 I_{m1}, I_{m2} — амплитудные значения токов во входном и выходном резонаторах
 K_p — коэффициент усиления клистрона по мощности
 $v_{ф}, v_{фn}$ — фазовая скорость электромагнитной волны и n -й пространственной гармоники
 $v_{гр}$ — групповая скорость электромагнитной волны
 $\lambda_{э0}, K_{э0}$ — длина и волновое число замедленной основной (нулевой) пространственной гармоники
 $R_{св}$ — сопротивление связи электронного потока с замедляющей системой
 $\varepsilon_{zm,n}$ — амплитуда напряженности продольного поля n -й пространственной волны
 Φ_e — относительный угол пролета электронов
 v_e — скорость электронного потока
 $G_{вх},$
 $G_{вых}$ — входная и выходная проводимости ЛБВО
 K — коэффициент усиления
 $K_{ш}, P_{ш}$ — коэффициент и мощность шумов
 P_c — мощность входного сигнала

- $v_{\Phi(-1)}$ — фазовая скорость первой обратной гармоники
 $U_{кр}$
 $B_{кр}$ — напряжение анода и индукция магнитного поля критического режима магнетрона
 n — номер вида колебаний анодного блока магнетрона
 $v_{\Phi p n}$ — фазовая скорость p -й пространственной гармоники для определенного n -го вида колебаний
 U_p, I_p — напряжение и ток разряда
 W_{12}
 W_{21} — вероятности индуцированных переходов при поглощении и излучении квантов
 B_{12}
 B_{21} — коэффициенты Эйнштейна для индуцированного поглощения и излучения
 A_{21} — коэффициент Эйнштейна для спонтанного излучения
 w_{21} — вероятность спонтанного излучения
 n_1, n_2 — населенности уровней E_1 и E_2
 g_1, g_2 — статистические веса, или кратности вырождения уровней E_1 и E_2
 N_1, N_2 — общее число переходов поглощения и излучения соответственно
 ρ_v — спектральная плотность энергии
 β — постоянная, характеризующая величину потерь на единицу длины активной среды
 α — коэффициент усиления на единицу длины активной среды
 ω_q, λ_q — частота и длина волны q -го вида колебаний
 q — номер вида колебаний
 L — длина активной среды
 r_1, r_2 — коэффициенты отражения зеркал
 ξ — коэффициент дифракционных потерь
 η_l — КПД лазера
 $E_{эл}$
 $E_{кол}$
 $E_{вр}$ — энергии электронных, колебательных и вращательных состояний

Оглавление

Предисловие	3
Введение	6

Раздел 1

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ

Глава 1. ФИЗИКА ПОЛУПРОВОДНИКОВ	12
1.1. Равновесная концентрация свободных носителей заряда	12
1.2. Неравновесные носители заряда	20
1.3. Электропроводность полупроводников	23
1.4. Законы движения носителей заряда в полупроводниках	27
Глава 2. КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ. ЭЛЕКТРИЧЕСКИЕ ПЕРЕХОДЫ	31
2.1. Основные определения. Классификация электрических переходов	31
2.2. Физические процессы в электронно-дырочных переходах	32
2.3. Вольт-амперная характеристика $p-n$ -перехода	41
2.4. Электрическая модель $p-n$ -перехода	51
2.5. Переходные процессы в $p-n$ -переходе	54
2.6. Контакты металл—полупроводник	58
2.7. Гетеропереходы	64
Глава 3. ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ	71
3.1. Общие сведения и классификация диодов	71
3.2. Выпрямительные диоды	76
3.3. Импульсные диоды	82
3.4. Стабилитроны	84
3.5. Варикапы	87
3.6. Туннельные диоды	90
3.7. Лавинно-пролетные диоды	95
3.8. Диоды Ганна	104
Глава 4. ВИПОЛЯРНЫЕ ТРАНЗИСТОРЫ	115
4.1. Общие вопросы. Устройство, режимы работы транзисторов	115

