БИБЛИОТЕКА «ПРИБОРОСТРОЕНИЕ»

В.В. Андреев, В.П. Жалнин, А.А. Столяров

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОЭЛЕКТРОНИКИ



BAUMANPRESS
ИЗДАТЕЛЬСТВО МГТУ им. Н.Э. БАУМАНА

ПРИОРИТЕТ2030^
ЛИДЕРАМИ СТАНОВЯТСЯ

БИБЛИОТЕКА «ПРИБОРОСТРОЕНИЕ»

Серия основана в 2022 году

Том 3

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

В.А. Шахнов	–	главный редактор, д-р техн. наук, проф., чл.-корр. РАН, зав. кафедрой МГТУ им. Н.Э. Баумана
А.И. Власов	–	зам. главного редактора, канд. техн. наук, доцент МГТУ им. Н.Э. Баумана
В.А. Быков	–	д-р техн. наук, профессор, президент Нанотехнологического общества России
М.Ф. Булатов	–	д-р физ.-мат. наук, профессор, директор Научно-технологического центра уникального приборостроения РАН
Дун Гэ	–	профессор Университета Цинхуа (КНР)
А.Н. Козырев	–	начальник производства ООО «Резонит»
А.С. Комшин	–	д-р техн. наук, профессор МГТУ им. Н.Э. Баумана
И.Г. Мироненко	–	д-р техн. наук, профессор СПбГЭТУ «ЛЭТИ»
С.Б. Нестеров	–	д-р техн. наук, профессор, президент Российского научно-технического вакуумного общества имени С.А. Векшинского
В.В. Одиноков	–	заместитель генерального директора по науке АО НИИ точного машиностроения
И.А. Покровский	–	исполнительный директор Ассоциации разработчиков и производителей электроники
Ю.С. Сахаров	–	д-р техн. наук, профессор Университета «Дубна»
А.Л. Стемпковский	–	д-р техн. наук, профессор, академик РАН, научный руководитель Института проблем проектирования в микроэлектронике РАН
В.Б. Стешенко	–	канд. техн. наук, доцент, заместитель генерального конструктора по электронной компонентной базе АО «Российские космические системы»

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»

В.В. Андреев, В.П. Жалнин, А.А. Столяров

Физические основы микро- и наноэлектроники

Допущено Федеральным учебно-методическим объединением в системе высшего образования по укрупненной группе специальностей и направлений подготовки высшего образования 09.00.00 «Информатика и вычислительная техника» в качестве учебного пособия

BAUMANPRESS
Москва
ИЗДАТЕЛЬСТВО
МГТУ им. Н. Э. Баумана
2023

УДК 621.382
ББК 32.85
А65

Издание доступно в электронном виде по адресу
https://bmstu.press/catalog/item/7900/

Издание учебно-методического комплекса по укрупненным группам специальностей и направлений «Информатика и вычислительная техника» и «Приборостроение» подготовлено в рамках Программы стратегического академического лидерства «Приоритет–2030» МГТУ им. Н.Э. Баумана

Рецензенты:
д-р техн. наук, профессор *Н.К. Юрков*, заведующий кафедрой «Конструирование и производство радиоаппаратуры» Пензенского государственного университета, заслуженный деятель науки РФ;
д-р техн. наук, профессор *Д.Ю. Муромцев,* кафедра «Конструирование радиоэлектронных и микропроцессорных систем» Тамбовского государственного технического университета

Андреев, В. В.

А65 Физические основы микро- и наноэлектроники / В. В. Андреев, В. П. Жалнин, А. А. Столяров. — Москва : Издательство МГТУ им. Н. Э. Баумана, 2023. — 292, [2] с. : ил. (Библиотека «Приборостроение». Т. 3).

ISBN 978-5-7038-6098-4
ISBN 978-5-7038-6042-7 (т. 3)

Учебное пособие по дисциплине «Физические основы микро- и наноэлектроники» содержит рекомендации по организации и проведению лекций, практических занятий, семинаров, лабораторных работ и деловых игр, перечень учебных видео- и аудиоматериалов, слайдов, типовых плакатов и другие дидактические материалы для работы профессорско-преподавательского состава и студентов по данной дисциплине.

Соответствует стандартам WorldSkills по компетенции «Производство и инженерные технологии», а также минимуму содержания и уровню подготовки студентов по укрупненным группам специальностей и направлений «Информатика и вычислительная техника» и «Приборостроение». Может представлять интерес для специалистов, занимающихся вопросами приборостроения и созданием цифровых систем.

Материалы находятся в свободном доступе — их можно читать, загружать, копировать, распространять, печатать и ссылаться на их полные или частичные тексты с указанием авторства без каких-либо ограничений. Тип лицензии CC: Attribution 4.0 International (CC BY 4.0).

УДК 621.382
ББК 32.85

Уважаемые читатели! Пожелания, предложения, а также сообщения о замеченных опечатках и неточностях Издательство просит направлять по электронной почте: info@bmstu.press

ISBN 978-5-7038-6098-4
ISBN 978-5-7038-6042-7 (т. 3)

© Андреев В.В., Жалнин В.П., Столяров А.А., 2023
© Оформление. Издательство МГТУ им. Н.Э. Баумана, 2023

Предисловие

Учебно-методические комплексы (УМК) дисциплин по направлению «Приборостроение» разработаны в рамках адаптации образовательных программ высшего образования к требованиям Перечня специальностей и укрупненных направлений подготовки высшего образования — бакалавриата и магистратуры (Приказ Министерства образования и науки РФ от 12 сентября 2013 г. № 1061 «Об утверждении перечней специальностей и направлений подготовки высшего образования» (с изменениями и дополнениями)). Учебно-методические комплексы направлены на изучение и освоение компетенции Future Skills в рамках актуализированных образовательных программ высшего образования в соответствии с запросом цифровой трансформации промышленности.

В настоящее время приборы различных классов, назначения, функциональной сложности используются в различных отраслях промышленности, медицины, сельского хозяйства, оборонной техники, в быту. Это микроэлектронная элементная база, сенсорная техника, приборы для научных исследований, радиоэлектронные и электронно-вычислительные приборы и устройства, системы управления разного уровня, бытовая электроника.

Основной особенностью предметной области направления «Приборостроение» является междисциплинарный характер, который требует особых методических приемов и подбора соответствующего научного и учебного материала. Современные образовательные программы должны обеспечивать приобретение студентами профессиональных навыков и компетенций, необходимых для эффективной и самостоятельной работы в приборной индустрии. В связи с этим актуальной задачей является разработка и издание УМК, которые обеспечат учебно-методическую поддержку подготовки бакалавров и магистров по основным образовательным программам высшего профессионального образования по тематическому направлению «Приборостроение» образовательными учреждениями высшего профессионального образования на территории РФ.

Цель создания данного комплекта УМК — повышение эффективности междисциплинарной подготовки бакалавров и магистров путем распространения передового опыта среди вузов, осуществляющих подготовку по укрупненным группам специальностей и направлений «Информатика и вычислительная техника» и «Приборостроение», а также внедрения компонентов вариативного маршрутного обучения на базе адаптированного, сетевого учебно-методического комплекса дисциплин.

Авторы выражают благодарность рецензентам: научному руководителю кафедры «Конструирование радиоэлектронных и микропроцессорных систем» Тамбовского государственного технического университета Д.Ю. Муромцеву и заведующему кафедрой «Конструирование и производство радиоаппаратуры» Пензенского государственного университета Н.К. Юркову, замечания которых позволили улучшить содержание УМК.

Входящие в УМК учебные пособия обеспечат обучение студентов по основным образовательным программам высшего профессионального образования для направления подготовки «Приборостроение» образовательными учреждениями высшего профессионального образования на территории РФ.

Сетевые версии УМК дисциплин на основе Web-версии, соответствующие стандарту SCORM 2004, 3rd edition, размещены на портале Электронной образовательной системы МГТУ им. Н.Э. Баумана (e-learning.bmstu.ru).

Авторы будут признательны читателям за все замечания по содержанию УМК, которые следует направлять по адресу: 105005, Москва, 2-я Бауманская ул., ИУ-4, МГТУ им. Н.Э. Баумана.

Список сокращений

БИС — большая электронная схема
ВАХ — вольт-амперная характеристика
ГМС — гигантское магнитосопротивление
ГФЭ — газофазная эпитаксия
ДНК — дезоксирибонуклеиновая кислота
ДЭГ — двумерный электронный газ
МДП — металл–диэлектрик–полупроводник
МЛЭ — молекулярно-лучевая эпитаксия
МОП — металл–оксид–полупроводник
ОПЗ — область поверхностного заряда
ОПОП — основная профессиональная образовательная программа
ПАНИ — полианилин
ПЭО — полиэтиленоксид
ПЗС — прибор с зарядовой связью
РТД — резонансно-туннельный диод
СУОС — самостоятельно устанавливаемый образовательный стандарт
СЭМ — сканирующий электронный микроскоп
ТМС — туннельное магнитосопротивление
ТЭМ — термоэлектрический модуль
УГО — условно-графическое изображение
УНТ — углеродная нанотрубка
ФСС — фосфорно-силикатное стекло
ЭДС — электродвижущая сила
ЯМР — ядерный магнитный резонанс

Термины и определения

Боровский радиус — радиус электронной орбиты в атоме водорода в модели Бора, соответствующий нижайшему энергетическому уровню

Вырожденный газ — газ высокой плотности, в котором распределение частиц газа лимитировано квантово-механическим принципом тождественности частиц (принципом Паули для фермионов). Заполнение состояний такого газа описывается распределением Ферми

Длина волны де Бройля — длина волны $\lambda = h/p$, сопоставляемая в волновой механике с частицей с импульсом p (h – постоянная Планка)

Длина свободного пробега — среднее расстояние, проходимое носителем заряда между последовательными актами рассеяния на дефектах

Инверсионный слой — приповерхностная область полупроводника МДП-структуры, в которой концентрация неосновных (по отношению к объему полупроводника) носителей превышает концентрацию заряженных примесей

Кантилевер	—	механический зондовый датчик для АСМ-методик, представляющий собой балку, жестко закрепленную с одной стороны и имеющую острую иглу с радиусом закругления порядка 10 нм
Квантовая нить	—	наноструктура с одномерным электронным газом, в которой движение носителей пространственно ограничено по двум степеням свободы (1D-объект)
Квантовая точка	—	наноструктура, в которой движение носителей ограничено по всем пространственным степеням свободы. Локализованные в такой трехмерной квантовой яме носители образуют «нульмерный» газ с полностью дискретным спектром, подобно атомному спектру
Квантовая яма	—	наноструктура в виде пленки, толщина которой соизмерима с длиной волны де Бройля (2D-объект)
Контактная разность потенциалов	—	разность потенциалов, возникающая при контакте двух материалов в гомо- или гетероструктурах. Равна разности работ выхода образующих гетеропереход материалов
Мемри́стор	—	пассивный электронный элемент, двухполюсник, способный изменять свое сопротивление в зависимости от протекшего через него электрического заряда
Молетроника	—	электроника, в которой в качестве элементов микроэлектронных схем используются отдельные органические молекулы или даже их фрагменты
Невырожденный газ	—	газ малой плотности, в котором можно пренебречь влиянием квантово-механического принципа тождественности частиц. В случае *фермионного газа* это соответствует слабому влиянию принципа Паули на распределение частиц газа по состояниям, и газ с хорошей точностью может быть описан классической статистикой Больцмана
Обедненный слой	—	область вблизи p–n-перехода в гетероструктурах или в МДП-структурах, в которой концентрация свободных носителей мала по сравнению с концентрацией заряженных примесей
Обогащенный слой	—	приповерхностная область полупроводника, образующаяся при приложении электрического поля, притягивающего к поверхности основные носители заряда
Плотность состояний	—	число энергетических состояний в системе размерности D, приходящихся на единичный интервал энергии в расчете на единицу D-мерного объема
Подвижность	—	отношение дрейфовой скорости носителей к электрическому полю
Спинтроника	—	направление наноэлектроники, в котором для использования в электронных устройствах применяются не только заряд электрона в твердом теле, но и его собственный квантово-механический момент — спин

1. КОНСПЕКТ ЛЕКЦИЙ

Основная цель дисциплины: изучение физических законов и явлений, на которых основано функционирование элементов современной микро- и наноэлектроники.

Задачи дисциплины — формирование знаний, навыков и умений по следующим направлениям:

– элементы зонной теории твердых тел;

– статистика электронов и дырок в полупроводниках;

– электропроводность металлов и полупроводников;

– электрофизические свойства *p*–*n*-переходов и структур металл–диэлектрик–полупроводник;

– основные принципы работы полевых транзисторов;

– методы исследования физических процессов и определения параметров полупроводниковых материалов и структур.

При освоении дисциплины планируется формирование компетенций, предусмотренных ОПОП на основе СУОС по направлению подготовки 12.03.01 «Приборостроение» (уровень бакалавриата).

В конце каждой лекции приведены тесты для контроля знаний студентов, состоящие из вопросов и трех ответов на них, из которых следует выбрать один правильный.

1.1. ВВЕДЕНИЕ В ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОЭЛЕКТРОНИКИ

Цель лекции: введение в физические основы микро- и наноэлектроники, ознакомление с содержанием курса и его структурой.

Приборостроение — это отрасль науки и техники, занимающаяся исследованиями, проектированием и совершенствованием методов разработки, производства и применения интегрированных приборных систем. К разделам приборостроения относятся также высокоточное приборостроение, высокоточное машиностроение и биотехнологии с разрешающей способностью допусков размерных цепей объектов, составляющей десятки единиц нанометров. Электроника, включая микро- и наноэлектронику, в свою очередь, являются основой создания элементной базы приборостроения.

Началом развития электроники можно считать 1906 г., когда в США Ли де Форестом была изобретена трехэлектродная лампа. В 1950-х годах появились транзисторы, в 1960-х — интегральные микросхемы.

Микроэлектроника как область электроники, изучающая проблемы создания электронных устройств в миниатюрном исполнении, появилась в начале 1960-х годов. Развитие микроэлектроники показало, что она является одним из важнейших факторов, определяющих не только научно-технический прогресс, но и социальное развитие общества в целом. В частности, рынок электронного оборудования в последние годы динамично развивался со средней исторической скоростью 7 % в год. Это как минимум в 2 раза превышает среднюю скорость роста ВВП в развитых странах, поэтому высказываются прогнозы, что электронная промышленность станет самой большой отраслью в мире в XXI в. Так, мировой рынок электронного оборудования превысил 1,2 триллиона долл., обогнав рынок автомобилей.

Развитие микроэлектроники сопровождалось увеличением степени интеграции при одновременном уменьшении минимального размера элементов интегральных схем (рис. 1.1).

Основные особенности развития микроэлектроники во второй половине XX в.: резкое снижение себестоимости (применительно к элементу обработки единицы информации) при беспрецедентном снижении энергопотребления с одновременным улучшением характеристических параметров.

Гигантский рост производства изделий микроэлектроники основывался во многом на увеличении диаметра полупроводниковых пластин, что само по себе представляло чрезвычайно сложную материаловедческую задачу, требующую решения большого комплекса научно-технических задач на каждом

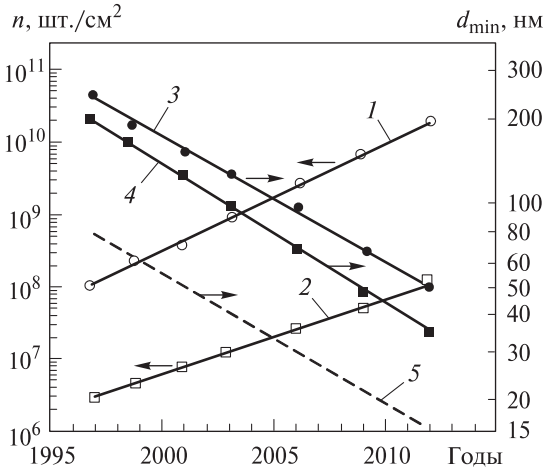


Рис. 1.1. Тенденции изменения плотности упаковки и минимального размера элементов интегральных схем:

1, *3* — динамическая память; *2*, *4* — микропроцессоры; *5* — результаты прикладных исследований

этапе увеличения диаметра. В настоящее время широко применяемый диаметр кристаллографически совершенных монокристаллов кремния, обладающих высокой чистотой и однородностью (менее одного атома примеси на 10^7 атомов кремния), применяемых в производстве, составляет 300 мм. В ближайшее время планируется, а на отдельных предприятиях уже и осуществляется, переход на еще большие диаметры — 400 и 450 мм.

Магистральным направлением развития микроэлектроники является уменьшение топологических норм транзисторных структур, т. е. переход от микрометрового в нанометровый диапазон линейных размеров или создание малоразмерных структур. В этом случае говорят о наноэлектронике. Иная физическая картина получается при рассмотрении наноэлектронных устройств. Уменьшение размеров на несколько порядков практически изменяет физические основы работы наноэлементов.

В наноэлементах используются уже не электроны как частицы, переносящие электрический заряд, а их волновые функции. Процессы дрейфа и диффузии, характерные для микроэлектроники, в наноэлектронных элементах отсутствуют.

В качестве материалов изделий микроэлектроники используются легированные полупроводники. В наноэлектронике применяются гетероструктуры, наноструктурированные материалы, кластеры, органические материалы. Технология формирования наноструктур или инженерия волновых функций основана на процессах направленного роста, методах сканирующей туннельной микроскопии, атомной силовой микроскопии. Если плотность размещения активных элементов в интегральных схемах составляет порядка 10^8 см^2, то

в устройствах наноэлектроники она может достигать $10^9...10^{10}$ элементов на квадратный сантиметр. Под наноэлектроникой понимают направление электроники, в котором изучаются физические явления и процессы взаимодействия электронов с электромагнитными полями, а также разработка технологии (нанотехнология) создания приборов и устройств, в которых это взаимодействие используется для передачи, обработки и хранения информации. Наноэлектроника представляет собой логическое развитие микроэлектроники.

Наноэлектроника является, с одной стороны, неким продолжением развития традиционных идей микроэлектроники, связанным со схемотехническими методами обработки информации, с другой стороны, новые идеи в квантовых явлениях и эффектах открывают новые возможности создания приборов нового поколения.

Ожидается создание к 2023 г. транзисторов с шириной затвора 20 Å (21 нм) и менее. Традиционные полевые транзисторы, выполненные по субмикронной технологии, приобретают новые свойства. Уменьшение толщины слоя оксида, длины канала приводит к квантованию поперечного движения в канале. В результате образуется квазидвумерный газ носителей заряда, увеличивается их подвижность, увеличивается туннельный ток. При длинах затворов транзисторов до 20 нм они сравнимы с длиной когерентности и длиной волн де Бройля при комнатной температуре.

В квантовых структурах дискретность состояний системы определяется дискретностью безынерционного квантово-размерного энергетического спектра. Это фундаментальное свойство квантовых приборов. Как отмечалось ранее, в полупроводниковых квантовых приборах обработка, хранение и преобразование информации происходят путем контролируемой передислокации волновой функции в полупроводниковой наноструктуре. Один из способов перестройки волновой функции заключается в передислокации максимума вероятности нахождения электрона из одной части квантовой системы в другую. Эта передислокация осуществляется под действием внешнего напряжения, например, с помощью квантовых точек, разделенных туннельно-прозрачными барьерами.

Тесты к лекции 1.1

1. ***Что такое микроэлектроника?***

а) область электроники, изучающая проблему создания электронных устройств в миниатюрном исполнении;

б) область электроники, изучающая проблему создания электронных устройств на радиолампах;

в) область электроники, изучающая проблему создания макроволновых электронных устройств.

2. ***Что является магистральным направлением развития микроэлектроники?***

а) уменьшение топологических норм транзисторных структур;

б) разработка современного программного обеспечения;

в) уменьшение стоимости транзисторных структур.

1.2. ЭЛЕМЕНТЫ КВАНТОВОЙ МЕХАНИКИ И ФИЗИЧЕСКОЙ СТАТИСТИКИ

Цель лекции: ознакомление с основами квантовой механики и физической статистики.

1.2.1. ВОЛНОВЫЕ СВОЙСТВА ЧАСТИЦ. УРАВНЕНИЕ ШРЕДИНГЕРА

К началу XX в. было установлено, что атомные явления не могут быть описаны как движение частиц или как чисто волновые процессы. Так, в явлениях дифракции, интерференции проявляется волновая природа света. В фотоэлектрических явлениях, эффекте Комптона (изменение частоты или длины волны фотонов при их рассеянии электронами) свет ведет себя как частица. В 1924 г. французский физик де Бройль выдвинул гипотезу, что с каждым телом должна быть связана плоская волна:

$$\lambda = \frac{h}{p} = \frac{h}{mv},$$

где h — постоянная Планка, равная $6{,}6 \cdot 10^{-34}$ Дж·с; p — импульс; m — масса микрочастицы; v — скорость микрочастицы.

Гипотеза де Бройля получила убедительное экспериментальное подтверждение. На волновых свойствах микрочастиц основана электронная микроскопия, нейтронография. Микрочастицы — электроны, протоны — нельзя представить в виде дробинки, уменьшенной до соответствующих размеров. Качественным отличительным признаком микрочастиц является органическое сочетание в них корпускулярных и волновых свойств.

Поскольку микрочастицы обладают волновыми свойствами, закон их движения должен описываться волновым уравнением. Впервые такое уравнение было записано Э. Шредингером (Австрия) для микрочастицы, движущейся в силовом поле и обладающей потенциальной энергией $u(x, y, z, t)$, в виде

$$-i\hbar \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m}\left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2}\right) - u(x, y, z, t)\Psi,$$

где $i = \sqrt{-1}$; $\hbar$ — постоянная Планка, деленная на 2π.

Функция $\Psi(x, y, z, t)$ является решением этого уравнения и называется волновой функцией. Она имеет следующий физический смысл: произведе-

ние Ψ и функции Ψ^*, комплексно сопряженной с Ψ, пропорционально вероятности того, что в момент времени t микрочастица может быть обнаружена в выделенном объеме dV. Обозначим вероятность обнаружения микрочастицы в объеме dV через $\omega(x, y, z, t)dV$. Тогда

$$\omega(x, y, z, t)\,dV = \Psi(x, y, z, t)\Psi^*(x, y, z, t)\,dV.$$

Интеграл $\int \omega dV$, взятый по всему объему, равен единице, так как он выражает достоверный факт, что микрочастица находится в этом объеме, следовательно,

$$\int \Psi\Psi^* dV = 1.$$

Это условие называется условием нормировки, а волновые функции, удовлетворяющие ему, называются нормированными.

Закон движения микрочастицы постоянно определяется заданием функции Ψ в каждый момент времени в каждой точке пространства.

Потенциальная энергия u в общем случае является функцией координат и времени. Однако в большинстве практических задач u является функцией только координат. В этом случае волновую функцию $\Psi(x, y, z, t)$ представляют в виде произведения функций $\Psi(x, y, z)$ и $\varphi(t)$:

$$\Psi(x, y, z, t) = \Psi(x, y, z)\,\varphi(t). \tag{1.1}$$

Рассмотрим движение микрочастицы вдоль оси x, тогда уравнение Шредингера можно записать в следующем виде:

$$-i\hbar\frac{\partial\Psi(x,t)}{\partial t} = \frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} - u(x)\Psi(x,t). \tag{1.2}$$

Подставим уравнение (1.2) в уравнение (1.1):

$$-i\hbar\frac{\partial\big(\Psi(x)\varphi(t)\big)}{\partial t} = \frac{\hbar^2}{2m}\frac{\partial^2\Psi(x)\varphi(t)}{\partial x^2} - u(x)\big(\Psi(x)\varphi(t)\big).$$

Разделим обе части полученного выражения на $\big(\Psi(x)\varphi(t)\big)$:

$$-i\hbar\frac{1}{\varphi(t)}\frac{\partial\varphi(t)}{\partial t} = \frac{\hbar^2}{2m}\frac{1}{\Psi(x)}\frac{\partial^2\Psi(x)}{\partial x^2} - u(x).$$

Тогда левая часть полученного уравнения зависит только от t, а правая — только от x. Они могут быть равны друг другу в том случае, если каждая равна одной и той же постоянной величине E.

Можно показать, что E есть полная энергия частицы. Приравнивая левую и правую части уравнения $-E$, запишем

$$\frac{\hbar^2}{2m}\frac{1}{\Psi(x)}\frac{\partial^2\Psi(x)}{\partial x^2}-u(x)=-E,$$

откуда

$$\frac{\partial^2\Psi(x)}{\partial x^2}+\frac{2m}{\hbar^2}(E-u_x)\Psi(x)=0; \tag{1.3}$$

$$-i\hbar\frac{1}{\varphi(t)}\frac{\partial\varphi(t)}{\partial t}=-E.$$

Отсюда получим

$$\frac{\partial\varphi(t)}{\partial t}=-i\frac{E}{\hbar}\varphi(t). \tag{1.4}$$

В общем случае уравнение (1.3) будет содержать вторые производные по другим координатам:

$$\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}+\frac{\partial^2\Psi}{\partial z^2}+\frac{2m}{\hbar^2}(E-u)\Psi=0. \tag{1.5}$$

Через оператора Лапласа уравнение (1.5) можно записать в виде

$$\Delta\Psi+\frac{2m}{\hbar^2}(E-u)\Psi=0.$$

Функция $\Psi(x, y, z)$, зависящая только от координат, называется амплитудной волновой функцией, а уравнение (1.5) — амплитудным уравнением Шредингера.

Было доказано, что при движении микрочастицы в ограниченной области пространства амплитудное уравнение Шредингера имеет решение только при определенных значениях энергии E: $E_1, E_2, ..., E_n,$ называемых собственными значениями энергии частицы. Волновые функции $\Psi_1, \Psi_2, ..., \Psi_n,$ соответствующие этим энергиям, называются собственными волновыми функциями.

Решением уравнения (1.4) является

$$\Psi_n=\exp\left(-i\frac{E_n}{\hbar}t\right),$$

где E_n — одно из собственных значений энергии. Функция $\varphi(t)$ выражает зависимость волновой функции $\Psi(x, y, z, t)$ от времени. Эта зависимость является гармонической с частотой $\nu_n=\frac{E_n}{h}$ или $\omega_n=\frac{E_n}{\hbar}$.

Если потенциальная энергия является функцией только координат, то решение уравнения Шредингера может быть представлено в виде

$$\Psi(x, y, z, t) = \Psi(x, y, z)\exp(-i\omega t).$$

В этом случае вероятность обнаружения частицы в элементе объема

$$\omega dV = \Psi\Psi^* dV$$

и не зависит от времени. Следовательно, распределение вероятности в пространстве является стационарным. Состояния микрочастиц, удовлетворяющие этому условию, называются стационарными состояниями. Амплитудное уравнение описывает стационарное состояние микрочастиц.

1.2.2. ТЕОРЕМА ФЕЛИКСА БЛОХА

Ф. Блох (1905–1983) — один из основоположников современной физики, лауреат Нобелевской премии по физике за создание теории ядерного магнитного резонанса (1952).

С 1924 по 1927 г. он учился в Федеральном институте, где среди его учителей были П. Дебай и Э. Шредингер, затем продолжил обучение в Лейпцигском университете у В. Гейзенберга. В 1928 г. в Лейпциге он получил докторскую степень, защитив диссертацию, посвященную проводимости электронов в металлах. В этой диссертации он сформулировал теорему, определявшую вид волновых функций электрона в металлах (функции Блоха), которая лежит в основе одноэлектронной теории твердых тел.

Согласно теореме Блоха, волновые функции, являющиеся решениями одноэлектронного уравнения Шредингера с периодическим потенциалом, имеющим период решетки, представляют собой плоские стоячие волны, модулированные периодической функцией, т. е.

$$\psi_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r})e^{i\vec{k}\vec{r}},$$

где $u_{\vec{k}}(\vec{r})$ — некоторая периодическая функция с периодом решетки, зависящая от квазиволнового вектора $\vec{k}$, характеризующего квантовое состояние электрона в кристалле.

Электрон в твердом теле оказывается под влиянием всех близлежащих ядер и электронов, поэтому величина его энергетического уровня неопределенная.

Некоторые из этих энергетических уровней настолько близки друг другу, что возможно существование нескольких различных положений с практически одинаковой энергией. С учетом этого удобно рассматривать не отдельные уровни, а энергетические зоны, состоящие из множества уровней с близкими значениями энергий.

Хотя такие зоны образованы из разрешенных уровней, переходы электронов от уровня к уровню в пределах энергетической зоны подчинены законам квантовых чисел. Энергетические уровни, расположенные за преде-

лами зоны, т. е. в области неразрешенных уровней, недоступны электрону. Движение электронов в пределах энергетической зоны совершается только за счет приобретения или потерь энергии.

1.2.3. СООТНОШЕНИЯ НЕОПРЕДЕЛЕННОСТИ ГЕЙЗЕНБЕРГА

К микрочастицам, обладающим волновыми свойствами, применять понятия классической механики, например, понятия координат частицы и ее импульса, можно лишь в ограниченной степени.

Пусть частица движется вдоль оси x и обладает импульсом p_x. Такой частице соответствует волна $\lambda = \frac{h}{p_x}$, являющаяся по своей сущности протяженным объектом. Монохроматическая волна простирается по оси x от $-\infty$ до $+\infty$. Следовательно, интервал локализации микрочастицы Δx равен бесконечности, т. е. микрочастица, имеющая определенный импульс p_x, не имеет определенной координаты x. Можно показать, что микрочастица, имеющая определенную координату, не имеет определенного импульса. В отличие от классической частицы, состояние микрочастицы не может быть охарактеризовано заданием одновременно определенных координат и составляющих импульса. Задать состояние микрочастицы можно, лишь допуская неопределенность в значениях координат и значениях составляющих импульса. Количественно эта неопределенность описывается соотношениями, записанными Гейзенбергом в 1927 г.:

$$\Delta x \Delta p_x \geq h; \quad \Delta y \Delta p_y \geq h; \quad \Delta z \Delta p_z \geq h.$$

Поскольку $p = mv$,

$$\Delta x \Delta v_x \geq \frac{h}{m}; \quad \Delta y \Delta v_y \geq \frac{h}{m}; \quad \Delta z \Delta v_z \geq \frac{h}{m}.$$

Из соотношений неопределенности следует, что чем точнее определяются координаты микрочастицы, тем неопределеннее становятся составляющие импульса. Поэтому бессмысленно говорить о траектории движения микрочастицы, т. е. о совокупности положений движущейся частицы в пространстве.

Соотношение неопределенности существует и между энергией и временем:

$$\Delta E \Delta t \geq h,$$

где Δt — время, в течение которого частица обладает энергией $E \pm \Delta E$.

Из соотношения неопределенности для энергии и времени следует, что неопределенность энергии возрастает при уменьшении времени пребывания микрочастицы в данном энергетическом состоянии.

Тесты к лекции 1.2

1. ***В чем суть гипотезы де Бройля?***

а) с каждым телом должна быть связана плоская волна;

б) волновой вектор микрочастицы не определен;

в) плоская волна определяется волновым вектором.

2. ***Что описывает уравнение Шредингера?***

а) закон движения микрочастицы;

б) вероятность прохождения частицы через барьер;

в) энергетический спектр микрочастицы.

3. ***Чему равна вероятность обнаружения частицы в элементе объема?***

а) квадрату волновой функции;

б) квадрату волнового вектора;

в) квадрату координат.

4. ***В чем суть теоремы Блоха?***

а) волновые функции, являющиеся решениями одноэлектронного уравнения Шредингера с периодическим потенциалом, имеющим период решетки, представляют собой плоские стоячие волны, модулированные периодической функцией;

б) периодические функции одноэлектронного уравнения Шредингера с волновым потенциалом, перекрывающим поле решетки, представляют собой бегущие волны;

в) потенциальные функции с периодическим потенциалом, имеющим период решетки и не являющиеся решениями уравнения Шредингера, представляют собой волновые функции, модулированные периодическим потенциалом.

5. ***Что описывает амплитудное уравнение Шредингера?***

а) стационарное состояние микрочастиц;

б) движение микрочастиц в периодическом поле;

в) вероятность столкновения двух микрочастиц.

6. ***В чем суть соотношения неопределенности Гейзенберга?***

а) чем точнее определяются координаты микрочастицы, тем неопределеннее становятся составляющие импульса;

б) чем точнее определяются размеры микрочастицы, тем неопределеннее становятся ее координаты;

в) чем точнее определяется число частиц в некотором объеме, тем неопределеннее становятся составляющие импульса.

1.3. ПОТЕНЦИАЛЬНЫЕ БАРЬЕРЫ ПОЛУБЕСКОНЕЧНОЙ ТОЛЩИНЫ

Цель лекции: ознакомление с потенциальным барьером полубесконечной толщины для микрочастиц.

Потенциальные барьеры и ямы для микрочастиц возникают, например, вследствие электрического взаимодействия электронов с ионами решетки в твердом теле, на границах раздела тел. Изменение потенциальной энергии частицы в зависимости от ее координат представляет собой потенциальный рельеф для этой частицы в заданном объеме. В кристаллах наблюдается периодический потенциальный рельеф, который в простейшем случае можно представить в виде совокупности одномерных прямоугольных барьеров, разделенных прямоугольными ямами.

Прямоугольный барьер полубесконечной толщины. Пусть в области A микрочастица имеет потенциальную энергию, равную нулю, а в области B — равную u.

Запишем одномерное амплитудное стационарное уравнение Шредингера:

$$\frac{\partial^2\Psi}{\partial x^2}+\frac{2m}{\hbar^2}(E-u)\Psi=0.$$

Длина волны де Бройля для электрона

$$\lambda=\frac{h}{p}=\frac{h}{mv}\rightarrow \nu=\frac{h}{\lambda m},$$

где m, v — масса и скорость электрона.

Кинетическая энергия электрона

$$E_k=\frac{mv^2}{2}=\frac{mh^2}{2\lambda^2m^2}=\frac{h^2}{2m\lambda^2}.$$

Введем волновое число $k=\frac{2\pi}{\lambda}$, запишем

$$\lambda=\frac{2\pi}{k},$$

откуда

$$E_k = E - u = \frac{h^2 k^2}{2m(2\pi)^2} = \frac{h^2 k^2}{8\pi^2 m},$$

где $\hbar = \dfrac{h}{2\pi} = 1{,}05 \cdot 10^{-34}$ Дж · с.

Тогда

$$E - u = \frac{\hbar^2 k^2}{2m}. \tag{1.6}$$

Выразив из формулы (1.6) k, получим

$$k^2 = \frac{(E-u)2m}{\hbar^2} \to k = \frac{1}{\hbar}\sqrt{2mE_k}.$$

Следовательно, уравнение Шредингера можно записать в виде

$$\frac{\partial^2 \Psi}{\partial x^2} + k^2 \Psi = 0.$$

В области A: $E_k = E$, $u = 0$ (E — полная энергия частицы), $\Psi = \Psi_1$; в области B: $E_k = E - u$, $\Psi = \Psi_2$.

Запишем уравнения Шредингера для областей A и B:

$$A: \frac{\partial^2 \psi_1}{\partial x^2} + k_1^2 \psi_1 = 0; \tag{1.7}$$

$$B: \frac{\partial^2 \psi_2}{\partial x^2} + k_2^2 \psi_2 = 0, \tag{1.8}$$

где

$$k_1 = \frac{1}{\hbar}\sqrt{2mE}; \quad k_2 = \frac{1}{\hbar}\sqrt{2m(E-u)}.$$

Общие решения уравнений (1.7) и (1.8) можно записать в виде

$$\Psi_1 = A_1 \exp(ik_1 x) + B_1 \exp(-ik_1 x);$$
$$\Psi_2 = A_2 \exp(ik_2 x) + B_2 \exp(-ik_2 x),$$

где A_1 — амплитуда падающей волны; B_1 — амплитуда отраженной волны в области A; A_2 — амплитуда волны, распространяющейся в области B в направлении x; B_2 — амплитуда отраженной волны в области B; так как область B полубесконечна, то отраженной волны нет и $B_2 = 0$.

Вероятность нахождения микрочастицы в том или ином месте пространства пропорциональна квадрату амплитуды волны де Бройля, тогда

$$k = \frac{|B_1|^2}{|A_1|^2},$$

отношение квадрата амплитуды отраженной волны к квадрату амплитуды падающей волны называется коэффициентом отражения микрочастицы от барьера. Отношение квадрата амплитуды прошедшей волны к квадрату амплитуды падающей волны называется коэффициентом прозрачности:

$$D = \frac{|A_2|^2}{|A_1|^2}.$$

Можно показать, что

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2; \tag{1.9}$$

$$D = \frac{4k_1k_2}{(k_1 + k_2)^2}. \tag{1.10}$$

Если $E > u$, то R и D определяются из формул (1.9) и (1.10) и $R + D = 1$, т. е. в данном случае в отличие от классической частицы наблюдается отражение. Коэффициент отражения R будет тем больше, чем меньше k_2 и, следовательно, чем меньше разность $E - u$.

Если $E = u$, то из уравнения (1.8) следует, что $k_2 = 0$, и тогда $R = 1$, а $D = 0$, т. е. наблюдается полное отражение волны.

Если $E < u$, то k_2 (см. (1.8)) является мнимой величиной.

Обозначим вещественную величину

$$k = \frac{1}{\hbar}\sqrt{2m(u - E)},$$

тогда $k_2 = ik$,

$$R = \left|\frac{k_1 - ik}{k_1 + ik}\right|^2 = \frac{(k_1 - ik)(k_1 + ik)}{(k_1 + ik)(k_1 - ik)} = 1$$

(см. определение комплексно-сопряженных чисел);

$$D = 1 - R = 0.$$

Однако в этом случае

$$\Psi_2 = A_2 e^{ik_2 x} = A_2 e^{-kx}.$$

Вероятность обнаружения частицы в области B в зависимости от значения координат будет равна:

$$\omega = \Psi_2 \Psi_2^* = |\Psi_2|^2 = |A_2|^2 e^{-2kx} = |A_2|^2 e^{-\frac{2}{\hbar}\sqrt{2m(u-E)}x}. \quad (1.11)$$

Произведение комплексно-сопряженных чисел есть число действительное и равное квадрату модуля каждого из них.

Комплексно-сопряженными числами называются два числа, действительные части которых равны, а мнимые различаются только знаками.

Выражение (1.11), показывающее, что вероятность обнаружения частицы в области B отлична от нуля, не противоречит, однако, $D = 1 - R$, показывающему, что падающая волна полностью отражается от барьера. Это обусловлено тем, что некоторая часть частиц из падающего потока проникает в область B, а затем выходит в область A.

Тесты к лекции 1.3

1. ***В результате чего возникают потенциальные барьеры и ямы для микрочастиц в твердом теле?***

а) в результате электрического взаимодействия электронов с ионами решетки в твердом теле, на границах раздела тел;

б) в результате нагрева твердого тела до температуры плавления;

в) в результате нагрева твердого тела до температуры Дебая.

2. ***Какой формулой определяется волновое число?***

а) $k = \frac{2\pi}{\lambda}$;

б) $k = 2\omega\pi$;

в) $k = 2\pi\delta$.

3. ***Что называется коэффициентом отражения микрочастицы от барьера?***

а) отношение квадрата амплитуды отраженной волны к квадрату амплитуды падающей волны;

б) отношение амплитуды отраженной волны к амплитуде падающей волны;

в) отношение модуля амплитуды отраженной волны к модулю амплитуды падающей волны.

4. ***Что называется коэффициентом прозрачности при прохождении микрочастицы через барьер?***

а) отношение квадрата амплитуды прошедшей волны к квадрату амплитуды падающей волны;

б) отношение амплитуды прошедшей волны к амплитуде падающей волны;

в) отношение модуля амплитуды прошедшей волны к модулю амплитуды падающей волны.

1.4. ПОТЕНЦИАЛЬНЫЕ БАРЬЕРЫ КОНЕЧНОЙ ТОЛЩИНЫ. ФУНКЦИЯ ФЕРМИ

Цель лекции: ознакомление с потенциальными барьерами.

1.4.1. ПРЯМОУГОЛЬНЫЙ ПОТЕНЦИАЛЬНЫЙ БАРЬЕР

Пусть в областях A и C потенциальная энергия частицы равна нулю, а в области B — равна u. Ранее было показано, что существует вероятность нахождения частицы на некотором расстоянии x от прямоугольного барьера при $E < u$. Если $x > a$, то частицы будут из области A переходить в область C (a — ширина прямоугольного барьера).

Такое прохождение частиц через барьер получило название туннельного эффекта. Коэффициент прозрачности барьера в этом случае

$$D = D_0 \exp\left(\frac{-2a}{\hbar}\sqrt{2m(u-E)}\right),$$

где D_0 — коэффициент пропорциональности, приблизительно равный единице.

При туннельном прохождении барьера энергия микрочастиц не изменяется, они покидают барьер с той же энергией, с какой входят в него.

Вероятность прохождения электрона через барьер при $E < u$ быстро убывает с увеличением толщины барьера.

Глубокая прямоугольная потенциальная яма. В областях A и C потенциальная энергия частицы равна u, а в области B — нулю.

Ширину потенциальной ямы обозначим a.

Решение уравнения Шредингера показывает, что микрочастица, запертая в потенциальной яме, может иметь только квантованные значения энергии. Каждому значению энергии будет соответствовать собственная волновая функция:

$$\Psi_n = A_0 \sin\frac{\pi}{a}nx,$$

где A_0 — амплитуда волны; n — квантовое число.

Энергия частицы определяется выражением

$$E = n^2\frac{h^2}{8ma^2}.$$

1.4.2. ФУНКЦИЯ ФЕРМИ

В соответствии с принципом запрета Паули в элементарной фазовой ячейке может находиться не более двух электронов, имеющих антипараллельные спины. Ячейка может быть полностью занята, занята частично, полностью свободна.

Из статистической физики известно, что вероятность заполнения электронами фазовой ячейки, имеющей энергию E, определяется функцией Ферми:

$$f_F = \frac{1}{e^{\frac{E-E_F}{kT}}+1},$$

где E_F — энергия Ферми; k — постоянная Больцмана.

При абсолютном нуле функция Ферми f_F принимает лишь два значения: при $E < E_F = E_{F0}$ — $f_F = 1$, а при $E > E_F$ — $f_F = 0$. Если $T \neq 0$, то наибольшее изменение f_F наблюдается вблизи E_F, а при $E = E_F$ — $f_F = \frac{1}{2}$.

Интервал изменения f_F от 1 до 0 простирается по оси абсцисс в пределах нескольких kT влево и вправо от E_F. Приближенно можно считать, что при $E < E_F - kT$ — $f_F = 1$, а при $E > E_F + kT$ — $f_F = 0$.

В условиях термодинамического равновесия уровень Ферми оказывается постоянным в любой системе контактирующих тел. Необходимым условием термодинамического равновесия системы является постоянство уровня Ферми.

Электронный газ, свойства которого описываются функцией Ферми, называют вырожденным, или квантованным. При этом число состояний одного порядка оказывается равным числу частиц. Электронный газ называется невырожденным, если число разрешенных энергетических состояний значительно больше числа электронов, при этом используется классическая статистика Максвелла — Больцмана, поскольку квантово-механические запреты затушевываются.

Электронный газ можно считать невырожденным, если выполняется условие

$$f_{\text{М–Б}} = e^{\frac{E-E_F}{kT}} >> 1.$$

Это условие выполняется при $e^{(E-E_F)/kT} >> 2 \div 3$ на «хвосте» кривой выше E_F. Этот «хвост» соответствует распределению Максвелла — Больцмана.

Наоборот, при выполнении условия

$$e^{\frac{E-E_F}{kT}} << 1$$

газ является вырожденным.

1.4.3. СТРОЕНИЕ АТОМА

Состояния электронов в атомах характеризуются четырьмя квантовыми числами.

1. *Главное квантовое число* $n = 1, 2, 3, 4, \ldots$. Оно определяет порядковый номер разрешенной орбиты или энергетического уровня электрона.

2. *Орбитальное квантовое число* l связано с орбитальным механическим моментом количества движения p_l-электрона соотношением

$$p_l = \hbar\sqrt{l(l+1)}.$$

Моментом количества движения точки относительно полюса называется вектор, равный векторному произведению радиуса-вектора точки, проведенного из полюса, и ее количество движения.

3. *Моментное квантовое число* $m = 0, \pm1, \pm2, \ldots$ определяет проекцию момента p_i на некоторое направление, m принимает лишь целочисленные значения.

4. *Спиновое квантовое число* $S = \pm\frac{1}{2}$ выражает собственный момент количества движения электрона. Он может быть направлен либо параллельно, либо антипараллельно орбитальному моменту количества движения.

Совокупность электронов, обладающих одним и тем же значением главного квантового числа n, образует слой атома: $n = 1$, к-слой, $n = 2$, L-слой; $n = 3$, M-слой и т. д.

Внутри слоя электроны с различными орбитальными числами l образуют оболочки $s(l = 0)$, $p(l = 1)$, $d(l = 2)$, $f(l = 3)$ и т. д.

Согласно принципу Паули, в одном квантовом состоянии, характеризуемом четырьмя квантовыми числами, может находиться не более одного электрона.

При заполнении слоев и оболочек электроны стремятся занять более низкие энергетические уровни.

Тесты к лекции 1.4

1. ***Как изменяется энергия микрочастицы при туннельном прохождении потенциального барьера?***

а) не изменяется;

б) увеличивается;

в) уменьшается.

2. ***Какова вероятность прохождения микрочастицы сквозь прямоугольный потенциальный барьер, если ее энергия меньше высоты барьера?***

а) больше нуля;

б) нуль;

в) 50 %.

3. ***Какие значения энергии может иметь микрочастица, запертая в потенциальной яме?***

а) может иметь только квантованные значения энергии;

б) может иметь любые значения энергии;

в) не может иметь никаких значений энергии.

4. ***Сформулируйте принцип запрета Паули:***

а) в элементарной фазовой ячейке может находиться не более двух электронов, имеющих антипараллельные спины;

б) в элементарной фазовой ячейке может находиться не более четырех электронов, имеющих параллельные спины;

в) в элементарной фазовой ячейке не может находиться более восьми электронов, имеющих антипараллельные спины.

5. ***Что определяет функция Ферми?***

а) вероятность заполнения электронами фазовой ячейки, имеющей определенную энергию;

б) вероятность преодоления электронами потенциального барьера;

в) вероятность нахождения электрона в данной точке пространства.

6. ***Чему численно равно значение функции Ферми при температуре абсолютного нуля, если энергия частиц меньше энергии Ферми?***

а) единице;

б) нулю;

в) 0,5.

1.5. ВОЛНОВЫЕ ПРОЦЕССЫ В ТВЕРДОМ ТЕЛЕ

Цель лекции: ознакомление с волновыми процессами в твердом теле и низкоразмерными объектами.

1.5.1. ФАЗОВАЯ И ГРУППОВАЯ СКОРОСТИ, ФОНОНЫ

Атомы, находящиеся в узлах кристаллической решетки, при температуре, отличной от абсолютного нуля, совершают колебания вблизи положения равновесия. Носители заряда, движущиеся среди колеблющихся атомов, обмениваются с ними энергией, благодаря чему устанавливается термодинамическое равновесие между решеткой и электронным газом.

Вследствие сильного межатомного взаимодействия колебание каждого атома передается соседним, и в кристалле возбуждаются «коллективные» колебания атомов, распространяющиеся во всевозможных направлениях. Если на атом действует сила, пропорциональная его смещению из положения равновесия, то возникают гармонические колебания. Следовательно, колеблющийся атом можно считать гармоническим осциллятором. Колебания каждого гармонического осциллятора можно разложить на составляющие по трем направлениям. Атомы кристалла упруго связаны между собой, и колебания любого из них передаются соседям, такие колебания называются нормальными. В кристалле, содержащем N атомов, распространяются упругие волны от $3N$-осцилляторов, т. е. возможно $3N$ нормальных колебаний.

Волновой процесс характеризуется также фазовой и групповой скоростями распространения волны. Фазовая скорость представляет собой скорость перемещения в пространстве точки с заданной фазой:

$$v_{\text{ф}} = \frac{\omega}{k} = \frac{2}{k}\sqrt{\frac{\beta}{M}}\left|\sin\frac{ka}{2}\right|,$$

где β — жесткость связи; M — масса атома.

Скорость распространения волнового пакета (скорость переноса энергии в среде) определяется групповой скоростью:

$$v_{\text{гр}} = \frac{\partial\omega}{\partial k} = a\sqrt{\frac{\beta}{M}}\left|\cos\frac{ka}{2}\right|.$$

Если длина цепочки атомов L, то наиболее длинноволновые колебания, которые могут распространяться в такой цепочке, имеют длину волны

$\lambda_{max} = 2L$, а наиболее коротковолновые — $\lambda_{min} = 2a$. В последнем случае может быть найдена максимально возможная частота колебаний

$$\omega_{max} = \frac{2\pi v}{\lambda_{min}} = \frac{\pi v}{a} = kv,$$

где v — средняя скорость распространения нормальных колебаний.

Отсюда следует, что максимальная частота колебаний является постоянной для каждого конкретного материала. Например, для меди ($a \approx 3{,}6 \cdot 10^{-10}$ м, $v \approx 3{,}6 \cdot 10^3$ м/с) $\omega_{max} \approx 3 \cdot 10^{13}$, для оксида цинка ($a \approx 4{,}7 \cdot 10^{-10}$ м, $v \approx 2{,}7 \cdot 10^3$ м/с) $\omega_{max} \approx 1{,}8 \cdot 10^{13}$.

Из соотношений для фазовой $v_ф$ и групповой $v_{гр}$ скоростей следует, что для длинных волн $(k \to 0)$ $v_ф = v_{гр} = v_{зв}$, где $v_{зв}$ — средняя скорость распространения звука в кристалле. Для коротких волн $(k \to \pi/a)$ $v_{гр} \to 0$, $v_ф \to (2a/\pi)\sqrt{\beta/M}$.

Рассмотрим тепловые колебания решетки с квантовой точки зрения. Энергия каждого из $3N$ колебаний квантована. Разрешенные значения энергии

$$E_\omega = (n + 1/2)\hbar\omega,$$

где $h = 0, 1, 2, \ldots$. Отсюда минимально возможное изменение тепловой энергии равно $\hbar\omega$. Эта минимальная порция, или квант энергии колебания, называется фононом. Если представить решетку в виде фононного газа, то увеличение энергии $E_\omega = E_{\omega 0} + \Delta E$ означает увеличение концентрации фононов, уменьшение же энергии $E_\omega = E_{\omega 0} - \Delta E$ — уменьшение концентрации фононов. При этом акустическим колебаниям решетки соответствуют акустические фононы, оптическим — оптические фононы.

1.5.2. НИЗКОРАЗМЕРНЫЕ ОБЪЕКТЫ: КВАНТОВАЯ ЯМА, КВАНТОВАЯ НИТЬ, КВАНТОВАЯ ТОЧКА

При переходе вещества в наносостояние его свойства начинают значительно зависеть от размера составляющих его нанообъектов. *Размерный эффект* — зависимость свойств тела от его размера.

Размерные эффекты возникают, когда размер объекта соизмерим с каким-либо параметром вещества, оказывающим существенное влияние на физические процессы в веществе и его свойства. Такими параметрами могут быть длина свободного пробега носителей в веществе, диффузионная длина, диаметр траекторий скольжения дислокаций и т. п. Для квантовых размерных эффектов данным параметром является длина волны де Бройля.

Длина волны де Бройля для электрона, движущегося в кристаллической структуре,

$$\lambda = \frac{h}{m^* v} = \frac{h}{\sqrt{2m^* E}},$$

где h — постоянная Планка; m^* — эффективная масса электрона; E — кинетическая энергия; v — скорость электрона.

Поскольку длина волны де Бройля для свободных электронов в полупроводниках значительно больше, чем в металлах, то квантово-размерные эффекты технологически легче осуществить на полупроводниках. Поэтому исследование этих эффектов и формирование наноструктур для применения в электронике проводится преимущественно на полупроводниках.

Следует отметить, что квантово-размерные эффекты можно наблюдать при условии, что средняя длина свободного пробега электронов превышает размер рассматриваемой области, ее границы имеют высокую степень совершенства, а отражения волны де Бройля от границ можно считать зеркальными.

В 3D-объектах электроны могут свободно перемещаться во всех трех направлениях:

$$\lambda \ll L_x, L_y, L_z.$$

Примером такого объекта является неограниченный кристалл полупроводника.

Двумерный 2D-объект. Это тонкий слой кристалла, толщина которого d соизмерима с длиной волны де Бройля $(d \sim \lambda)$:

$$\lambda > L_z \text{ и одновременно } \lambda \ll L_x, L_y.$$

В этом случае электроны могут свободно перемещаться только в плоскости слоя, а в направлении, перпендикулярном плоскости слоя, их движение будет ограничено потенциальным барьером. Двигаясь в этом направлении, электрон не способен покинуть слой, так как его работа выхода (1...6 эВ) намного больше энергии теплового движения (~0,026 эВ при комнатной температуре). Электроны оказываются в глубокой потенциальной яме, и их энергия в этом направлении будет квантоваться в сочетании с непрерывными спектрами электронов в направлениях, находящихся в плоскости слоя. 2D-объект называют квантовой ямой (*quantumwell*).

Примерами 2D-объектов могут служить проводящие каналы в транзисторах из металла–диэлектрика–полупроводника (МДП) и узкозонные слои в гетероструктурах из соединений A_3B_5 для инжекционных лазеров. Системы близко расположенных параллельных квантовых ям, между которыми возможно туннелирование электронов, составляют сверхрешетки.

Если накладывается квантовое ограничение на движение электронов еще в одном направлении, то получим 1D-объект, который называют квантовой проволокой (*quantumwire*):

$$\lambda > L_y,\ L_z \text{ и одновременно } \lambda \ll L_x.$$

Потенциальная яма для свободных электронов в нити двумерна.

Если размеры объекта в трех измерениях будут близки к длине волны

де Бройля, то получим 0D-объект, называемый квантовой точкой:

$$\lambda \gg L_x, L_y, L_z.$$

Потенциальная яма для квантовой точки трехмерна. Электроны в квантовых точках могут иметь только дискретный набор энергетических состояний. Примером квантовых точек являются самоорганизующиеся структуры, возникающие при осаждении атомов или молекул полупроводникового материала на поверхность другого материала с более широкой запрещенной зоной.

Тесты к лекции 1.5

1. ***Что такое фазовая скорость?***

а) скорость перемещения в пространстве точки с заданной фазой;

б) скорость изменения фаз колебательного процесса;

в) скорость перераспределения фаз.

2. ***Что такое групповая скорость?***

а) скорость распространения волнового пакета (скорость переноса энергии в среде);

б) скорость распространения частиц в противофазе;

в) скорость затухания энергии в среде.

3. ***Максимальная частота акустических колебаний для каждого конкретного материала является постоянной величиной, линейной функцией или сложной функцией?***

а) постоянной;

б) линейной функцией;

в) сложной функцией.

4. ***Что называется фононом?***

а) минимальная порция, или квант энергии колебания;

б) квант световой энергии;

в) квант радиоактивного излучения.

5. ***Что такое размерный эффект?***

а) зависимость свойств тела от его размера;

б) зависимость размера тела от его свойств;

в) возможность получения ошибочного результата при неправильном использовании размерности.

1.6. ЭЛЕМЕНТЫ ЗОННОЙ ТЕОРИИ ТВЕРДЫХ ТЕЛ

Цель лекции: ознакомление с понятием электронного газа в периодическом потенциальном поле.

В твердом теле расстояния между атомами настолько малы, что каждый из них оказывается в достаточно сильном поле соседних атомов. На больших расстояниях между атомами взаимодействием между ними можно пренебречь. Пусть атомы находятся на расстоянии $r \gg a$, где a — постоянная решетки (рис. 1.2). Высота барьеров для электронов равна расстоянию от уровня, на котором они находятся, до уровня 0–0. Этот потенциальный барьер препятствует свободному перемещению электронов от атома к атому.

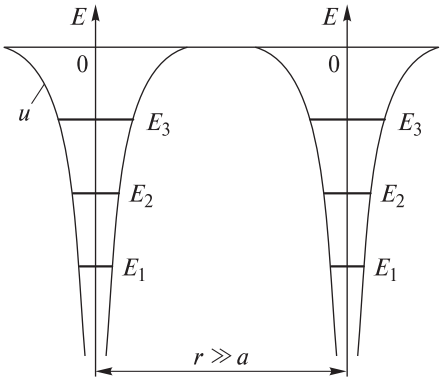


Рис. 1.2. Потенциальные ямы атомов, удаленных на значительное расстояние друг от друга

При сближении атомов взаимодействие между ними растет. Потенциальные кривые, отделяющие соседние атомы (на рис. 1.2 показаны частично), накладываются друг на друга и дают результирующие кривые, проходящие ниже линии 0–0 (рис. 1.3). Таким образом, сближение атомов оказывает двоякое действие на потенциальный барьер, оно уменьшает его толщину до значения порядка кристаллической решетки и понижает высоту. Высота потенциального барьера для верхних уровней оказывается ниже их положения в свободном атоме (E_3), и валентные электроны, находящиеся на этом уров-

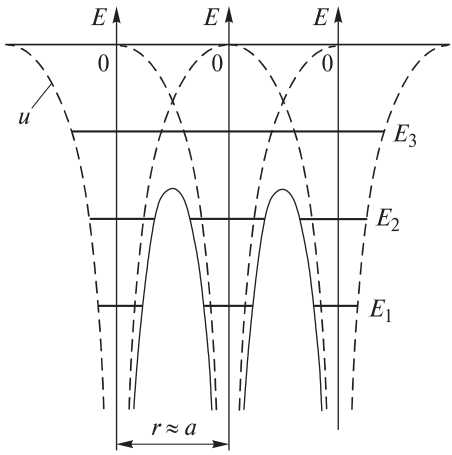


Рис. 1.3. Потенциальные ямы атомов, находящихся на расстоянии, равном постоянной решетки Q

не, получают возможность перемещаться по кристаллу от одного атома к другому. Такие электроны называются свободными, а их совокупность — электронным газом.

При расчете состояний электронного газа используют два крайних случая: приближение сильной и слабой связи.

Приближение сильной связи. Пусть электроны находятся в потенциальных ямах своих атомов. Уменьшение высоты и толщины барьера вследствие сближения атомов может привести к тому, что барьер окажется прозрачным для туннелирования электронов. Туннелируют преимущественно электроны на внешних уровнях, так как частота перехода от одного атома к другому

$$\gamma = \frac{v}{a} D = \frac{v}{a} e^{-\frac{2b}{\hbar}\sqrt{2m(u-E)}},$$

где v — скорость перемещения электрона в потенциальной яме; v/a — число подходов к барьеру в единицу времени; b — толщина барьера.

Тогда время пребывания электрона в атоме

$$\tau = \frac{1}{\gamma}.$$

Расчеты показывают, что при b, равном единицам ангстрем, τ очень мало: ~10–15 с. Тогда из соотношения неопределенностей $dEdt \geq h$ следует, что при уменьшении времени пребывания электрона в заданном состоянии увеличивается интервал энергии этого состояния, чему соответствует образование для электронов разрешенных зон (рис. 1.4). Наименьшее τ будет при этом у внешних электронов, и они становятся обобществленными. Более глубинные электроны имеют большее значение τ и меньшее dE.

Таким образом, каждому энергетическому уровню изолированного атома в кристалле соответствует зона разрешенных энергий. Зоны разрешенных энергий разделены запрещенными зонами. С увеличением энергии электрона в атоме ширина разрешенных зон увеличивается, а ширина запрещенных — уменьшается. В общем случае зоны, образованные отдельными уровнями, могут перекрываться, образуя гибридную зону.

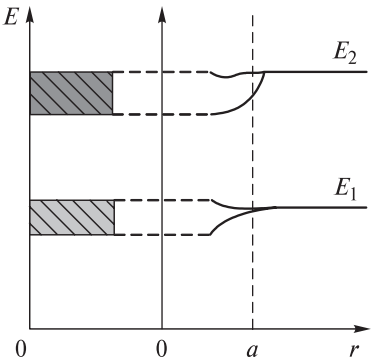


Рис. 1.4. Образование разрешенных энергетических зон при сближении атомов

Приближение слабой связи. Энергетический спектр электронов в кристалле, как было показано выше, имеет зонный характер. Определим, как энергия электронов зависит от импульса p. Рассмотрим свободные электроны, движущиеся в периодическом поле кристаллической решетки. В этом случае уравнение Шредингера решают при циклических граничных условиях:

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2m}{\hbar^2}(E-u)\Psi = 0.$$

Решение этого уравнения по Блоху можно записать в виде

$$\Psi(x) = u_1(x)e^{ikx},$$

где $u_1(x)$ — периодическая функция, период которой совпадает с периодом потенциала $u(x)$, т. е. равен d ($d = a + b$, где a — ширина потенциальной ямы; b — ширина потенциального барьера, согласно модели Кронига — Пенни (рис. 1.5)).

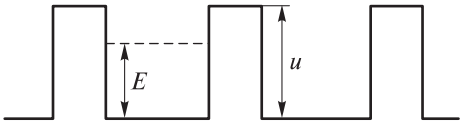


Рис. 1.5. Модель Кронига – Пенни

Можно показать, что с учетом граничных условий, подставив решение по Блоху в уравнение Шредингера, получим уравнение

$$\frac{mab}{\hbar^2} u \frac{\sin(\alpha a)}{\alpha a} + \cos(\alpha a) = \cos(ka), \tag{1.12}$$

где

$$\alpha = \left(\frac{2mE}{\hbar^2}\right)^{\frac{1}{2}}. \tag{1.13}$$

Если барьеры весьма высокие и тонкие, то можно принять $a \approx d$. Обозначим левую часть уравнения (1.12) через $y(\alpha a)$:

$$y(\alpha a) = \frac{mab}{\hbar^2} u \frac{\sin(\alpha a)}{\alpha a} + \cos(\alpha a). \qquad (1.14)$$

Уравнение (1.14) решаем графическим способом (по оси абсцисс откладываем αa, по оси ординат — $y(\alpha a)$). Качественно получаем зависимость, приведенную на рис. 1.6, где штриховые линии — часть кривой $y(\alpha a)$, выходящая за значения ±1. Эта часть кривой не может удовлетворять уравнению (1.12), поскольку $\cos(ka)$ не может быть по абсолютному значению больше единицы. Сплошной линией обозначена часть кривой $y(\alpha a)$, лежащая в пре-

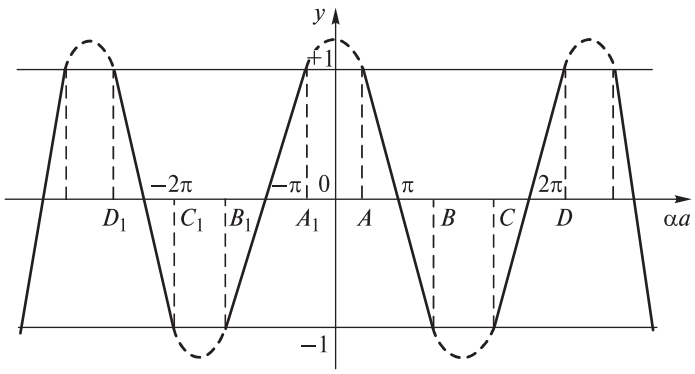


Рис. 1.6. Зависимость $y(\alpha a)$

делах значений ±1 и удовлетворяющая уравнению (1.14). Участки на оси абсцисс AB, CD, A_1B_1 и C_1D_1, на которых кривая не выходит за пределы значений ±1, соответствуют разрешенным зонам энергии, поскольку α и E связаны соотношением (1.13). При $\alpha a \to 0$ имеем $\frac{\sin(\alpha a)}{\alpha a} \to 1$, $y(\alpha a) > 1$, точка A определяется таким углом αa, при котором $y(\alpha a) = 1$. В точке B $y(\alpha a)$ достигает значения -1, следовательно, $\alpha a = \pi$. От точки B до точки C наблюдается запрещенная зона, затем разрешенная зона CD. В точке D $\alpha a = 2\pi$. Кривая $y(\alpha a)$ симметрична относительно нуля, следовательно, при отрицательном значении αa имеем аналогичные отрицательные углы. Ширина разрешенных зон увеличивается по мере роста αa. Из (1.12) следует, что в точке B $\cos(ka) = -1$, следовательно, $ka = \pi$. В точке D имеем $ka = 2\pi$ и т. д. Потолки разрешенных зон наблюдаются при условии

$$k = \frac{n\pi}{a}, \qquad (1.15)$$

где $n = 1, 2, 3, \ldots$.

На рис. 1.7 приведена кривая $E(k)$, соответствующая рассматриваемому случаю. Вблизи дна первой разрешенной зоны кривая $E(k)$ представляет со-

бой параболу. Однако к потолку зоны кривая $E(k)$ отклоняется от параболы. При $k = \pm\frac{\pi}{a}$ имеем потолок первой зоны и разрыв кривой, соответствующий запрещенной зоне. Далее имеются отрезки кривых, обусловленные более высокими разрешенными энергетическими зонами.

Таким образом, волновые законы движения электронов в периодическом потенциальном поле приводят к возникновению разрешенных и запрещенных зон энергии.

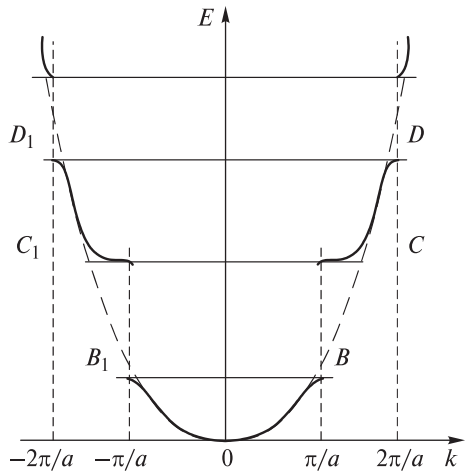


Рис. 1.7. Зависимость энергии электрона от волнового вектора k

Тесты к лекции 1.6

1. ***Какие существуют виды приближений?***

а) приближение сильной связи и приближение слабой связи;

б) приближение электронов к ядру атома при потере энергии;

в) приближение симметрии структуры кристалла.

2. ***К чему может привести уменьшение высоты и толщины энергетического барьера в кристаллической решетке вследствие сближения атомов?***

а) барьер окажется прозрачным для туннелирования электронов;

б) к объединению атомов;

в) к отталкиванию атомов.

3. ***Чем разделены зоны разрешенных энергий в твердом теле?***

а) запрещенными зонами;

б) магнитным полем;

в) электромагнитным полем.

4. ***Что такое гибридная зона?***

а) зона, образованная отдельными перекрывающимися энергетическими уровнями;

б) зона, образованная гибридными атомами;

в) зона, образованная гибридными молекулами.

1.7. ЗОНЫ БРИЛЛЮЭНА И ПОНЯТИЕ ЭФФЕКТИВНОЙ МАССЫ

Цель лекции: ознакомление с зонами Бриллюэна и понятием эффективной массы.

1.7.1. ЗОНЫ БРИЛЛЮЭНА

Ранее использовалось понятие пространства импульсов, которое определялось путем задания составляющих импульсов в декартовой системе координат. Модуль волнового вектора, или волновое число, $k = 2\pi/\lambda$, а $\lambda = h/p$ — длина волны де Бройля. Тогда $p = \hbar k$; если нет $\hbar$, то $p = hk/2\pi$, где $\hbar = h/2\pi$.

Следовательно, импульс пропорционален волновому вектору k: $p = \hbar k$; тогда вместо пространства импульсов можно рассматривать k-пространство, задаваемое составляющими k_x, k_y, k_z. Разрешенным энергетическим зонам в твердом теле соответствуют зоны в k-пространстве. Области значений волнового вектора k, в пределах которых энергия электрона $E(k)$, являющаяся периодической функцией k, испытывает полный цикл своего изменения, называют зонами Бриллюэна. На границах зон энергия претерпевает разрыв. Для одномерного кристалла первая зона Бриллюэна простирается от $k = -\pi/a$ до $k = \pi/a$ и имеет протяженность $2\pi/a$ (см. рис. 1.7).

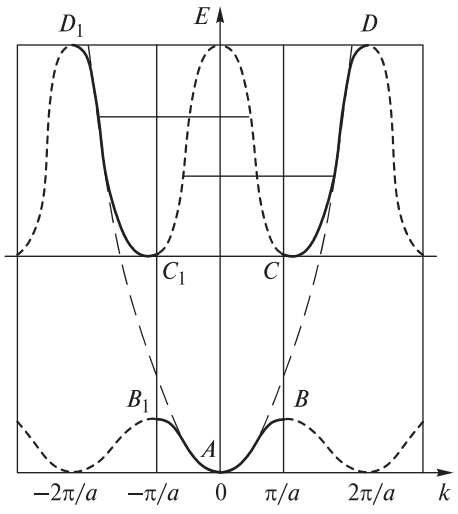


Рис. 1.8. Зависимость энергии электрона от волнового вектора k для разных энергетических зон

Как отмечалось выше, из решения уравнения Шредингера для электрона, находящегося в периодическом потенциальном поле кристаллической решетки, следует, что собственные значения (разрешенные) энергии электрона должны быть периодическими функциями k:

$$E(k_x) = E(k_x \pm n2\pi/a_x). \qquad (1.16)$$

Кривая $E(k_x)$ называется дисперсионной кривой. Как видно на рис. 1.8, для каждой из разрешенных зон справедливо соотношение (1.16), хотя кривые $E(k_x)$ для разрешенных зон отличаются друг от друга. С ростом E ширина разрешенных зон увеличивается.

Если отрезок CD сдвинуть на $2\pi/a_x$ влево, а отрезок C_1D_1 — на $2\pi/a_x$ вправо, то вторую зону Бриллюэна можно привести к первой. Первую зону Бриллюэна, куда перенесены функции $E(k_x)$ для разных энергетических зон, называют приведенной зоной Бриллюэна (рис. 1.9). В дальнейшем будут преимущественно рассматриваться лишь две верхние разрешенные энергетические зоны. Верхняя разрешенная зона называется зоной проводимости, нижняя — валентной зоной.

В реальных кристаллах направления составляющих волнового вектора k выбирают в соответствии с определенными кристаллографическими направлениями. Зависимости $E(k)$ у реальных кристаллов являются достаточно сложными.

В качестве примера рассмотрим зонную структуру кремния для двух направлений в k-пространстве. Минимум зависимости $E(k)$ или дисперсионной кривой называют дном энергетической зоны, максимум — потолком зоны. Как видно на рис. 1.10, дно зоны проводимости у кремния находится не в середине зоны Бриллюэна, а вблизи ее границы в направлении. Вершина валентной зоны расположена в середине зоны Бриллюэна.

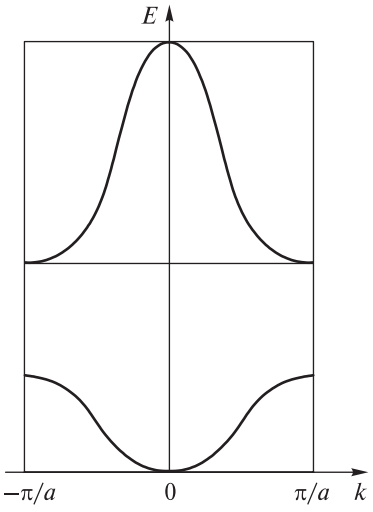


Рис. 1.9. Приведенная зона Бриллюэна

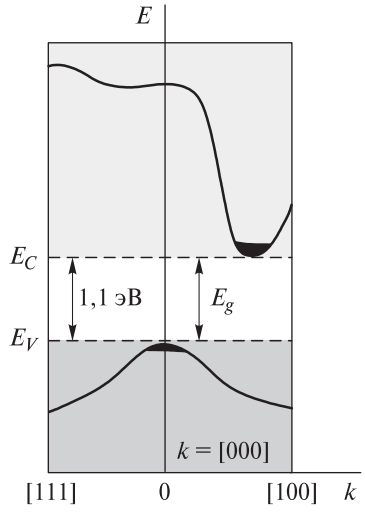


Рис. 1.10. Зонная структура кремния

Минимальный зазор между валентной зоной и зоной проводимости принимается за ширину запрещенной зоны E_g.

При упрощенном рассмотрении энергетической структуры полупроводников вместо истинных дисперсионных кривых $E(k)$, ограничивающих валентную зону и зону проводимости, проводят две параллельные прямые: одну — касательную к дну зоны проводимости E_C, вторую — касательную к вершине валентной зоны E_V.

Первую прямую принимают за нижнюю границу (дно) зоны проводимости, вторую — за верхнюю границу (потолок) валентной зоны.

1.7.2. ЭФФЕКТИВНАЯ МАССА ЭЛЕКТРОНА

Известно: $p = \hbar k = mv$,

$$v = \frac{\hbar}{m} k. \tag{1.17}$$

Кинетическая энергия свободного электрона $E_k = E$ и

$$E_k = \frac{mv^2}{2},$$

подставив в формулу (1.17), получим

$$E_k = \frac{\hbar^2 k^2}{2m}.$$

Продифференцируем E_k по k:

$$\frac{dE_k}{dk} = \frac{\pi^2}{m} k,$$

откуда

$$k = \frac{m}{\hbar^2} \frac{dE}{dk}. \tag{1.18}$$

После подстановки соотношения (1.18) в формулу (1.17) имеем

$$v = \frac{\hbar}{m} k = \frac{1}{\hbar} \frac{dE}{dk}. \tag{1.19}$$

Формула (1.19) справедлива не только для свободного электрона, но и для электрона, находящегося в потенциальном поле.

Пусть энергия зонного электрона изменяется под некоторым внешним воздействием:

$$dE = Fv \cdot dt, \tag{1.20}$$

где F — внешняя сила.

Подставим формулу (1.19) в уравнение (1.20):

$$dE = \frac{F}{\hbar} \frac{dE}{dk} dt,$$

откуда

$$\frac{dk}{dt} = \frac{F}{\hbar}. \tag{1.21}$$

Продифференцируем формулу (1.19) по времени:

$$\frac{dv}{dt} = \frac{1}{\hbar}\frac{d}{dt}\left(\frac{dE}{dk}\right) = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt}. \tag{1.22}$$

Подставив равенство (1.21) в соотношение (1.22), получим

$$\frac{dv}{dt} = \frac{F}{\hbar^2}\frac{d^2E}{dk^2} = a, \tag{1.23}$$

где a — ускорение.

Формула (1.23) связывает ускорение и силу, т. е. выражает второй закон Ньютона $F = ma;\ \ a = F/m$.

Из формулы (1.23) следует, что под действием внешней силы электрон в периодическом поле кристалла движется так, как двигался бы свободный электрон, обладающий массой

$$m^* = \frac{1}{\frac{1}{\hbar^2}\frac{d^2E}{dk^2}}. \tag{1.24}$$

Масса m^* называется эффективной массой электрона. Приписывая электрону, находящемуся в периодическом поле кристалла, массу m^*, можно считать этот электрон свободным и описывать его движение во внешнем поле так, как описывается движение свободного электрона. Эффективная масса, отражающая особенности движения электрона в периодическом поле, является своеобразной функцией. Она может быть как положительной, так и отрицательной, а по абсолютному значению как меньше, так и больше массы покоя электрона. Эффективная масса свободного электрона равна его массе покоя.

При движении электрона в периодическом потенциальном поле кристалла работа внешней силы может переходить как в кинетическую, так и потенциальную энергию электрона:

$$A = \Delta E_k + \Delta u.$$

Если часть работы внешней силы $F = -q\varepsilon$ переходит в потенциальную энергию, то скорость электрона возрастает медленнее, чем у свободного электрона и, следовательно, его эффективная масса больше его массы покоя. Если вся работа внешней силы переходит в потенциальную энергию, то скорость электрона изменяться не будет, и он будет вести себя, как частица с бесконечно большой массой.

В потенциальную энергию может переходить не только работа внешней силы, но и кинетическая энергия электрона. Скорость электрона будет в этом случае уменьшаться, т. е. он ведет себя как частица с отрицательной массой.

Возможен случай, когда в кинетическую энергию может переходить не только работа внешней силы, но и потенциальная энергия, тогда скорость

электрона будет расти быстрее, чем у свободного электрона, т. е. его эффективная масса будет меньше его массы покоя.

На рис. 1.11 показаны зависимости E, v, m^* от волнового вектора k. Вблизи дна разрешенной зоны энергия электрона положительна, а у потолка зоны — отрицательна.

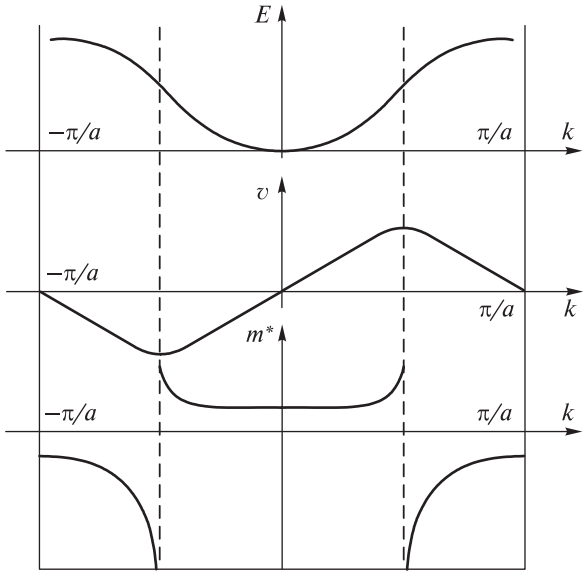


Рис. 1.11. Зависимости энергии, скорости и эффективной массы электрона от волнового вектора k

Точка A — точка перегиба зависимости $E(k)$, в этой точке dE/dk достигает максимума, вторая производная $d^2E/dk^2 = 0$, а $m^* \to \infty$. Поскольку $\vec{k}$ — вектор, то m^* зависит от направления движения электрона в кристалле.

Тесты к лекции 1.7

1. ***В каком диапазоне простирается первая зона Бриллюэна для одномерного кристалла?***

а) от $k = -\pi/a$ до $k = \pi/a$;
б) от $k = -\pi$ до $k = \pi$;
в) от $k = -2\pi$ до $k = 2\pi$.

2. ***Как называется верхняя разрешенная зона?***

а) зона проводимости;
б) валентная зона;
в) запрещенная зона.

3. ***Как называется нижняя разрешенная зона?***

а) валентная зона;

б) запрещенная зона;

в) зона проводимости.

4. ***Что такое запрещенная зона?***

а) уровень энергии, которым электроны не могут обладать;

б) пространство, в котором не могут быть электроны;

в) пространство, к котором не могут находиться электроны и дырки.

5. *Какое значение может принимать эффективная масса электрона?*

а) как больше, так и меньше нуля;

б) только больше единицы;

в) только меньше единицы.

6. ***Как изменится скорость электрона, если вся работа внешней силы переходит в потенциальную энергию?***

а) не будет изменяться;

б) будет увеличиваться;

в) будет уменьшаться.

1.8. ЗОННАЯ СХЕМА КРИСТАЛЛИЧЕСКИХ ТЕЛ И ПЛОТНОСТЬ СОСТОЯНИЙ

Цель лекции: ознакомление с зонной схемой кристаллических тел и плотностью состояний.

1.8.1. ЗОННАЯ СХЕМА КРИСТАЛЛИЧЕСКИХ ТЕЛ — ПРОВОДНИКИ, ДИЭЛЕКТРИКИ, ПОЛУПРОВОДНИКИ

Как отмечалось выше, зоны изображают горизонтальными прямыми, которые не являются осью отсчета какого-либо параметра. Часто по горизонтальной оси откладывают расстояние в трехмерном пространстве. При этом зоны могут наклоняться в электрическом поле и искривляться на границах раздела. В данном разделе с горизонтальным направлением не связывается какая-либо величина. В вертикальном направлении вверх откладывают энергию электронов, вниз — энергию дырок. Энергия электронов увеличивается при переходе на более высокий уровень, а энергия дырок, наоборот, уменьшается. Каждая энергетическая зона содержит ограниченное число энергетических уровней. При ограниченном числе электронов в атомах твердых тел заполненными окажутся лишь несколько наиболее нижних энергетических зон.

По характеру заполнения зон тела подразделяют на две группы.

Группа I — тела, у которых над полностью заполненной зоной располагается зона, заполненная лишь частично (рис. 1.12, *а*). Такая зона образуется из атомного уровня, заполненного в атоме не полностью, например, у щелочных металлов. Частично заполненной может оказаться зона и при наложении заполненных зон на пустые. Наличие частично заполненной зоны присуще металлам.

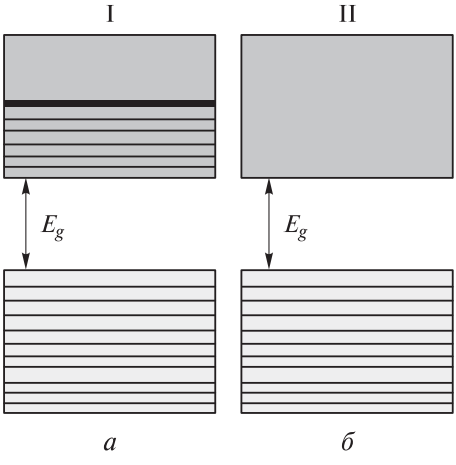


Рис. 1.12. Заполнение энергетических зон электронами

Группа II — тела, у которых над полностью заполненными зонами располагаются пустые зоны (рис. 1.12, *б*). Типичным примером таких тел являются элементы IV группы Периодической системы элементов Д.И. Менделеева: углерод C, кремний Si, германий Ge, а также оксиды, нитриды, карбиды и т. д.

Электроны внешних энергетических зон имеют практически одинаковую свободу движения независимо от того, являются ли тела металлами или диэлектриками. Движение осуществляется путем туннелирования. Однако электропроводность металлов и диэлектриков значительно отличается: $\rho_{VM} \approx 10^{-7}$ Ом · м; $\rho_{VD} \approx 10^{11}$ Ом · м. Следовательно, для появления у тел проводимости наличие свободных электронов на внешних уровнях не является достаточным. Кроме того, необходимо, чтобы в энергетической зоне, к которой принадлежат данные электроны, имелись незанятые состояния. Это связано с тем, что внешняя сила $F = -q\varepsilon$, действующая на электроны, ускоряет или замедляет их движение, а следовательно, изменяет их энергию, что обусловлено переходом в новые квантовые состояния. Это возможно только в неполностью укомплектованных зонах, когда даже слабое электрическое поле способно сообщить электронам импульс, достаточный для того чтобы перевести их на один из близлежащих свободных уровней. В теле появляется движение электронов, обусловливающее появление электрического тока. Такие тела являются хорошими проводниками.

Если валентная зона полностью занята, то ближайшая свободная зона отделена от нее запрещенной зоной E_g. Поэтому внешнее поле не способно изменить характер движения электронов, для этого необходимо увеличить их энергию на E_g, что возможно только в очень сильных полях. Отсутствие у тел в их энергетическом спектре энергетических зон, частично укомплектованных электронами, делает их непроводниками.

