

Министерство образования и науки Российской Федерации

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ
И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

ФАКУЛЬТЕТ ДИСТАНЦИОННОГО ОБУЧЕНИЯ (ФДО)

П. Н. Дробот

НАНОЭЛЕКТРОНИКА

Учебное пособие

Томск
2016

УДК 621.38-022.532(075.8)

ББК 32.85я73

Д 750

Рецензенты:

С. В. Мельченко, канд. физ.-мат. наук, директор Института инноваций Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики;

С. Г. Михальченко, д-р техн. наук, профессор, заведующий кафедрой промышленной электроники, заместитель директора НИИ промышленной электроники ТУСУР.

Дробот П. Н.

Д 750 Нанoeлектроника: учебное пособие / П. Н. Дробот. – Томск : ФДО, ТУСУР, 2016. – 286 с.

Учебное пособие целостно показывает весь материал в развитии – от полупроводниковой электроники к нанoeлектронике, от квантовой физики к квантовым свойствам электронов в нанoeлектронных приборах на квантовых эффектах. Рассмотрены новейшие материалы: молибденит, алмаз, антимониды и арсениды индия, углеродные нанотрубки, графен и приборы на их основе.

© Дробот П. Н., 2016

© Оформление.

ФДО, ТУСУР, 2016

Оглавление

Введение	6
1 История нанoeлектроники и нанотехнологий	9
1.1 Bell Telephone Company: от изобретения телефона до изобретения транзистора.....	9
1.2 Гордон Мур: от компании Шокли до компании Intel, закон Мура	14
1.3 Наноразмерные объекты в традиционной полупроводниковой электронике.....	19
1.4 Внедрение в микроэлектронику новых эпитаксиальных и ионно-лучевых (плазменных) технологий	22
1.5 Технологические проблемы на пути перехода от микро- к нанoeлектронике	30
1.6 Физические проблемы и новые подходы к созданию электроники наноразмерных элементов и наноструктур	35
2 Необходимые сведения из квантовой физики.....	40
2.1 Переход от классической физики к квантовой	40
2.2 Работы Эйнштейна.....	48
2.3 Работы Резерфорда и Бора	52
2.4 Волны де Бройля и квантово-волновой дуализм частиц вещества	58
2.5 Волновая механика и уравнение Шредингера	61
2.6 Принцип запрета Паули, вырождение состояний и снятие вырождения под внешним воздействием	70
2.7 Образование энергетических зон и их заполнение в кристаллах	74
2.8 Соотношение неточностей Гейзенберга.....	77
3 Квантовое поведение электронов в кристалле	82
3.1 Уравнение Шредингера для кристалла.....	82
3.2 Модель Кронига – Пенни	84
3.3 Переход к нанofизике полупроводников.....	89

3.4 Поведение носителей заряда в прямоугольной потенциальной яме. Квантовое ограничение	92
3.4 Квантовые структуры с двумерным электронным газом	96
3.5 Квантовые структуры с одномерным электронным газом	98
3.6 Квантовые структуры с нульмерным электронным газом	100
3.7 Баллистический транспорт.....	103
3.8 Туннелирование	106
4 Полупроводниковые гетероструктуры и сверхрешетки.....	110
4.1 Гетеропереходы и гетероструктуры.....	111
4.2 Сверхрешетки	117
5 Технологии формирования квантоворазмерных наноструктур.....	123
5.1 Формирование квантовых ям.....	125
5.2 Формирование квантовых точек.....	126
5.3 Формирование квантовых проволок (нитей)	131
6 Квантовые эффекты	135
6.1 Двумерный электронный газ в магнитном поле.....	135
6.2 Целочисленный и дробный квантовый эффект Холла.....	136
6.3 Эффект Ааронова – Бома	140
6.4 Квантово-размерный эффект Штарка.....	143
6.5 Квантово-размерный туннельный эффект	150
6.6 Эффект Джозефсона	153
6.7 Кулоновская блокада	162
6.7.1 Кулоновская блокада с одним туннельным переходом.....	163
6.7.2 Кулоновская блокада с двумя туннельными переходами	166
6.7.3 Сотуннелирование	168
6.8 Квантовое магнетосопротивление. Спинтроника.....	170
6.9 Эффект Кондо.....	177
7 Устройства наноэлектроники.....	183
7.1 Приборы на одноэлектронном туннелировании.....	183
7.1.1 Одноэлектронный транзистор	184
7.1.2 Одноэлектронный насос (стандарт силы тока). Одноэлектронная память	187

7.2 Приборы на резонансном туннелировании	192
7.2.1 Диоды на резонансном туннелировании	192
7.2.2 Транзисторы на резонансном туннелировании	196
7.2.3 Логические элементы на резонансно-туннельных приборах.....	198
7.3 Приборы на основе сверхрешеток.....	202
7.3.1 Инфракрасные фотоприемники.....	203
7.3.2 Сверхрешетки в лазерных структурах.....	205
7.3.3 Лавинные фотодиоды	210
7.3.4 Оптические модуляторы	212
7.4 Транзисторы для СВЧ-электроники.....	214
7.4.1 Транзисторы с высокой подвижностью (HEMT)	216
7.4.2 Транзисторы на основе SiGe-технологии.....	222
7.4.3 Транзисторы на основе технологии нитрида галлия (GaN) на подложке из карбида кремния (SiC)	225
7.5 Перспективные материалы и приборы	227
7.5.1 Молибденит MoS_2 – перспективный материал нанoeлектроники.....	227
7.5.2 Алмаз как материал для СВЧ-приборов.....	231
7.5.3 Транзисторы на антимонидах и арсенидах индия.....	233
7.5.4 Транзисторы на углеродных нанотрубках	236
8 Графеновая электроника.....	244
8.1 Получение и открытие свойств графена.....	244
8.2 В направлении к графеновой электронике.....	247
8.3 Первые транзисторы на графене	252
8.4 Создание графеновой транзисторной технологии GNR-FETs	253
8.5 Высокоскоростные графеновые транзисторы.....	255
Заключение.....	259
Литература.....	261
Глоссарий.....	276

Введение

Наноэлектроника – это раздел электроники, который происходит от предваряющего его раздела электроники – микроэлектроники, предшествующего как по хронологии развития, так и по малости размера элементов (диоды, транзисторы и другие) интегральных микросхем. В 1959 г. Р. Нойсом была изобретена первая микросхема, для изготовления которой впервые использовались *в сочетании* следующие технологии: планарный процесс, разделение элементов запертыми *p-n*-переходами и покрытие из двуокиси кремния SiO₂ для защиты *p-n*-переходов и других элементов в нужных местах кремниевой подложки [1, 2].

С этого момента технология микросхем постоянно совершенствовалась в целях дальнейшего уменьшения размеров элементов микросхемы. Количество этих элементов на единице площади микросхемы постоянно увеличивалось, а в 1965 г. Г. Муром был эмпирически установлен закон, в соответствии с которым число транзисторов на кремниевой подложке будет удваиваться каждые 24 месяца. Закон Мура продолжает в этом виде действовать поныне, с небольшим отклонением, немного ускорившись до удвоения за 18 месяцев [2]. Об этом и о многом другом говорится в первой главе, в частности об истоках этих достижений – потребности людей общаться друг с другом быстро и удобно, через любые расстояния, в любое время. Для этого был изобретен телефон Белла (Bell, 1876), создано инновационное телефонное предприятие «Компания телефонов Белла» (Bell Telephone Company, 1877), внутри которой подразделение Лаборатория телефонов Белла (Bell Telephone Laboratory) занималось исследованиями и разработками. Именно здесь в 1936 г. была начата работа по изготовлению твердотельных переключающих и усиливающих устройств для замены вакуумных ламп и электромеханических реле в узлах телефонной связи. Эти работы были прерваны Второй мировой войной, но в послевоенные годы руководство

Bell Telephone Laboratory формирует программу глобальной связи – быстрой и надежной связи каждого с каждым в планетарном масштабе.