4.2.	Физические процессы в нормальном активном режиме. Коэффициенты передачи тока	118
4.3.	Модель Эберса—Молла. Статические характеристики биполярных транзисторов	122
4.4.	Биполярный транзистор как линейный четырехполюсник. Параметры транзистора	132
4.5.	Эквивалентные схемы	139
4.6.	Переходные и частотные характеристики биполярного транзистора	142
4.7.	Импульсный режим работы. Транзисторный ключ ...	148
4.8.	Разновидности биполярных транзисторов	154
Глава 5.	ТИРИСТОРЫ	159
5.1.	Общие сведения. Устройство. Режимы работы	159
5.2.	Основные физические процессы. Принцип действия ..	161
5.3.	Переходные процессы и импульсные свойства тиристоров	167
5.4.	Разновидности тиристоров. Параметры и модели тиристоров	173
Глава 6.	ПОЛЕВЫЕ ТРАНЗИСТОРЫ	178
6.1.	Общие сведения	178
6.2.	Формирование канала в МДП-транзисторах	180
6.3.	Общие принципы управления проводимостью канала в полевых транзисторах. Статические вольт-амперные характеристики	185
6.4.	Моделирование полевых транзисторов	194
6.5.	Полевой транзистор как линейный четырехполюсник. Параметры транзисторов. Эквивалентные схемы ...	196
6.6.	Частотные и импульсные свойства полевых транзисторов	200
6.7.	Разновидности полевых транзисторов. Силовые комбинированные транзисторы	201

Раздел 2

ИНТЕГРАЛЬНЫЕ СХЕМЫ

Глава 7.	АКТИВНЫЕ И ПАССИВНЫЕ ЭЛЕМЕНТЫ ИНТЕГРАЛЬНЫХ СХЕМ	205
7.1.	Общие сведения. Термины и определения	205
7.2.	Электрическая изоляция элементов полупроводниковых ИС	208
7.3.	Особенности биполярных транзисторов ИС	211
7.4.	Транзисторы ИС типа $p-n-p$	217
7.5.	Интегральные диоды	219
7.6.	Полевые транзисторы ИС	222
7.7.	Пассивные элементы ИС	225

Глава 8.	АНАЛОГОВЫЕ ИНТЕГРАЛЬНЫЕ СХЕМЫ	228
8.1.	Общие сведения. Термины и определения	228
8.2.	Источники стабильного тока, напряжения и опорного напряжения	230
8.3.	Дифференциальные усилители	238
8.4.	Операционные усилители	246
Глава 9.	ЦИФРОВЫЕ ИНТЕГРАЛЬНЫЕ СХЕМЫ	249
9.1.	Особенности цифровых интегральных схем	249
9.2.	Элементарные (базовые) цифровые схемы на биполярных транзисторах	251
9.3.	Простейшие инверторные (ключевые) схемы на МДП-транзисторах	253
9.4.	Бистабильные схемы и триггеры	258
9.5.	Логические элементы на биполярных транзисторах...	260
9.6.	Логические элементы на полевых транзисторах	269
9.7.	Элементы полупроводниковых запоминающих устройств	271
Глава 10.	НАНОЭЛЕКТРОНИКА И ФУНКЦИОНАЛЬНАЯ ЭЛЕКТРОНИКА	276
10.1.	Общие положения. Возможности нанoeлектроники и функциональной электроники.	276
10.2.	Особенности нанoeлектронных приборов	279
10.3.	Приборы с зарядовой связью	285
10.4.	Элементы акустоэлектроники	293
10.5.	Элементы СБИС на цилиндрических магнитных доменах	296

Раздел 3

ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ С ЭЛЕКТРОСТАТИЧЕСКИМ И ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ

Глава 11.	ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ С ЭЛЕКТРОСТАТИЧЕСКИМ УПРАВЛЕНИЕМ	301
11.1.	Общие сведения	301
11.2.	Основы эмиссионной электроники	302
11.3.	Электронно-управляемые лампы	307
11.4.	Мощные генераторные и модуляторные лампы	317
11.5.	Эквивалентные схемы электронных ламп	321
11.6.	Вакуумные интегральные схемы	323
Глава 12.	ЭЛЕКТРОННО-ЛУЧЕВЫЕ ПРИБОРЫ	328
12.1.	Классификация, устройство и принцип действия электронно-лучевых приборов	328