По ширине запрещенной зоны тела, относящиеся ко второй группе, подразделяют на диэлектрики и полупроводники. У диэлектриков запрещенная зона, как правило, $E_g > 3$ эВ (Al_2O_3 – $E_g = 7$ эВ, SiO_2 – $E_g = 9$ эВ). У типичных полупроводников $E_g \le 3$ эВ (Ge – $E_g = 0{,}7$ эВ, Si – $E_g = 1{,}12$ эВ). Предположим, что валентная зона полностью заполнена электронами. Как было показано выше, электроны, располагающиеся на верхних уровнях зоны, имеют отрицательную эффективную массу и отрицательный заряд. Если с одного из таких уровней удалить электрон, то освободившийся уровень будет вести себя как частица, имеющая положительный заряд и обладающая эффективной массой, численно равной эффективной массе удаленного электрона. Такую фиктивную частицу, располагающуюся вблизи потолка зоны и имеющую положительную эффективную массу и положительный заряд, называют дыркой. Электроны и дырки называются подвижными носителями заряда, в отличие от неподвижных носителей заряда — ионов, располагающихся в узлах решетки.

1.8.2. ПЛОТНОСТЬ СОСТОЯНИЙ

Энергетический спектр — совокупность возможных значений энергии частицы в данных условиях. Если энергия квантуется, то энергетический спектр называется дискретным (квантовым), если может принимать непрерывный ряд значений — спектр называется сплошным (непрерывным).

Плотность состояний $g(E)$ — число квантовых состояний электронов на единицу объема, площади или длины (в зависимости от размерности объекта), отнесенное к единичному интервалу энергий. Согласно этому определению, плотность состояний

$$g(E) = \frac{dn(E)}{dE},$$

где $dn(E)$ — число возможных состояний в интервале энергий от E до $E + dE$.

Знание плотности состояний $g(E)$ и вероятности их заполнения электронами $f(E)$ позволяет установить распределение электронов рассматриваемой системы по квантовым состояниям и описать электрические, оптические и некоторые другие свойства системы. Электроны обладают полуцелым спином. Поэтому вероятность заполнения ими квантовых состояний определяется статистикой Ферми — Дирака и подчиняется принципу Паули.

Плотности состояний 3D-электронного газа. Электроны могут двигаться свободно в любом направлении. Как известно, соотношение неопределенностей характеризует область пространственной локализации микрочастицы при заданном интервале проекций импульса:

$$dx dp_x \geq h; \;\; dy dp_y \geq h; \;\; dz dp_z \geq h.$$

Перемножим левые и правые части неравенств:

$$dx\, dy\, dz\, dp_x\, dp_y\, dp_z \geq h^3.$$

Обозначим неопределенности, дающие при перемножении точно h^3, индексом «0»:

$$dx^0\, dy^0\, dz^0\, dp_x^0\, dp_y^0\, dp_z^0 = h^3. \qquad (1.25)$$

Равенство (1.25) характеризует минимальную, так называемую фазовую ячейку в шестимерном пространстве координат-импульсов, «разрешенную» частице.

Пусть dz — число минимальных (элементарных) фазовых ячеек в каком-либо объеме $dx\, dy\, dz\, dp_x\, dp_y\, dp_z$ шестимерного пространства координат-импульсов:

$$dz = \frac{dx\, dy\, dz\, dp_x\, dp_y\, dp_z}{dx^0\, dy^0\, dz^0\, dp_x^0\, dp_y^0\, dp_z^0},$$

или

$$dz = \frac{dV dp_x dp_y dp_z}{h^3}.$$

Рассмотрим единичный объем твердого тела $dV = 1$. Тогда число элементарных фазовых ячеек в единичном обычном объеме в некотором объеме трехмерного пространства импульсов

$$dz = \frac{dZ}{dV} = \frac{dp_x dp_y dp_z}{h^3}.$$

Найдем dz в шаровом слое пространства импульсов, ограниченном сферами с радиусами p и $p + dp$; при этом $dp_x dp_y dp_z = 4\pi p^2 dp$ — объем шарового слоя:

$$E_k = \frac{mv^2}{2},$$

откуда

$$p^2 = (mv)^2 = 2mE_k = 2mE. \tag{1.26}$$

Приравняем E_k к E, так как потенциальная энергия свободной частицы равна нулю:

$$E = \frac{m^* v^2}{2} = \frac{\bar{p}^2}{2m^*} = \frac{p_x^2 + p_y^2 + p_z^2}{2m^*} = \frac{\hbar^2}{2m^*}\left(k_x^2 + k_y^2 + k_z^2\right).$$

Из соотношения (1.26)

$$dp = \sqrt{\frac{m}{2E}}dE,$$

тогда

$$dp_x dp_y dp_z = 4\pi p^2 dp = 4\pi 2mE\sqrt{\frac{m}{2E}}dE = 2\pi(2m)^{3/2} E^{1/2} dE.$$

Отсюда

$$dz = \frac{dp_x dp_y dp_z}{h^3} = \frac{2\pi(2m)^{3/2}}{h^3} E^{1/2} dE.$$

Тогда число элементарных фазовых ячеек, приходящихся на единичный интервал энергии,

$$g(E) = \frac{dZ}{dE} = \frac{2\pi(2m)^{3/2}}{h^3} E^{1/2}.$$

Здесь $g(E)$ называется плотностью состояний и является полуквадратичной функцией энергии E (рис. 1.13). С увеличением энергии возрастает и $dZ = gdE$, так как возрастает объем шарового слоя заданной толщины dp с ростом p. Следовательно, увеличивается и число элементарных фазовых ячеек.

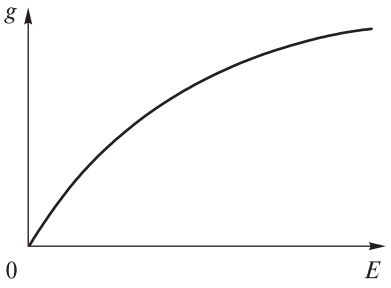


Рис. 1.13. Зависимость плотности состояний от энергии для 3D-электронного газа

Плотность состояний 2D-электронного газа. Движение электронов не ограничено вдоль двух осей x и y и ограничено отрезком d_y в направлении третьей оси z. Движение в направлении этой оси можно рассматривать как движение в одномерной глубокой прямоугольной потенциальной яме. Энергия такого движения квантуется и определяется по формуле

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2m^* d_y^2}.$$

Движение в одномерной потенциальной яме характеризуется квантовым числом n ($n = 1, 2, ...$). Если яма бесконечно глубока, то на ее ширине укладывается целое число $\frac{\lambda_n}{2}$:

$$d_y = \frac{n\lambda_n}{2}.$$

При этом устойчивыми могут быть только такие состояния движения электрона, которым соответствует стоячая волна, образованная падающей и отраженной от стенок ямы волной де Бройля. Величины E_n называются квантоворазмерными уровнями.

Энергия движений вдоль осей x и z не квантуется и определяется такими же выражениями, как для свободной частицы. Полную энергию электрона можно представить в виде

$$E = \frac{\hbar^2 k_x^2}{2m^*} + E_n + \frac{\hbar^2 k_z^2}{2m^*}.$$

Энергетический спектр электрона в квантовой яме двухмерного нанообъекта дискретно-непрерывный. Каждому размерному уровню E_n соответствует множество возможных значений E (подзон) за счет свободного движения электрона вдоль осей x и z. Эта совокупность энергий называется двумерной подзоной размерного квантования.

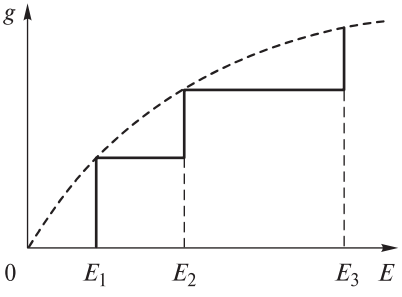


Рис. 1.14. Зависимость плотности состояний от энергии для 2D-электронного газа

Зависимость $g(E)$ имеет ступенчатый характер (рис. 1.14). Каждая размерная подзона вносит одинаковый вклад, равный $\frac{m^*}{\pi\hbar^2}$, в величину плотности состояний.

Модель бесконечно глубокой прямоугольной потенциальной ямы справедлива только для движения электрона в тонком слое при наноразмерной толщине d. Индивидуальный протяженный плоскопараллельный слой нанометровой толщины трудно реализовать. На практике в наноразмерных структурах создают области, где движение

носителей ограничено в одном измерении, и можно считать, что эти носители находятся в одномерной потенциальной яме, как, например, электроны в нанометровом слое узкозонного материала между двумя слоями широкозонного.

Плотность состояний 1D-электронного газа. Энергия электрона, связанная с движением вдоль осей y и z, должна квантоваться, как в одномерных потенциальных ямах шириной d_y и d_z. Полная энергия электрона равна

$$E = \frac{\hbar^2}{2m^*}k_x^2 + \frac{\hbar^2\pi^2}{2m^*}\frac{n^2}{d_y^2} + \frac{\hbar^2\pi^2}{2m^*}\frac{m^2}{d_z^2},$$

где m, n = 1, 2, 3, ..., т. е.

$$E = \frac{\hbar^2}{2m^*}k_x^2 + E_{mn},$$

где E_{mn} — энергия размерных уровней.

Положение каждого из них зависит от двух квантовых чисел m и n, а также от d_y и d_z. Зона проводимости в квантовой нити разбивается на одномерные подзоны. Плотность состояний на единицу длины $g(E)$ имеет ряд резких пиков, соответствующих размерным уровням (рис. 1.15).

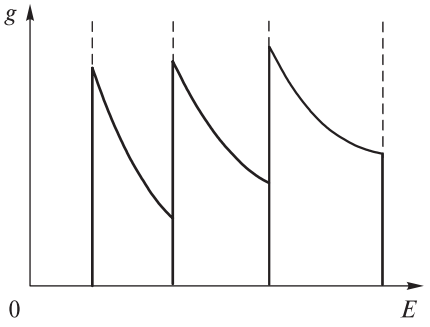


Рис. 1.15. Зависимость плотности состояний от энергии для 1D-электронного газа

Большинство электронов в подзоне имеет энергии вблизи соответствующего размерного уровня.

Плотность состояний 0D-электронного газа. Энергия свободных электронов должна квантоваться для движений во всех трех измерениях. Энергетический спектр электронов в квантовой точке полностью дискретен, как у отдельного атома. Энергия в этом случае

$$E_{lmn} = \frac{\hbar^2\pi^2}{2m^*}\cdot\frac{l^2}{d_x^2} + \frac{\hbar^2\pi^2}{2m^*}\cdot\frac{m^2}{d_y^2} + \frac{\hbar^2\pi^2}{2m^*}\cdot\frac{n^2}{d_z^2},$$

где l, m, n равны 1, 2, 3, ...; d_x, d_y, d_z — размеры области в трех измерениях.

Энергетический спектр электронов состоит из отдельных размерных уровней E_{lmn}. Величина E_{lmn} зависит от трех квантовых чисел и размеров d_x, d_y, d_z. График плотности состояний $g(E)$ в квантовой точке имеет так называемый δ-образный вид (рис. 1.16). Функция $g(E) = \infty$, если $E = E_{lmn}$ (E совпадает с размерным уровнем), функция $g(E) \to 0$, если $E \neq E_{lmn}$ (E лежит в промежутке между размерными уровнями).

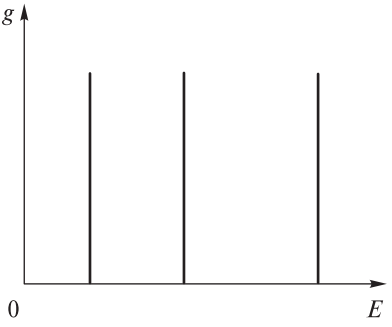


Рис. 1.16. Зависимость плотности состояний от энергии для 0D-электронного газа

Тесты к лекции 1.8

1. ***Как называются энергетические зоны в зонной теории твердых тел?***
а) валентная, проводимости и запрещенная;
б) ковалентная, потенциальная, кинетическая;
в) туннельная, генерации, рекомбинации.
2. ***Какая величина запрещенной зоны у диэлектриков?***
а) больше 3 эВ;
б) от 1 до 3 эВ;
в) менее 1 эВ.
3. ***Какой заряд имеет дырка в полупроводнике?***
а) положительный;
б) отрицательный;
в) нейтральный.
4. ***Какие носители заряда в полупроводнике являются подвижными?***
а) электроны и дырки;
б) ионы;
в) нейтроны.
5. ***Что представляет собой энергетический спектр частицы?***
а) совокупность возможных значений энергии частицы в данных условиях;
б) уровень энергии, достаточный для перехода частицы от одного атома к другому;
в) уровень энергии, достаточный для начала испускания частицей квантов света.
6. ***Какой вид имеет зависимость плотности состояний от энергии для 3D-электронного газа?***
а) параболический;
б) ступенчатый;
в) зубчатый.

1.9. СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

Цель лекции: ознакомление со статистикой электронов и дырок в полупроводниках.

1.9.1. СОБСТВЕННЫЕ И ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ

Собственные полупроводники. Химически чистые полупроводники называются собственными полупроводниками. К ним относятся как чистые химические элементы Ge и Si, так и химические соединения арсенид галлия GaAs, арсенид индия InAs, карбид кремния SiS и др. При температуре $T = 0$ валентная зона собственного полупроводника полностью занята электронами, а зона проводимости пуста (рис. 1.17, *а*), поэтому при $T = 0$ проводник обладает нулевой проводимостью и ведет себя как диэлектрик. При повышении температуры, вследствие термического возбуждения, часть электронов валентной зоны приобретает энергии, достаточные для преодоления запрещенной зоны и перехода в зону проводимости (рис. 1.17, *б*).

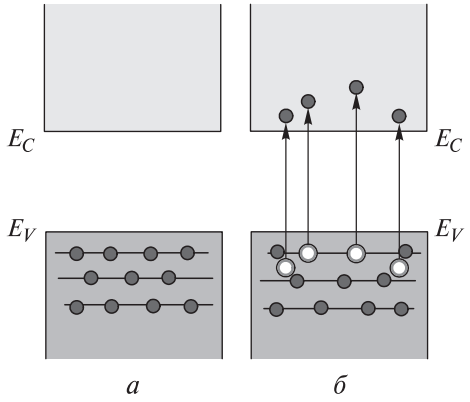


Рис. 1.17. Собственный полупроводник:
а — $T = 0$; *б* — $T \neq 0$

Это приводит к появлению в зоне проводимости свободных электронов, а в валентной зоне — свободных уровней, на которые могут переходить электроны этой зоны. Кристалл становится проводящим, так как в зоне проводимости появились свободные электроны, а в валентной зоне — уровни, на которые могут переходить электроны, т. е. валентная зона становится занятой не полностью.

Чем у́же запрещенная зона и выше температура кристалла, тем больше электронов переходит в зону проводимости, тем выше электропроводность.

Таким образом, проводимость полупроводников является возбужденной. Возбуждающим фактором могут быть нагрев, облучение светом, ионизирующим излучением.

Разделение тел на полупроводники и диэлектрики носит условный характер, так как их проводимость определяется температурой. Например, алмаз, являющийся диэлектриком при комнатной температуре, приобретает заметную проводимость при высокой температуре (~300 °С).

Примесные полупроводники. Полупроводники, как и все вещества, содержат примесные атомы, создающие свои собственные электрические уровни, называемые примесными уровнями. Эти уровни могут располагаться как в разрешенной, так и в запрещенной зонах полупроводника на различных расстояниях от дна зоны проводимости и потолка валентной зоны.

Донорные уровни. При замещении атома кремния Si атомом фосфора P заполняются все валентные связи и остается один лишний электрон, так как фосфор пятивалентен. Этот электрон вследствие воздействия электрических полей соседних атомов Si сравнительно слабо связан с атомом P и при сообщении ему незначительной энергии может покинуть атом и свободно перемещаться в кристаллической решетке Si. С точки зрения зонной теории этот процесс можно представить следующим образом. Лишнему электрону P соответствует энергетический уровень E_D (рис. 1.18).

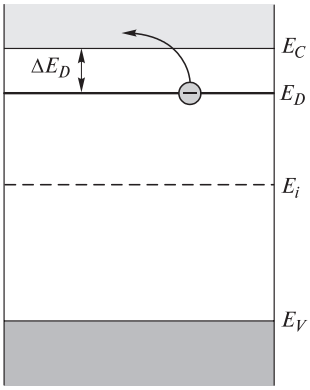


Рис. 1.18. Донорные уровни

Поскольку концентрация P невелика, лишние электроны не создают зону, а имеют один общий уровень энергии E_D, повторенный N_D раз (N_D — концентрация доноров, см$^{-3}$). Для каждого донора характерна своя энергия ионизации ΔE_D, соответствующая энергетическому зазору между E_D и дном зоны проводимости E_C. Чем меньше ΔE_D, тем легче (при более низкой температуре) можно ионизовать донор. В результате ионизации доноров, т. е. перехода примесных электронов с уровня энергии E_D в зону проводимости, образуются электроны проводимости. Образующиеся при этом положительные заряды локализуются на неподвижных атомах P и в электропроводности не участвуют.

Примеси, являющиеся источниками электронов проводимости, называются донорами, а энергетические уровни этих примесей — донорными уровнями. Полупроводники, содержащие донорную примесь, называются электронными полупроводниками, или полупроводниками n-типа, их также называют донорными полупроводниками.

Акцепторные уровни. При замещении атома Si атомом бора B образуются лишь три ковалентные связи, так как бор трехвалентен. Для образования четвертой связи не хватает одного электрона, который может заимствоваться

у атома кремния и тем самым заполнить связь с кремнием. При этом заимствованный электрон попадает на уровень акцептора E_A и закрепляется там, а у атома, от которого этот электрон перешел, образуется вакансия в валентной связи, чему соответствует образование свободной дырки вблизи потолка валентной зоны (рис. 1.19). Эта дырка может переходить на другие аналогичные вакантные уровни в валентной зоне под действием электрического поля, т. е. может участвовать в переносе заряда. Электроны, захваченные атомами бора, теряют способность перемещаться и в процессе проводимости не участвуют. Носителями заряда в акцепторных полупроводниках являются лишь дырки, возникающие в валентной зоне. Примеси, захватывающие электроны из валентной зоны полупроводника, называются акцепторными, а энергетические уровни этих примесей — акцепторными уровнями. Полупроводники, содержащие акцепторные примеси, называются дырочными полупроводниками, или полупроводниками *p*-типа.

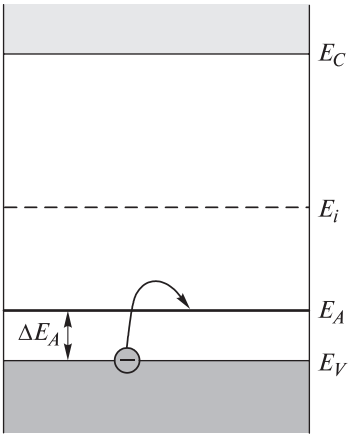


Рис. 1.19. Акцепторные уровни

1.9.2. ЗАВИСИМОСТЬ КОНЦЕНТРАЦИИ СВОБОДНЫХ НОСИТЕЛЕЙ В ПОЛУПРОВОДНИКЕ ОТ ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ

В отличие от металлов, в которых электронный газ является вырожденным и подчиняется статистике Ферми — Дирака, в собственных и слаболегированных примесных полупроводниках электронный (дырочный) газ является невырожденным, и распределение электронов по состояниям описывается классической статистикой Максвелла — Больцмана. Для таких полупроводников концентрация свободных носителей зависит от положения уровня Ферми и температуры T.

При температуре, отличной от нуля, в зоне проводимости полупроводника находятся электроны и в валентной зоне — дырки. Обозначим их концентрации соответственно n и p.

Определим равновесную концентрацию электронов проводимости в зоне проводимости. При этом за нулевой уровень отсчета энергии электронов принимаем потолок валентной зоны.

Концентрация электронов, энергия которых заключена в интервале от E до $E + dE$,

$$dn = g(E) \cdot 2 f_{\text{М–Б}} dE, \tag{1.27}$$

где $g(E)$ — число элементарных фазовых ячеек или состояний, приходящихся на единичный интервал энергии; $f_{\text{М–Б}}$ — функция распределения Максвелла — Больцмана:

$$f_{\text{М–Б}} = e^{\frac{E_F - E}{kT}},$$

выражающая вероятность заполнения электроном с энергией E фазовой ячейки; множитель «2» выражает принцип Паули.

Как известно,

$$g(E) = \frac{2\pi(2m)^{3/2}}{h^3} E^{1/2},$$

тогда выражение (1.27) можно записать в виде

$$dn = \frac{4\pi(2m_n)^{3/2}}{h^3} e^{\frac{E_F - E}{kT}} E^{3/2} dE. \tag{1.28}$$

Полное число электронов n, находящихся при температуре T в зоне проводимости, можно получить, интегрируя выражение (1.28) по всем значениям энергии, соответствующим зоне проводимости:

$$n = \frac{4\pi(2m_n)^{3/2}}{h^3} e^{\frac{E_F}{kT}} \int_E^{E_B} e^{-\frac{E}{kT}} E^{1/2} dE. \tag{1.29}$$

Проинтегрировав выражение (1.29), получим

$$n = \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{E_V - E_F}{kT}}; \tag{1.30}$$

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

где N_C — эффективная плотность состояний в зоне проводимости.

Если провести подобный расчет для дырок, то можно получить аналогичное выражение:

$$p = \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{\frac{E_V - E_F}{kT}}; \tag{1.31}$$

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

где N_V — эффективная плотность состояний в валентной зоне.

В выражениях (1.30) и (1.31) m_n, m_p — эффективные массы соответственно электронов и дырок.

Из выражений (1.30) и (1.31) следует, что концентрация свободных носителей заряда в данной зоне определяется расстоянием этой зоны от уровня Ферми E_F: чем больше это расстояние, тем ниже концентрация носителей.

Для любого невырожденного полупроводника произведение

$$np = 4\frac{(2\pi kT)^3 (m_n m_p)^{3/2}}{h^6} e^{\frac{E_V - E_C}{kT}}. \qquad (1.32)$$

Из формулы (1.32) следует, что при фиксированной температуре произведение концентраций электронов и дырок для данного полупроводника является постоянной величиной. Эта формула выражает так называемый закон действующих масс применительно к газу свободных носителей в полупроводниках.

1.9.3. УРОВЕНЬ ФЕРМИ И РАВНОВЕСНАЯ КОНЦЕНТРАЦИЯ НОСИТЕЛЕЙ В НЕВЫРОЖДЕННЫХ СОБСТВЕННЫХ ПОЛУПРОВОДНИКАХ

Если в полупроводнике отсутствуют примеси, так что $N_A = N_D = 0$, где N_A и N_D — концентрации соответственно акцепторной и донорной примесей, то такой полупроводник называется собственным, или чистым. В собственных полупроводниках концентрация электронов в зоне проводимости n_i равна концентрации дырок в валентной зоне p_i: $n_i = p_i$.

Индексами «*i*» обозначают равновесные концентрации в собственных полупроводниках. Приравнивая правые части соотношений (1.30) и (1.31), получаем

$$2\frac{(2\pi kT\overline{m}_n)^{3/2}}{h^3} e^{\frac{E_F - E_C}{kT}} = 2\frac{(2\pi kT\overline{m}_p)^{3/2}}{h^3} e^{\frac{E_V - E_F}{kT}};$$

$$e^{\frac{2E_F}{kT}} = \left(\frac{m_p}{m_n}\right)^{3/2} e^{\frac{E_C + E_V}{kT}},$$

откуда

$$E_F = \frac{E_C + E_V}{2} + kT \ln\left(\frac{m_p}{m_n}\right)^{3/4}.$$

Это соотношение определяет положение уровня Ферми в собственных полупроводниках. При абсолютном нуле ($T = 0$)

$$E_F = \frac{E_C + E_V}{2} = \frac{E_g}{2}, \qquad (1.33)$$

где E_g — ширина запрещенной зоны, т. е. при $T = 0$ уровень Ферми располагается посередине запрещенной зоны.

С повышением температуры он смещается вверх к дну зоны проводимости, если $m_p > m_n$, и вниз, если $m_p < m_n$ (рис. 1.20). Однако в большинстве случаев это смещение настолько незначительно, что им можно пренебречь

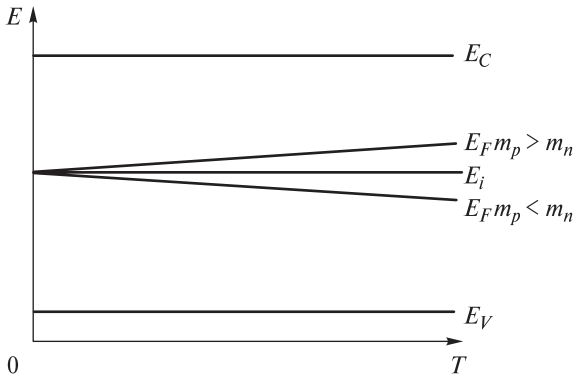


Рис. 1.20. Положение уровня Ферми в собственном полупроводнике

и считать, что уровень Ферми в собственных полупроводниках располагается посередине зоны.

Если подставить соотношение (1.33) в выражения (1.30) и (1.31), то получим

$$n_i = p_i = 2\frac{\left(2\pi\sqrt{m_n m_p}kT\right)^{3/2}}{h^3}e^{-\frac{E_g}{2kT}}. \tag{1.34}$$

Из выражения (1.34) следует, что равновесная концентрация носителей заряда в собственном полупроводнике определяется шириной запрещенной зоны E_g. Причем зависимость n_i от T и E_g значительная.

Из выражения (1.34) можно получить

$$\left(n_i\right)^2 = \frac{4\left(2\pi kT\right)^3\left(m_n m_p\right)^{3/2}}{h^6}e^{-\frac{E_g}{kT}}. \tag{1.35}$$

Правые части выражений (1.35) и (1.32) равны, следовательно, равны и левые:

$$np = n_i^2. \tag{1.36}$$

Соотношение (1.36) также выражает закон действующих масс, из него следует, что концентрации электронов и дырок в любом невырожденном полупроводнике таковы, что их произведение равно квадрату собственной концентрации электронов. Соотношение (1.36) справедливо при любой температуре:

$$n_{(T)}p_{(T)} = n_{i(T)}^2.$$

1.9.4. ПОЛОЖЕНИЕ УРОВНЯ ФЕРМИ И КОНЦЕНТРАЦИЯ НОСИТЕЛЕЙ В ПРИМЕСНЫХ ПОЛУПРОВОДНИКАХ

Область низких температур. При низких температурах энергия тепловых колебаний решетки значительно меньше ширины запрещенной зоны. Поэтому данные колебания не могут обеспечить переброс электронов из валентной зоны в зону проводимости. Однако этой энергии оказывается достаточно для переброса электронов с донорных уровней в зону проводимости, а дырок — с акцепторных уровней в валентную зону, так как для осуществления этих процессов необходима энергия на два порядка меньше, чем E_g. Следовательно, в области низких температур в примесных полупроводниках происходит возбуждение лишь «примесных» носителей заряда: электронов — в электронных полупроводниках и дырок — в дырочных полупроводниках.

Показано, что положение уровня Ферми в этой области температур определяет следующими соотношениями:

– для электронного полупроводника

$$E_F = \frac{E_C + E_D}{2} + \frac{kT}{2}\ln\left[\frac{N_D h^3}{2\left(2\pi m_n kT\right)^{3/2}}\right];$$

– для дырочного полупроводника

$$E_F = \frac{E_A + E_V}{2} + \frac{kT}{2}\ln\left[\frac{N_A h^3}{2\left(2\pi m_p kT\right)^{3/2}}\right],$$

где N_D, N_A — концентрации соответственно донорной и акцепторной примесей.

На рис. 1.21 показано изменение положения уровня Ферми с повышением температуры в примесных полупроводниках *n*- и *p*-типов соответственно.

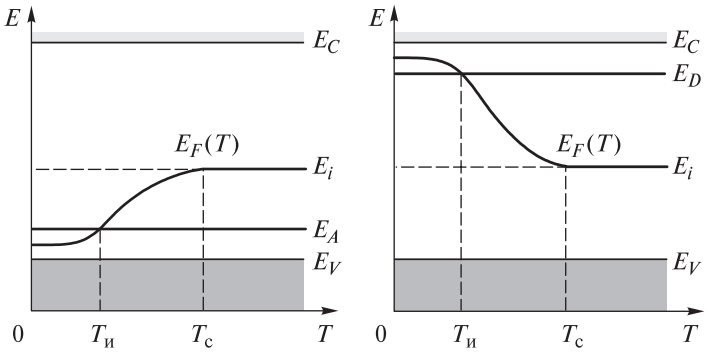


Рис. 1.21. Температурные зависимости положения уровня Ферми в примесных полупроводниках

Подставляя в соотношения (1.30) и (1.31) полученные выражения для положения уровня Ферми в примесных полупроводниках, соответственно получаем

$$n = \sqrt{N_D \frac{\left(2\pi m_n kT\right)^{3/2}}{h^3}} e^{-\frac{E_C - E_D}{2kT}};$$

$$p = \sqrt{N_A \frac{\left(2\pi m_p kT\right)^{3/2}}{h^3}} e^{-\frac{E_A}{2kT}}.$$

Данные выражения определяют концентрацию электронов в электронном полупроводнике и дырок — в дырочном полупроводнике.

Область истощения примесей. С увеличением температуры концентрация электронов в зоне проводимости возрастает, а концентрация электронов на донорных уровнях снижается, так как донорные уровни постепенно истощаются. Аналогично себя ведут и акцепторные уровни в дырочных полупроводниках.

При полном истощении примесей концентрация электронов в зоне проводимости у полупроводника *n*-типа становится практически равной концентрации донорной примеси N_D. Числом электронов, перешедших из валентной зоны, при этих температурах можно пренебречь, т. е.

$$n \approx N_D.$$

Аналогично для дырок

$$p \approx N_A.$$

Температура, при которой происходит полное истощение примесей $T_и$, тем выше, чем больше энергия ионизации примеси: $\Delta E_D = E_C - E_D$ и $\Delta E_A = E_A - E_V$ (для Ge $E_A = 0{,}01$ эВ, $T_и \approx 30$ К).

Область высоких температур. При дальнейшем повышении температуры начинается термическое возбуждение собственных носителей: в электронном полупроводнике все большее число электронов переходит из валентной зоны в зону проводимости. До тех пор пока концентрация собственных носителей (электронов, пришедших в зону проводимости из валентной зоны) остается меньше N_D $\left(n_i \ll N_D\right)$, суммарная концентрация электронов в зоне проводимости сохраняется практически постоянной и равной N_D. Однако с повышением температуры концентрация n_i собственных носителей увеличивается и может не только достичь N_D, но и значительно превзойти ее. При этом полупроводник все более приближается к состоянию собственного полупроводника $\left(n_i = p_i\right)$, вследствие чего уровень Ферми приближается к положению уровня Ферми в собственном полупроводнике. При $n_i \gg N_D$ $n = n_i + N_D \approx n_i$.

Это соответствует переходу к собственной проводимости полупроводника. Температура, при которой осуществляется переход к собственной проводимости полупроводника T_c, тем выше, чем больше ширина запрещенной зоны полупроводника и концентрация примеси в нем.

При температуре $T > T_c$ уровень Ферми в примесном полупроводнике совпадает с уровнем Ферми в собственном полупроводнике (см. рис. 1.21).

Качественные кривые зависимости логарифма концентрации электронов в зоне проводимости в *n*-полупроводнике от обратной температуры при различных концентрациях примеси приведены на рис. 1.22. На кривой можно выделить три участка: I — участок, соответствующий примесной проводимости полупроводника; II — участок, соответствующий области истощения примеси; III — участок собственной проводимости полупроводника. У полупроводников с высокой концентрацией примеси участок истощения примеси отсутствует (кривая N_{D4}).

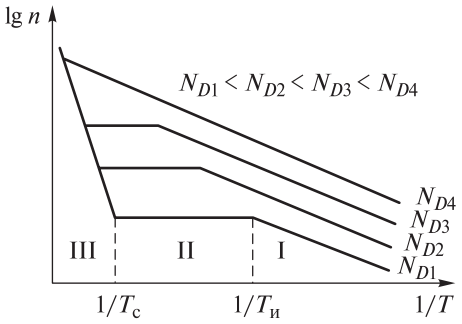


Рис. 1.22. Температурная зависимость концентрации носителей в примесных полупроводниках

Как отмечалось выше, в собственных полупроводниках проводимость осуществляется электронами зоны проводимости и дырками валентной зоны. В примесных полупроводниках проводимость обусловлена в основном носителями одного типа: в *n*-полупроводниках электронами зоны проводимости, так как их концентрация значительно превышает концентрацию дырок, в *p*-полупроводниках — дырками.

Электроны в *n*-полупроводнике и дырки в *p*-полупроводнике называют основными носителями заряда. Наряду с основными носителями в примесных полупроводниках содержатся неосновные носители заряда: дырки в *n*-полупроводниках и электроны в *p*-полупроводниках.

Основные носители обозначают n_{no}, p_{po} и неосновные — n_{po}, p_{no}. Таким образом, в *n*-полупроводнике концентрация электронов n_{no}, концентрация дырок p_{no}.

В таком случае закон действующих масс можно записать в следующем виде:
– для n-полупроводника

$$n_{no} p_{no} = n_i^2;$$

– для p-полупроводника

$$p_{po} n_{po} = n_i^2.$$

Следовательно, при легировании полупроводника примесью повышается концентрация основных носителей, что приводит к понижению концентрации неосновных носителей, так как их произведение должно оставаться неизменным.

Тесты к лекции 1.9

1. ***Какие полупроводники называются собственными?***
а) химически чистые;
б) содержащие собственную примесь;
в) сплавы различных полупроводников.
2. ***Какие полупроводники называются примесными?***
а) содержащие примеси;
б) сплав из двух полупроводников;
в) сплав из более чем двух полупроводников.
3. ***Какие примеси в полупроводниках называются донорными?***
а) примеси, являющиеся источниками электронов проводимости;
б) примеси, являющиеся источниками ионов в полупроводниках;
в) примеси, являющиеся источниками дырок в полупроводниках.
4. ***Какие примеси в полупроводниках называются акцепторными?***
а) примеси, захватывающие электроны из валентной зоны полупроводника;
б) примеси, захватывающие электроны из зоны проводимости полупроводника;
в) примеси, захватывающие дырки из запрещенной зоны полупроводника.
5. ***Оцените качественно наличие электронов в зоне проводимости собственного полупроводника при температуре абсолютного нуля:***
а) ни одного электрона;
б) имеется небольшое количество;
в) много.
6. ***Сформулируйте закон действующих масс в применении к газу свободных носителей в полупроводниках:***
а) произведение концентраций электронов и дырок для данного полупроводника при фиксированной температуре является постоянной величиной;
б) масса электронов и дырок неизменна при прохождении электрического тока через полупроводник;
в) суммарная масса электронов и дырок в полупроводнике пропорциональна величине тока, проходящего через полупроводник.

7. ***Что обусловливает электрический ток в примесных полупроводниках при низких температурах?***

а) примесные носители заряда;

б) ионы;

в) собственные носители заряда.

8. ***Что обусловливает электрический ток в примесных полупроводниках при высоких температурах?***

а) собственные носители заряда;

б) примесные носители заряда;

в) ионы.

1.10. НЕРАВНОВЕСНЫЕ НОСИТЕЛИ И УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Цель лекции: ознакомление с неравновесными носителями и уравнением непрерывности.

1.10.1. НЕРАВНОВЕСНЫЕ НОСИТЕЛИ, РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ

Наряду с процессом тепловой генерации электронов и дырок вследствие перехода электронов из валентной зоны в зону проводимости в полупроводниках протекает процесс рекомбинации свободных носителей. Если бы процесс генерации был единственным, то концентрация свободных носителей непрерывно возрастала с течением времени. Процесс генерации носителей уравновешивается рекомбинацией носителей. Электрон в зоне проводимости может потерять энергию и перейти вновь в валентную зону. При этом ликвидируется как электрон проводимости, так и дырка в валентной зоне.

При любой температуре в полупроводниках устанавливается равновесие между процессом тепловой генерации носителей и процессом рекомбинации. Этому равновесию соответствует равновесная концентрация носителей. Такие носители называются равновесными. Закон действующих масс применим только к равновесным носителям.

Если генерация носителей заряда осуществляется под действием других факторов, а не вследствие повышения температуры, таких как облучение светом, введение носителей заряда извне за счет электрического поля и т. п., то концентрация носителей будет отличаться от равновесной концентрации. В полупроводнике появятся избыточные носители, называемые неравновесными носителями. Концентрации таких носителей обозначают Δn и Δp. Полная концентрация носителей заряда

$$n = n_o + \Delta n;$$
$$p = p_o + \Delta p,$$

где n_o, p_o — равновесные концентрации.

Каждый неравновесный носитель, возникнув в полупроводнике, существует в нем определенное время, ограниченное его рекомбинацией. Для отдельных носителей это время различно, поэтому вводят среднее время жизни носителей: τ_n — для электронов, τ_p — для дырок.

Процесс генерации характеризуется скоростью генерации *G*, которая равна числу носителей, появляющихся в единицу времени в единичном объеме

полупроводника. Аналогично процесс рекомбинации характеризуют скоростью рекомбинации:

$$R_n = -\frac{dn}{dt} = -\frac{d\left(n_o + \Delta n\right)}{dt} = -\frac{d\left(\Delta n\right)}{dt}; \quad R_p = -\frac{d\left(\Delta p\right)}{dt},$$

знак «минус» указывает на то, что в процессе рекомбинации концентрация носителей уменьшается.

Виды рекомбинации. Процесс рекомбинации носителей, т. е. переход электрона из зоны проводимости в валентную зону, может происходить двумя путями. При межзонной рекомбинации электрон из зоны проводимости переходит в валентную зону. Возможен второй способ перехода электрона — сначала он переходит на примесный уровень $E_{\text{п}}$, а затем с примесного уровня — в валентную зону. Этот тип рекомбинации называется *рекомбинацией через примесный уровень*. В обоих случаях выделяется одна и та же энергия. Только в первом случае она выделяется сразу, а во втором — постепенно (рис. 1.23).

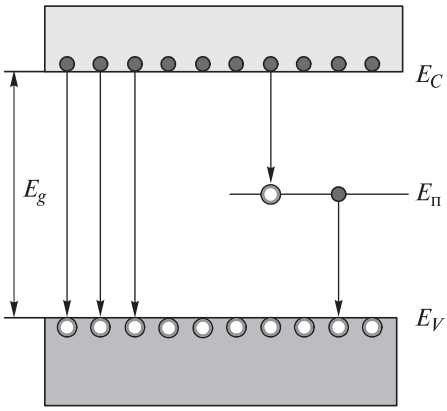


Рис. 1.23. Межзонная рекомбинация и рекомбинация через примесный уровень

Выделение энергии может происходить в виде квантов света $h\nu$ или в виде фотонов. Тепловые колебания кристаллической решетки квантованы, соответствующие им квазичастицы (аналогичны квантам света) называются фотонами $h\nu_g$, где ν_g — частота нормальных колебаний решетки. Если при рекомбинации происходит выделение кванта света, то она называется излучательной. Если вся энергия электронов при рекомбинации переходит в тепловую энергию, то она называется безызлучательной.

1.10.2. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

Процессы рекомбинации как в приповерхностном слое, так и в объеме полупроводника в принципе не отличаются. Однако, так как на поверхности полупроводника имеются добавочные поверхностные уровни, то возможности рекомбинации носителей на поверхности увеличиваются по сравнению с объемом. Поскольку активные области полупроводниковых приборов расположены вблизи поверхности полупроводника, влияние поверхности необходимо учитывать при их проектировании и анализе работы.

Обозначим поверхностное время жизни носителей τ_S, объемное — τ_V. Если рабочий участок полупроводникового прибора расположен частично

в объеме и частично в приповерхностном слое, используют эффективное время жизни носителей τ, которое определяется из выражения

$$\frac{1}{\tau} = \frac{1}{\tau_S} + \frac{1}{\tau_V}.$$

Поскольку $\tau_S < \tau_V$, эффективное время жизни ближе к τ_S.

Поверхностное время жизни трудно определить экспериментально или рассчитать, поэтому для характеристики поверхностной рекомбинации используется параметр s — скорость поверхностной рекомбинации, который впервые был введен Шокли.

Скорость поверхностной рекомбинации представляет собой коэффициент, определяющий скорость потока носителей заряда к поверхности. Направленные потоки носителей из объема к поверхности полупроводника возникают вследствие более высокой интенсивности рекомбинации у его поверхности:

$$J_n = qs\Delta n;$$
$$J_p = qs\Delta p,$$

где J_n, J_p — плотность тока соответственно электронов и дырок, текущего к поверхности полупроводника.

Данные токи являются паразитными, и их следует уменьшать, снижая скорость поверхностной рекомбинации всеми возможными средствами. Скорость поверхностной рекомбинации существенно зависит от способа и качества обработки кристалла и составляет $10^2...10^4$ см/с.

1.10.3. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Рассмотрим локальный объем полупроводника с одиночной площадью $S = 1$ см^2 и толщиной dx. Объем такой области $V = Sdx = dx$ (рис. 1.24). Пусть $n_{(x,t)}$ — концентрация электронов в данной области. Общее число электронов при этом составит $n_{(x,t)}dx$. Изменение числа электронов в этой локальной области за время dt

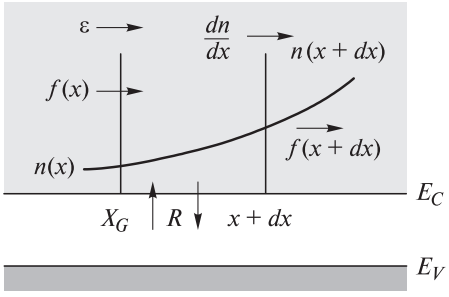


Рис. 1.24. Потоки носителей в полупроводнике

$$\left[n_{(x,t+dt)} - n_{(x,t)}\right]dx = \left(\frac{dn}{dt}\right)dt\,dx.$$

Изменение числа электронов может происходить вследствие процессов:

– термической генерации

$$n_G = Gdt\,dx,$$

где G — скорость генерации;

– рекомбинации носителей

$$n_R = -R \cdot dt\, dx;$$

– разности входящего и выходящего потоков электронов в локальную область и из нее под действием электрического поля и градиента концентрации

$$\left[f_n(x,t) - f_n(x + dx_1 t) \right] dt = -\left(\frac{df_n}{dx} \right) dx\, dt.$$

Общее изменение концентрации электронов

$$\left(\frac{dn}{dt} \right) dt\, dx = G dt\, dx - R dt\, dx - \left(\frac{df_n}{dt} \right) dx\, dt,$$

тогда

$$\left(\frac{dn}{dt} \right) = G - R - \left(\frac{df_n}{dx} \right).$$

Плотность тока электронов $j_n = -f_n q,$ где q — заряд электрона, следовательно, $f_n = -j/q,$ тогда

$$\left(\frac{dn}{dt} \right) = G - R + \frac{1}{q} \frac{dj_n}{dx}.$$

Это выражение называется уравнением непрерывности. Плотность тока j_n в общем случае складывается из дрейфовой и диффузионных составляющих.

Пусть $j_n = 0,$ тогда

$$\frac{dn}{dt} = G - R.$$

В равновесном состоянии $G = R\, dn/dt = 0.$ В неравновесном состоянии $G \neq R.$ Обозначим $G - R = -\Delta n(t)/\tau_n,$ где $\Delta n(t)$ — избыточная концентрация электронов, являющаяся функцией времени. Подставив ее в выражение для dn/dt, получим

$$\frac{dn}{dt} = -\frac{\Delta n(t)}{\tau_n}.$$

Решение этого уравнения будет иметь вид

$$\Delta n(t) = \Delta n(0) e^{-t/\tau_n},$$

где $\Delta n(0)$ — избыточная концентрация в начальный момент времени.

Величину τ_n называют временем жизни носителей, в данном случае электронов. Это среднее время, в течение которого избыточная концентрация уменьшается в e раз вследствие рекомбинации.

Тесты к лекции 1.10

1. ***Какие носители в полупроводнике называются неравновесными?***

а) избыточные;

б) основные;

в) неосновные.

2. ***Что называется генерацией носителей?***

а) переход электронов из валентной зоны в зону проводимости;

б) переход электронов из зоны проводимости в валентную зону;

в) переход дырок из валентной зоны в запрещенную зону.

3. ***От чего зависит скорость поверхностной рекомбинации носителей в полупроводнике?***

а) от способа и качества обработки полупроводникового кристалла;

б) от концентрации носителей;

в) от наличия примесей.

4. ***Что такое межзонная рекомбинация носителей?***

а) рекомбинация непосредственно между зоной проводимости и валентной зоной;

б) рекомбинация через один примесный уровень;

в) рекомбинация через два примесных уровня.

5. ***Что такое излучательная рекомбинация?***

а) выделение кванта света;

б) выделение донорных носителей;

в) выделение акцепторных носителей.

6. ***Перечислите все процессы, в результате которых в полупроводнике может происходить изменение числа носителей:***

а) термическая генерация, рекомбинация носителей, действие электрического поля и градиента концентрации;

б) термическая генерация, действие электрического поля и градиента концентрации;

в) термическая генерация и действие электрического поля.

7. ***Что такое время жизни носителей в полупроводнике?***

а) среднее время, в течение которого избыточная концентрация уменьшается в *e* раз вследствие рекомбинации;

б) время между взаимными столкновениями носителей;

в) время между столкновениями носителей с атомами кристаллической решетки.

1.11. ЭЛЕКТРОПРОВОДНОСТЬ ТВЕРДЫХ ТЕЛ

Цель лекции: ознакомление с электропроводностью твердых тел.

1.11.1. ДВИЖЕНИЕ ЭЛЕКТРОНОВ ПОД ДЕЙСТВИЕМ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

В полупроводнике, на который не действует электрическое поле, число электронов, движущихся в противоположных направлениях, всегда одинаково, поэтому их средняя скорость в любом направлении равна нулю, и электрический ток не возникает.

При приложении внешнего электрического поля к полупроводнику в нем возникает направленное движение электронов, т. е. электрический ток. Согласно закону Ома,

$$j = \sigma\mathcal{E},$$

где σ — удельная электропроводность [Ом$^{-1}\cdot$м$^{-1}$]; $\mathcal{E}$ — напряженность электрического поля.

Часто пользуются не удельной электропроводностью, а удельным сопротивлением:

$$\rho = 1/\sigma.$$

Направленное движение электронов под действием внешнего электрического поля называется дрейфом электронов. Средняя скорость этого движения называется скоростью дрейфа v_D.

При наложении внешнего поля на электрон действует сила $F = -q\mathcal{E}$, под действием которой электрон должен был бы двигаться ускоренно, непрерывно увеличивая свою скорость. Однако при движении электрон сталкивается с дефектами кристаллической решетки и вследствие этого теряет свою скорость. Под дефектами в данном случае понимают любые нарушения кристаллической решетки — примесные атомы, отклонения атомов от своего положения вследствие теплового движения и т. д. Действие решетки на электрон аналогично действию силы сопротивления, препятствующей движению электрона.

Тогда дрейфовая скорость электрона кристалла в электрическом внешнем поле

$$v_D = a\tau,$$

где a — ускорение электрона под действием внешнего электрического поля; τ — среднее время между двумя столкновениями электрона с решеткой, называемое средним временем свободного пробега.

Начальная дрейфовая скорость $v_{D0} = 0$, так как считается, что после каждого очередного столкновения направленное движение прекращается.

Поскольку $F = -q\mathcal{E}$; $F = m_n a$,

$$-q\mathcal{E} = ma;$$

$$-q\mathcal{E} = m\frac{v_D}{\tau};$$

$$v_D = -q\frac{e\tau}{m_n}\mathcal{E}.$$

Отношение скорости дрейфа к напряженности поля называется подвижностью носителей:

$$\mu_n = \frac{v_D}{\mathcal{E}} = -\frac{q\tau}{m_n}.$$

При более строгом рассмотрении τ можно определить как время релаксации, характеризующее процесс восстановления равновесного состояния, нарушенного электрическим полем, после снятия последнего. При этом считается, что время релаксации после выключения электрического поля, вызвавшего отклонения от равновесного состояния, равно среднему времени свободного пробега.

Можно считать, что электрон в кристалле движется прямолинейно до тех пор, пока не встретится с дефектом решетки и не рассеется. Средний отрезок пути λ, который проходит электрон между двумя последовательными актами рассеяния, принимают за длину свободного пробега. Если электрон уже в одном акте рассеяния полностью теряет скорость в данном направлении и возвращается в состояние хаотического движения, то

$$\lambda = v\tau,$$

где v — средняя скорость движения электронов; τ — просто время свободного движения электрона.

Однако для полной потери скорости в данном направлении требуется не одно, а в среднем ν столкновений с рассеивающими центрами. Время, в течение которого будет происходить ликвидация направленного движения, представляет собой время релаксации. Тогда средний путь, который пройдет электрон за это время,

$$L = \lambda\nu = v\tau,$$

где L — средняя транспортная длина свободного пробега.

Отсюда

$$\tau = \frac{\nu\lambda}{v}.$$

1.11.2. ЗАВИСИМОСТЬ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА ОТ ТЕМПЕРАТУРЫ

Рассеяние электронов и дырок проводимости может происходить на ионах и нейтральных атомах примеси, дефектах структуры, границах кристалла, тепловых колебаниях решетки.

В области высоких температур основное значение имеет рассеяние электронов на тепловых колебаниях решетки. Установлено, что в области высоких температур, в которой основное значение имеет рассеяние на фононах, т. е. на тепловых колебаниях решетки, подвижность невырожденного газа пропорциональна температуре в степени –3/2:

$$\mu_{\text{М–Б}} \sim T^{-3/2}.$$

Подвижность носителей вырожденного газа обратно пропорциональна температуре:

$$\mu_{\text{Ф–Д}} \sim T^{-1}.$$

В области низких температур основное значение имеет рассеяние на ионизированных примесных атомах. Ионы примеси отклоняют электроны, проходящие вблизи них, и тем самым уменьшают скорость их движения в первоначальном направлении.

Показано, что подвижность носителей заряда в области низких температур, обусловленная рассеянием на ионизированных примесях, пропорциональна $T^{3/2}$ для проводников с невырожденным газом и не зависит от T для проводников с вырожденным газом.

У чистых металлов концентрация примеси очень мала, и при низких температурах основным механизмом рассеяния остается рассеяние на колебаниях кристаллической решетки, поэтому подвижность вырожденного электронного газа в этих металлах пропорциональна T^{-5}.

1.11.3. ЭЛЕКТРОПРОВОДНОСТЬ ЧИСТЫХ МЕТАЛЛОВ

Вычислим плотность электрического тока, протекающего по кристаллу, независимо от того, является ли он полупроводником или металлом.

Пусть электроны обладают скоростью v_D. Построим мысленно внутри полупроводника цилиндр, площадь основания которого равна единице, а образующая — v_D (рис. 1.25). Все электроны, заключенные в этом цилиндре, за 1 с пройдут через основание.

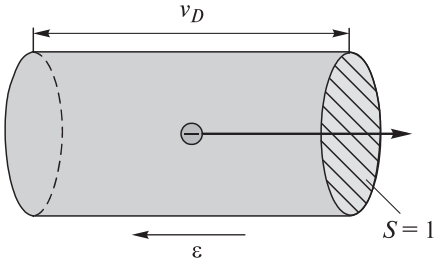


Рис. 1.25. Цилиндр, выделенный в кристалле

При этом плотность тока

$$j = qnv_D = qn\mu\mathcal{E};$$

так как $j = \sigma\mathcal{E}$, то

$$\sigma = qn\mu,$$

где q — заряд электрона; n — число электронов в единице объема, т. е. концентрация; μ — подвижность.

Электропроводность металлов обусловлена дрейфом свободных носителей заряда одного знака. У большинства металлов этими носителями являются электроны и лишь у некоторых — дырки (бериллий, цинк).

Металлы — это вырожденные проводники, поэтому концентрация электронного газа в них практически не зависит от температуры. Зависимость электропроводности от температуры полностью определяется температурной зависимостью подвижности электронов вырожденного газа, поэтому в области высоких температур

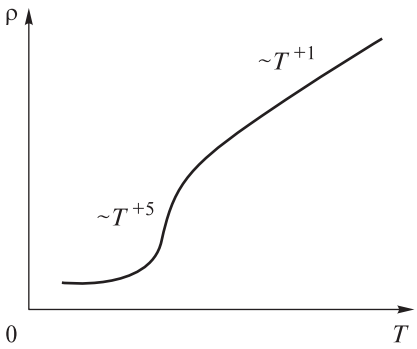


Рис. 1.26. Зависимость удельного сопротивления от температуры

$$\sigma = A/T,\ \rho_V = aT;$$

в области низких температур

$$\sigma = B/T^5,\ \rho_V = bT^5,$$

где A, B, a, b — коэффициенты пропорциональности.

Зависимость удельного сопротивления металлов от температуры приведена на рис. 1.26. В области абсолютного нуля основное значение даже в чистых металлах приобретает рассеяние на примесях. При этом ρ_V не зависит от температуры.

1.11.4. ЭЛЕКТРОПРОВОДНОСТЬ СОБСТВЕННЫХ ПОЛУПРОВОДНИКОВ

В собственных полупроводниках носителями заряда являются как электроны в зоне проводимости, так и дырки в валентной зоне. Электропроводность собственных полупроводников складывается из проводимости электронов: $\sigma_n = qn_i\mu_n$, имеющих концентрацию n_i и подвижность μ_n, и из проводимости дырок: $\sigma_p = qp_i\mu_p$, имеющих концентрацию p_i и подвижность μ_p, т. е.

$$\sigma_i = qn_i\mu_n + qp_i\mu_p.$$

Поскольку у собственных полупроводников $n_i = p_i$,

$$\sigma_i = qn_i\left(\mu_n + \mu_p\right).$$

Как известно,

$$n_i = \frac{2\left(2\pi\sqrt{m_n m_p}\,kT\right)^{3/2}}{h^3} e^{-\frac{E_g}{2kT}},$$

следовательно,

$$\sigma_i = q\left(\mu_n + \mu_p\right)\frac{2\left(2\pi\sqrt{m_n m_p}\,kT\right)^{3/2}}{h^3} e^{-\frac{E_g}{2kT}} = \sigma_0 e^{-\frac{E_g}{2kT}}.$$

Температурную зависимость электропроводности удобно представлять в полулогарифмических координатах:

$$\lg\sigma_i = \lg\sigma_0 - \frac{E_g}{2kT}.$$

Качественная зависимость $\lg\sigma_i = f\left(1/T\right)$ показана на рис. 1.27. Это прямая, тангенс угла наклона α которой составляет

$$\operatorname{tg}\alpha = \frac{E_g}{2k}.$$

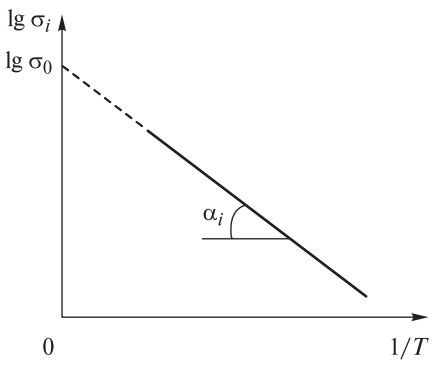


Рис. 1.27. Температурная зависимость электропроводности собственного полупроводника

Отсюда можно определить ширину запрещенной зоны полупроводника. Пересечение продолжения прямой $-\sigma_i = f\left(1/T\right)$ с осью абсцисс позволяет определить $\lg\sigma_0$, а следовательно, и σ_0. При данной температуре проводимость собственных полупроводников определяется шириной их запрещенной зоны.

Таким образом, между электропроводностью металлов и полупроводников существует принципиальное различие.

В металлах, в которых электронный газ является вырожденным и концентрация носителей заряда практически не зависит от температуры, температурная зависимость проводимости полностью определяется подвижностью носителей.

В полупроводниках, в которых электронный газ в невырожденном состоянии и его концентрация резко зависят от температуры, температурная зависимость проводимости практически полностью определяется температурной зависимостью концентрации носителей.

1.11.5. ЭЛЕКТРОПРОВОДНОСТЬ ПРИМЕСНЫХ ПОЛУПРОВОДНИКОВ

Температурная зависимость электропроводности невырожденных примесных полупроводников, как и собственных полупроводников, определяется температурной зависимостью концентрации носителей. Типичная кривая температурной зависимости электропроводности имеет три характерных участка (рис. 1.28).

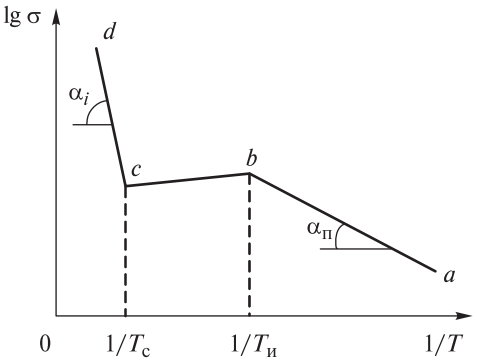


Рис. 1.28. Температурная зависимость электропроводности примесного полупроводника ($\alpha_п$ — угол, определяемый примесными носителями)

Участок *ab* простирается от области низких температур до температуры истощения примеси $T_и$. Как известно, концентрация носителей в этой области

$$n = \sqrt{N_D \left(\frac{2\pi m_n kT}{h^3} \right) e^{3/2}} - \frac{\Delta E_D}{2kT}.$$

Подвижность носителей в этой области температур определяется расстоянием на примесях и пропорциональна $T^{3/2}$, тогда

$$\sigma = qn\mu = \sigma_{п.о} e^{-\frac{\Delta E_D}{2kT}},$$

где $\sigma_{п.о}$ — коэффициент, слабо зависящий от температуры (по сравнению с экспоненциальной зависимостью), определяемый примесными основными носителями.

Тангенс угла наклона участка *ab* составляет

$$\mathrm{tg}(a_п) = \frac{\Delta E_D}{2k},$$

т. е. пропорционален энергии ионизации донорных уровней. На участке *ab* проводимость полупроводника определяется появлением примесных носителей, возникающих вследствие ионизации примесных атомов.

При температуре истощения примесей $T_и$ участок *ab* переходит в участок *bc*, на котором все примесные атомы ионизированы, но еще не происходит теплового возбуждения собственных носителей. На этом участке концентрация носителей практически постоянна и равна концентрации примеси: $n \approx N_D$, поэтому температурная зависимость электропроводности определяется температурной зависимостью подвижности. Если в области температур от $T_и$ до $T_с$ основным механизмом рассеяния является рассеяние на тепловых колебаниях решетки, для которого подвижность носителей пропорциональна $T^{-3/2}$, то проводимость на участке *bc* будет падать с ростом температуры.

Если же на участке *bc* основным механизмом рассеяния является рассеяние на примесях, т. е. $\mu \sim T^{3/2}$, то электропроводность будет увеличиваться с ростом T.

При температуре T_c участок *bc* переходит в участок *cd*, соответствующий собственной проводимости. На участке *cd* концентрация носителей приблизительно равна концентрации собственных носителей. Проводимость на этом участке

$$\sigma \approx \sigma_i = \sigma_0 e^{-\frac{E_g}{2kT}}.$$

Тангенс угла наклона участка *cd* пропорционален ширине запрещенной зоны:

$$\mathrm{tg}(\alpha_c) = \frac{E_g}{2k}.$$

1.11.6. ДИФФУЗИОННЫЕ УРАВНЕНИЯ

Пусть вдоль полупроводника имеется градиент концентрации свободных носителей заряда, создание которого возможно с помощью освещения образца, его неравномерного нагрева и т. д.

В общем случае ток проводимости состоит из геометрической суммы дрейфового и диффузионного токов.

Дрейфовые составляющие плотности тока проводимости определяют по закону Ома:

$$j_{n\,\text{др}} = \sigma_n \mathcal{E} = qn\mu_n \mathcal{E} = -qn\mu_n \,\mathrm{grad}(\varphi);$$

$$j_{p\,\text{др}} = \sigma_p \mathcal{E} = qp\mu_p \mathcal{E} = -qp\mu_p \,\mathrm{grad}(\varphi),$$

где n, p — концентрации электронов и дырок, которые в общем случае не являются равновесными; φ — электрический потенциал.

Диффузионная составляющая тока зависит от градиента концентрации свободных носителей заряда:

$$j_{n\,\text{диф}} = qD_n \,\mathrm{grad}(n);$$

$$j_{p\,\text{диф}} = -qD_p \,\mathrm{grad}(p).$$

Диффузионный ток создает объемные заряды, поле которых вместе с внешним электрическим полем обусловливает дрейфовый ток.

В представленном на рис. 1.29 случае, когда $p_1 > p_2$, левая часть полупроводника будет заряжаться отрицательно, а правая — положительно и внутреннее электрическое поле полупроводника, обусловленное объемными зарядами, направлено против внешнего электрического поля.

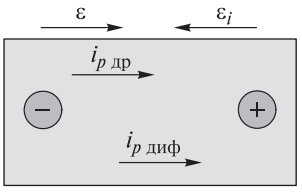


Рис. 1.29. Объемные заряды в полупроводнике

В общем случае ток носителей одного вида в полупроводнике равен сумме диффузионного и дрейфового токов:

$$j_n = qn\mu_n\mathcal{E} + qD_n \operatorname{grad}(n) = -qn\mu_n \operatorname{grad}(\varphi) + qD_n \operatorname{grad}(n);$$

$$j_p = qp\mu_p\mathcal{E} + qD_p \operatorname{grad}(p) = -qp\mu_p \operatorname{grad}(\varphi) - qD_p \operatorname{grad}(p).$$

Если рассматривать токи в каком-либо одном направлении, то эти уравнения можно записать в виде:

$$j_{nx} = qn\mu_n\mathcal{E}_x + qD_n\frac{dn}{dx} = -qn\mu_n\frac{d\varphi}{dx} + qD_n\frac{dn}{dx};$$

$$j_{px} = qp\mu_p\mathcal{E}_x + qD_p\frac{dp}{dx} = -qp\mu_p\frac{d\varphi}{dx} - qD_p\frac{dp}{dx}.$$

Полная плотность тока

$$j = j_n + j_p = j_{n\,\text{др}} + j_{n\,\text{диф}} + j_{p\,\text{др}} + j_{p\,\text{диф}} =$$

$$= -q\left(n\mu_n + p\mu_p\right)\mathcal{E} + q\left(D_n\frac{dn}{dx} - D_p\frac{dp}{dx}\right).$$

Направления токов определяют обычно в зависимости от направления градиентов потенциала φ и концентраций. Знаки токов учитывают при записи уравнений.

Связь между дрейфовой подвижностью носителей и коэффициентами диффузии выражается с помощью уравнений Эйнштейна:

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

Подвижность носителей пропорциональна коэффициенту диффузии.

Тесты к лекции 1.11

1. ***Что такое дрейф электронов?***
а) направленное движение электронов под действием внешнего поля;
б) перемещение электронов в полупроводнике между дислокациями;
в) броуновское движение электронов в результате нагрева полупроводника.
2. ***Что называется подвижностью носителей?***
а) отношение скорости дрейфа к напряженности поля;
б) скорость перемещения носителей по полупроводнику;
в) ускорение носителей в полупроводнике.

3. ***Чем обусловлена электропроводность металлов?***

а) дрейфом свободных носителей заряда одного знака;

б) движением ионов;

в) тепловым колебанием кристаллической решетки.

4. ***Чем обусловлена электропроводность собственных полупроводников?***

а) электронами в зоне проводимости и дырками в валентной зоне;

б) донорными носителями заряда;

в) акцепторными носителями заряда.

5. ***От чего зависит концентрация носителей в собственных полупроводниках?***

а) от температуры;

б) от размера полупроводника;

в) от влажности.

6. ***Что такое температура истощения примесей?***

а) температура, при которой все примесные носители переходят в зону проводимости или валентную зону;

б) температура, при которой все примеси испаряются из полупроводника;

в) температура, при которой все примеси растворяются в полупроводнике.

7. ***Какая температура называется температурой собственной проводимости?***

а) температура, при которой проводимость примесного полупроводника начинает обусловливаться собственными носителями;

б) температура, при которой собственный полупроводник перестает быть таковым;

в) температура, при которой примесный полупроводник начинает проводить ток.

1.12. КОНТАКТНЫЕ ЯВЛЕНИЯ

Цель лекции: ознакомление с понятием контактных явлений.

1.12.1. КОНТАКТ ЭЛЕКТРОННОГО И ДЫРОЧНОГО ПОЛУПРОВОДНИКОВ

Соответствующим введением примесей в полупроводник можно создать такое их распределение, что одна часть кристалла будет полупроводником *n*-типа, а другая — полупроводником *p*-типа.

Электронно-дырочным переходом (*p*–*n*-переходом) называют слой полупроводника, расположенный по обе стороны от границы раздела *p*- и *n*-областей.

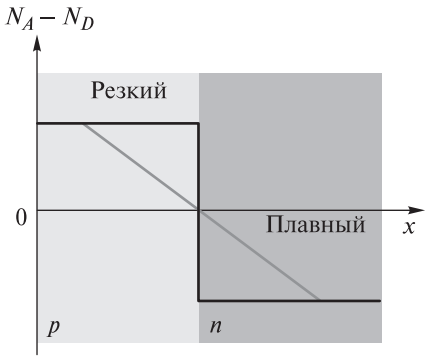


Рис. 1.30. Плавный и резкий *p*–*n*-переходы

В зависимости от характера распределения примесей различают резкий и плавный *p*–*n*-переходы. В резком *p*–*n*-переходе концентрация акцепторов и доноров изменяется скачкообразно на границе *p*- и *n*-областей (рис. 1.30).

В плавном переходе концентрация акцепторов и доноров является линейной функцией расстояния (см. рис. 1.30).

Относительно резкий *p*–*n*-переход можно создать при вплавлении примеси, плавный — при диффузии.

Поскольку на границе раздела *p*- и *n*-областей имеется градиент концентрации свободных носителей заряда, будет происходить процесс диффузии электронов в *p*-область и дырок — в *n*-область. Это приводит к обеднению основными носителями заряда приграничных слоев и возникновению объемных зарядов противоположного знака. В *p*-полупроводнике в приграничном слое снижается концентрация дырок, в *n*-полупроводнике — концентрация электронов.

В резком *p*–*n*-переходе создаются обедненные слои ступенчатого объемного заряда (рис. 1.31), в плавном — линейного объемного заряда (см. рис. 1.31). Для случая на рис. 1.31 и 1.32 концентрация акцепторов в *p*-области выше концентрации доноров в *n*-области. Толщины слоев обратно пропорциональны концентрациям примесей в областях полупроводника. Однако при любых соотношениях концентрации примесей в областях полупроводника сумма объемных зарядов в *p*- и *n*-областях равна нулю, т. е. площади под кривыми $\rho(x)$ равны между собой.

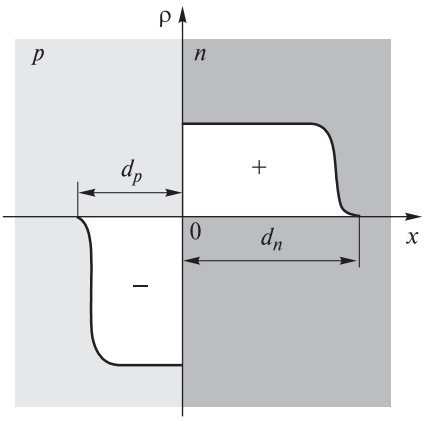


Рис. 1.31. Объемный заряд в резком *p*–*n*-переходе

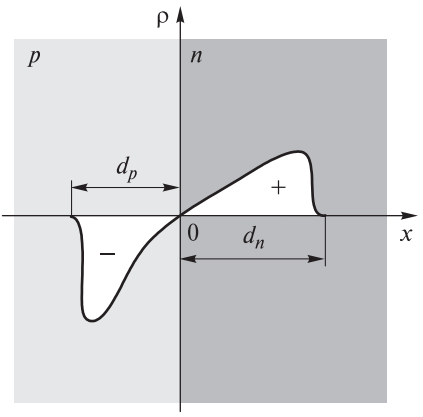


Рис. 1.32. Объемный заряд в плавном *p*–*n*-переходе

Считают, что концентрация равновесных основных носителей заряда в полупроводнике вне *p*–*n*-перехода равна концентрации примесей, т. е. примеси полностью ионизованы. Тогда равновесная концентрация электронов n_{no} в нейтральной части *n*-полупроводника равна N_D, а равновесная концентрация дырок p_{po} в нейтральной части *p*-области равна N_A:

$$n_{no} = N_D;\ p_{no} = N_A.$$

Зависимости концентраций основных и неосновных носителей заряда по обе стороны от резкого *p*–*n*-перехода и в самом резком *p*–*n*-переходе приведены на рис. 1.33.

Для равновесных концентраций всегда справедлив закон действующих масс, поэтому произведение концентраций основных и неосновных носителей заряда в обеих частях *p*–*n*-перехода всюду одинаково и равно n_i^2:

$$n_{no}\,p_{no} = p_{po}\,n_{po} = n_i^2.$$

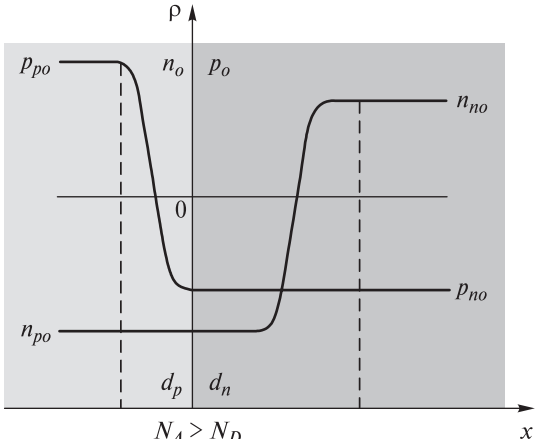


Рис. 1.33. Зависимости концентраций основных и неосновных носителей от расстояния в *p*–*n*-переходе ($N_A > N_D$):

n_{no}, p_{no} — соответственно концентрация основных электронов и неосновных (дырок) носителей заряда в *n*-области; p_{po}, n_{po} — соответственно концентрация основных (дырок) неосновных (электронов) носителей заряда в *p*-области

1.12.2. РАВНОВЕСНОЕ СОСТОЯНИЕ *p*–*n*-ПЕРЕХОДА

Концентрация дырок в *p*-области на несколько порядков превосходит концентрацию их в *n*-области, а концентрация электронов в *n*-области намного превосходит концентрацию электронов в *p*-области. Такое различие в концентрации однотипных носителей в контактирующих областях полупроводника приводит к возникновению диффузионных потоков электронов из *n*-области в *p*-область и диффузионного потока дырок из *p*-области в *n*-область. При этом область n, из которой диффундировали электроны, заряжается положительно, а область p, из которой диффундировали дырки, — отрицательно. Область *p*–*n*-перехода обеднена основными носителями заряда. Неосновные носители заряда рекомбинируют с основными. Поскольку область *p*–*n*-перехода обеднена основными носителями заряда, она будет обладать бóльшим сопротивлением, чем электронейтральные *p*–*n*-слои. В целом переход электронейтрален, так как положительный и отрицательный заряды в смежных слоях одинаковы. Поэтому наличие различий в концентрации примесей в смежных слоях приводит к различию в ширине областей, занимаемых пространственными зарядами. В слое с меньшей концентрацией примеси ширина области пространственного заряда больше.

Нескомпенсированные заряды ионов примесей вызывают появление электрического поля, направленного от положительного заряда к отрицательному, т. е. из слоя *n* в слой *p*. Это поле будет препятствовать дальнейшей диффузии. В равновесном состоянии диффузионные токи уравновешиваются дрейфовыми токами. Полный ток при этом через *p*–*n*-переход равен нулю. Возникновение электрического поля в *p*–*n*-переходе приводит к появлению разности потенциалов между смежными слоями, которая называется контактной разностью потенциалов.

Рассмотрим зонную схему для равновесного состояния *p*–*n*-перехода. Как было показано ранее, уровень Ферми является общим при контакте тел, находящихся в термодинамическим равновесии. Поэтому уровни Ферми в *n*- и *p*-областях должны находиться на одинаковом уровне, что вызывает искривление энергетических зон.

Зонные схемы *p*- и *n*-полупроводников до соприкосновения показаны на рис. 1.34. Зонная схема *p*–*n*-перехода в равновесном состоянии приведена на рис. 1.35.

Образующаяся в *p*–*n*-переходе контактная разность потенциалов $U_к$ создает в *p*–*n*-переходе потенциальный барьер $qU_к = \varphi_0$, препятствующий переходу электронов из *n*-области в *p*-область, а дырок — из *p*-области в *n*-область:

$$qU_к = E_{cp} - E_{cn} = E_{ip} - E_{in} = E_{vp} - E_{vn}; \tag{1.37}$$

$$qU_к = E_{ip} - E_{in} = \left(E_{ip} - E_F\right) + \left(E_F - E_{in}\right), \tag{1.38}$$

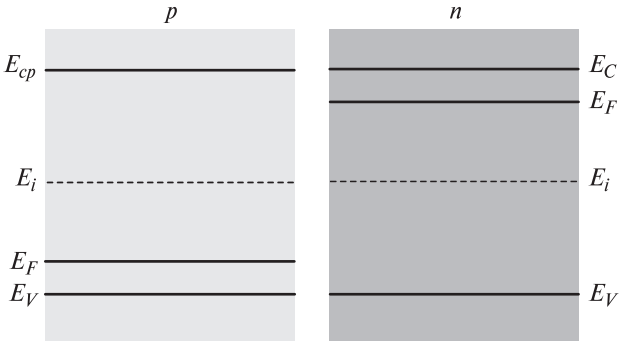


Рис. 1.34. Зонные диаграммы полупроводников *p*- и *n*-типов

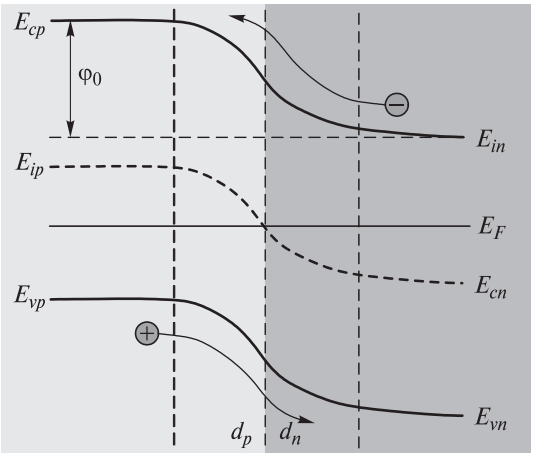


Рис. 1.35. Зонная диаграмма *p*–*n*-перехода

где E_{cp}, E_{cn} — уровень энергии дна зоны проводимости в полупроводнике соответственно *p*- и *n*-типов; E_{ip}, E_{in} — уровень энергии середины запрещенной зоны в полупроводник соответственно *p*- и *n*-типов; E_{vp}, E_{vn} — уровень энергии потока валентной зоны в полупроводнике соответственно *p*- и *n*-типов; E_F — уровень энергии Ферми.

Ход электростатического потенциала противоположен ходу зон (рис. 1.36):

$$\varphi = d\mathcal{E}/dx.$$

Из закона действующих масс следует

$$n_{no} = n_i \exp\left(\frac{E_F - E_{in}}{kT}\right); \qquad (1.39)$$

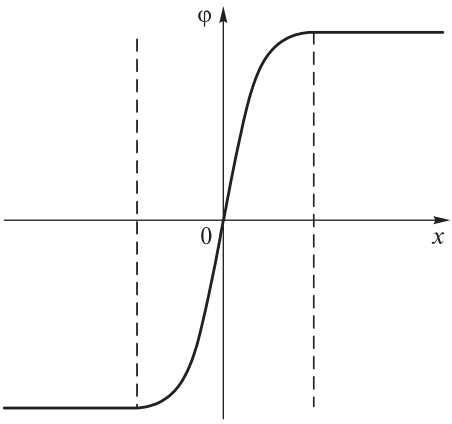


Рис. 1.36. Зависимость потенциала от расстояния в *p*–*n*-переходе

$$p_{po} = n_i \exp\left(-\frac{E_F - E_{ip}}{kT}\right). \qquad (1.40)$$

Из выражений (1.39) и (1.40) можно получить

$$E_F - E_{in} = kT \ln \frac{n_{no}}{n_i}; \qquad (1.41)$$

$$E_{ip} - E_F = kT \ln \frac{p_{po}}{p_i}. \qquad (1.42)$$

Подставим выражения (1.41) и (1.42) в соотношения (1.38):

$$qU_к = kT \ln \frac{n_{no} p_{po}}{n_i^2} = \\ = kT \ln \frac{n_{no}}{n_{po}} = kT \ln \frac{p_{po}}{p_{no}}. \qquad (1.43)$$

Следовательно, чем сильнее легированы области полупроводника, т. е. чем больше $n_{no} = N_D$ и $p_{po} = N_A$, тем больше контактная разность потенциалов.

Из соотношений (1.43) можно получить формулы, выражающие равновесные концентрации неосновных носителей заряда через равновесные концентрации основных носителей заряда в противоположных областях:

$$n_{po} = n_{no} e^{-\frac{qU_к}{kT}};$$

$$p_{no} = p_{po} e^{-\frac{qU_к}{kT}}.$$

1.12.3. ЗОННАЯ ДИАГРАММА *p*–*n*-ПЕРЕХОДА ПРИ ПРИЛОЖЕНИИ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Рассмотрим изменение зонной диаграммы при приложении к *p*–*n*-переходу внешнего напряжения. При этом следует учитывать, что «плюс» внешнего источника опускает уровни энергии в зонной диаграмме, «минус» — поднимает. При приложении внешнего электрического поля нарушается термодинамическое равновесие, и уровни Ферми заменяются на квазиуровни Ферми, т. е. уровни Ферми при определенном поле. Полем в толще полупроводника пренебрегаем, т. е. считаем, что энергетические зоны идут без наклона (горизонтально).

Это означает, что практически все внешнее напряжение падает на переходе, а падение напряжения на электронейтральных частях полупроводника

близко к нулю. На рис. 1.37 приведены зонные диаграммы *p*–*n*-перехода в равновесном состоянии (*а*), при прямом (*б*) и обратном (*в*) включении. При наложении внешнего электрического поля в прямом направлении изгиб зон в *p*–*n*-переходе уменьшается и становится равен

$$q(U_к - U),$$

где U — внешнее смещение в прямом направлении.

Квазиуровни Ферми для основных носителей заряда в *n*-области (E_{Fn}) и основных носителей в *p*-области (E_{Fp}) сдвигаются относительно друг друга на qU, т. е.

$$E_{Fn} - E_{Fp} = qU.$$

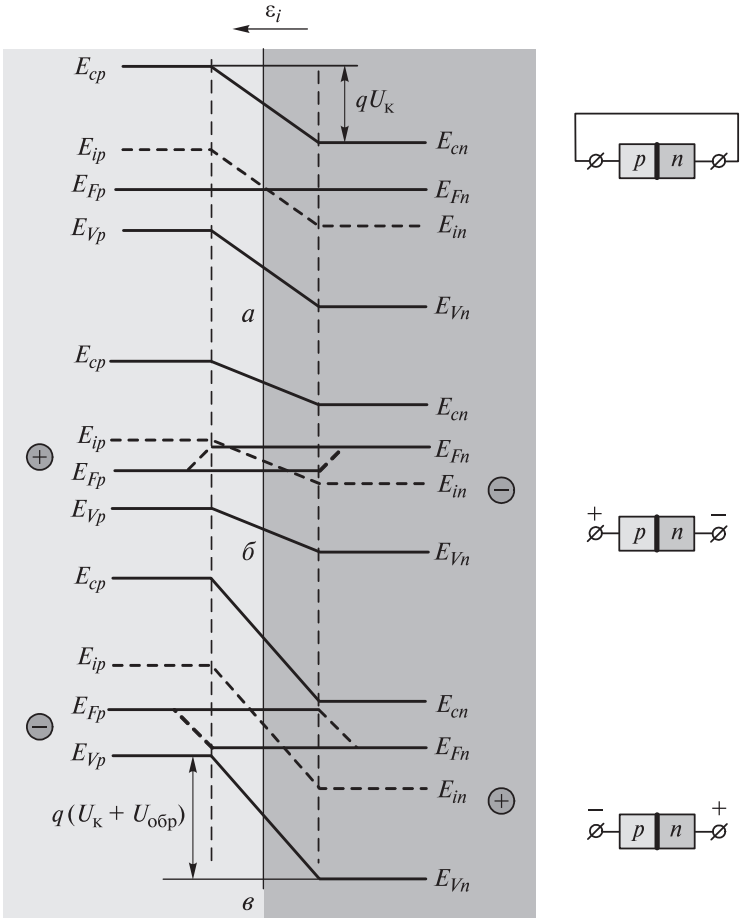


Рис. 1.37. Зонные диаграммы *p*–*n*-перехода:
а — в равновесном состоянии; *б* — при прямом смещении; *в* — при обратном смещении

Для тонкого *p*–*n*-перехода допустимо считать квазиуровень E_{Fn} неизменным во всей *n*-области, а также во всем переходе. Квазиуровень E_{Fp} можно считать одинаковым в *p*-области и во всем *p*–*n*-переходе. Штриховыми линиями обозначены на рис. 1.37 условно квазиуровни Ферми для неосновных носителей заряда в каждой из областей. Вдали от *p*–*n*-перехода они совпадают с квазиуровнями для основных носителей заряда, т. е. с уровнями Ферми для каждой из областей.

При наложении внешнего электрического поля в обратном направлении (см. рис. 1.37) изгиб зон на *p*–*n*-переходе увеличивается и составит

$$q\left(U_{\text{к}} - U_{\text{обр}}\right),$$

где $U_{\text{обр}}$ — часть внешнего напряжения, падающего на переходе.

1.12.4. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ТОНКОГО *p*–*n*-ПЕРЕХОДА

Тонким *p*–*n*-переходом называется электронно-дырочный переход, толщина которого столь мала, что можно пренебречь процессами рекомбинации и генерации носителей заряда в области объемного заряда *p*–*n*-перехода.

Если известна плотность потока дырок j_p или электронов j_n в каком-либо сечении *p*–*n*-перехода, то она такова же в любом другом сечении *p*–*n*-перехода.

Из формул (1.39) и (1.40) можно получить концентрацию электронов $n_{p\text{гр}}$ на границе запирающего слоя в *p*-области и концентрацию дырок $p_{n\text{гр}}$ на границе запирающего слоя в *n*-области:

$$n_{p\text{гр}} = n_i e^{\frac{E_{Fn} - E_{ip}}{kT}}; \tag{1.44}$$

$$p_{n\text{гр}} = n_i e^{-\frac{E_{Fp} - E_{in}}{kT}}, \tag{1.45}$$

при этом E_{ip} и E_{in} принимают значения, соответствующие границам.

Для неосновных носителей заряда вдали от *p*–*n*-перехода можно записать:

$$n_{po} = n_i e^{\frac{E_{Fp} - E_{ip}}{kT}}; \tag{1.46}$$

$$p_{no} = n_i e^{\frac{-E_{Fn} - E_{in}}{kT}}. \tag{1.47}$$

В выражениях (1.46) и (1.47) произошла замена квазиуровней по сравнению с формулами (1.44), (1.45), так как в электронейтральной части полупроводника вдали от *p*–*n*-перехода в *p*-области $E_{Fp} = E_{Fn}$, а в *n*-области $E_{Fn} = E_{Fp}$.