Динамика уменьшения размеров элементов микросхемы на уровне 100 нанометров обозначает рубеж перехода в эру нанотехнологий. Возможно, рубеж 100 нм выглядит условным. Но если вдуматься в смысл чисел, то с формализованной точки зрения все становится понятным. Если до некоторого момента мы имеем дело с микроразмерами элементов, с микроэлектроникой, то обозначить переход к существенно другой области размеров можно по критерию уменьшения размеров микроэлементов. То есть так, как это принято в физике и технике – как уменьшение на порядок. В физике и технике изменение на порядок означает изменение в 10 раз и это существенное изменение.

Поэтому рубеж в 100 нм, или 0,1 мкм, означает уменьшение на порядок относительно микронных размеров. Такое существенное уменьшение размеров, вероятно, должно приводить к существенным качественным изменениям как в нанoeлектронике, так и в нанотехнологиях. Изучению этих качественных изменений и новых возможностей электроники и технологий посвящены главы настоящей книги.

Однако изучение наномира, в котором расстояния и размеры элементов сопоставимы с размером атома и с межатомным расстоянием в кристаллах полупроводников, невозможно без хорошего понимания квантовой физики. Это – сложный предмет, законы квантовой физики непривычны. Но ведь квантовая физика стала возникать на базе классической физики. Постепенное изучение квантовой физики в той последовательности, в которой она развивалась, позволит успешно разобраться в закономерностях квантового наномира. Для этого предназначена глава 2, а в главе 3 законы квантовой физики применяются к пониманию поведения электронов в полупроводниках.

Соглашения, принятые в учебном пособии

Для улучшения восприятия материала в данном пособии используются следующие пиктограммы и специальное выделение важной информации.



Эта пиктограмма означает определение или новое понятие.



Эта пиктограмма означает теорему.



Этот блок означает «Внимание!». Здесь выделена важная информация, требующая акцента на ней. Автор здесь может поделиться с читателем опытом, чтобы помочь избежать некоторых ошибок.



В блоке «На заметку» автор может указать дополнительные сведения или другой взгляд на изучаемый предмет, чтобы помочь читателю лучше понять основные идеи.



Эта пиктограмма означает цитату.



Контрольные вопросы по главе

1 История нанoeлектроники и нанотехнологий

1.1 Bell Telephone Company: от изобретения телефона до изобретения транзистора

«Все очень просто, если знать историю». Не бывает хорошего живописца, не знающего историю живописи, не бывает хорошего музыканта, который не знает историю музыки. Точно так же не может быть хорошего ученого или инженера, не знающего историю своей профессии.

Отправной вехой на пути к транзистору, микро- и нанoeлектронике является изобретение телефона и два определяющих понятия: «телефон» и «глобальная связь».

В 1876 г. Александр Грэхем Белл изобрел звуковой электрический телефон, а через год он основал собственную компанию и назвал ее, как это часто бывает, собственным именем – Bell Telephone Company. Основной бизнес – услуги связи и производство электротехники для этого. Через год был основан первый городской телефонный узел, и такие узлы появились в крупных городах США, для их обслуживания была создана компания Bell System.

В 1885 г. была организована «дочка» Bell Telephone Company – American Telephone and Telegraph (АТТ), она развивалась так бурно, что через 15 лет в 1899 г. смогла выкупить Bell Telephone Company и стать владельцем компании Bell System. АТТ долго оставалась монополистом рынка связи в Америке. Внутри АТТ имелись электротехническое подразделение Western Electric Company и исследовательское подразделение, слияние которых в 1925 г. привело к образованию исследовательского центра АТТ – Bell Telephone Laboratories, сокращенно Bell Labs. И здесь в названии R&D центра АТТ присутствует имя Белла. Bell Labs расположилась в городке Мюррей Хилл в штате Нью-Джерси на северо-востоке США.



В 1996 г. от АТТ отделилась компания Lucent Technologies, и АТТ потеряла и Bell Labs, и Western Electric. Тем не менее, в Мюррей Хилл в 2003 г. открыта Bell Labs Nanofabrication – Лаборатория нанотехнологий Нью-Джерси. И снова имя Белла уже рядом с нанотехнологиями.

В 2006 г. Lucent Technologies слилась с французской компанией Alcatel – образована Alcatel-Lucent, а Bell Labs теперь – R&D центр Alcatel-Lucent. В 2008 г. в Alcatel-Lucent свернули исследования по физике полупроводников и полупроводниковой электронике, а Bell Labs занимается исследованиями в коммерческих интересах Alcatel-Lucent. Но за многие десятилетия исследовательский центр Bell Labs внес выдающийся вклад в развитие электроники – и микро-, и нано-. Открытия и инновации Bell Labs были критическими и в глобальной связи и оказали значительное влияние на технологическую базу мировой экономики. Например, это изобретение транзистора в Bell Labs в 1947 г., путь к которому начался на 10 лет раньше, пять лет из которых были бесплодно поглощены Второй мировой войной.



В 1936 г. директор Bell Labs Мервин Келли формирует группу ученых для исследований по замене в узлах телефонной связи ламповых усилителей полупроводниковыми [2].

Группу возглавил Дж. Беккер, привлечший теоретика У. Шокли и великолепного экспериментатора У. Браттейна. Шокли окончил докторантуру в Массачусетском технологическом институте (MIT) и поступил на работу в Bell Labs.

В 1931 г. А. Вильсон опубликовал теорию полупроводников, построенную на новейшей квантовой физике [3], и стало возможным объяснить многие эффекты в полупроводниках, например выпрямление. Поэтому в 1938 г.

Б. И. Давыдов (Физико-технический институт АН СССР под руководством А. Ф. Иоффе, г. Ленинград) [4–6], Н. Мотт (Бристольский университет, Великобритания) [7] и В. Шоттки («Сименс&Хальске», Германия) [8] независимо объясняют эффект выпрямления и предлагают теорию контактных явлений в полупроводниках.

Шокли, вдохновленный исследованием купроксных (оксид меди) выпрямителей в Bell Labs и трактовками выпрямления Давыдова, Мотта и Шоттки, взялся за дело. В 1938 г. в возрасте 26 лет Шокли делает набросок твердотельного усилителя. Идея проста – вакуумный триод, но в виде полупроводникового стержня с двумя торцевыми контактами и третьим электродом посередине (аналог сетки), напряжение на котором управляет проводимостью стержня. Шокли приступает к теоретическим расчетам, а в декабре 1939 г. он записал: «Сегодня мне открылось, что усилитель, использующий полупроводники, а не вакуум, в принципе возможен» [2]. Но попытки построить твердотельный усилитель ни к чему не приводят.



.....