12.2.	Электронный прожектор с электростатической фокусировкой	330
12.3.	Электронный прожектор с магнитной фокусировкой	336
12.4.	Отклоняющие системы	338
12.5.	Экраны электронно-лучевых трубок	342
12.6.	Особенности электронно-лучевых приборов различного назначения	347
Глава 13.	ЭЛЕКТРОВАКУУМНЫЕ ПРИБОРЫ СВЧ С ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ И ПРОДОЛЬНОМ ВЗАИМОДЕЙСТВИЕМ — ПРИБОРЫ ТИПА О	361
13.1.	Общие сведения	361
13.2.	Клистроны	363
13.3.	Лампы бегущей волны. Устройство ЛБВО. Замедляющие системы	377
13.4.	Физические процессы в ЛБВО	384
13.5.	Параметры и характеристики ЛБВО	386
13.6.	Лампы обратной волны типа О	390
Глава 14.	ЭЛЕКТРОННЫЕ ПРИБОРЫ СВЧ С ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ И СКРЕЩЕННЫМИ ПОЛЯМИ — ПРИБОРЫ ТИПА М	395
14.1.	Общие сведения	395
14.2.	Магнетроны. Движение электронов в скрещенных полях	396
14.3.	Колебательная система магнетрона	400
14.4.	Характеристики и параметры магнетронов	405
14.5.	Приборы магнетронного типа	410

Раздел 4

ПРИБОРЫ ОТОБРАЖЕНИЯ ИНФОРМАЦИИ. ОПТОЭЛЕКТРОННЫЕ ПРИБОРЫ

Глава 15.	ГАЗОРАЗРЯДНЫЕ ПРИБОРЫ И ИНДИКАТОРЫ	416
15.1.	Общие сведения	416
15.2.	Элементарные процессы в газовых разрядах	418
15.3.	Приборы тлеющего разряда. Плазменные панели	422
15.4.	Приборы дугового разряда	433
15.5.	Электривакуумные и электролюминесцентные индикаторы	434
15.6.	Пассивные индикаторы. Жидкокристаллические индикаторы	439
15.7.	Сравнение индикаторов различного типа	449
Глава 16.	ОПТОЭЛЕКТРОННЫЕ ПРИБОРЫ	450
16.1.	Общие сведения	450
16.2.	Оптические явления в полупроводниках	452
16.3.	Светодиоды	458

16.4.	Полупроводниковые фотоприемники	465
16.5.	Оптопары	491
16.6.	Солнечные преобразователи	496
16.7.	Электровакуумные фотоприемники	508

Раздел 5

ПРИБОРЫ КВАНТОВОЙ ЭЛЕКТРОНИКИ

Глава 17.	ОСНОВЫ КВАНТОВОГО УСИЛЕНИЯ	521
17.1.	Индукцированные и спонтанные переходы	521
17.2.	Усиление в квантовых системах	526
17.3.	Основные элементы устройства квантовых генераторов	526
17.4.	Условия баланса мощности и фаз в лазерах (оптических квантовых генераторах)	534
17.5.	Спектр и характеристики излучения квантовых генераторов	535
Глава 18.	КВАНТОВЫЕ ПРИБОРЫ СВЧ-ДИАПАЗОНА (МАЗЕРЫ) ..	540
18.1.	Активное вещество твердотельных мазеров. Парамагнитные уровни энергии	540
18.2.	Квантовые парамагнитные усилители	546
18.3.	Охлаждение мазеров и их параметры	552
18.4.	Квантовые генераторы СВЧ-диапазона	554
Глава 19.	ГАЗОВЫЕ ЛАЗЕРЫ	558
19.1.	Общие сведения	558
19.2.	Процессы создания инверсии населенностей в газовых лазерах	559
19.3.	Атомарные газовые лазеры. Гелий-неоновый лазер ...	561
19.4.	Ионные лазеры	566
19.5.	Молекулярные газовые лазеры	568
19.6.	Разновидности газовых лазеров	570
Глава 20.	ТВЕРДОТЕЛЬНЫЕ И ЖИДКОСТНЫЕ ЛАЗЕРЫ	574
20.1.	Особенности активных сред	574
20.2.	Системы накачки	577
20.3.	Рубиновый лазер	583
20.4.	Лазеры на средах, активированных неодимом	586
20.5.	Жидкостные лазеры	589
20.6.	Режимы работы и основные характеристики твердотельных лазеров	594
Глава 21.	ПОЛУПРОВОДНИКОВЫЕ ИНЖЕКЦИОННЫЕ ЛАЗЕРЫ ..	603
21.1.	Полупроводниковые материалы, используемые в источниках излучения	603