Разделим левые и правые части выражений (1.44) и (1.46) друг на друга:

$$\frac{n_{p\text{гр}}}{n_{po}} = e^{\frac{E_{Fn}-E_{Fp}}{kT}},$$

откуда

$$n_{p\text{гр}} = n_{po}e^{\frac{E_{Fn}-E_{Fp}}{kT}} = n_{po}e^{\frac{qU}{kT}}.$$

Разделим левые и правые части выражений (1.45) и (1.47) друг на друга:

$$\frac{p_{n\text{гр}}}{p_{no}} = e^{\frac{E_{Fn}-E_{Fp}}{kT}},$$

откуда

$$p_{n\text{гр}} = p_{no}e^{\frac{E_{Fn}-E_{Fp}}{kT}} = p_{no}e^{\frac{qU}{kT}}.$$

При $E_{Fn} - E_{Fp} > 0$, т. е. при прямом включении p–n-перехода, $n_{p\text{гр}} > n_{no}$, $p_{n\text{гр}} > p_{no}$.

Избыточная концентрация неосновных носителей заряда на границах в этом случае

$$(\Delta n)_o = n_{p\text{гр}} - n_{po} = n_{po}\left(e^{qU/kT} - 1\right); \tag{1.48}$$

$$(\Delta p)_o = p_{n\text{гр}} - p_{no} = p_{no}\left(e^{qU/kT} - 1\right). \tag{1.49}$$

Эта избыточная концентрация появляется вследствие инжекции носителей заряда через p–n-переход; Δn и Δp являются функциями:

$$\Delta n_p = f(x_p);$$
$$\Delta p_n = f(x_n).$$

Зависимость избыточной концентрации неосновных носителей от расстояния от области p–n-перехода приведена на рис. 1.38.

Электроны, инжектированные в p-область, притягивают к себе дырки из объема этой области, так что вне p–n-перехода сохраняется электронейтральность. Недостаток носителей заряда в объеме пополняется через контакт. Аналогично дырки, инжектированные в n-область, притягивают к себе электроны из объема этой области, где также сохраняется электронейтральность. Глубина проникновения инжектированных носителей определяется их рекомбинацией. Неосновные носители рекомбинируют с основными, поэтому концентрация неосновных носителей снижается с расстоянием:

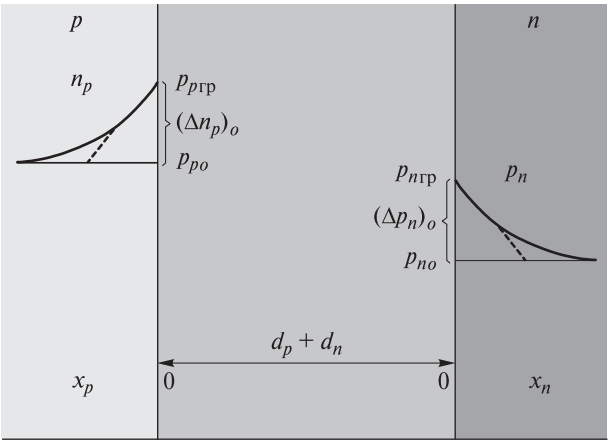


Рис. 1.38. Зависимость избыточной концентрации неосновных носителей от расстояния в области *p*–*n*-перехода

$$\Delta n_p = \left(\Delta n_p\right)_o e^{-x_p/L_n}; \tag{1.50}$$

$$\Delta p_n = \left(\Delta p_n\right)_o e^{-x_n/L_p}, \tag{1.51}$$

где L_n, L_p — расстояние от границы *p*–*n*-перехода, на котором концентрации неосновных носителей уменьшаются в *e* раз, называемое диффузионной длиной. Отметим, что при приложении прямого смещения запирающий слой сужается. При обратном смещении *p*–*n*-перехода он расширяется, и $E_{Fn} - E_{Fp} = qU < 0$.

Тогда для обратного включения $n_{p\text{гр}} < n_{po}$. В этом случае наблюдается проникновение неосновных носителей заряда через *p*–*n*-переход из областей прилегания к *p*–*n*-переходу. Выражения (1.48), (1.49) сохраняются и в этом случае. Однако избыточная концентрация здесь имеет отрицательный знак $\left(\Delta n_p\right)_o < 0;\ \left(\Delta p_n\right)_o < 0,$ т. е. в этом случае мы имеем дело не с избыточной, а с недостаточной концентрацией неосновных носителей. Следует отметить, что *p*–*n*-переход можно считать тонким, если $d < L_n$ и $d < L_p$, т. е. если толщина перехода меньше диффузионных длин носителей заряда.

Плотность диффузионного тока через границы при любой полярности внешнего напряжения:

$$\left(j_n\right)_{x_p=0} = qD_n\left(\frac{dn_p}{dx_p}\right)_{x_p=0} = qD_n\left[\frac{d\left(\Delta n_p\right)}{dx_p}\right]_{x_p=0}; \tag{1.52}$$

$$\left(j_p\right)_{x_n=0} = qD_p\left(\frac{dp_n}{dx_n}\right)_{x_n=0} = qD_p\left[\frac{d(\Delta p_n)}{dx_n}\right]_{x_n=0}. \qquad (1.53)$$

Подставим в выражения (1.52) и (1.53) выражения (1.50) и (1.51) и получим с учетом соотношений (1.48) и (1.49:

$$\left(j_n\right)_{x_p=0} = \frac{qD_n\left(\Delta n_p\right)_0}{L_n}\left[e^{-x_p/L_n}\right]_{x_p=0} = \frac{qD_n n_{p0}}{L_n}\left[e^{qU/kT}-1\right];$$

$$\left(j_p\right)_{x_n=0} = \frac{qD_p\left(\Delta p_n\right)_0}{L_p}\left[e^{-x_n/L_p}\right]_{x_n=0} = \frac{qD_p n_{n0}}{L_p}\left[e^{qU/kT}-1\right].$$

Поскольку предполагалось, что плотности электронов и дырок одинаковы в любом сечении p–n-перехода и на его границах, а также поскольку движение электронов и дырок противоположно друг другу, ВАХ p–n-перехода можно представить в виде

$$j = j_n + j_p = \left(\frac{qD_n n_{po}}{L_n} + \frac{qD_p p_{no}}{L_p}\right)\left(e^{qU/kT}-1\right). \qquad (1.54)$$

С учетом

$$L_n = \sqrt{D_n\tau_n};$$

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

$$j = \left(\frac{qL_n n_{po}}{\tau_n} + \frac{qL_p p_{no}}{\tau_p}\right)\left(e^{qU/kT}-1\right).$$

Проанализируем выражение (1.54).

При обратном смещении на p–n-переходе $\left(U<0\right)$ $e^{-qU/kT}\to 0,$ а скобка $\left(e^{-qU/kT}-1\right)\to -1,$ следовательно, обратный ток p–n-перехода стремится к

$$j_{\text{н}} = q\left(\frac{L_n n_{po}}{\tau_n} + \frac{L_p p_{no}}{\tau_p}\right) = \left(\frac{qD_n n_{po}}{L_n} + \frac{qD_p n_{no}}{L_p}\right). \qquad (1.55)$$

Эту плотность тока называют плотностью тока насыщения. Она достигается уже при $U \approx -0{,}1$ В; $j_{\text{н}}$ определяется концентрацией неосновных носителей и у германиевых p–n-переходов больше, чем у кремниевых.

При прямых смещениях плотность тока через p–n-переход растет по экспоненте и уже при небольших напряжениях достигает большого значения.

Подставляя в выражение (1.55) соотношение (1.54), получим

$$j = j_{\text{н}}\left(e^{qU/kT}-1\right).$$

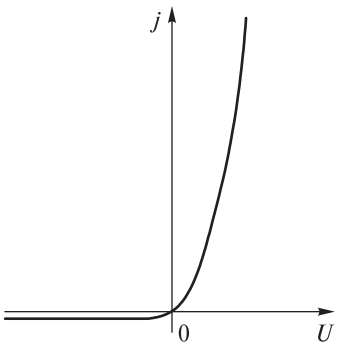


Рис. 1.39. Вид ВАХ *p–n*-перехода

На рис. 1.39 показана ВАХ *p–n*-перехода. Как видно из рисунка, *p–n*-переход обладает ярко выраженной односторонней (униполярной) проводимостью, т. е. проявляет высокие выпрямляющие свойства.

Следует отметить, что перенос заряда через *p–n*-переход осуществляется с помощью неосновных носителей, так как область *p–n*-перехода обеднена основными носителями.

Тесты к лекции 1.12

1. ***Что такое электронно-дырочный переход (p–n-переход)?***

а) слой полупроводника, расположенный по обе стороны от границы раздела *p*- и *n*-областей;

б) интерфейс для подключения разъема к полупроводнику *p*- или *n*-типа;

в) металлургическая граница в полупроводнике, определяющая ионную проводимость.

2. ***Что такое резкий p–n-переход?***

а) скачкообразное изменение концентрации акцепторов и доноров на границе *p*- и *n*-областей;

б) скачкообразное изменение тока на границе *p*- и *n*-областей;

в) скачкообразное изменение напряжения на границе *p*- и *n*-областей.

3. ***Что такое плавный p–n-переход?***

а) концентрация акцепторов и доноров является линейной функцией расстояния;

б) ток плавно изменяется на границе *p*- и *n*-областей;

в) напряжение плавно изменяется на границе *p*- и *n*-областей.

4. ***Что такое контактная разность потенциалов p–n-перехода?***

а) потенциальный барьер между *p*- и *n*-областями;

б) напряжение источника питания, прикладываемое к *p–n*-переходу;

в) разность потенциалов, возникающая на *p–n*-переходе при подключении источника питания.

5. ***Как зависит контактная разность потенциалов p–n-перехода от степени легирования областей полупроводника?***

а) чем сильнее легированы области полупроводника, тем больше контактная разность потенциалов;

б) чем меньше легированы области полупроводника, тем больше контактная разность потенциалов;

в) никак не зависит.

6. ***Чем обусловлен обратный ток p–n-перехода?***

а) неосновными носителями;

б) основными носителями;

в) ионами.

7. При каком примерно обратном напряжении возникает ток насыщения в p–n-переходе?

а) –0,1 В;

б) –1 В;

в) –10 В.

1.13. ЕМКОСТЬ *p–n*-ПЕРЕХОДА

Цель лекции: ознакомление с понятием емкости *p–n*-перехода, механизмами пробоя, биполярными транзисторами и гетероструктурами.

1.13.1. БАРЬЕРНАЯ И ДИФФУЗИОННАЯ ЕМКОСТЬ *p–n*-ПЕРЕХОДА

Значения объемных зарядов в *p–n*-переходе изменяются в зависимости от прикладываемого обратного напряжения U вследствие изменения толщины обедненного слоя. Следовательно, *p–n*-переход обладает электрической емкостью.

Барьерная емкость (зарядовая) *p–n*-перехода

$$C_{\text{бар}} = \frac{dQ_{\text{об}}}{dU} = \frac{\varepsilon_0 \varepsilon_n S}{d(U)},$$

где $Q_{\text{об}}$ — заряд обедненного слоя *p–n*-перехода; ε_0 — диэлектрическая постоянная; ε_n — относительная диэлектрическая проницаемость кристалла полупроводника *n*-типа; S – площадь *p–n*-перехода.

Барьерная емкость совпадает с емкостью плоского конденсатора, расстояние между обкладками которого равно толщине обедненного слоя, с ростом обратного смещения уменьшается вследствие увеличения $d(U)$. У *p–n*-переходов с большей концентрацией примеси барьерная емкость больше.

Вольт-фарадная характеристика (ВФХ) $C_{\text{бар}} = f(U)$ (рис. 1.40) имеет достаточно сложный вид, поэтому применяют аппроксимацию:

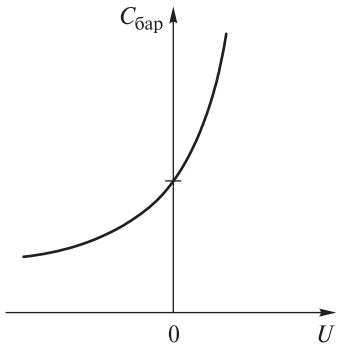


Рис. 1.40. Зависимость барьерной емкости от напряжения

$$C_{\text{бар}} = \frac{C_{\text{бар}(0)}}{(1 - U/U_{\text{к}})^m},$$

где m = 0,3...0,5; $U_{\text{к}}$ — контактная разность потенциалов.

Барьерная емкость слабо возрастает с увеличением температуры за счет $U_{\text{к}}$. Температурный коэффициент емкости уменьшается с ростом обратного напряжения.

При прямом смещении происходит инжекция неосновных носителей в *p*- и *n*-областях. При изменении прикладываемого напряжения изменяется и концентрация инжектированных носителей, а следовательно, и заряды областей,

что можно рассматривать как действие емкости. Так как эта емкость появляется за счет диффузионной составляющей тока, то ее называют диффузионной. Если толщина базы $W \gg L_p$, то диффузная емкость

$$C_D = \frac{1}{2}\left(\frac{q}{kT}\right) I \tau_p,$$

где τ_p — среднее время диффузии носителей в p-области.

При $W \ll L_p$

$$C_D = \left(\frac{q}{kT}\right) I \tau_D,$$

где τ_D — среднее время диффузии носителей через n-область, $\tau_D = \frac{W^2}{2D_p}$; D_p — коэффициент диффузии дырок.

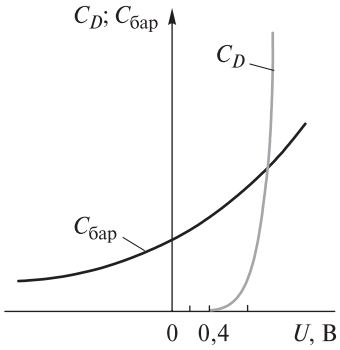


Рис. 1.41. Зависимости барьерной и диффузионной емкостей от напряжения

При больших прямых смещениях, когда ток инжекции значительно превышает ток рекомбинации, C_D экспоненциально возрастает с ростом напряжения и значительно превышает $C_{бар}$ (рис. 1.41). При малых напряжениях диффузионная емкость меньше барьерной.

1.13.2. МЕХАНИЗМЫ ПРОБОЯ *p–n*-ПЕРЕХОДОВ

При достаточно большом обратном смещении p–n-переход «пробивается» и через него протекает очень большой ток. Различают три основных механизма пробоя: туннельный эффект, лавинное умножение, тепловая неустойчивость.

Туннельный пробой. В основе данного вида пробоя лежит туннельный эффект, т. е. проникновение электронов сквозь потенциальный барьер, когда его толщина достаточно мала (рис. 1.42). Когда электрическое поле в германиевых или кремниевых p–n-переходах достигает ~10^6 В/см, начинают протекать токи, обусловленные туннельными переходами между зонами. Для получения таких полей концентрации примеси в p- и n-областях должны быть достаточно высокими. Туннельный пробой происходит в германиевых и кремниевых p–n-переходах при напряжениях пробоя, меньших $4E_g/q$. При напряжениях пробоя $(4...6)E_g/q$ пробой обусловлен как туннельным, так и лавинным механизмами. При напряжениях пробоя бóльших $6E_g/q$ пробой связан с лавинным механизмом. Поскольку ширина запрещенной зоны Ge, Si, GaAs уменьшается с увеличением температуры, напряжение пробоя, связанное с туннельным эффектом, имеет отрицательный температурный коэффициент.

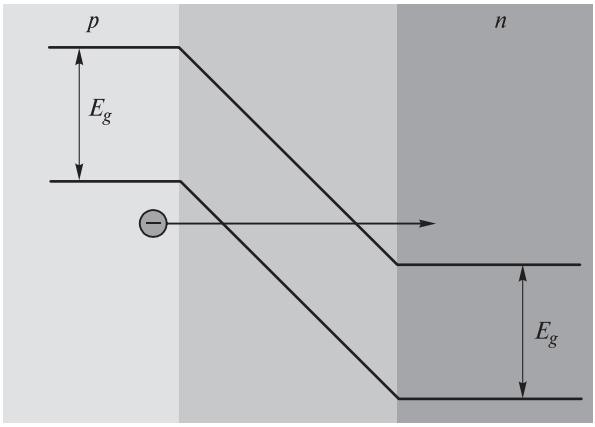


Рис. 1.42. Схема межзонного туннелирования

Лавинное умножение. Напряжение лавинного пробоя определяет верхний предел обратного напряжения диодов и коллекторного напряжения биполярного транзистора. Процесс лавинного пробоя полупроводника аналогичен ударной ионизации в газе. Электрон или дырка, ускоренные полем, могут ионизировать атом полупроводника, в результате чего получится новая пара электрон–дырка, а ток через *p–n*-переход возрастает. При достаточно большой напряженности электрического поля ионизация может приобрести лавинный характер. При лавинном пробое пробивное напряжение с ростом температуры увеличивается.

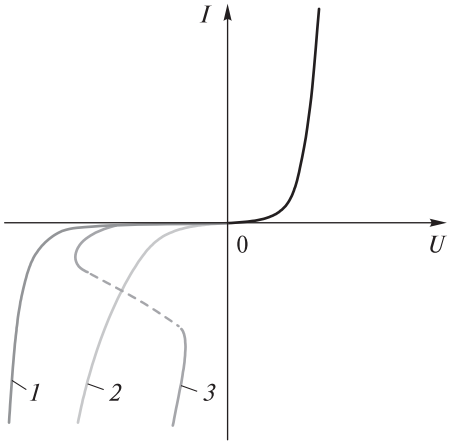


Рис. 1.43. Вид ВАХ кремниевых *p–n*-переходов:
1 — лавинный пробой; *2* — туннельный пробой; *3* — тепловой пробой

Для резкого несимметричного *p–n*-перехода напряжение пробоя

$$U_{\text{пр}} = \frac{\varepsilon_{\text{пр}} W}{2} = \frac{\varepsilon_n \varepsilon_0 \varepsilon_m^2}{2qN_D},$$

где N_D — концентрация примеси в слаболегированной области; ε_m — максимальное электрическое поле.

Тепловой пробой. Тепловой механизм пробоя обусловлен выделением теплоты в *p–n*-переходе при протекании обратного тока $I_{\text{об}}$. Рассеиваемая в переходе мощность $P = UI_{\text{об}}$. Под действием этой мощности увеличивается температура перехода, а следовательно, растет и обратный ток. Вследствие теплового нагрева на обратной ветви ВАХ появляется участок с отрицательным дифференциальным сопротивлением, ток резко возрастает, и диод выходит из строя. На рис. 1.43 показаны ВАХ кремниевых *p–n*-переходов с лавинным, туннельным и тепловым пробоями.

Тесты к лекции 1.13

1. ***Как барьерная емкость зависит от приложенного к p–n-переходу напряжения?***

а) чем больше обратное напряжение, тем барьерная емкость меньше;

б) чем меньше обратное напряжение, тем барьерная емкость меньше;

в) чем больше прямое напряжение, тем барьерная емкость меньше.

2. ***Как диффузионная емкость зависит от приложенного к p–n-переходу напряжения?***

а) чем больше прямое напряжение, тем диффузионная емкость больше;

б) чем больше прямое напряжение, тем диффузионная емкость меньше;

в) никак не зависит.

3. ***Что такое туннельный пробой p–n-перехода?***

а) электрический пробой *p–n*-перехода, вызванный туннельным эффектом;

б) тепловой пробой *p–n*-перехода;

в) электромагнитный пробой *p–n*-перехода.

4. ***Что такое лавинный пробой p–n-перехода?***

а) электрический пробой *p–n*-перехода, вызванный лавинным размножением носителей заряда под действием сильного электрического поля;

б) пробой *p–n*-перехода, вызванный сильным нагревом;

в) пробой *p–n*-перехода, вызванный большим прямым приложенным напряжением.

5. ***Что такое тепловой пробой p–n-перехода?***

а) пробой, обусловленный выделением теплоты в *p–n*-переходе при протекании обратного тока в результате приложения большого обратного напряжения;

б) пробой, обусловленный поглощением теплоты в *p–n*-переходе при протекании прямого тока;

в) пробой, обусловленный лавинным размножением фононов в полупроводнике.

1.14. БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

Цель лекции: ознакомление с физическими основами реализации полупроводниковых транзисторов.

1.14.1. ОБЩИЕ СВЕДЕНИЯ

Полупроводниковый триод, получивший впоследствии название транзистор, разработали американские ученые Дж. Бардин и В. Браттейн в 1948 г. Создание этого полупроводникового прибора имело огромное значение для развития электронной и вычислительной техники. Для работы транзистора требуются меньшие напряжения, чем для радиоламп, транзисторы имеют существенно меньшие габаритные размеры, отличаются более высокой надежностью, стоимость их производства значительно ниже. В настоящее время биполярные транзисторы являются базовыми элементами во многих усилительных схемах.

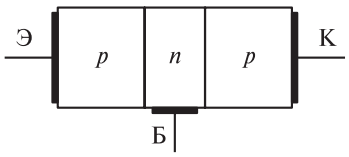


Рис. 1.44. Схема биполярного транзистора *p*–*n*–*p*-типа: Э —эмиттер; Б — база; К — коллектор

Биполярным транзистором называется полупроводниковый прибор, имеющий трехслойную структуру и состоящий из двух *p*–*n*-переходов, ток переносится двумя типами носителей заряда: электронами и дырками. Слои такой структуры называются эмиттером, базой и коллектором. Промышленность выпускает два типа биполярных транзисторов: *p*–*n*–*p*-типа и *n*–*p*–*n*-типа. На рис. 1.44 представлена схема биполярного транзистора *p*—*n*–*p*-типа.

Назначение эмиттера — инжектирование в слой базы носителей заряда. В коллектор попадают носители заряда, прошедшие из эмиттера через базу. Конструктивно слой базы выполняют намного тоньше, чем остальные слои, чтобы носители заряда, прошедшие через базу, не успевали рекомбинировать и достигали коллектора. Переход, образующийся между базой и эмиттером, получил название эмиттерного перехода, между базой и коллектором — коллекторного.

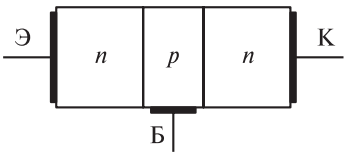


Рис. 1.45. Схема биполярного транзистора *n*–*p*–*n*-типа: Э — эмиттер; Б — база; К — коллектор

Схема биполярного транзистора *n*–*p*–*n*-типа приведена на рис. 1.45. Эмиттер изготовляют с высоким содержанием примеси, чтобы в нем было большое количество основных носителей заряда, благодаря чему становится возможна инжекция этих носителей в слой базы.

Основная идея работы транзистора заключается в том, что, управляя небольшим током базы, можно управлять большим током, протекающим через коллектор.

Транзисторы выпускают как в пластмассовых, так и в металлических корпусах, но все они имеют три вывода: эмиттер, базу и коллектор. База — управляющий вывод транзистора, коллектор для *n*–*p*–*n*-транзисторов находится под положительным потенциалом, для *p*–*n*–*p*-транзисторов — под отрицательным.

Эмиттер для *n*–*p*–*n*-транзисторов находится под отрицательным потенциалом, для *p*–*n*–*p*-транзисторов — под положительным.

1.14.2. ПРИНЦИП ДЕЙСТВИЯ БИПОЛЯРНОГО ТРАНЗИСТОРА

Каждый переход транзистора можно включить либо в прямом, либо в обратном направлении. Поэтому возможны следующие режимы работы транзистора.

1. Режим отсечки — оба *p*–*n*-перехода закрыты. В этом случае между эмиттером и коллектором идет относительно небольшой ток. Управление транзистором в этом режиме невозможно.

2. Режим насыщения — оба *p*–*n*-перехода открыты и через транзистор идет сравнительно большой ток. Управление транзистором в этом режиме также невозможно.

3. Активный режим — один из *p*–*n*-переходов открыт, а другой закрыт. В этом режиме возможно эффективное управление транзистором.

Различают два варианта включения транзистора: нормальное и инверсное. При нормальном включении на эмиттерном переходе напряжение подключается в прямом направлении, а на коллекторном — в обратном. При инверсном включении полярность обратная.

В активном режиме работы эмиттерный переход транзистора смещен в прямом направлении (открыт), а коллекторный переход — в обратном. Рассмотрим *n*–*p*–*n*-транзистор (рис. 1.46), все рассуждения будут аналогичны для случая *p*–*n*–*p*-транзистора, с заменой слова «электроны» на «дырки», а «дырки» на «электроны», а также с заменой всех напряжений на противоположные по знаку.

Предположим, что транзистор *n*–*p*–*n*-типа (см. рис. 1.46) работает в активном режиме без нагрузки (статический режим). Напряжения на эмиттерном и коллекторном переходах задаются внешними источниками постоянного напряжения: $E_Б$ и $E_К$ соответственно. Их полярность и значения напряжения обеспечивают смещение эмиттерного перехода в прямом направлении, а коллекторного перехода — в обратном. В активном режиме работы $E_Б$ составляет десятые доли вольта, $E_К$ — единицы или даже сотни вольт, в зависимости от марки транзистора. Потенциал базы меньше потенциала коллектора, следовательно, коллекторный переход смещен в обратном направлении, при этом сопротивление эмиттерного перехода мало́, а коллекторного перехода велико́.

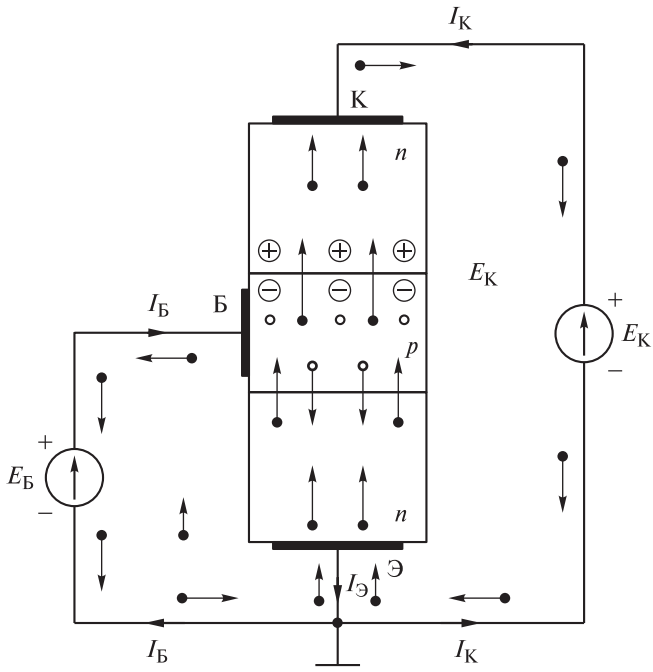


Рис. 1.46. Схема работы транзистора *n*–*p*–*n*-типа

Поскольку эмиттерный переход смещен в прямом направлении, потенциальный барьер перехода понижен, поэтому электроны его легко преодолевают и инжектируются из эмиттера в базу. Небольшая часть электронов (несколько процентов) в базе успевает рекомбинировать с дырками и в результате возникает сравнительно небольшой базовый ток $I_Б$ (дырок в базе мало, так как толщина базы мала), а бóльшая часть оставшихся электронов достигает коллекторного перехода. Поскольку коллекторный переход смещен в обратном направлении, на этом переходе образуется область пространственных зарядов (подобно обычному *p*–*n*-переходу при обратном напряжении). На рис. 1.46 эта область показана в виде кружочков со знаками «плюс» и «минус». Направление электрического поля в области пространственного заряда (от «плюса» к «минусу») способствует продвижению электронов из эмиттера через коллекторный переход. Пройдя коллекторный переход, электроны стремятся через коллектор к положительному полюсу источника питания. Эти электроны и создают коллекторный ток. Ток коллектора $I_К$ меньше тока эмиттера $I_Э$ на величину тока базы $I_Б$. В соответствии с законом Кирхгофа между токами всегда справедливо соотношение: $I_Э = I_К + I_Б$, так как $I_Б << I_Э$, то $I_Э \approx I_К$.

Поскольку базу выполняют очень тонкой, уменьшая тем самым концентрацию примесей (дырок), ток базы $I_Б$ будет небольшим.

Коэффициент α, связывающий ток эмиттера и ток коллектора ($I_К = \alpha I_Э$), называется коэффициентом передачи тока эмиттера. Численное значение

коэффициента $\alpha = 0{,}9...0{,}999$, чем больше коэффициент α, тем лучше транзистор. Этот коэффициент мало зависит от напряжения коллектор — база и база — эмиттер.

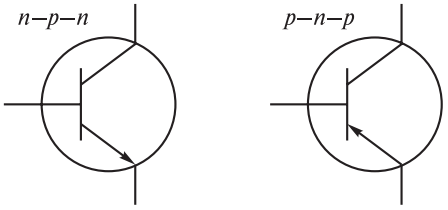


Рис. 1.47. Условные обозначения транзисторов:
а — транзистор *n–p–n*; *б* — транзистор *p–n–p*

Вследствие обратного смещения коллекторного перехода возникает обратный ток коллектора, который состоит из неосновных носителей заряда коллектора. Значение обратного тока коллектора определяется процессами термогенерации неосновных носителей, т. е. зависит от температуры коллектора. Обратный ток коллектора также называют тепловым током.

Качество биполярного транзистора значительно зависит от теплового тока: чем он меньше, тем транзистор качественнее.

При увеличении прямого входного напряжения между базой и эмиттером понижается потенциальный барьер в эмиттерном переходе и, соответственно, возрастает ток через этот переход, при уменьшении напряжения ток уменьшается. Таким образом, изменяя ток в цепи база — эмиттер, можно управлять выходным током коллектора. Причем незначительное изменение тока базы приводит к значительному изменению коллекторного тока. На рис. 1.47 представлены условные обозначения транзисторов на принципиальных схемах.

1.14.3. СХЕМЫ ВКЛЮЧЕНИЯ БИПОЛЯРНЫХ ТРАНЗИСТОРОВ

Транзистор имеет три вывода и для его нормальной работы требуется два источника питания. Один вывод транзистора будет подключен к двум источникам питания одновременно, отсюда следует, что существуют три возможных варианта подключения биполярного транзистора: ОБ — с общей базой, ОЭ — с общим эмиттером, ОК — с общим коллектором. Каждый вариант подключения обладает как преимуществами, так и недостатками. Соответствующий вариант подключения выбирают в зависимости от области применения и требуемых характеристик. Название варианта включения определяется по выводу (электроду) транзистора, который является общим для входной и выходной цепей по переменному току. Для каждой из схем включения транзистора имеются свои характеристики. Наиболее распространены ВАХ для схем ОЭ и ОБ, которые обычно приводятся в справочниках.

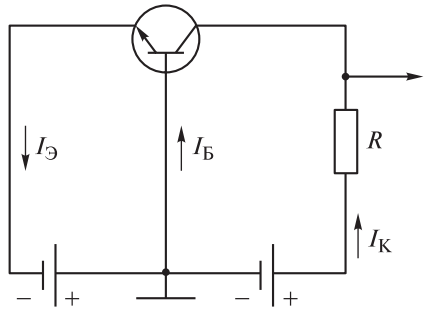


Рис. 1.48. Схема включения с ОБ

Схема включения с ОБ. Данный вариант подключения (рис. 1.48) обеспечивает хорошую частотную характеристику, одна-

ко не позволяет получить такого большого усиления сигнала по напряжению, как при включении ОЭ.

Усиления по току не возникает, поэтому данную схему включения иногда называют «токовый повторитель» из-за того, что она имеет свойство стабилизации тока. В схеме с общей базой не происходит изменение фазы выходного сигнала относительно входного.

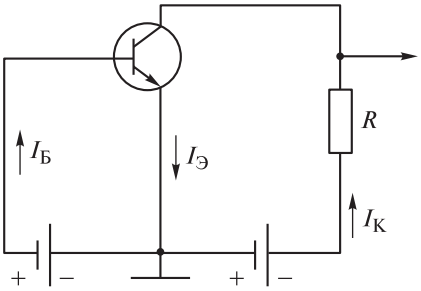


Рис. 1.49. Схема включения с ОЭ

Схема включения с ОЭ. Данный вариант подключения (рис. 1.49) позволяет получить максимальное усиление как по току, так и по напряжению, а следовательно, и по мощности. Особенностью данного варианта подключения является смещение по фазе выходного переменного напряжения на 180° относительно входного. Основной недостаток данной схемы — это низкая частотная характеристика, что накладывает определенные ограничения на использование при высокочастотном входном сигнале.

В случае напряжения между коллектором и эмиттером $U_{КЭ} = 0$ входная ВАХ p–n–p-транзистора (рис. 1.50, *а*) представляет собой прямую ветвь ВАХ двух параллельно включенных диодов. При $U_{КЭ} < 0$ кривые на графике сдвигаются вправо и вниз по сравнению с предыдущим случаем. Смещение кривых вниз обусловлено протеканием навстречу вносимому базовому току обратного тока $I_{К0}$, и при $U_{БЭ} = 0$ значение $I_Б < 0$. Смещение кривых вправо вызвано эффектом модуляции базы, а именно: с ростом $U_{КЭ}$ растет напряже-

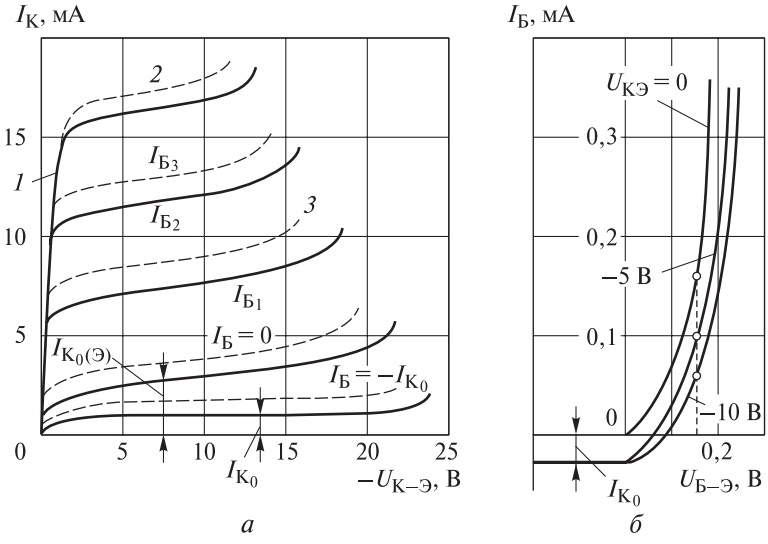


Рис. 1.50. Входная (*а*) и выходная (*б*) ВАХ p–n–p-транзистора с общим эмиттером

ние между коллектором и базой $U_{КБ}$, поэтому толщина обратно смещенного коллекторного перехода увеличивается, а толщина базы уменьшается. В результате в базе рекомбинирует меньше носителей, и ток базы $I_Б$ уменьшается.

Рассмотрим выходные характеристики (рис. 1.50, *б*) режима отсечки, когда $U_{БЭ} < 0$ соответствует самая нижняя кривая. При этом эмиттерный и коллекторный переходы заперты ($I_Б = -I_{К0}$) и из коллектора в базу течет тепловой ток $I_{К0}$. Вторая снизу кривая, соответствующая $I_Б = 0$, равносильна режиму с «оторванной» базой, т. е. $I_К = \beta I_Б + I_{К0(Э)} = I_{К0(Э)} = (1+\beta)I_{К0}$. В этом случае из коллектора в эмиттер течет сквозной ток $I_{К0(Э)}$. На графике выходных характеристик можно выделить три области.

1. Режим насыщения. При малых значениях напряжения $U_{КЭ} > 0$ и $U_{БЭ} > 0$ коллекторный переход смещен в прямом направлении. При этом наблюдается нелинейная область с сильной зависимостью $I_К$ от $U_{КЭ}$.

2. Активный режим. В этом режиме коллекторный переход смещен в обратном направлении, наблюдается слабая зависимость $I_К$ от $U_{КЭ}$. Это так называемая линейная область. Некоторый подъем кривых объясняется эффектом модуляции базы. При увеличении $U_{КЭ}$ происходит уменьшение толщины базы, которое приводит к уменьшению тока $I_Б$, но поскольку ток базы необходимо поддерживать постоянным ($I_Б = \text{const}$), приходится увеличивать $U_{БЭ}$. За счет этого ток $I_Б$ увеличивается, и ток $I_К = I_Б\beta$ также возрастает.

3. Область пробоя коллекторного перехода. В подавляющем большинстве случаев это нерабочая область за исключением транзисторов специальных типов.

Схема включения с ОК. Данный вариант подключения (рис. 1.51) имеет усиление по току почти такое же, как при подключении с ОЭ, а усиление по напряжению немного меньше единицы. Смещение напряжения по фазе отсутствует, поэтому данную схему называют «эмиттерный повторитель», так как напряжение на выходе соответствуют входному напряжению.

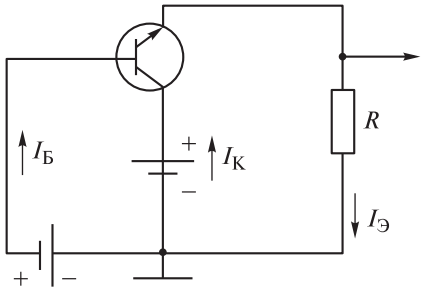


Рис. 1.51. Схема включения с ОК

Вольт-амперные характеристики биполярного транзистора в активном режиме. Выходные характеристики биполярного транзистора в активном режиме, включенного по схеме с ОБ, представлены на рис. 1.52. Видно, что кривая зависимости тока эмиттера от приложенного напряжения между эмиттером и базой значительно зависит от этого напряжения (рис. 1.52, *а*). На рис. 1.52, *б* видно, что выходные ВАХ имеют две характерные области: левая часть графика — сильная зависимость $I_К$ от $U_{КБ}$; далее горизонтальный участок с небольшим подъемом кривых — слабая зависимость $I_К$ от $U_{КБ}$.

Семейства входных и выходных ВАХ биполярных транзисторов используются для оценки их работы в широком диапазоне импульсных и постоянных токов, мощностей и напряжений, а также для расчета цепи смещения и стабилизации режима.

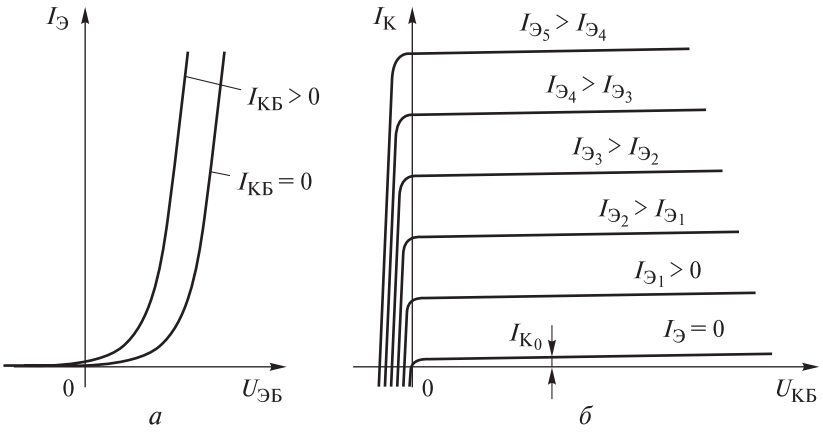


Рис. 1.52. Выходные характеристики биполярного транзистора в активном режиме, включенного по схеме с ОБ ($I_{Э_1}...I_{Э_5}$ — соответственно первое — пятое значение тока эмиттера)

1.14.4. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называется переход, образующийся на границе двух полупроводников с различной шириной запрещенной зоны. На границе гетероперехода происходят изменения свойств полупроводника — изменяются структура энергетических зон, ширина запрещенной зоны, подвижности и эффективные массы.

На возможность использования специфических свойств такого контакта для повышения эффективности инжекции эмиттеров биполярных транзисторов указывалось в патенте Шокли в 1948 г. В настоящее время гетеропереходы находят столь широкое применение при создании высокочастотных транзисторов и оптоэлектронных приборов, что работы российского ученого Ж.И. Алферова и Г. Кремера (США) в области исследования гетеропереходов были отмечены присуждением им совместно с изобретателем интегральной схемы Дж. Килби Нобелевской премии по физике в 2000 г.

Различают анизотропные и изотропные гетеропереходы. Анизотропные переходы создаются в результате контакта полупроводников с дырочной и электронными типами проводимости. Изотропные гетеропереходы возникают при контакте полупроводников одного типа проводимости.

В гетеропереходах наблюдаются скачки энергии в зонах проводимости и валентной зоне. Гетероструктуры представляют собой комбинации различных гетеропереходов.

Вследствие разрывов зон высота потенциальных барьеров, которые при инжекции преодолевают электроны и дырки, становится различной, что приводит к изменению соотношения токов инжектируемых электронов и дырок. Отношение тока инжекции электронов к току инжекции дырок существенно возрастает. Наличие разрывов в положении краев энергетических зон при-

водит к тому, что в гетеропереходах происходит практически односторонняя инжекция из широкозонного в узкозонный полупроводник.

Если из-за разрыва в положении края зоны инжектируемые носители при переходе в узкозонный полупроводник увеличивают свою кинетическую энергию, то при больших прямых смещениях концентрация носителей, инжектированных в узкозонный полупроводник, может даже превышать концентрацию свободных носителей в широкозонном слое. Данные свойства гетероперехода широко используются в полупроводниковых лазерах, некоторых транзисторах и других приборах, принцип работы которых требует однонаправленной инжекции носителей.

Различие параметров кристаллической решетки контактирующих материалов приводит к появлению на гетерогранице оборванных связей, дислокаций несоответствия и других дефектов, энергетические уровни которых выступают как центры рекомбинации. Поэтому задача выбора материалов для создания гетеропереходов оказывается достаточно сложной. Необходимо, чтобы различие параметров решетки двух контактирующих полупроводников при температуре выращивания было невелико при близких температурных коэффициентах линейного расширения.

Рассмотрим зонную диаграмму между арсенидом галлия и арсенидом галлия алюминия GaAs–Al_xGa_xAs (рис. 1.53). Величина x характеризует содержание алюминия. С ростом x увеличивается ширина запрещенной зоны твердого раствора Al_xGa_xAs. Для $x = 0{,}3$ ширина запрещенной зоны твердого раствора $Al_{0,3}Ga_{0,7}As$ равна 1,8 эВ.

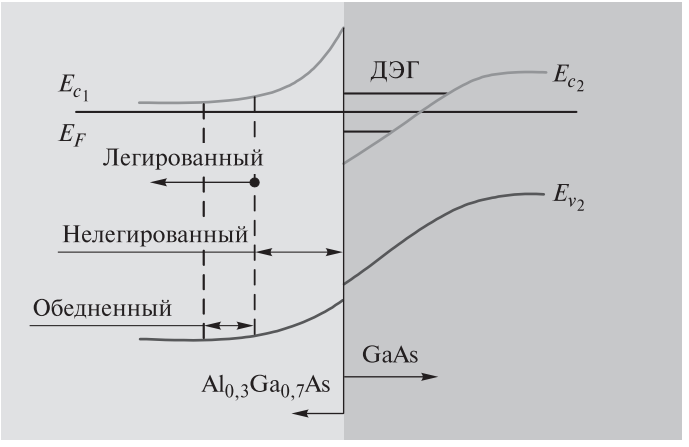


Рис. 1.53. Зонная диаграмма гетероперехода

У границы раздела двух полупроводников образуется квантовый колодец, или зона двумерного электронного газа (ДЭГ) (2D-газ). Двумерный электронный газ представляет собой систему электронов, энергетические уровни которых дискретны. Они могут перемещаться только в плоскости ДЭГ, не покидая потенциальной ямы. Характерным свойством ДЭГ является то, что

можно регулировать плотности электронов путем воздействия поперечного электрического поля. Электроны ДЭГ имеют повышенную эффективную концентрацию и подвижность.

При температуре 300 К для гетеропереходной пары GaAs–$Al_{0,3}Ga_{0,7}As$ величина развыва зоны проводимости и валентной зоны составит соответственно

$$\Delta E_c \approx 0,62\Delta E_g \approx 0,23 \text{ эВ};$$

$$\Delta E_v \approx 0,38\Delta E_g \approx 0,14 \text{ эВ}.$$

В этом случае отношения токов инжектируемых электронов и дырок составят $\exp\left(\Delta E_g / kT\right) \approx 2 \cdot 10^6$. Столь значительный эффект дает основание говорить о действительно одностороннем характере инжекции в гетеропереходе.

Тесты к лекции 1.14

1. ***Что такое гетеропереход?***

а) переход, образующийся на границе двух полупроводников с различной шириной запрещенной зоны;

б) переход, образующийся на границе двух полупроводников с одинаковой шириной запрещенной зоны;

в) переход, образующийся на границе соприкосновения металла и полупроводника.

2. ***Что такое анизотропный гетеропереход?***

а) гетеропереход, который получается в результате контакта полупроводников с дырочной и электронными типами проводимости;

б) гетеропереход, который получается в результате контакта полупроводников с одинаковым типом проводимости;

в) гетеропереход, который получается в результате контакта двух различных металлов.

3. ***Что такое изотропный гетеропереход?***

а) гетеропереход, возникающий при контакте полупроводников одного типа проводимости;

б) гетеропереход, возникающий при контакте полупроводников разного типа проводимости;

в) гетеропереход, возникающий при контакте полупроводников с благородными металлами.

4. ***Назовите схемы включения биполярного транзистора:***

а) с общей базой, общим эмиттером и общим коллектором;

б) с параллельной базой, общим эмиттером и общим коллектором;

в) с общей базой, общим эмиттером и последовательным коллектором.

5. ***Что представляет собой зона двумерного электронного газа в гетеропереходе?***

а) систему электронов, энергетические уровни которых дискретны, они могут перемещаться только в плоскости ДЭГ, не покидая потенциальной ямы;

б) облако ионов, энергетические уровни которых непрерывны, они могут перемещаться не только в плоскости ДЭГ, но и покидать потенциальную яму;

в) систему электронов, энергетические уровни которых непрерывны, они не могут перемещаться в плоскости ДЭГ, а покидают потенциальную яму.

6. ***Какой прибор называется биполярным транзистором?***

а) полупроводниковый прибор, имеющий трехслойную структуру и состоящий из двух *p*–*n*-переходов, в котором ток переносится двумя типами носителей заряда — электронами и дырками;

б) полупроводниковый прибор, состоящий из трех *p*–*n*-переходов, в котором ток переносится двумя типами носителей заряда — электронами и дырками;

в) полупроводниковый прибор, имеющий четырехслойную структуру и состоящий из четырех *p*–*n*-переходов, в котором ток переносится двумя типами носителей заряда — электронами и дырками.

7. ***Каково назначение эмиттера в биполярном транзисторе?***

а) инжектирование в слой базы носителей заряда;

б) усиление тока в слое базы;

в) управление током в слое базы.

1.15. ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ

Цель лекции: ознакомление с понятием поверхностных явлений.

1.15.1. ОСНОВНЫЕ ПОНЯТИЯ О ПОВЕРХНОСТНЫХ СОСТОЯНИЯХ

Рассмотрим процессы на поверхности полупроводника. В 1932 г. советским ученым И.Е. Таммом было показано, что обрыв кристаллической решетки приводит к возникновению разрешенных уровней в запрещенной зоне. Уровни Тамма возможны для идеальной поверхности полупроводника. Реальные поверхности покрыты слоем адсорбированных атомов и молекул. Эти примеси также создают поверхностные уровни, которые могут быть как донорными, так и акцепторными. Роль примесей могут играть и различные дефекты решетки. Разрешенные уровни в запрещенной зоне полупроводника на его поверхности называются поверхностными уровнями, или поверхностными состояниями.

При высокой плотности поверхностных состояний они, взаимодействуя друг с другом, могут размыться в поверхностную зону. Электроны в этой зоне движутся только вдоль поверхности.

Поверхностные состояния могут захватывать электроны или, наоборот, отдавать их, заряжаясь положительно или отрицательно. Заряжение поверхности полупроводника при заполнении поверхностных состояний сопровождается возникновением у поверхности слоя объемного заряда, нейтрализующего поверхностный заряд. Нейтрализация происходит путем притяжения к поверхности носителей со знаком заряда, противоположным знаку заряда поверхности, и отталкивания носителей одного знака. Вследствие этого поверхностный слой обедняется носителями того знака, которым заряжена поверхность, и обогащается носителями противоположного знака. При этом наблюдается изгиб зон и возникает поверхностный объемный заряд. Если поверхность полупроводника заряжена отрицательно, то в приповерхностном слое зоны изгибаются вверх, а если положительно, то вниз.

1.15.2. ЭФФЕКТ ПОЛЯ. МЕТАЛЛ–ДИЭЛЕКТРИК–ПОЛУПРОВОДНИК-СТРУКТУРЫ

Измерять объемный поверхностный заряд полупроводника можно с помощью электрического поля, перпендикулярного поверхности. Для создания такого поля в полупроводнике, как правило, используют МДП-структуры (рис. 1.54).

Если к идеальной МДП-структуре приложить положительное или отрицательное напряжение, то в приповерхностной области полупроводника могут возникнуть три состояния: обеднение, инверсия и обогащение этой области носителями заряда.

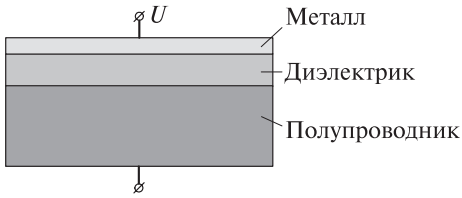


Рис. 1.54. Схема МДП-структуры

Обедненная основными носителями область появляется в случае, когда на металлический электрод подается потенциал, по знаку совпадающий с основными носителями заряда (рис. 1.55, *а*, *б*). Вызванный таким потенциалом изгиб зон приводит к увеличению расстояния от уровня Ферми до дна зоны проводимости в полупроводнике *n*-типа и до потолка валентной зоны в полупроводнике *p*-типа. Увеличение этого расстояния сопровождается обеднением приповерхностной области основными носителями.

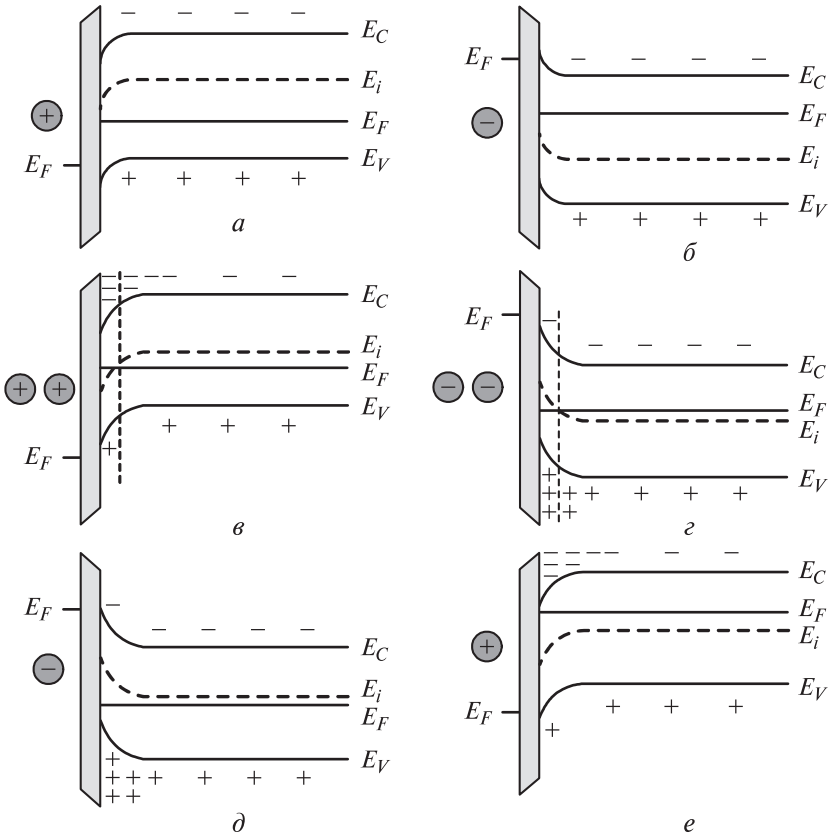


Рис. 1.55. Зонные диаграммы МДП-структур (выделены тоном) в режимах обеднения, инверсии и обогащения

Если на металлический электрод подается достаточно большой потенциал, совпадающий по знаку с основными носителями (рис. 1.55, *в*, *г*), то расстояние от уровня Ферми до потолка валентной зоны в полупроводнике *n*-типа оказывается меньше расстояния до дна зоны проводимости (рис. 1.55, *г*), вследствие чего концентрация неосновных носителей заряда (дырок) у поверхности полупроводника становится выше концентрации основных носителей, и тип проводимости этой области изменяется. Изменение типа проводимости полупроводника называется инверсией, а слои, в которых она наблюдается, называются инверсионными слоями.

Если знак потенциала металлического электрода противоположен знаку заряда основных носителей тока в полупроводнике, то происходит притяжение основных носителей к поверхности и обогащение ими приповерхностного слоя (рис. 1.55, *д*, *е*).

С изменением концентрации основных носителей под действием внешнего поля в приповерхностном слое изменяется также проводимость. Явление изменения поверхностной проводимости под действием поперечного поля называется эффектом поля.

1.15.3. ВОЛЬТ-ФАРАДНАЯ ХАРАКТЕРИСТИКА

Рассмотрим емкость МДП-структуры с полупроводником *n*-типа при изменении напряжения смещения на металлическом электроде при наложении переменного сигнала.

В режиме обогащения в МДП-структуре с полупроводником *n*-типа (плюс на металлическом электроде) электроны притягиваются к поверхности, и в приповерхностном слое полупроводника возникает заряд Q_S притянутых электронов, по абсолютной величине равный заряду Q_M, индуцируемому на металлическом электроде (рис. 1.56). С увеличением напряжения толщина заряженного слоя сохраняется практически неизменной, и рост заряда Q_S происходит за счет повышения объемной плотности, поэтому емкость структуры не зависит от приложенного напряжения.

В состоянии обеднения электроны отталкиваются от поверхности полупроводника, и в приповерхностном слое образуется неподвижный заряд ионизированных доноров $Q_C = qN_Д d_{ОПЗ}$, где $d_{ОПЗ}$ — толщина области поверхностного заряда (ОПЗ) (рис. 1.57). Емкость такой структуры будет складываться из емкостей двух последовательно соединенных конденсаторов: конденсатора, образованного диэлектрическим слоем $d_Д$, и конденсатора, образованного обедненным слоем $d_{ОПЗ}$.

Емкости этих конденсаторов при единичной площади

$$C_Д = \frac{\varepsilon_0 \varepsilon_Д}{d_Д};\ C_{ОПЗ} = \frac{\varepsilon \varepsilon_П}{d_{ОПЗ}}.$$

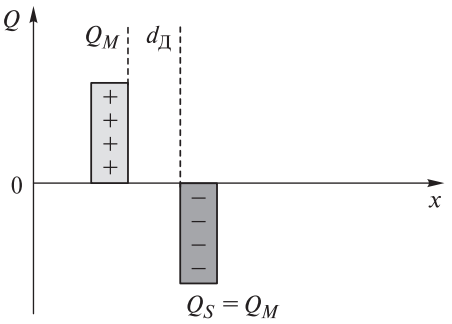


Рис. 1.56. Структура зарядов в режиме обогащения

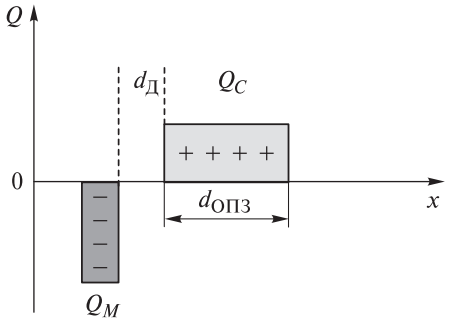


Рис. 1.57. Структура зарядов в режиме обеднения

Емкость $C_{ОПЗ}$ зависит от падающей на полупроводнике разности потенциалов, так как с ее изменением изменяется толщина обедненного слоя $d_{ОПЗ}$, играющая роль расстояния между обкладками конденсатора.

Емкость структуры в целом составит

$$\frac{1}{C} = \frac{1}{C_Д} + \frac{1}{C_{ОПЗ}};$$

$$C = \frac{C_Д C_{ОПЗ}}{C_Д + C_{ОПЗ}}.$$

С ростом напряжения смещения ширина ОПЗ увеличивается, вследствие чего емкость $C_{ОПЗ}$ уменьшается. При этом уменьшается и общая емкость. Это происходит до наступления инверсии. В состоянии инверсии к поверхности полупроводника притягиваются дырки, образуя положительный заряд Q_P (рис. 1.58). До наступления инверсии рост заряда в полупроводнике $Q_П$ определяется ростом заряда ОПЗ и увеличением $d_{ОПЗ}$, а после наступления инверсии — ростом заряда Q_P. Поскольку толщина инверсного слоя, образованного дырками, почти не зависит от приложенного напряжения, в состоянии инверсии емкость структуры не зависит от напряжения.

На рис. 1.59 приведена вольт-фарадная характеристика МДП-структуры. Минимум кривой $C = f(U)$ соответствует образованию инверсного слоя и характеризуется напряжением инверсии $U_{ин}$.

При высокочастотном (ВЧ) сигнале емкость полупроводника определяется

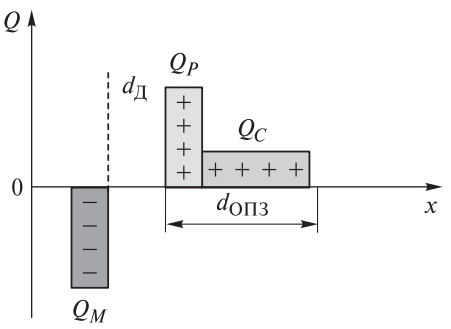


Рис. 1.58. Структура зарядов в режиме инверсии

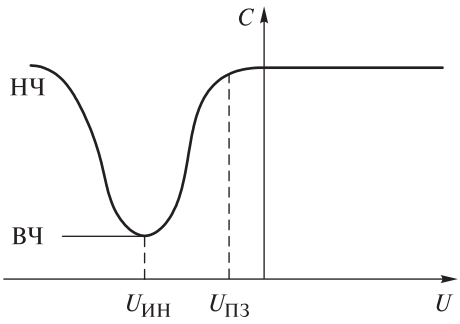


Рис. 1.59. Вольт-фарадная характеристика МДП-структуры с полупроводником *n*-типа

емкостью ОПЗ и не зависит от заряда дырок, а так как после образования инверсного слоя ширина ОПЗ не изменяется, то $C = f(U)$ остается постоянной. При низкочастотном (НЧ) сигнале емкость полупроводника определяется в этом случае зарядом дырок в инверсном слое, который быстро растет с напряжением. Следовательно, возрастает и общая емкость структуры, вплоть до $C \cong C_{Д}$. Затем C не изменяется с ростом напряжения, так как емкость структуры определяется емкостью $C_{Д}$, которая в этом случае меньше емкости полупроводника $C_{П}$.

Кроме рассмотренных выше, в МДП-структуре могут возникать заряд на поверхностных состояниях Q_{SS} и заряд в диэлектрике, которые приводят к сдвигу ВФХ вправо или влево по оси напряжений.

1.15.4. ПОЛЕВЫЕ ТРАНЗИСТОРЫ

Основные принципы работы полевого транзистора были разработаны Лилиенфельдом и Хейлом в начале 1930-х годов. Первый МДП-транзистор был изготовлен Кангом и Аталлой в 1960 г. Полевой транзистор с управляющим *p*–*n*-переходом разработан Шокли в 1952 г. и изготовлен Дейки и Россом в 1953 г. Полевые транзисторы с диодом Шоттки впервые были изготовлены на основе арсенида галлия в 1967 г.

Полевые транзисторы содержат три полупроводниковые области: исток, сток и канал, а также управляющий электрод — затвор. Исток и сток — сильно легированные области полупроводника.

По структуре и способу управления проводимостью канала различают три типа полевых транзисторов:

– с изолированным затвором, между металлическим затвором и каналом расположен слой диэлектрика (МДП-транзисторы);

– с управляющим переходом металл–полупроводник (с диодом Шоттки), металлический электрод затвора образует выпрямляющий контакт с каналом, на который в рабочем режиме подается обратное напряжение;

– с управляющим *p*–*n*-переходом, в качестве затвора используют слой полупроводника, образующий с каналом *p*–*n*-переход, в рабочем режиме имеющий обратное включение.

Полевые транзисторы по типу проводимости канала подразделяют на транзисторы с каналом *n*- или *p*-типа.

Если канал *n*-типа, то рабочими носителями являются электроны, и на сток подается положительный потенциал. В случае канала *p*-типа рабочими носителями являются дырки, и на сток подается отрицательный потенциал.

В полевых транзисторах используется движение носителей заряда одного знака, которые под действием электрического поля, созданного вдоль канала, перемещаются от истока к стоку.

Характерной особенностью полевых транзисторов является малый ток затвора. Входное сопротивление полевых транзисторов на постоянном токе составляет 10^8...10^{10} Ом. Поэтому полевые транзисторы являются приборами, управляемыми с помощью напряжения, в отличие от биполярных транзисторов, которые управляются с помощью тока.

Характерные признаки МДП-транзистора как объекта наноэлектроники:

– толщины подзатворных диэлектриков от единиц нанометров до 100 нм;

– длина канала от 20 нанометров до единиц микрометров;

– квантование электронного газа в инверсионном слое;

– механизмы переноса заряда через наноразмерные диэлектрические слои: сильнополевая туннельная инжекция по Фаулеру — Нордгейму (туннелирование носителей через треугольные потенциальные барьеры толщиной от единиц до десятков нанометров) и прямое туннелирование через диэлектрические слои толщиной менее 4 нм;

– интерференционные явления при переносе электронов через диэлектрические пленки наноразмерной толщины.

Полевые транзисторы с изолированным затвором. Структура МДП-транзистора с индуцированным каналом *n*-типа показана на рис. 1.60. В подложке *p*-типа с невысокой концентрацией акцепторов (10^{15} см$^{-3}$) методами диффузии или ионного легирования сформированы сильно легированные истоковая и стоковая области с концентрацией примеси 10^{19} см$^{-3}$.

Расстояние между стоковой и истоковой областями называется длиной канала L и составляет от десятков нанометров до нескольких микрометров. На поверхности полупроводника создается тонкий слой подзатворного диэлектрика толщиной 10...100 нм. В большинстве случаев это термически выращенный диоксид кремния. Металлические слои образуют выводы стока, истока и затвора.

Работа транзистора. Предположим, что заряд в диэлектрике равен нулю. Тогда при равенстве нулю напряжения между затвором и истоком ($U_{зи} = 0$) проводящий канал в полупроводнике *p*-типа будет отсутствовать и между стоком и истоком оказываются два встречно включенных *p*–*n*-перехода. Поэтому при подаче напряжения на сток ток, протекающий через транзистор, будет ничтожно мал.

При подаче на затвор отрицательного напряжения ($U_{зи} < 0$) поверхностный слой обогащается дырками, а ток, протекающий через транзистор, остается ничтожно малым. Транзистор закрыт. При подаче на затвор возрастающего положительного потенциала в подложке сначала образуется слой, обедненный дырками, а затем инверсионный слой электронов, образующий проводящий канал. Через транзистор начинает протекать ток, зависящий от напряжения на затворе.

Каналы, образующиеся под действием внешнего напряжения и отсутствующие в равновесном состоянии, называются индуцированными. Толщи-

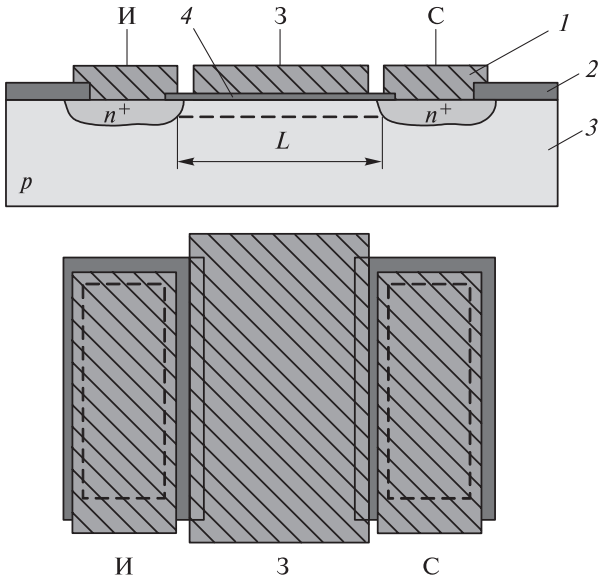


Рис. 1.60. Структура МДП-транзистора:
1 — металлизация; *2* — толстый оксид; *3* — кристалл кремния; *4* — подзатворный диэлектрик; И — исток; З — затвор; С — сток

на индуцированного канала составляет 1...2 нм и практически не изменяется. Модуляция проводимости канала осуществляется за счет изменения концентрации носителей. Напряжение на затворе, при котором образуется проводящий канал, называется пороговым напряжением.

Если поверхностный слой полупроводника между стоком и истоком легировать донорами, то проводящий канал будет существовать и при нулевом напряжении на затворе. Пороговое напряжение при этом будет отрицательным. Такие транзисторы называются МДП-транзисторами со встроенным каналом.

Обозначения МДП-транзисторов и полярности рабочих напряжений для включения по схеме с общим истоком показаны на рис. 1.61.

В МДП-транзисторах с *p*-каналом области истока и стока p^+-типа формируются в подложке *n*-типа. Если МДП-транзистор выполнен со встроенным каналом, то он работает как при положительном, так и при отрицательном напряжении на затворе, тогда как МДП-транзистор с индуцированным каналом работает только при одной полярности напряжения на затворе.

Статические ВАХ МДП-транзисторов. Выходные характеристики МДП-транзисторов показаны на рис. 1.62. На выходные ВАХ существенное влияние оказывают изменения в структуре канала, возникающие с ростом напряжения на стоке.

При $U_И = U_С = 0$ электрическое поле в диэлектрике и полупроводнике будет однородным, и толщина канала будет одинаковой от истока до стока.

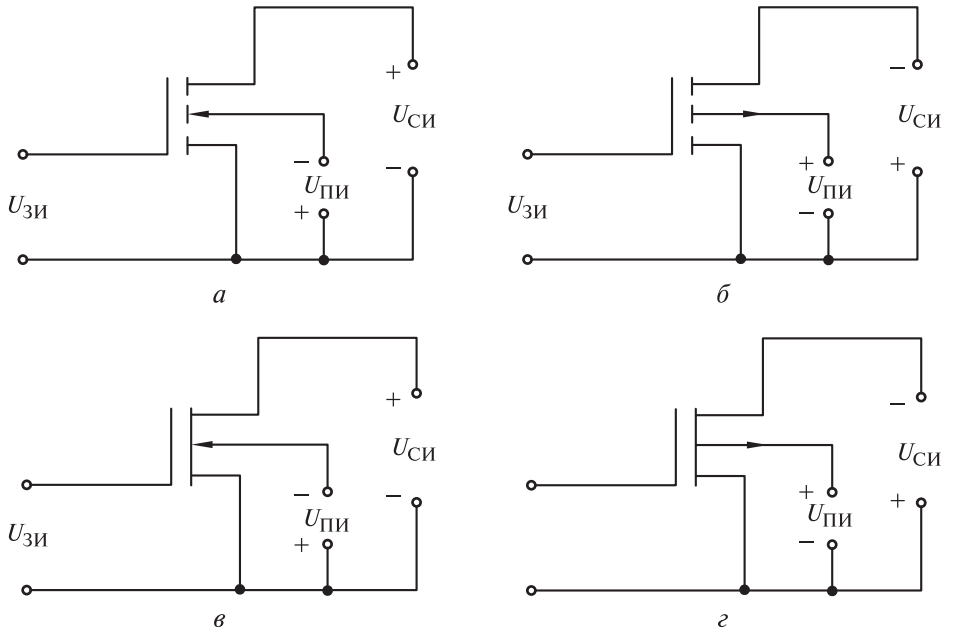


Рис. 1.61. Схемы МДП-транзисторов с индуцированным *n*-каналом (*а*), с индуцированным *p*-каналом (*б*), со встроенным *n*-каналом (*в*), со встроенным *p*-каналом (*г*) ($U_{ПИ}$ — напряжение между подложкой и истоком)

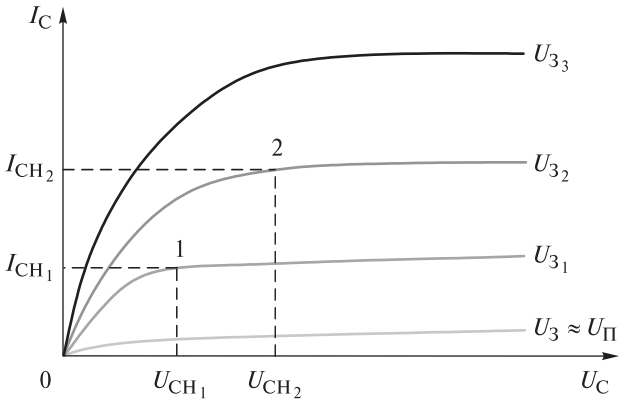


Рис. 1.62. Зависимость тока стока от напряжения сток–исток при различных напряжениях на затворе ($U_{З_3} > U_{З_2} > U_{З_1} > U_З > U_П$; точки 1 и 2 соответствуют переходу транзистора в режим насыщения)

Если напряжение $U_{СИ} > 0$ и невелико, то канал ведет себя как обычное сопротивление. Ток стока увеличивается пропорционально напряжению стока. Эту область ВАХ называют линейной областью работы транзистора.

С ростом напряжения $U_{СИ}$ будут увеличиваться ток стока и потенциал поверхности полупроводника в направлении от истока к стоку. Вследствие этого разность потенциалов между затвором и поверхностью полупроводника будет уменьшаться в направлении стока. Соответственно сечение канала начинает сужаться в направлении стока. При напряжении на стоке, равном напряжению насыщения $U_{СН}$, разность потенциалов между затвором и поверхностью полупроводника будет равна нулю у стока. Толщина канала у стока становится равной нулю, МДП-транзистор переходит в режим отсечки канала. При напряжении $U_{СИ} > U_{СН}$ точка отсечки сдвигается к истоку и происходит укорочение канала на ΔL (рис. 1.63). На участке ΔL обедненный слой выходит на поверхность полупроводника.

После отсечки канала ток стока не будет зависеть от потенциала стока. Эта область ВАХ называется областью насыщения тока стока.

На острие канала, в точке его перекрытия, концентрируется электрическое поле, напряженность которого становится выше критической $\varepsilon_{кр}$, и наступает режим насыщения скорости дрейфа электронов ($v_D = \mu_n \varepsilon_{кр} = \text{const}$), инжектированных из острия канала в обедненный слой, следовательно, плотность тока $j_n = e\mu_n n\varepsilon_{кр} = \text{const}$, где e — заряд электрона μ_n, n — подвижность и концентрация электронов.

Основные параметры МДП-транзисторов

1. *Удельная емкость затвор–полупроводник* — определяет степень модуляции проводимости канала:

$$C_0 = \frac{\varepsilon_д \varepsilon_0}{d},$$

где $\varepsilon_д$, d — относительная диэлектрическая проницаемость и диэлектрическая толщина подзатворного диэлектрика.

2. *Пороговое напряжение* (напряжение затвор–исток, при котором возникает проводящий канал в транзисторе):

$$U_п = U_{п.з} + U_0,$$

где $U_{п.з}$ — напряжение плоских зон; U_0 — напряжение изгиба зон.

3. *Крутизна сток–затворной характеристики* (чем больше S, тем чувствительнее реакция транзистора на изменение напряжения на затворе):

$$S = \left.\frac{dI_С}{dU_{ЗИ}}\right|_{U_{СИ}=\text{const}}.$$

4. *Внутреннее (выходное) сопротивление* (сопротивление канала для переменного тока при постоянном напряжении затвор–исток):

$$R_{\text{С}} = \left.\frac{dU_{\text{СИ}}}{dI_{\text{С}}}\right|_{U_{\text{ЗИ}}=\text{const}}.$$

5. Коэффициент усиления (отношение изменения напряжения исток–сток к изменению напряжения затвор–исток при постоянном токе стока):

$$K = \left.\frac{dU_{\text{СИ}}}{dU_{\text{ЗИ}}}\right|_{I_{\text{С}}=\text{const}};$$

$$K = SR_{\text{С}}.$$

Обычно $K = 50...200$.

1.15.5. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С УПРАВЛЯЮЩИМ *p–n*-ПЕРЕХОДОМ

Структура такого транзистора приведена на рис. 1.64. На подложке *p*-типа формируется эпитаксиальный *n*-слой, в котором методами диффузии создаются области истока, стока n^+-типа и затвора p^+-типа. Управляющий *p–n*-переход образуют области p^+ и *n*. Токопроводящим каналом является эпитаксиальный слой *n*-типа, расположенный между затвором и подложкой. При работе транзистора управляющий *p–n*-переход должен быть включен в обратном направлении.

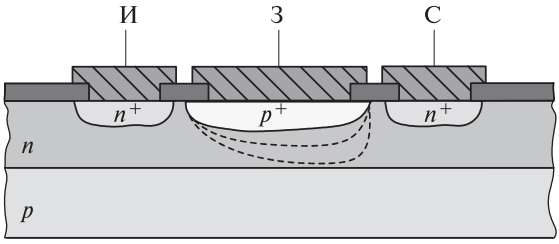


Рис. 1.64. Структура полевого транзистора с управляющим переходом

Глубина обедненного слоя управляющего *p–n*-перехода тем больше, чем больше обратное напряжение на затворе. Толщина канала будет также соответственно меньше. Следовательно, с изменением обратного напряжения будет изменяться поперечное сечение канала, а следовательно, и его сопротивление. При наличии напряжения между стоком и истоком, изменяя обратное напряжение на затворе, можно управлять выходным током транзистора.

Входным током транзистора является обратный ток *p–n*-перехода, составляющий для кремниевых приборов $10^{-9}...10^{-11}$ А.

На сток транзистора подается положительное напряжение, *p–n*-переход между эпитаксиальным *n*-слоем и подложкой включается в обратном направлении, поэтому к подложке прикладывается отрицательное относительно

истока напряжение. Иногда подложка используется в качестве второго затвора. В некоторых транзисторах подложка соединяется с затвором и не имеет отдельного вывода.

Статические выходные и проходные характеристики полевых транзисторов с управляющим *p–n*-переходом показаны на рис. 1.65 и 1.66.

Выходные характеристики имеют участок насыщения тока, связанный, как и у МДП-транзисторов, с образованием «горловины» канала вблизи стока.

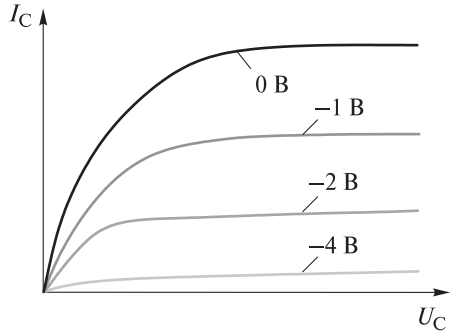


Рис. 1.65. Выходные ВАХ полевого транзистора с управляющим переходом

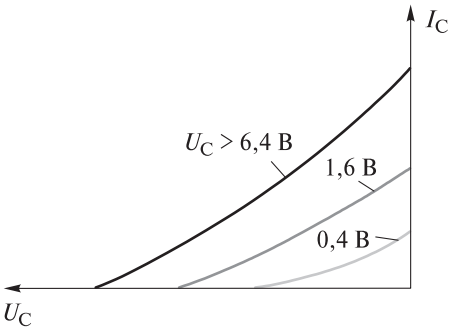


Рис. 1.66. Проходные ВАХ полевого транзистора с управляющим переходом

Напряжением отсечки полевого транзистора с управляющим *p–n*-переходом называется напряжение на затворе, при котором практически полностью перекрывается канал, и ток стока стремится к нулю.

1.15.6. МЕХАНИЗМЫ ПЕРЕНОСА ЗАРЯДА ЧЕРЕЗ ТОНКИЕ ДИЭЛЕКТРИЧЕСКИЕ ПЛЕНКИ

Основными механизмами переноса заряда, определяющими проводимость изолирующих пленок, являются: термополевая эмиссия Шоттки, эмиссия Пула – Френкеля и сильнополевая туннельная инжекция по Фаулеру – Нордгейму.

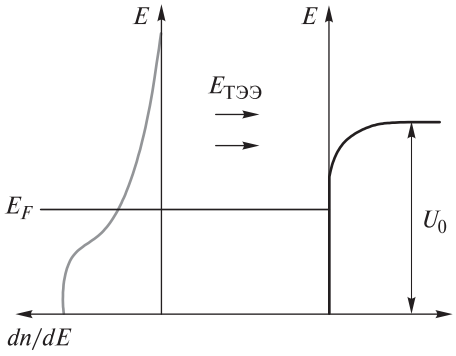


Рис. 1.67. Термоэлектронная эмиссия

Термополевая эмиссия по Шоттки. Термоэлектронная эмиссия электронов объясняется наличием высокоэнергетического «хвоста» в распределении электронов по энергии. Термоэлектронной эмиссией называется испускание электронов нагретым телом в вакууме (рис. 1.67).

При наложении внешнего электрического поля происходит понижение

потенциального барьера, разделяющего металл и диэлектрик. Это явление называется эффектом Шоттки.

Снижение потенциального барьера происходит в результате сложения потенциала внешнего поля $U_\varepsilon = q\mathcal{E}_x$ с потенциалом U_0 на границе металл–диэлектрик, определяемым силами зеркального изображения (рис. 1.68).

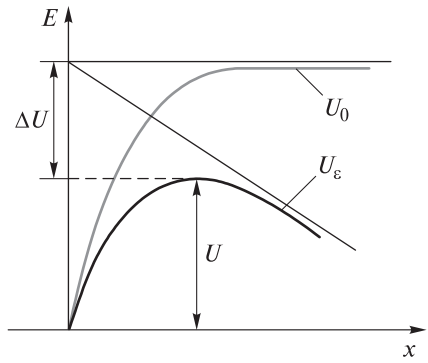


Рис. 1.68. Термополевая эмиссия

Термополевая эмиссия Шоттки представляет собой надбарьерную термоэлектронную эмиссию, облегченную за счет наложения внешнего электрического поля. Эмиссия Шоттки является одним из основных механизмов переноса зарядов в системах металл–диэлектрик–металл и металл–полупроводник–металл.

Плотность тока эмиссии по Шоттки

$$j_{\text{ш}} = AT^2 e^{-\frac{q\varphi_{\text{в}}}{kT}} e^{\frac{q}{kT}\sqrt{\frac{qU}{4\pi\varepsilon\varepsilon_0 d}}},$$

где $A = 12 \cdot 10^5$ А/(см$^2 \cdot$ град^2) — постоянная Ричардсона; $\varphi_{\text{в}}$ — высота потенциального барьера.

Эмиссия Пула — Френкеля. В диэлектрических слоях, содержащих большое количество структурных дефектов и примесей, имеется высокая концентрация ловушек, способных захватывать носители.

Сильное электрическое поле может вызвать активацию свободных носителей заряда внутри самой диэлектрической пленки (рис. 1.69). Эмиссия Пула — Френкеля — это ускоренный электрическим полем процесс термовозбуждения электронов с ловушек в зону проводимости диэлектрика.

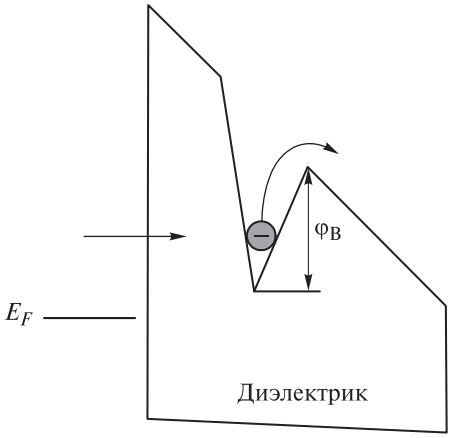


Рис. 1.69. Эмиссия Пула – Френкеля через тонкие диэлектрические пленки

Плотность тока эмиссии по Пулу — Френкелю равна:

$$j_{\text{ф}} = B\mathcal{E} e^{-\frac{q\varphi_{\text{в}}}{kT}} e^{\frac{2q}{kT}\sqrt{\frac{qU}{4\pi\varepsilon\varepsilon_0 d}}},$$

где B — коэффициент пропорциональности; $\varphi_{\text{в}}$ — глубина ловушки.

Зависимости плотностей токов эмиссий по Шоттки и Пулу — Френкелю спрямляются в координатах $\lg j = f\sqrt{U}$.

Токи эмиссий Шоттки и Пула — Френкеля в значительной степени зависят от температуры.

Тесты к лекции 1.15

1. ***Что такое МДП-структура?***

а) структура, состоящая из металла, диэлектрика и полупроводника;

б) структура, состоящая из металла, диполя и подложки;

в) структура, состоящая из металла, донорной примеси и полупроводника.

2. ***Что называется эффектом поля?***

а) явление изменения поверхностной проводимости под действием поперечного поля;

б) явление изменения относительной проводимости транзистора под действием продольного поля;

в) явление изменения барьерной проводимости под действием поперечного магнитного потока.

3. ***Какое явление в полупроводнике называется инверсией?***

а) изменение типа проводимости полупроводника;

б) появление инверсионных следов на поверхности полупроводника;

в) инверсионное повреждение *p*—*n*-перехода.

4. ***Какие три электрода содержат полевые транзисторы?***

а) исток, сток и затвор;

б) эмиттер, база, коллектор;

в) эмиттер, сток, коллектор.

5. ***Какие типы проводимости каналов бывают у полевых транзисторов?***

а) тип проводимости *n*- или *p*-типа;

б) с дискретным и непрерывным типом проводимости;

в) с фотонным и инверсионным типом проводимости.

6. ***Какие каналы в полевых транзисторах называются индуцированными?***

а) каналы, образующиеся под действием внешнего напряжения;

б) каналы, образующиеся при отключении внешнего напряжения;

в) каналы, образующиеся вследствие туннельного эффекта.

7. ***Какое напряжение называется напряжением отсечки в полевом транзисторе с управляющим p—n-переходом?***

а) напряжение на затворе, при котором практически полностью перекрывается канал, и ток стока стремится к нулю;

б) напряжение на затворе, при котором практически полностью открывается канал, и ток стока стремится к насыщению;

в) напряжение между истоком и стоком, при котором практически полностью открывается канал, и ток стока стремится к нулю.

8. ***Что такое термоэлектронная эмиссия?***

а) испускание электронов нагретым телом в вакууме;

б) испускание ионов нагретым телом в вакууме;

в) испускание атомов нагретым телом в вакууме.

9. ***Что такое термополевая эмиссия Шоттки?***

а) надбарьерная термоэлектронная эмиссия, облегченная за счет наложения внешнего электрического поля;

б) эмиссия, образованная за счет эффекта Шоттки — Пула;

в) эмиссия электронов за счет воздействия магнитного поля.

10. ***Что такое эмиссия Пула — Френкеля?***

а) ускоренный электрическим полем процесс термовозбуждения электронов с ловушек в зону проводимости диэлектрика;

б) ускоренный электрическим полем процесс перехода электронов из зоны проводимости на ловушки в диэлектрике;

в) ускоренный электрическим полем процесс перехода электронов через *p*–*n*-переход.