Параллельно работе У. Шокли в Bell Labs в 1940 г. произошло важное событие – электрохимик Р. Ол и его коллега Дж. Скафф совершенно случайно, экспериментально, при воздействии светом, обнаружили в кремниевой пластине (из кремния, полученного в их лаборатории) наличие двух граничащих областей с разным типом проводимости. Ойл и Скафф назвали их электронной и дырочной, а «барьер», где эти области встречались, стали называть «*p-n*-переход» [9, 10]. Свет, падающий на границу *p*- и *n*-областей, стимулировал электроны для перетекания из *n*-области в *p*-область (электрический ток).

Независимо от Р. Ола и одновременно советский физик В. Е. Лашкарев впервые экспериментально обнаружил *p-n*-переход в закиси меди [11]. Он установил, что стороны «запорного слоя», расположенного параллельно границе «медь – закись меди», имели

противоположные знаки носителей тока. Также Лашкарев раскрыл механизм инжекции, лежащей в основе действия полупроводниковых диодов и транзисторов.

.....

Работы по твердотельному усилителю в Bell Labs прерываются с началом Второй мировой войны. Шокли откомандирован в распоряжение Министерства обороны, где работал по 1945 г. над проблемами, далекими от твердотельного усилителя [2]. Следует отметить и положительный для полупроводниковой электроники факт, связанный с радиолокацией, – интенсивное развитие технологий контролируемого выращивания кристаллов полупроводников *p*- и *n*-типов. По кремнию работали в Пенсильванском университете под руководством Ф. Зейтца [12, 13]. По германию – в университете Пэрдью (Purdue University, Индиана) под руководством К. Ларк-Горовица [14, 15]. В Bell Labs такие работы тоже велись во главе со Скаффом [16]. Эти работы помогли понять, как наличие примесей приводит к электронным и дырочным полупроводникам [16].



.....

В послевоенные годы фирма Bell Labs формирует программу глобальной связи – быстрой и надежной связи каждого с каждым в планетарном масштабе [2]. В то время в цепях связи для коммутации использовались электронная лампа и электромеханическое реле. Они были громоздки, малонадежны, срабатывали медленно и потребляли много энергии. Новой идее – глобальной связи – нужны новые средства реализации, и снова Bell Labs возвращается к идее использования полупроводников, вновь создается группа с научным руководителем У. Шокли. В группу входят будущие соизобретатели транзистора экспериментатор У. Браттейн и теоретик Дж. Бардин.

.....

Но идея Шокли об управлении током в полупроводнике электрическим полем никак не реализуется, разочарование, усталость заставляют физика уехать в отпуск в Европу. В его отсутствие Браттейн совершенно случайно экспериментально обнаруживает эффект усиления в 100 раз в трехэлектродной полупроводниковой структуре, причем эффекта поля как не было так и нет, а усиление возникает совсем по иной причине, с которой успешно разобрался теоретик Бардин. Новое устройство назвали германиевым триодом. Бардину и Браттейну название не понравилось. Они хотели, чтобы название заканчивалось на «тор», как резистор, термистор. Их коллега Дж. Пирс по аналогии с параметром «крутизна характеристики» («transconductance») вакуумного триода предложил назвать аналогичный параметр твердотельного усилителя «transresistance», а сам усилитель – транзистор [2].

19 декабря 1947 г. появился точечный транзистор, 23 декабря 1947 г. состоялась его презентация руководству Bell Labs. Шокли срочно возвратился в Америку, он назвал это «великолепным рождественским подарком», но успех Бардина и Браттейна глубоко задел самолюбие Шокли. Он стоял у истоков полупроводникового усилителя, но на соавторство в патенте претендовать не мог. В течение месяца он проводит интенсивные теоретические исследования и 23 января 1948 г. предлагает принципиально другую конструкцию транзистора, которую изобрел в новогоднюю ночь 1947 г., и теорию транзистора с технологически встроенными *p-n*-переходами. В этом транзисторе активные области разделены технологическими плоскостями *p-n*-переходов. В июне 1948 г. Шокли подал заявку на патент на плоскостной транзистор и в 1949 г. издал его теорию [17]. Более эффективный транзистор со слоеной структурой уравнил Шокли в правах на открытие транзисторного эффекта.



До сих пор малоизвестна страница истории о независимом от американцев изобретении транзистора в Европе в 1948 г. Два немецких физика, Г. Матаре и Г. Велкер, участвовавших ранее в военной немецкой радарной программе, после войны работали в Па-

риже, в европейском филиале американской компании Westinghouse. Независимо от американцев Велкер и Матаре создали германиевый точечно-контактный транзистор [18].

.....

1.2 Гордон Мур: от компании Шокли до компании Intel, закон Мура

А. Бекман – американский химик, изобретатель рН-метра, основатель компании Beckman Instruments, в 1956 г. профинансировал первое предприятие по производству кремниевых приборов в том местечке Калифорнии, которое ныне известно как Кремниевая долина. В сентябре 1955 г. Шокли и Бекман создали компанию Shockley Semiconductor Laboratory как отделение Beckman Instruments. Цель – «участвовать быстро и энергично в деятельности, связанной с полупроводниками». Шокли арендовал здание в г. Маунтин-Вью в Калифорнии и набрал «самую творческую команду в мире для разработки и производства транзисторов». Он привлек чрезвычайно способных инженеров, включая Г. Мура и Р. Нойса, которые разработали технологии, связанные с кремнием и диффузией. В декабре 1956 г. Шокли, Бардин и Браттейн получили Нобелевскую премию за изобретение транзистора, но сотрудники стали разочаровываться в Шокли из-за его высокомерного стиля руководства и считали, что нужно заниматься транзисторами, а не тиристорными структурами, которые давным-давно задумал Шокли.

В сентябре 1957 г. восемь служащих Шокли – Г. Мур, Р. Нойс, Дж. Блэнк, В. Гринич, Ж. Эрни, Дж. Кляйнер, Дж. Ласт и Ш. Робертс – покинули Шокли (их окрестили «вероломной восьмеркой»). «Восьмерка» при финансовой поддержке компании Fairchild Camera and Instrument организовала компанию Fairchild Semiconductor Corporation там же, в Калифорнии, в городке Пало Альто [2].

Вначале Fairchild Semiconductor производила диффузионные транзисторы с мезаструктурами, но вскоре перешла на более технологичные планарно-

диффузионные. В 1958 г. создали мезатранзистор, IBM остановила свой выбор на Fairchild Semiconductor и заказала 100 транзисторов по \$150 за штуку. Не было другого производителя, удовлетворявшего жестким требованиям к высоковольтному транзистору для запуска электронной памяти на магнитных сердечниках в компьютере сверхзвукового бомбардировщика В-70. Параллельно две команды разрабатывали *n-p-n*- (Мур) и *p-n-p*-транзистор (Эрни) [19, 20].

Только через 5 месяцев отладили выращивание кристаллов, затем разработали фотолитографию, используя 16-мм линзы кинокамеры (Ласт, Нойс), отладили нанесение алюминия для электрических связей (Мур) и собственное производство испытательной аппаратуры (Блэнк, Гринич, Кляйнер). Разработали первую технологию двойной диффузии для изготовления и эмиттера, и базы мезатранзистора. Команда Мура успешно выдала *n-p-n*-транзистор 2N697 и признание на выставке Wescon в августе 1958 г. [21–24].