- $I_{Э}, I_{Б}, I_{К}$ — токи эмиттера, базы и коллектора
 $U_{Э}, U_{Б}, U_{К}$ — напряжения эмиттера, базы и коллектора
 α, β — коэффициенты передачи тока в схеме с общей базой и общим эмиттером
 α_N, α_I — нормальный и инверсный коэффициенты передачи тока в схеме с общей базой
 $I_{КВО}, I_{ЭВО}$ — обратный токи коллекторного и эмиттерного переходов в схеме с общей базой
 $U_{ЭБ}, U_{КБ}$ — напряжения на эмиттерном и коллекторном переходах в схеме с ОБ
 $U_{БЭ}, U_{КЭ}$ — напряжения на базе и коллекторе в схеме с ОЭ
 U_A — напряжение Эрли
 $h_{11}, h_{12}, h_{21}, h_{22}$ — h -параметры транзистора как четырехполюсника
 $r_{Э}, r_{К}$ — дифференциальное сопротивление эмиттерного и коллекторного переходов
 $y_{11}, y_{12}, y_{21}, y_{22}$ — y -параметры транзистора как четырехполюсника
 $C_{Эдиф}, C_{Кдиф}$ — диффузионные емкости эмиттерного и коллекторного переходов
 $C_{Эбар}, C_{Кбар}$ — барьерные емкости эмиттерного и коллекторного переходов
 $\tau_{Кэф}$ — эффективное время жизни неосновных носителей в коллекторе
 $t_{прБ}$ — время пролета носителей через базу
 $R_{К}$ — сопротивление в цепи коллектора
 $\Phi_{пов}$ — поверхностный потенциал
 $U_{И}, U_{З}, U_{С}$ — напряжения истока, затвора и стока
 $Q_{пов}$ — поверхностный заряд
 $\epsilon_{пов}$ — поверхностная напряженность электрического поля
 $I_{И}, I_{С}, I_{З}$ — токи истока, стока и затвора
 $U_{Снас}$ — напряжение насыщения стока
 $U_{Зпор}$ — пороговое напряжение на затворе
 $\Phi_{пор}$ — пороговый потенциал

$R_{зи}$	—	активное сопротивление затвора—исток и затвор—сток
$R_{зс}$	—	активное сопротивление затвора—исток и затвор—сток
$C_{зи}$	—	емкость перекрытия затвора с истоком и стоком соответственно
$C_{зс}$	—	емкость перекрытия затвора с истоком и стоком соответственно
r_k	—	сопротивление канала
S	—	статическая крутизна
R_i	—	внутреннее сопротивление
$C_{вых}$	—	выходная емкость
$g_0, g_{кэ}$	—	дифференциальная выходная проводимость
$U_{вх}$	—	входное напряжение
$U_{вых}$	—	выходное напряжение
F	—	коэффициент обратной связи операционного усилителя
U^0, U^1	—	напряжения (уровни) логического нуля и единицы
$U_{л}$	—	логический перепад
$U_{оп}$	—	опорное напряжение
$Q_{п макс}$	—	максимальная величина зарядового пакета
$f_{т мин}$	—	минимальная тактовая частота
$P_{рас}$	—	рассеиваемая мощность
$\lambda_{ак}$	—	длина акустической волны
$U_{вкл}$	—	напряжение включения
$I_{вкл}$	—	ток включения
$U_{откр}$	—	напряжение в открытом состоянии
$U_{уд}, I_{уд}$	—	напряжение и ток удержания соответственно
U_A, I_A	—	напряжение и ток тиристора
U_y	—	напряжение управляющего электрода
M	—	коэффициент размножения носителей
ν, λ	—	частота и длина волны оптического излучения
R, T, α	—	коэффициенты отражения, прохождения и поглощения оптического излучения
$\sigma_{ф}$	—	проводимость при облучении полупроводника светом
$I_{ф}$	—	фототок
K_R	—	коэффициент усиления фоторезистора
$n_{ф}$	—	объемная плотность фотоносителей
$t_{нр (сп)}$	—	время нарастания (спада) фототока
D^*	—	обнаружительная способность
$\sqrt{i_m^2}$	—	среднеквадратическое значение шумового тока
I_T	—	тепловой ток фотоприемника
η	—	квантовая эффективность, светоотдача и КПД
$P_{ф}$	—	мощность оптического излучения
P_m, U_m	—	максимальные значения выходных мощности, напряжения и тока фотоэлемента
I_m	—	максимальные значения выходных мощности, напряжения и тока фотоэлемента