1.16. ТУННЕЛЬНЫЕ ЭФФЕКТЫ И ПОЛУПРОВОДНИКОВЫЕ СВЕРХРЕШЕТКИ

Цель лекции: ознакомление с туннельными эффектами и полупроводниковыми сверхрешетками.

1.16.1. ОСНОВНЫЕ ПОНЯТИЯ О ТУННЕЛЬНЫХ ЭФФЕКТАХ

Туннельный эффект, как было показано ранее, заключается в прохождении микрочастиц сквозь потенциальный барьер, высота которого больше энергии налетающей частицы. Вероятность прохождения микрочастицы сквозь барьер называется коэффициентом прозрачности D. Значение D тем больше, чем меньше ширина барьера и разность между его высотой и энергией частицы. При прохождении частицей потенциального барьера посредством туннелирования ее энергия не изменяется. Вероятность туннельного эффекта велика, если ширина потенциального барьера соизмерима с длиной волны де Бройля электрона. Туннельный эффект широко используется во многих приборах микро- и наноэлектроники (туннельные диоды, обращенные переключающие диоды, диоды Зенера, сверхрещетки и др.). Он определяет пределы функционирования элементов интегральных схем на основе традиционных принципов. В частности, толщина подзатворного диэлектрика МДП-транзисторов ограничивается значением прямого туннельного тока.

Уникальными свойствами обладает резонансный туннельный эффект, который проявляется в двух- или многобарьерной периодической структуре и состоит в резком увеличении вероятности прохождения частицы сквозь барьеры, если ее энергия совпадает с каким-либо размерным уровнем энергии в потенциальной яме, разделяющей барьеры. Резонансное туннелирование сквозь ряд барьеров возникает только в случае, если ширина ям и барьеров имеет порядок длины волны де Бройля. В этом эффекте время прохождения структуры электроном включает помимо времени туннелирования время пребывания электрона в яме, разделяющей барьеры, т. е. время его жизни на резонансном уровне.

На основе двухбарьерных структур могут работать СВЧ-приборы в диапазоне сотен гигагерц и переключатели с задержкой менее 1 пс. Созданы приборы на основе двухбарьерной структуры — резонансно-туннельные диоды и транзисторы. Разработаны и находят все более широкое применение многобарьерные структуры, которые называются сверхрешетками.

1.16.2. СИЛЬНОПОЛЕВАЯ ТУННЕЛЬНАЯ ИНЖЕКЦИЯ

Основным механизмом переноса заряда в МДП-структурах является сильнополевая туннельная инжекция носителей по Фаулеру — Нордгейму. При малых толщинах слоя оксида может осуществляться прямое туннелирование через слой диэлектрика. Граница между прямым туннелированием и сильнополевой инжекцией по Фаулеру — Нордгейму лежит в диапазоне 3,5...4,0 нм.

В сильных электрических полях в МДП-структурах возникает треугольный потенциальный барьер, образующийся за счет изгиба зон в диэлектрике, через который осуществляется квантово-механический туннельный перенос электронов (рис. 1.70).

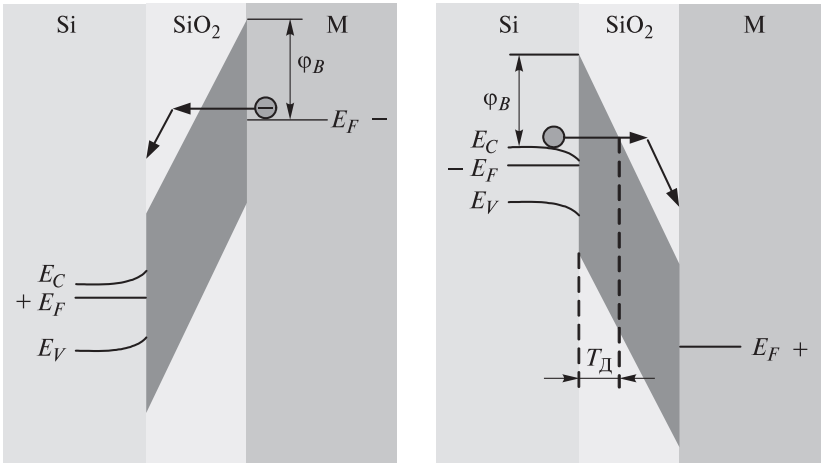


Рис. 1.70. Зонные диаграммы МДП-структур при инжекции электронов из полупроводника (*а*) и металла (*б*)

При положительной полярности напряжения на металлическом затворе треугольный потенциальный барьер образуется на границе $Si–SiO_2$ (см. рис. 1.70, *а*), при отрицательной — на границе $SiO_2–M$ (рис. 1.70, *б*).

Зависимость плотности тока сильнополевой туннельной инжекции по Фаулеру — Нордгейму описывается следующим выражением:

$$j = \frac{A}{\varphi_B}\mathcal{E}^2 \exp\left(-\frac{B\varphi_B^{3/2}}{\mathcal{E}}\right),$$

где A, B — соответствующие коэффициенты; φ_B — высота потенциального барьера на инжектирующей границе.

Высота потенциального барьера на границе $Si–SiO_2$ составляет 2,8...3,19 эВ.

Экспериментальные зависимости тока сильнополевой туннельной инжекции рассматривают в координатах Фаулера — Нордгейма: $\ln\left(j/\mathcal{E}^2\right) = f\left(1/\mathcal{E}\right)$.

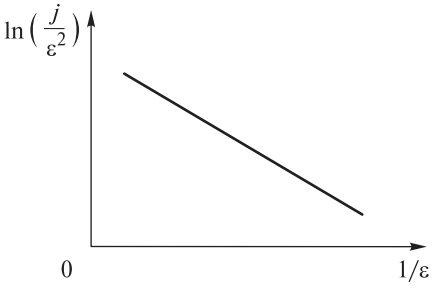


Рис. 1.71. Вольт-амперная характеристика в координатах Фаулера — Нордгейма

По наклону прямой, построенной в координатах Фаулера — Нордгейма, определяют высоту потенциального барьера на инжектирующей границе (рис. 1.71).

Сильнополевая туннельная инжекция по Фаулеру — Нордгейму может использоваться для модификации (целенаправленного изменения) характеристик МДП-структур. Сильнополевой инжекцией электронов в диэлектрике, содержащем электронные ловушки, можно изменять зарядовое состояние подзатворной системы. В процессе сильнополевой инжекции осуществляется заполнение инжектированными электронами электронных ловушек, и в диэлектрике образуется отрицательный заряд, сохраняющийся после прекращения инжекции. Это позволяет изменять пороговое напряжение МДП-транзисторов.

Инжекционная модификация применяется в полевых приборах на основе МДП-структур, параметры которых можно изменять после их изготовления с помощью сильнополевой инжекции.

1.16.3. ТУННЕЛЬНЫЙ ЭФФЕКТ В *p*—*n*-ПЕРЕХОДЕ, ОБРАЗОВАННОМ ВЫРОЖДЕННЫМИ ПОЛУПРОВОДНИКАМИ

Вырождение полупроводников и малые значения ширины *p*—*n*-перехода достигаются за счет сильного легирования *p*—*n*-областей. Уровень Ферми в этом случае располагается в валентной области полупроводника *p*-типа и в зоне проводимости полупроводника *n*-типа.

Зонные диаграммы контакта двух вырожденных полупроводников приведены на рис. 1.72. Малая ширина *p*—*n*-перехода позволяет электронам туннелировать из зоны проводимости *n*-области на свободные уровни в валентной зоне, а дыркам — из валентной области в зону проводимости *n*-области.

При отсутствии внешнего смещения встречные потоки дырок и электронов будут равны, а результирующий ток через *p*—*n*-переход будет равен нулю (рис. 1.72, *а*). При подаче прямого смещения на *p*—*n*-переход зона проводимости *n*-области сместится вверх, и заполненные уровни *n*-области окажутся напротив свободных уровней валентной зоны *p*-области (рис. 1.72, *б*). Вероятность перехода электронов из зоны проводимости *n*-области в *p*-область увеличивается, а вероятность перехода дырок из валентной зоны *p*-области уменьшается. В результате через *p*—*n*-переход потечет прямой ток, увеличивающийся с ростом прикладываемого напряжения и достигающий максимума, когда заполненная часть зоны проводимости *n*-области будет располагаться напротив незаполненной части валентной зоны.

При дальнейшем увеличении прямого смещения перекрытие этих частей зон уменьшается, и ток туннелирования падает до нуля. При $U > U_B$ через

p–n-переход протекает обычный диффузионный ток или ток рекомбинации (рис. 1.72, *в*). При обратном смещении (рис. 1.72, *г*) туннельный ток возрастает с увеличением напряжения, так как все большее число электронов из валентной зоны p-области переходит в зону проводимости n-области.

Вольт-амперная характеристика туннельного диода показана на рис. 1.73. Участок *OA* — возрастание туннельного тока, участок *AB* — уменьшение, *BC* — диффузионная ветвь ВАХ. На участке отрицательного сопротивления *AB* действует положительная обратная связь по напряжению; U_m — напряжение, при котором дифференциальная проводимость первый раз обращается в нуль и соответствует пику прямого тока и началу участка *AB* с отрицательным дифференциальным сопротивлением; $U_в$ — напряжение, при котором дифференциальная проводимость второй раз обращается в нуль и соответствует окончанию участка *AB* с отрицательным дифференциальным сопротивлением. Увеличение внешнего напряжения (рис. 1.74) приводит к уменьшению туннельного тока и, следовательно, к уве-

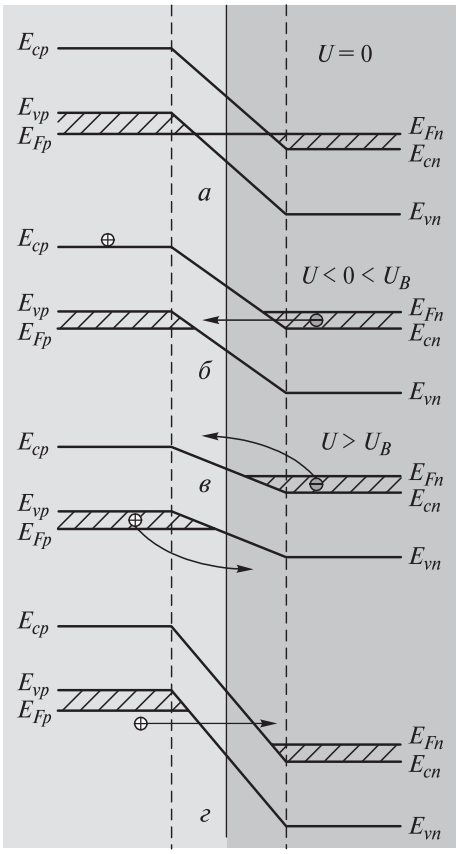


Рис. 1.72. Зонные диаграммы туннельного диода

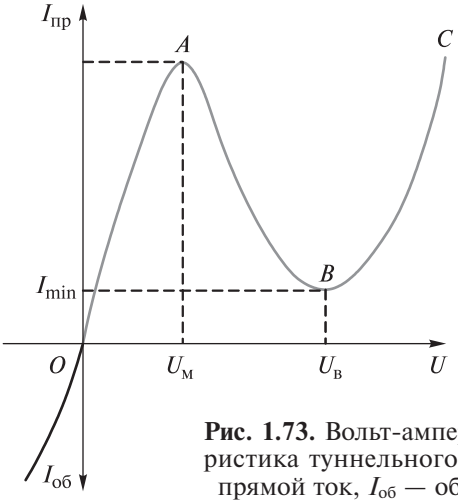


Рис. 1.73. Вольт-амперная характеристика туннельного диода ($I_{пр}$ — прямой ток, $I_{об}$ — обратный ток)

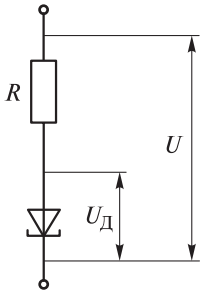


Рис. 1.74. Последовательное включение резистора и туннельного диода

личению сопротивления туннельного диода. Вследствие перераспределения внешнего напряжения падение напряжения на диоде $U_д$ возрастает, что приводит к еще большему увеличению сопротивления.

Туннельные диоды обладают высоким быстродействием (время переключения — единицы наносекунд). Различают усилительные, генераторные и переключающиеся туннельные диоды. Усилительные туннельные диоды применяются в усилителях и гетеродинах приемных устройств, в схемах детекторов и смесителей диапазона СВЧ, генераторные — в СВЧ генераторах диапазона волн 1...10 см, переключательные — в импульсных схемах наносекундного диапазона длин волн.

1.16.4. РЕЗОНАНСНО-ТУННЕЛЬНЫЕ ДИОДЫ

Резонансно-туннельный диод (РТД) представляет собой диод с несколькими (как правило, с двумя) потенциальными барьерами, в котором реализуются условия для резонансного туннелирования. Физический смысл работы РТД состоит в том, что электронная волна проникающего через первый барьер электрона попадает в потенциальную яму, в которой она практически полностью «отражается» от стенок и интерферирует с волной продолжающего туннелировать электрона. В случае если фазы падающей и дважды отраженной от стенок волн совпадают, то амплитуда волны в яме резко возрастает, что вызывает резкое резонансное возрастание тока, протекающего через структуру. Условием возникновения резонанса является совпадение энергии туннелирующего электрона с уровнем энергии в квантовой яме.

Структура РТД представлена на рис. 1.75. Сильнолегированные слои n^+–GaAs составляют эмиттер и коллектор. Слой слаболегированного n–GaAs образует квантовую яму толщиной 3...10 нм. Слои $Al_xGa_{1-x}As$ толщиной 2...5 нм образуют потенциальные барьеры. Высота барьеров *2* и *4* зависит от концентрации алюминия. Высоту барьера *3* и ширину потенциальной ямы d_3 подбирают таким образом, чтобы в яме образовался только один размерный квантовый уровень E_1.

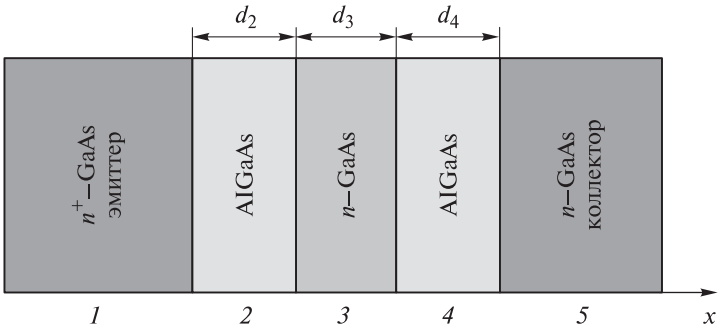


Рис. 1.75. Структура РТД:
1–5 — потенциальные барьеры; d_2–d_4 — расстояния между квантовыми ямами

На рис. 1.76 показаны зонные диаграммы РТД в различных режимах (заполненные уровни выделены тоном).

В равновесном состоянии напряжение между эмиттерами и коллектором $U = 0$, ток в структуре отсутствует (рис. 1.76, *а*).

Если постепенно повышать напряжение на структуре, то через нее пойдет слабый ток, обусловленный обычным туннелированием электронов через потенциальный барьер *2* (участок $0 - U_1$ на рис. 1.77). При этом энергетические уровни эмиттера будут подниматься относительно уровней коллектора. При напряжении U_1, когда уровень E_{F_1} сравняется с уровнем E_1, через структуру потечет значительно бóльший ток, связанный с резонансным туннелированием (см. рис. 1.76). С увеличением напряжения на структуре $U_1 < U < U_2$ ток резко возрастает до тех пор, пока с уровнем E_1 потенциальной ямы не сравняется дно зоны проводимости E_{C_1} (см. рис. 1.76). При дальнейшем увеличении напряжения $U > U_2$ уровень E_1 опускается ниже дна зоны проводимости в запрещенную зону, где электронов нет, и ток резко падает (см. рис. 1.76), что приводит к появлению на ВАХ участка с отрицательным дифференциальным сопротивлением (см. рис. 1.77).

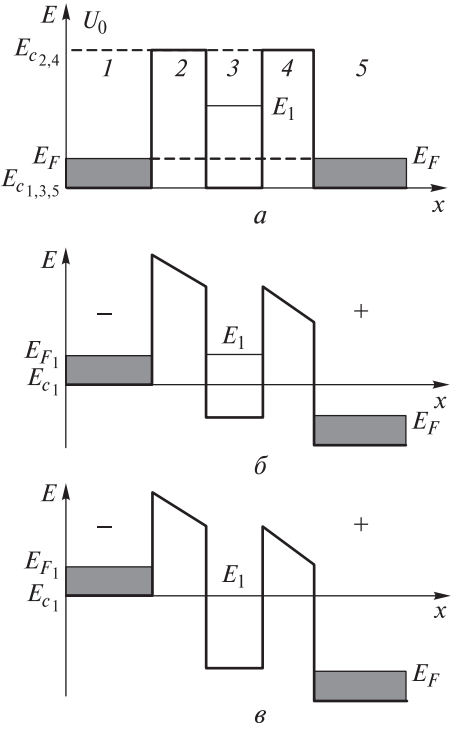


Рис. 1.76. Зонные диаграммы РТД: *а* — равновесное состояние; *б* — $U < U_1$; *в* — $U_1 < U < U_2$

Аналогичные резонансные явления могут наблюдаться в структурах с тремя барьерами и в периодических сверхрешетках.

Наличие на ВАХ РТД участка с отрицательным дифференциальным сопротивлением позволяет использовать данный прибор для генерации СВЧ-колебаний с частотами выше (на порядок и более), чем в обычных туннельных диодах, до 700 ГГц. Физическими факторами, ограничивающими быстродействие РТД, являются время жизни электронного состояния в яме и время пролета электроном обедненного слоя. Отрица-

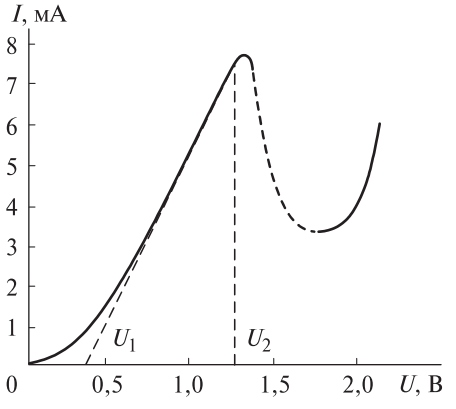


Рис. 1.77. Вольт-амперная характеристика РТД

тельное дифференциальное сопротивление в РТД сохраняется по крайней мере до частоты ~2,5 ТГц. Более высокое быстродействие РТД в основном обусловлено существенным улучшением условий для туннельного преодоления барьера и меньшей емкостью структуры. Обычные туннельные диоды работают при плотности тока $10^2...10^3$ А/см^2, которая ограничена невысокой прозрачностью туннельного барьера, а в РТД из-за практически 100%-ного прохождения барьера электронами в определенном интервале энергий плотность тока может достигать $4 \cdot 10^5$ А/см^2. Удельная барьерная емкость в РТД вследствие присутствия истощенного слоя в несколько раз меньше, чем в туннельных диодах.

Резонансно-туннельный диод преобразуется в транзистор, если к слою потенциальной ямы структуры подвести электрод, регулирующий положение размерного уровня E_1 с помощью электрического поля. На основе РТД разработаны аналого-цифровые преобразователи, работающие на частотах в несколько гигагерц, логические элементы, запоминающие устройства, которые можно использовать в процессорах, и другие цифровые устройства для быстродействующей электроники.

1.16.5. ПОЛУПРОВОДНИКОВЫЕ СВЕРХРЕШЕТКИ

Сверхрешетки — синтезированные кристаллические структуры, в которых кроме периодического потенциала кристаллической решетки имеется другой периодический потенциал, период которого значительно превышает постоянную решетки. Параметры этого потенциала можно изменять, что дает возможность управлять энергетическим спектром носителей.

Для создания требуемой искусственной периодической структуры применяют два способа.

1. Используют один и тот же полупроводник, но легируют его попеременно, создавая слои *n*- и *p*-типов (так называемые nipi-сверхрешетки, или легированные сверхрешетки). Донорные атомы в *n*-слоях отдают электроны, которые связываются акцепторными атомами в *p*-слоях. Дополнительный периодический потенциал создают чередующиеся заряды ионизированных акцепторов и доноров.

2. Используют чередующиеся слои двух различных полупроводников, в которых запрещенная зона одного материала перекрывает запрещенную зону другого (так называемые композиционные сверхрешетки). Используют гетероструктуры из чередующихся слоев различного состава и ширины запрещенной зоны, но с близкими значениями постоянной решетки, например, $Al_xGa_{1-x}As–GaAs$; $In_xGa_{1-x}As–GaAs$; $In_xGa_{1-x}As–InP$; ZnS–ZnSe и др. Дополнительный периодический потенциал создается за счет периодического изменения ширины запрещенной зоны.

Оба типа сверхрешеток получают с использованием технологии молекулярно-лучевой эпитаксии (МЛЭ) и газофазной эпитаксии (ГФЭ) из паров

металлоорганических соединений. Оба метода позволяют осуществлять послойное эпитаксиальное наращивание атомных слоев заданного состава на монокристаллическую подложку. Метод молекулярно-лучевой эпитаксии обычно используется для получения экспериментальных образцов гетероструктур, ГФЭ — для их массового производства. Период повторения слоев составляет от нескольких нанометров до десятков нанометров (для сравнения — постоянные решетки кристаллов Si и GaAs равны примерно 0,5 нм).

Особенностью сверхрешеток является возможность искусственно формировать их электронный спектр. Если при выращивании сверхрешетки расстояние между квантовыми ямами d_2 сделать небольшим, чтобы электроны могли туннелировать из одной ямы в другую через потенциальный барьер, образованный широкозонным полупроводником, то уровни размерного квантования, отвечающие движению электрона перпендикулярно стенкам ямы, размываются в так называемую мини-зону (рис. 1.78). При этом ширина мини-зоны определяется перекрытием волновых функций электронов в соседних ямах, а расстояние между разными мини-зонами — шириной квантовой ямы. Это позволяет, изменяя толщины слоев в сверхрешетке, направленно изменять электронный спектр полупроводника и тем самым создавать новый искусственный полупроводниковый материал, обладающий уникальными свойствами. Совокупность методов получения материалов с модифицированной зонной структурой лежит в основе так называемой зонной инженерии.

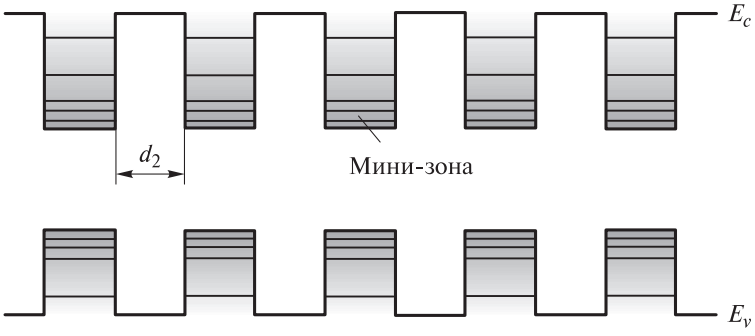


Рис. 1.78. Энергетическая диаграмма сверхрешетки

Можно сделать следующие качественные выводы относительно энергетической структуры сверхрешетки. Потенциал сверхрешетки периодичен, поэтому размерные уровни расщепляются в зоны. Спектр имеет зонный характер; в зоне столько уровней, сколько ям в структуре. Поскольку период сверхрешетки значительно больше ее постоянной, получающиеся сверхрешеточные зоны представляют собой более мелкое дробление энергетических зон исходных полупроводниковых кристаллов и называются мини-зонами.

Чем меньше ширина ямы, тем больше расстояние между мини-зонами и больше эффективная ширина запрещенной зоны; чем меньше ширина

барьера, тем шире мини-зоны. Таким образом, можно перестраивать энергетический спектр сверхрешетки простым изменением толщины слоев, что легко осуществить в методе МЛЭ. График плотности состояний $g(E)$ имеет ступенчатый вид, как и $g(E)$ для квантовой ямы, но с иной формой ступеней.

Сверхрешетки используются в ряде полупроводниковых приборов, например, в лазерах, светодиодах, фотоприемниках, а также в транзисторах и других устройствах с отрицательным дифференциальным сопротивлением. Принципы действия подобных приборов основаны на специфических для сверхрешеток явлениях: квантовом ограничении носителей заряда в потенциальных ямах, пространственном разделении электронов и дырок (большие времена жизни неравновесных носителей) или электронов и доноров (высокие подвижности), резонансном туннелировании, малых временах туннельных переходов, возможности перестройки энергетического спектра.

Тесты к лекции 1.16

1. ***Как называется вероятность прохождения микрочастицы сквозь потенциальный барьер?***

а) коэффициент прозрачности;

б) коэффициент вероятности;

в) коэффициент прохождения.

2. ***Что такое резонансный туннельный эффект?***

а) резкое увеличение вероятности прохождения частицы сквозь барьеры, если ее энергия совпадает с каким-либо размерным уровнем энергии в потенциальной яме, разделяющей барьеры;

б) резкое уменьшение вероятности прохождения частицы сквозь барьеры, если ее энергия совпадает с каким-либо размерным уровнем энергии в потенциальной яме, разделяющей барьеры;

в) резкое увеличение вероятности прохождения частицы сквозь барьеры, если ширина потенциальной ямы кратна периоду кристаллической решетки полупроводника.

3. ***Для чего может использоваться сильнополевая туннельная инжекция по Фаулеру — Нордгейму?***

а) для модификации (целенаправленного изменения) характеристик МДП-структур;

б) для увеличения граничной частоты СВЧ-устройств;

в) для сглаживания пульсаций в источниках питания.

4. ***В чем особенность ВАХ туннельного диода?***

а) наличие области с отрицательным дифференциальным сопротивлением;

б) наличие области с горизонтальным участком;

в) наличие области в виде фигуры Лиссажу.

5. *Что такое резонансно-туннельный диод?*

а) диод с несколькими потенциальными барьерами, в котором реализуются условия для резонансного туннелирования;

б) диод с одним потенциальным барьером, который используется в выпрямительных устройствах;

в) диодная сборка в виде резонансного выпрямительного моста.

6. *Что такое сверхрешетки?*

а) синтезированные кристаллические структуры, в которых кроме периодического потенциала кристаллической решетки имеется другой периодический потенциал, период которого значительно превышает постоянную решетки;

б) полупроводниковая кристаллическая структура в виде решетки с отверстиями, кратными периоду кристаллической решетки;

в) синтезированные кристаллы, которые изменяют периодический потенциал кристаллической решетки при механическом воздействии.

7. *Что такое nipi-сверхрешетка?*

а) сверхрешетка, которую получают из одного и того же типа полупроводника путем попеременного легирования, создавая слои *n*- и *p*-типов;

б) сверхрешетка, которую получают из разных типов полупроводников путем попеременного нанесения слоев *n*- и *p*-типов;

в) сверхрешетка, которую получают путем поочередного нанесения тонких пленок никеля Ni с чередованием диэлектрика.

8. *В чем заключается особенность сверхрешеток?*

а) в возможности искусственно формировать их электронный спектр;

б) в возможности простого определения резонансной частоты проходящей электронной волны;

в) в создании лучших условий для эмиссии донорных атомов.

1.17. ОДНОЭЛЕКТРОННЫЕ УСТРОЙСТВА, УСТРОЙСТВА МОЛЕКУЛЯРНОЙ И СПИН-ЭЛЕКТРОНИКИ

Цель лекции: ознакомление с одноэлектронными устройствами, устройствами молекулярной и спин-электроники.

1.17.1. КУЛОНОВСКАЯ БЛОКАДА

Кулоновской блокадой называется отсутствие тока через туннельный переход при наличии внешнего напряжения, если туннелированию электронов препятствует их кулоновское взаимодействие.

Рассмотрим плоский конденсатор с туннельно-прозрачным диэлектриком. Если перенести заряд электрона с одной обкладки на другую, то конденсатор окажется заряженным. Для переноса заряда требуется энергия ΔE, так как перенос происходит против сил взаимодействия с положительным зарядом, возникающим на обкладке, с которой удален электрон. Эта энергия равна энергии заряженного конденсатора:

$$\Delta E = \frac{q^2}{2C_{\text{к}}},$$

где $C_{\text{к}}$ — емкость конденсатора.

Приложим к конденсатору возрастающее напряжение. При этом до некоторого значения $U = U_k$ туннельный ток через переход не возникает, так как работа сил поля источника напряжения при перемещении электрона с одной обкладки на другую остается меньше работы против кулоновских сил отталкивания. Протеканию туннельного тока препятствует кулоновская блокада. Когда напряжение достигнет значения $U = U_k$, достаточного для преодоления кулоновских сил отталкивания, еще один электрон перейдет с одной обкладки на другую. Напряжение U_k определяется из соотношения

$$\Delta E = qU_k = \frac{q^2}{2C},$$

откуда

$$U_k = \frac{q}{2C}.$$

Электроны туннелируют через переход по одному. Наблюдение эффекта кулоновской блокады возможно в условиях, когда энергия теплового движения электронов недостаточна для преодоления блокады: $\Delta E \gg kT$.

Пусть полупроводниковая квантовая точка связана с двумя резервуарами электронов туннельными переходами (рис. 1.79). Увеличим число электронов в квантовой точке за счет туннелирования в точку одного электрона из левого резервуара. Для этого необходимо передать электрону потенциальную

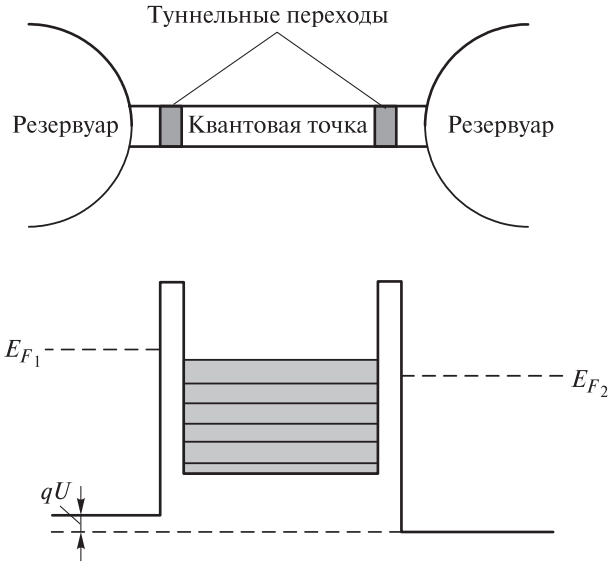


Рис. 1.79. Структура с двумя туннельными переходами

энергию qU от внешнего источника. Потенциальная энергия при этом составляет $Q^2/2C$, где Q, C — исходный заряд квантовой точки и ее емкость соответственно. Следовательно, электрону для попадания в точку необходимо придать потенциальную энергию, бóльшую или равную $q^2/2C$. Необходимое для этого напряжение составляет $q/2C$. Туннелирование невозможно, если

$$|U| < \frac{q}{2C}.$$

Данное неравенство означает, что на ВАХ структуры существует диапазон напряжений (от $-q/2C$ до $q/2C$), в котором ток через точку не может протекать ни при каких условиях (рис. 1.80). При $U = U_k$ электроны туннелируют и проходят через квантовую точку по одному. Пока в точке есть один добавочный электрон, другой не может туда проникнуть из-за кулоновского отталкивания. При увеличении напряжения возможны мультиэлектронные переходы. Но если один из барьеров будет значительно более толстым, то

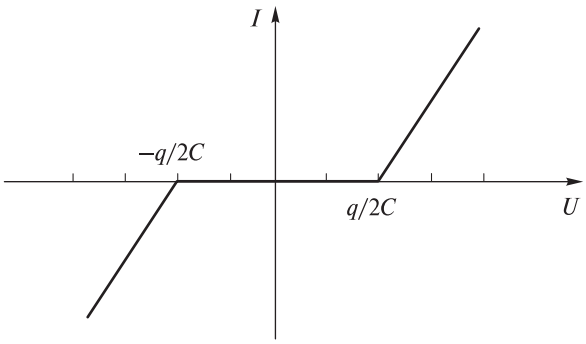


Рис. 1.80. Вольт-амперная характеристика структуры с эффектом кулоновской блокады

электроны снова будут проходить через квантовую точку по одному. Когда один из барьеров значительно толще другого, ВАХ структуры имеет вид, приведенный на рис. 1.81, и называется кулоновской лестницей.

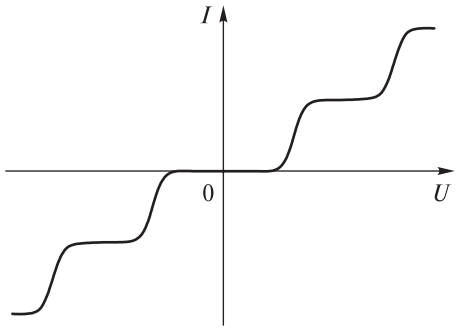


Рис. 1.81. Кулоновская лестница

Значения протекающего через точку тока будут изменяться скачком при значениях приложенного напряжения

$$U = \left(\frac{1}{2C}\right)(2n+1)q, \quad n = 0, 1, 2, \ldots.$$

Подаваемое напряжение измеряется единицами милливольт, одноэлектронный туннельный ток составляет десятки наноампер.

Из полученных уравнений следует, что по мере уменьшения размеров квантовой точки (и соответствующего уменьшения C) возрастает значение энергии, необходимой для изменения числа электронов в квантовой яме.

1.17.2. ОДНОЭЛЕКТРОННЫЙ ТРАНЗИСТОР

Одноэлектронный транзистор содержит область проводимости, соединенную с истоковыми и стоковыми электродами туннельными барьерами, которые имеют емкостную связь с затвором. Изменяя напряжение на затворе, можно управлять током, протекающим между истоком и стоком через область проводимости. Если создать несколько областей с взаимной емкостной связью, то можно реализовать перемещение через эти области дискретных электронов.

Одноэлектронный транзистор работает на эффекте кулоновской блокады. Схема включения одноэлектронного транзистора показана на рис. 1.82. Квантовая точка с полным зарядом Nq подсоединена к истоку и стоку двумя туннельными барьерами. Такая квантовая точка называется кулоновским островом. На сток и затвор подается небольшое положительное напряжение относительно истока.

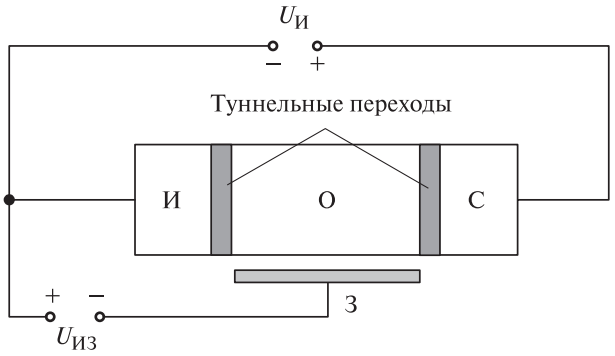


Рис. 1.82. Схема включения одноэлектронного транзистора

На рис. 1.83 приведены энергетические диаграммы, поясняющие работу транзистора. Цифры над уровнями Ферми острова (–1, 1, 2, ...) означают число электронов, перешедших на кулоновский остров с истока. При этом заряд области проводимости составляет $-q, -2q, -3q, \ldots$ соответственно; (–1) означает, что электрон удален с острова. Заряд острова в таком случае равен $+q$.

Если напряжение на затворе равно нулю, кулоновская блокада подавляет туннелирование электронов с истока, поэтому ток между истоком и стоком равен нулю (рис. 1.83, *а*). Поле положительно заряженного затвора смещает уровни острова вниз (уменьшает энергию электронов острова). При $U_З = U_{ЗК}$ кулоновская блокада прорывается, электроны туннелируют с истока на остров, а затем на сток (рис. 1.83, *б*). Между истоком и стоком появляется ток. При дальнейшем увеличении напряжения $U_З$ и смещении уровня (1) вниз снова возникает блокада (рис. 1.83, *в*), ток отсутствует, но на острове находится один добавочный электрон. При $U = U_{З_2}$ ток возникает вновь (рис. 1.83, *г*).

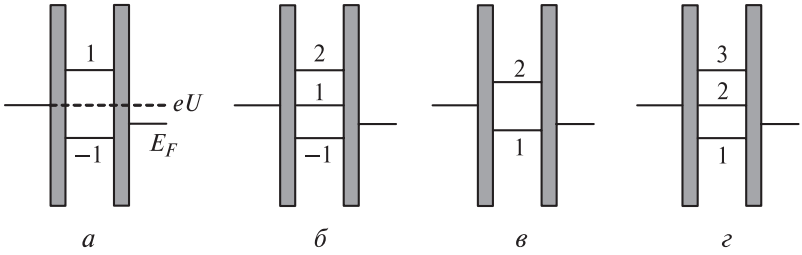


Рис. 1.83. Зонные диаграммы, поясняющие работу одноэлектронного транзистора:
а — $U_З = 0$; *б* — $U_З = U_{ЗК}$; *в* — $U_{ЗК} < U_З < U_{З_1}$; *г* — $U_З = U_{З_2}$

По мере увеличения напряжения на затворе блокада периодически наступает и прорывается, но с каждым разом увеличивается на единицу число электронов на острове. Эти электроны находятся на острове стабильно, а туннелирующие электроны проходят через остров по одному.

На рис. 1.84 показан вид зависимости тока от напряжения на затворе одноэлектронного транзистора. Возникающие скачки тока называются осцилляциями кулоновской блокады. Напряжение $U_З$ имеет порядок единиц милливольт, ток составляет десятки пикоампер.

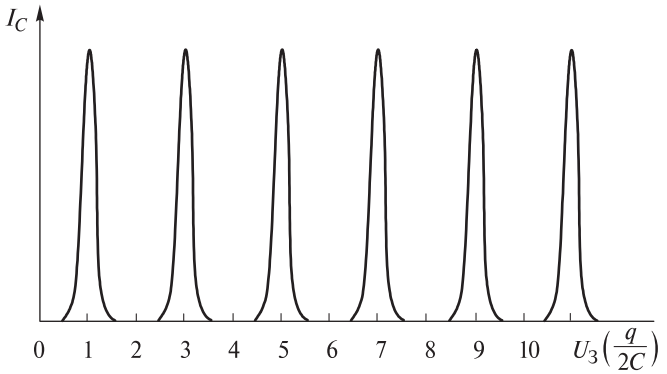


Рис. 1.84. Зависимость тока от напряжения одноэлектронного транзистора

Периодичность пиков напряжения определяется изменением числа электронов, содержащихся на квантовом острове на единицу $\Delta U = q/C$. Это дает возможность определить емкость квантового острова путем измерения разности напряжений на затворе между соседними пиками.

Одноэлектронный транзистор может находиться в состояниях «включено» и «выключено» в зависимости от напряжения на затворе. Поэтому одноэлектронный транзистор можно использовать как переключатель. При прохождении электрона сменяются состояния «включено» и «выключено».

Теоретически время туннелирования равно ~10^{-15} с, этим определяется верхний предел быстродействия прибора (ТГц). Поскольку одноэлектронный транзистор может находиться в двух состояниях, на его основе возможно создание ячеек памяти с малыми размерами, высоким быстродействием и низким энергопотреблением.

Одноэлектронный транзистор, разработанный Кембриджским университетом и компанией Japan Science and Technology Corporation, представлен на рис. 1.85. Размеры острова не должны превышать 10 нм. Высота потенциального барьера равна 0,173 эВ. Материалом для острова служит отдельный кластер аморфного кремния, поверхность которого оксидирована при низкой температуре для создания тонкого барьерного слоя.

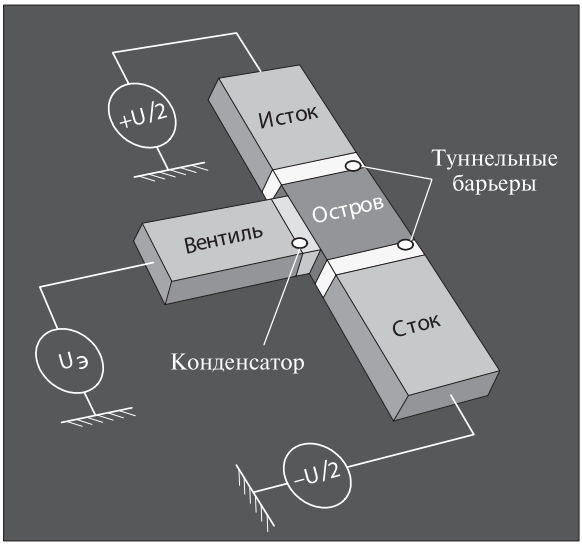


Рис. 1.85. Одноэлектронный транзистор (Japan Science and Technology Corporation)

Одной из перспективных конструкций одноэлектронных устройств является Т-образная транзисторная структура, состоящая из двух одноэлектронных транзисторов, связанных между собой через туннельный конденсатор. На основе Т-образных транзисторных структур могут быть реализованы одноэлектронный коммутатор, одноэлектронный насос, одноэлектронный инвертор и др.

1.17.3. УСТРОЙСТВА МОЛЕКУЛЯРНОЙ ЭЛЕКТРОНИКИ

Молекулярная электроника — наука об электронных свойствах молекулярных материалов и возможности их применения в электронике как элементов и приборов для записи, хранения и передачи информации. В молекулярной электронике выделяют два основных направления: микромолекулярную элек-

тронику (молекулярную электронику — молетронику) и макромолекулярную электронику, или органическую электронику.

Макромолекулярная электроника — электроника, в которой в качестве элементов схем используются тонкие (20...200 нм) пленки органических материалов.

Молекулярная электроника, или молетроника, — электроника, в которой в качестве элементов микроэлектронных схем используются отдельные органические молекулы или даже их фрагменты.

Можно выделить два основных свойства молекулярных материалов:

1) локализацию электронов на молекуле;

2) однозначное расположение атомов в молекуле (в отличие от расположения примесей в кристаллической решетке полупроводников).

Следует отметить, что к молекулярным материалам относятся как неорганические, так и органические материалы, однако последних существует намного больше, поэтому чаще под молекулярными материалами понимают органические материалы.

Главной проблемой при использовании отдельных молекул является отсутствие соответствующей схемотехники. Молекулярные устройства должны представлять собой сложные разветвленные цепи из различных атомных группировок. Методы синтеза таких устройств пока не разработаны.

Отдельные виды молекул и их совокупности могут выполнять функции диэлектриков, проводников, активных элементов — диодов, транзисторов и логических элементов.

Примером молекул-изоляторов могут служить алканы C_nH_{2n+2}. Связи C–C и C–H образованы локализованными молекулярными σ-орбиталями. Такие молекулы ток не проводят.

Для создания молекулярной электроники необходимо прежде всего иметь материалы с высокой проводимостью. Как известно из теории твердого тела, металлическими свойствами могут обладать вещества с определенной зонной структурой, а именно с частично заполненной верхней разрешенной зоной. Другими словами, для получения хорошей проводимости необходимо иметь дополнительные носители заряда в разрешенных зонах.

Роль таких проводников выполняют длинные стержнеобразные молекулы с чередующимися одинарными и двойными (или тройными) связями углерода. Главной особенностью атома C (так же, как и O и N) является способность к гибридизации и образованию соединений с двойными и тройными связями, в которых присутствует π-связь. В таких молекулах внешние, так называемые молекулярные, π-орбитали электронов атомов углерода распространяются по всей молекуле (делокализованные орбитали). Это и обеспечивает возможность переноса электронов проводимости вдоль всей молекулы, что обусловливает ее высокую электропроводность.

В 1974 г. ведущие инженеры фирмы IBM А. Авирам и М. Ратнер предложили модель выпрямителя (диода), состоящего из одной органической молекулы. Две половинки этой молекулы обладают противоположными свойствами по отношению к электрону: одна может только отдавать электрон

(донор), а другая — только принимать (акцептор). Если поместить такую асимметричную молекулу между двумя металлическими электродами, то вся система будет проводить ток только в одном направлении.

Аналогом *p*–*n*-перехода в данном случае является донорно-акцепторная молекула. Модельная молекула, состоящая из акцептора А (тетрацианохинодиметан), донора Д (тетратиофульвален) и соединительной системы метиленовых мостиков С представлена на рис. 1.86.

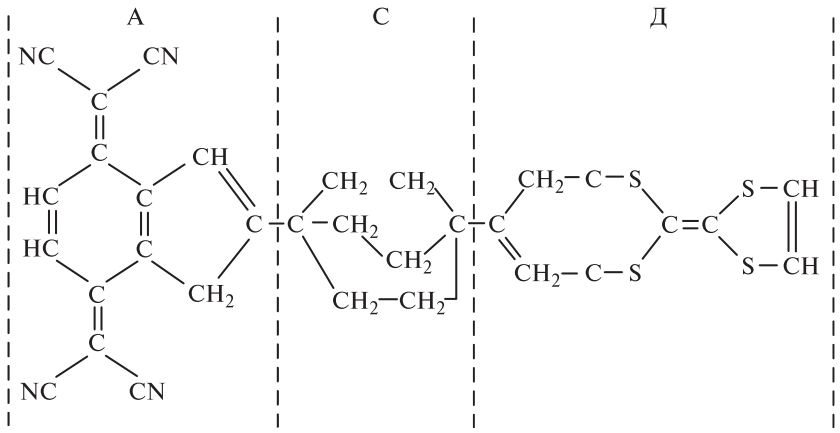


Рис. 1.86. Молекула, обладающая выпрямляющими функциями

Акцептор А имеет низколежащую свободную орбиталь π_1, донор Д — свободную орбиталь с более высокой энергией π_2. Донор и акцептор разделены изолирующей подгруппой. Молекула помещена между двумя металлическими электродами с уровнями Ферми E_{F_1} и E_{F_2} (рис. 1.87). Если на электрод *2* подать «минус», а на электрод *1* — «плюс», то уровень E_{F1} поднимется и электроны будут переходить из металлического электрода *1* на свободный уровень π_1. Уровень E_{F_2} понизится, на него будут уходить элек-

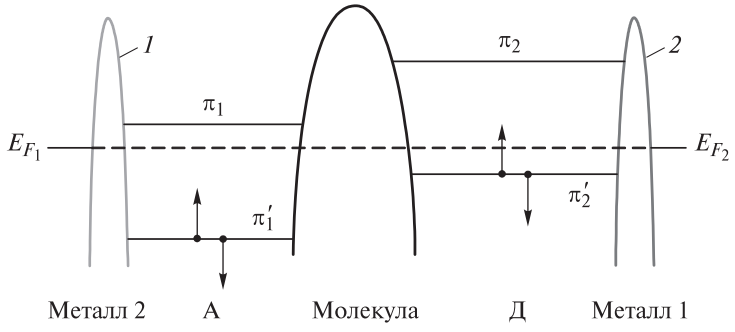


Рис. 1.87. Энергетическая диаграмма выпрямителя в равновесном состоянии

троны с занятого уровня π_2. На освободившийся уровень туннелируют электроны, перешедшие на уровень π_1. Ток течет от электрода *1* к электроду *2*, электроны при движении понижают энергию на каждой ступени.

Если изменить полярность электродов, то на свободный уровень π_2 электроны из проводника *2* перейдут при значительно бóльшем напряжении, так как уровень π_2 расположен высоко. Таким образом, реализуется односторонняя проводимость, т. е. рассмотренная структура обладает выпрямляющими свойствами.

В настоящее время существует несколько вариантов молекулярных транзисторных устройств — аналогов полевого транзистора и способов их изготовления:

– молекулярный одноэлектронный транзистор, в котором квантовым островом служит ион кобальта, между ионом и золотыми электродами имеются туннельные барьеры;

– на основе молекулы C_{60} и туннельного микроскопа;

– на основе одностеночной полупроводниковой нанотрубки и кремниевой подложки.

Существуют классы молекул, которые могут находиться в двух стабильных состояниях с различными значениями электропроводности. Такие молекулы могут быть использованы в качестве переключателей, логических элементов или элементов памяти.

Молекулярный логический элемент «НЕ–И» с двумя входами показан на рис. 1.88, *а*. Электрон будет проходить через молекулярную цепочку только

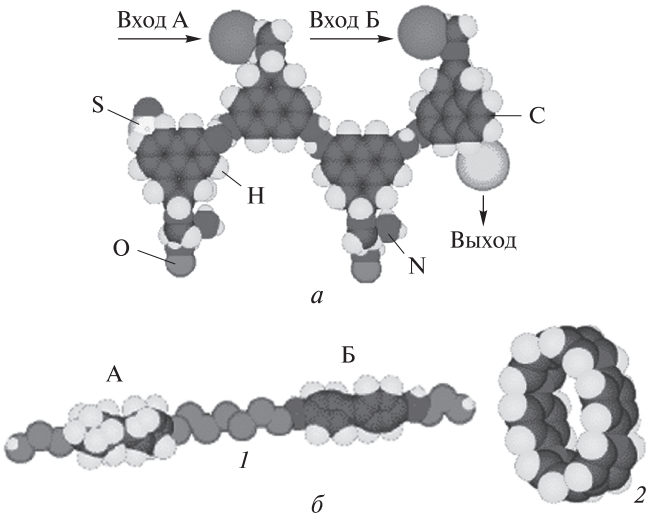


Рис. 1.88. Молекулярные логические элементы:
1 — линейный компонент; *2* — циклический компонент бусины;
C — атом углерода; H — атом водорода; N — атом азота;
S — атом серы; O — атом кислорода
(http://nauka.relis.ru/07/0012/07012064.htm)

в те моменты, когда сигнал на обоих входах отсутствует. Логический элемент на основе молекул ротаксана представлен на рис. 1.88, *б*. В рабочем положении бусина нанизана на нить. Переключение элемента проводится за счет перемещения бусины из положения А в положение Б (и обратно).

Имея молекулы — проводники, изоляторы, диоды, транзисторы, логические элементы и переключатели, можно разрабатывать молекулярные интегральные схемы. Размер молекулярного транзистора равен ~1 нм. Если создать ИМС из 10^9 таких транзисторов, то она будет размером с песчинку. При этом ее производительность возрастет в 10^2–10^3 раз, а энергопотребление уменьшится до весьма малых значений.

Теоретические основы молетроники уже достаточно хорошо разработаны, и созданы прототипы практически всех элементов логических схем, однако на пути их реального использования встают значительные сложности. Главная проблема молекулярной электроники — интеграция молекул в схему. Молекулярное устройство должно представлять собой сложные разветвленные цепи из атомных группировок. Подходы к созданию базовых элементов схем хорошо разработаны, но проблема их интеграции в порядке, обеспечивающем работу схемы, еще далека от решения. Основным путем решения данной проблемы является процесс самосборки, основанный на молекулярном распознавании взаимно дополняющих структур. В настоящее время разработаны технологии некоторых простых процессов самосборки: формирование упорядоченных самоорганизованных пленок; синтез по методу Мэррифилда, в котором соединяются «выходы» одних молекул с «входами» других; получение трехмерных молекулярных структур типа «решеток», «лестниц» и крестообразных структур (из молекул ДНК).

1.17.4. САМОСБОРКА И САМООРГАНИЗАЦИЯ

Классический процесс проектирования предполагает, что сначала выполняют чертеж детали. Затем в соответствии с чертежом удаляют из металлической заготовки все лишнее либо вручную, либо при помощи металлообрабатывающего станка. Такая технология получила название «сверху-вниз». В недалеком будущем нанотехнологии позволят выполнять изделия из материалов с нуля, т. е. начиная с атомов или молекул. Такая технология называется «снизу-вверх». Следует отметить, что складывать атом к атому «вручную» не будут, предполагается, что возможно использовать явление самоорганизации, самосборки наноструктур и наноустройств.

Самосборка (*self-assembly*) — это термин для описания процессов, в результате которых неорганизованные системы благодаря специфическому, местному взаимодействию компонентов систем приходят к упорядоченному состоянию. Самосборка — реализация метода распределенного управления, при котором управляющие функции реализованы во внутренней структуре элементов, участвующих в процессе, а управляющая информация, детерминирующая процесс, распределена по всем элементам.

Самоорганизация в классических терминах может быть описана как спонтанная и обратимая организация молекулярных единиц в упорядоченную структуру с помощью нековалентных взаимодействий. Спонтанность означает, что взаимодействия, ответственные за образование самособранной системы, проявляются в локальных масштабах. Другими словами, наноструктура строит саму себя, подобно тому как это делают живые организмы.

Узнавание структур для последующей самосборки должно быть основано на комбинации так называемых узнающих модулей, что следует из геометрической комплементарности и химического узнавания. Сборка систем происходит, как правило, без использования исходных матриц. Самосборка идет за счет многоточечных нековалентных взаимодействий.

1.17.5. УСТРОЙСТВА СПИН-ЭЛЕКТРОНИКИ

Спинтроника — направление наноэлектроники, в котором для использования в электронных устройствах применяются не только заряды электрона в твердом теле, но и его собственный квантово-механический момент — спин.

Спин — собственный момент количества движения элементарной частицы, имеющий квантовую природу и не связанный с перемещением частицы как целого. Спин электрона может принимать одно из двух значений: $+1/2$ и $-1/2$. Спин имеет размерность (Дж·с), но так как он квантован и кратен постоянной Планка, его обычно изображают в виде целого (или полуцелого, как в случае с электроном) числа (рис. 1.89). Несмотря на то что спин был искусственно введен для объяснения экспериментальных данных, его существование теоретически обосновано в рамках развития релятивистской квантовой механики — наличие спина является одним из следствий уравнения Дирака для волновой функции свободной частицы.

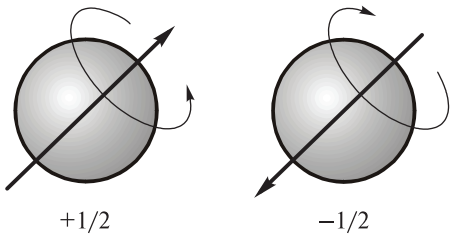


Рис. 1.89. Схема электронов с различными спинами

Основу спин-электроники составляют два основных эффекта, являющиеся следствием особенностей транспорта заряда, контролируемого спином электронов в наноразмерных структурах: гигантское магнитосопротивление и туннельное магнитосопротивление.

Гигантское магнитное сопротивление (ГМС). Наличие у электрона спина обусловливает не только его электрические, но и магнитные свойства. Соответственно, электроны с различным спином движутся в магнитном поле по-разному.

Эффект ГМС основан на том, что электроны с различным направлением спина (и, соответственно, собственного магнитного момента) под влиянием внешнего магнитного поля движутся по-разному. Магнитная головка ГМС состоит из четырех слоев (рис. 1.90). Верхний слой, состоящий из антиферро-

магнетика, называется обменным (*exchange*) и предназначен для того, чтобы зафиксировать магнитное поле второго слоя, который так и называется — иксирующим (*pinned*). Для обеспечения необходимых магнитных свойств второй слой выполняют из ферромагнетика (сплавы никеля, железа, кобальта). Магнитное поле фиксирующего слоя всегда направлено в одну сторону. Третий слой — проводящий (*conducting*) — обычно медный — служит для разделения фиксирующего и чувствительного слоев. Последний, чувствительный (*sensing*) слой также выполняют из ферромагнетика. В отличие от фиксирующего слоя, направление магнитного поля в нем определяется внешним магнитным полем. В данном случае это магнитное поле ячейки жесткого диска, которая содержит один бит информации. В зависимости от состояния ячейки изменяется ориентация магнитного поля в чувствительном слое.

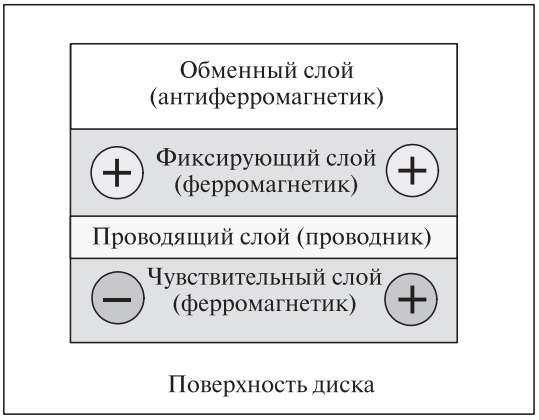


Рис. 1.90. Схематическое изображение ГМС головки

Действие считывающих ГМС-головок основано на фиксировании ГМС-структуры. Если ориентации магнитного поля в чувствительном и в фиксирующем слоях очень малых изменений внешних полей, например, при переходе от магнитного домена одной ориентации к домену противоположной ориентации, совпадают, то сопротивление сенсора уменьшается до минимального значения. Это обусловлено тем, что те электроны, спины (и, соответственно, магнитные моменты) которых совпадают с направлением магнитного поля, не испытывают значительного сопротивления и легко проходят сквозь все слои. Электроны, спины которых направлены против магнитного поля, испытывают значительное сопротивление в обоих ферромагнитных слоях.

Если ориентации магнитного поля в чувствительном и в фиксирующем слоях противоположны друг другу, то независимо от ориентации спина электроны будут испытывать значительное сопротивление в одном из ферромагнитных слоев. Этот эффект имеет чисто квантовую природу, так как его существование напрямую связано с присутствием у электрона спина. Взаимодействие спинов электронов с магнитным полем обеспечивает магнитным головкам на основе эффекта ГМС в 2 раза большую чувствительность, чем традиционным магнитным головкам.

Намагниченность ГМС-элемента после переключения сохраняется до нового переключения, поэтому его можно использовать как элемент памяти. Такие элементы применяют для ячеек магнитной памяти произвольного доступа MRAM (рис. 1.91). Шина записи с помощью тока перемагничивает

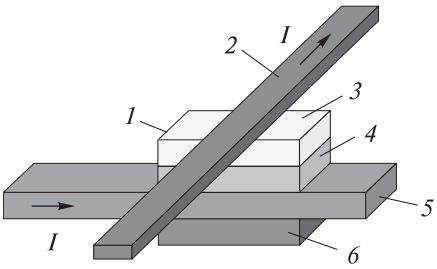


Рис. 1.91. Ячейка MRAM (размер около 200×200 нм)на основе ГМС-элементов:
1 — шина считывания; *2* — шина записи; *3* — изолятор; *4* — мягкий ферромагнетик (Fe); *5* — немагнитный материал (Cu); *6* — жесткий ферромагнетик (Co)

магнитомягкий материал, записывая бит, а шина считывания проходит через немагнитный материал и по сопротивлению определяет состояние бита.

Магнитная память произвольного доступа MRAM на основе ГМС-элементов имеет следующие основные преимущества: энергонезависимость, большая скорость доступа, субмикронный размер, сравнимый с DRAM и SRAM.

Туннельное магнитосопротивление (ТМС). Туннельное магнитосопротивление — эффект спин-зависимого туннелирования электронов через нанометровый слой диэлектрика или полупроводника, помещенный между двумя ферромагнетиками (рис. 1.92). Электроны туннелируют слой диэлектрика, создавая ток, проходящий из одного ферромагнетика в другой.



Рис. 1.92. Структура, обладающая ТМС

В таких структурах наряду с туннельным эффектом используется эффект различного движения электронов с разными спинами в магнитном поле. Если магнитное поле в обоих ферромагнитных слоях клапана совпадает по направлению, то те электроны, спин которых направлен по вектору магнитного поля, легко туннелируют из одного ферромагнитного слоя в другой. Таким образом, сопротивление структуры электрическому току оказывается низким. Если магнитное поле в этих слоях противонаправлено, то туннельный переход затрудняется, так как в одном из слоев электрон будет двигаться против магнитного поля. Вследствие этого сопротивление клапана возрастает на 20...40 %.

На использовании эффекта ТМС основано перспективное направление — магнитная оперативная память MRAM (*MagneticRandom Access Memory*). Ее особенность в том, что информация в ней сохраняется не в виде электрических зарядов на обкладках конденсатора, как в динамической памяти DRAM, и не в виде состояния триггера, как в статической памяти SRAM, а в виде намагниченности слоя ферромагнетика. В качестве достоинств такого типа памяти можно отметить весьма малое энергопотребление, высокое быстродействие, гигантскую плотность хранения информации. Поддержание состояния намагниченности вообще не требует затрат энергии, как это происходит в жестких дисках. Высокая плотность хранения обусловлена небольшими

размерами элементарной ячейки, для которой не требуется большого числа транзисторов, как для SRAM, и отдельных схем регенерации, как для DRAM.

Спиновый транзистор. В спиновом транзисторе состояния «включен» и «выключен» зависят от направления спинов электронов, участвующих в токе. Любое спин-электронное устройство, в том числе и спиновый транзистор, должно содержать три основных элемента:

1) инжектор — механизм для электрического инжектирования спин-поляризованных электронов в полупроводник;

2) средства для управления спиновым током в полупроводнике;

3) детектор — электрическую схему для прецизионного детектирования (измерения) результирующего спинового тока.

Показано, что кремний в настоящее время может быть использован для совершения многочисленных манипуляций над спином на масштабе в несколько сот микрометров и в течение времени, достаточного для осуществления нескольких тысяч логических операций (десятки наносекунд), что обусловливает в перспективе использование спин-электронных систем на базе кремния.

Вид кремниевого спинового транзистора, принцип действия и зонная диаграмма (диаграмма энергетических состояний барьеров, которые встречают электроны при прохождении через вещества) приведены на рис. 1.93.

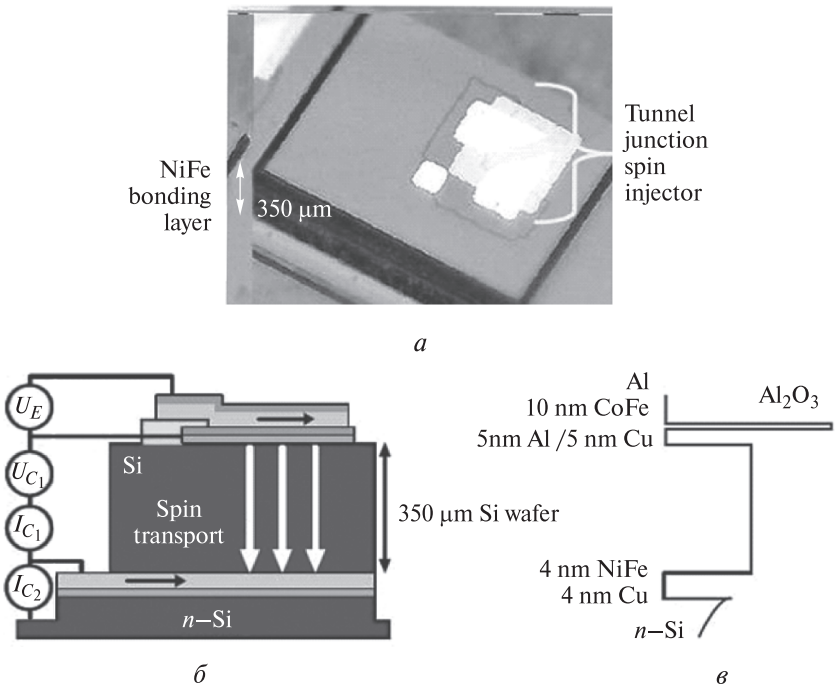


Рис. 1.93. Устройство кремниевого спинового транзистора:
а — изображение; *б* — схема работы; *в* — зонная диаграмма компонентов
(Physical Review Letters 99, 177209, 2007)

На первом этапе при приложенном напряжении U_E неполяризованные электроны инжектируются из алюминиевого эмиттера (источника) в ферромагнитный слой $Co_{84}Fe_{16}$. Благодаря спин-зависимому рассеиванию электронов в магнитном слое электроны с выделенным направлением спина (например, «спин-вниз») отсеиваются, так как направление намагниченности слоя $Co_{84}Fe_{16}$ не совпадает с направлением спинов. Отобранные электроны с однонаправленными спинами туннелируют через тонкий слой Al_2O_3. В данном случае туннельный барьер проходят только «горячие» электроны (с энергией, достаточно высокой для преодоления энергетических барьеров), создавая эмиттерный ток (ток источника). «Горячие» электроны необходимы для увеличения эффективности прибора. После прохождения барьера Шоттки (потенциальный барьер, возникающий на границе металл–полупроводник) в беспримесный монокристаллический слой кремния электроны занимают свободные места в зоне проводимости полупроводника и под действием приложенного к нему напряжения U_{C_1} начинается их упорядоченное движение. При этом возникает коллекторный ток I_{C_1} (ток на детекторе). После прохождения через слой кремния толщиной 350 мкм спин-поляризованные электроны детектируются вторым спиновым транзистором. Ферромагнитный слой $Ni_{80}Fe_{20}$ регистрирует спины электронов, которые инжектируются в кремний *n*-типа (основными носителями тока в котором являются электроны) для увеличения чувствительности детектора (в зоне проводимости *n*-типа кремния имеются избыточные электроны, которые усиливают спиновый ток), создавая коллекторный ток I_{C_2}. Спиновый ток зависит от относительной намагниченности обоих ферромагнитных слоев.

Механизм работы детектора приведен на рис. 1.94. При параллельном направлении намагниченностей в слоях $Co_{84}Fe_{16}$ и $Ni_{80}Fe_{20}$ (рис. 1.94, *а*) ток выше, чем при антипараллельном их направлении (рис. 1.94, *б*).

Развитие спинтроники должно существенно повысить быстродействие компьютеров и плотность записи информации.

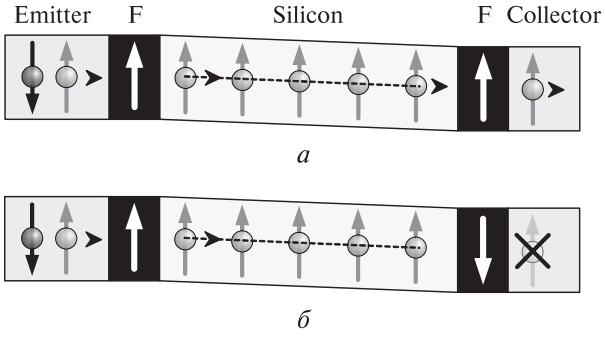


Рис. 1.94. Механизм работы инжектора и детектора: *а* — слои $Co_{84}Fe_{16}$ и $Ni_{80}Fe_{20}$ намагничены параллельно; *б* — антипараллельно (emmiter — источник тока; *F* — первый и второй ферромагнитные слои; silicon — кремниевая прослойка; collector — приемник спинового тока) (Physical Review Letters 99, 177209, 2007)

Тесты к лекции 1.17

1. ***Что называется кулоновской блокадой?***

а) отсутствие тока через туннельный переход при наличии внешнего напряжения, если туннелированию электронов препятствует их кулоновское взаимодействие;

б) явление отталкивания одноименных зарядов;

в) кулоновское взаимодействие электрически заряженных частиц.

2. ***При каких условиях возможно наблюдение эффекта кулоновской блокады?***

а) в условиях, когда энергия теплового движения электронов недостаточна для преодоления блокады;

б) в условиях, когда энергия теплового движения электронов больше энергии блокады;

в) в условиях наличия сильных магнитных полей.

3. ***На каком эффекте работает одноэлектронный транзистор?***

а) на эффекте кулоновской блокады;

б) на эффекте прямого перехода;

в) на эффекте обратного перехода.

4. ***Как может использоваться одноэлектронный транзистор?***

а) как переключатель;

б) как усилитель;

в) как генератор.

5. ***Что такое молекулярная электроника?***

а) наука об электронных свойствах молекулярных материалов и возможности их применения в электронике как элементов и приборов для записи, хранения и передачи информации;

б) технология органической химии, на основе которой строятся различные электронные устройства;

в) технология биофизики, использующая органические материалы как конструкционные.

6. ***В чем заключается главная проблема молекулярной электроники?***

а) в интеграции молекул в схему;

б) в отсутствии базовых логических элементов;

в) в отсутствии токопроводящих органических материалов.

7. ***Что такое спинтроника?***

а) направление электроники, в котором для использования в электронных устройствах применяются не только заряд электрона в твердом теле, но и его собственный квантово-механический момент;

б) направление электроники, в котором для использования в электронных устройствах применяется спинальный эффект;

в) направление электроники, в котором для использования в электронных устройствах применяется только заряд электрона в твердом теле.

8. ***Что такое спин электрона?***

а) собственный момент количества движения элементарной частицы, имеющий квантовую природу и не связанный с перемещением частицы как целого;

б) скорость движения элементарной частицы вокруг своей оси, имеющая квантовую природу и не связанная с перемещением частицы как целого;

в) частота вращения элементарной частицы, имеющая квантовую природу и не связанная с перемещением частицы как целого.

9. ***На чем основан эффект гигантского магнитного сопротивления?***

а) электроны с различным направлением спина (и, соответственно, собственного магнитного момента) под влиянием внешнего магнитного поля движутся по-разному;

б) электроны с различным направлением спина (и, соответственно, собственного магнитного момента) под влиянием внешнего магнитного поля меняют свое направления движения на противоположное;

в) электроны с различным направлением спина (и, соответственно, собственного магнитного момента) под влиянием внешнего магнитного поля осуществляют эмиссию в вакуум.

10. ***Что такое туннельное магнитосопротивление?***

а) эффект спин-зависимого туннелирования электронов через нанометровый слой диэлектрика или полупроводника, помещенный между двумя ферромагнетиками;

б) эффект туннельного отражения электронов от зеркального слоя диэлектрика или полупроводника, помещенный между двумя ферромагнетиками;

в) эффект туннелирования дырок через нанометровый слой диэлектрика или полупроводника, помещенный между двумя ферромагнетиками.

1.18. ФИЗИЧЕСКИЕ ОСНОВЫ НИЗКОРАЗМЕРНЫХ СТРУКТУР

Цель лекции: ознакомление с основными понятиями низкоразмерных структур.

Начало нанотехнологической эре положил в 1959 г. Р. Фейнман в лекции There's Plenty of Room at the Bottom («Там внизу — много места»). Понятие «нанотехнология» было введено в обиход в 1974 г. Н. Танигучи (Япония).

Наноинженерия (*nanotechnologica lengineering*) — междисциплинарная область фундаментальной и прикладной науки и техники, предметом которой являются исследования, проектирование и совершенствование методов производства и применения интегрированных систем, основанных на законах и принципах нанотехнологий и микро- и наносистемной техники.

Наноэлектроника — область современной электроники, занимающаяся разработкой физических и технологических основ создания интегральных электронных схем и устройств на их основе с размерами элементов менее 100 нм. В табл. 1.1 приведено место нанообъектов в окружающем мире. Компоненты больших электронных схем (БИС) имеют размер менее 100 нм и относятся к наномиру.

Таблица 1.1

Место нанообъектов в окружающем мире

Размерная область (характерный размер)	Объект
Макромир (10 м)	Кашалот
1–2 м	Человек
10 см = 10^{-3} м	Страусиное яйцо
1 см = 10^{-2} м	Пчела
1 мм = 10^{-3} м	Песчинка
Микромир (100 мкм = 10^{-4} м)	Толщина полиэтиленовой пленки, яйцеклетка
10 мкм = 10^{-5} м	Толщина человеческого волоса
1 мкм = 10^{-6} м	Эритроцит крови, кишечная палочка
Наномир (100 нм = 10^{-7} м и менее)	Размер компонентов БИС
10 нм = 10^{-8} м	Вирус
1 нм = 10^{-9} м	Белковая молекула, диаметр спирали ДНК

Базисом для возникновения нанотехнологий стали многолетние исследования физиков, химиков, механиков, материаловедов, биологов и специалистов в других областях науки. Для того чтобы понять, почему наномир начинается от 100 нм и далее в сторону уменьшения размера, необходимо рассмотреть зависимость изменения числа атомов на поверхности от размера частиц, составляющих данное количество материала (рис. 1.95).

Как видно на рис. 1.95, при уменьшении размера менее 100 нм происходит резкое возрастание числа поверхностных атомов, вклад которых в свойства нанообъекта становится определяющим. Поверхностные атомы — это, как правило, атомы с оборванными ковалентными связями, поэтому они намного более активны, чем атомы внутри частицы. Кроме того, на данном размерном уровне начинает проявляться действие законов квантовой механики. Все это в комплексе обусловливает совершенно иные физико-химические свойства частиц размером менее 100 мкм по сравнению с макро- и микрочастицами.

Наиболее характерными особенностями наноматериалов являются:

– появление нетрадиционных видов симметрии структуры и особых видов сопряжения границ раздела фаз;

– ведущая роль процессов самоорганизации в структурообразовании, доминирующих над процессами искусственного упорядочения;

– высокая полевая активность и каталитическая избирательность поверхности наночастиц и их ансамблей;

– особый характер протекания процессов передачи энергии, заряда и пространственных организаций, отличающихся низким энергопотреблением, высокой скоростью и наличием синергетических признаков и ярко выраженной анизотропией свойств.

На рис. 1.96 представлена классификация наноматериалов по Р. Зигелю.

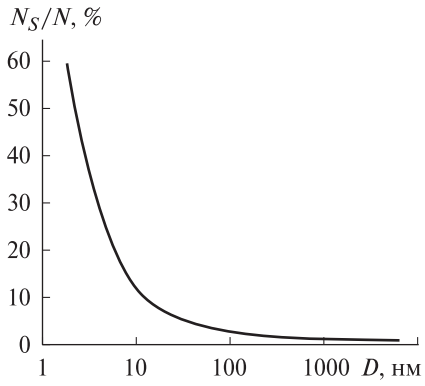


Рис. 1.95. Зависимость изменения числа атомов на поверхности N_S от размера частиц D, составляющих данное количество материала (N – общее число атомов частицы)

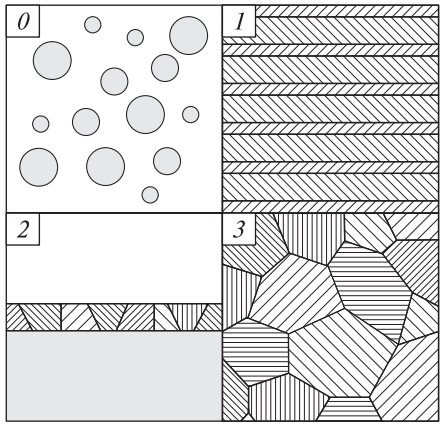


Рис. 1.96. Классификация наноматериалов по Р. Зигелю:

0 — атомные кластеры и наночастицы; *1* — многослойный материал; *2* — наноструктурное покрытие; *3* — объемные наноструктурные материалы

Наноматериалы могут быть очень мелкими частицами в виде порошков или коллоидных взвесей, многослойных материалов, каждый слой которых имеет наноразмерную толщину. Также это могут быть различные покрытия с наноструктурированным поверхностным слоем внутри какого-либо материала, например, сплава металлов, могут образовываться объемные наноструктурные материалы.

Физические свойства наноматериалов в зависимости от уменьшения размера структурного элемента:

– фазовые превращения — понижение температуры фазовых превращений, в том числе температуры плавления;

– кинетические — аномально высокие значения коэффициентов диффузии, повышение теплоемкости, снижение теплопроводности;

– электрические — повышение электрического сопротивления, возрастание диэлектрической проницаемости;

– магнитные — возрастание коэрцетивной силы, магнитосопротивления, появление супермагнетизма;

– механические — повышение предела текучести, твердости, вязкости разрушения, износостойкости, проявление сверхпластичности при высоких температурах.

Уменьшение размера структурного элемента какого-либо материала приводит к значительным изменениям свойств этих материалов, что позволяет использовать их для создания различных технических устройств с улучшенными параметрами.

Тесты к лекции 1.18

1. ***Что такое наноэлектроника?***

а) область современной электроники, занимающаяся разработкой физических и технологических основ создания интегральных электронных схем и устройств на их основе с размерами элементов менее 100 нм;

б) область современной электроники, занимающаяся разработкой физических и технологических основ создания интегральных электронных схем и устройств на их основе с размерами элементов менее 10 нм;

в) область современной электроники, занимающаяся разработкой физических и технологических основ создания интегральных электронных схем и устройств на их основе с размерами элементов менее 1 нм.

2. ***Что такое нанолуковица?***

а) сферическая наночастица с концентрической множественной структурой раковины;

б) куполообразная наночастица с множеством ячеек внутри;

в) гибридная наночастица в виде стрелки лука.

3. ***Что происходит с наночастицей при уменьшении ее размера менее 100 нм?***

а) резкое возрастание числа поверхностных атомов;

б) резкое уменьшение числа поверхностных атомов;

в) резкое уменьшение ее объема.

4. ***На каком размерном уровне начинает проявляться действие законов квантовой механики?***

а) на размерном уровне менее 100 нм;

б) на размерном уровне менее 100 мкм;

в) на размерном уровне менее 100 мм.

5. ***Какие науки являлись базисом для возникновения нанотехнологий?***

а) физика, химия, механика, материаловедение, биология;

б) геофизика, геохимия, гидравлика, электротехника;

в) лингвистика, археология, ботаника.

1.19. УГЛЕРОДНЫЕ НАНОСТРУКТУРЫ

Цель лекции: ознакомление с углеродными наноструктурами.

1.19.1. ОБЩИЕ СВЕДЕНИЯ ОБ УГЛЕРОДНЫХ НАНОТРУБКАХ

Развитие нанотехнологии в значительной степени связано с открытием, изучением и уже начавшимся практическим использованием трех углеродных наноструктур: фуллеренов, углеродных нанотрубок и графена (рис. 1.97).

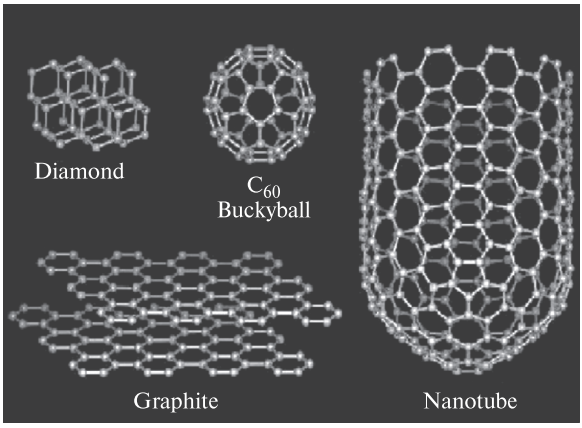


Рис. 1.97. Углеродные наноструктуры (https://chemintinity.com/sducation/614/)

Углерод имеет следующие аллотропные формы: алмаз, графит, карбин (3D); графен (2D); УНТ (1D); фуллерены (0D).

Интерес к исследованию подобных структур обусловлен, с одной стороны, их уникальными физико-химическими характеристиками, с другой — значительными перспективами прикладного применения.

1.19.2. УГЛЕРОДНЫЕ НАНОТРУБКИ

Нанотрубка представляет собой цилиндр, полученный при свертывании плоской гексагональной сетки графита без швов (рис. 1.98). Минерал графит имеет слоистую структуру. Каждый слой построен из гексагональных ячеек — шестиугольников, в вершинах которых расположены атомы углерода. Угле-

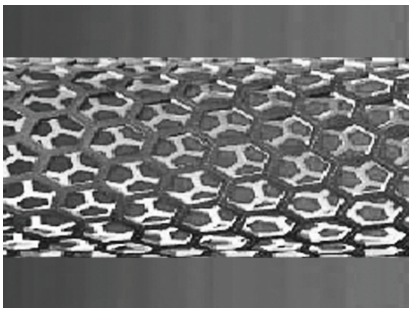


Рис. 1.98. Одностенная УНТ

родные нанотрубки (УНТ) представляют собой однослойные или многослойные полые наноскопические цилиндры диаметром от 0,5 нм.

Если сворачивать трубки относительно различных направлений, то можно получать УНТ с принципиально различающейся структурой и, как следствие, с разными физическими свойствами. Взаимная ориентация гексагональной сетки и оси нанотрубки определяется основной характеристикой нанотрубки, которая называется хиральностью.

Направление сворачивания листа графена задается вектором $\vec{C}$, который описывается набором целых чисел (n, m), являющихся его координатами в базисе локальных векторов $\vec{a}_1$ и $\vec{a}_2$, т. е. $C = na_1 + ma_2$ (рис. 1.99). Ось УНТ перпендикулярна направлению сворачивания листа, т. е. вектору $\vec{C}$. Диаметр УНТ

$$D = \frac{|C|}{\pi} = \frac{a\sqrt{3}}{\pi}\sqrt{n^2 + nm + m^2},$$

где a — расстояние между атомами в гексагональной ячейке, a = 0,142 нм.

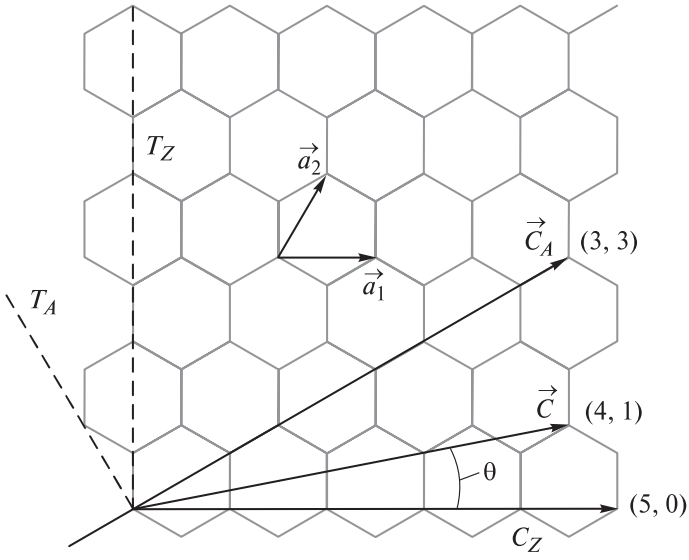
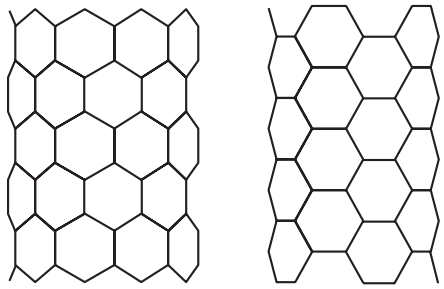


Рис. 1.99. Углы хиральности

Хиральность можно описать также с помощью угла хиральности θ, который задается как угол между векторами $\vec{C}$ и $\vec{a}_1$ (см. рис. 1.99):

$$\theta = \operatorname{arctg}\left(\frac{m\sqrt{3}}{2n+m}\right) = \arccos\left(\frac{2n+m}{2\sqrt{n^2+nm+m^2}}\right).$$

Угол хиральности изменяется от 0 до 30°, и его предельным значениям соответствуют две особые конфигурации нанотрубок. На рис. 1.99 штриховыми линиями показаны направления оси УНТ для этих двух случаев. При $C_z = na_1$ получившаяся трубка будет характеризоваться числами (n, 0). Такие трубки называют зигзагообразными (*zigzag*) (рис. 1.100, *а*), в них ось трубки направлена параллельно одной из пар С–С-связей в шестиугольниках решетки. При $C_A = n(a_1 + a_2)$ ось УНТ перпендикулярна одной из пар С–С-связей. Такие трубки называют кресельными (*armchair*) (рис. 1.100, *б*).



а *б*

Рис. 1.100. Зигзагообразные (*а*) и кресельные (*б*) УНТ

Нанотрубки существуют не только в форме прямых цилиндров. Разработаны методы получения Т-образных и У-образных нанотрубок.