Через 5 лет после изобретения транзистора возникла идея интегральной схемы (английский электронщик Дж. Даммер, 1952), но не был указан способ, как это сделать [25].

В 1958 г. Дж. Килби из Texas Instruments создал схему на основе германиевого мезатранзистора *p-n-p*-типа, части которого он химически вытравливал, чтобы получить отдельно транзистор, конденсатор и резисторы. Затем элементы соединялись золотыми проволочками в схему генератора. Технология *flying-wire* была непрактичной и неудобной в массовом производстве, с низким процентом годной продукции на выходе [26, 27].



В это время в Fairchild Semiconductor Эрни работал над повышением надежности мезатранзисторов, он повторно обратился к идее 1957 г. о новой технологии, в которой окисную пленку оставляют в нужном месте на кремниевой пластине, чтобы ею предохранить *p-n*-переходы. В 1957 г. Fairchild, сосредоточившаяся на первых транзисторах, не уделяла внимания этой технологии. В новой планарной технологии SiO_2 успешно защищал все переходы и спаи.

Сначала Эрни оформил патент в январе 1959 г., а в марте 1959 г. продемонстрировал работающий планарный транзистор. Окисная пленка успешно выполняла свою защитную функцию [28–30].

.....

.....



Приборы по планарной технологии имели лучшие характеристики, значительно снизились токи утечки, важные в проектировании компьютерной логики. *И все узлы интегральной схемы изготавливались с одной стороны пластины.* В апреле 1960 г. Fairchild запустила в производство планарный транзистор 2N1613 и лицензировала технологию. Производство интегральных схем (ИС) в миллиард транзисторов в наше время основано на прорывной идее Эрни. Один историк назвал это «самым важным новшеством в истории полупроводниковой промышленности» [32].

.....

Еще за месяц до того, как Texas Instruments сообщила об изобретении микросхемы Килби, Нойс набросал в лабораторном журнале собственную интегральную схему [32]. В 1959 г. изготовлена работающая микросхема Нойса. Во-первых, он воспользовался в ней изобретением, которое сделал К. Леховец из Sprague Electric Company [33], находившейся далеко от Калифорнии в штате Массачусетс, в 5 000 км от Пало-Альто. У Леховца компоненты разделялись *p-n*-переходами, через которые ток мог проходить только в одном направлении. В устройстве Нойса *p-n*-переходы сочетались с планарным процессом Эрни и покрытием из SiO₂. Проводящие соединения между элементами делались напылением металла в выгравированные на поверхности крошечные бороздки [32]. Такой метод был гораздо эффективнее трудоемкого впаивания тончайших проводничков, которое проводилось под микроскопом вручную при изготовлении микросхемы Килби.

В 1960 г. в мир вошла структура металл-окисел-полупроводник (МОП) и первые транзисторы на ее основе. В 1959 г. Дж. Аталла и Д. Канг в Bell Labs

получили первый полевой транзистор с изолированным затвором. Они обнаружили, что число поверхностных состояний на границе Si–SiO₂ может быть заметно уменьшено. Этот прибор достаточно медленный и малопригодный для переключателей в телефонных системах, и его производство не расширялось. Но Канг указал на простоту его изготовления и возможность применения в интегральных схемах [34, 35]. В Fairchild и RCA признали эти преимущества МОП-транзистора, и в 1960 г. К. Цайнингер и Ч. Мюллер создали МОП-транзистор в RCA, а Р. Саа из Fairchild построил управляемый МОП-тетрод. Ф. Хейман и С. Хофштейн создали в RCA в 1962 г. экспериментальную ИС с 16 транзисторами [36, 37].

Новые технологии привело к тому, что ИС Нойса были настолько практичнее ИС Килби, что даже Texas Instruments приняла их на вооружение. В 1962 г. обе фирмы массово производили ИС, вскоре прозванных «чипами». В 1960-е гг. с уменьшением размеров компонентов количество их на одном чипе возрастало очень быстро, примерно удваиваясь каждый год. Изобретение ИС было отмечено Нобелевской премией в 2000 г., ее получил Килби. К сожалению, Нойс до этого дня не дожил.



В 1965 г., через шесть лет после изобретения ИС, Мур, работавший директором отдела разработок Fairchild Semiconductors, дал эмпирический прогноз развития микроэлектроники на ближайшие десять лет – число транзисторов на кристалле будет удваиваться каждые 24 месяца. Представив в виде графика рост производительности запоминающих микросхем, он обнаружил закономерность: новые модели микросхем разрабатывались спустя приблизительно одинаковые периоды – 18–24 мес. – после их предшественников, а число компонентов в них возрастало каждый раз примерно вдвое. Это наблюдение получило название закон Мура [38].

В июле 1968 г. Нойс и Мур основали небольшую фирму, ставшую впоследствии корпорацией INTEL, с головным офисом, называемым сейчас Robert

Noyce Building, в г. Санта-Клара, все в той же Калифорнии. Вскоре к ним присоединился сослуживец по компании Шокли химик венгерского происхождения Э. Гроув. Первый бизнес-план уместился на странице А4. Спустя 2 года они создали Intel 1103 – первую запоминающую микросхему DRAM (dynamic random access memory), производимую в коммерческих масштабах.

В 1969 г. их бывшие сотрудники по работе в компании Шокли во главе с У. Сандерсом основали конкурирующую компанию AMD (Advanced Micro Devices) со штаб-квартирой в городе Саннивел (также в Кремниевой долине в Калифорнии), ставшую в США второй по значимости в разработке и производстве микроэлектронных изделий. Сандерс стал в ней бессменным ответственным руководителем и исполнительным директором.

В 1968 г. Burroughs выпустила первый компьютер на ИС, а в 1970 г. Intel начала продавать ИС памяти. В 1971 г. был сделан еще один важный шаг к персональному компьютеру (ПК): Intel выпустила ИС, аналогичную по функциям процессору большой ЭВМ. Так появился первый микропроцессор Intel-4004. Через год выпущен процессор Intel-8008, который работал в два раза быстрее Intel-4004.

В 1975 г. на конференции International Electron Devices Meeting Мур отметил, что за прошедшее 10-летие количество элементов на кристаллах удваивалось каждый год, но в будущем сложность чипов возрастет, удвоение числа транзисторов (и уменьшение их размеров) в ИС будет происходить медленнее – каждые два года.

Это новое предсказание также сбылось и закон Мура продолжает в этом виде (удвоение за два года) действовать поныне, в последнее время немного ускорившись до удвоения за 18 месяцев. Это наглядно прослеживается на примере деятельности Intel, создающей все более и более мощные микропроцессоры для ПК (рис. 1.1).

Термины «интегральная схема» (ИС), «интегральная электроника» (ИЭ) подчеркивают идею интеграции элементов, позже за ней прочно закрепилось название «микроэлектроника», которое отражает идею микроминиатюризации.