21.2.	Инжекционные полупроводниковые лазеры на основе гомопереходов	606
21.3.	Инжекционные лазеры на основе гетеропереходов ...	611
21.4.	Разновидности полупроводниковых лазеров	617

Раздел 6

ВОПРОСЫ ПРИМЕНЕНИЯ И ЭКСПЛУАТАЦИИ ЭЛЕКТРОННЫХ ПРИБОРОВ

Глава 22.	ШУМЫ ЭЛЕКТРОННЫХ ПРИБОРОВ	622
22.1.	Общие сведения	622
22.2.	Источники шумов	623
22.3.	Методы описания шумов	629
22.4.	Шумы электронных приборов различного типа	633
Глава 23.	ЭКСПЛУАТАЦИОННЫЕ УСЛОВИЯ РАБОТЫ, РЕЖИМЫ И НАДЕЖНОСТЬ ЭЛЕКТРОННЫХ ПРИБОРОВ ..	644
23.1.	Эксплуатационные условия, параметры и режимы работы электронных приборов	644
23.2.	Радиационная стойкость электронных приборов	647
23.3.	Надежность электронных приборов	673
Приложение 1.	Основные параметры Si, GaAs и Ge	679
Приложение 2.	Основные уравнения, используемые для анализа работы электронных приборов	681
Литература		690
Список основных использованных обозначений		692

Учебное издание

Шишкин Геннадий Георгиевич
Шишкин Алексей Геннадиевич

ЭЛЕКТРОНИКА

Учебник для вузов

Зав. редакцией *Т. Д. Гамбурцева*
Редактор *И. Я. Ицхоки*
Художественное оформление *Ю. В. Христин*
Технический редактор *И. В. Грибкова*
Компьютерная верстка *А. В. Маркин*
Корректор *Г. И. Мосякина*

Санитарно-эпидемиологическое заключение
№ 77.99.60.953.Д.010105.09.08 от 22.09.2008.

Подписано к печати 15.04.09. Формат 60×90 ¹/₁₆.
Бумага типографская. Гарнитура «Школьная». Печать офсетная.
Усл. печ. л. 44,0. Тираж 3000 экз. Заказ № 6779.

ООО «Дрофа». 127018, Москва, Суцеский вал, 49.

**По вопросам приобретения продукции
издательства «Дрофа» обращаться по адресу:**
127018, Москва, Суцеский вал, 49.
Тел.: (495) 795-05-50, 795-05-51. Факс: (495) 795-05-52.

Торговый дом «Школьник».
109172, Москва, ул. Малые Каменчики, д. 6, стр. 1А.
Тел.: (495) 911-70-24, 912-15-16, 912-45-76.

Сеть магазинов «Переплетные птицы»:
Тел.: (495) 912-45-76;

140408, Московская обл., г. Коломна,
Голутвин, ул. Октябрьской революции, 366/2.
Тел.: (495) 741-59-76.

Интернет-магазин: <http://www.drofa.ru>

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Можайский полиграфический комбинат».
143200, г. Можайск, ул. Мира, 93.

Данная книга — первый учебник, по мнению рецензентов, полностью соответствующий Государственному стандарту высшего профессионального образования и программе дисциплины «Электроника» по направлению 210300 «Радиотехника», не имеющий аналогов в современной учебной литературе.

Книга охватывает все разделы современной электроники: физику полупроводников и электрических переходов; физические процессы, устройство и характеристики полупроводниковых диодов, биполярных и полевых транзисторов, тиристоров, фотоэлектрических и оптоэлектронных приборов, электровакуумных приборов, включая ЭВП СВЧ с динамическим управлением и различные типы индикаторов. В данный учебник входят такие актуальные вопросы, как базовые элементы аналоговых и цифровых интегральных схем, принципы и устройства квантовой электроники, нанoeлектроники и функциональной электроники. Рассматриваются проблемы шумов, радиационной стойкости и надежности приборов.

Учебник предназначен для подготовки бакалавров, магистров и инженеров (специалистов) широкого профиля в области электронной техники, радиотехники и связи.

Авторы учебника — известные в нашей стране и за рубежом ученые, круг научных интересов которых связан с электроникой, физикой плазмы, электродинамикой заряженных сред и др.

Шишкин Геннадий Георгиевич — профессор МАИ (ТУ), заслуженный деятель науки и техники РФ, почетный работник профессионального высшего образования РФ.

Шишкин Алексей Геннадиевич — старший научный сотрудник факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова.