Основными методами получения УНТ являются:

– разрядно-дуговой метод, основанный на термическом распылении графитового анода в плазме дугового разряда, горящего в атмосфере гелия;

– термическое распыление в дуге в присутствии катализаторов;

– лазерное распыление;

– химическое осаждение из пара.

В настоящее время нет способа выращивания нанотрубок с заданной хиральностью; в одном процессе синтеза образуются нанотрубки с различными хиральностями. Диаметр и длину трубок можно варьировать изменением условий синтеза.

Нанотрубки после процесса получения обычно закрыты многослойными полусферическими головками, каждый слой которых состоит из пяти- и шестиугольных ячеек. Химическая активность головки и стенок различна; головки менее устойчивы, поэтому специальной химической обработкой можно удалять головки и получать нанотрубки с открытыми концами.

Однослойные нанотрубки получают специальными методами синтеза. В остальных методах бóльшая часть получаемых трубок многослойные, они содержат от двух до десятков слоев. Нанотрубки обладают хорошими механическими характеристиками. Они в 50–1000 раз прочнее стали и имеют в 6 раз меньшую плотность (1,3...1,4 г·см$^{-3}$). Предел прочности однослойных УНТ, по разным оценкам, составляет 50...150 ГПа. Относительное удлинение УНТ до разрушения 10...15 %, т. е. они обладают и достаточно высокой пластичностью.

Бóльшая часть получаемых трубок многослойные, они содержат от двух до десятков слоев. Многослойные УНТ имеют некоторые отличительные особенности. Считается, что предел прочности и упругость многослойных УНТ выше по сравнению с однослойными, однако их гибкость падает с ростом числа слоев. Внутренняя нанотрубка способна выдвигаться из многослойной УНТ, а затем втягиваться обратно под действием сил Ван-дер-Ваальса подобно пружине. Такое перемещение внутренней УНТ происходит практически без трения. Силы трения, возникающие при вращении внутреннего слоя относительно внешних, также очень малы, что позволяет рассматривать многослойную УНТ как почти идеальный подшипник.

Высокие механические параметры УНТ обусловлены двумя основными факторами. Во-первых, в нанотрубках концентрации дефектов, приводящих к ухудшению механических свойств, значительно ниже по сравнению с традиционными материалами, а во-вторых, УНТ обладают способностью к перестройке углерод-углеродных связей под нагрузкой, в частности при изгибе. За счет последнего эффекта шестиугольные ячейки стенок УНТ могут перестраиваться в пятиугольные или семиугольные, что приводит к изменению не только механических, но и электрических свойств УНТ.

Нанотрубки могут проявлять металлические или полупроводниковые свойства. Проводимость УНТ зависит от угла хиральности. Считается, что чисто металлическими свойствами обладают кресельные УНТ, а остальные, удовлетворяющие приведенному выше соотношению, — полуметаллическими. Электронные состояния последних не образуют одной широкой непрерывной энергетической зоны проводимости, а разбиваются на отдельные подзоны. Ширина запрещенной зоны полупроводниковых УНТ обратно пропорциональна диаметру трубки, поэтому изменение диаметра может привести к снижению или увеличению проводимости УНТ. Многослойные УНТ обычно проявляют металлические и полуметаллические свойства, даже если они состоят из полупроводниковых нанотрубок. Это обусловлено влиянием сил взаимодействия между трубками, которое приводит к значительному сокращению энергетической щели между валентной зоной и зоной проводимости.

Нанотрубки могут выступать в роли полупроводников и использоваться для изготовления транзисторов (рис. 1.101). Ожидается, что плотность размещения транзисторов будет в 10 раз больше, чем в современных чипах, и они также будут более экономичными.

Нанотрубки с металлическим типом проводимости могут пропускать ток плотностью до 10^9 А/см^2, что примерно в 1000 раз больше по сравнению с традиционными металлическими проводниками. Это объясняется тем, что УНТ при не слишком большой их длине являются баллистическими проводниками.

Уникальные свойства нанотрубок делают их одним из основных материалов наноэлектроники. Решение проблемы широкого прикладного использования УНТ во многом зависит от стоимости производства нанотрубок в макроскопических масштабах. Несмотря на высокую стоимость производства УНТ, такие их уникальные свойства, как сверхминиатюрные размеры, высо-

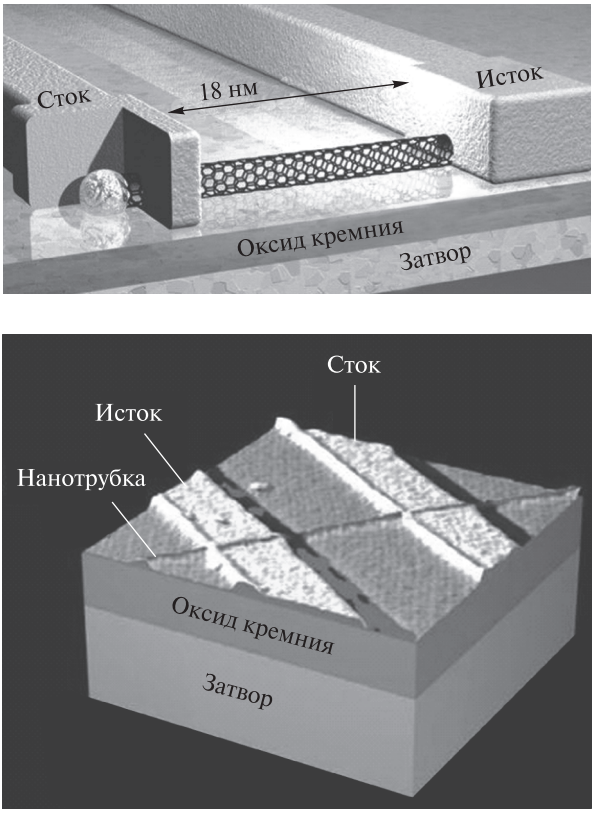


Рис. 1.101. Транзисторы на основе нанотрубок (http://faculty.ifmo.ru/mv/foto/foto/nauka)

кая механическая прочность, устойчивость к химическим воздействиям, хорошая электропроводность и др. указывают на возможность их эффективного применения в областях наноэлектроники, водородной энергетики, химических технологий.

1.19.3. ГРАФЕН

Графен (*graphene*) представляет собой изолированный моноатомный слой графита (рис. 1.102). Это слой атомов углерода, размещенных в узлах гексагональной двухмерной кристаллической решетки. Графен можно рассматривать в качестве исходного материала для построения других углеродных форм.

Графен относится к наиболее прочным материалам и имеет очень высокий коэффициент теплопроводности при комнатной температуре — около $5 \cdot 10^3$ Вт/(м·К), что почти в 15 раз выше по отношению к меди и в 1,5 раза превышает аналогичный показатель для УНТ.

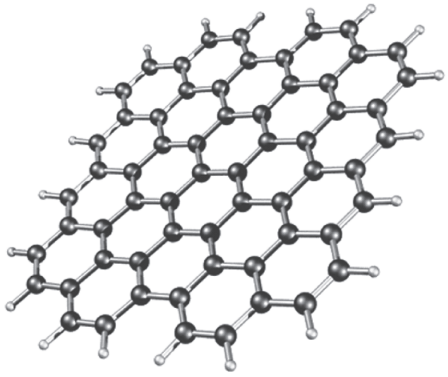


Рис. 1.102. Графен (technolenta.ru)

Электрические свойства графена определяются особенностями его электронной структуры. У графена валентная зона касается зоны проводимости в точках, соответствующих вершинам шестиугольника, т. е. узлам кристаллической решетки. Вблизи этих точек энергия электронов линейно зависит от волнового вектора в отличие от параболического закона дисперсии для периодических структур, согласно которому энергия носителя пропорциональна квадрату волнового вектора. При обычном туннелировании вероятность проникновения частицы в область за барьером уменьшается с ростом его высоты и ширины. В графене электрон или дырка, движущиеся перпендикулярно барьеру, преодолевают его с единичной вероятностью вне зависимости от значений высоты и ширины. Подвижность носителей заряда достигает в графене $(1{,}5–2{,}0)\cdot 10^4$ см/(В·с), что приблизительно на порядок выше по сравнению с кремнием.

Вырезая ленты с кресельной или зигзагообразной структурой, можно изменять их проводимость, т. е. получать ленты с металлической или полупроводниковой проводимостью подобно рассмотренным выше УНТ. Для полупроводниковых графеновых нанолент экспериментально подтверждена обратная зависимость ширины энергетической щели от поперечного размера ленты.

В настоящее время графеновые наноленты рассматриваются как наиболее перспективный заменитель кремния в качестве полупроводникового материала для наноэлектроники. Такие наноленты могут использоваться наряду с УНТ и фуллеренами в качестве добавок при создании новых материалов.

1.19.4. ФУЛЛЕРЕНЫ И ФУЛЛЕРИТЫ

Атомы углерода, испарившиеся с разогретой поверхности графита, соединяясь друг с другом, могут образовывать не только нанотрубки, но и другие молекулы, представляющие собой выпуклые замкнутые многогранники, например в виде сферы или эллипсоида. В этих молекулах атомы углерода расположены в вершинах правильных шести- и пятиугольников, из которых составлена поверхность сферы или эллипсоида.

Все эти молекулярные соединения атомов углерода названы фуллеренами по имени американского инженера, дизайнера и архитектора Р. Фуллера, применявшего для постройки куполов своих зданий пяти- и шестиугольники, являющиеся основными структурными элементами молекулярных каркасов всех фуллеренов. Фуллерены представляют собой шарообразные молекулы, состоящие из 60, 70, 76 (C_{60}, C_{70}, C_{76} соответственно) и более атомов

углерода. Самым изученным является фуллерен C_{60}, который был открыт в 1985 г. в экспериментах по лазерному испарению графитовой мишени. Масс-спектрометрический анализ продуктов испарения показал, что в их составе есть частицы с массовым числом 720, соответствующим 60 атомам С.

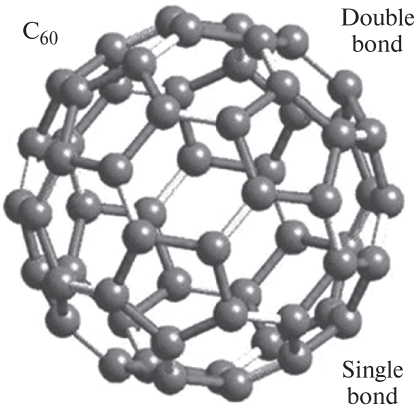


Рис. 1.103. Фуллерен (www4.nau.edu)

Поверхность молекулы C_{60} представляет собой многогранник, состоящий из 20 шестиугольных и 12 пятиугольных граней (рис. 1.103). Диаметр молекулы C_{60} составляет около 0,7 нм, а диаметр ее внутренней полости — около 0,5 нм.

Фуллерен C_{60} способен присоединять атомы многих элементов, которые могут размещаться в его внутренней полости. Это позволяет направленно влиять на свойства фуллерена. Так, при легировании его атомами щелочных металлов он приобретает достаточно высокую электропроводность. Фуллерены находят применение при производстве аккумуляторных батарей; изучаются возможности создания на их основе оптических затворов, устройств для записи информации и элементов солнечных батарей. Молекулы C_{60} пригодны также для применения в качестве добавок при создании различных наноматериалов и присадок для ракетного топлива.

Из фуллеренов при определенных условиях могут образовываться молекулярные кристаллы — фуллериты (рис. 1.104).

Фуллериты имеют низкую плотность (1,7 г/см^3) по сравнению с графитом и алмазом (2,3 г/см^3 и 3,5 г/см^3 соответственно), но их твердость сопоставима с твердостью алмаза. Фуллериты, как и фуллерены, легко поддаются легированию различными элементами, что позволяет, в частности, широко варьировать их электрические свойства. Обычный кристалл фуллерита является полупроводником с шириной запрещенной зоны 1,5 эВ. Легирование щелочными металлами увеличивает проводимость фуллерита на несколько порядков. Получены даже сверхпроводящие кристаллические пленки C_{60}, для которых эффект сверхпроводимости наблюдался при температурах около 10...40 К. Основные направления применения фуллеритов схожи с направлениями применения фуллеренов.

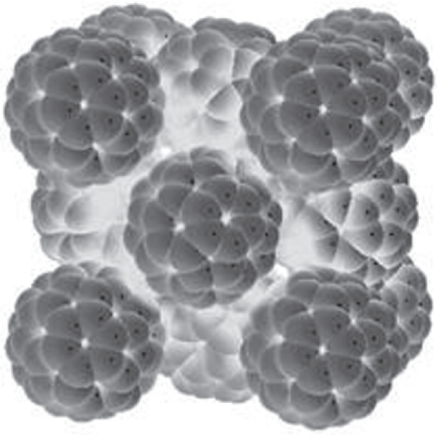


Рис. 1.104. Кристаллическая структура фуллерита (modificator.ru)

Тесты к лекции 1.19

1. ***Что такое нанотрубка?***

а) цилиндр, полученный при свертывании плоской гексагональной сетки графита без швов;

б) цилиндр, полученный при выдавливании графита через пресс-форму;

в) цилиндр, полученный при вырезании из куска графита специальным инструментом.

2. ***Что такое хиральность нанотрубки?***

а) взаимная ориентация гексагональной сетки и оси нанотрубки;

б) возможность совмещаться со своим зеркальным отражением;

в) взаимная ориентация ячеек гексагональной сетки.

3. ***Чем обусловлены высокие механические параметры углеродной нанотрубки?***

а) во-первых, в нанотрубках концентрации дефектов, приводящих к ухудшению механических свойств, значительно ниже по сравнению с традиционными материалами, во-вторых, УНТ обладают способностью к перестройке углерод-углеродных связей под нагрузкой, в частности при изгибе;

б) высокой прочностью атома углерода;

в) высокой прочностью атомарных связей в углеродных структурах.

4. ***Влияет ли угол хиральности на свойства нанотрубок?***

а) да, сильно;

б) нет;

в) да, слабо.

5. ***Что представляют собой фуллерены?***

а) шарообразные молекулы, состоящие из нескольких десятков атомов углерода;

б) крупные органические молекулы;

в) фуллерены представляют собой крупные неорганические молекулы.

1.20. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКОВЫХ СТРУКТУРАХ

Цель лекции: ознакомление с основными термоэлектрическими эффектами в полупроводниках.

1.20.1. ЭФФЕКТ ЗЕЕБЕКА

К термоэлектрическим явлениям относятся три эффекта: эффект Зеебека, эффект Пельтье и эффект Томсона.

Эффект Зеебека, или термоэлектрический эффект, был открыт в 1821 г. немецким физиком Т.И. Зеебеком (Thomas Johann Seebeck). Данный эффект заключается в возникновении электродвижущей силы (ЭДС) в замкнутой электрической цепи, состоящей из одного или нескольких последовательно соединенных разнородных проводников, грани которых находятся при различных температурах (рис. 1.105).

Появление ЭДС обусловлено разным уровнем Ферми в различных материалах. При их соприкосновении уровень Ферми выравнивается (в одном материале понижается, в другом — повышается). Этот процесс происходит за счет перехода электронов через контакт, что и приводит к появлению ЭДС. Следует отметить, что значение ЭДС невелико (порядка нескольких десятых долей вольта). В небольшом интервале температур термоЭДС E можно считать пропорциональной разности температур T_2 и T_1:

$$E = \alpha_{12}\left(T_2 - T_1\right),$$

где $\alpha_{1,2}$ — коэффициент термоЭДС.

Значение коэффициента термоЭДС зависит от природы контактирующих материалов. Для металлов эти значения равны единицам и десяткам мкВ/К, для полупроводников они составляют сотни мкВ/К. Полупроводники

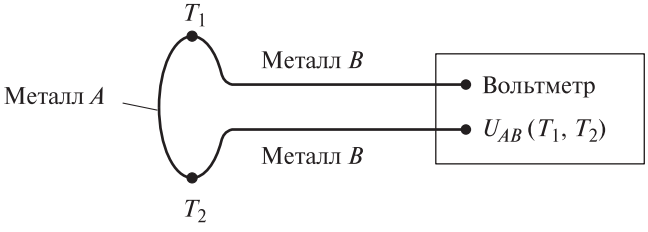


Рис. 1.105. Схема измерения термоЭДС:
T_1, T_2 — различные температуры

обладают на порядок бóльшей термоэлектрической силой, чем металлы. Это обусловлено более сильной зависимостью характеристик полупроводников от температуры (проводимость, концентрация носителей заряда); КПД эффекта Зеебека является чрезвычайно низким (порядка 10 %).

Эффект Зеебека используют в термопарах для измерения температур разных объектов, также применяют для измерения плотности лучистой (электромагнитной) энергии.

На космических аппаратах эффект Зеебека применяется для генерации электричества. В запускаемых зондах для исследования окраин нашей Солнечной системы или космоса за ее пределами используется эффект Зеебека для питания электронных устройств, находящихся на их борту. Это стало возможным благодаря радиационному термоэлектрическому генератору.

1.20.2. ЭФФЕКТ ПЕЛЬТЬЕ

Эффект Пелетье заключается в нагревании одного контакта и охлаждении другого, если рассмотренную выше цепь подсоединить к внешнему источнику тока, т. е. эффекты Зеебека и Пельтье являются противоположными. Эффект Пельтье открыт в 1834 г. французским физиком Ж.Ш. Пельтье (Jean-Charles Peltier).

Значение перемещенной энергии и направление ее переноса зависят от вида контактирующих веществ и от направления и силы протекающего электрического тока:

$$Q = \Pi It,$$

где Q — количество выделенного или поглощенного количества теплоты; Π — коэффициент Пельтье; I — сила тока; t — время протекания тока.

Эффект Пельтье более заметен у полупроводников, это свойство используется в элементах Пельтье (рис. 1.106).

Эффект Пельтье объясняется следующим образом. При контактировании двух разнородных веществ появляется контактная разность потенциалов, которая создает внутреннее контактное электрическое поле. Если через контакт протекает электрический ток, то это поле будет либо способствовать, либо препятствовать прохождению тока. Если ток преодолевает контактное поле, то внешний источник питания затрачивает дополнительную энергию, которая выделяется в контакте, что приводит к его нагреву. Если же ток идет таким образом, что поддерживается контактным полем, то оно и совершает работу по перемещению зарядов. Необходимая для этого энергия отбирается у вещества, что приводит к его охлаждению в месте контакта.

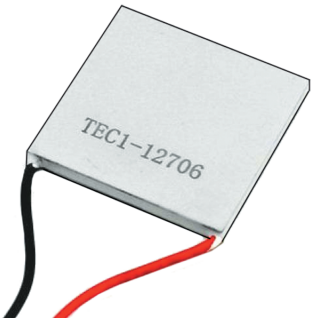


Рис. 1.106. Элемент Пельтье (термоэлектрический модуль)

Основные направления практического использования эффекта Пельтье в полупроводниках — получение холода для создания термо-

электрических охлаждающих устройств, подогрев для целей отопления, термостатирование, управление процессом кристаллизации в условиях постоянной температуры. При пропускании тока через элемент Пельтье одна его сторона будет нагреваться, а другая охлаждаться.

1.20.3. ЭФФЕКТ ТОМСОНА

Эффект открыт У. Томсоном в 1851 г. и относится к термоэлектрическим эффектам. При пропускании электрического тока через проводник или полупроводник, вдоль которого существует градиент температуры (рис. 1.107), в нем, кроме джоулевой теплоты, в зависимости от направления тока будет выделяться или поглощаться дополнительное количество теплоты (теплота Томсона). Количество теплоты Томсона пропорционально силе тока, времени и перепаду температур, зависит от направления тока. Объяснение эффекта заключается в следующем. В условиях, когда вдоль проводника, по которому протекает ток, существует градиент температуры, причем направление тока соответствует движению электронов от горячего конца проводника к холодному, при переходе из более горячего сечения в более холодное электроны передают избыточную энергию окружающим атомам (выделяется теплота), а при обратном направлении тока, проходя из более холодного участка в более горячий, пополняют свою энергию за счет окружающих атомов (теплота поглощается).

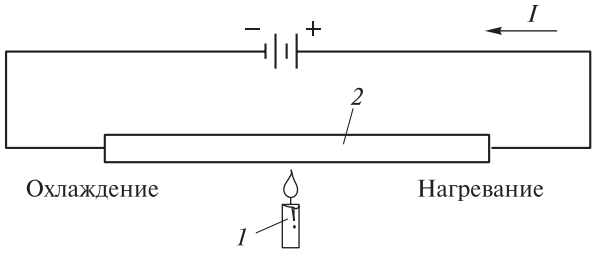


Рис. 1.107. К объяснению эффекта Томсона:
1 — нагреватель; *2* — металлический проводник

В полупроводниках важным является то, что концентрация носителей в них значительно зависит от температуры. Если полупроводник нагрет неравномерно, то концентрация носителей заряда в нем будет больше там, где выше температура, поэтому градиент температуры приводит к градиенту концентрации, вследствие чего возникает диффузионный поток носителей заряда. Это приводит к нарушению электронейтральности. Разделение зарядов порождает электрическое поле, препятствующее разделению. Таким образом, если в полупроводнике имеется градиент температуры, то в нем имеется объемное электрическое поле $E_{об}$.

Предположим, что через такой образец пропускают электрический ток под действием внешнего электрического поля $E_{вн}$. Если ток идет против внут-

реннего объемного поля $E_{об}$, то внешнее поле совершает дополнительную работу при перемещении зарядов относительно поля $E_{об}$, что приведет к выделению количества теплоты, дополнительного к ленц-джоулевым потерям. Если ток (или внешнее поле $E_{вн}$) направлен по объемному полю $E_{об}$, то это поле совершает работу по перемещению зарядов для создания тока. В данном случае внешний источник тратит меньшую энергию для поддержания тока, чем в том случае, когда внутреннего поля $E_{об}$ нет. Работа поля $E_{об}$ может совершаться только за счет тепловой энергии самого проводника, поэтому он охлаждается. Явление выделения или поглощения теплоты в проводнике, обусловленное градиентом температуры, при прохождении тока называется эффектом Томсона. Теплота Томсона (Q_m) пропорциональна плотности тока (J), времени (t) и перепаду температур вдоль проводника (ΔT):

$$Q_m = \tau \Delta T J t,$$

где τ — коэффициент Томсона, зависящий от свойств материала.

Таким образом, вещество нагревается, когда поля $E_{вн}$ и $E_{об}$ противоположно направлены, и охлаждается, когда их направления совпадают.

Тесты к лекции 1.20

1. ***В чем заключается эффект Зеебека?***

а) в возникновении ЭДС в замкнутой электрической цепи, состоящей из одного или нескольких последовательно соединенных разнородных проводников, грани которых находятся при различных температурах;

б) в аномальном повышении сопротивления места контакта двух проводников при их нагреве;

в) в нагреве одного проводника и охлаждении другого при пропускании через их контакт электрического тока.

2. ***Что является причиной возникновения эффекта Зеебека?***

а) разный уровень Ферми в разных материалах;

б) разный уровень проводимости контактируемых материалов;

в) различная плотность фононного газа.

3. ***От чего зависит значение коэффициента термоЭДС?***

а) от природы контактирующих материалов;

б) от температуры;

в) от значения проходящего тока.

4. ***Какое примерно значение имеет КПД эффекта Зеебека?***

а) 10 %;

б) 20 %;

в) 30 %.

5. ***В чем заключается эффект Пельтье?***

а) в нагревании одного контакта и охлаждении другого, если через замкнутую цепь двух или более контактов разнородных проводников пропустить электрический ток;

б) в возникновении ЭДС в замкнутой цепи двух контактов разнородных проводников при их нагреве;

в) в аномальном падении сопротивления в месте контакта двух разнородных проводников при охлаждении.

6. ***В чем заключается эффект Томсона?***

а) при пропускании электрического тока через проводник или полупроводник, вдоль которого существует градиент температуры, в нем кроме джоулевой теплоты в зависимости от направления тока будет выделяться или поглощаться дополнительное количество теплоты;

б) при пропускании электрического тока через два проводника или полупроводника, расположенных параллельно друг другу, между ними возникает сила отталкивания или притяжения в зависимости от направления тока;

в) при пропускании электрического тока через два проводника или полупроводника, расположенных перпендикулярно друг другу, между ними возникает сила отталкивания или притяжения в зависимости от направления тока.

7. ***От каких параметров зависит количество теплоты Томсона?***

а) от силы тока, времени, перепада температур, направления тока;

б) от вида контактируемых материалов;

в) от сопротивления материала.

1.21. ВЫСОКОТЕМПЕРАТУРНАЯ ПОЛУПРОВОДНИКОВАЯ ЭЛЕКТРОНИКА

Цель лекции: ознакомление с полупроводниками, предназначенными для работы при высоких температурах.

1.21.1. ОБЩИЕ СВЕДЕНИЯ О ВЫСОКОТЕМПЕРАТУРНЫХ ПОЛУПРОВОДНИКАХ

Во многих современных динамично развивающихся отраслях гражданской промышленности, военной и космической техники существует потребность в электронике, способной надежно работать в жестких условиях эксплуатации, например при крайне высоких температурах. Как известно, классический подход инженеров при проектировании электроники, которая должна успешно работать в повышенном диапазоне температур, заключается в использовании активного или пассивного охлаждения. Однако в ряде случаев охлаждение оказывается весьма затруднительным или даже невозможным. Более того, в целях повышения надежности или удешевления оборудования работа при повышенной температуре иногда даже предпочтительна.

Для обеспечения надежной работы электронного оборудования при повышенной температуре необходимо контролировать целый ряд ключевых параметров, а также обеспечить решение ряда проблем. Одной из таких наиболее важных и широко известных проблем является рост тока утечки через подложку при повышенных температурах. Ток утечки в таких кристаллах удваивается при росте температуры на каждые 10 °С, что неприемлемо во многих видах электронной аппаратуры.

Рассмотрим физические явления, происходящие в полупроводнике при повышении температуры. Как отмечалось выше, энергетические уровни электронов в полупроводниках и диэлектриках могут находиться в одной из двух зон: валентной зоне или зоне проводимости. Между этими зонами расположена запрещенная зона. Энергетические уровни электронов в запрещенной зоне находиться не могут. Отличие диэлектриков от полупроводников заключается только в ширине запрещенной зоны, считается, что у полупроводников она меньше 3 эВ.

При температуре абсолютного нуля все электроны располагаются в валентной зоне, а в зоне проводимости нет ни одного электрона, поэтому собственный полупроводник не проводит ток. При нагревании полупроводника электроны начинают переходить в зону проводимости. Чем выше температура, тем больше электронов переходит в зону проводимости, именно поэтому сопротивление полупроводника падает, а ток утечки увеличивается.

В собственном полупроводнике концентрация свободных носителей n_i определяется двумя основными параметрами: шириной запрещенной зоны E_g и температурой T:

$$n_i = \sqrt{N_C N_V} \cdot e^{-\frac{E_g}{2kT}},$$

где N_C, N_V — эффективная плотность состояний электронов и дырок соответственно.

Чем больше ширина запрещенной зоны E_g, тем меньше собственная концентрация свободных носителей при одной и той же температуре. Зависимость концентрации свободных носителей от температуры для некоторых полупроводников представлена на рис. 1.108. Видно, что концентрация свободных носителей при увеличении температуры быстрее всего увеличивается у кремния, а меньше всего у нитрида галлия и карбида кремния. Ниже приведена ширина запрещенной зоны и концентрация собственных носителей для граничной температуры, где $T_{гр}$ — граничная температура, при которой значение концентрации n_i сравняется со значением концентрации основных носителей, определяемой легирующей концентрацией N_d.

Материал	Si	GaAs	GaP	SiC(4H)	GaN
E_g, эВ	1,1	1,4	2,8	3,0	3,4
n_i, см$^{-3}$	$1{,}6\cdot10^{10}$	$1{,}1\cdot10^{7}$	$1{,}2\cdot10^{1}$	$1{,}1\cdot10^{-4}$	$9{,}2\cdot10^{-10}$
$T_{гр}$, °C	270	470	620	900	1300

Если полупроводниковому кристаллу не обеспечить эффективный теплоотвод, начинается процесс, при котором разогрев кристалла приводит к увеличению тока утечки, в результате чего кристалл еще больше разогревается и т. д., до выхода полупроводникового прибора из строя.

Одним из эффективных способов борьбы с токами утечки является использование полупроводников с широкой запрещенной зоной.

Чем шире запрещенная зона полупроводника, тем меньше вероятность перехода электронов из валентной зоны в зону проводимости. Если рассмотреть исторический аспект развития электроники, особенно силовой, то можно заметить, что развитие электроники постепенно движется в направлении использования полупроводников со все более широкой запрещенной зоной. Первые полупроводниковые приборы были на основе германия, но от них быстро отказались, так как этот материал имел ширину запрещенной зоны около 0,7 эВ. Затем стали использовать кремний, у которого ширина запрещенной зоны составляла 1,12 эВ. Карбид кремния в зависимости от типа кристаллической решетки может иметь ширину запрещенной зоны 2,2...3,3 эВ, что позволяет обеспечить на порядок более высокое сопротивление исток–сток у полевого транзистора в закрытом состоянии. Для сравнения, приборы на основе кремния выдерживают температуру до 125 °C, а приборы из карбида кремния — теоретически до 600 °C (на практике до 200 °C, и только потому, что температуру больше 200 °C не выдерживают корпусы приборов). Особенностью карбида кремния также является многообразие форм кристал-

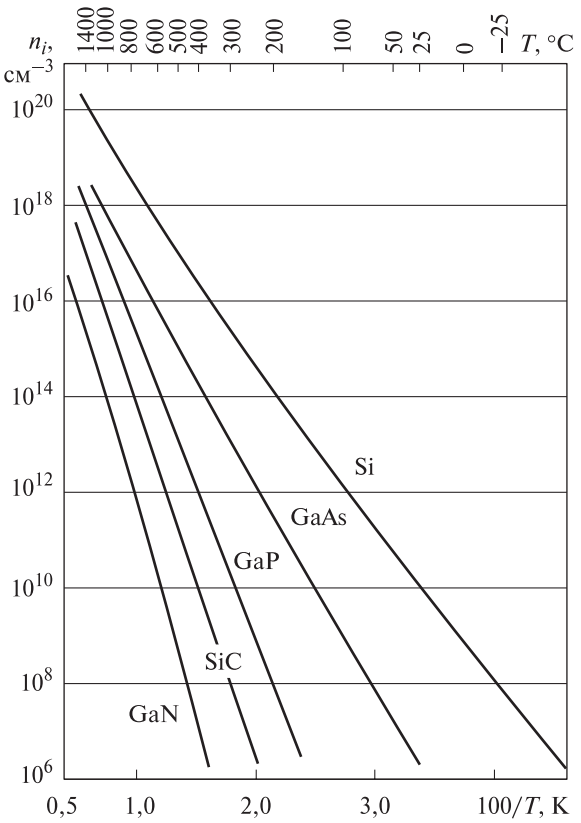


Рис. 1.108. Зависимость концентрации свободных носителей от температуры (https://studfile.net/html/528/114/html_EC5fgBlXIV.B2Ld/img-T7psPE.png)

лической решетки. Для целей электроники на практике пока применяются только два варианта формы кристаллической решетки: 4H и 6H. Кроме того, карбид кремния обладает в 3 раза большей теплопроводностью по сравнению с кремнием. Это обеспечивает лучший отвод теплоты от кристалла.

1.21.2. ТВЕРДОТЕЛЬНЫЕ ПРИБОРЫ НА КАРБИДЕ КРЕМНИЯ

Карбид кремния SiC — бинарное полупроводниковое соединение с большой шириной запрещенной зоны: 2,2...3,3 эВ (в зависимости от модификации). Модификаций карбида кремния насчитывается около 170, но только две из них — 4H-SiC и 6H-SiC — применяются в настоящее время при производстве полупроводниковых приборов. Материал устойчив к окислению вплоть до температур 1400 °C. При комнатной температуре SiC не взаимодействует ни с какими кислотами. Эти свойства обусловливают технологи-

ческие трудности при создании приборов твердотельной электроники на его основе.

Приборы на основе карбида кремния могут работать при высоких температурах вплоть до 600 °С, обладают стабильностью электрических характеристик при воздействии температуры и времени. Карбид кремния имеет теплопроводность в 3 раза больше, чем у Si, и в 10 раз больше, чем у GaAs, устойчив к воздействию радиации, что очень важно для создания приборов специального назначения. Эти качества определяют большую перспективность карбида кремния для изготовления полупроводниковых приборов. В настоящее время разработаны и производятся на основе карбида кремния мощные высоковольтные диоды Шоттки, мощные ВЧ- и СВЧ-транзисторы и светодиоды повышенной светоотдачи. Из сравнения ВАХ диода Шоттки на основе карбида кремния при различных температурах следует, что диод на основе карбида кремния имеет хорошие характеристики при повышенных температурах (рис. 109).

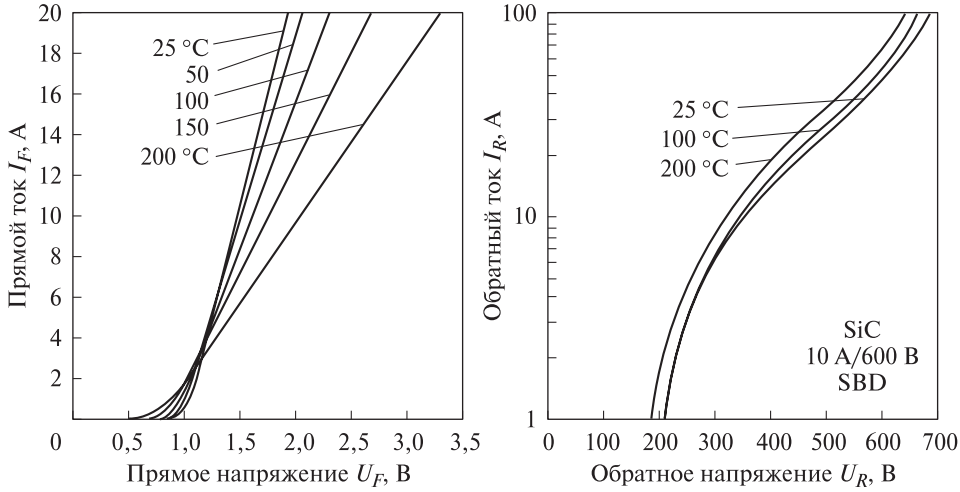


Рис. 1.109. Вольт-амперные характеристики диода Шоттки на основе карбида кремния при различных температурах (https://www.icquest.ru/media/7539/641159813813096.jpg, https://www.icquest.ru/media/7540/587646695169060.jpg)

Рассмотрим некоторые технологические проблемы, связанные с производством приборов на основе карбида кремния. Такой материал, как карбид кремния, человечество использует уже более 100 лет, в основном в качестве материала для изготовления шлифовальных инструментов вследствие очень высокой твердости, сопоставимой с алмазом. Поскольку карбид кремния является очень твердым материалом, его сложно обрабатывать. Второй проблемой является сложность его очистки от примесей, а для целей электроники необходим очень чистый карбид кремния. Технологические трудности были преодолены только в 2010-х годах. Еще одной проблемой, характерной для транзисторов из карбида кремния, является сложность управления такими приборами.

Как известно, кремниевый МОП-транзистор открывается при подаче на затвор напряжения 1...4 В относительно истока, в зависимости от марки. Если на затворе нулевое напряжение, то такой транзистор будет находиться в закрытом состоянии. Большинство транзисторов, изготовленных из карбида кремния, управляются следующим образом. Для открытия транзистора на затвор требуется подать напряжение 20...25 В относительно истока, а для закрытия требуется отрицательное напряжение: около –5 В. Данное обстоятельство существенно усложняет конструкцию узла, управляющего мощным транзистором из карбида кремния, что побудило конструкторов и исследователей искать пути решения.

В ноябре 2020 г. американская компания United SiC начала серийный выпуск МОП-транзисторов из карбида кремния, у которых напряжение открытия составляет 12 В, а напряжение закрытия равно нулю, как у кремниевых транзисторов. Преимуществом транзисторов на базе карбида кремния является и более высокое быстродействие, обусловленное тем, что они в процессе работы не входят в режим насыщения.

Более низкие токи утечки определяют меньший нагрев SiC-транзисторов, а это значит, что систему теплоотвода можно сделать компактной.

Современные транзисторы на основе карбида кремния при комнатной температуре имеют сопротивление в закрытом состоянии до 350 МОм против 15 МОм у кремниевых аналогов, а максимальное напряжение между истоком и стоком может достигать 15 кВ. Это позволяет применять такие транзисторы для коммутации в средневольтных распределительных сетях постоянного тока.

Второй высокотемпературный полупроводник — нитрид галлия GaN. Нитрид галлия имеет ширину запрещенной зоны еще бóльшую, чем карбид кремния (табл. 1.2). Поэтому высокая рабочая температура твердотельных устройств на основе GaN так же, как и для карбида кремния, является одним из преимуществ этого материала.

Таблица 1.2

Характеристики некоторых полупроводниковых материалов

Параметр	Материал		
	GaN	Si	SiC
Ширина запрещенной зоны, эВ	3,4	1,12	3,2
Диэлектрическая проницаемость	9,5	11,4	9,7
Подвижность, $см^2/(В \cdot с)$	990–2000	1500	650
Критическая напряженность, МВ/см	3,3	0,3	3,5
Дрейфовая скорость насыщения электронов, $\times 10^7$ см/с	2,5	1	2

Высокая критическая напряженность поля у GaN дает потенциальную возможность реализовать более высоковольтные приборы. Большая ширина запрещенной зоны обеспечивает высочайшую стабильность свойств при изменении температуры или воздействии радиации, что чрезвычайно важно в

первую очередь для космической и военной электроники, а также для устройств, работающих в жестких условиях. Высокая подвижность электронов и дрейфовая скорость определяют значительно меньшее сопротивление в проводящем состоянии и высокую удельную мощность по сравнению с карбидом кремния.

В результате GaN-транзисторы имеют на порядок бóльшую удельную мощность, что должно приводить к значительному уменьшению габаритных размеров приборов по сравнению с традиционными кремниевыми транзисторами.

Из нитрида галлия изготавливают большинство типов современных светодиодов. В настоящее время электронной промышленностью также освоено поколение устройств на основе нитрида галлия — транзисторы типа HEMT (*High Electron Mobility Transistor*) с высокой подвижностью электронов. В целях удешевления производства их изготавливают на кремниевых подложках, а не на подложках из карбида кремния или в виде чистой нитрид-галлиевой структуры.

В России Госкорпорация «Ростех» серийно выпускает мощные GaN-транзисторы, предназначенные для передатчиков базовых станций мобильной связи. Следут отметить, что НИИ микроэлектроники (г. Воронеж, входит в состав Госкорпорации «Ростех») и Компания «Миландр» (г. Москва) также занимаются разработкой силовых транзисторов на основе нитрида галлия.

Значительное преимущество GaN-транзисторов — возможность создания компактной аппаратуры, построенной с их использованием. Поэтому они считаются перспективными для использования в электромобилях, солнечных панелях, зарядных устройствах, базовых станциях мобильной связи, систем управления трехфазными электродвигателями и т. п.

Тесты к лекции 1.21

1. ***Какая из проблем считается очень важной при работе полупроводников при высоких температурах?***

а) рост тока утечки через подложку при повышенных температурах;

б) туннельный пробой *p*–*n*-перехода;

в) лавинный пробой *p*–*n*-перехода.

2. ***Какими основными параметрами в собственном полупроводнике определяется концентрация свободных носителей?***

а) двумя основными параметрами: шириной запрещенной зоны и температурой;

б) количеством донорных или акцепторных примесей;

в) двумя основными параметрами: шириной зоны проводимости и шириной валентной зоны.

3. ***В какой последовательности у полупроводников быстрее всего увеличивается концентрация свободных носителей при увеличении температуры?***

а) кремний, карбид кремния, нитрид галлия;

б) кремний, нитрид галлия, карбид кремния;

в) нитрид галлия, кремний, карбид кремния.

4. ***Какую проблему позволяет решить использование полупроводников с широкой запрещенной зоной?***

а) уменьшить токи утечки при нагреве;

б) уменьшить себестоимость полупроводниковых приборов;

в) уменьшить количество донорных или акцепторных примесей.

5. ***Как зависит вероятность перехода электронов из валентной зоны в зону проводимости от ширины запрещенной зоны?***

а) чем шире запрещенная зона полупроводника, тем меньше вероятность перехода электронов из валентной зоны в зону проводимости;

б) чем уже запрещенная зона полупроводника, тем меньше вероятность перехода электронов из валентной зоны в зону проводимости;

в) никак не зависит.

6. ***Какими преимуществами обладают приборы на основе нитрида галлия?***

а) возможность работать при высокой температуре;

б) низкая себестоимость;

в) высокая доступность.

1.22. ФОТОЧУВСТВИТЕЛЬНЫЕ ЭЛЕМЕНТЫ ПОЛУПРОВОДНИКОВОЙ ЭЛЕКТРОНИКИ

Цель лекции: ознакомление с основными полупроводниковыми фоточувствительными элементами.

1.22.1. ФОТОПРИЕМНИКИ

Фотодетекторы или фотоприемники — полупроводниковые приборы, регистрирующие оптическое излучение и преобразующие оптический сигнал на входе в электрический сигнал на выходе. Рассмотрим некоторые наиболее важные параметры фотоприемников.

Вольтовая чувствительность S_V показывает, насколько изменится напряжение ΔU на выходе фотоприемника при единичном изменении мощности ΔP падающего оптического излучения:

$$S_V = \frac{\Delta U}{\Delta P} \ [\text{В/Вт}].$$

Токовая чувствительность S_i показывает, насколько изменится ток ΔI в цепи фотоприемника при единичном изменении мощности ΔP падающего оптического излучения:

$$S_i = \frac{\Delta I}{\Delta P} \ [\text{А/Вт}].$$

Пороговая чувствительность — минимальная энергия оптического излучения P_m, которая вызовет на выходе фотоприемника сигнал, находящийся в заданном отношении m к шуму:

$$P_m = \frac{m\sqrt{\bar{U}_m^2}}{S_V};$$

$$m = \frac{\Delta U}{\sqrt{\bar{U}_m^2}} = \frac{P_m S_V}{\sqrt{\bar{U}_m^2}},$$

где $\sqrt{\bar{U}_m^2}$ — среднеквадратичное значение шумового напряжения.

Величина, обратная пороговой чувствительности, называется обнаружительной способностью:

$$D = \frac{1}{P_m}.$$

Зависимость обнаружительной способности от длины волны излучения представлена на рис. 1.110. Видно, что обнаружительная сопособность значительно зависит от материала фотоприемника. У каждого типа полупроводника имеется своя граничная длина волны, зависящая от ширины запрещенной зоны. При длинах волн, бóльших граничной, энергии квантов света оказываются недостаточными для генерации электронно-дырочных пар, и в этом случае обнаружительная способность становится нулевой.

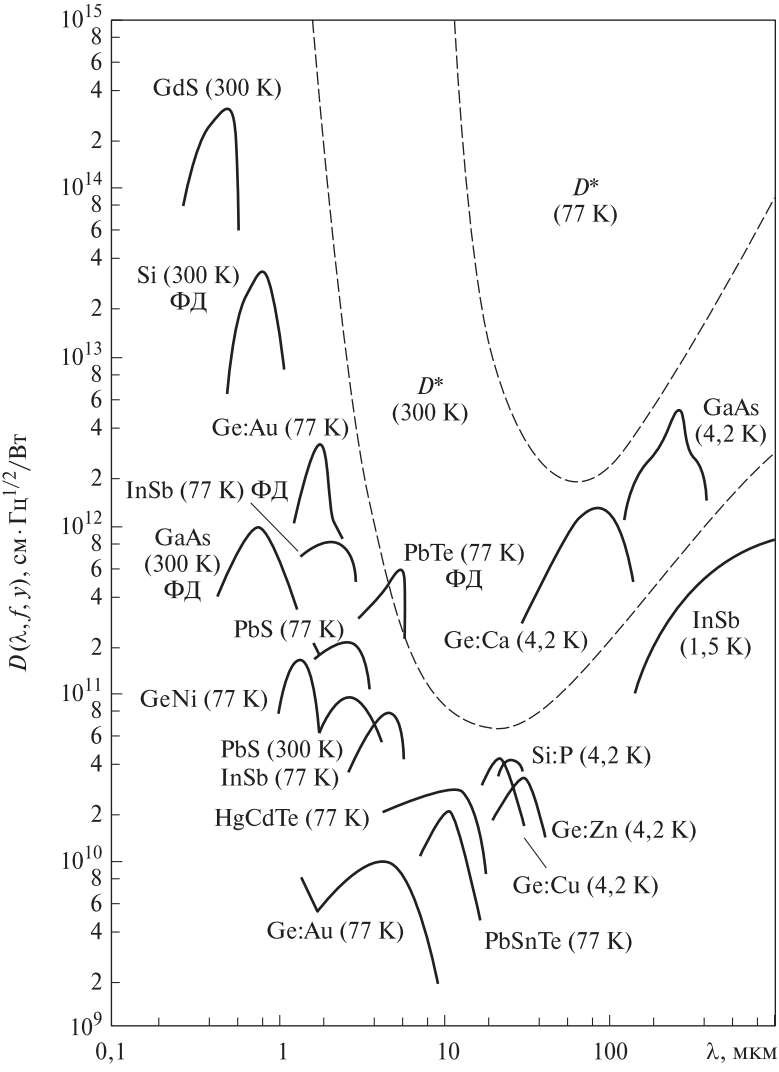


Рис. 1.110. Обнаружительная способность фотоприемников из различных материалов в зависимости от длины волны излучения (D — идеальная)

Граничная длина волны

$$\lambda_{гр} = \frac{1{,}22}{E_g} \ [\text{мкм}].$$

Значения длины волны для некоторых полупроводников приведены ниже.

Материал	SiC	GaP	GaAs	Si	Ge	InAs	InSb	KPT CdHgTe
E_g, эВ	3,0	2,26	1,42	1,12	0,66	0,36	0,17	0,10
$\lambda_{гр}$, мкм	0,4	0,54	0,86	1,09	1,85	3,4	7,2	12,2

1.22.2. ФОТОРЕЗИСТОРЫ

Фоторезисторами называются полупроводниковые резисторы, изменяющие свое электрическое сопротивление под действием оптического излучения. Условно-графическое изображение (УГО) фоторезистора на схемах представлено на рис. 1.111.

На рис. 1.112 представлено схематическое устройство фоторезистора. На диэлектрическую подложку *1* нанесен слой полупроводника *2*, на который нанесено два металлических контакта *3*. Контакты подключены к внешнему источнику питания и нагрузке. Когда световой поток Ф отсутствует, то в цепи течет небольшой, так называемый темновой ток. При освещении полупроводника *2* световым потоком Ф в полупроводнике происходит генерация

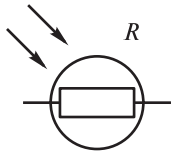


Рис. 1.111. Условно-графическое изображение фоторезистора

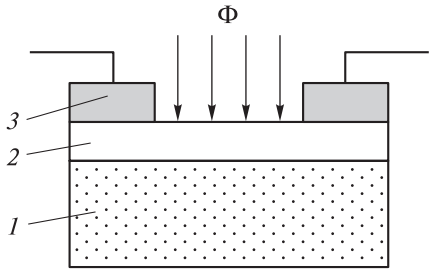


Рис. 1.112. Схематическое устройство фоторезистора

электронно-дырочных пар, т. е. число носителей тока увеличивается, а сопротивление полупроводника падает. Это приводит к увеличению тока в цепи и изменению напряжения на нагрузке.

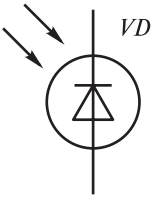


Рис. 1.113. Условно-графическое изображение фотодиода

1.22.3. ФОТОДИОДЫ

Фотодиодом называется полупроводниковый диод, в основе принципа действия которого лежит явление внутреннего фотоэффекта (рис. 1.113).

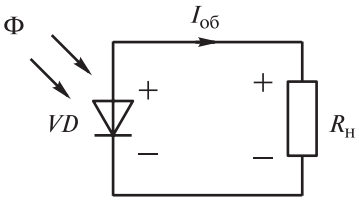


Рис. 1.114. Схема включения фотодиода в гальваническом режиме:
($I_{об}$ — обратный ток фотодиода; $R_н$ — сопротивление нагрузки)

Фотодиод может работать в двух режимах: фотогальваническом и фотодиодном. В фотогальваническом режиме фотодиод работает без внешнего источника напряжения и создает собственную ЭДС под действием энергии света. Схема включения фотодиода в гальваническом режиме представлена на рис. 1.114. Физический принцип возникновения собственной ЭДС заключается в возникновении в полупроводнике под воздействием светового потока неравновесных носителей заряда — электронов и дырок, которые формируют фотоэлектродвижущую силу.

Дырки перемещаются в зону *p*, а электроны в зону *n*. Таким образом, дырки заряжают *p*-область положительно, а электроны — *n*-область отрицательно; образуется разность потенциалов и возникает обратный ток. Чем выше освещенность, тем больше обратный ток. В фотодиодном режиме необходимо обеспечить смещение фотодиода в обратном направлении (рис. 1.115).

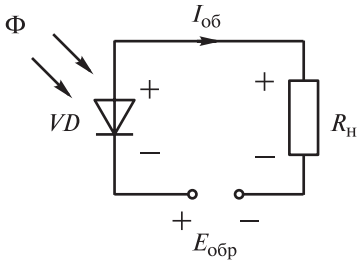


Рис. 1.115. Схема подключения фотодиода в фотодиодном режиме

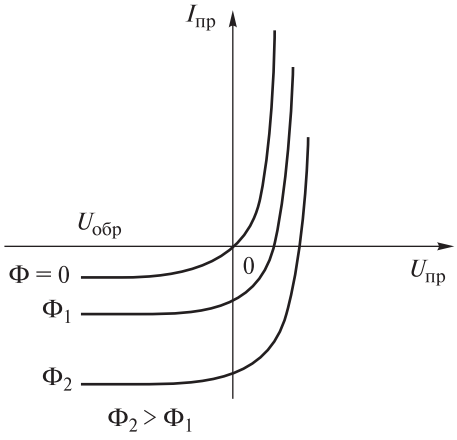


Рис. 1.116. Зависимость изменения тока фотодиода $I_{пр}$ от светового потока Ф

При отсутствии освещения (Ф = 0) сопротивление обратно смещенного фотодиода велико и через него протекает темновой обратный ток небольшой величины (рис. 1.116). Под действием светового потока ток увеличивается.

1.22.4. ЛАВИННЫЙ ФОТОДИОД

Высокочувствительный полупроводниковый прибор, преобразующий свет в электрический сигнал за счет фотоэффекта и обеспечивающий внутреннее усиление посредством эффекта лавинного умножения, называется лавинным

фотодиодом. Он может применяться в условиях незначительного светового потока. Схема устройства лавинного фотодиода приведена на рис. 1.117. В кристалле полупроводника *1* формируют *p*- и *n*-области таким образом, чтобы на область *p*–*n*-перехода мог падать световой поток Ф. Также наносят омические контакты *2* для подключения внешнего источника питания, антиотражающее покрытие *3*, которое выполняет функцию своеобразного фильтра. Лавинный фотодиод включают в схему с обратным смещением. При подаче сильного обратного смещения (близкого к напряжению лавинного пробоя, обычно порядка нескольких сотен вольт для кремниевых приборов) происходит усиление фототока (примерно в 100 раз) за счет ударной ионизации (лавинного умножения) генерированных светом носителей заряда.

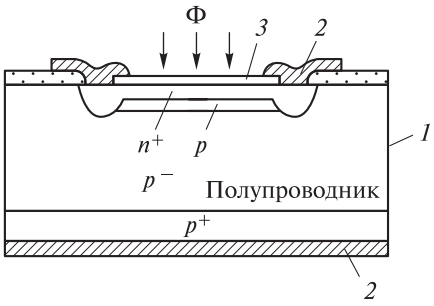


Рис. 1.117. Схема устройства лавинного фотодиода

1.22.5. ФОТОТРАНЗИСТОР

Фототранзистор — оптоэлектронный полупроводниковый прибор, вариант биполярного или полевого транзистора. Условное графическое обозначение биполярного фототранзистора представлено на рис. 1.118.

Если фототранзистор создан на базе биполярного, то область базы доступна для светового облучения, за счет этого появляется возможность управлять усилением электрического тока с помощью оптического излучения. Если фототранзистор создан на базе полевого, то падающий световой поток генерирует в *n*-канале и *p*–*n*-переходе (канал–затвор) электроны и дырки. Концентрация электронов в *n*-канале повышается, и уменьшается его сопротивление, а ток стока возрастает. Увеличение дырок в *p*-области вызывает появление фототока в цепи затвора.

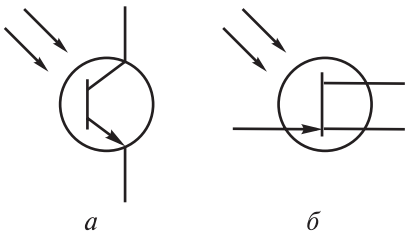


Рис. 1.118. Условное графическое изображение биполярного фототранзистора *n*–*p*–*n*-типа (*а*) и полевого фототранзистора с *p*–*n*-переходом и каналом *n*-типа (*б*)

1.22.6. ОПТРОНЫ

Оптрон или оптопара — полупроводниковый прибор, в котором конструктивно объединены источник и приемник излучения, имеющие между собой оптическую связь.

Оптрон состоит из оптического передатчика (излучателя), оптического канала и приемника оптического сигнала. Фотопередатчик преобразовывает

электрический сигнал в оптический. Передатчиком в большинстве случаев служит светодиод.

Оптический сигнал передается по оптическому каналу к приемнику. Канал бывает закрытым, когда излученный передатчиком свет не выходит за пределы корпуса оптрона и сигнал, сформированный приемником, синхронизирован с сигналом на входе передатчика. Такие каналы бывают воздушными или заполненными специальным оптическим компаундом. Также есть «длинные» оптроны, каналом в которых служит оптоволокно. Если оптрон сконструирован так, что генерируемое излучение, перед тем как попасть на приемник, покидает пределы корпуса, такой канал называется открытым. С его помощью можно регистрировать препятствия, возникающие на пути светового луча. Фотоприемник производит обратное преобразование оптического сигнала в электрический. В качестве приемника чаще всего применяются фотодиоды, фоторезисторы, фототранзисторы.

Широкое применение получили оптронные микросхемы — сборки оптронов с обвязкой в одном корпусе. Такие оптроны используют как переключающие устройства.

На рис. 1.119, *а* приведено УГО диодной оптопары. Диодные оптопары применяются в каналах передачи информации, для гальванической развязки и т. п.

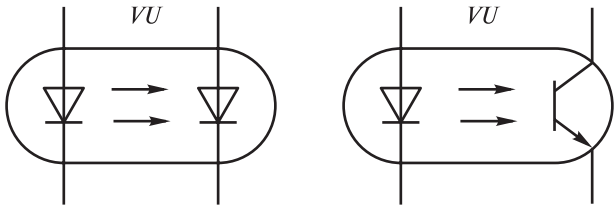


Рис. 1.119. Условно-графические изображения диодной (*а*) и транзисторной (*б*) оптопар

На рис. 1.119, *б* приведено УГО транзисторной оптопары. Транзистор обычно работает в ключевом режиме, и такая оптопара применяется в коммутирующих устройствах.

1.22.7. ПРИБОРЫ С ЗАРЯДОВОЙ СВЯЗЬЮ

Прибор с зарядовой связью (ПЗС) — полупроводниковый прибор, действие которого основано на накоплении и хранении локализованного заряда (зарядового пакета) в потенциальных ямах, образуемых в полупроводниковом кристалле под действием внешнего электрического поля, и последующей направленной передаче зарядовых пакетов из одной потенциальной ямы в

другую при изменении напряжения на внешних электродах. Конструктивно ПЗС выполняются в виде интегральной схемы.

Прибор с зарядовой связью применяют в качестве приемников оптического излучения для преобразования оптических изображений в последовательность электрических импульсов, т. е. для формирования видеосигнала.

Прибор с зарядовой связью изобретен американскими физиками У. Бойлом и Дж. Смитом в 1969 г.

Упрощенно ПЗС можно рассматривать как матрицу близко расположенных МДП-конденсаторов.

С физической точки зрения ПЗС интересны тем, что электрический сигнал в них представлен не током или напряжением, как в большинстве других твердотельных приборах, а зарядом. При соответствующей последовательности тактовых импульсов напряжения на электродах МДП-конденсаторов зарядовые пакеты можно переносить между соседними элементами прибора. Поэтому такие приборы получили название приборов с переносом заряда или с зарядовой связью.

Фрагмент структуры линейного ПЗС в режиме накопления показан на рис. 1.120. Структура состоит из слоя кремния *p*-типа (подложка), изолирующего слоя диоксида кремния и набора пластин-электродов. Один из электродов смещен более положительно, чем остальные два, и именно под ним происходит накопление заряда. Полупроводник *p*-типа получают легированием кристалла кремния акцепторной примесью, которая создает в кристалле полупроводника свободные положительно заряженные носители — дырки. Дырки в полупроводнике *p*-типа являются основными носителями заряда, а свободных электронов там очень мало, поскольку они являются неосновными носителями. Если подать небольшой положительный потенциал на один из электродов ячейки трехфазного ПЗС, а два других электрода оставить под нулевым потенциалом относительно подложки, то под положительно смещенным электродом образуется область, обедненная основными носителями — дырками. Они будут оттеснены вглубь кристалла. Это означает, что под электродом формируется потенциальная яма.

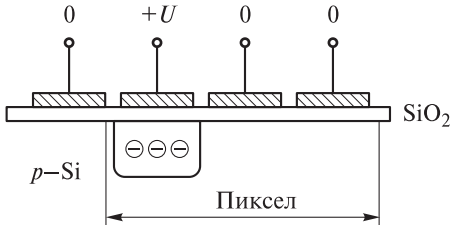


Рис. 1.120. Фрагмент структуры линейного ПЗС

В основе работы ПЗС лежит явление внутреннего фотоэффекта. Когда в кремнии поглощается фотон, генерируется пара носителей заряда: электрон и дырка. Электростатическое поле в области пиксела вытесняет дырку вглубь кремния, а неосновные носители заряда — электроны — будут накапливаться в потенциальной яме под электродом, к которому подведен положительный потенциал. Здесь они могут храниться достаточно длительное время, поскольку дырок в обедненной области нет и электроны не рекомбинируют.

Заряд, накопленный под одним электродом, в любой момент может быть перенесен под соседний электрод, если его потенциал будет увеличен, в то время как потенциал первого электрода будет уменьшен (рис. 1.121). Пере-

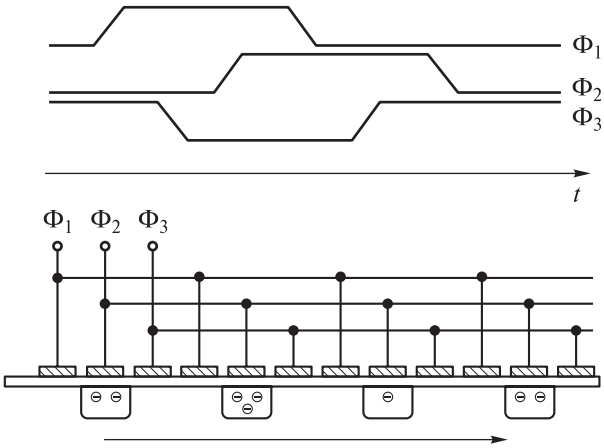


Рис. 1.121. Схема переноса зарядов в ПЗС

нос в ПЗС можно выполнить в одном из двух направлений: влево или вправо. Все зарядовые пакеты линейки пикселов будут переноситься в ту же сторону одновременно. Перемещением накопленного в потенциальных ямах заряда управляют подачей на электроды ПЗС-матрицы определенной последовательности импульсов напряжения, так называемых тактовых импульсов, смещенных по фазе Φ_1, Φ_2 и Φ_3 относительно друг друга на 120°.

В качестве преобразователей оптического изображения в видеосигнал используют ПЗС-матрицы. Они являются твердотельным аналогом передающих электронно-лучевых приборов в телевизионных передающих камерах.

Тесты к лекции 1.22

1. ***Какие приборы называются фотоприемниками?***

а) приборы, регистрирующие оптическое излучение и преобразующие оптический сигнал на входе в электрический сигнал на выходе;

б) приборы, фотографирующие изображения в цветном или черно-белом варианте;

в) приборы, регистрирующие акустические сигналы и преобразующие их в видимое изображение.

2. ***Что такое вольтовая чувствительность?***

а) вольтовая чувствительность показывает, насколько изменится напряжение на выходе фотоприемника при единичном изменении мощности падающего оптического излучения;

б) это минимальный уровень напряжения в вольтах, которое необходимо приложить к фотоприемнику, чтобы он начал регистрировать оптическое излучение;

в) вольтовая чувствительность определяет диапазон рабочих напряжений фотоприемника.

3. *Что такое токовая чувствительность фотоприемника?*

а) токовая чувствительность показывает, насколько изменится ток в цепи фотоприемника при единичном изменении мощности падающего оптического излучения;

б) минимальное значение тока, который должен протекать через фотоприемник, чтобы он начал регистрировать оптическое излучение;

в) токовая чувствительность определяет диапазон рабочих токов фотоприемника.

4. Что такое пороговая чувствительность?

а) минимальная энергия оптического излучения, которая вызовет на выходе фотоприемника сигнал, находящийся в заданном отношении к шуму;

б) средняя энергия оптического излучения, которая вызовет на выходе фотоприемника сигнал, находящийся в заданном отношении к шуму;

в) максимальная энергия оптического излучения, которая вызовет на выходе фотоприемника сигнал, находящийся в заданном отношении к шуму.

5. *Что называется обнаружительной способностью фотоприемника?*

а) величина, обратная пороговой чувствительности;

б) величина, равная логарифму пороговой чувствительности;

в) среднеквадратичное отклонение пороговой чувствительности.

6. *Что такое фоторезистор?*

а) полупроводниковый резистор, изменяющий свое электрическое сопротивление под действием оптического излучения;

б) полупроводниковый резистор, излучающий световую энергию при пропускании через него тока;

в) полупроводниковый резистор, изготавливаемый методом фотонабора.

7. *Что такое фотодиод?*

а) полупроводниковый диод, в основе принципа действия которого лежит явление внутреннего фотоэффекта;

б) полупроводниковый диод, в основе принципа действия которого лежит явление внешнего фотоэффекта;

в) полупроводниковый диод, в основе принципа действия которого лежит явление излучения квантов света при пропускании через него тока.

8. *Из каких основных элементов состоит оптрон?*

а) из излучателя, оптического канала и приемника оптического сигнала;

б) из эмиттера, оптического канала и коллектора;

в) из истока, оптического канала и стока.

9. *Каким образом в ПЗС управляют перемещением заряда, накопленного в потенциальных ямах?*

а) управляют подачей на электроды ПЗС-матрицы определенной последовательности тактовых импульсов напряжения;

б) управляют последовательным освещением ПЗС-матрицы слева направо;

в) управляют включением яркой вспышки с энергией, достаточной для генерации пары электрон–дырка.

1.23. МЕМРИСТОРЫ

Цель лекции: ознакомление с новым типом пассивных электронных элементов — мемристором.

1.23.1. ОСНОВНЫЕ ПОНЯТИЯ МЕМРИСТИВНЫХ ТЕХНОЛОГИЙ

Мемри́стор — пассивный электронный элемент, двухполюсник, способный изменять свое сопротивление в зависимости от протекшего через него электрического заряда.

В настоящее время технологии КМОП-структур получили значительное развитие. Вместе с тем идет поиск и разработка электронных компонентов, основанных на новых физических принципах работы. Научно-исследовательские работы по искусственному интеллекту как направление микроэлектроники очень быстро развиваются. Появляется все большее количество информации, для хранения которой необходимы ячейки памяти. Стандартные компьютеры уже не могут удовлетворить возрастающие потребности. В 2008 г. Компанией HP был предложен мемристор — новый элемент для хранения и обработки информации, который, по-видимому, способен решить обозначенные выше проблемы.

Конструкция мемристора напоминает конструкцию конденсатора. Между двух электродов помещен функциональный слой нанометровой толщины.

Мемристор изменяет свою проводимость в зависимости от протекающего через него тока. Перспективные устройства на его основе будут быстрее и более емкими, именно эти характеристики смогут сделать компьютеры недорогими и малых габаритов. Основное свойство мемристора — сохранение информации. Сопротивление мемристора сохраняется при отключении тока. Однако вследствие отсутствия полноценной теории работы мемристора учеными всего мира осуществляется экспериментальный поиск наилучшей архитектуры мемристора с наилучшими параметрами. Существенной проблемой является подбор материалов для мемристоров, обладающих определенными свойствами. Рассмотрим несколько видов мемристоров, изготовленных из разных материалов с применением разных технологий, требования, предъявляемые к каждому из них, и сравним характеристики этих мемристоров.

Функция мемристора заключается в запоминании своего сопротивления, значение которого зависит от его последнего состояния в данный момент. Когда напряжение на мемристоре оказывается равным нулю, его сопротивление останется таким, каким оно было в момент отключения напряжения, поэтому его признают запоминающим элементом. Этот пассивный элемент обладает двумя выводами. Единица сопротивления мемристора Ом и зависит от прошедшего через него заряда q.

Основные положительные черты мемристора:

– в выключенном состоянии не потребляет питания, обладает большой скоростью переключения и памятью;

– имеет простую структуру, напоминающую конденсатор и только два вывода;

– имеет широкую область применения: энергонезависимая память, реконфигурируемая электроника и даже нейроморфные системы;

– возможность объединения мемристоров в кроссбар-архитектуре и создания матриц памяти с произвольным доступом;

– технология создания мемристора хорошо компонуется с отработанной технологией создания КМОП-структур.

К недостаткам мемристора относятся:

– как сравнительно новый элемент, мемристор технически еще далек от совершенства, быстродействие и стабильность его характеристик находятся пока на недостаточно высоком уровне;

– требуются дополнительные исследования для доведения мемристора до коммерчески успешного изделия.

Мемристоры на основе оксидов металлов. Среди многих оксидов металлов для резистивного слоя мемристоров изучают такие материалы, как TiO_x, HfO_x, AlO_x, TaO_x и ZrO_x. Одним из первых был предложен мемристор на основе диоксида титана TiO_2, который может изменять свое сопротивление при перемещении вакансий кислорода под действием внешнего напряжения (рис. 1.122).

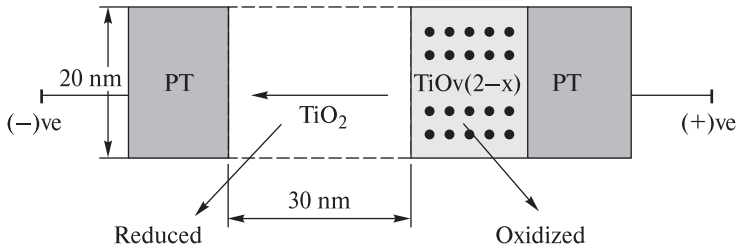


Рис. 1.122. Схема мемристора на базе оксида титана (*Strukov D.B.*, *Snider G.S.*, *Stewart D.R.*, *Williams R.S.* The missing memristor found // Nature letters. 2008. Vol. 453. Pp. 80–83)

Резистивный материал мемристора TiO_2, предложенного фирмой HP, расположен между двумя платиновыми электродами. При подаче напряжения на платиновые электроды кислородные вакансии, содержащиеся в диоксиде титана, будут перемещаться вправо или влево в зависимости от направления тока, при этом изменяется сопротивление резистивного слоя. К этой же группе можно отнести мемристоры на основе манганита MnO(OH). Мемристор на основе манганита обладает повышенной стабильностью и долговечностью работы.

Спиновые мемристоры. Эти мемристоры зависят от характеристик спина электрона, благодаря чему может проявляться гистерезисное поведение. Принцип действия таких мемристоров основан на том, что при прохожде-

нии тока через тонкую пленку, находящуюся под действием магнитного поля одного из конструктивных элементов мемристора, спины электронов, составляющих ток, будут ориентированы вдоль магнитных силовых линий. Это явление известно как спин-поляризация, и сопротивление прохождению тока будет минимальным. Если магнитное поле изменит свое направление на противоположное, то сопротивление току резко возрастет. Данный физический механизм может быть непосредственно использован для построения спинтронного мемристора.

Сегнетоэлектрические мемристоры. Такой мемристор состоит из сегнетоэлектрической пленки нанометровой толщины, находящейся между двумя металлическими электродами. В нанометровой структуре присутствует перестраиваемое гистерезисное и энергонезависимое резистивное переключение, которое аналогично мемристивному характеру классического мемристора. Наиболее известная в настоящее время сегнетоэлектрическая архитектура мемристора включает в себя пленку $BaTiO_3$, нанесенную на слой $La_xSr_yMnO_3$, которая находится между двумя металлическими электродами. Резистивное переключение сегнетоэлектрического мемристора рассматривается как результат изменения поляризации диэлектрика в сегнетоэлектриках. Следует отметить, что для тонкой сегнетоэлектрической пленки, которая, однако, может поддерживать свою сегнетоэлектричность, поверхностный заряд в сегнетоэлектрике не полностью экранируется электродами, что позволяет создать ненулевое деполяризующее электрическое поле в сегнетоэлектрическом слое. В результате в районе одного из электродов могут накапливаться кислородные вакансии на границе раздела оксид металла — сегнетоэлектрик. Благодаря этому высота энергетического барьера и сопротивление «включенного» состояния могут быть эффективно уменьшены. Отрицательное смещение индуцирует диссипацию накопленных зарядов на границе раздела и (или) восстановление оксида металла на границе раздела, переводя гетероструктуру в высокоомное состояние.

Резонансно-туннельные диодные мемристоры. При исследовании квантово-размерных резонансно-туннельных диодов было обнаружено, что они обладают свойствами мемристора. Резонансно-туннельный диод может быть изготовлен из различных полупроводниковых материалов и иметь различные резонансные туннельные структуры, такие как двойной и тройной барьеры, квантовая проволока и квантовая точка. Архитектура такого резонансно-туннельного диодного мемристора имеет вид двух контактов из сильно легированных полупроводниковых материалов с узкой запрещенной зоной. Такая архитектура образует область эмиттера, квантовую яму между двумя барьерами и область коллектора. Особенностью данной конструкции является то, что квантовая яма изготовлена с использованием полупроводника с небольшой запрещенной зоной, в то время как барьеры выполнены из материалов с большой запрещенной зоной. Когда на устройство подается напряжение, электроны, ранее захваченные в эмиттере, получают энергию от внешнего электрического поля и поэтому могут туннелировать через квантовую яму к коллектору, что приводит к увеличению тока. Экспериментально было уста-

новлено, что такое устройство имеет ВАХ в виде петли гистерезиса, аналогичной классическому мемристору. Однако никаких существенных достижений в этой области пока не достигнуто.

Мемристоры на основе халькогенидов. Халькогениды — бинарные химические соединения халькогенов (элементов шестой группы Периодической системы элементов Д.И. Менделеева) с металлами. Под действием электрического тока они могут изменять свое фазовое состояние — переходить из поликристаллической фазы в аморфную и обратно, демонстрируя при этом большое различие между значениями электрического сопротивления. Механизмом фазового перехода в халькогенидах является нагрев и охлаждение, переход происходит достаточно быстро, что позволяет использовать их в качестве диэлектрика в составе трехслойной структуры мемристора. Программирование такого элемента памяти происходит путем подачи электрического импульса соответствующей амплитуды и длительности.

Принципиальная пригодность конструкции на основе халькогенида доказана. Однако ей присущи недостатки, свойственные системам, использующим нагрев для процесса перехода: многократный нагрев приводит к изменению физических свойств из-за разрушения внутренней структуры вещества. Поэтому наиболее перспективными выглядят системы на основе оксидов металлов, работающие за счет перемещения вакансий кислорода в нанометровом диэлектрическом слое оксидов металлов при приложении электрического поля.

Органические мемристоры. Одним из первых и наиболее изученных органических мемристоров был прибор на основе структуры, состоящей из слоя полианилина в контакте с твердым электролитом на основе полиэтиленоксида. Мемристивная функция полимерного устройства достигается путем переключения электропроводности полимерной пленки между высокорезистивным (также называемым восстановленным) состоянием и высокопроводящим (также называемым окисленным) состоянием посредством хорошо известной окислительно-восстановительной реакции. Тонкий проводящий полимерный слой — полианилин (ПАНИ) наносят на изолирующую твердую подложку с двумя предварительно сформированными металлическими электродами. Дополнительная пленка полиэтиленоксида (ПЭО), легированного перхлоридом лития, образуется на центральной области слоя ПАНИ таким образом, чтобы обеспечить ионный поток в слой ПАНИ (и из него) на границе раздела. В основной области ПАНИ–ПЭО происходит окислительно-восстановительная реакция, и поэтому она называется активной зоной. Кроме того, поскольку для окислительно-восстановительной реакции требуется опорный потенциал, используют третий электрод — серебряную проволоку, присоединяемую к слою ПЭО, которая является электродом сравнения.

Известно, что в качестве резистивного слоя, расположенного между двумя электродами, был использован фиброин из коконов шелкопряда Bombyxmori. Механизм переключения сопротивления связан с образованием и разрушением проводящих филаментов в слое фиброина, поэтому в данной конструкции оказалось достаточно только двух электродов.

1.23.2. ОСНОВНЫЕ СВОЙСТВА МЕМРИСТОРА

Важной особенностью мемристора является его ВАХ в виде петли гистерезиса (рис. 1.123).

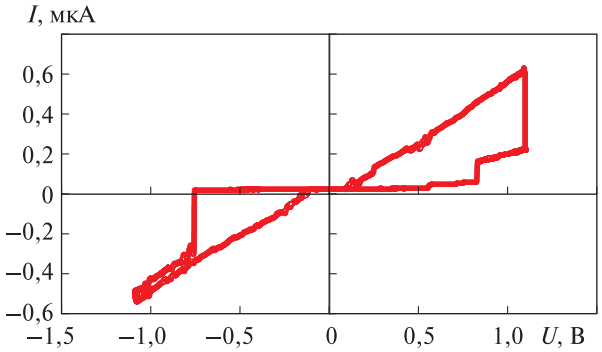


Рис. 1.123. Вольт-амперная характеристика мемристора в виде петли гистерезиса

Наличие петли гистерезиса обусловливает свойство этого элемента хранить информацию.

Следующее свойство мемристора состоит в зависимости площади петли гистерезиса от частоты переключения сигнала. При увеличении частоты площадь петли уменьшается и превращается на ВАХ в прямую линию, как у обычного резистора. Таким образом, на относительно высоких частотах мемристор ведет себя как обычный резистор.

Для мемристоров на основе оксидов металлов изменение сопротивления мемристора $M(q)$ в зависимости от прошедшего через него заряда $q(t)$ за время t описывается выражением

$$M(q) = R_{\text{off}}\left(1 - \frac{\mu_v R_{\text{on}}}{D} q(t)\right),$$

где R_{off} — высокое сопротивление мемристора; R_{on} — низкое сопротивление мемристора; μ_v — средняя подвижность ионов в резистивном слое; D — толщина резистивного слоя.

Из приведенного выражения следует, что заряд, проходящий через мемристор, не накапливается в нем, как в конденсаторе, а изменяет сопротивление устройства. Поэтому мемристор не является элементом, накапливающим энергию, и по сути, это нелинейный резистор. В табл. 1.3 представлены некоторые сравнительные характеристики ряда мемристоров.

Как следует из табл. 1.3, по параметру $R_{\text{on}}/R_{\text{off}}$ наиболее высокие показатели у мемристоров на основе оксидов металлов, что обусловливает возможность их использования для создания устройств памяти, имеющих не только два состояния «0» и «1», а несколько. Таким образом, теоретически одна

Таблица 1.3

Сравнительные характеристики некоторых мемристоров[1–5]

Тип мемристора	Отношение сопротивлений R_{on}/R_{off}	Время переключения, нс	Максимальное число циклов переключения
На основе оксидов металлов	2000	10	10^9
Полимерные	100	25	10^8
Сегнетоэлектрические	300	10	10^{14}
Спинтронные	5	10	10^{16}
Халькогенидные	>1000	50	10^8

мемристорная ячейка памяти может хранить в разы больше информации на той же площади, что и современные ячейки памяти. Мемристоры на основе оксидов металлов также имеют наименьшее время переключения из состояния высокого сопротивления в низкое. Однако такие мемристоры выдерживают не более 10^9 циклов переключения, что недостаточно. Поэтому необходимо проводить исследования по подбору материалов и архитектуры таких мемристоров.

Полимерные мемристоры в настоящее время имеют не самые лучшие характеристики, и требуется проведение многих исследований в этой области.

Сегнетоэлектрические мемристоры обладают хорошими характеристиками, а спинтронные мемристоры — одними из лучших для создания ячеек памяти, что обусловливает их применение.

Быстродействие халькогенидных мемристоров не очень высокое, тем не менее они также подходят для создания ячеек памяти, не требующих высокого быстродействия.

Поскольку мемристоры на основе оксидов металлов отличаются наиболее высокими показателями по ряду параметров, а технология их изготовления достаточно хорошо отработана, проведем анализ материалов, используемых для их изготовления.

---

[1] *Zhalnin V.P.*, *Vlasov A.I.*, *Gudoshnikov I.V.*, *Kadyr A.T.*, *Shakhnov V.A.* Market for memristors and data mining memory structures for promising smart systems // Entrepreneurship and Sustainability Issues. 2020. Vol. 8. No. 2. Pp. 98–115.

[2] *Zhalnin V.P.*, *Shakhnov V.A.*, *Vlasov A.I.* Methods for improvement of the consistency and durability of the inorganic memristor structures // International Journal of Nanotechnology. 2019. Vol. 16. No. 1–3. Pp. 187–195.

[3] *Zhalnin V.P.*, *Vlasov A.I.*, *Rasuk A.A.*, *Kosyanov O.V.* Composite nano-materials for implementation of promising memristive structures // Key Engi-neering Materials. 2022. Vol. 910 KEM. Pp. 703–712.

[4] *Ситников А.В.* Прикладная электроника: учебник. А.В. Ситников, И.А. Ситников. М.: Инфра-М, 2018. 576 с.

[5] *Strukov D.B.*, *Snider G.S.*, *Stewart D.R.*, *Williams R.S.* The missing memristor found // Nature letters, 2008. Vol. 453. Pp. 80–83.

1.23.3. АНАЛИЗ МАТЕРИАЛОВ МЕМРИСТИВНЫХ СТРУКТУР

В зависимости от материала свойства мемристоров могут отличаться. К основным свойствам мемристоров относят рабочие токи и напряжения, размеры устройств, термостабильность, запас шума, свойство изменения фазы и др. Поскольку эти свойства отличаются, области применения мемристоров могут быть различными. В табл. 1.4 представлены рабочие параметры некоторых мемристоров на основе оксидов металлов.

Как следует из табл. 1.4, наиболее энергоэффективными являются мемристоры на основе диоксида кремния, оксида никеля и оксида железа. Мемристоры на основе оксида железа больше подходят для изготовления аналоговых устройств вследствие довольно плавного изменения сопротивления. Мемристоры на основе оксида никеля больше подходят для цифровых устройств.

Диапазон применения мемристоров очень широк благодаря высокой масштабируемости и уникальным физико-химическим возможностям. Существует множество материалов, проявляющих мемристивный эффект: оксиды, халькогениды, полимеры, сегнетоэлектрики и т. п.

Таблица 1.4

Рабочие параметры некоторых мемристоров на основе оксидов металлов

Тип резистивного слоя	Рабочее напряжение, В	Рабочий ток, мА	Мощность, мкВт	Особенности свойств
Диоксид титана TiO_2	2	0,2	400	Высокое быстродействие, низкий уровень шума, R_{on}/R_{off} ~2000
Диоксид кремния SiO_2	2	0,00025	0,5	R_{on}/R_{off} ~10^5
Оксид железа Fe_2O_3	1,8	10	18	Характерно плавное изменение сопротивления. Больше подходит для аналоговых устройств
Оксид никеля NiO	1,5	5	7,5	Больше подходит для цифровых устройств вследствие особенностей переключения
Оксид ванадия VO_2	50	6	300	R_{on}/R_{off} ~10^4. Подходит для устройств с многоуровневой логикой
Оксид графита (графена) CO	6	100	600	Высокая термостабильность

Наиболее подходящими областями для применения мемристоров являются создание ячеек памяти, синапсов, нейроморфных архитектур, реконфигурируемая логика и многие другие. Однако технология мемристивных устройств является новой и создает множество проблем, таких как поиск нового материала, проблемы совместимости аналоговых и цифровых свойств мемристора и т. п. Ряд мемристоров обладает свойством обеспечивать способность хранения многоуровневой информации в одной ячейке, что позво-

лит изготовлять устройства хранения информации более компактными. Кроме того, мемристоры могут одновременно иметь аналоговое поведение, подобное синапсу мозга. Из этого следует, что мемристоры можно использовать как для цифровых, так и для аналоговых устройств, а также для создания нейроморфных систем.

Свойства мемристоров — рабочие токи и напряжения, размеры устройств, термостабильность, запас шума, свойство изменения фазы, масштабируемость — зависят от материала, из которого изготовлен компонент. Использование таких элементов памяти, как мемристор, является перспективным для применения в устройствах хранения и обработки информации. Мемристоры на базе оксидов металлов являются в настоящее время наиболее перспективными для первоочередного внедрения в промышленность, поскольку технология их изготовления хорошо отработана и они обладают приемлемыми характеристиками.

Тесты к лекции 1.23

1. ***Что такое мемристор?***

а) пассивный электронный элемент, двухполюсник, способный изменять свое сопротивление в зависимости от протекшего через него электрического заряда;

б) разновидность операционного усилителя;

в) разновидность аналого-цифрового преобразователя.

2. ***Какой параметр мемристора изменяется при прохождении через него тока?***

а) проводимость;

б) степень поляризации;

в) индуктивность.

3. ***В чем заключается основная функция мемристора?***

а) в запоминании своего сопротивления;

б) в изменении индуктивности;

в) в изменении степени поляризации.

4. ***Какое примерно напряжение необходимо оставить на контактах мемристора, чтобы его параметры не изменялись?***

а) 0 В;

б) 2 В;

в) 3 В.

5. ***В чем заключается принцип действия спинового мемристора?***

а) спины электронов, составляющих ток, ориентируются вдоль магнитных силовых линий;

б) спины электронов, составляющих ток, ориентируются поперек магнитных силовых линий;

в) спины электронов, находящихся в одной фазовой ячейке мемристора, имеют одинаковое направление.

6. ***В чем заключается принцип действия мемристора на основе диоксида титана?***

а) кислородные вакансии, содержащиеся в диоксиде титана, будут перемещаться вправо или влево в зависимости от направления тока, при этом изменяется сопротивление резистивного слоя;

б) доменные структуры, содержащиеся в диоксиде титана, будут ориентироваться вдоль электрического поля в зависимости от направления тока, при этом изменяется сопротивление резистивного слоя;

в) при прохождении тока через слой диоксида титана последний нагревается, и в результате этого изменяется сопротивление резистивного слоя.