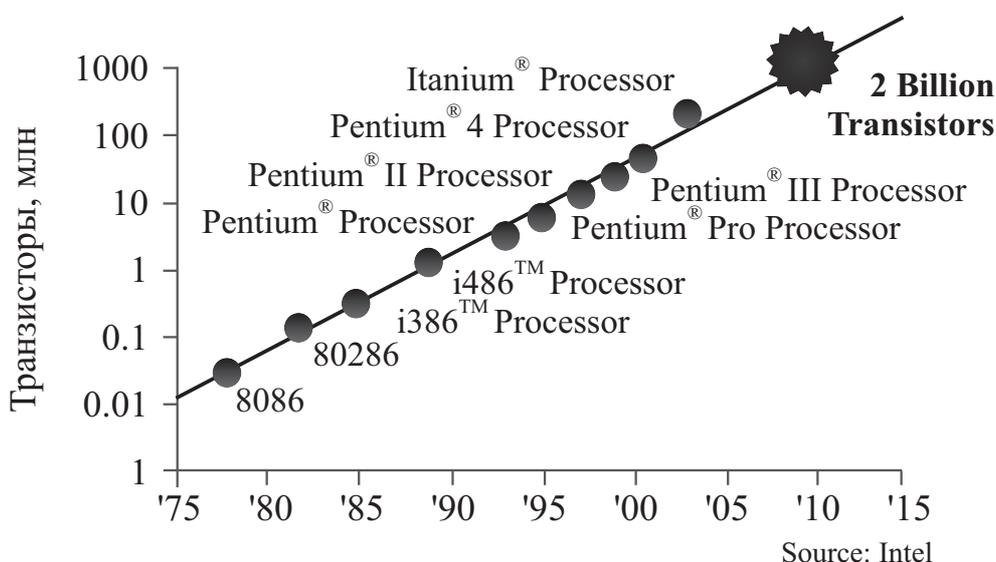


Рис. 1.1 – Закон Мура – рост числа транзисторов в чипе во времени

Отметим, что современные микропроцессоры Intel производятся по технологии 22 и 14 нм и воплощают в себе достижения нанотехнологий и нанoeлектроники, преемственно развившейся из микроэлектроники.

1.3 Наноразмерные объекты

в традиционной полупроводниковой электронике



Оказывается, в традиционной полупроводниковой электронике много примеров объектов с наноразмерами – в первую очередь это ширина области объемного заряда *p-n*-перехода в диодах, стабилитронах, туннельных диодах.

Электронно-дырочный, или *p-n-переход*, – это тонкий слой на границе двух полупроводников с разным типом проводимости, который содержит объемные заряды ионов примесей, обеднен подвижными носителями тока и обладает высоким сопротивлением. Ширина *p-n*-перехода – это ширина приграничной области, обедненной носителями заряда, которая располагается в *p*- и *n*-областях: $l_{p-n} = l_p + l_n$.

$$l_p = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_k}{e} \frac{1}{N_A}}; \quad (1.1)$$

$$l_n = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_k}{e} \frac{1}{N_D}}; \quad (1.2)$$

отсюда

$$l_{p-n} = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_k}{e} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}, \quad (1.3)$$

где N_A и N_D – концентрации акцепторов и доноров в дырочной и электронной областях, соответственно. Если $N_A = N_D$, то $l_p = l_n$ и p - n -переход называется симметричным, если $N_A > N_D$ ($N_A < N_D$), то $l_p < l_n$ ($l_p > l_n$) и p - n -переход называется несимметричным, причем он большей частью располагается в области полупроводника с меньшей концентрацией примеси.

Из выражений (1.1–1.3) нетрудно рассчитать ширину l_{p-n} при следующих данных: $\varepsilon_0 = 8,85 \cdot 10^{-12}$ Ф/м; $e = 1,6 \cdot 10^{-19}$ Кл; для обычных диодов $N_A = 10^{23}$ м⁻³; $N_D = 10^{22}$ м⁻³. Для Si $\varepsilon = 11,8$; $\varphi_k = 0,8$. Далее рассчитаем l_{p-n} для резкого асимметричного перехода, когда $N_A < N_D$.

$$l_{p-n} = \sqrt{\frac{2 \cdot 11,8 \cdot 8,85 \cdot 0,8}{1,6 \cdot 10^{-7}} \cdot \frac{1}{10^{23}}} = 1,04 \cdot 10^{-7} \text{ м}. \quad (1.4)$$

То есть для обычного диода $l_{p-n} = 104$ нм или сотни нанометров, в зависимости от величины концентрации атомов примеси.

Из соотношений (1.1–1.3) видно, что на l_{p-n} существенное влияние оказывает концентрация примесных атомов. Увеличение ее сужает запирающий слой, а уменьшение расширяет его. Это часто используется для придания полупроводниковым приборам требуемых свойств.

Стабилитроны – это диоды, предназначенные для стабилизации напряжения на нагрузке при изменении питающего напряжения или сопротивления нагрузки и т. п. Концентрация примесных атомов в стабилитроне гораздо выше, чем в обычных диодах.

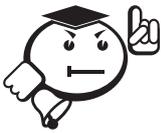
Для стабилитрона с туннельным пробоем p - n -переход образован полупроводниками с более высокой концентрацией примесей, чем в обычном диоде ($10^{23} \div 10^{25} \text{ м}^{-3}$). Оценим l_{p-n} в стабилитроне по формуле (1.2):

$$l_{p-n} = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_k}{e} \left(\frac{1}{N_A} \right)} = 10,2 \cdot 10^{-9} \text{ м.} \quad (1.5)$$

То есть в стабилитроне l_{p-n} на порядок меньше, чем в обычном диоде, — десятки нанометров, в зависимости от концентрации примесных атомов.

Для туннельного диода, в котором за счет туннельного эффекта прямая ветвь ВАХ имеет область с отрицательным дифференциальным сопротивлением, p - n -переход образован полупроводниками с еще более высокой концентрацией примесей ($10^{25} \div 10^{27} \text{ м}^{-3}$). Рассчитаем ширину p - n -перехода туннельного диода по формуле:

$$l_{p-n} = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_k}{e} \left(\frac{1}{N_A} \right)} = 1,02 \cdot 10^{-9} \text{ м.} \quad (1.6)$$



В туннельном диоде l_{p-n} составляет уже единицы нанометров, в зависимости от концентрации атомов примеси.

Возможность управления типом проводимости с помощью легирования различными примесями и идея инжекции неравновесных носителей заряда — это два важнейших элемента развития полупроводниковой электроники. Но большинство новых эффектов возникает при значительно меньших размерах структур, так как диаметры атомов полупроводников составляют лишь примерно 0,25 нм, а типичные межатомные расстояния в кристаллах близки к 0,5 нм. Дальнейший прогресс, связанный с уходом электроники в наномир, обусловлен прогрессирующим развитием полупроводниковой технологии, способной выращивать слои различных полупроводниковых кристаллов и технологически соединять их в многослойные структуры.

1.4 Внедрение в микроэлектронику новых эпитаксиальных и ионно-лучевых (плазменных) технологий

Эпитаксия (от гр. *epi* – на, над, при и *taxis* – расположение, порядок) – процесс наращивания монокристаллических слоев вещества на подложку (кристалл), при котором кристаллографическая ориентация наращиваемого слоя повторяет кристаллографическую ориентацию подложки. Эпитаксия позволяет получать тонкие однородные монокристаллические слои – эпитаксиальные слои – любого типа проводимости и любого удельного электрического сопротивления, какие невозможно создать иным способом. Различают гетероэпитаксию, когда вещества подложки и наращиваемого слоя различны по химическому составу и кристаллической структуре, и гомоэпитаксию (автоэпитаксию), когда подложка и наращиваемый слой одинаковы по химическому составу или отличаются только примесным составом.