7. ***В чем заключается принцип действия сегнетоэлектрического мемристора?***

а) в изменении поляризации диэлектрика в сегнетоэлектрике и, как результат, изменении сопротивления;

б) в изменении толщины слоя кислородных вакансий и, как результат, изменении сопротивления;

в) в изменении индуктивности диэлектрика в сегнетоэлектрике и, как результат, изменении сопротивления.

1.24. ДАТЧИКИ

Цель лекции: ознакомление с принципами работы различных датчиков.

1.24.1. ОСНОВНЫЕ ПОНЯТИЯ В ОБЛАСТИ ИЗМЕРЕНИЯ ФИЗИЧЕСКИХ ВЕЛИЧИН

Ранее в значении измерения физических величин использовался преимущественно русский термин «датчик». Иностранный термин «сенсор» вначале употреблялся только по отношению к устройствам, реагирующим на прикосновение человека («сенсорные устройства»: кнопки и клавиатура — *touchpad* и пр.) — в таком смысле в настоящее время верен именно этот вариант. Однако термин «сенсор» расширенно стали применять и в значении «датчик», поэтому будем считать, что сенсор и датчик в большинстве случаев одно и то же.

Датчик — устройство или элемент, принимающий и преобразующий измеряемое воздействие в электрический сигнал. Под измеряемым воздействием подразумевается не электрическая величина, а некоторая физическая, например, температура, давление или световой поток, а также действие различных по химическому составу веществ или биологических объектов.

На рис. 1.124 представлена классификация датчиков по природе физического эффекта.

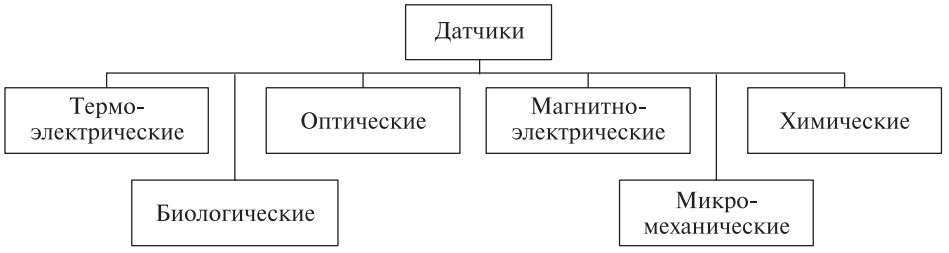


Рис. 1.124. Классификация датчиков

Ниже приведены характеристики датчиков.

Передаточная функция датчика показывает зависимость его выходного сигнала S от входного воздействия x. Эта зависимость может быть выражена в виде таблицы, графика или математического выражения $S = f(x)$.

Данная функция может быть как линейной, так и нелинейной (например, логарифмической, экспоненциальной или степенной).

Диапазон измеряемых значений — динамический диапазон внешних воздействий, который датчик может воспринять. Эта величина показывает максимально возможное значение входного сигнала, которое датчик преобразует в электрический сигнал, не выходя за пределы допустимых погрешностей.

Диапазон выходных значений — разность между электрическими выходными сигналами, измеренными при максимальном и минимальном внешнем воздействии.

Погрешность измерений — величина максимального расхождения между показаниями реального и идеального датчиков.

На точность датчиков влияют следующий характеристики: шумы, гистерезис, мертвая зона, параметры калибровки, повторяемость датчиков от партии к партии и воспроизводимость погрешностей.

1.24.2. ПЬЕЗОЭЛЕКТРИЧЕСКИЕ ДАТЧИКИ

В датчиках этого вида используется пьезоэлектрический эффект, который заключается в образовании зарядов в кристаллическом материале при приложении к нему механических напряжений. Этот эффект наблюдается в природных кристаллах, таких как кварц SiO_2, в поляризованных керамических материалах и некоторых полимерах. Слово «пьезо» происходит из греческого *piezen*, означающего давление.

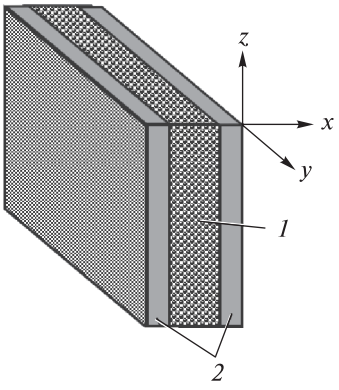


Рис. 1.125. Схема пьезоэлектрического датчика:
1 — кристалл; *2* — электроды

На рис. 1.125 представлена схема пьезоэлектрического датчика. Кристалл кварца в виде параллелепипеда имеет с двух противоположных сторон металлические электроды. Когда кристалл сдавливается, на электродах появляется разность потенциалов U.

Пьезоэлектрические кристаллы являются прямыми преобразователями механической энергии в электрическую.

Значение заряда Q_x, генерируемого на поверхности пьезоэлектрического кристалла, пропорционально силе F_x, приложенной, например, в направлении оси x:

$$Q_x = K_0 F_x,$$

где K_0 — пьезоэлектрическая постоянная.

Кристалл с нанесенным на него электродами представляет собой конденсатор, обладающий емкостью C, на котором напряжение определяется следующим выражением:

$$U = Q_x/C = K_0 F_x/C.$$

Тензорезистивные датчики. Действие тензорезистивных датчиков основано на изменении электрического сопротивления проводящего материала при его механической деформации. Это явление называется тензорезистивным эффектом.

На рис. 1.126, *а* представлено сечение датчика, содержащего тонкую кремниевую мембрану, сформированную анизотропным травлением в нагретой щелочи KOH, на рис. 1.126, *б* — схема расположения тензорезистивных датчиков. При приложении давления тонкая диафрагма деформируется, изменяя электрическое сопротивление расположенного на ней поликремниевого резистора.

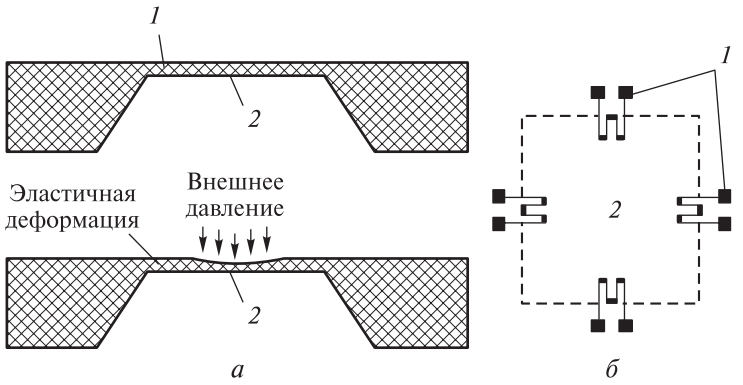


Рис. 1.126. Конструкция тензорезистивного датчика:
1 — тензорезистор; *2* — тонкая диафрагма

Мостовая измерительная схема (рис. 1.127), используемая в датчиках давления, содержит четыре резистора, соединенных в кольцевой замкнутый контур. Между верхней и нижней точками моста подключается источник напряжения $U_п$. Между левой и правой точками измеряется выходной сигнал с помощью внешнего усилителя или чувствительного регистрирующего прибора.

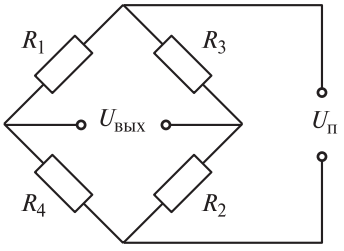


Рис. 1.127. Мостовая схема тензорезистивного датчика давления

Преимущества использования мостовой схемы заключаются в исключении влияния внешнего источника напряжения на выходной сигнал и устранении значительной по величине постоянной составляющей за счет предварительной балансировки моста, что важно при измерении слабых сигналов.

Акустические датчики. К акустическим датчикам можно отнести микроскопические микрофоны. Наиболее распространенными конструкциями микроскопических микрофонов являются основанные на конденсаторном принципе (рис. 1.128).

Важных элементов в таком микрофоне всего два: гибкая обкладка — мембрана и более толстая, неподвижная обкладка. Под воздействием давления воздуха мембрана смещается, изменяется емкость между обкладками, при постоянном заряде изменяется напряжение. Эти данные пересчитываются в амплитуды и частоты звуковой волны.

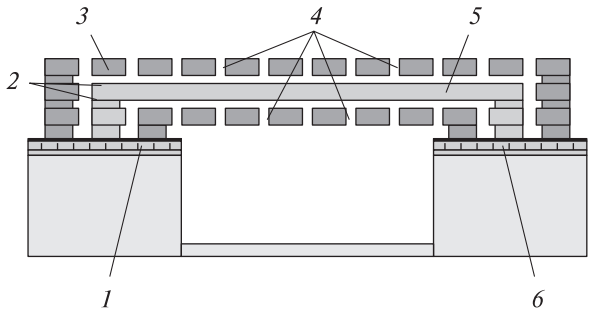


Рис. 1.128. Схема конденсаторного микрофона:
1 — верхняя неподвижная обкладка; *2* — отверстия в неподвижных обкладках; *3* — гибкая обкладка (диафрагма); *4* — зазоры; *5* — нижняя неподвижная обкладка; *6* — вентиляционное отверстие

Электрохимические датчики. Эти датчики являются самыми универсальными среди всех химических датчиков. В зависимости от принципа действия их подразделяют на датчики, измеряющие напряжение (потенциометрические) и измеряющие проводимость или сопротивление (кондуктометрические). В конструкции любого электрохимического датчика присутствуют как минимум два специальных электрода, между которыми протекает химическая реакция или происходит перемещение зарядов, образующихся в процессе этой реакции.

При проведении измерений электрическая цепь датчика замыкается через внешнюю измерительную цепь.

Принцип действия прибора (рис. 1.129) основан на явлении протекания специфичной химической реакции (электрохимической реакции) в электрохимической ячейке, представляющей собой емкость с раствором электролита с электродами (анодом и катодом).

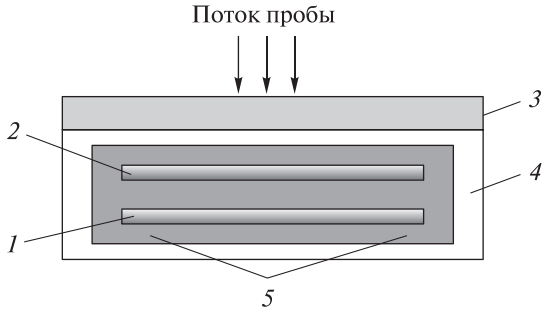


Рис. 1.129. Схема электрохимического датчика:
1 — считывающий электрод; *2* — рабочий электрод; *3* — тефлоновая мембрана; *4* — электронный блок; *5* — кислотный электролит

Поток пробы анализируемого вещества поступает в датчик через тефлоновую мембрану и вступает в химическую реакцию с электролитом, заполняющим ячейку. В результате в растворе возникают заряженные ионы, между электродами начинает протекать электрический ток, пропорциональный концентрации анализируемого компонента в пробе. Электронная схема обрабатывает возникающий электрический сигнал.

Термокаталитические датчики. На чувствительном элементе таких датчиков в результате протекания каталитических реакций происходит высвобождение тепла, а происходящее при этом изменение температуры внутри устройства регистрируется с помощью соответствующих детекторов температуры. Если одну из спиралей (рис. 1.130) покрыть слоем активного катализатора, а вторую — пассивным слоем, то находящийся в атмосфере монооксид углерода CO будет реагировать с кислородом воздуха на активном катализаторе, образуя диоксид углерода CO_2.

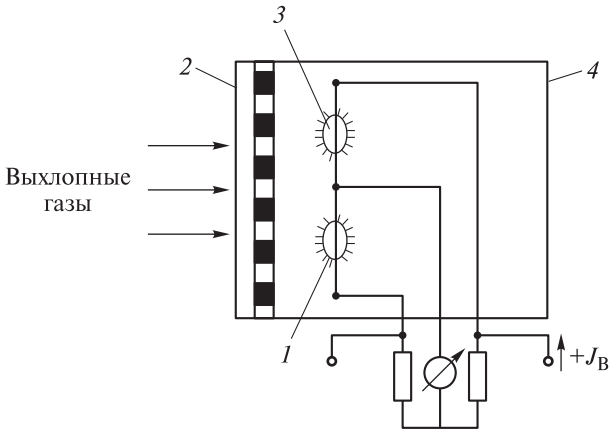


Рис. 1.130. Схема термокаталитического датчика:
1 — активный катализатор; *2* — металлокерамика; *3* — пассивный катализатор; *4* — измерительная камера

Выделяющаяся в результате этой реакции тепловая энергия вызывает повышение сопротивления активной спирали и в результате заметную разбалансировку моста.

Адсорбционные датчики. В самых простых адсорбционных датчиках используется изменение электрического сопротивления некоторых материалов, возникающее вследствие адсорбции газа (рис. 1.131).

На керамическом корпусе находятся два электрода, между которыми наносят слой полупроводящего оксида металла. Если газ проходит над этим активированным слоем, то проводимость оксида металла изменяется. С помощью мостовой схемы изменение проводимости преобразуется в изменение напряжения.

Чувствительные элементы адсорбционного датчика для различных анализируемых веществ:

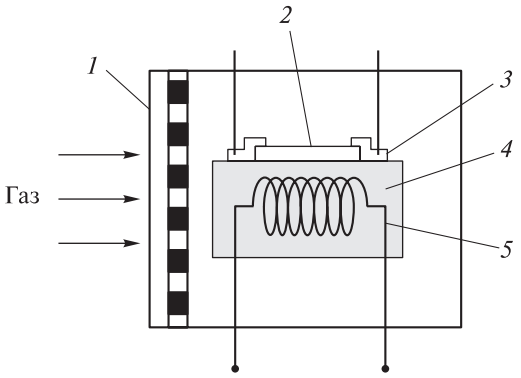


Рис. 1.131. Схема адсорбционного датчика:
1 — металлокерамика; *2* — оксид металла; *3* — электроды; *4* — керамический корпус; *5* — нагреватель

– оксид углерода CO — слои SnO_2 при рабочей температуре 250...500 °C;
– водород H_2 — слой SnO_2 с добавкой 1 % Sb_2O_3;
– этанол C_2H_5OH — слои SnO_2 на кварцевых подложках;
– сероводород H_2S — слой SnO_2 с добавкой алюминия;
– изобутан C_4H_{10} — слой легированного ZnO на подложке из Al_2O_3 и слой катализатора из соединения платины.

Биологические датчики. Эти датчики являются особым видом химических датчиков. Эволюция развития живых существ привела к созданию очень чувствительных органов, реагирующих на присутствие в воздухе даже нескольких молекул определенных веществ. Датчики на основе биологически активных материалов не обладают такой высокой чувствительностью, но в комбинации с другими детекторами (например, амперометрическими или тепловыми) могут качественно и количественно определять ряд биологических компонентов, таких как простейшие организмы, клетки, ткани, органеллы, мембраны, ферменты, рецепторы, антитела и нуклеиновые кислоты.

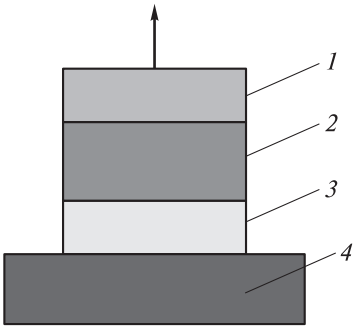


Рис. 1.132. Схема энзимного биодатчика (стрелка показывает выход электрического сигнала на регистрирующую аппаратуру):
1 — чувствительный элемент; *2* — энзимный слой; *3* — буферный слой; *4* — исследуемый образец

Энзимные датчики, являющиеся одной из разновидностей биобиодатчиков (рис. 1.132), обладают высокой избирательной способностью. В качестве чувствительного элемента может использоваться либо тепловой, либо электрохимический, либо оптический детектор. Энзимы действуют только в водной среде, поэтому их, как правило, размещают в гидрогеле. Принцип действия таких детекторов заключается в том, что ферменты (катализаторы), удерживаемые в специальном слое, вступают в реак-

цию с элементами подложки, проникающими в этот слой за счет процессов диффузии.

Термоэлектрические датчики. Принцип работы этой группы датчиков основан на рассмотренном ранее эффекте Зеебека и заключается в том, что в замкнутых контурах проводников или полупроводников возникает электрический ток, если места спайки различаются по температуре. Для измерения температуры один конец термопары помещают в среду измерения, а другой служит для снятия значений.

Терморезистивные датчики. Как следует из названия, этот тип датчиков работает по принципу изменения сопротивления проводника при изменении его температуры. Благодаря простой и надежной конструкции датчики этого типа широко применяются в электронике и машиностроении. Достоинством этих измерителей является высокая точность, чувствительность и простые устройства считывания.

Полупроводниковые термодатчики. Этот тип датчиков работает на принципе изменения характеристик *p*–*n*-перехода под воздействием температуры. Так как зависимость напряжения на транзисторе от температуры всегда пропорциональна, можно сделать датчик с высокой точностью измерения. Несомненными плюсами такого решения являются дешевизна, высокая точность данных и линейность характеристик на всем диапазоне измерения. Кроме того, их можно монтировать прямо на полупроводниковой подложке, что делает этот тип датчиков незаменимым для микроэлектронной промышленности.

Оптический датчик. Это оптическое электронное устройство, в котором используются рассмотренные ранее различные типы детектора оптического излучения, например, фотодиод, фоторезистор, фототранзистор и т. п. Диапазон длин волн, в котором используются такие датчики, от инфракрасного до ультрафиолетового. Принимаемый датчиком оптический сигнал преобразуется в электрический. Для работы датчика необходимо наличие излучателя, излучение от которого должно быть направлено на датчик. Датчик регистрирует изменение интенсивности излучения.

Микроакселерометры (датчики ускорения). Датчики ускорения относятся к инерциальным датчикам. В зависимости от способов съема информации они бывают оптическими, емкостными, тензочувствительными и другими датчиками. Обычно датчики ускорения (скорости) имеют сравнительно сложную пространственную структуру, содержащую инерционную массу, элементы ее подвеса, датчики смещения массы, другие элементы. Рассмотрим принцип работы конденсаторных акселерометров (рис. 1.133). Инерционная масса подвешена к корпусу акселерометра на двух пружинных подвесах.

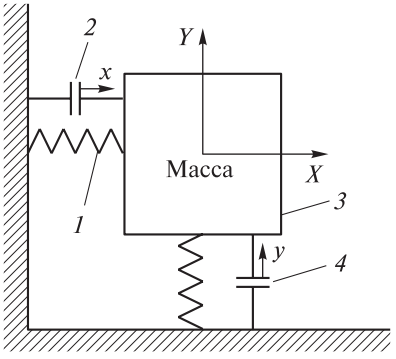


Рис. 1.133. Схема акселерометра:
1 — подвесы; *2* — конденсатор по оси *x*; *3* — масса; *4* — конденсатор по оси *y*

Между массой и корпусом установлены по двум осям конденсаторы, расстояние между обкладками которых может изменяться при перемещении массы.

В состоянии покоя расстояния между обкладками конденсатора не меняются и, следовательно, емкость конденсаторов также не меняется. При воздействии силы на корпус акселерометра эти расстояния меняются, что в дальнейшем фиксируется специальной аналоговой измерительной системой.

Тесты к лекции 1.24

1. ***Что такое датчик?***

а) устройство или элемент, принимающий и преобразующий измеряемое воздействие в электрический сигнал;

б) устройство или элемент, усиливающий сигнал, принятый детектором;

в) устройство или элемент, излучающий световой поток при воздействии на него.

2. ***Какие из перечисленных величин могут быть зафиксированы датчиком?***

а) температура, давление;

б) напряжение, ток;

в) емкость, индуктивность.

3. ***Что показывает передаточная функция датчика?***

а) зависимость его выходного сигнала от входного воздействия;

б) зависимость входного сигнала от выходного воздействия;

в) зависимость его выходного сигнала от напряжения питания.

4. ***Что такое диапазон измеряемых датчиком значений?***

а) динамический диапазон внешних воздействий, который датчик может воспринять;

б) диапазон выходных значений, который датчик может выдать;

в) диапазон напряжений питания, в котором датчик может работать.

5. ***Какой вид воздействия пьезоэлектрический датчик преобразует в электрический сигнал?***

а) механический;

б) химический;

в) оптический.

6. ***Какой вид воздействия тензоэлектрический датчик преобразует в электрический сигнал?***

а) механический;

б) радиационный;

в) биохимический.

7. ***На каком принципе работают полупроводниковые термодатчики?***

а) на принципе изменения характеристик p–n-перехода под воздействием температуры;

б) на принципе изменения ширины запрещенной зоны под воздействием температуры;

в) на принципе изменения степени легирования под воздействием температуры.

1.25. ОСНОВЫ КВАНТОВЫХ ВЫЧИСЛЕНИЙ

Цель лекции: ознакомление с физическими основами построения квантовых вычислительных систем.

1.25.1. ОСНОВНЫЕ ПОНЯТИЯ В ОБЛАСТИ КВАНТОВЫХ ТЕХНОЛОГИЙ

В настоящее время перспективная общая архитектура процессоров, даже с применением углеродных нанотрубок, проявляющих квантовые свойства, не отличается от используемых в современной полупроводниковой вычислительной технике. Такой подход имеет предел развития из-за ограничения минимальных размеров элемента процессора. Увеличение числа транзисторов при этом предполагает линейный рост производительности процессора. Идея предельного распараллеливания массива информации в современной вычислительной технике в целях повышения производительности еще далека от эффективной реализации.

Принципиально новые — квантовые вычисления — были предложены в 1980-х годах Ю. Маниным и Р. Фейнманом. Высокая эффективность устройств достигается за счет использования квантовых эффектов. В квантовых вычислениях для выполнения логической операции используется кубит, который находится не в двух состояниях (0 и 1), как в случае традиционных битовых вычислений, а во всех состояниях одновременно, т. е. суперпозиции состояний. Подобное состояние не позволяет подобрать аналога в классической физике и является чистым квантово-механическим эффектом. Основное достоинство квантового компьютера состоит в так называемом квантовом параллелизме, т. е. вычисление происходит одновременно по 2^N (N — число кубитов) состояниям, тогда как в обычном компьютере требуется $2N$ последовательных операций (рис. 1.134). Такая особенность позволяет проводить вычисления параллельно, что обеспечивает экспоненциальное увеличение эффективности при решении задач по оптимизации и факторизации чисел.

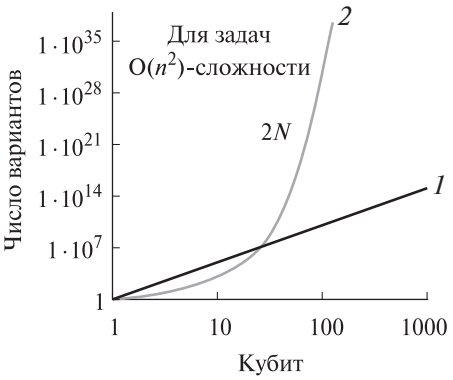


Рис. 1.134. Эффективность классических (1) и квантовых (2) вычислений

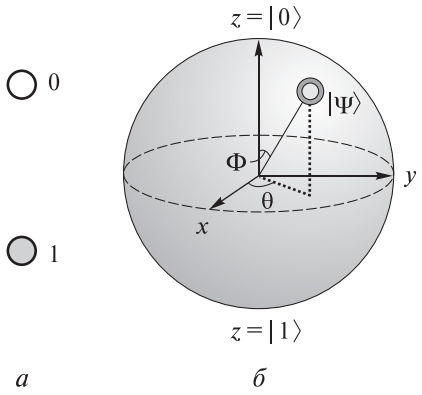


Рис. 1.135. Модель классического (*а*) и квантового (*б*) бита информации (https://ai4wifi.com)

Рассмотрим модель квантового бита информации. Представим шар, похожий на глобус, в котором направление по радиусу на север представлено нулем, а направление на юг — единицей. Это обычные биты. Любые другие положения соответствуют квантовым суперпозициям «0» и «1». Например, полужирной линией обозначен кубит, имеющий конкретные координаты, которые могут кодировать бесконечную последовательность цифр (рис. 1.135).

Следует отметить, что это не соответствует бесконечному количеству информации, которую содержит кубит. Квантовая механика требует, чтобы информация из кубита была извлечена путем измерения. Как и классический бит, кубит представляет единицу или ноль. Однако особенностью кубита, как и любого объекта квантовой механики, является огромный диапазон суперпозиций. В этом диапазоне содержатся в различных соотношениях оба исходных состояния. Другими словами, в кубите содержится топологическое пространство состояний между 0 и 1, поэтому кубит — базовый ресурс квантовой теории информации.

В одном кубите можно закодировать бесконечное количество классической битовой информации, однако извлечь ее обратно весьма проблематично. Если попытаться определить состояния кубита с помощью прямого измерения, то в результате прямого измерения получим ноль или единицу. Важнейшая особенность при этом — вероятность нуля или единицы зависит от широты исходного состояния (см. рис. 1.136). Таким образом, у кубита можно прочесть только один бит с определенной вероятностью. Кубит необходимо рассматривать как единицу квантовой информации, а не как бесконечное число квантовых битов.

Кубит обладает рядом уникальных особенностей. В квантовой системе состояние квантовой частицы (кубита) выражается через суперпозицию базисных состояний $\left(|0\rangle \text{ и } |1\rangle\right)$, имеющих матрицы плотности. Сами квантовые состояния объединяются с помощью умножения тензоров, в результате которого образуется пространство состояний из n квантовых частиц с размерностью 2^n. Это означает, что если один кубит может быть в двух супепозиционных состояниях: $\left(|0\rangle \text{ и } |1\rangle\right)$, то два кубита — уже в четырех (00, 01, 10, 11). На k кубитах можно провести математические операции с 2^k числами одновременно. Таким образом, имея в распоряжении всего сотню кубитов, можно оперировать с числами, превосходящими число атомов во Вселенной.Отсюда следует принцип квантового параллелизма. Квантовый параллелизм представляет собой вычислительный процесс, при котором в квантовом компьютере определяются выходные значения всех входных состояний практически сразу, без необходимости последовательного перебора или расчета.

Такой результат говорит об очень высокой перспективности работ по созданию квантовых компьютеров.

Поразительные возможности квантового компьютера являются результатом развития идей квантовой физики и квантовой информатики. На квантовом уровне микромир может быть математически описан уравнением Шредингера:

$$\frac{i\hbar\partial|\Psi\rangle}{\partial t} = \tilde{H}|\Psi\rangle,$$

где линейный оператор $\tilde{H}$ обладает свойством суперпозиции состояний

$$\tilde{H}\left(a|\Psi_1\rangle + b|\Psi_2\rangle\right) = a\tilde{H}|\Psi_1\rangle + b\tilde{H}|\Psi_2\rangle.$$

Другими словами, если система может существовать в состояниях $a|\Psi_1\rangle$ и $b|\Psi_2\rangle$, то она может существовать и в состоянии

$$|a|^2 + |b|^2 = 1,$$

где a, b — комплексные амплитуды.

Этот принцип является основой построения элементной базы квантовых компьютеров. Кубит представляет собой когерентную суперпозицию двух базисных состояний: $a|0\rangle$ и $b|1\rangle$. Для записи вектора состояния обычно используют дираковские обозначения. В самом элементарном случае двухуровневой системы (например, кубита) вектор состояния имеет вид

$$|\Psi_1\rangle = a|0\rangle + b|1\rangle.$$

Довольно часто в качестве синонима словосочетания «вектор состояния» используют термин «волновая функция».

Кубит — единица количества квантовой информации. Амплитуды a и b определяют вероятность нахождения системы в одном из базисных состояний, удовлетворяющих условию нормировки:

$$\int_V \Psi_1\Psi_2 dv = 1.$$

Другими словами, суммарная вероятность обнаружения частицы во всем пространстве объема V равна единице. В общем случае интегрирование следует проводить по всем переменным, кроме времени, от которых волновая функция явно зависит в данном представлении.

1.25.2. АРХИТЕКТУРА КВАНТОВОГО ПРОЦЕССОРА

Квантовые компьютеры представляют собой сопроцессоры или ускорители классических компьютеров, которые питают их и управляют ими. Квантовый компьютер всегда управляется классическим компьютером, как и графический процессор для видеоигр или для обучения нейронных сетей глубокому обучению. Эти обычные компьютеры используются для запуска

программ, управляющих квантовым процессором с физическими операциями, которые должны выполняться над кубитами, и интерпретируют результаты считывания кубитов.

Проведение кубитовых операций осуществляется под управлением обычного компьютера. Вычислительный процесс в квантовом компьютере носит интерференционный характер прежде всего потому, что амплитуды базисных состояний являются комплексными числами. Именно создание интерференции состояний позволяет приумножить вычислительную мощь квантового компьютера. Под континуальной средой, составляющей топологическое пространство состояний квантового компьютера, понимают физическую среду в любом агрегатном состоянии, в которой должно быть сосредоточено достаточное число кубитов, а также могут быть введены данные и на выходе измерено состояние кубитов.

Рассмотрим требования для физической реализации квантового компьютера.

1. Необходимо создать континуальную среду, содержащую определенное число кубитов, например из двухуровневых частиц.

2. Для осуществления заданного алгоритма вычислительного процесса должна быть реализована возможность физического воздействия на каждый кубит либо на любую их пару.

3. Необходимо технически реализовать возможность приготовления по командам управляющего компьютера требуемого числа кубитов входного регистра в исходном базисном состоянии $|01, 02, 03, ..., 0L\rangle$.

4. В процессе квантовых вычислений следует исключить эффекты, связанные с разрушением квантовой когерентности и квантовых состояний при взаимодействии с внешней средой, а также исключить помехи.

5. Физические свойства континуальной среды, в которой используются выбранные кубиты, должны обеспечить нелинейные взаимодействия для выполнения одно- и двухкубитовых операций.

6. На выходе квантового компьютера должно быть обеспечено надежное и точное измерение состояний квантовой вычислительной системы.

Последний пункт является наиболее трудной задачей при создании квантового компьютера, поскольку, как известно, при попытке прямого измерения состояния микрочастицы наблюдается коллапс ее волновой функции.

Рассмотрим схему квантового компьютера (рис. 1.136). Квантовый компьютер состоит из трех основных блоков: квантового регистра состояний, блока контроллеров квантовых вентилей и блока измерения состояния кубитов.

Квантовые регистры хранят информацию, обрабатываемую в компьютере, и используют принцип суперпозиции и запутанности, позволяя выполнять одновременные операции с большим числом значений регистров. Если провести параллель с классическими вычислениями, это память. Но обработка проводится непосредственно в памяти.

Контроллеры квантовых вентилей — физические устройства, которые воздействуют на кубиты квантовых регистров как для их инициализации, так и для выполнения на них квантовых вентилей. Эти команды применяются итеративно, в соответствии с алгоритмами, которые должны быть выполнены.

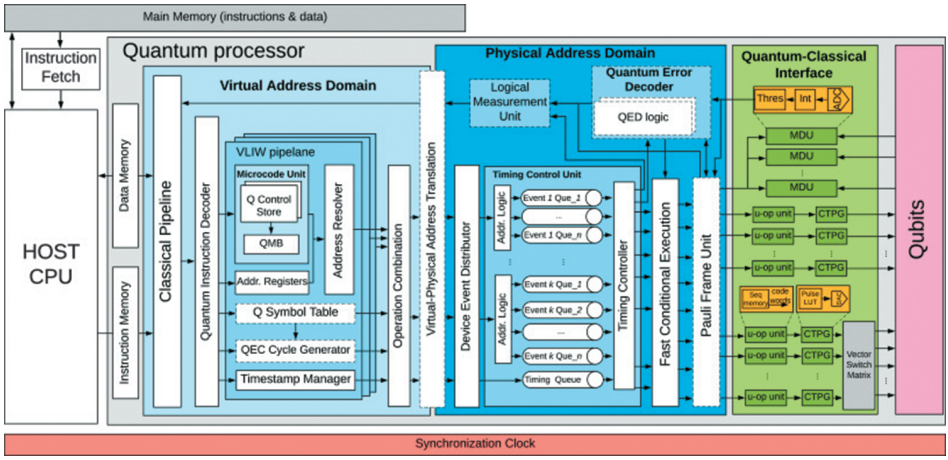


Рис. 1.136. Схема квантового компьютера (*Fu X.*, *Lao L.*, *Bertels K.*, *Almudever C.G.* A control microarchitecture for fault-tolerant quantum computing. 2019. Vol. 70. Pp. 21–30)

Их также можно использовать для управления кодами исправления ошибок. Квантовые вентили передают регистры как данными, так и инструкциями. Это не отдельные операции, как в классических микропроцессорах.

Измерение состояния кубита используется для получения результата в конце последовательного выполнения квантовых вентилей алгоритма и для оценки синдромов ошибок во время квантовой коррекции ошибок. Этот цикл инициализации, расчета и измерения обычно применяется несколько раз для оценки результата алгоритма. Затем результат усредняется до значения от 0 до 1 для каждого кубита в регистрах квантового компьютера. Значения, считанные физическими считывающими устройствами, затем преобразуются в цифровые значения и передаются на обычный компьютер, который управляет всем и выполняет интерпретацию результатов. В обычных случаях, таких как D-Wave и IBM, вычисления повторяются не менее 1000 раз.

Квантовые микросхемы обычно включают в себя квантовые регистры, элементы управления квантовыми вентилями и измерительные устройства (в случае сверхпроводящих кубитов или кубитов с электронным спином управление осуществляется посредством подачи микроволнового сигнала). Другие типы кубитов, как, например, на ионных ловушках или холодных атомах, используют лазеры для инициализации и измерения кубитов. Существующие на данный момент чипы схем имеют размер полнокадрового или двухформатного фотосенсора. Каждый кубит относительно велик, их размер измеряется в микронах для сверхпроводящих кубитов или до 100 нм для кубитов с электронным спином, тогда как современные КМОП-транзисторы процессоров имеют размеры транзисторов около 5 нм.

Чип для сверхпроводящих и электронных спиновых кубитов имеет размеры несколько квадратных сантиметров. Обычно он интегрируется в мед-

ный корпус, оснащенный коаксиальными разъемами, необходимыми для управления вентилями кубитов. В новейших сверхпроводниковых процессорах от IBM и Google с 53 кубитами требуется более 160 таких разъемов. Чип интегрирован в два небольших концентрических корпуса из алюминия и Cryoperm экраны (MμShield) с магнитной изоляцией.

Ключевые показатели. Самый известный набор ключевых показателей производительности квантовых компьютеров был создан Д. Ди Винченцо в 2000 г., когда он был исследователем IBM. Хотя отдельных кубитов практически не существовало, он определил основные технические характеристики универсального квантового компьютера на основе вентилей следующим образом.

1. Хорошо охарактеризованные кубиты.

Квантовые компьютеры используют кубиты, использующие квантовые объекты, которые могут иметь два различных и измеримых состояния. Их физические характеристики хорошо известны. Архитектура является масштабируемой в том смысле, что может использовать большое число физических кубитов, а затем и логических кубитов.

2. Инициализируемые кубиты.

В общем значению $|0\rangle$, часто называемому основным состоянием для ассоциированных квантовых объектов, соответствует, например, низший уровень энергии элементарной частицы или искусственного атома, как для сверхпроводящих кубитов.

3. Время когерентности.

Время, в течение которого кубиты находятся в когерентном состоянии, должно быть больше времени активации квантовых вентилей, чтобы можно было выполнить алгоритм, содержащий достаточно длинную последовательность квантовых вентилей. Преимущество кодов исправления ошибок, использующих большое число физических кубитов, заключается в увеличении полезного времени вычислений.

4. Набор универсальных квантовых вентилей.

Квантовое оборудование должно позволять создавать универсальный набор вентилей. Это зависит от технологии кубитов. Для этого требуется минимальный набор однокубитных вентилей, позволяющий создавать любое вращение в сфере Блоха, дополненное двухкубитным вентилем CNOT.

5. Измерение.

Возможность измерения состояния кубитов в конце вычислений. Это измерение не должно влиять на состояние других кубитов в системе. В идеале частота ошибок измерения должна быть значительно ниже 0,1 %.

Критерии Ди Винченцо довольно просты. С практической и эксплуатационной точки зрения квантовые компьютеры также можно охарактеризовать другим набором параметров.

6. Число кубитов.

Это обусловливает доступную вычислительную мощность, поскольку эта мощность теоретически увеличивается экспоненциально с числом кубитов, это ключевой параметр. По состоянию на середину 2021 г. коммерческий рекорд составлял 65 кубитов с самой большой системой IBMQ, доступной

в облаке. Число кубитов следует оценивать с точки зрения его масштабируемости. Некоторые технологии легче миниатюризировать и масштабировать, чем другие. В эту миниатюризацию необходимо интегрировать как чипы квантовых кубитов, так и элементы, которые ими управляют. Кроме того, необходимо следить за тем, чтобы декогеренция и шум не увеличивались по мере роста числа кубитов. Кубиты в ионных ловушках обладают превосходной точностью, но плохо масштабируются. Вероятно, сверхпроводящие кубиты лучше масштабируются, но их точность не стабильна по мере роста числа кубитов. Кубиты на холодных атомах масштабируются немного лучше, но с некоторыми практическими ограничениями по числу управляемых атомов. Теоретически кубиты с электронными спинами могут масштабироваться лучше всего.

7. Возможность подключения кубитов.

Это обусловливает скорость выполнения квантовых алгоритмов. Чем больше эта физическая связь, тем быстрее будет выполнение кода. При низкой связности компилятору квантового кода придется добавить гораздо больше операций для связывания кубитов вместе, особенно полагаясь на шлюзы SWAP. Эта связь сильно варьируется от одной технологии к другой. В 2D-технологиях, как и в случае со сверхпроводящими и кремниевыми кубитами, она ограничена соседними кубитами.

8. Параллельность вычислений.

То, как кубитовые вентили могут быть распараллелены в разных зонах кубитов без нарушения работы, также будет определять скорость выполнения квантовых алгоритмов.

9. Точность выполнения логических операций.

При выполнении квантовых вентилей и считывании их состояния точность кубитов обусловливает возможность выполнения длинных алгоритмов. Это напрямую влияет на глубину поддерживаемого алгоритма, а также на способность запускать коды квантовой коррекции ошибок и создавать логические кубиты с произвольным уровнем точности.

10. Длительность операций как для квантовых вентилей, так и для измерения состояния кубита.

Первое, очевидно, важно для того, чтобы алгоритмы работали как можно быстрее, но второе не менее важно, потому что кубит участвует в кодах исправления ошибок и, следовательно, обусловливает время выполнения всех алгоритмов.

11. Рабочая температура.

Для процессора и его оборудования многое зависит от типа кубитов. Действующие в настоящее время квантовые компьютеры на основе сверхпроводников работают при очень низкой, близкой к абсолютному нулю температуре, но предполагается, что некоторые типы кубитов все еще находятся на стадии исследований. Так, например, предназначены для работы при комнатной температуре, схемы на фотонах и NV-центрах. Однако это не обязательно относится к сопутствующему оборудованию, такому как фотонные генераторы и однофотонные детекторы. Работа при очень низкой температуре — это способ сохранить когерентность кубитов. Но чем ниже температура, тем меньше энергии могут излучать кубиты и их управляющая электроника.

12. Общее потребление энергии включает в себя потери на все компоненты квантового компьютера: сам процессор, всю его управляющую электронику, а также задействованные криогенные системы.

1.25.3. ФИЗИЧЕСКИЕ ПЛАТФОРМЫ ДЛЯ РЕАЛИЗАЦИИ КВАНТОВЫХ ВЫЧИСЛЕНИЙ

Рассмотрим варианты реализации квантового компьютера. Для создания реального квантового компьютера необходимо получить практические результаты по ряду направлений исследований, в частности, по созданию регистров кубитов. В качестве кубита может быть выбрана почти любая квантовая система, например, можно выбрать двухуровневые квантовые ячейки со спином, равным 1/2. Это могут быть электронные или ядерные спины, можно также выбрать орбитальные состояния электрона в квантовых ямах или квантовых точках. Состояния «0» и «1» электрона в квантовых точках разделить потенциальным барьером; управление динамикой кубита осуществлять лазерными импульсами через возбужденные уровни энергии электрона. Одиночный фотон может также служить кубитом. Любые два состояния фотона с ортогональными поляризациями интерпретируются как состояния кубита $|0\rangle$ и $|1\rangle$. В качестве кубита можно использовать два фотонных состояния, различающихся фазой. Кроме двухуровневой реализации кубитов, возможна их реализация в трехуровневой системе. Биты в таких системах получили название «кутриты» (*q-trits*), в многоуровневых системах единицы информации называются кудитами (*q-dits*).

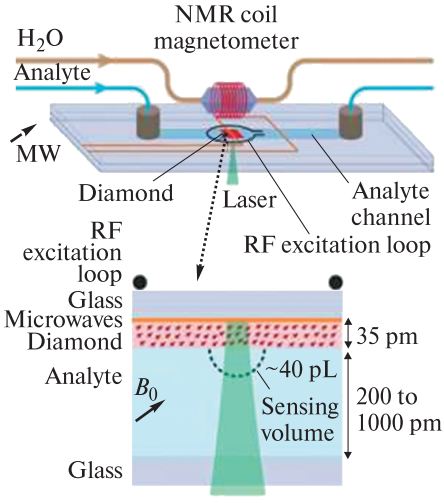


Рис. 1.137. Схема квантового компьютера на основе ЯМР-резонанса на МОП-структуре (*Hensen B.* et al. A silicon quantum-dot-coupled nuclear spin qubit // Nat. Nanotechnol., 2020. Vol. 15. No. 1. Pp. 13–17)

Квантовые компьютеры на основе ядерного магнитного резонанса (ЯМР). Такие устройства представляют собой континуальную среду с ядерными спинами-кубитами. Считается, что реализовать ЯМР-компьютер возможно, если число кубитов в компьютере будет больше одной тысячи.

Материал подложки выполняют из бесспинового изотопа кремния ^{28}Si. В подложку в приповерхностный слой внедрены донорные атомы фосфора ^{31}P, замещающие атомы кремния в узлах кристаллической решетки (рис. 1.137). Такие доноры обладают ядерными спинами $I = 1/2$. Спины не взаимодействуют с окружающими атомами кремния, но могут взаимодействовать между собой за счет частичного перекрывания электронных волновых функций. Регулярно расположенные в полупроводниковой структуре донорные атомы с

ядерными спинами могут быть использованы в качестве кубитов. Глубина залегания доноров ~20 нм.

Кубиты располагают на расстоянии l (~20 нм) один от другого, а над ними формируют решетку из управляющих затворов с периодом l. Решетка из A-затворов служит для управления резонансной частотой ядерно-спинового кубита. Подобная решетка из J-затворов служит для управления взаимодействием электронов соседних ядерных спинов.

Квантовые компьютеры на квантовых точках. Квантовые точки удерживают один или несколько электронов или дырок основного материала, которые достаточно малы, чтобы квантовые эффекты были значительными, и достаточно велики, чтобы их можно было контролировать. В качестве основного материала обычно используют полупроводники. Изначально использовался GaAs, так как на его основе возможно сделать высококачественные квантовые ямы для улавливания электронов. Однако и Ga, и As имеют общий ядерный спин, который взаимодействует с электронным спином и вызывает быстрое разрушение когерентного состояния. В настоящее время предпочтительными полупроводниковыми материалами являются Si, Ge, SiGe и изотоп ^{28}Si. Преимуществом такого подхода является возможность использования передовых полупроводниковых технологий.

Цепочка квантовых точек формируется в подзатворном диэлектрике кремниевого МОП-транзистора (рис. 1.138). Для индивидуального управления кубитами наряду с общим затвором внутри диэлектрика создают систему индивидуальных управляющих затворов G_i, действующих непосредственно на каждый кубит.

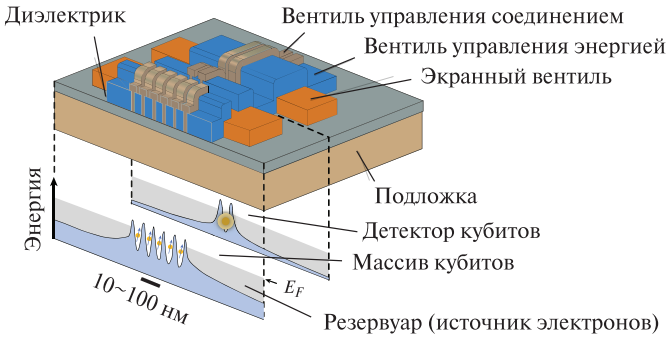


Рис. 1.138. Схема квантового компьютера на квантовых точках

При приложении к общему верхнему затвору большого напряжения происходит заполнение состояний квантовых точек и унифицируются состояния кубитов. Входные и выходные сигналы поступают через индивидуальные затворы каждого кубита. Систему кубитов в МОП-транзисторе можно рассматривать как последовательный ряд МОП-структур, которые соответствуют связанным квантовым точкам.

При приложении напряжения между истоком и стоком начнет формироваться обедненная область, ширина которой будет увеличиваться в сторону стока. Из условия уравнения непрерывности следует, что токи на разных участках должны быть равны. Если напряжение на первом кубите равно V_1, а на втором — V_2, то, вычисляя ток, можно определить перераспределение зарядов и сдвиг порогового напряжения.

Рабочие температуры кубитов на квантовых точках могут достигать 1 K, что позволяет интегрировать управляющую электронику на одном чипе. Архитектуры квантовых компьютеров на основе квантовых точек в настоящее время являются активной областью исследований.

Квантовые компьютеры на ионных ловушках. Кубиты на ионных ловушках кодируют квантовую информацию на электронных энергетических уровнях ионов, таких как Ca^+ или Yb^+. Благодаря тому что ион обладает электрическим зарядом, возможно получить два практических явления. Во-первых, ионы можно «подвесить» во внешнем электромагнитном поле в так называемой ионной ловушке. Во-вторых, ионы отталкиваются друг от друга вследствие действия кулоновских сил и выстраиваются в цепочку на расстоянии 5...10 мк друг от друга. Благодаря этому же кулоновскому взаимодействию реализуется запутанное состояние между ними (рис. 1.139). Для контроля и считывания состояния кубитов обычно используют лазеры и детекторы фотонов.

Работа системы при криогенных температурах помогает достичь необходимых условий вакуума и снижает электронные шумы. Так, работа при температуре 50...100 K дает значительные преимущества.

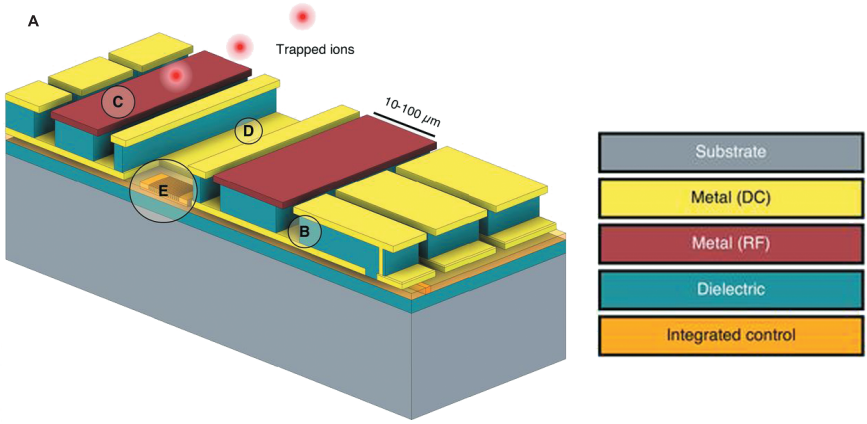


Рис. 1.139. Схема квантового компьютера на ионных ловушках (*Fu X.*, *Lao L.*, *Bertels K.*, *Almudever C.G.* A control microarchitecture for fault-tolerant quantum computing. 2019. Vol. 70. Pp. 21–30)

Линейные ионные ловушки хорошо зарекомендовали себя и позволяют получить запутанные состояния для 10...20 кубитов. Текущие усилия сосредоточены на масштабировании до 50...100 кубитов. При большем числе кубитов необходим поиск нового метода физического перемещения ионов между линейными ловушками или передачи квантовой информации между модулями. Возникают проблемы с масштабируемостью систем с несколькими линейными ловушками.

Квантовые компьютеры на фотонах. Кубиты на основе фотонов имеют ряд преимуществ. Во-первых, в отличие от остальных систем, квантовые состояния света слабо подвержены декогеренции, или она может быть эффективно подавлена, во-вторых, оптические фотоны являются универсальными носителями информации, пригодными как для обработки (реализации алгоритмов вычислений), так и для передачи (в том числе и по квантовым каналам связи) информации. Однако практическое применение такого рода систем ограничено вероятностным характером двухкубитных оптических гейтов, обусловленным слабым взаимодействием одиночных фотонов друг с другом.

Ключевыми компонентами (модулями) линейно-оптического квантового компьютера являются источники одиночных неразличимых фотонов, фотонные интегральные схемы с предельно малыми потерями и высокоэффективные однофотонные детекторы.

В настоящее время реализована многофотонная запутанность в кремниевых устройствах и квантовая телепортация, которые могут обеспечить масштабируемую платформу для квантовых вычислений или более общей квантовой обработки информации.

Квантовые компьютеры на базе сверхпроводников. Наиболее перспективной платформой создания квантового компьютера являются сверхпроводниковые кубиты. Первая реализация такого типа кубита была представлена Компанией NEC в 1999 г. Сверхпроводниковый кубит по сути является нелинейным устройством. Конструкционно кубит представляет собой *LC*-контур, где в качестве индуктивности используется джозефсоновский переход — нелинейный бездиссипативный элемент. В качестве емкости может быть использован встречно-штыревой либо плоскопараллельный конденсатор. Для связи между кубитами, а также для считывания состояния кубита используются линейные резонансные структуры (резонаторы). Как правило, резонаторы выполнены в виде компланарной линии.

Благодаря особенности джозефсоновского перехода, заключающейся в нелинейной индуктивности данного элемента, дополнительно шунтируя переход конденсатором, возможно создать особый колебательный контур. Энергетические уровни данного контура эквидистантны, что дает возможность разделять и четко определять состояния системы. Типовая схема сверхпроводникового кубита представлена на рис. 1.140. Для возбуждения состояния кубита используется *XY*-управление, для настройки частоты кубита — *Z*-управление.

Преимуществом рассматриваемой платформы является совместимость с технологиями микроэлектронной промышленности. За счет планарной реа-

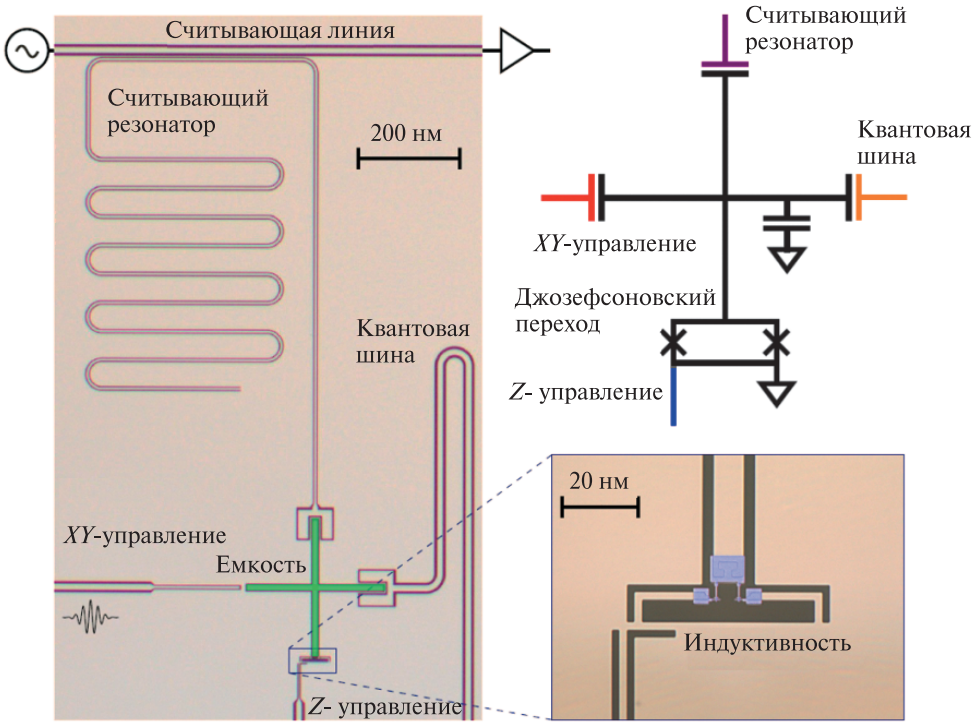


Рис. 1.140. Схема квантового компьютера на сверхпроводящих кубитах (*Pino J.M.* et al. Demonstration of the trapped-ion quantum CCD computer architecture // Nature, 2021. Vol. 592. No. 7853. Pp. 209–213)

лизации кубитов также возможна масштабируемость. Работа со сверхпроводниковыми структурами подразумевает работу при сверхнизких температурах, для этого используют криостат растворения.

Традиционный подход к управлению сверхпроводящими кубитами и реализации квантовых вентилей предполагает использование резонансных микроволновых импульсов с тщательно контролируемой амплитудой и фазой. Генерация микроволновых сигналов для сверхпроводниковых кубитов требует значительных аппаратных затрат на оборудование, работающее при комнатной температуре за пределами криостата, включая микроволновые источники, генераторы сигналов произвольной формы и усилители сигнала.

В 2019 г. специалистами Google был опубликован доклад о достижении квантового превосходства. В научном эксперименте исследовали квантовый компьютер Google на базе процессора, содержащего 53 кубита, получивший название Sycamore. Данный результат наглядно продемонстрировал перспективность реализации квантового компьютера на базе сверхпроводников.

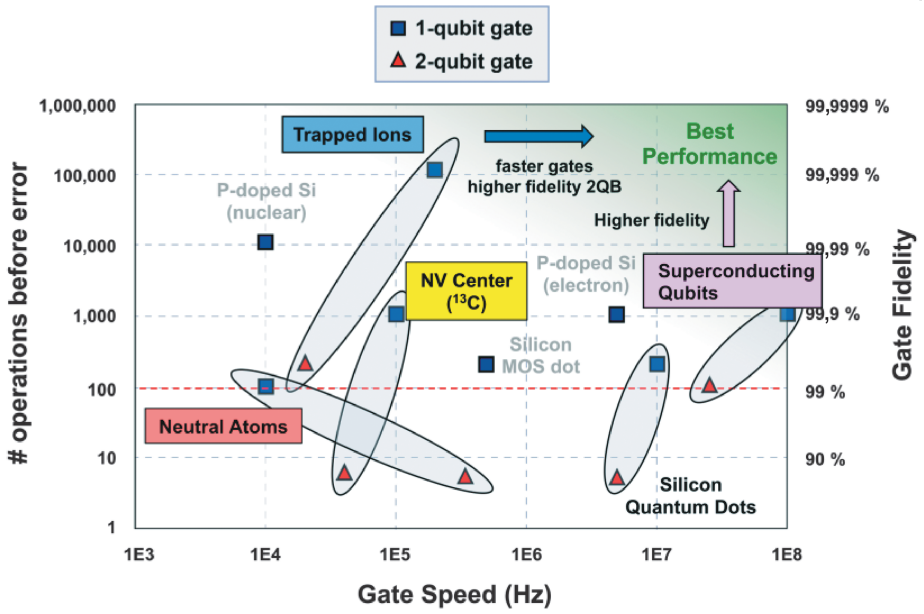


Рис. 1.141. Тенденции развития различных типов кубитов

На рис. 1.141 представлено сравнение кубитов по трем параметрам: тактовой частоте квантовых вентилей (число вентилей, которые могут выполняться в секунду), по числу операций до возникновения ошибок и точности квантовых вентилей. В последнее время наблюдается стремительное развитие кубитов различных типов.

1.25.4. ОБЛАСТИ ПРИМЕНЕНИЯ КВАНТОВЫХ КОМПЬЮТЕРОВ

Ключевым отличием квантовых вычислений от классических является их значительная вычислительная мощность. Квантовые компьютеры могут быть использованы для решения NP-полных задач в области информационной безопасности, машинного обучения, квантовой химии, моделирования новых материалов и лекарств.

Теоретическое обоснование концепции использования квантовой механики для повышения производительности вычислений было дано при публикации алгоритма Шора[6], обеспечивающего экспоненциальное ускорение

[6] *Shor P.W.* Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM J. Comput., 1997. Vol. 26. No. 5. Pp. 1484–1509.

для факторизации числа. Спустя семь лет с помощью этого алгоритма компанией IBM было экспериментально показано разложение числа 15 с использованием кубитов на основе явления ядерного магнитного резонанса. Другим подтверждением эффективности является алгоритм Гровера, обеспечивающий полиномиальное ускорение для неструктурированного поиска. В настоящее время доступно множество алгоритмов, в том числе алгоритмы машинного обучения, оптимизации на основе алгоритма квантовой приближенной оптимизации (QAOA) и, в частности, применения в теоретической химии и материаловедении.

Несмотря на то что для создания универсального квантового компьютера необходимо решить огромное количество фундаментальных и научно-прикладных задач, квантовые вычисления уже находят применение в жизни. Такие компании, как Google и Microsoft, начали работу над квантово-безопасными алгоритмами шифрования. Хотя в настоящее время они находятся на стадии тестирования, ожидается, что квантово-безопасные алгоритмы помогут защитить банковские транзакции, военную связь, медицинские записи и т. д. В сотрудничестве с Recruit Communication компания D-Wave Systems стала первой в применении квантовых вычислений для оптимизации маркетинга, рекламы и коммуникации. Концерн Volkswagen продемонстрировал использование квантовых вычислений для оптимизации трафика. Компания The Weather Company в сотрудничестве с Национальным центром атмосферных исследований и Университетской корпорацией атмосферных исследований в США IBM создает модель квантовых вычислений, которая может оценивать грозы на региональном уровне.

В развитии квантовых вычислений выделяют несколько этапов.

1. Квантовое превосходство — академическое определение способности построить квантовый компьютер, который превосходит возможности классического компьютера при решении какой-либо задачи. В настоящее время квантовое превосходство принадлежит компаниям Google на сверхпроводниковой платформе и Научно-техническому университету Китая (USTC) на фотонных интегральных схемах.

2. Квантовые симуляторы — системы, позволяющие моделировать другие квантовые системы, что может быть использовано при исследовании материалов, комбинаторной оптимизации и обработке больших данных. Узконаправленные вычислители моделируют системы, чтобы предоставить информацию об их поведении. Например, система из одиночных холодных атомов в оптической решетке, создаваемой лазерными лучами, может моделировать процессы, происходящие в сложных твердотельных системах, таких как высокотемпературные сверхпроводники. Кроме того, с помощью квантовых симуляторов уже в настоящее время можно рассчитать свойства простых материалов, исследовать свойства атомного ядра или разработать более эффективные алгоритмы для компенсации ошибок, отработать настройки квантовой системы и т. д.

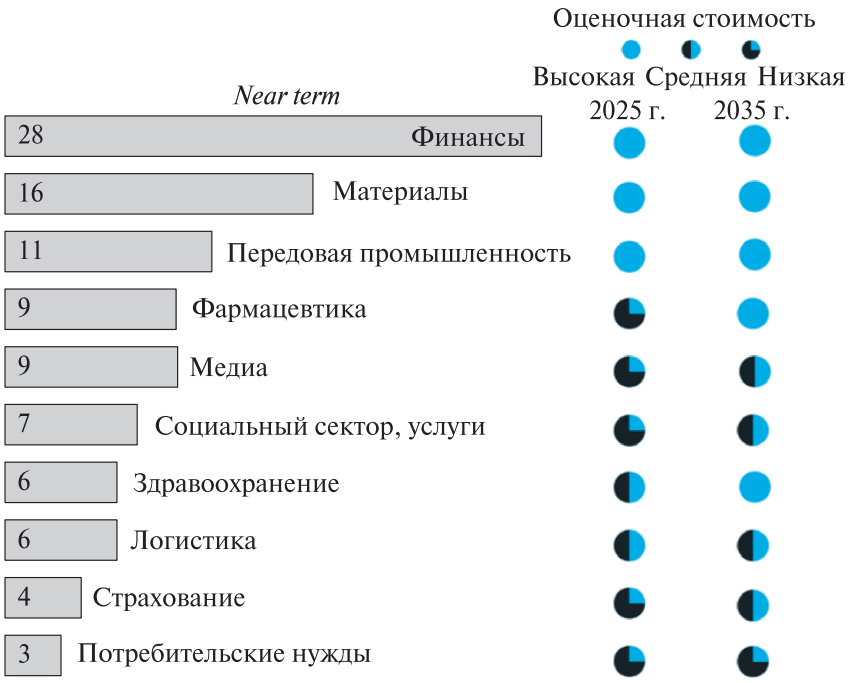


Рис. 1.142. Прогноз Компании McKinsey & Company

3. Универсальный квантовый компьютер, на его создание могут потребоваться десятки лет. Ключевым отличием квантовых вычислений от классических является их огромная вычислительная мощность. Квантовые компьютеры могут быть использованы для решения NP-полных задач в области информационной безопасности, машинного обучения, квантовой химии, моделирования новых материалов и лекарств.

Инфографика распределения квантовых вычислений в различных областях (прогноз от Компании McKinsey & Company) представлена на рис. 1.142.

Оптимизация. Самый известный пример проблемы, допускающей полиномиальное квантовое ускорение, это неструктурированный поиск, нахождение отмеченного элемента из списка в базе данных, что необходимо для оптимизации цепочек поставок и производства. Оптимизация включает в себя выбор наилучшего из многих и важна для решения множества практических задач в таких областях, как инженерное проектирование, логистика, производство и финансы. Сложность решения оптимизационных задач возрастает с количеством элементов, возможных взаимодействий и ограничений. Ожидается, что квантовые компьютеры обеспечат наибольшую пользу именно при решении сложных задач оптимизации. Например, квантовые вычисления могут быть использованы для снижения затрат на производ-

ственные процессы и сокращения времени цикла сложных процессов. Еще одно применение — квантовая оптимизация кредитных портфелей, чтобы кредиторы могли высвободить капитал, снизить процентные ставки и улучшить свои предложения. Оптимизация является основной подзадачей в приложениях машинного обучения и требует множества примеров оптимальных и почти оптимальных решений. Процессоры квантового отжига от коммерческой Компании D-Wave Systems позволяют получить доступ в режиме онлайн в модели облачных вычислений. Решение задач оптимизации трафика и поездок включено в пилотные программы компаний D-Wave и Volkswagen для оптимизации транспортных потоков в Пекине, Барселоне и Лиссабоне. В последнем случае парк автобусов двигался по разным маршрутам, адаптированным к условиям движения в реальном времени с помощью квантового алгоритма, который Концерн Volkswagen продолжает настраивать после каждого пробного запуска.

Криптография. Предполагается, что квантовые компьютеры достаточного размера могут использоваться для взлома существующих криптографических протоколов (в частности, шифрования RSA). Целочисленная факторизация, которая лежит в основе безопасности криптографических систем с открытым ключом, считается вычислительно невыполнимой с обычным компьютером для больших целых чисел, если они являются произведением нескольких простых чисел (например, произведения двух 300-значных простых чисел). Для сравнения: квантовый компьютер мог бы эффективно решить эту проблему, используя алгоритм Шора, чтобы найти ее факторы. Эта способность позволит квантовому компьютеру взломать многие криптографические системы, используемые в настоящее время. В частности, большинство популярных шифров с открытым ключом основаны на сложности факторизации целых чисел или проблеме дискретного логарифмирования, которые могут быть решены с помощью алгоритма Шора.

Квантовая сеть. Коммуникационная сеть будет защищать передаваемые данные с использованием фундаментальных законов квантовой механики. Квантовая сеть является практической реализацией так называемой квантовой криптографии. Квантовые сети формируют важный элемент квантовых вычислений и квантовых систем криптографии. Они допускают транспортировку квантовой информации между физически разделенными квантовыми системами. В распределенных квантовых вычислениях сетевые узлы в сети могут обрабатывать информацию, выполняя функцию квантовых вентилей. Безопасная передача данных может быть реализована с помощью алгоритмов квантового распределения ключей.

Машинное обучение, поддерживаемое или улучшенное квантовыми вычислениями, является относительно новой областью применения, находящейся на стадии оценки и разработки. Квантовый отжиг может быть использован при дискретной оптимизации и выборке, в частности, при выборке Больцмана, которая является основным методом машинного обучения. Из-

вестные приложения включают в себя распознавание изображений и вывод образов для транспортных средств и здравоохранения, нейронные сети и системы рекомендаций. Достижения в алгоритмах квантового машинного обучения также привели к усовершенствованию классических алгоритмов, что уменьшило потенциальные выгоды.

В 2020 г. Компания Google в сотрудничестве с Концерном Volkswagen и Университетом Ватерлоо запустила программу TensorFlow Quantum для ускорения разработки квантовых вычислений, которая представляет собой библиотеку с открытым исходным кодом, используемую для прототипирования моделей квантового машинного обучения.

Разработка лекарств входит в список особенно важных применений квантовых вычислений. Фармацевтическим компаниям может потребоваться более 10 лет и затраты миллиардов долларов, чтобы открыть или разработать новое лекарство, для которого ученые проводят сотни миллионов сравнений на обычных компьютерах. Однако возможности обработки традиционных компьютеров ограничены, поскольку они могут анализировать молекулы только до определенного размера. Эту проблему можно решить с помощью квантовых вычислений. По мере того как алгоритмы и аппаратное обеспечение квантовых вычислений становятся все более распространенными, они сделают возможным сравнение гораздо более крупных молекул. Это может резко сократить время и затраты, связанные с разработкой лекарств, что позволит ученым находить лекарства от различных болезней быстрее, чем ожидалось.

Реклама и маркетинг. Применение квантовых вычислений в рекламе и маркетинге является революционным. Квантовые алгоритмы могут помочь создавать и доставлять более качественную рекламу, интерпретируя ассоциации, влияющие на модели покупок. Вместо того чтобы просто использовать историю браузера для показа рекламы, квантовые алгоритмы фокусируются на том, как пользователи себя чувствуют после просмотра рекламы и какие типы рекламы могут помочь установить долгосрочные отношения с клиентами.

В сотрудничестве с Компанией Recruit Communication компания D-Wave Systems стала первой в применении квантовых вычислений для оптимизации маркетинга, рекламы и коммуникации. Они стремятся использовать квантовые вычисления в рекламе для анализа сложных данных за меньшее время и повышения эффективности доставки рекламы своим целевым клиентам.

Интеллектуальная деятельность. Гигантская вычислительная мощность квантовых компьютеров позволит использовать их для осуществления интеллектуальной деятельности различного характера, например, создать на их основе экспертную систему. В памяти такой экспертной системы может храниться огромное количество научных и технических сведений — стандартов, патентов, образцов промышленного оборудования и др.; экспертная система под контролем специалистов будет выдавать готовые чертежи экспериментальных разработок.

1.25.5. КВАНТОВЫЕ КОМПЬЮТЕРЫ В РОССИИ

Российские ученые внесли огромный вклад в развитие квантовых вычислений: от зарождения идеи вычислений Ю. Манина в 1980 г. до реализации первого кубитного процессора. Так, группа ученых из Гарвардского университета и Массачусетского технологического института под руководством М. Лукина, профессора физики из Гарварда и сооснователя Российского квантового центра, создала программируемый квантовый вычислитель на базе 51 кубита, став, таким образом, лидером среди участников «квантовой гонки». Принцип его работы заключался в формировании кубитов на основе «холодных атомов», которые удерживались оптическими «пинцетами» — специальным образом организованными лазерными лучами (рис. 1.143). С помощью этого квантового вычислителя была решена задача моделирования поведения квантовых систем из множества частиц, практически нерешаемая с помощью классических компьютеров.

В 2021 г. специалисты Национальной квантовой лаборатории совместно с Российским квантовым центром (РКЦ) и ФИАН сообщили о создании прототипа квантового компьютера на платформе из 20 ионов, захваченных электромагнитной ловушкой. Российские ученые также работают и над специализированным облачным софтом. Собственные универсальные облачные платформы квантовых вычислений представили Центр квантовых тех-

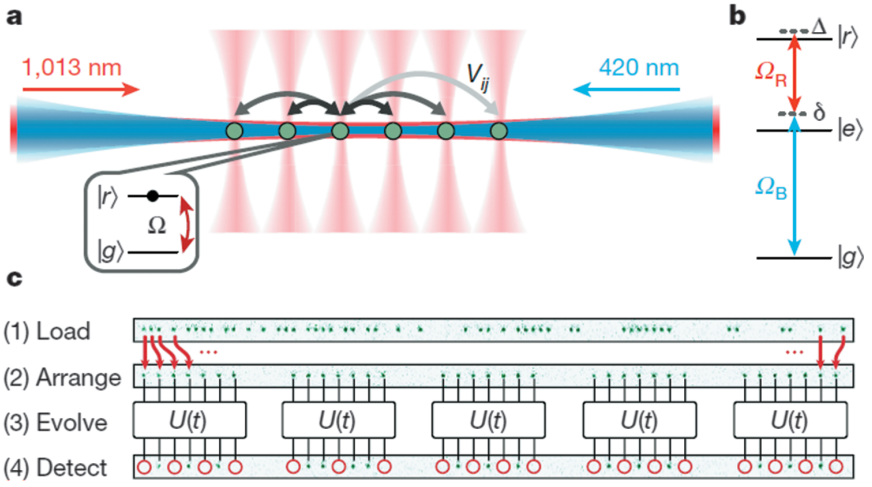
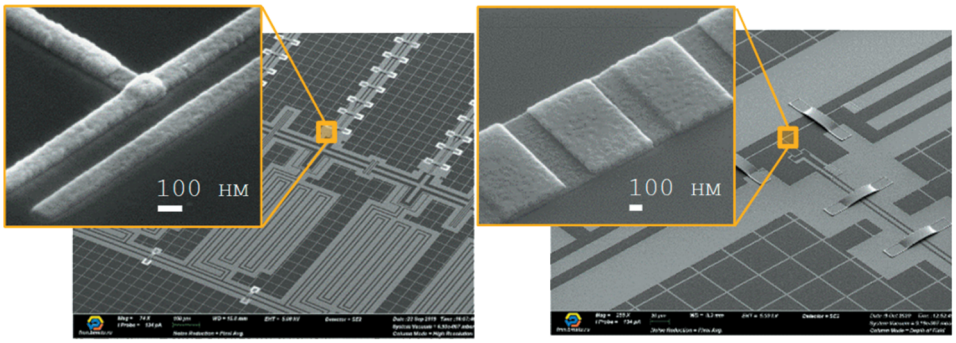


Рис. 1.143. Схема квантового вычислителя на основе «холодных атомов» (*Perdomo-Ortiz A.*, Benedetti *M.*, *Realpe-Gómez J.*, *Biswas R.* Opportu-nities and challenges for quantum-assisted machine learning in near-term quantum computers // Quantum Sci. Technol., 2018. Vol. 3. No. 3)



а *б*

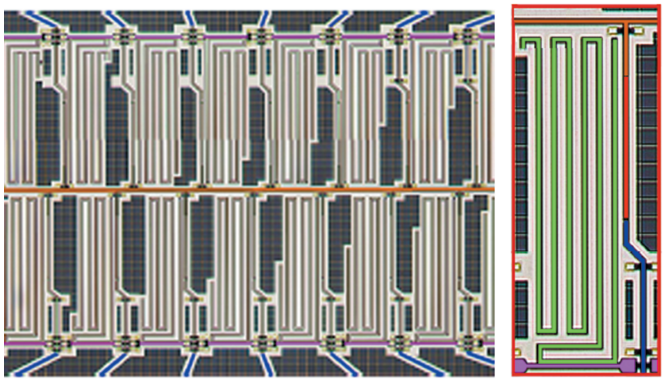
Рис. 1.144. Изображения схем зарядовых (*а*) и потоковых (*б*) кубитов, изготовленных в НОЦ «ФМНС» МГТУ им. Н.Э. Баумана, полученные на СЭМ

нологий МГУ им. М.В. Ломоносова и РКЦ. В 2022 г. была запатентована новая архитектура квантового процессора на основе кудитов, которые в отличие от кубитов имеют более двух энергетических уровней.

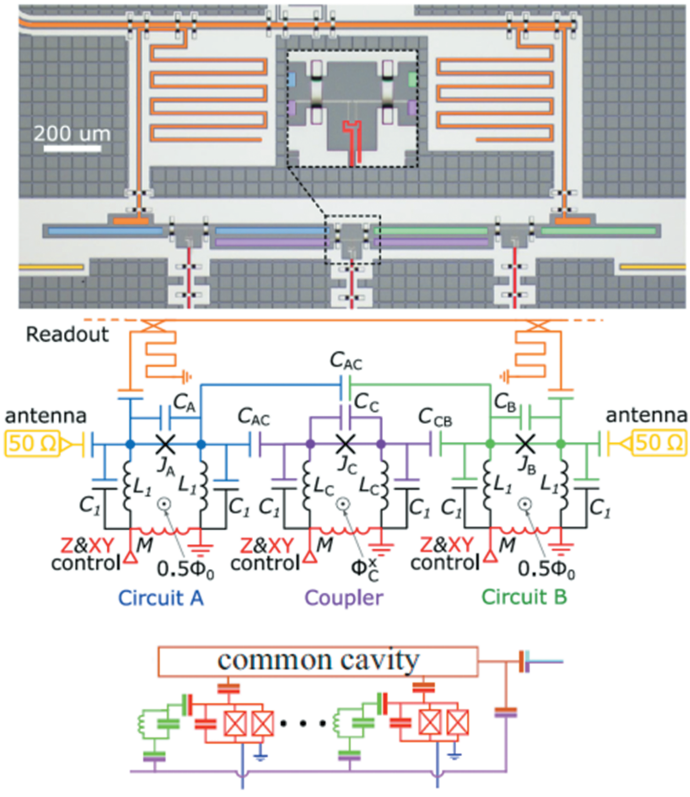
Развитие элементной и измерительной баз для наиболее перспективной — сверхпроводниковой платформы, требующей огромных материальных ресурсов, передового технологического оборудования и высокой квалификации кадров, началось только в 2016 г. при поддержке Фонда перспективных исследований. Исполнителями проекта выступили научный консорциум ведущих российских вузов (МФТИ, МИСиС, НГТУ, МГТУ им. Н.Э. Баумана), ИФТТ РАН, РКЦ и ФГУП «ВНИИА им. Н.Л. Духова», являющийся головной организацией консорциума.

В результате выполнения проекта был сформирован перспективный научный и технологический задел по созданию квантовых компьютеров на основе сверхпроводящих кубитов. Были разработаны уникальные технологии создания элементной базы для сверхпроводящих кубитов с толщинами пленок 20...300 нм, параметром шероховатости не более 1,5 нм, контролируемой толщиной туннельного слоя менее 2 нм и высокой воспроизводимостью линейных размеров в широком диапазоне размеров от 0,08 до 2 мкм (среднеквадратическая ошибка не ниже 5 нм). Изображения изготовленных схем кубитов-трансмонов и флаксониумов, полученные с разрешением до 1 нм на сканирующем электронном микроскопе (СЭМ), представлены на рис. 1.144.

Одной из ключевых характеристик кубитов, необходимой для выполнения логических операций, является «время жизни» кубита. Разработанная в рамках проекта технология создания сверхпроводниковых кубитов-трансмонов позволила обеспечить характерное время когерентности более 50 мкс и высокую точность выполнения операций, что приближает российские разработки в этой области к лучшим мировым аналогам.



а



б

Рис. 1.145. Оптические изображения (*а*) и электрическая схема чипов с 25 перестраиваемыми кубитами-трансмонами (*б*) (*Mazhorin G.S.* et al. Cavity-QED of a quantum metamaterial with tunable disorder // Phys. Rev. A 105. 033519. 2022. Pp. 1–12) и кубитов-флаксониумов с высокой точностью выполнения логических операций (*Moskalenko I.N.*, *Simakov I.A.*, *Abramov N.N.* et al. High fidelity two-qubit gates on fluxoniums using a tunable coupler // Quantum Inf 8, 130. 2022. https://doi.org/10.1038/s41534-022-00644-x), изготовленных в НОЦ «ФМНС» МГТУ им. Н.Э. Баумана

В результате проведенной работы в 2019 г. был реализован алгоритм Гровера на двух кубитах, продемонстрировавший возможность реализации всех необходимых логических операций для универсального квантового процессора: инициализации, одно- и двухкубитных операций и считывания с удовлетворительным для небольших алгоритмов уровнем ошибок. В 2021 г. был продемонстрирован сверхпроводящий чип с 25 перестраиваемыми кубитами-трансмонами, а также получены высокоточные одно- и двухкубитные вентили с точностью соответственно 99,55 и 99,23 % на кубитах-флаксониумах. Оптические изображения (*а*) и эквивалентная электрическая схема устройства (*б*) приведена на рис. 1.145, где различным цветам соответствуют разные элементы схемы (емкостные и индуктивные элементы).

В 2021 г. были созданы квантовые симуляторы на основе сверхпроводящих кубитов для решения квантово-механических задач. Так, квантовая симуляция распространения фотонов в рамках модели Бозе — Хаббарда была выполнена за 2 ч, в то время как на классическом 138-ядерном вычислительном кластере на это ушло почти семь дней машинного времени, а также был создан квантовый симулятор на основе массива из 25 сверхпроводящих кубитов для демонстрации затухания мезоскопической флуктуации, имитирующей приближение к термодинамическому пределу[7].

Тесты к лекции 1.25

1. ***Что такое кубит?***

а) единичный носитель информации в квантовом компьютере;

б) бит информации в кубе;

в) название квантовой пленки.

2. ***В чем заключается основное достоинство квантового компьютера?***

а) в так называемом квантовом параллелизме;

б) в его очень небольшом размере;

в) в очень небольшом энергопотреблении.

3. ***Что является особенностью кубита?***

а) диапазон суперпозиций. В этом диапазоне содержатся в различных соотношениях оба исходных состояния;

б) возможность создания квантового компьютера с очень небольшим энергопотреблением;

в) возможность организации последовательных вычислений.

4. С каким количеством чисел можно одновременно произвести математические операции, если число кубитов равно k?

[7] *Fedorov G.P.* et al. Photon transport in a Bose-Hubbard chain of super-conducting artificial atoms // Physical Review Letters. 2021. Vol. 126. No. 18. Pp. 180503.

а) можно провести математические операции с 2^k числами одновременно;

б) можно провести математические операции с $2k$ числами одновременно;

в) можно провести математические операции с k^2 числами одновременно.

5. *Что представляет собой квантовый параллелизм?*

а) вычислительный процесс, при котором в квантовом компьютере определяются выходные значения всех входных состояний;

б) параллельное выделение квантовых частиц в микромире;

в) процесс образования двух пар частиц, связанных между собой.

2. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Ниже приведены методические указания по выполнению лабораторных работ.

ЛАБОРАТОРНАЯ РАБОТА № 1. ОПРЕДЕЛЕНИЕ ШИРИНЫ ЗАПРЕЩЕННОЙ ЗОНЫ ПОЛУПРОВОДНИКА

Цель лабораторной работы: изучение методики экспериментального определения ширины запрещенной зоны полупроводников на основе исследования температурной зависимости электропроводности.

Задание по лабораторной работе.

1. Получить задание.
2. Определить ширину запрещенной зоны полупроводника.
3. Проанализировать результаты лабораторной работы, сформулировать краткие выводы.

Теоретическая часть

В изолированных атомах электроны могут находиться только в определенных квантовых состояниях и на определенных энергетических уровнях. Разрешенные уровни энергии отделены друг от друга областями запрещенных энергий.

В твердых телах атомы находятся на относительно небольшом расстоянии друг от друга и взаимодействие между атомами достигает значительной величины. В результате электроны каждого атома оказываются в электрическом поле не только своего атомного ядра, но и во внешнем поле соседних ядер (рис. 2.1).

Такие энергетические взаимодействия приводят к расщеплению уровней энергии электронов и образованию энергетической зоны, состоящей из дискретных подуровней, на каждом из которых может находиться по два электрона. Зоны разрешенных энергий отделяются друг от друга зонами запрещенных энергий (запрещенными зонами).

В зависимости от степени заполнения зон электронами и шириной запрещенной зоны ΔE возможны четыре случая заполнения зон электрона (рис. 2.2).

Энергетическая зона, образованная из энергетических уровней внешних валентных электронов атомов, называется валентной зоной. Выше валент-

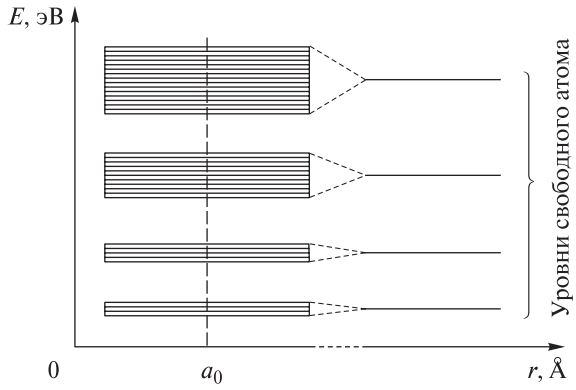


Рис. 2.1. Схема возникновения зон энергии из энергетических уровней атома при их сближении (образовании твердого тела) (a_0 — расстояние между атомами твердого тела)

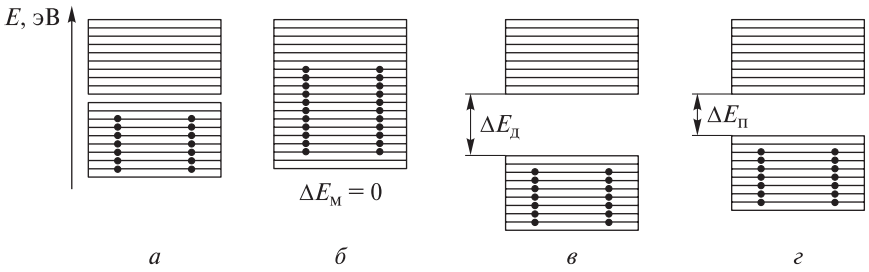


Рис. 2.2. Случаи заполнения зон электронами:

а — валентная зона полностью заполнена электронами, зона проводимости пуста; *б* — валентная зона заполнена полностью, но перекрывается с зоной проводимости; *в* — валентная зона полностью заполнена электронами, зона проводимости пуста и между ними находится широкая запрещенная зона более 3 Эв; *г* — валентная зона полностью заполнена электронами, зона проводимости пуста и между ними находится неширокая запрещенная зона менее 3 Эв

ной зоны в энергетическом спектре электронов располагаются зоны, образованные из пустых энергетических уровней атомов (зоны проводимости).

В общем случае валентная зона полностью заполнена электронами, а зона проводимости пуста (рис. 2.2, *а*, *в*, *г*). В зависимости от ширины запрещенной зоны ΔE твердые тела, имеющие такое зонное строение, являются проводниками, диэлектриками или полупроводниками.

Валентная зона заполнена полностью, но перекрывается с зоной проводимости (рис. 2.2, *б*). В этом случае электроны могут легко переходить на энергетические уровни зоны проводимости и участвовать в электропроводимости. Это характерно для металлов.

Если ширина запрещенной зоны $\Delta E_д$ намного больше энергии теплового движения электронов, определяемой температурой, то электроны не могут

перейти в зону проводимости (т. е. для перехода в зону проводимости тепловой энергии электронов оказывается недостаточно). Такой кристалл является диэлектриком (рис. 2.2, *в*). Если запрещенная зона $\Delta E_{\text{п}}$ достаточно узкая, так что переброс электронов из валентной зоны в зону проводимости может быть осуществлен сравнительно легко даже путем теплового возбуждения, то кристалл является полупроводником (рис. 2.2, *г*).

Электроны в полупроводниках могут осуществлять переход из валентной зоны в зону проводимости из-за относительно небольшой ширины запрещенной зоны ΔE_g ($\Delta E_g < 3$ эВ). Такой переход возможен в результате воздействия внешнего электрического поля, тепловой и световой энергий, ионизирующего излучения и т. п.

В этом случае в валентной зоне возникает свободный энергетический уровень, в зоне проводимости появляется свободный электрон. Незанятое электроном энергетическое состояние в валентной зоне называется *дыркой*, а процесс — генерацией пары носителей зарядов. Генерация носителей зарядов приводит к тому, что электроны могут перемещаться в зоне проводимости, переходя на близлежащие свободные энергетические уровни, а дырки — в валентной зоне, что эквивалентно перемещению положительных зарядов, равных по абсолютной величине зарядам электронов. Перемещение дырки можно представить как заполнение свободных энергетических уровней в валентной зоне электронами близлежащих занятых энергетических уровней.

Электропроводность, обусловленная генерацией пар носителей заряда электрон–дырка, называется собственной проводимостью. Процесс возвращения возбужденных электронов из зоны проводимости в валентную зону, в результате которого пара носителей заряда электрон–дырка исчезает, называется рекомбинацией. Оба процесса — генерация и рекомбинация пар носителей заряда — происходят одновременно, вследствие чего в полупроводнике устанавливается динамическое равновесие, определяющее равновесную концентрацию электронов и дырок. Чем выше температура, тем больше равновесная концентрация, поэтому с повышением температуры удельная проводимость полупроводника должна увеличиваться.

В запрещенной зоне полупроводника возможно расположение энергетических уровней, обусловленных нарушением периодичности кристаллической решетки и называемых локальными уровнями. Локальные уровни, обусловленные примесью, называются примесными. Эти уровни размещаются вблизи как зоны проводимости, так и валентной зоны. В первом случае вероятен переход электрона с занятого примесного уровня в зону проводимости. Такой тип дефекта кристаллической решетки называется *донором*, а создающая его примесь — *донорной*. Во втором случае вероятен переход (захват) электрона из валентной зоны на незанятый примесный уровень и образование дырки проводимости. Дефект кристаллической решетки подобного типа называется акцептором, а соответствующая ему примесь — акцепторной. Электропроводность полупроводника, обусловленная ионизацией атомов акцепторной или донорной примесью, называется примесной электропроводностью. В зависимости от типа примеси возможна электронная или дырочная электропроводность, обусловленная соответственно перемещением электронов и дырок.

Следует отметить, что примесная электропроводность зависит от температуры так же, как и собственная электропроводность.

В примесных полупроводниках наряду с примесной электропроводностью существует собственная проводимость. В собственном полупроводнике концентрации электронов и дырок равны. В примесном полупроводнике преобладает концентрация электронов (для донорной примеси) или дырок (для акцепторной примеси). Подвижные носители зарядов с преобладающей концентрацией называются основными. Так, в полупроводнике с донорной примесью основные носители заряда — электроны, поэтому его называют полупроводником *n*-типа (от англ. *negative* — отрицательный). Дырки в таком полупроводнике являются неосновными носителями заряда. В полупроводнике с акцепторной примесью основные носители зарядов — дырки, поэтому его называют полупроводником *p*-типа (от англ. *positive* — положительный). В этом случае электроны являются неосновными носителями заряда.

Концентрация основных носителей заряда определяет удельную электрическую проводимость полупроводника: чем выше концентрация, тем больше удельная электрическая проводимость.

Температурная зависимость электропроводности полупроводника определяется температурной зависимостью как концентрации носителей заряда, так и их подвижности.

В общем случае удельная электропроводность полупроводника

$$\sigma = N_0 e \mu \text{ [Ом/м]},$$

где N_0 — концентрация носителей заряда; e — заряд электрона; μ — подвижность носителей заряда.

Под подвижностью понимают дрейфовую скорость частицы V_D в электрическом поле напряженностью $E = 1$ [В/см], т. е.

$$\mu = V_D / E \text{ [м}^2/(\text{В} \cdot \text{с})].$$

В полупроводнике под влиянием внешнего электрического поля дрейфовое движение совершают как электроны, так и дырки. Несмотря на то что электроны и дырки движутся в противоположных направлениях, так как их заряды противоположны по знаку, направления электронной и дырочной составляющих дрейфового тока совпадают.

Поэтому для собственного полупроводника плотность дрейфового тока

$$j = j_n + j_p = eE(N_i \mu_n + P_i \mu_p) \text{ [А/м}^2],$$

где N_i, μ_n, P_i, μ_p — концентрации и подвижности электронов и дырок соответственно.

Собственная электропроводность

$$\sigma = eN_i \mu_n + eP_i \mu_p.$$

Концентрация электронов и дырок в собственном полупроводнике, как отмечалось выше, $N_i = P_i$ и определяются по формуле

$$N_i = (N_c N_v)^{1/2} \exp(-\Delta E_g / 2kT) \text{ [1/м}^3],$$

где N_c, N_v — эффективные плотности квантовых состояний в зоне проводимости и валентной зоне соответственно; k — постоянная Больцмана, $k = 1{,}38 \cdot 10^{-23}$ [Дж/К].

Подвижность носителей заряда в значительно меньшей степени зависит от температуры по сравнению с температурной зависимостью их концентрации. Поэтому можно считать, что электропроводность полупроводника растет с температурой примерно по тому же закону, что и концентрация электронов и дырок:

$$\sigma = \sigma_0 \exp(-\Delta E_0 / 2kT)\ [1/\text{м}^3],$$

где σ_0 — электропроводность при $T \Rightarrow \infty$; ΔE_0 — термическая ширина запрещенной зоны.

Понятие *термической ширины запрещенной зоны* обусловлено тем, что при определении этого фундаментального параметра полупроводника не учитывается реальная температурная зависимость подвижности электронов и дырок, а также зависимость самой ширины запрещенной зоны от температуры.

Если прологарифмировать выражение для σ, то оно примет вид

$$\ln\sigma = \ln\sigma_0 - (\Delta E_0 / 2k)\ [1/T].$$

В координатах $\ln\sigma - 1/T$ эта зависимость имеет вид прямой с углом наклона, тангенс которого пропорционален $\Delta E_0 / 2k$ (рис. 2.3).

В примесном полупроводнике зависимость $\sigma(T)$ более сложная. При низких температурах концентрация носителей заряда определяется интенсивностью процесса ионизации примесей, а $\mu \sim T^{3/2}$, при высоких температурах происходит процесс генерации пар, а $\mu \sim T^{-3/2}$. В этом случае электропроводность примесного полупроводника можно выразить как сумму проводимости основной решетки $\sigma_{\text{осн}}$ и проводимости, обусловленной примесью $\sigma_{\text{пр}}$, т. е.

$$\sigma = \sigma_{\text{осн}} + \sigma_{\text{пр}} = \sigma_0 \exp(-\Delta E_0 / 2kT) + \sigma_1 \exp(-\Delta E / 2kT),$$

где ΔE — энергия активации примесных носителей заряда.

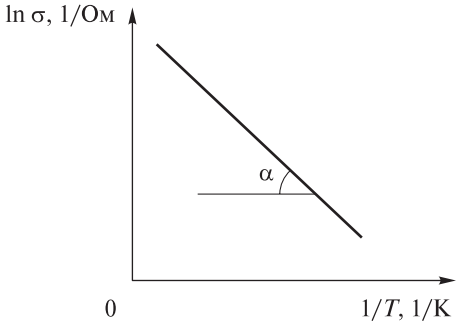


Рис. 2.3. Температурная зависимость электропроводности собственного полупроводника ($\text{tg}\alpha = \Delta E_0/2k$)

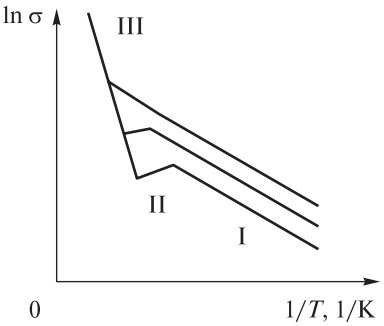


Рис. 2.4. Температурная зависимость электропроводности примесного полупроводника

Зависимость $\ln\sigma - 1/T$ для примесного полупроводника представлена на рис. 2.4.

При низких температурах участок I на кривой соответствует примесной электропроводности. После того как примеси исчерпаны (участок II), электропроводность может несколько уменьшаться за счет падения μ. Участок III соответствует собственной электропроводности.

Если концентрация примеси достаточно велика, участок II отсутствует. Процесс генерации пар начинается, когда примеси еще не исчерпаны.

Ниже приведены значения ширины запрещенной зоны ΔE для некоторых полупроводников.

Материал	Ge	Si	GaAs	CdS	CdSe	PbS	InSb
ΔE, эВ.....	0,72	1,12	1,42	2,42	1,70	0,41	0,17

Экспериментальная часть

Лабораторная работа проводится в следующей последовательности.

Подготовить оборудование к выполнению лабораторной работы. Схема лабораторного макета показана на рис. 2.5. Макет состоит из электрического нагревателя *3* в виде полого цилиндра, внутри которого размещены терморезистор *2* и термодатчик *4*. Термодатчик подключен к цифровому дисплею *3*, показывающему температуру внутри цилиндрического нагревателя. Терморезистор изготовлен из собственного полупроводника, и его выводы подключены к омметру *1*.

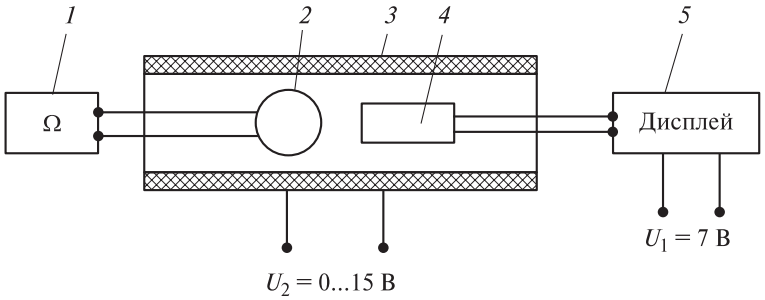


Рис. 2.5. Схема лабораторного макета для измерения ширины запрещенной зоны полупроводников:

1 — омметр; *2* — терморезистор; *3* — электрический нагреватель; *4* — термодатчик; *5* — дисплей термодатчика (цифровой термометр)

Все элементы макета, кроме омметра, смонтированы в пластмассовом корпусе, снабженном разъемами для подключения нагревателя и дисплея к источникам питания. Для подключения омметра на корпусе установлены отдельные разъемы.

Перед выполнением лабораторной работы необходимо собрать лабораторный стенд в следующей последовательности:

1) для обеспечения питания дисплея подключить проводами разъемы макета с обозначением «Дисплей U = 7 В» к первому каналу источника питания, соблюдая полярность;

2) для обеспечения питания электрического нагревателя подключить проводами разъемы макета с обозначением «Нагреватель» ко второму каналу источника питания;

3) подключить проводами омметр к соответствующим разъемам макета. Полярность подключения значения не имеет. В качестве измерительного прибора можно использовать также мультиметр (тестер) или иной универсальный измерительный прибор со встроенным омметром;

4) установить диапазон измерений омметра (измерительного прибора), килоом. Если в процессе измерения сопротивления образца выяснится, что удобнее работать в другом диапазоне, переключить прибор в другой диапазон измерений.

Порядок выполнения экспериментальной части лабораторной работы.

1. Проверить правильность собранной схемы измерений (см. рис. 2.5).

2. Включение питания разрешается только после проверки преподавателем правильности собранной схемы.

3. Включить измерительный прибор и установить режим измерения сопротивления.

4. Установить на источнике питания все ручки регулировки напряжения в положение «0» (в крайнее правое положение против хода часовой стрелки).

5. Включить источник питания.

6. Установить на источнике питания на канале, подключенном к дисплею, напряжение 7 В. Во избежание выхода из строя дисплея категорически запрещается устанавливать напряжение более 7 В.

7. Снять первое значение сопротивления на омметре и значение температуры на дисплее и записать в таблицу журнала.

8. Установить на источнике питания на канале, подключенном к электрическому нагревателю, напряжение 4 В и следить за изменением сопротивления и температуры в течение примерно 10 мин. За это время температура нагревателя повысится и стабилизируется. О стабилизации свидетельствует прекращение изменения одновременно и сопротивления и температуры. Снять показания омметра и дисплея и записать их в таблицу.

9. Повторить п. 8 для напряжений 8 В и 12 В. Во избежание выхода из строя макета категорически запрещается устанавливать напряжение более 12 В.

Завершение работы

После проведения измерений выключить оборудование, соблюдая приведенный ниже порядок:

1) выключить источник питания;
2) выключить измерительный прибор;
3) разобрать схему измерений;
4) сдать лабораторный макет преподавателю.

Содержание отчета

Отчет должен быть оформлен в виде журнала лабораторных работ (прил. 1). Содержание отчета:

1) краткая теоретическая часть в виде ответов на контрольные вопросы;
2) схема стенда для измерения электропроводности полупроводников;
3) таблица с экспериментальными данными и расчеты;
4) график зависимости $\ln\sigma$ от $1/T$ и рассчитанное значение ширины запрещенной зоны;
5) выводы.

Контрольные вопросы и задания

1. Каким образом с точки зрения зонной модели возникают свободные носители заряда в разрешенных зонах?
2. Как объяснить температурную зависимость подвижности в полупроводниках?
3. Чем объясняется различный вид температурной зависимости сопротивления у металлов и полупроводников?
4. Какова природа температурной зависимости ширины запрещенной зоны полупроводников?
5. Дайте классификацию твердых тел по электрическим свойствам на основе зонной теории.
6. Что называется подвижностью носителей заряда и как она зависит от температуры?
7. По каким характерным признакам твердые тела подразделяют на металлы, полупроводники и диэлектрики?
8. Что называется зоной проводимости?
9. Что называется запрещенной зоной?
10. Что называется валентной зоной?
11. Что такое «дырки»?
12. Что понимают под собственной электропроводностью?
13. Что понимают под примесной электропроводностью?
14. Можно ли температурную зависимость электропроводности представить графически в виде прямой линии? Если да, то в каких координатах?
15. Что определяет угловой коэффициент α?

ЛАБОРАТОРНАЯ РАБОТА № 2. ИССЛЕДОВАНИЕ ТЕРМОЭЛЕКТРИЧЕСКИХ ЯВЛЕНИЙ В ПОЛУПРОВОДНИКАХ

Цель лабораторной работы: практическое и теоретическое изучение термоэлектрических явлений (прямого и обратного термоэлектрического эффекта).

Задание по лабораторной работе.

1. Получить задание.
2. Рассчитать параметры прямого и обратного термоэлектрического эффекта.
3. Провести исследование эффекта Пельтье и эффекта Зеебека в полупроводниках.
4. Проанализировать результаты лабораторной работы, сформулировать краткие выводы.

Теоретическая часть

Если привести в соприкосновение два разных металла или металл и полупроводник либо два полупроводника разных типов, то между ними возникнет разность потенциалов, которая называется контактной. Контактная разность потенциалов обусловлена тем, что при соприкосновении часть электронов переходит из одного контактирующего элемента в другой, а у полупроводников переходят также и дырки. Физическая граница контакта является потенциальным барьером. Вследствие различия значений энергии, необходимой для преодаления барьера в каждом из контактируемых элементов, диффузный поток электронов (и дырок в случае полупроводников) через границу не будет уравновешенным, и элемент с меньшей работой выхода зарядится положительно.

В результате различия концентрации носителей зарядов слева и справа от границы контактирующих элементов возникнет скачок потенциала в месте контакта. Возникшее в месте контакта поле объемного заряда будет тормозить дальнейший переход зарядов через границу, и поэтому дальнейший переход зарядов прекратится, а в месте контакта установится некоторая разность потенциалов, называемая внутренней контактной разностью потенциалов.

Термоэлектрические явления в полупроводниках обусловлены эффектами Пельтье, Томсона и Зеебека.

Эффект Пельтье заключается в том, что во время пропускания через контакт двух разнородных полупроводников (а также проводников) электрического тока в зависимости от его направления в контакте выделяется или поглощается количество теплоты, дополнительное к джоулевой теплоте.

Эффект Пельтье проявляет свои свойства в процессах выделения или поглощения количества теплоты ($Q_П$) областью контакта двух разнородных полупроводников или полупроводника и металла.

Численное значение $Q_П$ определяется значением электрического тока I, протекающего через структуру контакта:

$$Q_П = \pm П I t = \pm П q,$$

где П — коэффициент Пельтье; t — время протекания тока; q — количество прошедшего электричества.

Из приведенной формулы следует, что коэффициент Пельтье — это тепловая энергия, переносимая через место контакта, приходящаяся на электрон в единицу времени.

Коэффициентом теплопроводности называется физическая величина, равная количеству теплоты, которое переносится через единичную поверхность за 1 с, если градиент температур равен единице.

В Международной системе единиц (СИ) размерность коэффициента теплопроводности Вт/(м · К).

Возникновение эффекта Пельтье обусловлено различием условий установления равновесного состояния электронов или дырок в контактирующих материалах. При подключении внешнего источника питания к контактирующим материалам возникает ток. При протекании тока через такую структуру внешнее электрическое поле переносит электроны или дырки из одного материала в другой. Свободные заряды, перейдя в другую область контакта, имеют избыток (или недостаток) энергии, которая отдается кристаллической решетке материала. В результате температура контакта растет. Если же заряды обладают меньшей энергией, то кристаллическая решетка отдает им свою энергию и ее температура падает. Предположим, что имеется контакт металла с полупроводником электронного типа. Пусть работа выхода электрона в металле больше, чем в полупроводнике электронной проводимости ($A_м > A_П$). В этом случае электроны из полупроводника будут переходить в металл, оставляя в области своего ухода некомпенсированные положительные ионы, т. е. в области контакта возникает потенциальный барьер (рис. 2.6).

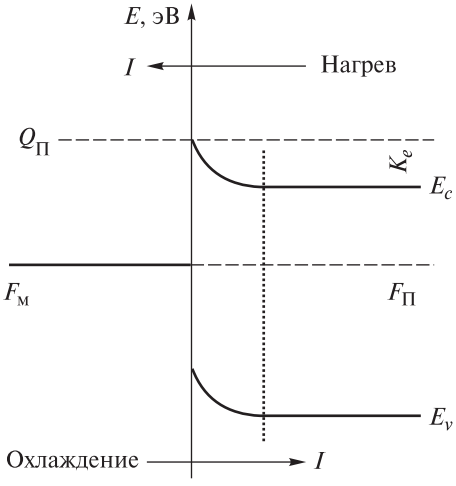


Рис. 2.6. Схема возникновения эффекта Пельтье в контакте электронного полупроводника с металлом

Если на электроды такого контакта подать внешнее напряжение («плюс» — на металл, «минус» — на полупроводник), то электроны, переходя в металл, опускаются до уровня Ферми $F_П$. В этом процессе они взаимодействуют с фононами, дефектами, друг с другом, отдают избыток энергии решетке в виде коли-

чества теплоты, называемого теплотой Пельтье $\Delta Q_{\text{П}}$. Численное значение $\Delta Q_{\text{П}}$ определяют из равенства

$$\Delta Q_{\text{П}} = K_e - F_{\text{П}}.$$

При смене полярности внешнего напряжения (на металл — «минуса», а на полупроводник — «плюса») электроны, находящиеся в металле на энергетических уровнях, располагающихся выше дна зоны проводимости, переходят из металла в полупроводник, получая дополнительную энергию за счет атомов решетки, расположенных вблизи контакта. Происходит поглощение теплоты Пельтье, приводящее к охлаждению контакта. Количество поглощенной теплоты определяют, используя равенство для $\Delta Q_{\text{П}}$.

При использовании невырожденных полупроводников средняя кинетическая энергия электронов в полупроводниках *n*-типа

$$K_e = (\beta + 2)kT.$$

Тогда, подставляя K_e в $\Delta Q_{\text{П}}$, получим выражение для вычисления коэффициента Пельтье:

$$\Pi = \frac{K_e - F_{\text{П}}}{e} = \frac{\left[(\beta + 2)kT - F_{\text{П}}\right]}{e},$$

где β — коэффициент, определяемый типом рассеяния; k — коэффициент Больцмана; T — температура полупроводника.

Уровень Ферми

$$F_{\text{П}} = kT \ln \frac{n}{N_C},$$

где n — концентрация электронов в полупроводнике;

$$N_C = 2\frac{\left(2\pi m_n^*\right)^{3/2}}{h^2},$$

m_n^* — эффективная масса электрона; h — постоянная Планка.

Подставляя в выражение для П значение $F_{\text{П}}$ для электронного полупроводника и учитывая, что

$$\alpha_1 = \frac{k}{e\left(\beta + 2 + \dfrac{\ln N_C}{n}\right)},$$

получаем:

$$\Pi = \alpha_1 T.$$

Это равенство выведено Томсоном из первого и второго начала термодинамики для свободно движущихся частиц, поэтому аналогичные рассуждения справедливы и для контактов дырочного полупроводника с металлом при прохождении через него тока.

Эффект Пельтье возникает и в *p*–*n*-переходах полупроводниковых приборов. Так, если на *p*–*n*-переход подать внешнее напряжение в прямом направлении, то электроны *n*-области и дырки *p*-области *p*–*n*-перехода, двигаясь навстречу друг другу, взаимодействуют (рекомбинируют). Причем электрон, попав в *p*-область *p*–*n*-перехода, занимает в валентной зоне место дырки. Энергия, высвобождаемая в результате рекомбинации, поглощается решеткой, в результате область контакта нагревается. Величина этой энергии существенно выше, чем при контакте двух металлов, поэтому наиболее сильно эффект Пельтье наблюдается в случае использования полупроводников *p*- и *n*-типов проводимости.

При пропускании через *p*–*n*-переход тока в запорном направлении электроны *p*-области переходят в *n*-область, в которой их энергия становится больше, т. е. на образование электронно-дырочной пары затрачивается энергия, заимствованная у решетки, область контакта охлаждается.

Таким образом, при прохождении через *p*–*n*-переход электрического тока различие средней энергии электронов в областях *n*- и *p*-типов проводимости приводит к выделению или поглощению теплоты Пельтье (рис. 2.7).

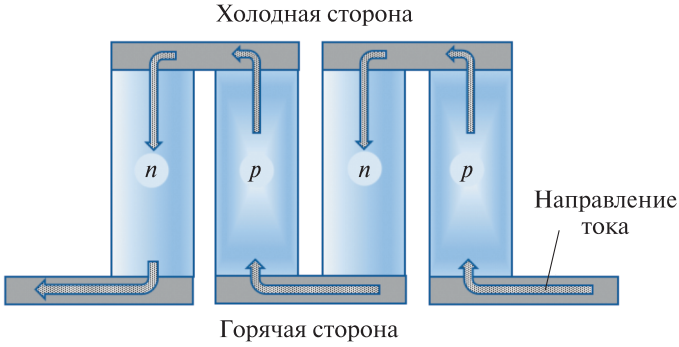


Рис. 2.7. Схема, поясняющая выделение или поглощение теплоты Пельтье на контакте двух полупроводников с различными типами проводимости

Электрический способ охлаждения при прохождении тока через контакт двух полупроводников получил практическое применение в различных охлаждающих устройствах: микрохолодильниках, термостатах, термоэлектрических гигрометрах и др.

Эффект Томсона проявляется через выделение или поглощение количества теплоты в однородном полупроводнике при наличии продольного градиента температуры dT/dx и протекании электрического тока через образец. Таким образом, количество теплоты Томсона пропорционально прошедшему через полупроводник количеству электричества и градиенту температуры вдоль полупроводника:

$$G_\tau = \tau It\frac{dT}{dx} = \tau e\frac{dT}{dx},$$

где τ — коэффициент Томсона, зависящий от материала полупроводника.

Основой возникновения эффекта Томсона служит значительное увеличение тепловой скорости электронов в области максимальной температуры. Это обусловливает их уход в более холодную область полупроводника, где скорость электронов меньше. Уменьшение скорости сопровождается снижением амплитуды колебаний теплового движения электронов, что позволяет высокоэнергетичным электронам легко проникать в холодную область. Это приводит к тому, что при прохождении электрического тока в образце в направлении возрастания температуры электроны в полупроводнике *n*-типа проводимости будут переходить в обратном направлении, т. е. из мест с более высокой температурой в область с более низкой температурой. Избыток энергии электронов в холодной области поглощается решеткой, т. е. происходит ее нагрев теплотой Томсона. При изменении направления тока происходит поглощение теплоты Томсона и кристалл охлаждается.

Поскольку основой эффекта Томсона служат особенности теплового движения электронов, коэффициент Томсона имеет связь с коэффициентом термоЭДС:

$$\tau = T\frac{d\alpha}{dT},$$

где T — средняя температура полупроводника; $d\alpha$ — изменение коэффициента термоЭДС при изменении температуры на dT.

Эффект Зеебека — возникновение ЭДС (термоЭДС) в электрической цепи, состоящей из последовательно соединенных разнородных проводников или полупроводников при наличии градиента температур в контактах металл–полупроводник, полупроводников с разным типом проводимости или с разной шириной запрещенной зоны (гетеропереходы). Эффект Зеебека применяют в термопарах для измерения температуры, в энергетических установках космического назначения и т. д. Пусть в стержне невырожденного полупроводника контакты имеют разную температуру, причем $T_2 > T_1$ (рис. 2.8).

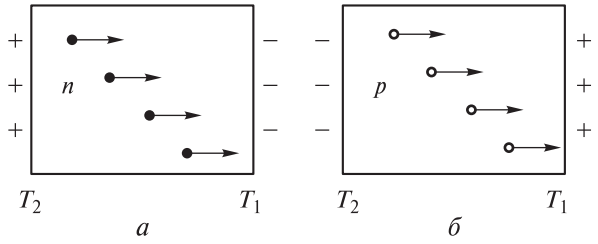


Рис. 2.8. Схема передвижения свободных носителей заряда в области градиента температур:
а — донорный полупроводник (*n*-типа); *б* — акцепторный полупроводник (*p*-типа)

В стержне с градиентом температуры действуют два механизма движения свободных электронов:

1) в нагретой области стержня происходит увеличение концентрации и энергии свободных электронов. Поэтому, стремясь ослабить градиент, они под действием теплового движения начинают диффундировать в холодную область полупроводникового стержня;

2) в области низких температур (значительно меньше комнатной) к этому механизму добавляется процесс увлечения свободных электронов фононами, движущимися от горячего конца стержня к холодному, так как их концентрация в этом случае значительна, и пренебрегать влиянием фононов на характер движения электронов уже нельзя.

Диффузионное перемещение электронов приводит к образованию разности потенциалов, называемой термоэлектрической силой (термоЭДС). В условиях термодинамического равновесия термоЭДС формирует поток электронов равный, но противоположно направленный потоку диффузионных электронов. Тогда при одинаковой температуре металла и полупроводника контактные разности потенциалов в области контакта будут равны по величине и направлены в противоположные стороны. Если же выполняется неравенство $T_2 > T_1$, то возникает разность контактных разностей потенциалов $\Delta\varphi_k$, вносящая свой вклад в термоЭДС полупроводника.

Таким образом, возникновение термоЭДС E_T обусловлено двумя механизмами: объемной диффузией носителей заряда от горячего конца к холодному и контактной диффузией, возникающей из-за температурной зависимости контактной разности потенциалов:

$$E_T = \left(\frac{\partial \varphi_k}{\partial T} + \frac{\partial \varphi_{\text{об}}}{\partial T} \right),$$

$$\Lambda T = E_k + E_{\text{об}},$$

где φ_k — разность потенциалов за счет контактной диффузии; $\varphi_{\text{об}}$ — разность потенциалов за счет объемной диффузии носителей заряда от горячего конца стержня к холодному; E_k — термоЭДС за счет контактной диффузии; $E_{\text{об}}$ — термоЭДС за счет объемной диффузии носителей заряда от горячего конца стержня к холодному.

При объемной диффузии наличие градиента температуры вызывает одностороннюю диффузию потока электронов от горячего конца полупроводника к холодному, обусловленную градиентом концентрации свободных носителей [$n = f(x)$] и тепловым движением при постоянном коэффициенте диффузии электронов D_n (см. рис. 2.8).

Это условие позволяет использовать равенство

$$P_n = D_n \frac{dn}{dx},$$

где P_n — плотность потока электронов при градиенте концентрации; n – число электронов; x — расстояние.

И наоборот, при постоянной концентрации свободных носителей и изменении температуры происходит изменение коэффициента $D_n = f(T)$, которое обусловливает возникновение потока электронов:

$$P_D = n\frac{dD_n}{dx},$$

где P_D — плотность потока электронов, обусловленная изменением коэффициента диффузии.

Коэффициент диффузии электронов D_n зависит от средней кинетической энергии K_e, которая определяется механизмом рассеяния носителей заряда и температуры:

$$D_n = CK_e^{\beta+1/2},$$

где C — коэффициент пропорциональности.

Если коэффициент D_n и концентрация n температурозависимы, то в случае динамического равновесия свободных носителей заряда, например, электронов, в изолированном полупроводнике справедливо равенство

$$D_n\frac{dn}{dx} + n\frac{dD_n}{dx} = E\mu_n n,$$

где E — напряженность поля, создаваемого термоЭДС; μ_n — подвижность электронов.

Пусть коэффициент диффузии будет постоянной величиной ($D_n = \text{const}$), а концентрация свободных электронов будет температурозависимой ($n = f(T)$), тогда последнее уравнение можно записать в следующем виде:

$$D_n\frac{dn}{dx} = D_n\frac{\delta_n}{\delta T}\frac{\delta T}{\delta x} = E\mu_n n.$$

Подвижность свободных электронов в условиях равновесия можно определить из уравнения Эйнштейна:

$$\frac{D}{\mu} = \frac{kT}{e}.$$

Решив совместно последние два выражения, получим выражение для коэффициента термоЭДС:

$$\alpha_n = \frac{Edx}{dT} = \frac{k}{e}T\frac{\delta\ln(n)}{\delta T}.$$

При постоянной концентрации свободных электронов ($n = \text{const}$) и температурозависимом коэффициенте диффузии ($D_n = f(T) = f(T_e)$) уравнение для α_n упрощается до вида

$$n\frac{dD_n}{dx} = n\frac{\delta D_n \delta K_e}{\delta K_e dx} = E\mu_n n.$$

Подставив в уравнение для α_n среднюю энергию электронов, представленную в предыдущем уравнении, получим выражение для определения численного значения коэффициента термоЭДС:

$$\alpha_D = \frac{Edx}{dT} = \frac{k}{e}\left(\beta + \frac{1}{2}\right).$$

Коэффициенты α_n и α_D позволяют определить конкретные значения объемной термоЭДС ($E_{об}$).

При тепловом нагреве контактной пары зависимостью работы выхода металла от температуры пренебрегают вследствие ее малости, поэтому изменение контактной термоЭДС полностью определяется работой выхода полупроводника, выражаемой через изменение уровня Ферми.

Контактная разность потенциалов

$$\varphi_k = \frac{A_м - A_п}{e} = \frac{F_м - F_п}{e}.$$

Дифференцируя выражение для φ_k по температуре, можно получить выражение для коэффициента контактной термоЭДС:

$$\alpha_k = \frac{\delta\varphi_k}{\delta T} = -\frac{1}{e}\frac{\delta F_n}{\delta T}.$$

Для невырожденных полупроводников при $k_T >> \Delta E_д$, где $\Delta E_д$ — диапазон энергий примесных атомов донорного типа, уровень Ферми

$$F_П = kT\ln\frac{n}{N_C};$$

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

Подставляя N_C в $F_П$, получаем выражение для расчета контактной разности потенциалов:

$$\alpha_k = -\frac{k}{e}\left(-\frac{3}{2} + \frac{F_n}{kT} + T\frac{\delta\ln n}{\delta T}\right).$$

В режиме $n >> p$ и $kT >> \Delta E_д$ суммарный коэффициент термоЭДС электронного полупроводника

$$\alpha_1 = \alpha_n + \alpha_D + \alpha_k = \frac{k}{e\left(2 + \beta + \frac{\ln N_C}{n}\right)}.$$

Проведя аналогичные рассуждения для полупроводника дырочной проводимости при условии $p >> n$ и $kT >> \Delta E_д$ и с учетом только диффузионного механизма, получим выражение для коэффициента термоЭДС:

$$\alpha_2 = \alpha_p + \alpha_D + \alpha_k \frac{k}{e\left(2 + \beta + \frac{\ln N_V}{p}\right)}.$$

Анализ выражений для α_1 и α_2 показывает, что в примесных полупроводниках коэффициенты слабо зависят от температуры, т. к. они пропорциональны $\ln(T)^{3/2}$ и обратно пропорциональны логарифму концентрации носителей. В области полной ионизации атомов примеси, наблюдаемой при высоких температурах, выражения для α_1 и α_2 можно записать в следующем виде:

$$\alpha_1 = \frac{k}{e\left(2 + \beta + \frac{\ln N_C}{N_д}\right)};$$

$$\alpha_2 = \frac{k}{\left(2 + \beta + \frac{\ln N_V}{N_a}\right)}.$$

Из последних равенств следует, что значение α в примесных полупроводниках растет с увеличением температуры при условии неполной ионизации атомов примеси и уменьшается с увеличением концентрации примеси донорного (N_d) или акцепторного (N_a) типов.

Для нахождения численного значения коэффициента α в режиме перехода от истощения примеси к собственной проводимости необходимо учитывать носители обоих знаков, тогда выражения, полученные для примесного полупроводника, можно обобщить на случай собственной проводимости, используя выражение

$$\alpha = \frac{\alpha_2 \mu_p p - \alpha_1 \mu_n n}{\mu_p p + \mu_n n}.$$

В собственных полупроводниках электрический ток переносится электронами и дырками, поэтому, как следует из выражения для α, их термоЭДС значительно ниже, чем в случае примесных полупроводников. Это означает, что в процессе истощения примеси и при переходе к собственной проводимости численное значение коэффициента α уменьшается. Уровень Ферми в собственных полупроводниках располагается в области середины запрещенной зоны. С изменением температуры изменяется в равном количестве концентрация электронов и дырок. Уровень Ферми практически не изменяет своего положения в запрещенной зоне при выполнении равенства $m_n^* \approx m_p^*$, поэтому контактную термоЭДС можно принять равной нулю в объеме полупроводника.

В полупроводниках *p*-типа проводимости при температурах, меньших температуры истощения примесных атомов (T_1), выполняется неравенство $p >> n$, поэтому возникновение термоЭДС обусловлено концентрацией дырок, что следует также из выражения для α. При повышении температуры от T_1 до T_2 термоЭДС медленно увеличивается, так как возрастает концентрация собственных носителей заряда в полупроводнике. Увеличение температуры до T_3 приводит к переходу от примесной проводимости к собственной и, согласно выражению для α, термоЭДС изменяет знак. При дальнейшем увеличении температуры она уже не изменяется до температуры T_4, так как преобладает электронная составляющая проводимости полупроводника. С дальнейшим ростом температуры от T_4 до T_5 термоЭДС вновь уменьшается.

В большинстве практических случаев для определения термоЭДС можно использовать следующие свойства.

1. В небольшом диапазоне температур значение термоЭДС можно считать пропорциональным разности температур ΔT:

$$E_T = \alpha \Delta T.$$

2. Если термоэлектрическая цепь составлена из последовательно соединенных электронного и дырочного полупроводников, то термоЭДС отдельных элементов

$$E_{12} = (\alpha_1 + \alpha_2)\Delta T = \alpha_{12}\Delta T,$$

где α_{12} — коэффициент дифференциальной термоЭДС, характеризующий пару полупроводников.

Таким образом, коэффициент α_{12} зависит в первую очередь от материала ветвей термопары и от интервала температур, в котором используется термопара.

Структура кристаллической решетки металлов обладает системой полной ионизации матричных атомов, концентрация которых не изменяется при изменении температуры. Из этого следует, что и контактная разность потенциалов также практически не зависит от температуры. Поэтому термоЭДС металлов практически не превышает нескольких микровольт на градус.

В полупроводниках наблюдается значительная зависимость от температуры:

1) кинетической энергии и концентрации свободных носителей заряда;

2) контактного потенциала и связанного с ним уровня Ферми.

Это приводит к увеличению термоЭДС до 1 мкВ/градус и в случае контакта металл–полупроводник возникновению термоЭДС исключительно в объеме кристаллической структуры полупроводника. Это, в свою очередь, позволяет по направлению термоЭДС определять тип проводимости кристалла. Таким образом, при отрицательно заряженном холодном конце термопары свободными носителями заряда являются электроны (полупроводник *n*-типа проводимости), а если заряд положительный, то дырки (полупроводник *p*-типа проводимости).

При слабых токах все три термоэлектрических явления считаются обратимыми и при стационарном тепловом режиме изменяют знак с изменением

разности температур и направления тока. В этом случае, используя закон сохранения энергии и разность температур между контактами dT, можно определить связь между коэффициентами α, Π и τ в замкнутой цепи из двух материалов, анализируя выражение

$$\alpha_{12} dTdI = \frac{\delta \Pi_{12}}{\delta T} dTdI + (\tau_1 - \tau_2) dTdI.$$

Это равенство позволяет определить соотношение между коэффициентами α_{12}, Π_{12} и τ_1, τ_2 при контакте двух полупроводников:

$$\alpha_{12} = \frac{\delta \Pi_{12}}{\delta T} + (\tau_1 - \tau_2).$$

Таким образом, это равенство аналитически объединяет в общую группу термоэлектрические эффекты.

Термоэлементы на основе полупроводников намного более эффективные, чем на основе металлов. На рис. 2.9 представлено конструктивное исполнение и внешний вид термоэлемента, который используется в данной лабораторной работе.

Термоэлектрический модуль (ТЭМ), или элемент Пельтье, представленный на рис. 2.9, — это конструктивно завершенное устройство, состоящее из большого числа (например, 127) термопар, соединенных между собой коммутационными пластинами, расположенных в прямоугольной форме и помещенных между двумя тонкими керамическими пластинами.

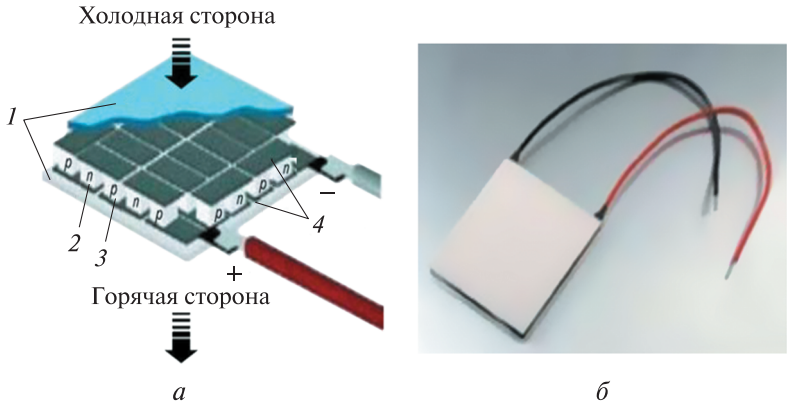


Рис. 2.9. Конструктивное исполнение (*а*) и внешний вид (*б*) термоэлектрического модуля (https://kit-e.ru/wp-content/uploads/120_7-5.jpg):
1 — изолятор (керамика); *2* — полупроводник *n*-типа; *3* — полупроводник *p*-типа; *4* — проводник (медь)

Число термопар может быть изменено в широких пределах: от единиц до сотен, что позволяет конструировать ТЭМ с холодильной мощностью от десятых долей до сотен ватт с рабочим напряжением от десятых долей до десятков вольт.

Экспериментальная часть

Принципиальная схема экспериментальной установки для изучения характеристик термоэлемента представлена на рис. 2.10.

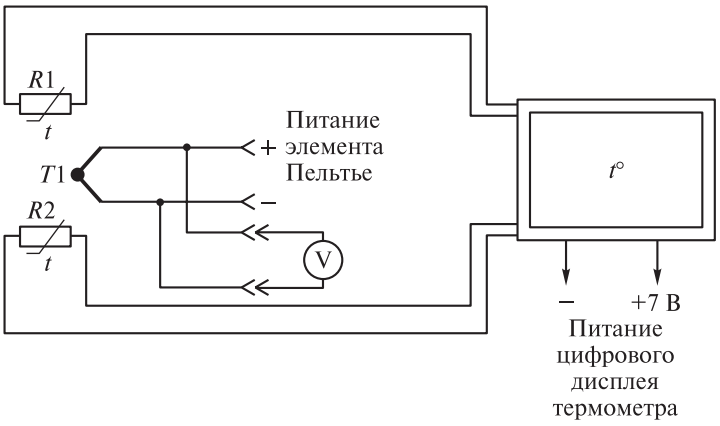


Рис. 2.10. Принципиальная схема экспериментальной установки

На термоэлемент *T*1, представляющий собой модуль Пельтье, подается напряжение от источника питания. В результате прохождения тока через термоэлемент одна сторона модуля нагревается, а противоположная сторона охлаждается. Температура сторон модуля измеряется с помощью терморезисторов *R*1 и *R*2, которые подсоединены к цифровому дисплею.

Для измерения термоЭДС на отдельные контакты подключается милливольтметр, при этом источник питания должен быть отсоединен от макета.

Объектом исследования является термоэлемент Пельтье типа ТЕС1-12706 или его аналог, установленный на корпусе макета. На обеих сторонах термоэлемента установлены термодатчики, подсоединенные к двухканальному цифровому термометру, снабженному цифровым дисплеем.

Используемые приборы и оборудование:

– источник постоянного тока типа АТН-2031 или его аналог — 1 шт.;
– лабораторный макет — 1 шт.;
– цифровой вольтметр или мультиметр — 1 шт.;
– соединительные провода;
– источник теплоты (фен или лампа накаливания).

Для выполнения лабораторной работы собирают стенд, состоящий из источника постоянного тока (рис. 2.11), цифрового вольтметра или мультиметра, лабораторного макета и соединительных проводов.

ВНИМАНИЕ!

Правый канал источника питания служит для питания цифрового дисплея макета 7 В, левый — для питания элемента Пельтье.

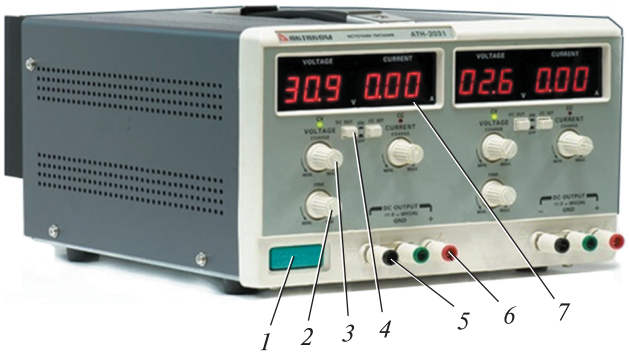


Рис. 2.11. Двухканальный источник питания, работающий в режимах однополярного и двухполярного питания:

1 — кнопка включения (выключения) питания; *2* — ручка точной регулировки выходного параметра; *3* — ручка грубой регулировки выходного параметра; *4* — кнопка переключения режима стабилизации напряжения; *5* — гнездо для подключения отрицательного напряжения; *6* — гнездо для подключения положительного напряжения; *7* — цифровой индикатор

Для выполнения лабораторной работы требуется собрать электрическую схему в соответствии с принципиальной схемой, приведенной на рис. 2.10, для чего выполнить следующие операции:

1) подключить питание дисплея на стороне макета, обозначенного словом «Дисплей», для этого плюсовое гнездо правого канала источника питания соединить проводом с соответствующим плюсовым гнездом макета. Минусовое гнездо правого канала источника питания соединить с соответствующим минусовым выводом макета;

2) подключить питание термоэлемента на стороне макета, обозначенного словом «Пельтье», для этого плюсовое гнездо левого канала источника питания соединить проводом с соответствующим плюсовым гнездом макета. Минусовое гнездо левого канала источника питания соединить с соответствующим минусовым выводом макета;

3) подключить цифровой вольтметр или мультиметр к соответствующим гнездам макета на стороне макета, обозначенного «mV»;

4) на источнике питания все ручки регулятора напряжения и тока повернуть против хода часовой стрелки до упора (на ноль). Отжать на источнике питания кнопки подключения напряжения на обоих каналах;

5) после завершения сборки схемы преподаватель проверит правильность подключения.

Категорически запрещается включать источник питания без разрешения преподавателя.

Порядок выполнения лабораторной работы при изучении эффекта Пельтье.

1. Включить источник питания. Перед включением убедиться, что обе ручки регулятора напряжения повернуты против часовой стрелки до упора и отжаты кнопки подачи напряжения на обоих каналах.

2. Установить на источнике питания напряжение 7 В на правом канале и ограничение тока 0,1 А и нажать кнопку подачи напряжения, при этом должны появиться цифры на цифровом дисплее макета. Верхний ряд дисплея макета показывает температуру горячей стороны элемента Пельтье (красный цвет). Нижний ряд показывает температуру холодной стороны элемента Пельтье (синий цвет).

3. Записать значения температур на дисплее макета при токе через элемент Пельтье, равном «0».

4. Установить на источнике питания на левом канале ограничение тока 1,2 А и нажать на кнопку подачи напряжения на левом канале источника питания.

5. Плавно поворачивая ручку установки напряжения, установить напряжение, при котором значение тока на дисплее источника питания составит 0,1 А (±0,01 А).

6. Наблюдать за измененим показаний цифрового термометра примерно 2...3 мин до стабилизации температуры. Когда температура стабилизируется, записать значение тока, температуры горячей и холодной сторон элемента Пельтье.

7. Повторить п. 5–6 для тока 0,2 А; 0,3 А; 0,4 А; 0,5 А; 0,6 А; 0,7 А (±0,01 А).

8. После завершения измерений повернуть ручку установки напряжения на левом канале источнике питания против хода часовой стрелки до упора и отсоединить провода левого канала источника питания от макета.

Порядок выполнения лабораторной работы при изучении эффекта Зеебека.

1. Установить температуру на обеих сторонах элемента Пельтье практически одинаковой (допустимы расхождения в две-три десятых доли градуса).

2. Подсоединить вольтметр или мультиметр к контактам лабораторного макета, обозначенным «ЭДС». Диапазон измерения установить =V.

3. Записать значения температуры горячей и холодной сторон элемента и значение ЭДС.

4. Включить фен, установить температуру 100 град. Допускается вместо фена использовать лампу накаливания.

5. Установить фен на расстоянии примерно 60 см от лабораторного макета и направить поток горячего воздуха на одну сторону элемента Пельтье.

6. Подождать примерно 2...3 мин до стабилизации температуры на горячей стороне, снять показания температуры и ЭДС и записать их значения.

7. Провести серию экспериментов при расстояниях, равных 50, 40, 30 и 20 см, при этом записать соответствующие значения температур и ЭДС.

Завершение работы

После завершения эксперимента выключить питание всех приборов, соблюдая следующий порядок:

1) убедиться, что обе ручки ругулировки напряжения и тока на блоке питания повернуты в крайнее левое положение (против хода часовой стрелки до упора);

2) выключить блок питания и все измерительные приборы;

3) выключить фен;

4) разобрать схему.

Обработка результатов измерений

После проведения измерений выполнить необходимые работы.

1. По результатам измерений рассчитать коэффициент дифференциальной термоЭДС, характеризующий пару полупроводников для каждой разности температур.

2. Построить график изменения разности температуры сторон термоэлемента в зависимости от тока, проходящего через термоэлемент.

3. Построить график изменения термоЭДС в зависимости от разности температуры сторон термоэлемента по данным эксперимента.

4. Рассчитать теоретическую зависимость изменения термоЭДС от разности температур сторон термоэлемента, построить график зависимости рядом с экспериментальной кривой. Объяснить, чем обусловлено несовпадение кривых.

5. В выводах объяснить, чем обусловлено расхождение значения коэффициента термоЭДС для каждого напряжения.

Содержание отчета

Отчет должен быть оформлен в виде журнала лабораторных работ (прил. 1). Содержание отчета:

1) цель и задачи работы;

2) ответы на контрольные вопросы;

3) принципиальная электрическая схема макета;

4) расчетные значения;

5) графики;

6) выводы.

Контрольные вопросы и задания

1. В чем суть явлений термоэлектричества?

2. Что называется контактной разностью потенциалов (КРП)?

3. Что такое коэффициент теплопроводности и в каких единицах он измеряется?

4. Что такое эффект Пельтье?

5. Почему происходит нагревание (охлаждение) спаев термопары при пропускании через них тока?

6. Что такое коэффициент Пельтье?

7. Где используется эффект Пельтье?

8. Почему в полупроводниках эффект Пельтье имеет большее значение, чем в металлах?

9. В чем заключается эффект Зеебека?

10. При каком условии появляется термоЭДС в цепи, состоящей из разнородных материалов?

11. Где можно использовать эффект Зеебека?

ЛАБОРАТОРНАЯ РАБОТА № 3.
ИССЛЕДОВАНИЕ ЗАВИСИМОСТИ БАРЬЕРНОЙ И ДИФФУЗИОННОЙ ЕМКОСТИ *p–n*-ПЕРЕХОДА ОТ ПРИЛОЖЕННОГО НАПРЯЖЕНИЯ

Цель лабораторной работы: изучение влияния обратного напряжения на изменение барьерной емкости *p–n*-перехода и прямого напряжения на диффузионную емкость *p–n*-перехода.

Задание по лабораторной работе.

1. Получить задание.
2. Проанализировать влияние обратного напряжения на изменение барьерной емкости *p–n*-перехода и прямого напряжения на диффузионную емкость *p–n*-перехода.
3. Обобщить результаты лабораторной работы, сформулировать краткие выводы.

Теоретическая часть

В результате образования контакта между полупроводниками различного типа получаю *p–n*-переходы. Типичная энергетическая диаграмма *p–n*-перехода в равновесном состоянии (без внешнего электрического поля) и распределение носителей заряда представлены на рис. 2.12. Поскольку электронные средства и ширина запрещенной зоны контактирующих материалов одинаковые, разрыв зон ΔE_C и ΔE_V в плоскости металлического контакта равен нулю; потенциальных барьеров здесь два, и они одинаковые по величине.

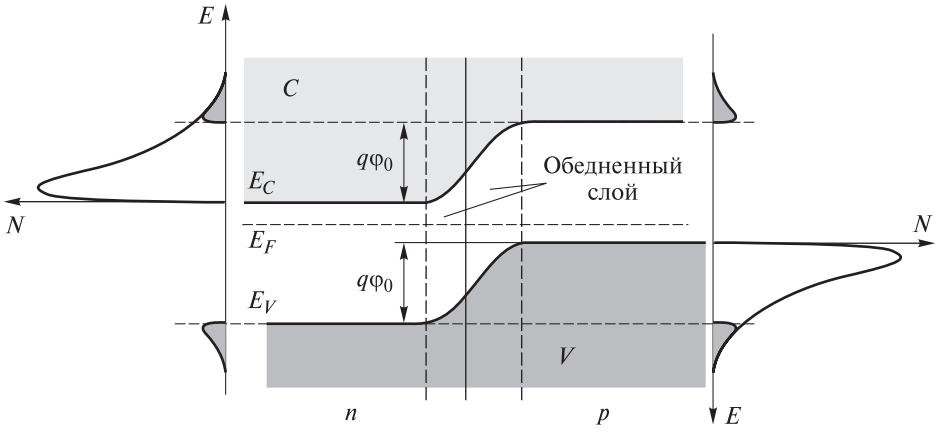


Рис. 2.12. Энергетическая диаграмма и концентрация носителей заряда в равновесном *p–n*-переходе (https://avatars.mds.yandex.net/i?id=2a00000179e2f2bd4aaba18678e04532dce1-4735414-images-thumbs&n=13)

Видно, что для основных носителей заряда (дырок для *p*-области и электронов для *n*-области) существует потенциальный барьер высотой $q\varphi_0$, для неосновных носителей потенциального барьера не существует, и они проходят через *p*–*n*-переход беспрепятственно.

Контактная разность потенциалов

$$q\varphi_0 = kT\ln\frac{N_D N_A}{n_i^2}.$$

Решая уравнение Пуассона, можно получить распределение напряженности электрического поля и потенциала в *p*–*n*-переходе (рис. 2.13):

$$E = \frac{q}{\varepsilon_s\varepsilon_0}\begin{cases} N_D\left(x+\delta_n\right), & x<0; \\ -N_A\left(x-\delta_p\right), & x>0. \end{cases}$$

$$\phi = \begin{cases} -\dfrac{q}{\varepsilon\varepsilon_0}N_D\left(\dfrac{x^2}{2}+\delta_n x\right)+\phi_n-\dfrac{q}{\varepsilon\varepsilon_0}N_D\dfrac{1}{2}\delta_n^2; \\ \dfrac{q}{\varepsilon\varepsilon_0}N_A\left(\dfrac{x^2}{2}-\delta_p x\right)+\phi_p+\dfrac{q}{\varepsilon\varepsilon_0}N_A\dfrac{1}{2}\delta_p^2. \end{cases}$$

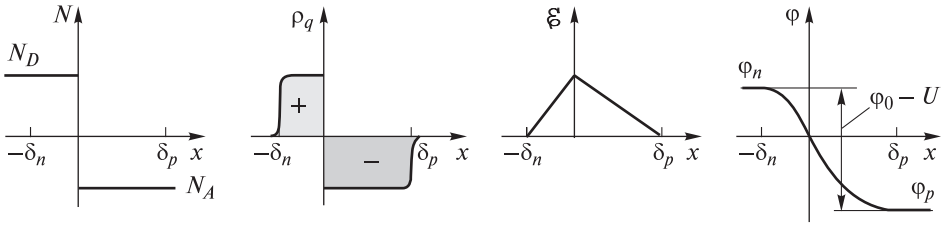


Рис. 2.13. Распределение соответственно примеси, заряда, напряженности электрического поля и потенциала в *n*-переходе (https://studfile.net/html/2706/40/html_Znp03Qk9nM.95H2/ img-wjHVhb.png)

Из уравнений для φ можно рассчитать общую толщину обедненной области:

$$\delta = \delta_n + \delta_p = \sqrt{\frac{2\varepsilon_s\varepsilon_0\left(N_D+N_A\right)}{qN_D N_A}\left(\phi_0 - U\right)}.$$

Для описания емкости пространственный заряд берется в одной из областей полупроводника (*n*- или *p*-):

$$Q_s = qN_D\delta_n S = S\sqrt{\frac{2\varepsilon_s\varepsilon_0 qN_D N_A}{\left(N_D+N_A\right)}\left(\phi_0 - U\right)},$$

тогда емкость составит

$$C = S\sqrt{\frac{\varepsilon_S \varepsilon_0 q N_D N_A}{2(N_D + N_A)(\phi_0 - U)}} = \frac{\varepsilon_S \varepsilon_0 S}{\delta}.$$

Обычно диоды на основе *p*–*n*-переходов изготавливают с несимметричным легированием. Если, например,

$$N_A << N_D,$$

выражение для емкости упрощается до вида

$$C = S\sqrt{\frac{\varepsilon_S \varepsilon_0 q N_A}{2(\phi_0 - U)}},$$

т. е. значение емкости определяется только примесью слаболегированной области.

Рассмотренная емкость получила название барьерной. Ее значение определяется неподвижными зарядами ионов легирующих примесей, без учета подвижных носителей заряда (электронов и дырок). Однако в *p*–*n*-переходе в отличие от барьера Шоттки происходит инжекция (впрыскивание) неосновных носителей заряда в соседние области. Это приводит к перераспределению основных носителей заряда, что происходит за счет процессов диффузии (рис. 2.14). Емкость, обусловленная этими зарядами, называют диффузионной.

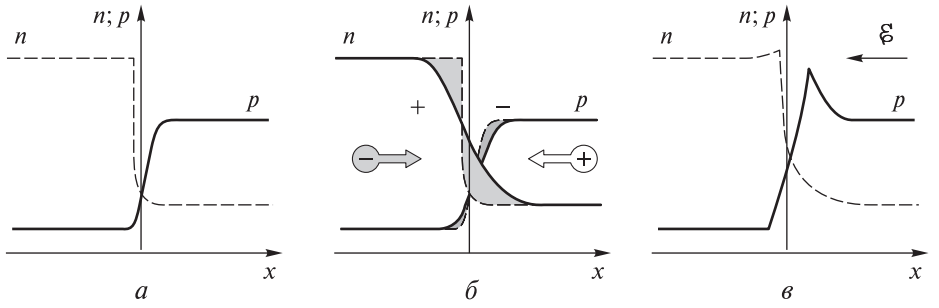


Рис. 2.14. Распределение носителей при изменении напряжения:

а, *в* — установившиеся состояния; *б* — нейтрализация заряда инжектированных носителей (https://ulsu.ru/media/documents/Gurin_Issled_poluprovod_struktur_metodom_volt-farad_kharakteristik_2016.pdf)

Несмотря на то что в установившемся состоянии примыкающие к *p*–*n*-переходу области не заряжены, диффузионную емкость можно связывать с зарядом инжектированных носителей, так как инжектированные носители и нейтрализующие их основные носители не исчезают.

Для сравнения учтем, что и обычный конденсатор в целом электрически нейтрален. Но в обычном конденсаторе положительный и отрицательный заряды пространственно разделены (то же самое можно сказать и о барьерной емкости *p*–*n*-перехода), в то время как при инжекции и положительный,

и отрицательный заряды оказываются в одной и той же области и пространственно не разделяются, в результате невозможно обнаружить область, где проходят токи смещения. В этом существенное отличие диффузионной емкости от барьерной емкости *p–n*-перехода и от емкости обычного конденсатора.

Основные понятия

Варикап — обратносмещенный полупроводниковый диод, предназначенный для применения в качестве элемента с электрически управляемой емкостью. У варикапов нормируют (и обеспечивают при производстве) емкость *p–n*-перехода при определенном напряжении смещения на нем и добротность. При увеличении обратного напряжения емкость варикапа уменьшается по закону

$$C_{\text{бар}} = \frac{C_{\text{бар}(0)}}{\left(1 + \dfrac{U_{\text{обр}}}{U_k}\right)^m},$$

где $C_{\text{бар}}$ — барьерная емкость *p–n*-перехода; $C_{\text{бар}(0)}$ — емкость *p–n*-перехода при нулевом обратном напряжении; U_k — контактная разность потенциалов; $U_{\text{обр}}$ — обратное напряжение на варикапе; m — коэффициент, зависящий от типа варикапа, $m = 0{,}3...0{,}5$.

К основным параметрам варикапа относят следующие.

Общая емкость варикапа $C_{\text{в}}$ — емкость, измеренная при определенном обратном напряжении.

Коэффициент перекрытия по емкости: $K_{\text{п}} = C_{\text{в max}}/C_{\text{в min}}$ — отношение емкостей варикапа при двух крайних значениях обратного напряжения ($K_{\text{п}} = 5...8$ раз).

Обратный ток $I_{\text{обр}}$ — постоянный ток, протекающий через варикап в обратном направлении при заданном обратном напряжении. На рис. 2.15 приведена ВФХ *p–n*-перехода варикапа.

При увеличении обратного напряжения барьерная емкость уменьшается. Диффузионная емкость варикапа увеличивается при возрастании прямого напряжения (см. рис. 2.15).

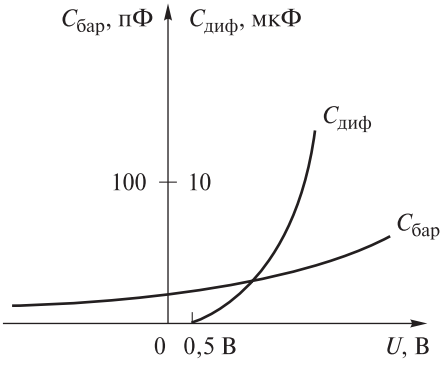


Рис. 2.15. Вольт-фарадные характеристики *p–n*-перехода варикапа

Экспериментальная часть

Принципиальная схема лабораторного макета для измерения ВФХ *p–n*-перехода приведена на рис. 2.16.

На варикап *VD*1 с источника питания через резистор *R*1 подается напряжение сначала в обратном направлении, т. е. на разъем *XT*1 подается «плюс», а на разъем *XT*2 «минус». При этом барьерная емкость варикапа уменьшается с увеличением обратного напряжения. Изменение барьерной емкости варикапа измеряют с помощью измерителя емкости,

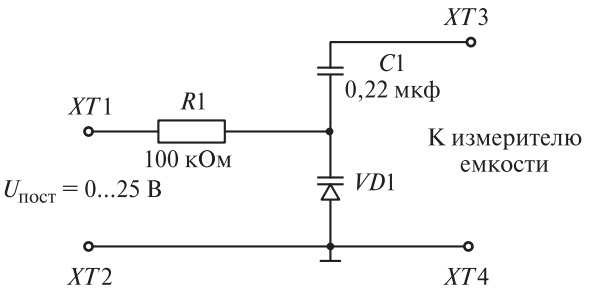


Рис. 2.16. Принципиальная схема лабораторного макета

например, мультиметра типа MastechMS8268 или его аналога. После получения экспериментальных данных при изменении обратного напряжения на варикапе от 0 до 25 В проводят измерения емкости при прямом напряжении. Для этого изменяют полярность подключения источника питания и проводят измерение при изменении напряжения в прямом направлении от 0 до 1 В.

Подготовить оборудование к выполнению лабораторной работы.

Объектом исследования является вариакап типа КВ103А или его аналог, установленный на макетной плате вместе с нагрузочным резистором *R*1 и разделительной емкостью большого номинала *C*1 (см. рис. 2.16). Разделительная емкость *C*1 служит для предотвращения поступления постоянного тока на ее измеритель. Поскольку емкость *C*1 подсоединена последовательно с варикапом и имеет большой номинал, ее влияние на процесс измерения будет незначительным.

Используемые приборы и оборудование:

– источник питания типа АТН-2031 или его аналог — 1 шт.;
– лабораторный макет — 1 шт.;
– мультиметр типа Mastech MS8268 или его аналог;
– соединительные провода.

Для выполнения лабораторной работы собирают стенд, состоящий из источника постоянного тока (см. рис. 2.11), лабораторного макета, мультиметра и соединительных проводов.

Перед выполнением лабораторной работы необходимо подготовить источник питания: включить его в сеть, убедиться, что отжаты все кнопки стабилизации напряжения и тока, ручки повернуты против хода часовой стрелки до упора.

Собрать электрическую схему в соответствии с принципиальной схемой, представленной на рис. 2.16:

1) подключить источник питания к лабораторному макету, для этого плюсовое гнездо блока питания соединить проводом с соответствующим выводом макета, обозначенным *XT*1 ($U_{обр}$ = 0...25 В). Минусовое гнездо источника постоянного тока соединить с соответствующим выводом макета, обозначенным как «Земля» ⊥ (*XT*2);

2) подключить мультиметр к лабораторному макету, для этого один провод мультметра подключить к выводу макета, обозначенному как «Земля» ⊥(*XT*2), а второй провод мультиметра подключить к выводу макета, обозначенному *XT*3 (U_c), и переключить мультиметр в режим измерения емкости. При этом необходимо убедиться, что провода вставлены в мультиметр в соответствующие гнезда (один провод в гнездо COM, второй — в гнездо, имеющее обозначение емкости).

После завершения сборки схемы преподаватель проверит правильность подключения.

Категорически запрещается включать источник питания без разрешения преподавателя.

Порядок выполнения лабораторной работы при измерении барьерной емкости.

1. Включить источник питания. Перед включением убедиться, что обе ручки регулятора напряжения повернуты против хода часовой стрелки до упора.

2. Нажать на источнике питания кнопку включения ограничения тока DCSET. Ручку регулировки тока повернуть по ходу часовой стрелки и установить ограничение тока 0,1 А. При этом убедиться, что напряжение на источнике питания равно нулю.

3. Записать в журнал значение емкости, показываемое мультиметром при нулевом обратном напряжении.

4. Нажать на источнике питания копку подключения напряжения к гнездам источника питания и установить напряжение 2 В.

5. Записать в журнал значение емкости, показываемое мультиметром при установленном обратном напряжении.

6. Далее увеличивать обратное напряжение на источнике питания в соответствии со значениями, указанными в табл. № 1 журнала лабораторных работ, до достижения 25 В и записывать значения емкости в журнал для каждого напряжения.

7. После измерения последнего значения отжать на источнике питания кнопку DCOUT и повернуть ручку регулировки напряжения против хода часовой стрелки до упора (установить 0).

Порядок выполнения лабораторной работы при измерении диффузионной емкости.

1. Поменять полярность напряжения, подаваемого на варикап путем переключения соединительных проводов. Убедиться, что напряжение на источнике питания равно нулю.

2. Нажать на источнике питания кнопку DCOUT.

3. Провести замеры изменения емкости варикапа при подаче прямого напряжения от 0 до 0,8 В с шагом 0,1 В. Для плавной регулировки напряжения использовать нижнюю ручку регулировки напряжения на источнике питания. Результаты измерения внести в журнал.

Во избежание выхода из строя варикапа категорически запрещается устанавливать напряжение более 0,8 В.

Завершение работы

После завершения эксперимента выключить питание всех приборов, соблюдая следующий порядок:

1) убедиться, что обе ручки ругулировки напряжения на блоке питания повернуты в крайнее левое положение (против хода часовой стрелки до упора);

2) выключить блок питания;

3) выключить мультиметр;

4) разобрать схему.

Обработка результатов измерений

После проведения измерений выполнить необходимые расчеты и построения в следующем порядке.

1. Рассчитать коэффициент перекрытия по емкости K_C.

2. Рассчитать теоретическое значение изменения барьерной емкости для всех экспериментальных значений обратного напряжения.

3. Построить ВФХ исследуемого варикапа для рассчитанных и экспериментальных значений (первая кривая — экспериментальная для барьерной емкости, вторая — теоретическая для барьерной емкости, третья — экспериментальная для диффузионной емкости).

4. В выводах объяснить возможные причины расхождения теоретических и экспериментальных кривых на графике.

Содержание отчета

Отчет должен должен быть оформлен в виде журнала лабораторных работ. Содержание отчета:

1) цель и задачи работы;

2) ответы на контрольные вопросы;

3) принципиальная электрическая схема макета;

4) расчетные значения;

5) график ВФХ;

6) выводы.

Контрольные вопросы и задания

1. Какие полупроводниковые диоды называются варикапами?

2. Объясните назначение варикапа, поясните принцип его работы.

3. Перечислите основные параметры варикапов.

4. Как изменится емкость варикапа при увеличении обратного напряжения?

5. Объясните физическую суть изменения емкости варикапа при увеличении обратного напряжения.

6. Что такое коэффициент перекрытия?

7. Можно ли использовать в качестве варикапа обычный полупроводниковый диод или, например, коллекторный (эмиттерный) переход биполярного транзистора?

8. Как экспериментально снять вольт-фарадную характеристику варикапа?

9. Как изменится емкость варикапа при увеличении прямого напряжения?

10. Как называется емкость, измеряемая при прямом напряжении на варикапе?

11. Как называется емкость, измеряемая при обратном напряжении на варикапе?

ЛАБОРАТОРНАЯ РАБОТА № 4.
ОПРЕДЕЛЕНИЕ КОНЦЕНТРАЦИИ И ПОДВИЖНОСТИ НОСИТЕЛЕЙ ТОКА В ПОЛУПРОВОДНИКЕ

Цель лабораторной работы: изучение эффекта Холла в полупроводниках, определение коэффициента Холла, концентрации и подвижности носителей тока.

Задание по лабораторной работе.

1. Получить задание.
2. Определить коэффициенты Холла, концентрации и подвижности носителей тока.
3. Проанализировать результаты лабораторной работы, сформулировать краткие выводы.

Теоретическая часть

В технике датчики Холла широко используются для измерения магнитных полей. Датчики Холла часто применяют для бесконтактного измерения токов в отдельных участках цепи, а также для измерения тока в пучках заряженных частиц. Существуют возможности создания на основе эффекта Холла магнитометров и усилителей.

Эффектом Холла называется возникновение в образце с током плотностью J, помещенном в магнитное поле напряженностью B, электрического поля напряженностью E в направлении, перпендикулярном B и J.

В опытах по изучению этого явления в образцах различных веществ измеряемой величиной является разность потенциалов Холла, однозначно связанная с электрическим полем.

Свойства полупроводников в основном определяются концентрацией носителей тока. Наиболее точный метод определения концентрации носителей заряда основан на эффекте Холла.

Пусть по пластине с площадью сечения S протекает электрический ток J в направлении, указанном на рис. 2.17 (направление движения электронов обратное). Разность потенциалов при отсутствии магнитного поля между электродами a и c, лежащими на одном из эквипотенциальных сечений пластины, равна нулю. Если создать магнитное поле B, перпендикулярное направлению тока, то между боковыми гранями (электродами a и c) возникает разность потенциалов Холла.

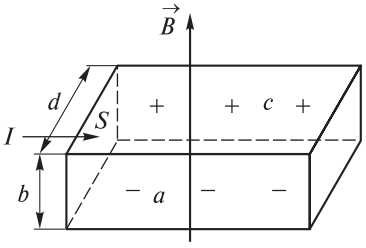


Рис. 2.17. Иллюстрация эффекта Холла

Экспериментально установлено, что разность потенциалов Холла U пропорциональна индукции магнитного поля B, силе тока I и обратно пропорциональна толщине пластины b:

$$U = R_X \frac{JB}{b}.$$

Множитель R_X, зависящий от вида вещества, называется коэффициентом Холла. Рассмотрим механизм появления разности потенциалов U. Как известно, на движущийся с дрейфовой скоростью V заряд e в магнитном поле действует сила Лоренца, перпендикулярная скорости движения и направлению поля: $F = q_e VB$.

Под действием силы Лоренца носители тока отклоняются к ближней грани пластины, заряжая ее отрицательно. На дальней грани при этом возникает равный по величине положительный заряд. Внутри пластины появляется поперечное электрическое поле ε, а между гранями — разность потенциалов U.

Отклонение носителей продолжается до тех нор, пока возбужденное ими поле ε не уравновесит отклоняющее действие магнитного поля, т. е. $\varepsilon q_e = q_e VB$.

Умножив обе части этого равенства на концентрацию носителей n, получим: $\varepsilon q_e n = q_e VBn = JB$, где J — плотность тока в образце.

Учитывая, что

$$j = \frac{J}{S} = \frac{J}{bd},$$

а $U = \varepsilon d$, имеем

$$U = \frac{1}{nq_e} \frac{JB}{b} = R_X \frac{JB}{b}.$$

Коэффициент Холла

$$R_X = \frac{1}{nq_e}.$$

Это выражение справедливо для металлов или вырожденных полупроводников, у которых носители тока подчиняются распределению Ферми — Дирака.

Учет статистического распределения скоростей носителей тока в невырожденных полупроводниках приводит к появлению добавочного множителя A, называемого холл-фактором. Значение A зависит от механизма рассеяния носителей тока. При рассеянии на ионизированных примесных атомах $A = 1{,}93$, при рассеянии на тепловых колебаниях решетки $A = 3\pi/8$.

В используемом датчике Холла из арсенида индия

$$R_X = \frac{3\pi}{8nq_e}.$$

Зная R_X, можно найти концентрацию носителей:

$$n = \frac{3\pi}{8R_X q_e}.$$

Одновременное измерение R_X и электропроводности σ позволяет вычислить подвижность носителей u. Действительно, электропроводность

$$\sigma = q_e n u.$$

Исключив концентрацию носителей, после преобразований получим

$$u = \frac{8R_X\sigma}{3\pi}.$$

Кроме возможности определения концентрации носителей, эффект Холла позволяет определить знак носителей и тип проводимости. Для электронных полупроводников $R_X < 0$, для дырочных $R_X > 0$.

При выполнении работы, как правило, не удается расположить электроды *a* и *c* (см. рис. 2.17) на эквипотенциальном сечении образца. Вследствие этого имеется некоторая разность потенциалов U_0 между электродами *a* и *c* при отсутствии магнитного поля *B*. Ее знак не зависит от направления *B*. Для исключения U_0 проводят два измерения U при противоположных направлениях *B*. Для этого изменяют направление тока через электромагнит и берут среднее значение U:

$$U' = U + U_0 \text{ (прямое поле)};$$

$$-U'' = -U + U_0 \text{ (обратное поле)}.$$

Вычитая из первого выражения второе, получаем, что холловская разность потенциалов

$$U = \frac{|U'| + |U''|}{2}.$$

Работа сводится к измерению значений U' и U'' для различных токов, пропускаемых через электромагнит, и вычислению среднего значения U. По значению U находят коэффициент Холла:

$$R_X = \frac{Ub}{JB}.$$

Концентрация носителей

$$n = \frac{3\pi}{8R_X q_e}.$$

По прилагаемым к конкретному лабораторному макету характеристикам датчика Холла (данные можно получить у преподавателя) находят электропроводность:

$$\sigma = \frac{l}{RS},$$

где l, R, S — соответственно длина, сопротивление и площадь поперечного сечения датчика.

Зная R_X и σ, по формуле

$$u = \frac{8R_X\sigma}{3\pi}$$

вычисляют подвижность носителей u.

Области приложения эффекта Холла в науке и технике достаточно разнообразны. В научной практике эффект Холла используют для определения концентрации и типа носителей, для оценки ширины запрещенной зоны, энергии ионизации носителей (по температурной зависимости R_X в полупроводниках). В комплексе с измерениями электропроводности эффект Холла позволяет вычислять ее температурную зависимость и подвижность носителей тока.

Экспериментальная часть

Объектом исследования является полупроводниковый датчик Холла, установленный на магнитопроводе электромагнита.

Используемые приборы и оборудование:

– источник постоянного тока типа АТН-2031 или его аналог — 1 шт.;
– лабораторный макет — 1 шт.;
– цифровой вольтметр или мультиметр — 1 шт.;
– соединительные провода.

Для выполнения лабораторной работы используют стенд (рис. 2.18), состоящий из источника постоянного тока ИП, электромагнита NS, цифрового вольтметра, миллиамперметра и соединительных проводов.

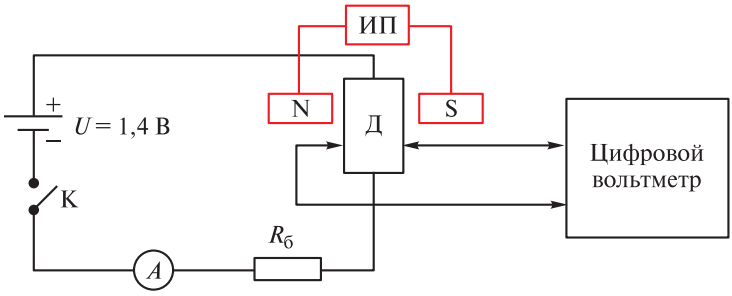


Рис. 2.18. Схема стенда для определения концентрации и подвижности носителей тока в полупроводнике

Перед выполнением лабораторной работы необходимо собрать электрическую схему в соответствии с рис. 2.18.

После завершения сборки схемы преподаватель проверяет правильность подключения.

Категорически запрещается включать питание источника постоянного тока без разрешения преподавателя.

Внимание!
Перед включением питания источника постоянного тока обе ручки регулятора напряжения (грубо и точно) повернуть против хода часовой стрелки до упора.

Порядок выполнения лабораторной работы

1. Включить питание источника постоянного тока. Перед включением убедиться, что обе ручки регулятора напряжения повернуты против хода часовой стрелки до упора.

2. Установить напряжение на источнике питания, питающего датчик Холла, 1,4 В, при этом через датчик Холла начнет протекать постоянный ток, значение которого необходимо рассчитать по закону Ома, учитывая, что сопротивление датчика Холла 354 Ом.

3. Установить напряжение на источнике питания, питающего электромагнит, 1 В.

4. Замерить напряжение U' и записать его значение в журнал.

5. Далее выполнить серию экспериментов по определению U', последовательно увеличивая напряжение, питающее электромагнит, с шагом 1 В до достижения последнего значения 10 В, записывая измеренные значения.

6. Поменять на блоке питания полярность подключения электромагнита путем перестановки проводов.

7. Замерить напряжение U'' при напряжении, питающем электромагнит, равном 10 В, записать измеренное значение.

8. Выполнить серию экспериментов по определению U'', последовательно уменьшая напряжение, питающее электромагнит, с шагом 1 В до достижения последнего значения 1 В, записывая измеренные значения.

Завершение работы

После завершения эксперимента выключить питание всех приборов, соблюдая следующий порядок:

1) убедиться, что обе ручки регулировки напряжения на источнике постоянного тока повернуты в крайнее левое положение (против хода часовой стрелки до упора);

2) выключить питание источника постоянного тока;

3) выключить питание цифрового вольтметра;

4) разобрать схему.

После проведения измерений выполнить обработку результатов измерений.

1. Заполнить в журнале таблицу с экспериментальными данными.

2. Построить график зависимости напряжения, снимаемого с датчика Холла, от тока электромагнита.

3. Рассчитать постоянную Холла, концентрацию и подвижность носителей тока в полупроводнике.

Содержание отчета

Отчет должен должен быть оформлен в виде журнала лабораторных работ. Содержание отчета:

1) цель и задачи работы;
2) ответы на контрольные вопросы;
3) схема стенда;
4) графики зависимости напряжения, снимаемого с датчика Холла, от тока электромагнита;
5) расчетные значения;
6) выводы.

Контрольные вопросы и задания

1. В чем заключается эффект Холла?
2. Какие типы проводимости имеются в полупроводниках?
3. Какие данные о проводниках и полупроводниках можно получить на основе исследования эффекта Холла?
4. Поясните схему включения приборов.
5. Почему на боковых гранях проводника с током в поперечном магнитном поле появляются электрические заряды?
6. Как исключить влияние разности потенциалов между электродами датчика при отсутствии магнитного поля?
7. Что называется подвижностью носителей тока? Какова ее размерность?
8. Как вычисляют концентрацию и подвижность носителей тока?
9. Приведите примеры применения эффекта Холла.
10. Эффект Холла характерен только для полупроводников, или он наблюдается также для металлов? Ответ обосновать.

3. НОРМАТИВНАЯ ДОКУМЕНТАЦИЯ

3.1. ПРИМЕРНАЯ БАЗОВАЯ ПРОГРАММА ДИСЦИПЛИНЫ



Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

Утверждаю
Первый проректор —
проректор по учебной работе
МГТУ им. Н.Э. Баумана
_____ Б.В. Падалкин
«___» _____ 20__ г.

Факультет «Информатика и системы управления»
Кафедра ИУ4 «Проектирование и технология производства
электронной аппаратуры»

Факультет «Информатика и управление»
Калужский филиал МГТУ им. Н.Э. Баумана
Кафедра ИУК1 «Проектирование и технология производства электронных приборов»

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Физические основы микро- и наноэлектроники»

для направления подготовки (уровень бакалавриата)
12.03.01 «Приборостроение»

Авторы программы:
Андреев В.В., проф., д-р техн. наук
Жалнин В.П., доцент, канд. техн. наук, эл. адрес: zhalninvp@bmstu.ru
Столяров А.А., проф., д-р техн. наук

Москва

Рабочая программа дисциплины определяет требования к знаниям и умениям студента, а также содержание и виды учебных занятий и отчетности.

Программа разработана в соответствии с основной профессиональной образовательной программой (ОПОП) и учебным планом МГТУ им. Н.Э. Баумана, составленными на основе самостоятельно устанавливаемого образовательного стандарта (СУОС) по направлению подготовки 12.03.01 «Приборостроение».

При освоении дисциплины планируется формирование компетенций, предусмотренных ОПОП на основе СУОС по направлению подготовки 12.03.01 «Приборостроение» (уровень бакалавриата).

Для категорий «знать, уметь, владеть» планируется достижение результатов обучения (РО), вносящих на соответствующих уровнях вклад в формирование компетенций, предусмотренных ОПОП.

Цели и задачи дисциплины

Цель дисциплины: изучение физических законов и явлений, на которых основано функционирование элементов современной микроэлектроники, освоение методов расчета параметров полупроводниковых структур, приобретение практических навыков исследования физических процессов, протекающих в полупроводниковых приборах.

Задачей дисциплины является изучение:

– основ квантовой механики и физической статистики;
– основ зонной теории твердых тел;
– положений статистической теории электронов и дырок в полупроводниках;
– положений теории электропроводности твердых тел;
– основ контактных и поверхностных явлений;
– положений теории туннельных эффектов и устройств на их основе.

Изучение дисциплины базируется на следующих курсах (разделах курсов):

– физика;
– математический анализ;
– материаловедение и материалы электронных средств;
– функциональная логика и теория алгоритмов;
– электротехника.

После освоения данной дисциплины студент подготовлен для изучения следующих курсов учебного плана:

– Основы аналого-цифровой схемотехники;
– Схемотехническое проектирование электронных средств;
– Конструкторское проектирование электронных средств;
– Технологические процессы микроэлектроники;
– Выполнения курсовых и дипломных проектов.

Знания, умения и навыки, получаемые после освоения дисциплины

Студент должен знать:

– основные понятия и термины, в том числе на английском языке, описывающие изделия, процессы и оборудование микроэлектроники;

– основы происходящих физических процессов в полупроводниковых и наноструктурах;

– принципы работы приборов на основе полупроводниковых структур;

– физические модели для расчета и анализа параметров полупроводниковых и наноструктур в целях выявления их оптимальных параметров.

Студент должен уметь:

– проводить измерения физических параметров полупроводниковых приборов;

– проводить анализ параметров статистическими методами, выявлять систематические и случайные погрешности;

– применять современные компьютерные программы для оформления текстовых и графических материалов лабораторных работ, домашних заданий, обзоров;

– применять соответствующее программное обеспечение для проведения расчетов параметров полупроводниковых и наноструктур;

– применять современное оборудование для определения и исследования различных параметров полупроводниковых и наноструктур;

– анализировать полученные результаты с применением современной вычислительной техники;

– анализировать причинно-следственные связи в рассмотренных процессах, обозначить и обосновать наиболее эффективные способы повышения микроминиатюризации.

Студент должен владеть:

– навыками работы в команде при выполнении лабораторных работ, групповых домашних заданий;

– приемами регулярной самостоятельной проработки и освоения модулей дисциплины, самоконтроля в достижении запланированных результатов обучения, поиска и освоения дополнительных источников информации по дисциплине при регулярном текущем контроле;

– приемами регулярной самостоятельной проработки и освоения модулей дисциплины, самоконтроля достижения запланированных результатов обучения, поиска и освоения дополнительных источников информации по дисциплине при регулярном текущем контроле;

– навыками применения условно-графических изображений приборов на основе полупроводниковых структур;

– навыками использования современного программного обеспечения для расчета параметров полупроводниковых структур.

Ниже приведены объем изучаемой дисциплины по семестрам (табл. 3.1), содержание дисциплины (табл. 3.2), часы, отведенные на изучение дисциплины по модулям (табл. 3.3), на самостоятельную работу (табл. 3.4) и на курсовую работу и курсовой проект (табл. 3.5).