Свойства эпитаксиальных слоев определяются сопряжением кристаллических решеток наращиваемого слоя и подложки, и существенно их структурно-геометрическое соответствие: легче сопрягаются вещества с одинаковыми или близкими кристаллическими структурами. Эпитаксия легко осуществляется, если разность постоянных решеток не превышает 10%; в этом случае тонкий эпитаксиальный слой продолжает атомные плоскости подложки. При больших расхождениях сопрягаются наиболее плотно упакованные плоскости. При различных решетках в сопрягаемых слоях возникают структурные дефекты – дислокации несоответствия, их плотностью можно управлять, меняя параметры решетки растущего кристалла (например, введением примесей) и получая бездислокационные слои с высокой подвижностью и малой плотностью носителей заряда. Кроме этого, сопряжение зависит от температуры, степени пресыщения осаждаемого вещества, совершенства подложки и чистоты ее поверхности. Поэтому подложку предварительно подвергают механической, химической или радиационной обработке. Эпитаксиальный слой растет за счет атомов и молекул, составляющих адсорбционный слой, и скорость роста зависит от пресыщения в этом слое.

Эпитаксия возможна из любой фазы: газовой (газофазная эпитаксия – ГФЭ), жидкой (жидкостная, или жидкофазная, эпитаксия – ЖФЭ) и твердой (твердофазная эпитаксия – ТФЭ). Преимущественное развитие получили ГФЭ и ЖФЭ.

Методы ГФЭ делятся на химические и физические. Химические методы ГФЭ основаны на осаждении из газовой фазы вещества, полученного в результате следующих химических реакций: восстановления хлоридов Si и Ge водородом; пиролитического разложения моносилана; диспропорционирования диоксидов и дихлоридов Si и Ge. Процессы осуществляются в реакторах, газовая система подает в реакторную камеру газовую смесь требуемого состава. Добавляя к газовой смеси соединения легирующих элементов, выращивают слои *n*- или *p*-типа соответственно. Температура процесса определяется кинетикой химической реакции и обычно в пределах 800–1300°C.



.....

Еще в 1951 г. Е. Тил и Г. Кристенсен в Bell Labs разработали технологию осаждения эпитаксиального слоя [39]. В СССР в 1957 г. Н. Н. Шефталь, Е. Ю. Кокориш и А. В. Красилов описали подобную работу на германии и кремнии [40].

.....

Пленки германия и кремния, образующиеся в результате восстановления водородом их хлористых соединений (GeCl_2 и SiCl_4) при высокой температуре, кристаллизуются на подложке из монокристалла германия или кремния; при этом в газовую фазу вводятся желаемые примеси в виде газообразных хлористых соединений, которые также восстанавливаются водородом и кристаллизуются вместе с германием или кремнием [40].

В 1960 г. команда Bell Labs во главе с Г. Тойрером использовала ГФЭ для наращивания пленки кремния между подложкой и коллектором транзистора. Технологию быстро перенесли в Western Electric для производства кремниевых транзисторов для телефонных переключателей в Bell System.



.....

Так в 1960-е гг. появляется эпитаксиальная технология, которая увеличила на порядок напряжение пробоя транзистора и резко увеличила его скорость переключения – важные характеристики для проектирования ИС. Добавленная стоимость производства была более чем возмещена улучшением рабочих характеристик.

.....

В это трудно поверить, но, как это обычно бывает, нечто новое, революционное, выходящее за рамки традиционных представлений, всегда встречается в штыки сторонниками привычных подходов и решений. Так и идеи миниатюризации электроники поначалу активно критиковались.



.....

Ситуация в корне изменилась благодаря подготовке полета американского космолета «Аполлон» на Луну и разработке первой в мире твердотопливной межконтинентальной ракеты «Минитмен». Использование ИС в этих программах определило всеобщее признание интегральной электроники [41, 42].

.....

На развитие советской микроэлектроники в 1960–70-е гг. сильно влияла блокирующая деятельность КОКОМ (Координационный комитет по экспортному контролю), по многостороннему контролю над экспортом в СССР и другие соцстраны (стратегия *контролируемого технологического отставания*) [43, 44].



.....

В отечественной практике освоение планарной технологии было начато только в 1962 г. *с нулевого уровня* [45] (толчком послужили американские ИС для системы наведения ракеты «Минитмен»).

.....

В Зеленограде в 1962 г. был открыт НИИ микроприборов (НИИМП) и через полгода были получены первые результаты по тонкопленочной технологии [46].

К физическим методам ГФЭ относят методы термического осаждения из молекулярных пучков в вакууме – молекулярно-пучковая эпитаксия (МПЭ); мгновенного испарения, «горячей стенки», а также методы катодного распыления и осаждения. По методу термического осаждения из молекулярных пучков испаряемое вещество нагревается до требуемой температуры (выше или ниже температуры плавления испаряемого вещества в зависимости от упругости пара в точке плавления) в сверхвысоком вакууме ($\sim 1,3 \cdot 10^{-8}$ Па), при этом его атомы и молекулы попадают на подложку, где и происходит их конденсация.



МПЭ разработана в 70-е гг. XX в. и представляет собой результат фантастического усовершенствования старого способа, широко применявшегося для приготовления металлических пленок, – термического испарения в вакууме. Использование чистых источников напыляемых материалов, сверхвысокий вакуум, точный контроль температуры подложки, различные методы диагностики растущей пленки в сочетании с компьютерной системой управления параметрами процесса – все это вместе взятое привело к созданию качественно новой технологии, способной решать задачи, сама постановка которых раньше казалась невыполнимой.

МПЭ превратилась сегодня в одну из важнейших технологий для выращивания гетероструктур на основе соединений $A^{III}B^V$, прежде всего благодаря пионерской работе А. Чо [47].

МПЭ происходит в сверхвысоком вакууме, поэтому ее можно контролировать *in situ* тонкими методами: дифракция отраженных быстрых электронов, электронная Оже-спектроскопия (ЭОС); вторично-ионная масс-спектрометрия (ВИМС); рентгеновская фотоэлектрическая спектроскопия (ФЭС) и т. д.

В систему помещают соответствующую аппаратуру вместе с квадрупольным масс-анализатором для контроля интенсивности пучков и ионной пушкой для очистки поверхности. Эти возможности контроля и анализа дают МПЭ существенные преимущества перед другими методами.

Перечислим решение важнейших задач, обеспечиваемых МПЭ:

- а) получение монокристаллов высокой чистоты – за счет роста в сверхвысоком вакууме и высокой чистоты потоков веществ;
- б) выращивание сверхтонких структур с резкими изменениями состава на границах – за счет относительно низких температур роста, препятствующих взаимной диффузии;
- в) получение гладких бездефектных поверхностей для гетероэпитаксии – за счет ступенчатого механизма роста, исключающего возможности зародышеобразования;
- г) получение сверхтонких слоев полупроводников, диэлектриков и металлов (10–100 нм) с контролируемой толщиной – за счет точности управления потоками и относительно малых скоростей роста;
- д) создание структур со сложными профилями состава и (или) легирования, сверхрешетки с толщиной 5–10 нм.

В методе МПЭ испарение происходит из специальных тиглей, называемых эффузионными ячейками, которые нагреваются индукционным способом до нужной температуры испарения. Наиболее совершенным является электронно-лучевой способ нагрева, отчего такой метод получил название *молекулярно-лучевая эпитаксия* (МЛЭ) (а также за счет перевода на русский слова «пучок» – *beam* – обычно, как «луч»). Это два названия одного и того же метода. На основе пленок, полученных методом МПЭ (МЛЭ), создают оптоэлектронные ИС, сверхбыстродействующие большие ИС, фотоприемники и лазеры на гетероструктурах и другие приборы.

Методу МЛЭ присущи и недостатки: сложность получения высокого вакуума; высокая трудоемкость; высокая стоимость оборудования; низкая скорость роста пленок; сложность получения высоколегированных слоев.

Кроме методов МПЭ (МЛЭ) прогресс в электронике достигался с помощью ионно-лучевых и ионно-плазменных методов.

Ионно-лучевое осаждение – способ нанесения на подложку тонких пленок металлов, полупроводников или диэлектриков с помощью сфокусированных пучков (лучей) ионов низких энергий (5–100 эВ). При ионно-лучевом осаждении под действием ионного пучка повышается эффективная температура поверхности конденсации (за счет рассеивания кинетической энергии падающих ионов), возрастает десорбция загрязнений, активизируются процессы миграции и химического взаимодействия атомов осаждаемой фазы и подложки, увеличивается число адсорбционно-активных центров зародышеобразования и повышается нетермическая активация поверхности и приповерхностного слоя подложки. Кроме того, благодаря способности ускоренных ионов проникать в тело подложки на глубину нескольких атомных слоев резко увеличивается адгезия осаждаемых пленок к подложке.

Ионно-лучевое осаждение позволяет равномерно наносить пленку на поверхность объемных тел без сложных перемещений их в пространстве, выращивать эпитаксиальные слои при обычной температуре подложки, получать пленки с высокой плотностью. При торможении ионов в приповерхностном слое подложки или осаждаемом веществе возникают микрообласти, где в течение очень короткого интервала времени $\sim 10^{-11}$ с может развиваться давление до 10^8 Па, а температура достигать 100–1000°C, что позволяет получать пленки веществ, не существующих в свободном состоянии в природе или находящихся при нормальных условиях в метастабильном состоянии, а также создавать пленки с уникальными свойствами. Например, тонкие пленки углерода обладают хорошими изоляционными свойствами, высокой химической и радиационной стойкостью, повышенной адгезией к различным подложкам. Такие пленки могут быть использованы в качестве подзатворного слоя в МДП-структурах, для обеспечения надежного контакта мощных транзисторов с теплоотводящим радиатором и т. д. Важной областью применения ионно-лучевого осаждения

является получение пленок полупроводниковых соединений типа CdS, CdTe, CdSe и других и создание на их основе фотоприемников, солнечных батарей.

Ионно-плазменная технология – совокупность приемов и способов получения и обработки материалов и изделий электронной техники с использованием компонентов газоразрядной низкотемпературной плазмы. Применяется в производстве полупроводниковых приборов, например для очистки поверхности подложек и нанесения на них тонких пленок металлов, полупроводников или диэлектриков посредством ионно-плазменного распыления, ионного или плазмохимического осаждения; ионного или ионно-лучевого травления тонких пленок для получения ИС определенной топологии; удаления различных резистов (фото-, рентгено-, ионо- и электронрезистов) посредством ионно-химического и плазмохимического травления; внедрения примесей в металлы, полупроводники и диэлектрики бомбардировкой их ускоренными ионами легирующих веществ; ионно-лучевой обработки ионорезистов в процессах ионной литографии.

Ионно-плазменная технология нанесения тонких пленок представляет собой ионное распыление – разрушение поверхности твердых тел в результате бомбардировки их ионами в вакууме. Первоначально называлось катодным испарением, поскольку под действием бомбардирующих частиц катод «испарялся» даже при температуре ниже температуры плавления его материала. Позднее применительно к технологическим процессам утвердился термин «катодное распыление», так как на обрабатываемое тело подавался отрицательный потенциал, ускорявший движение положительных ионов в направлении поверхности этого тела. Однако термин «катодное распыление» применим лишь к части процессов распыления, часто отрицательный потенциал на обрабатываемое тело не подается и разрушение его поверхности достигается с помощью сформированного и ускоренного до необходимой энергии потока ионов.

Ионное распыление используется в основном для травления (очистки) поверхности подложки (мишени), а также для получения тонких пленок осаждением на подложку распыленного вещества мишени. Обычно используются

ионы инертных газов (He^+ , Ne^+ , Ar^+ , Kr^+ , Xe^+) с энергией 0,1–10 кэВ. Наиболее эффективно процесс протекает при бомбардировке ионами с энергией 0,1–0,5 кэВ. При этом глубина проникновения ионов в материал ограничивается несколькими атомными слоями, что создает наиболее благоприятные условия для передачи атомам энергии, достаточной для их вылета с бомбардируемой поверхности.

Интенсивность ионного распыления характеризуется коэффициентом распыления, который определяется как отношение числа выбитых атомов материала к числу падающих на него ионов и зависит от энергии, угла падения, массы и атомного номера бомбардирующих ионов, а также от плотности, массы, атомного номера и энергии поверхностной связи атомов распыляемого материала. Так, наибольшее значение коэффициент распыления атомов Ag, Cu, Mo при бомбардировке их ионами Ar^+ с энергией $10\text{--}5 \cdot 10^4$ эВ составляет соответственно 12, 7 и 2 атом/ион.

Распыляется материал в основном в виде отдельных нейтральных атомов, средняя энергия которых составляет 3–5 эВ, то есть в 10–100 раз превышает энергию испаряемых частиц, поэтому адгезия к подложке у пленок, осажденных из распыленного вещества, существенно выше, чем у пленок, полученных вследствие испарения. Количество распыленных конгломератов атомов (кластеров) незначительно (несколько процентов); в небольших количествах (~1%) вещество мишени распыляется в виде ионов.

Существует несколько способов ионного распыления. Наиболее простой из них – «диодное» распыление. Устройство для «диодного» распыления представляет собой вакуумную камеру (заполняемую после откачки инертным газом, напр. Ar, до давления 5–15 Па) с двумя электродами – катодом и анодом, отстоящими друг от друга на расстоянии 4–5 см. Обычно катодом служит распыляемая мишень, а роль анода выполняет металлическая пластина (подложкодержатель) с подложками, подготовленными для осаждения распыляемого вещества мишени. При подаче на катод (мишень) отрицательного (относительно

анода) напряжения 4–5 кВ зажигается аномальный тлеющий разряд, и в результате возникает поток ионов, бомбардирующих мишень.

В установках для ионного распыления «со смещением» подложкодержатель не является анодом, а сам находится под отрицательным потенциалом (100–200 В) относительно анода, что вызывает ионную бомбардировку осаждаемых пленок на подложках, удаляющую с их поверхности атомы и молекулы примесей.

Для уменьшения рабочего давления (до 10^{-2} Па) и повышения эффективности ионизации газа часто используют вспомогательный низковольтный дуговой разряд – источник ионизирующих электронов, который образуется между вспомогательным катодом и анодом. Такой способ называется «триодным» распылением.