Таблица 3.1

Объем дисциплины по видам учебных занятий

Вид учебной работы	Объем по семестрам, ч	
	Всего	1 семестр
	180	180
Аудиторная работа:	85	85
– лекции (Л)	51	51
– семинары (С)	17	17
– лабораторные работы (ЛР)	17	17
Самостоятельная работа обучающихся (СР):	95	95
– проработка учебного материала лекций	6,25	6,25
– подготовка к семинарам	2	2
– подготовка к лабораторным работам	6	6
– подготовка к экзамену	30	30
– подготовка к рубежному контролю	9	9
– другие виды самостоятельной работы	41,75	41,75
Вид промежуточной аттестации		Экзамен

Таблица 3.2

Содержание дисциплины (4-й семестр)

№	Тема модуля	Объем, ч			
		Л	С	ЛР	СР
1	Элементы квантовой механики и физической статистики. Элементы зонной теории твердых тел	18	6	6	23
2	Статистика электронов и дырок в полупроводниках. Электропроводность твердых тел	16	6	6	19
3	Контактные и поверхностные явления. Туннельные эффекты и устройства на их основе	17	5	5	23
4	Экзамен				
	ИТОГО:	51	17	17	95

Таблица 3.3

Модуль, содержание, продолжительность изучения

№	Наименование, содержание	Часы
1	Элементы квантовой механики и физической статистики. Элементы зонной теории твердых тел	
Лекции		18
Л1.1	Введение. Основные понятия и термины наноинженерии	2
Л1.2	Волновые свойства частиц, уравнение Шредингера. Соотношения неопределенности Гейзенберга, потенциальные барьеры для микрочастиц	4

Продолжение табл. 3.3

№	Наименование, содержание	Часы
Л1.3	Функция Ферми, фазовая и групповая скорости, фононы	4
Л1.4	Низкоразмерные объекты: квантовая яма, квантовая нить, квантовая точка. Электронный газ в периодическом потенциальном поле	2
Л1.5	Зоны Бриллюэна. Эффективная масса электрона. Зонная схема кристаллических тел — проводники, диэлектрики, полупроводники. Плотность состояний	6
Семинары		6
С1.1	Туннельный эффект на примере расчета некоторых параметров полупроводникового туннельного диода с использованием MATLAB	2
С1.2	Расчет собственной концентрации свободных носителей заряда в кремнии, германии, арсениде галлия и антимониде индия при различных температурах с использованием MATLAB	2
С1.3	Семинар-коллоквиум. Рубежный контроль	2
Лабораторные работы		4
ЛР1.1	Определение ширины запрещенной зоны полупроводника	4
Самостоятельная работа		23
СР1.1	Проработка учебного материала лекций	2,25
СР1.2	Подготовка к семинарам	0,75
СР1.3	Подготовка к лабораторным работам	2
СР1.4	Подготовка к рубежному контролю	3
СР1.5	Другие виды самостоятельной работы	15
2	Статистика электронов и дырок в полупроводниках. Электропроводность твердых тел	
Лекции		16
Л2.1	Собственные и примесные полупроводники. Зависимость концентрации свободных носителей в полупроводнике от положения уровня Ферми	4
Л2.2	Неравновесные носители, рекомбинация носителей. Уравнение непрерывности	4
Л2.3	Движение электронов под действием внешнего поля. Электропроводность чистых металлов	4
Л2.4	Электропроводность собственных полупроводников. Электропроводность примесных полупроводников. Диффузионные уравнения	4
Семинары		6
С2.1	Анализ объемного положения уровня Ферми относительно середины запрещенной зоны в собственных полупроводниках индия при различных температурах с использованием MATLAB	2
С2.2	Расчет собственного удельного сопротивления кремния, германия, арсенида галлия и антимонида индия при различных температурах с использованием MATLAB	2
С2.3	Семинар-коллоквиум. Рубежный контроль	2

Окончание табл. 3.3

№	Наименование, содержание	Часы
Лабораторные работы		4
ЛР2.1	Исследование термоэлектрических явлений в полупроводниках	4
Самостоятельная работа		19
СР2.1	Проработка учебного материала лекций	2
СР2.2	Подготовка к семинарам	0,75
СР2.3	Подготовка к лабораторным работам	2
СР2.4	Подготовка к рубежному контролю	3
СР2.5	Другие виды самостоятельной работы	11,25
3	Контактные и поверхностные явления. Туннельные эффекты и устройства на их основе	
Лекции		17
Л3.1	Контакт электронного и дырочного полупроводников. ВАХ тонкого *p*–*n*-перехода. Механизмы пробоя *p*–*n*-переходов	4
Л3.2	Эффект поля. МДП-структуры. Сильнополевая туннельная инжекция. Резонансно-туннельные диоды. Полупроводниковые сверхрешетки	4
Л3.3	Кулоновская блокада. Одноэлектронный транзистор. Устройства молекулярной электроники	2
Л3.4	Углеродные нанотрубки. Графен. Фуллерены и фуллериты	2
Л3.5	Микроэлектромеханические системы (МЭМС). Мемристоры	5
Семинары		5
С3.1	Расчет плотности тока электронов при отсутствии электрического поля в однородных слаболегированных полупроводниках Si, Ge, AsGa, InSb *n*-типа с использованием MATLAB	2
С3.2	Оценка высоты потенциального барьера в диоде Шоттки для различных металлов контакта при термодинамическом равновесии с использованием MATLAB	2
С3.3	Рубежный контроль	1
Лабораторные работы		9
ЛР3.1	Исследование зависимости барьерной и диффузионной емкости *p*–*n*-перехода от напряжения	4
ЛР3.2	Определение концентрации и подвижности носителей тока в полупроводнике	5
Самостоятельная работа		23
СР3.1	Проработка учебного материала лекций	2
СР3.2	Подготовка к семинарам	0,5
СР3.3	Подготовка к лабораторным работам	2
СР3.4	Подготовка к рубежному контролю	3
СР3.5	Другие виды самостоятельной работы	15,5
СР4.1	Подготовка к экзамену	30
4	Экзамен	30

Таблица 3.4

Самостоятельная работа (4-й семестр)

№	Тема самостоятельной работы	Объем, ч
1	Проработка учебного материала лекций	6,25
1	Подготовка к семинарам	2
2	Подготовка к лабораторным работам	6
2	Подготовка к экзамену	30
2	Подготовка к рубежному контролю	9
2	Другие виды самостоятельной работы	41,75

Самостоятельная проработка учебного материала лекций

Самостоятельная проработка курса лекций проводится по рекомендованной литературе и по материалам видеотеки (https://vk.com/video/playlist/-242540_1).

Самостоятельная подготовка к семинарам

Самостоятельная подготовка к семинарам проводится по рекомендованной литературе и по материалам видеотеки (https://vk.com/video/playlist/-242540_1).

Самостоятельная подготовка к лабораторным работам

Самостоятельная подготовка к лабораторным работам проводится по литературе и по материалам видеотеки (https://vk.com/video/playlist/-242540_1).

Самостоятельная подготовка к аттестациям

Самостоятельная подготовка к аттестациям проводится по рекомендованной литературе и по материалам видеотеки (https://vk.com/video/playlist/-242540_1).

Таблица 3.5

Курсовой проект, курсовая работа (4-й семестр)

№	Тема курсового проектирования, курсовой работы	Объем, ч	Литература

3.2. СТРУКТУРА И СОСТАВ ФОНДОВ ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ

Фонды оценочных средств по дисциплине представляют собой варианты экзаменационных билетов, перечень вопросов для рейтинговых и контрольных мероприятий.

Московский государственный технический университет им. Н.Э. Баумана
ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1 По курсу «Физические основы микро- и наноэлектроники» 1. Приближение сильной связи. 2. Контакт электронного и дырочного полупроводников. 3. Устройства спинэлектроники.
Утверждаю В.А. Шахнов
Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана
ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 2 По курсу «Физические основы микро- и наноэлектроники» 1. Приближение слабой связи. 2. *p*–*n*-переход в равновесном состоянии. 3. Углеродные нанотрубки.
Утверждаю В.А. Шахнов
Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 3

По курсу «Физические основы микро- и наноэлектроники»

1. Зоны Бриллюэна.
2. Зависимость контактной разности потенциалов от концетрации примеси в полупроводниках.
3. Классификация наноматериалов.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 4

По курсу «Физические основы микро- и наноэлектроники»

1. Эффективная масса.
2. *p*–*n*-переход при прямом смещении.
3. Причины резкого изменения свойств материалов при переходе к наноразмеру.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 5

По курсу «Физические основы микро- и наноэлектроники»

1. Зонная схема кристаллических тел, полупроводники, проводники, диэлектрики.
2. *p*–*n*-переход при обратном смещении.
3. Полупроводниковые сверхрешетки.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 6

По курсу «Физические основы микро- и наноэлектроники»

1. Зависимость концентрации свободных носителей в полупроводнике от положения уровня Ферми.
2. ВАХ тонкого *p*–*n*-перехода. Зависимость избыточной концентрации неосновных носителей от расстояния от *p*–*n*-перехода.
3. Нульмерный 0D-объект.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 7

По курсу «Физические основы микро- и наноэлектроники»

1. Положение уровня Ферми и концентрация носителей в собственных полупроводниках.
2. ВАХ тонкого *p*–*n*-перехода. Формула вольт-амперной характеристики для прямого и обратного тока диода.
3. Одномерный 1D-объект.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 8

По курсу «Физические основы микро- и наноэлектроники»

1. Положение уровня Ферми и концентрация носителей в примесных полупроводниках.
2. Тепловой ток и дифференциальное сопротивление *p*–*n*-перехода.
3. Двумерный 2D-объект.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 9

По курсу «Физические основы микро- и наноэлектроники»

1. Зависимость концентрации носителей в примесных полупроводниках от температуры. Закон действующих масс.
2. ВАХ *p*–*n*-перехода.
3. Соотношение неопределенности Гейзенберга.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 10

По курсу «Физические основы микро- и наноэлектроники»

1. Неравновесные носители. Рекомбинация носителей.
2. Поверхностные состояния.
3. Волновые свойства частиц. Уравнение де Бройля.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 11

По курсу «Физические основы микро- и наноэлектроники»

1. Рекомбинация носителей. Виды рекомбинации.
2. Пробой *p*–*n*-перехода.
3. Дрейфовые и диффузионные составляющие тока в полупроводниках.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 12

По курсу «Физические основы микро- и наноэлектроники»

1. Уравнение непрерывности.
2. Эффект поля. МДП-структуры.
3. Влияние сближения атомов на потенциальные барьеры.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 13

По курсу «Физические основы микро- и наноэлектроники»

1. Движение электронов под действием внешнего электрического поля.
2. Вольт-фарадные характеристики МДП-структур.
3. Электропроводность примесных полупроводников.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 14

По курсу «Физические основы микро- и наноэлектроники»

1. Зависимость подвижности носителей от температуры.
2. Полевые транзисторы. Общие сведения.
3. Отличия температурных зависимостей электропроводности металлов и полупроводников.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 15

По курсу «Физические основы микро- и наноэлектроники»

1. Основные механизмы рассеяния носителей. Подвижность носителей.
2. Полевые транзисторы с изолированным затвором.
3. Температурный метод определения ширины запрещенной зоны полупроводника.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 16

По курсу «Физические основы микро- и наноэлектроники»

1. Электропроводность чистых металлов.
2. Статические характеристики МДП-транзисторов.
3. Основные и неосновные носители в полупроводниках.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 17

По курсу «Физические основы микро- и наноэлектроники»

1. Электропроводность собственных полупроводников.
2. Полевые транзисторы с управляющим p–n-переходом.
3. Температурная зависимость концентрации носителей в примесных полупроводниках.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 18

По курсу «Физические основы микро- и наноэлектроники»

1. Электропроводность примесных полупроводников.
2. Перенос носителей через тонкие диэлектрические пленки. Эмиссия Шоттки, эмиссия Пула — Френкеля.
3. Концентрация носителей в примесных полупроводниках в зависимости от температуры.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 19

По курсу «Физические основы микро- и наноэлектроники»

1. Уравнение токов в полупроводниках.
2. Сильнополевая туннельная инжекция.
3. Положение уровня Ферми в примесных полупроводниках в зависимости от температуры.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 20

По курсу «Физические основы микро- и наноэлектроники»

1. Составляющие плотности тока в полупроводниках. Электрическое поле объемных зарядов.
2. Эффект поля. МДП-структуры.
3. Уровень Ферми и равновесная концентрация носителей в невырожденных собственных полупроводниках.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 21

По курсу «Физические основы микро- и наноэлектроники»

1. Кулоновская блокада.
2. Углеродные нанотрубки.
3. Закон действующих масс в полупроводниках.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 22

По курсу «Физические основы микро- и наноэлектроники»

1. Одноэлектронный транзистор.
2. Графен.
3. Зависимость положения уровня Ферми в собственном полупроводнике от температуры.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 23

По курсу «Физические основы микро- и наноэлектроники»

1. Устройства молекулярной электроники.
2. Фуллерены и фуллериты.
3. Собственные и примесные полупроводники.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 24

По курсу «Физические основы микро- и наноэлектроники»

1. Устройства спинэлектроники.
2. Основные виды наносенсоров и области их применения.
3. Положение уровня Ферми в собственном полупроводнике.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

Московский государственный технический университет им. Н.Э. Баумана

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 25

По курсу «Физические основы микро- и наноэлектроники»

1. Зонная схема кристаллических тел.
2. Необходимые условия для возникновения проводимости.
3. Понятие дырки в полупроводнике. Механизм образования дырки.

Утверждаю В.А. Шахнов

Билет рассмотрен и утвержден на заседании кафедры «__» _____ г.

3.3. СПЕЦИФИКАЦИЯ УЧЕБНЫХ ВИДЕО- И АУДИОМАТЕРИАЛОВ, СЛАЙДОВ, ЭСКИЗОВ, ПЛАКАТОВ И ДРУГИХ ДИДАКТИЧЕСКИХ МАТЕРИАЛОВ

В данном подразделе приведен типовой пример оформления слайдов к лекциям по курсу «Физические основы микроэлектроники». На первой лекции необходимо ознакомить студентов с целью и задачами курса, а также дать краткую информацию по структуре курса. Рекомендуемая структура для оформления схемы слайда приведена на рис. 3.1.

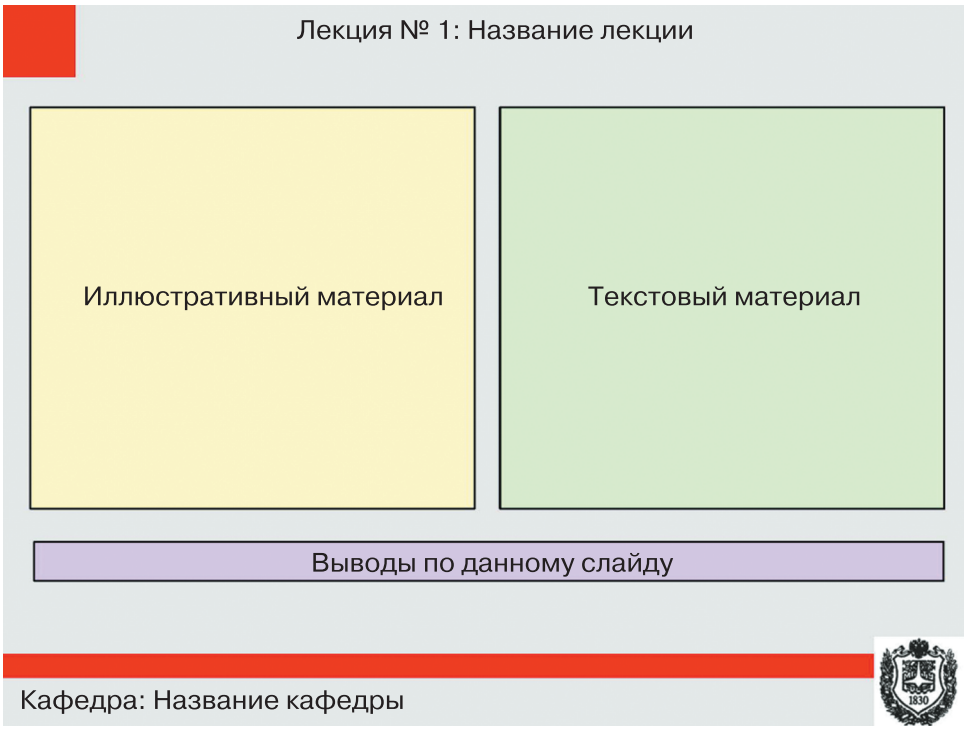


Рис. 3.1. Концептуальная схема слайда к лекции

В типовую структуру слайда рекомендуется включать название лекции, название кафедры, университета. Спецификация слайдов к лекциям приведена в табл. 3.6, примеры оформления — на рис. 3.2–3.21. С мультимедийными материалами к лекции можно ознакомиться на портале e-learning.bmstu.ru.

Таблица 3.6

Спецификация слайдов к лекциям

Название лекции	Число слайдов
1. Элементы квантовой механики и физической статистики	**29**
1.1. Волновые свойства частиц, уравнение Шредингера	5
1.2. Соотношение неопределенности Гейзенберга	3
1.3. Потенциальные барьеры для микрочастиц	7
1.4. Функция Ферми	4
1.5. Фазовая и групповая скорости, фононы	6
1.6. Низкоразмерные объекты: квантовая яма, квантовая нить, квантовая точка	4
2. Элементы зонной теории твердых тел	**38**
2.1. Электронный газ в периодическом потенциальном поле	9
2.2. Зоны Бриллюэна	7
2.3. Эффективная масса электрона	6
2.4. Зонная схема кристаллических тел – проводники, диэлектрики, полупроводники	7
2.5. Плотность состояний	9
3. Статистика электронов и дырок в полупроводниках	**37**
3.1. Собственные и примесные полупроводники	5
3.2. Зависимость концентрации свободных носителей в полупроводнике от положения уровня Ферми	8
3.3. Уровень Ферми и равновесная концентрация носителей в невырожденных собственных полупроводниках	4
3.4. Положение уровня Ферми и концентрация носителей в примесных полупроводниках	8
3.5. Неравновесные носители, рекомбинация носителей	3
3.6. Поверхностная рекомбинация	5
3.7. Уравнение непрерывности	4
4. Электропроводность твердых тел	**31**
4.1. Движение электронов под действием внешнего поля	5
4.2. Зависимость подвижности носителей заряда от температуры	6
4.3. Электропроводность чистых металлов	4
4.4. Электропроводность собственных полупроводников	3
4.5. Электропроводность примесных полупроводников	7
4.6. Диффузионные уравнения	6
5. Контактные явления	**24**
5.1. Контакт электронного и дырочного полупроводников	3
5.2. Равновесное состояние *p*–*n*-перехода	5
5.3. Зонная диаграмма *p*–*n*-перехода при приложении внешнего поля	4
5.4. ВАХ тонкого *p*–*n*-перехода	5
5.5. Барьерная и диффузионная емкость *p*–*n*-перехода	7

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОЭЛЕКТРОНИКИ

Доцент каф. ИУ4 Жалнин Владимир Петрович

Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

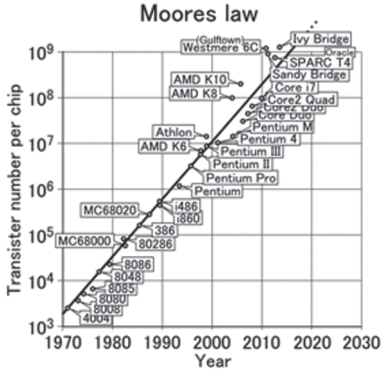
Рис. 3.2. Пример информационного слайда к вводной лекции

Вводная лекция

Лекция № 1

Исторические аспекты развития микроэлектроники

Развитие микроэлектроники сопровождалось ростом степени интеграции при одновременном уменьшении минимального размера элементов интегральных схем.





Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

Рис. 3.3. Примеры информационного слайда к лекции № 1

«Элементы квантовой механики и физической статистики»

Лекция № 3

Уравнение Шрёдингера

«Если электрон можно описать как волну, то как выглядит его волновое уравнение?» Питер Девай, 1926 г.



Нобелевский лауреат
Erwin Rudolf Josef Alexander Schrodinger (1887–1961)

Для микрочастицы, движущейся в силовом поле и обладающей потенциальной энергией $u(x, y, z, t)$, уравнение имеет вид

$$-i\hbar\frac{\partial\Psi}{\partial t}=\frac{\hbar^2}{2m}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}+\frac{\partial^2\Psi}{\partial z^2}\right)-u(x,y,z,t)\Psi,$$

где $i=\sqrt{-1}$; $\hbar$ – постоянная Планка, деленная на 2π (приведенная постоянная Планка или постоянная Дирака);

$\Psi(x, y, z, t)$ – является решением этого уравнения и называется волновой функцией.

Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

Рис. 3.4. Примеры информационного слайда к лекции № 3

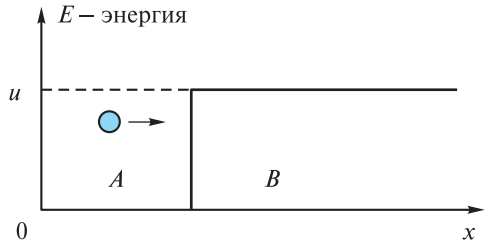
«Элементы квантовой механики и физической статистики»

Лекция № 4

Потенциальные барьеры для микрочастиц

Прямоугольный барьер полубесконечной толщины

Пусть в области A микрочастица имеет потенциальную энергию, равную нулю, а в области B – равную u



Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

Рис. 3.5. Примеры информационного слайда к лекции № 4

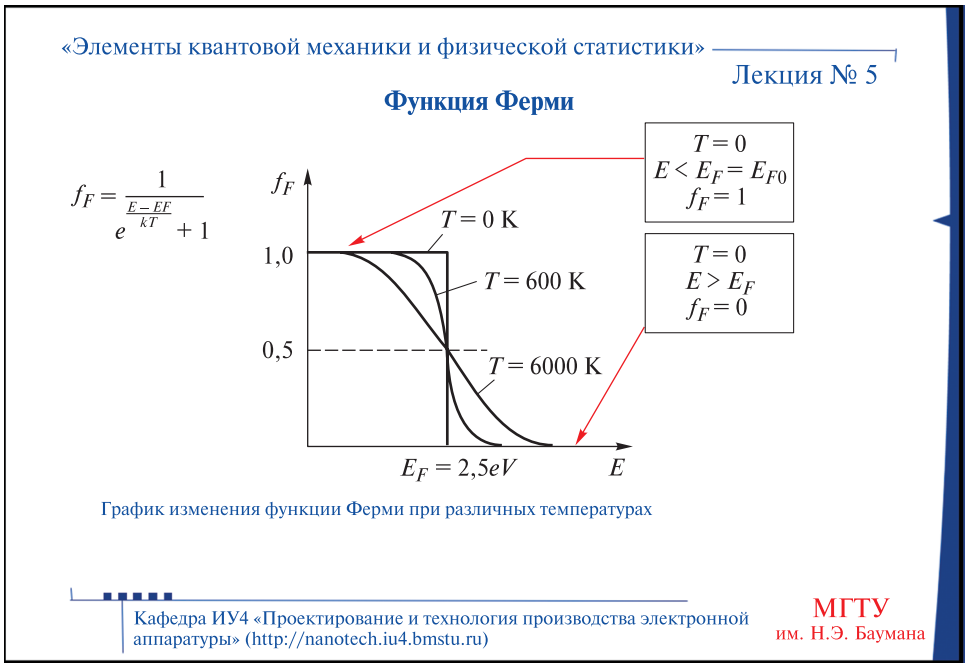


Рис. 3.6. Примеры информационного слайда к лекции № 5



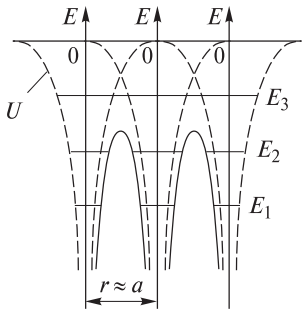
Рис. 3.7. Примеры информационного слайда к лекции № 6

«Элементы квантовой механики и физической статистики»

Лекция № 7

Элементы зонной теории твердых тел

Электронный газ в периодическом потенциальном поле



Потенциальные ямы атомов, находящихся на расстоянии, равном постоянной решетки

Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

а

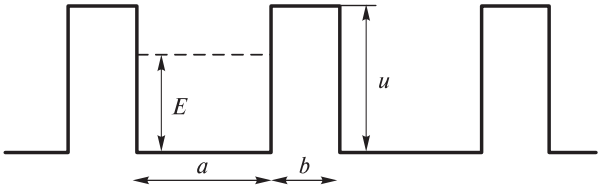
«Элементы квантовой механики и физической статистики»

Лекция № 7

Элементы зонной теории твердых тел

Приближение слабой связи

В рамках приближения слабой связи описывается состояние квазисвободиого электрона, движущегося в периодическом поле кристаллической решетки



Модель потенциального рельефа Кронига – Пенни

Уравнение Шрёдингера для циклических граничных условий:

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2m}{\hbar^2}(E-u)\Psi = 0$$

Кафедра ИУ4 «Проектирование и технология производства электронной аппаратуры» (http://nanotech.iu4.bmstu.ru)

МГТУ им. Н.Э. Баумана

б

Рис. 3.8. Примеры информационного слайда к лекции № 7

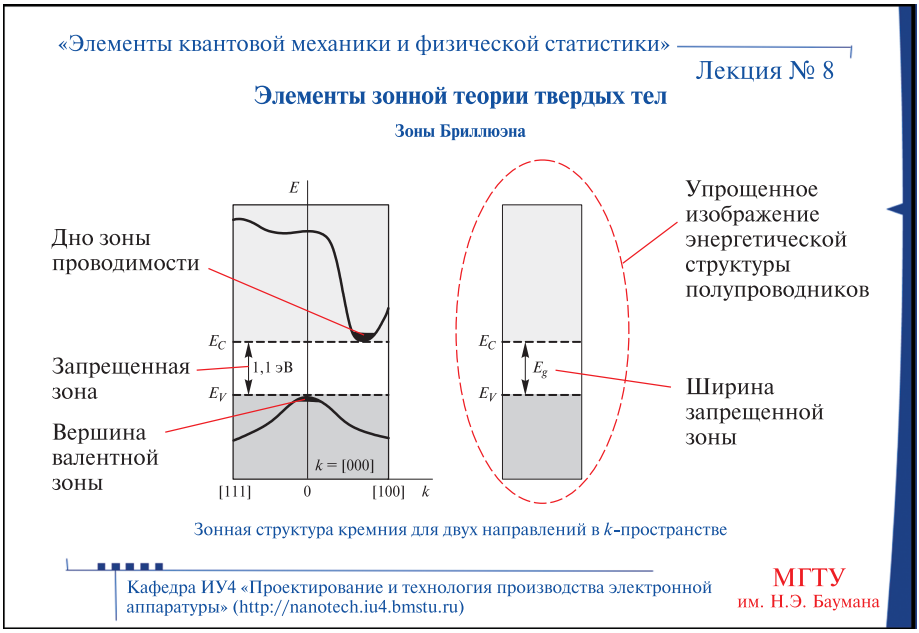


а



б

Рис. 3.9 (начало). Примеры информационного слайда к лекции № 8

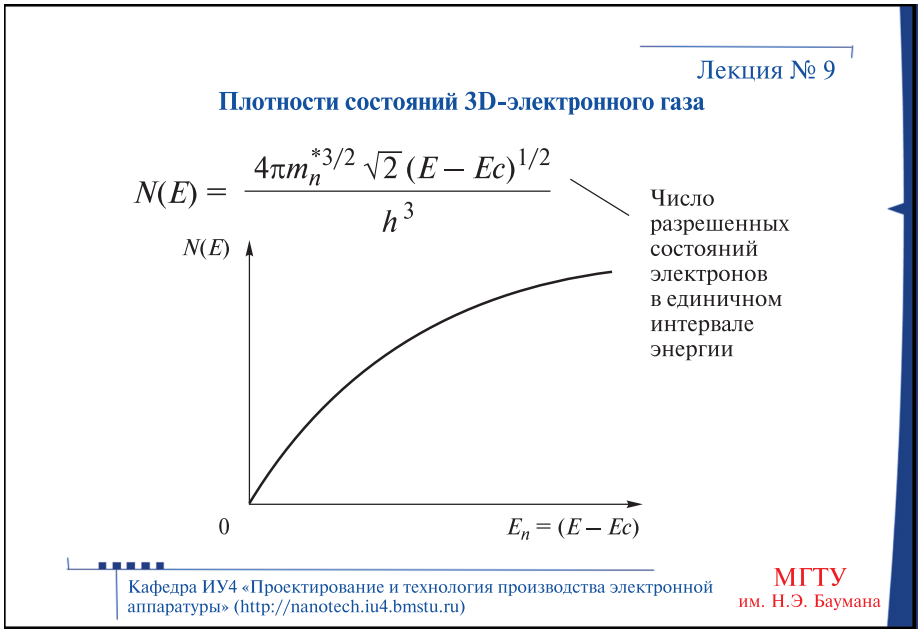


в

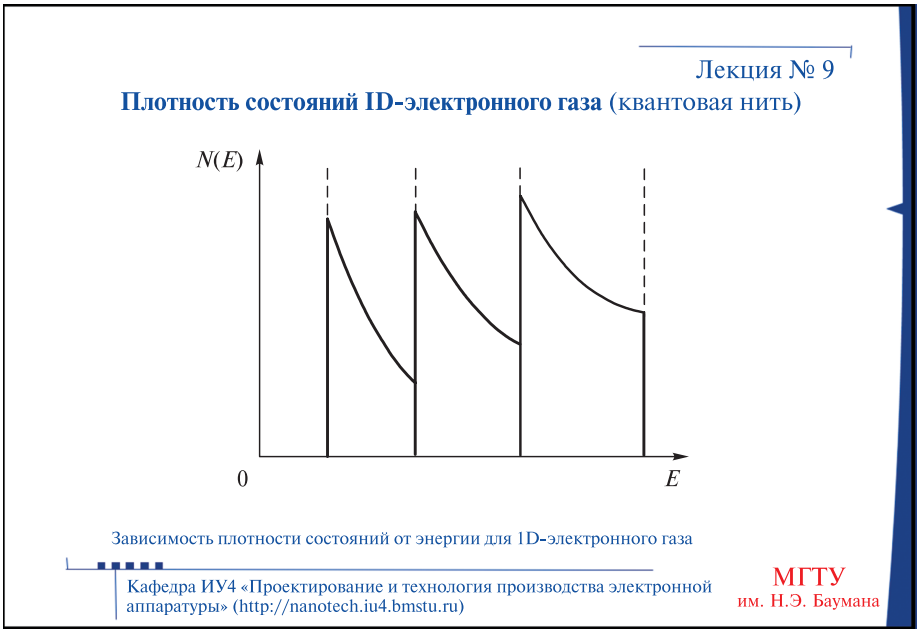


г

Рис. 3.9 (окончание). Примеры информационного слайда к лекции № 8



а



б

Рис. 3.10. Примеры информационного слайда к лекции № 9

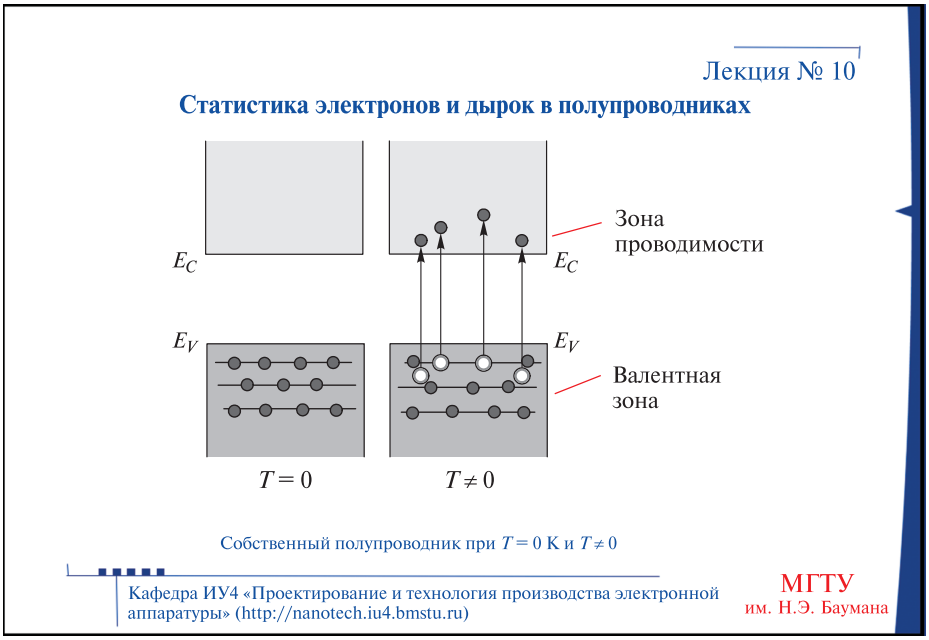


Рис. 3.11. Пример информационного слайда к лекции № 10

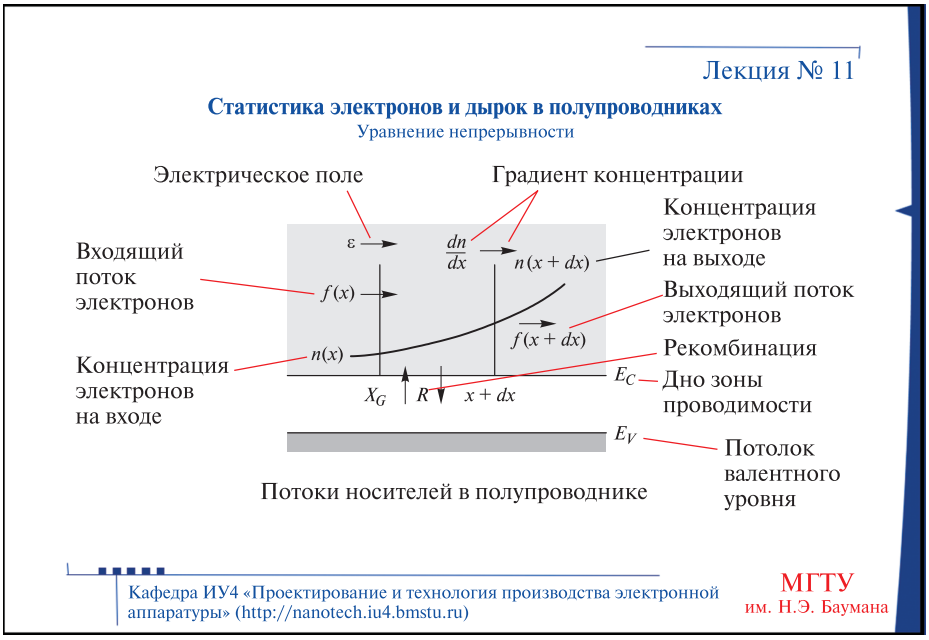
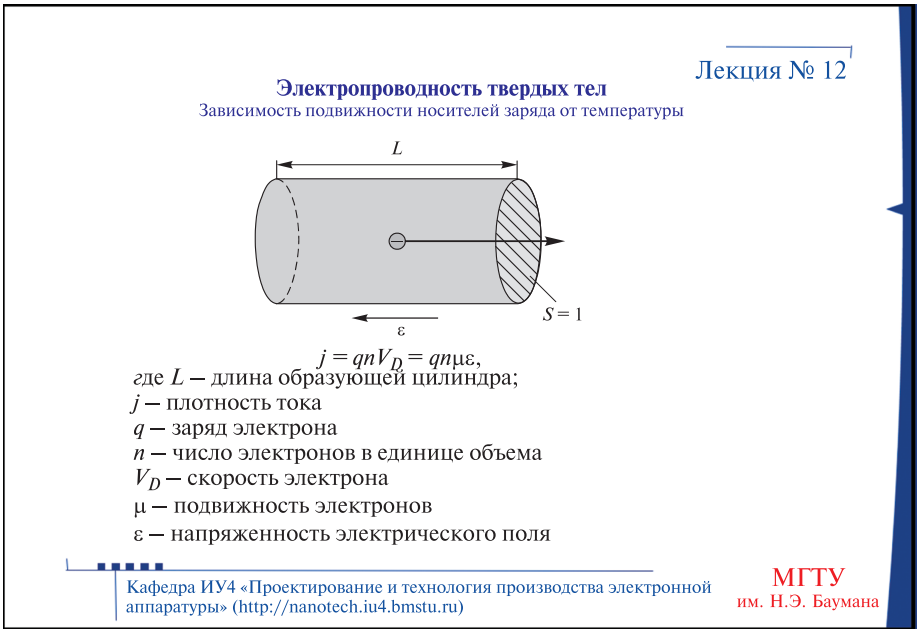
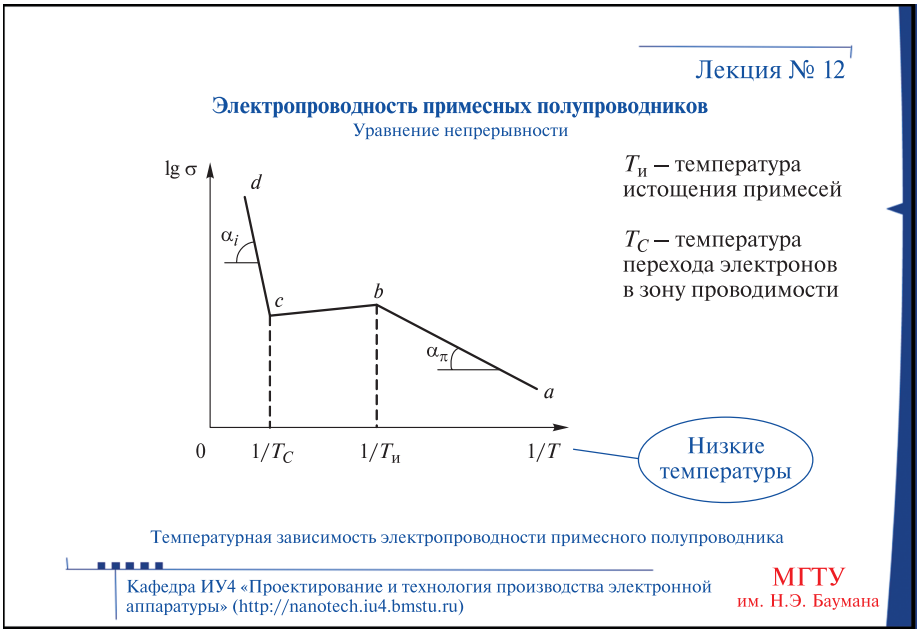


Рис. 3.12. Пример информационного слайда к лекции № 11

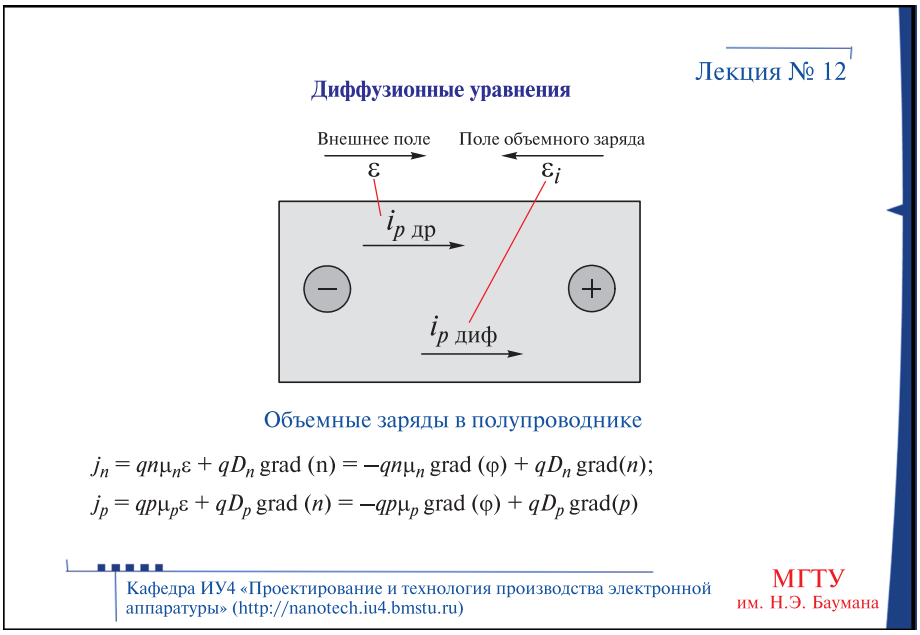


а



б

Рис. 3.13 (начало). Примеры информационного слайда к лекции № 12



в

Рис. 3.13 (окончание). Примеры информационного слайда к лекции № 12

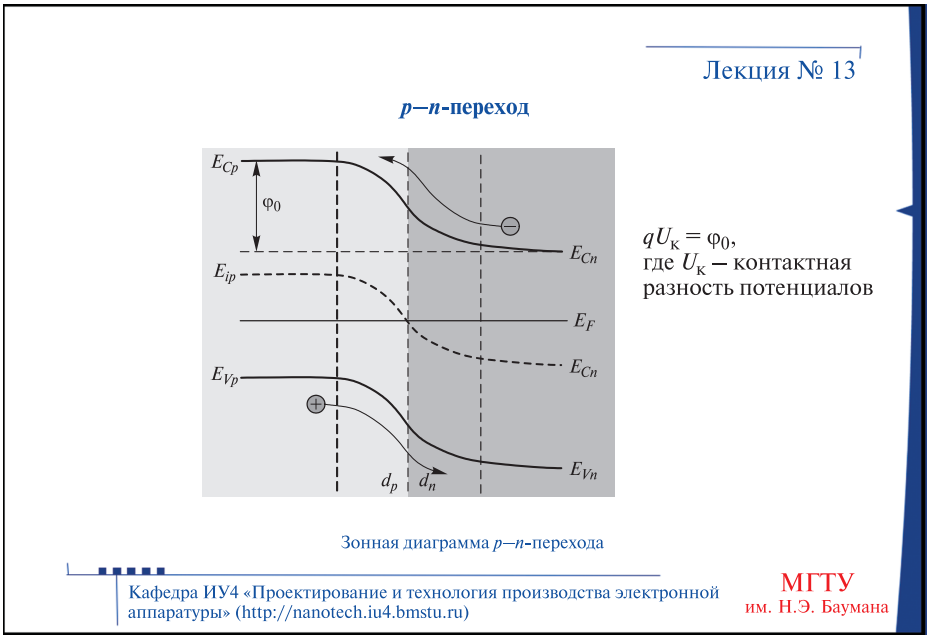
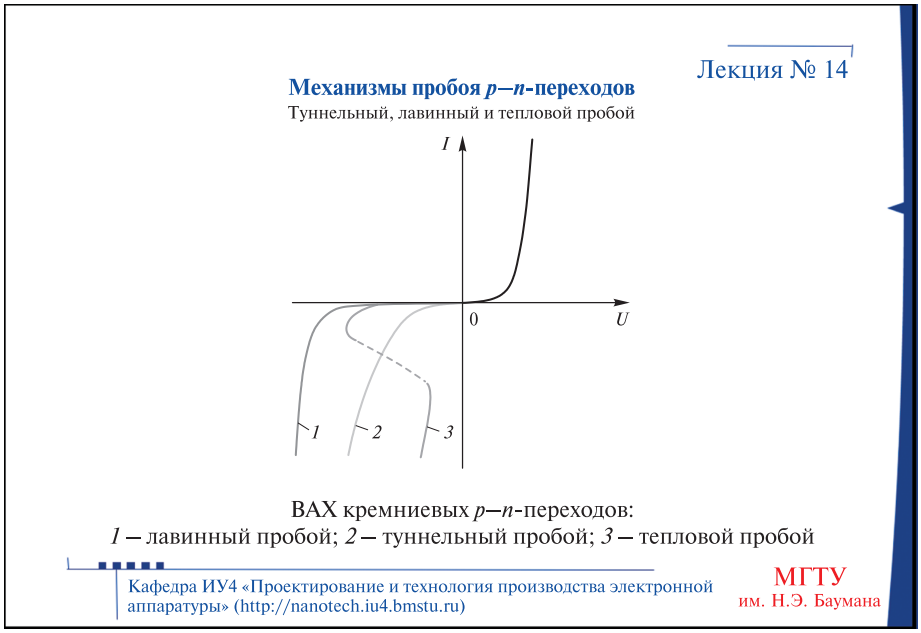
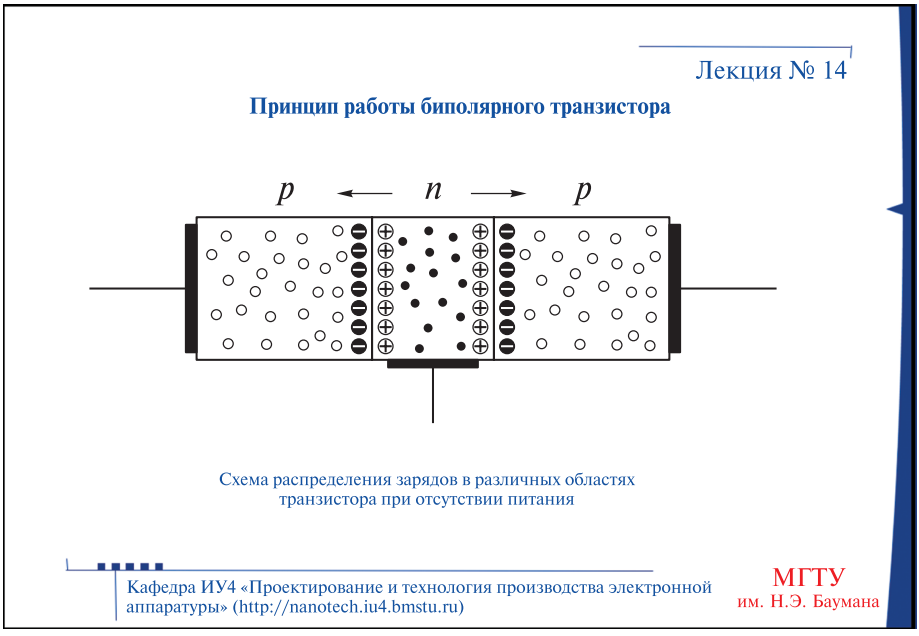


Рис. 3.14. Примеры информационного слайда к лекции № 13



а



б

Рис. 3.15. Примеры информационного слайда к лекции № 14

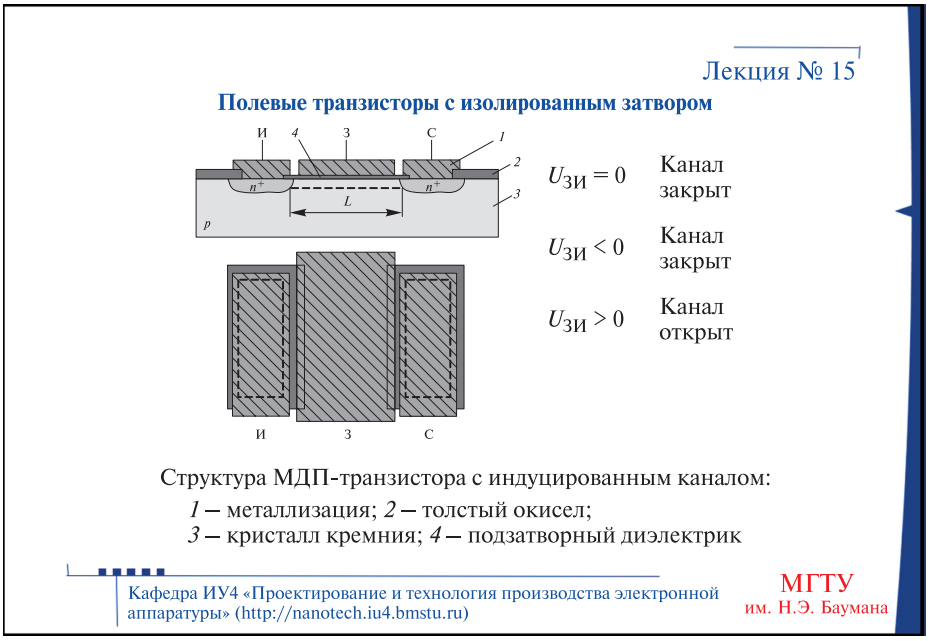


Рис. 3.16. Пример информационного слайда к лекции № 15

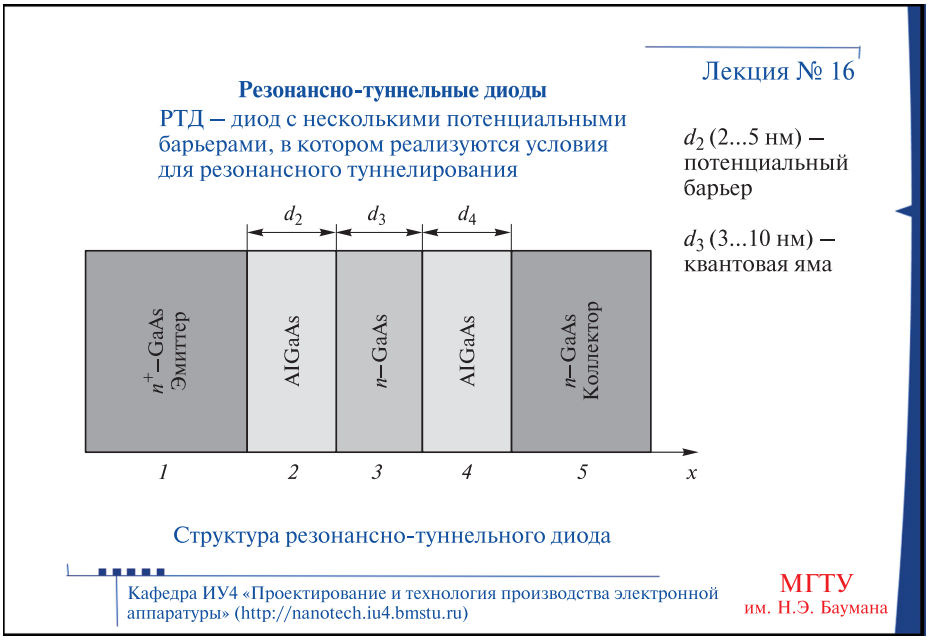
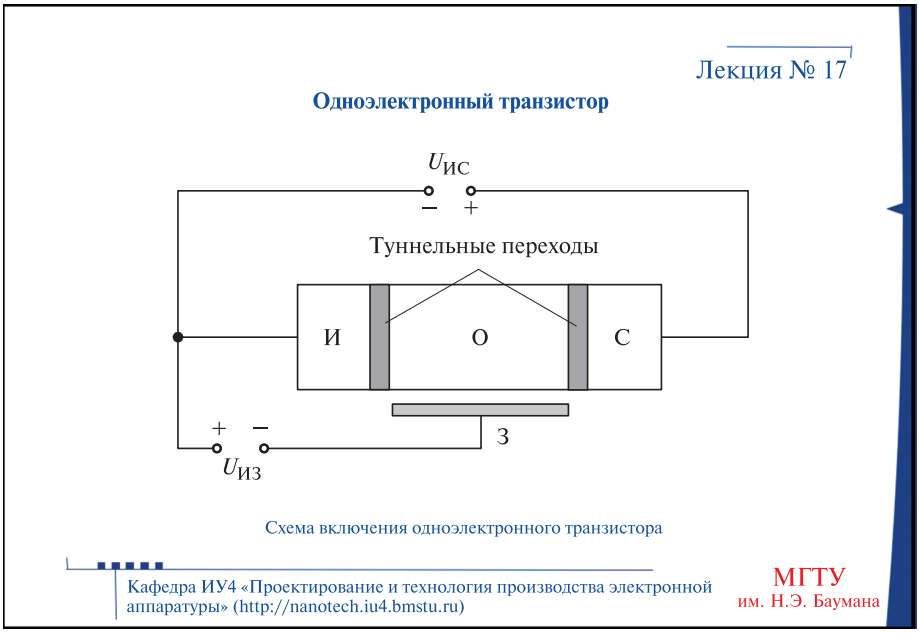
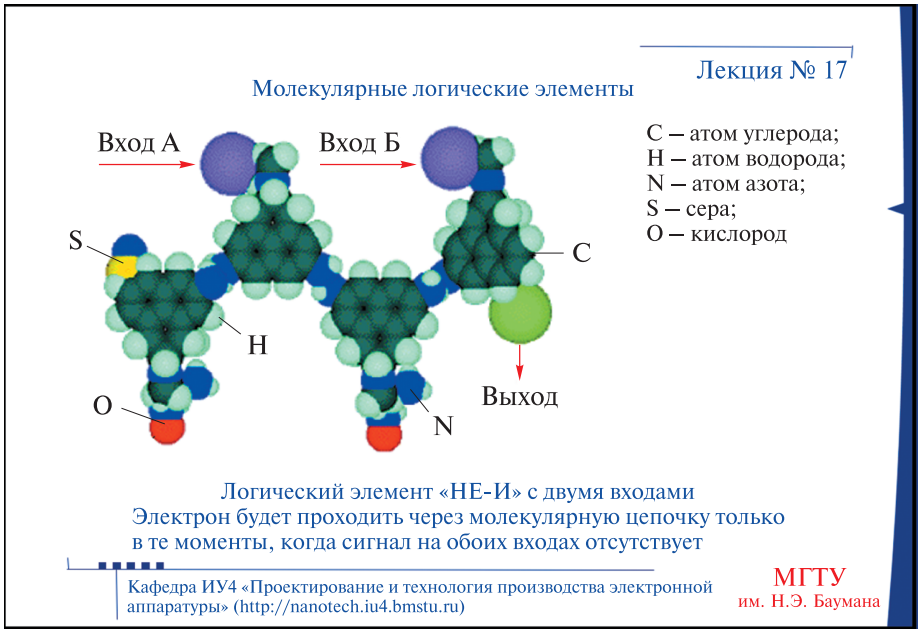


Рис. 3.17. Пример информационного слайда к лекции № 16

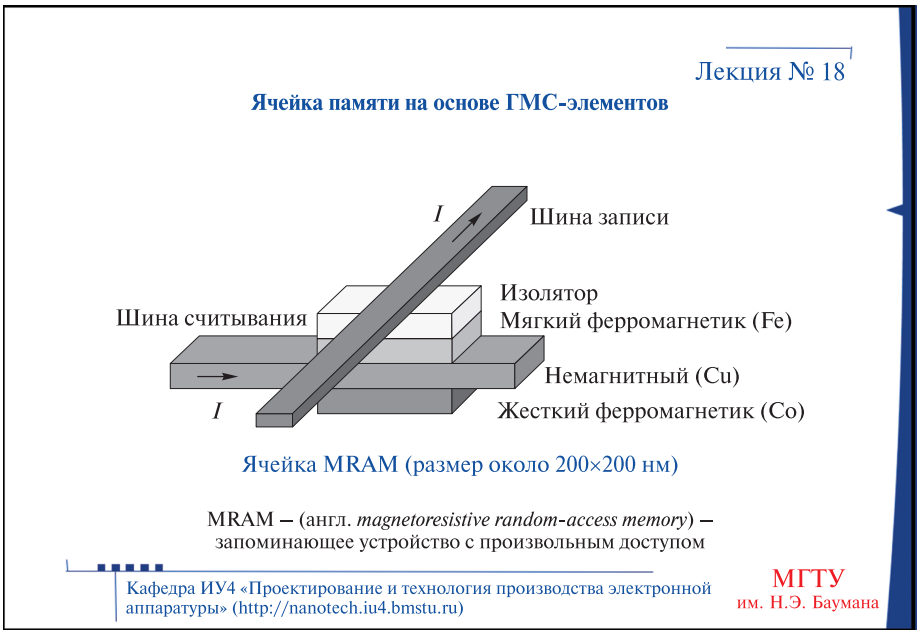


а

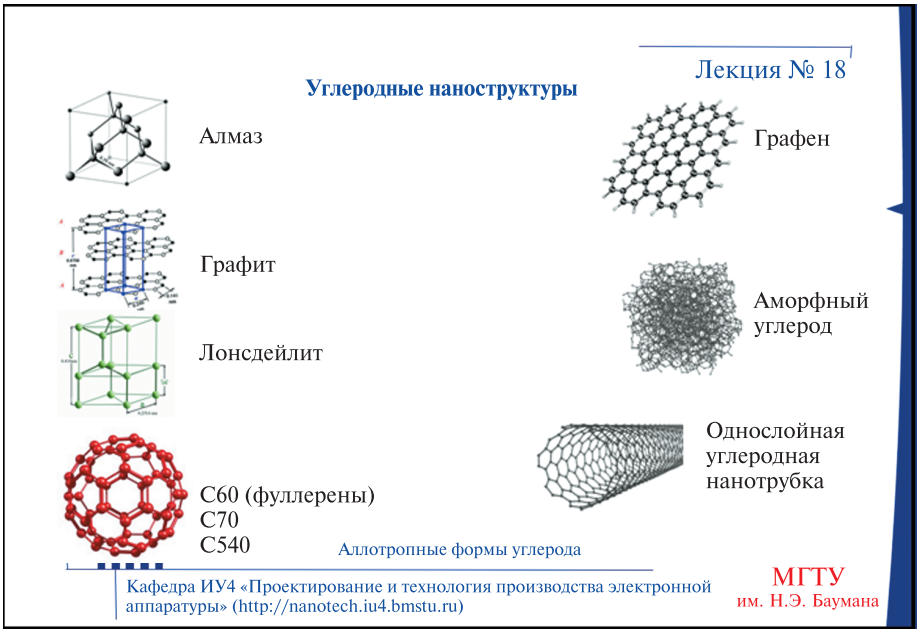


б

Рис. 3.18. Примеры информационного слайда к лекции № 17

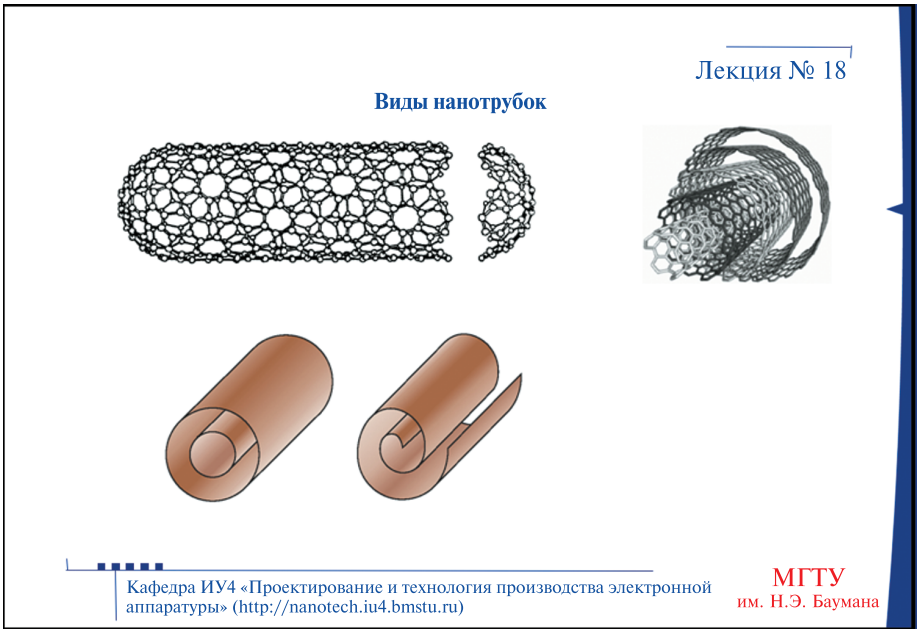


а

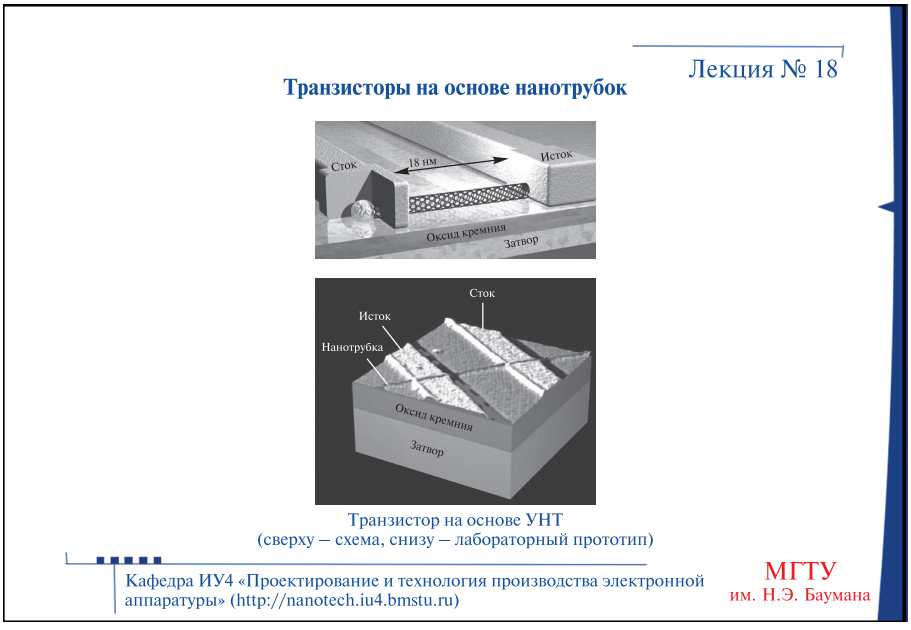


б

Рис. 3.19 (начало). Примеры информационного слайда к лекции № 18



в



г

Рис. 3.19 (окончание). Примеры информационного слайда к лекции № 18

Текст лекции, ее продолжительность, язык и насыщенность фактами и аргументами должны учитывать возрастной и социальный состава аудитории. Не рекомендуется писать весь материал лекции дословно. Бывает, что перед самым началом лекции приходится вносить срочные изменения и сделать их легче в конспекте, чем в дословном тексте. При подготовке текста особенное внимание рекомендуется уделять его началу и окончанию. Начало не должно быть слишком затянутым, не следует начинать с общих, всем известных фраз. Необходимо сразу заинтересовать слушателей, показать связь предмета лекции с реальными, известными из повседневной жизни событиями и объектами.

При выступлении необходимо стремиться избегать канцеляризмов, слов типа «значит», «так сказать» и т. д., жаргонных выражений. Живости, ясности выступления способствуют народные поговорки и пословицы, цитаты из литературных произведений и «крылатые слова» великих людей. Но и здесь необходимо знать меру. Считается, что при получасовом выступлении возможно использование двух-трех цитат и до пяти округленных цифр. Как в случае с подготовкой материалов в печать, текст выступления рекомендуется подготавливать за два-три дня в черновом варианте, а затем вновь вернуться к нему.

Психологи отмечают, что внимание аудитории к лектору постепенно начинает снижаться уже через 15 минут после начала лекции. В связи с этим не рекомендуется слишком затягивать свое выступление. Доказано также, что восприятие материала возрастает при комплексности воздействия на органы чувств человека. Поэтому желательно демонстрировать на лекции цветные слайды, видеофильмы, образцы, красочные плакаты и т. п.

Литература

Андреев В.В., Столяров А.А. Физические основы наноинженерии: учеб. пособие / под ред. В.А. Шахнова. М.: МГТУ им. Н.Э. Баумана, 2011.

Барыбин А.А. Электроника и микроэлектроника: физико-технологические основы. М.: Физматлит, 2006.

Варламов П.И., *Елсуков К.А.*, *Макарчук В.В.* Технологические процессы в наноинженерии. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Власов А.И., *Денисов А.А.*, *Елсуков К.А.* Бионаноинженерия: учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Власов А.И., *Елсуков К.А.*, *Косолапов И.А.* Оптическая микроскопия. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Власов А.И., *Елсуков К.А.*, *Панфилов Ю.В.* Методы микроскопии: учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Власов А.И., *Елсуков К.А.*, *Косолапов И.А.* Электронная микроскопия. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Власов А.И., *Назаров А.В.* Основы моделирования микро- и наносистем: учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Гуртов В.А. Твердотельная электроника: учеб. пособие для вузов. 3-е изд., доп. М.: Техносфера, 2008.

Епифанов Г.И., *Мома Ю.А.* Твердотельная электроника. М.: Высшая школа, 1986.

Ефимов И.Е., *Козырь И.Я.*, *Горбунов Ю.И.* Микроэлектроника. М.: Высшая школа, 1986.

Забродин Ю.С. Промышленная электроника. М.: Альянс, 2014.

Зи С. Физика полупроводниковых приборов. Т. 1. М.: Мир, 1984.

Игнатов А.Н., Фадеев Н.Е., *Савиных В.Л.* Классическая электроника и наноэлектроника. М.: Флинта, 2009.

Кальнов В.А., *Шахнов В.А.*, *Денисов А.А.* Проектирование наносенсоров. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Киселев В.Ф., *Козлов С.Н.*, *Зотеев А.В.* Основы физики поверхности твердого тела. М.: МГУ, 1999.

Кузовкин В.А. Электроника: учебник. М.: Логос, 2011.

Макарчук В.В., *Родионов И.А.*, *Цветков Ю.Б.* Методы литографии в наноинженерии. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011.

Орешкин П.Т. Физика полупроводников и диэлектриков. М.: Высшая школа, 1977.

Пасынков В.В., *Чиркин Л.К.* Полупроводниковые приборы. СПб.: Лань, 2001.

Петров К.С. Радиоматериалы, радиокомпоненты и электроника. СПб.: Питер, 2003.

Протасов Ю.С., *Чувашев С.Н.* Твердотельная электроника: учеб. пособие для вузов. М.: Изд-во МГТУ им. Н.Э. Баумана, 2003.

Проведение научных экспериментов в наноинженерии / А.А. Адамова, В.М. Башков, А.И. Власов, А.А. Жуков, Д.В. Козлов, Е.В. Кульгашев, Ю.М. Миронов, Е.В. Резчикова, В.А. Соловьев, В.А. Шахнов, В.М. Школьников. М.: Изд-во МГТУ им. Н.Э. Баумана, 2015.

Свистова Т.В. Микроэлектроника: учеб. пособие. Воронеж, Воронежский гос. технический ун-т, 2011.

Ситников А.В., Ситников И.А. Прикладная электроника: учебник. М.: Инфра-М, 2018.

Старосельский В.И. Физика полупроводниковых приборов микроэлектроники: учеб. пособие. М.: Юрайт, 2011.

Степаненко И.П. Основы микроэлектроники. М.: Лаборатория базовых знаний, 2004.

Столяров А.А. Физические основы микроэлектроники. М.: Изд-во МГТУ им. Н.Э. Баумана, 2006.

Троян П.Е. Микроэлектроника: учеб. пособие. Томск: ТУСУР, 2007.

Щука А.А., Сигов А.С. Электроника. В 4 ч. Ч. 2. Микроэлектроника: учебник для академического бакалавриата / А.С. Сигов, отв. ред. 2-е изд., испр. и доп. М.: Изд-во Юрайт, 2019.

Acín A. et al. The quantum technologies roadmap: A European community view // New J. Phys., 2018. Vol. 20. No. 8.

Arute F. et al. Quantum supremacy using a programmable superconducting processor // Nature, 2019. Vol. 574. No. 7779. Pp. 505–510.

Bernien H. et al. Probing many-body dynamics on a 51-atom quantum simulator // Nature. 2017. T. 551. No. 7682. Pp. 579–584.

Biamonte J., *Wittek P.*, *Pancotti N.*, *Rebentrost P.*, *Wiebe N.*, *Lloyd S.* Quantum machine learning // Nature, 2017. Vol. 549. No. 7671. Pp. 195–202.

Divincenzo D.P. Implementation of Quantum Computation. 2000. Pp. 1–9.

Fedorov G. P. et al. Photon transport in a Bose-Hubbard chain of super-conducting artificial atoms // Physical Review Letters. 2021. Vol. 126. No. 18. Pp. 180503.

Fu X., *Lao L.*, *Bertels K.*, *Almudever C.G.* A control microarchitecture for fault-tolerant quantum computing. 2019. Vol. 70. Pp. 21–30.

Gimeno-Segovia M., *Shadbolt P.*, *Browne D.E.*, *Rudolph T.* From Three-Photon Greenberger-Horne-Zeilinger States to Ballistic Universal Quantum Computation // Phys. Rev. Lett., 2015. Vol. 115. No. 2. Pp. 1–5.

Glattli D.C., *Roulleau P.S.* Levitons for electron quantum optics // Phys. Status Solidi Basic Res., 2017. Vol. 254. No. 3.

Göppl M. et al. Coplanar waveguide resonators for circuit quantum electrodynamics // J. Appl. Phys., 2008. Vol. 104. No. 11.

Groves T.R. Efficiency of electron-beam proximity effect correction // J. Vac. Sci. Technol. B Microelectron. Nanom. Struct., 1993. Vol. 11. No. 6. Pp. 2746.

Hendrickx N.W., *Franke D.P.*, *Sammak A.*, *Scappucci G.*, *Veldhorst M.* Fast two-qubit logic with holes in germanium // Nature, 2020. Vol. 577. No. 7791. Pp. 487–491.

Hensen B. et al. A silicon quantum-dot-coupled nuclear spin qubit // Nat. Nanotechnol., 2020. Vol. 15. No. 1. Pp. 13–17.

Humble T.S., *Thapliyal H.*, *Muñoz-Coreas E.*, *Mohiyaddin F.A.*, *Bennink R.S.* Quantum Computing Circuits and Devices // IEEE Des. Test, 2019. Vol. 36. No. 3. Pp. 69–94.

Lawrie W.I.L. et al. Quantum dot arrays in silicon and germanium // Appl. Phys. Lett., 2020. Vol. 116. No. 8.

Lucero E. et al. High-Fidelity Gates in a Single Josephson Qubit. vol. 247001, 2008. No. June. Pp. 1–4.

Mazhorin G.S. et al. Cavity-QED of a quantum metamaterial with tunable disorder // Phys. Rev. A 105. 033519. 2022. Pp. 1–12.

Moskalenko I.N., Simakov I.A., Abramov N.N. et al. High fidelity two-qubit gates on fluxoniums using a tunable coupler // Quantum Inf 8, 130. 2022. https://doi.org/10.1038/s41534-022-00644-x

Murali P., *Debroy D.M.*, *Brown K.R.*, *Martonosi M.* Architecting Noisy Intermediate-Scale Trapped Ion Quantum Computers // Proc. - Int. Symp. Comput. Archit., 2020. Vol. May. Pp. 529–542.

Nakamura Y., *Pashkin Y.A.*, *Tsai J.S.* Coherent control of macroscopic quantum states in a single-Cooper-pair box // Nature, 1999. Vol. 398. No. 6730. Pp. 786–788.

Niels A. et al. Understanding Quantum Technologies Fourth edition. No. September, 2021.

Orús R., *Mugel S.*, *Lizaso E.* Quantum computing for finance: Overview and prospects // Rev. Phys., 2019. Vol. 4. No. February.

Perdomo-Ortiz A., *Benedetti M.*, *Realpe-Gómez J.*, *Biswas R.* Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers // Quantum Sci. Technol., 2018. Vol. 3. No. 3.

Pino J.M. et al. Demonstration of the trapped-ion quantum CCD computer architecture // Nature, 2021. Vol. 592. No. 7853. Pp. 209–213.

Quantum Computing opens new front in Cloud! – AI4WIFI.URL:.https: // ai4wifi.com/quantum-computing-opens-new-front-in-cloud / (дата обращения 27.11.2022).

Rotta D., *Sebastiano F.*, *Charbon E.*, *Prati E.* Quantum information density scaling and qubit operation time constraints of CMOS silicon-based quantum computer architectures // npj Quantum Inf., 2017. Vol. 3. No. 1. Pp. 1–13.

SenZhong H. et al. Quantum computational advantage using photons // Science, 2020. Vol. 370. No. 6523. Pp. 1460–1463.

Shor P.W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM J. Comput., 1997. Vol. 26. No. 5. Pp. 1484–1509.

Strukov D.B., *Snider G.S.*, *Stewart D.R.*, *Williams R.S.* The missing memristor found // Nature Letters, 2008. Vol. 453. Pp. 80–83.

Strukov D.B., *Snider G.S.*, *Stewart D.R.*, *Williams R.S.* The missing memristor found // Nature letters, 2008. Vol. 453. Pp. 80–83.

Vandersypen L.M.K., *Chuang I.L.* NMR techniques for quantum control and computation 1069 // Rev. Mod. Phys., 2005. Vol. 76. No. October 2004. Pp. 1037–1069.

Veldhorst M., *Eenink H.G.J.*, *Yang C.H.*, *Dzurak A.S.* Silicon CMOS architecture for a spin-based quantum computer // Nat. Commun., 2017. Vol. 8. No. 1.

Wang J., *Sciarrino F.*, *Laing A.*, *Thompson M.G.* Integrated photonic quantum technologies // Nat. Photonics, 2020. Vol. 14. No. 5. Pp. 273–284.

Warring U., *Hakelberg F.*, *Kiefer P.*, *Wittemer M.*, *Schaetz T.* Trapped Ion Architecture for MultiDimensional Quantum Simulations // Adv. Quantum Technol., 2020. Vol. 3. No. 11. Pp. 1900137.

Wright K. et al. Benchmarking an 11-qubit quantum computer // Nat. Commun., 2019. Vol. 10. No. 1. Pp. 1–6.

Zhalnin V.P., *Shakhnov V.A.*, *Vlasov A.I.* Methods for improvement of the consistency and durability of the inorganic memristor structures // International Journal of Nanotechnology. 2019. Vol. 16. No. 1–3. Pp. 187–195.

Zhalnin V.P., *Vlasov A.I.*, *Rasuk A.A.*, *Kosyanov O.V.* Composite nanomaterials for implementation of promising memristive structures // Key Engineering Materials. 2022. Vol. 910 KEM. Pp. 703–712.

Zhalnin V.P., *Vlasov A.I.*, *Gudoshnikov I.V.*, *Kadyr A.T.*, *Shakhnov V.A.* Market for memristors and data mining memory structures for promising smart systems // Entrepreneurship and Sustainability Issues. 2020. Vol. 8. No. 2. Pp. 98–115.

Zhang X., *Li H.O.*, *Cao G.*, *Xiao M.*, *Guo G.C.*, *Guo G.P.* Semiconductor quantum computation // Natl. Sci. Rev., 2019. Vol. 6. No. 1. Pp. 32–54.

Содержание

Предисловие 5
Список сокращений 7
Термины и определения 7
1. КОНСПЕКТ ЛЕКЦИЙ 9
1.1. Введение в физические основы микро- и наноэлектроники 10
Тесты к лекции 1.1 12
1.2. Элементы квантовой механики и физической статистики 13
1.2.1. Волновые свойства частиц. Уравнение Шредингера 13
1.2.2. Теорема Феликса Блоха 16
1.2.3. Соотношения неопределенности Гейзенберга 17
Тесты к лекции 1.2 18
1.3. Потенциальные барьеры полубесконечной толщины 19
Тесты к лекции 1.3 22
1.4. Потенциальные барьеры конечной толщины. Функция Ферми 23
1.4.1. Прямоугольный потенциальный барьер 23
1.4.2. Функция Ферми 24
1.4.3. Строение атома 25
Тесты к лекции 1.4 25
1.5. Волновые процессы в твердом теле 27
1.5.1. Фазовая и групповая скорости, фононы 27
1.5.2. Низкоразмерные объекты: квантовая яма, квантовая нить, квантовая точка 28
Тесты к лекции 1.5 30
1.6. Элементы зонной теории твердых тел 31
Тесты к лекции 1.6 35
1.7. Зоны Бриллюэна и понятие эффективной массы 36
1.7.1. Зоны Бриллюэна 36
1.7.2. Эффективная масса электрона 38
Тесты к лекции 1.7 40
1.8. Зонная схема кристаллических тел и плотность состояний 42
1.8.1. Зонная схема кристаллических тел — проводники, диэлектрики, полупроводники 42
1.8.2. Плотность состояний 43
Тесты к лекции 1.8 48

1.9. Статистика электронов и дырок в полупроводниках 49
1.9.1. Собственные и примесные полупроводники 49
1.9.2. Зависимость концентрации свободных носителей в полупроводнике от положения уровня Ферми 51
1.9.3. Уровень Ферми и равновесная концентрация носителей в невырожденных собственных полупроводниках 53
1.9.4. Положение уровня Ферми и концентрация носителей в примесных полупроводниках 55
Тесты к лекции 1.9 58
1.10. Неравновесные носители и уравнение непрерывности 60
1.10.1. Неравновесные носители, рекомбинация носителей 60
1.10.2. Поверхностная рекомбинация 61
1.10.3. Уравнение непрерывности 62
Тесты к лекции 1.10 64
1.11. Электропроводность твердых тел 65
1.11.1. Движение электронов под действием внешнего электрического поля 65
1.11.2. Зависимость подвижности носителей заряда от температуры ... 67
1.11.3. Электропроводность чистых металлов 67
1.11.4. Электропроводность собственных полупроводников 68
1.11.5. Электропроводность примесных полупроводников 70
1.11.6. Диффузионные уравнения 71
Тесты к лекции 1.11 72
1.12. Контактные явления 74
1.12.1. Контакт электронного и дырочного полупроводников 74
1.12.2. Равновесное состояние *p*–*n*-перехода 76
1.12.3. Зонная диаграмма *p*–*n*-перехода при приложении внешнего электрического поля 78
1.12.4. Вольт-амперная характеристика тонкого *p*–*n*-перехода 80
Тесты к лекции 1.12 84
1.13. Емкость *p*–*n*-перехода 86
1.13.1. Барьерная и диффузионная емкость *p*–*n*-перехода 86
1.13.2. Механизмы пробоя *p*–*n*-переходов 87
Тесты к лекции 1.13 89
1.14. Биполярные транзисторы 90
1.14.1. Общие сведения 90
1.14.2. Принцип действия биполярного транзистора 91
1.14.3. Схемы включения биполярных транзисторов 93
1.14.4. Гетеропереходы 96
Тесты к лекции 1.14 98
1.15. Поверхностные явления 100
1.15.1. Основные понятия о поверхностных состояниях 100
1.15.2. Эффект поля. Металл–диэлектрик–полупроводник-структуры 100
1.15.3. Вольт-фарадная характеристика 102

1.15.4. Полевые транзисторы 104
1.15.5. Полевые транзисторы с управляющим *p*–*n*-переходом 109
1.15.6. Механизмы переноса заряда через тонкие диэлектрические пленки 110
Тесты к лекции 1.15 112
1.16. Туннельные эффекты и полупроводниковые сверхрешетки 114
1.16.1. Основные понятия о туннельных эффектах 114
1.16.2. Сильнополевая туннельная инжекция 115
1.16.3. Туннельный эффект в *p*–*n*-переходе, образованном вырожденными полупроводниками 116
1.16.4. Резонансно-туннельные диоды 118
1.16.5. Полупроводниковые сверхрешетки 120
Тесты к лекции 1.16 122
1.17. Одноэлектронные устройства, устройства молекулярной и спин-электроники 124
1.17.1. Кулоновская блокада 124
1.17.2. Одноэлектронный транзистор 127
1.17.3. Устройства молекулярной электроники 129
1.17.4. Самосборка и самоорганизация 133
1.17.5. Устройства спин-электроники 134
Тесты к лекции 1.17 139
1.18. Физические основы низкоразмерных структур 141
Тесты к лекции 1.18 143
1.19. Углеродные наноструктуры 145
1.19.1. Общие сведения об углеродных нанотрубках 145
1.19.2. Углеродные нанотрубки 145
1.19.3. Графен 149
1.19.4. Фуллерены и фуллериты 150
Тесты к лекции 1.19 152
1.20. Термоэлектрические явления в полупроводниковых структурах 153
1.20.1. Эффект Зеебека 153
1.20.2. Эффект Пелетье 154
1.20.3. Эффект Томсона 155
Тесты к лекции 1.20 156
1.21. Высокотемпературная полупроводниковая электроника 158
1.21.1. Общие сведения о высокотемпературных полупроводниках 158
1.21.2. Твердотельные приборы на карбиде кремния 160
Тесты к лекции 1.21 163
1.22. Фоточувствительные элементы полупроводниковой электроники 165
1.22.1. Фотоприемники 165
1.22.2. Фоторезисторы 167
1.22.3. Фотодиоды 167
1.22.4. Лавинный фотодиод 168

1.22.5. Фототранзистор ... 169
1.22.6. Оптроны ... 169
1.22.7. Приборы с зарядовой связью ... 170
Тесты к лекции 1.22 ... 172
1.23. Мемристоры ... 174
1.23.1. Основные понятия мемристивных технологий ... 174
1.23.2. Основные свойства мемристора ... 178
1.23.3. Анализ материалов мемристивных структур ... 180
Тесты к лекции 1.23 ... 181
1.24. Датчики ... 183
1.24.1. Основные понятия в области измерения физических величин ... 183
1.24.2. Пьезоэлектрические датчики ... 184
Тесты к лекции 1.24 ... 190
1.25. Основы квантовых вычислений ... 191
1.25.1. Основные понятия в области квантовых технологий ... 191
1.25.2. Архитектура квантового процессора ... 193
1.25.3. Физические платформы для реализации квантовых вычислений ... 198
1.25.4. Области применения квантовых компьютеров ... 203
1.25.5. Квантовые компьютеры в России ... 208
Тесты к лекции 1.25 ... 211
2. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ... 213
Лабораторная работа № 1. Определение ширины запрещенной зоны полупроводника ... 213
Контрольные вопросы и задания ... 220
Лабораторная работа № 2. Исследование термоэлектрических явлений в полупроводниках ... 221
Контрольные вопросы и задания ... 235
Лабораторная работа № 3. Исследование зависимости барьерной и диффузионной емкости p–n-перехода от приложенного напряжения ... 237
Контрольные вопросы и задания ... 243
Лабораторная работа № 4. Определение концентрации и подвижности носителей тока в полупроводнике ... 245
Контрольные вопросы и задания ... 250
3. НОРМАТИВНАЯ ДОКУМЕНТАЦИЯ ... 251
3.1. Примерная базовая программа дисциплины ... 251
3.2. Структура и состав фондов оценочных средств по дисциплине ... 258
3.3. Спецификация учебных видео- и аудиоматериалов, слайдов, эскизов, плакатов и других дидактических материалов ... 267
Литература ... 285

Учебное издание

Андреев Владимир Викторович
Жалнин Владимир Петрович
Столяров Александр Алексеевич

Физические основы микро- и наноэлектроники

Редактор *Л.Т. Мартыненко*
Художник *Я.М. Асинкритова*
Корректор *Ю.Н. Морозова*
Компьютерная графика *О.В. Левашовой*
Компьютерная верстка *Е.В. Жуковой*

Оригинал-макет подготовлен
в Издательстве МГТУ им. Н.Э. Баумана.

В оформлении использованы шрифты
Студии Артемия Лебедева.

Подписано в печать20.05.2023. Формат 70×100/16.
Усл. печ. л.23,73. Тираж 100 экз.

Издательство МГТУ им. Н.Э. Баумана.
105005, г. Москва, улица 2-я Бауманская, д. 5, стр. 1.
info@bmstu.press
https://bmstu.press

Отпечатано в типографии МГТУ им. Н.Э. Баумана.
105005, г. Москва, улица 2-я Бауманская, д. 5, стр. 1.
baumanprint@gmail.com