Наибольшая эффективность ионного распыления достигается с применением магнитного поля, увеличивающего длину пробега электронов и, следовательно, степень ионизации рабочего газа, позволяя снизить давление в камере и увеличить мощность разряда. Этот способ распыления наиболее удачно реализуется в магнетронных распылительных системах. Магнетронное распыление обеспечивает скорость осаждения пленок до 10–30 нм/с и позволяет наносить пленки любых материалов, в том числе диэлектриков, с высокой степенью воспроизводимости, что обусловило широкое применение этого метода в производстве электронной техники.

1.5 Технологические проблемы на пути перехода от микро- к нанoeлектронике

В соответствии с рисунком 1.2 переход к технологии 90 нм – это переход к нанотехнологиям.

С момента появления закона Мура за истекшие 50 лет скептики сотни раз предсказывали ему скорую «кончину», однако ученые и инженеры открытиями и трудом снова подтверждали провидение и безупречность выводов основателя

Intel. Закон Мура продолжает действовать и в соответствии с ним Intel вводит новые техпроцессы каждые два года.

В 2005 г. началось производство чипов по технологии 65 нм, в 2007 г. перешли на 45 нм, в 2009 г. – 32 нм, в 2011 г. – 22 нм, в 2015 г. – 14 нм.

В 2015 г. мир отметил 50-летие выполнения закона Мура (рис. 1.3), и в связи с этим появилась масса публикаций об этом, например [48, 49].

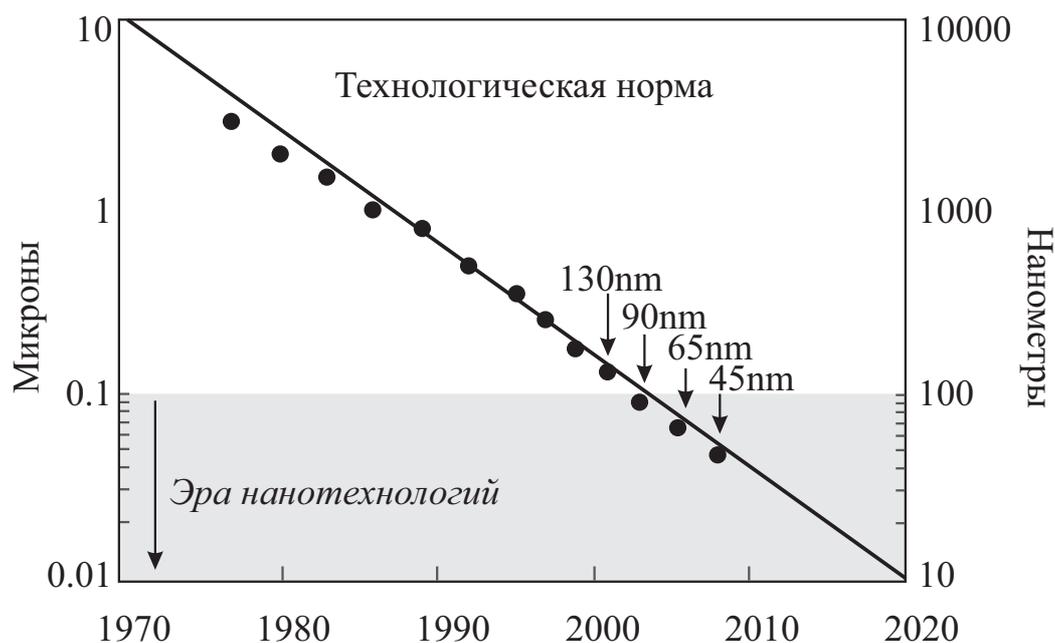


Рис. 1.2 – Динамика уменьшения размеров элементов микросхемы. Переход от микро- к нанотехнологиям, начиная со 100 нм в 2003 г.



Рис. 1.3 – Гордон Мур, 2015 г. [49]. Фото: Paul Sakuma / AP / East News

Но ценой каких технологических усилий удастся разработчикам на протяжении полувека сохранять действие закона Мура? Очевидно, что уменьшение размеров электронных элементов ИС будет неизбежно сталкиваться с физико-техническими и технологическими проблемами [50, 51].

1. Проблема с проводящими фрагментами ИС.

Традиционный алюминий при малой толщине полупроводника или металла термодинамически несовместим с используемым диэлектриком SiO_2 . Необходима замена алюминия на материал с более высокой проводимостью.

2. Проблема подзатворного диэлектрика полевых транзисторов.

Полевой транзистор состоит из канала, образованного областями стока и истока, и управляющего электрода – затвора, отделенного от канала подзатворным диэлектриком. В приборах наноразмера необходим материал с большей диэлектрической постоянной ϵ , чем у SiO_2 . Уже при 65-нм технологии толщина подзатворного диэлектрика составляет 1,2 нм (всего лишь 5 атомных слоев кремния) и дальнейшее ее уменьшение неизбежно приводит к катастрофическому росту токов утечки. Чтобы избежать этой утечки и сохранить емкость материала, которая определяет скорость переключения транзистора, необходимо перейти на другой материал, который немного толще, но при этом имеет схожую емкость.

3. Проблема изолятора в металлической разводке проводящих фрагментов. Изолятор должен иметь меньшее значение ϵ .

Проблема № 1. Для формирования новых проводящих слоев применялись варианты химического метода ГФЭ с использованием летучих металлоорганических соединений: импульсный МОСВД (metalorganic chemical vapour deposition), СВД (chemical vapour deposition) с ультрафиолетовой (УФ) стимуляцией и атомное послойное нанесение АЛД (atomic layer deposition).

Были получены как металлические пленки и покрытия Ir, Ru, Cu, Au, так и наночастицы Au, Cu. В качестве исходных соединений для осаждения металлических слоев Ir, Ru, Cu были использованы летучие комплексы металлов с бета-дикетонатными производными, для осаждения пленок и наночастиц Au –

карбоксилаты и салицилальдиминаты диметилзолота. Для исходных прекурсоров разработаны методы синтеза с высоким выходом конечных продуктов, исследованы и определены режимы осаждения [50].

Проблема № 2. Новым подходящим материалом оказался оксид гафния и исследования направлены на создание тонких слоев диэлектрика на основе HfO_2 , например работы лабораторий Института неорганической химии им. А. В. Николаева СО РАН (г. Новосибирск) [50] и лабораторий Intel [51]. Выбор HfO_2 определяется его высокой $\epsilon \sim 25$, большой шириной запрещенной зоны ($\sim 5,6$ эВ). Слои готовились методом ГФЭ с использованием в качестве исходных веществ дипивалоилметаната гафния, $\text{C}(\text{CH}_3)_3\text{COCHCO}(\text{CH}_3)_3$, и дициклопентадиэнил гафний бисдиэтиламина, $(\text{C}_5\text{H}_5)_2\text{Hf}(\text{N}(\text{C}_2\text{H}_5)_2)_2$ [50]. Важные проблемы: стабильность композиции HfO_2/Si и достаточно высокие пробивные напряжения образующихся слоев диэлектрика. По стабильности был проведен термодинамический анализ системы Hf-Si-O , показавший совместимость оксида гафния с кремнием [50].

Переход к новому диэлектрику с существенно большей ϵ позволил достичь требуемой емкости при приемлемых толщинах диэлектрика (рис. 1.4). В англоязычной литературе ϵ обозначается символом « k », фактически, греческая буква «каппа», отсюда термин *high-k*.

Емкость плоского конденсатора $C = \frac{\epsilon\epsilon_0 S}{d}$, где S – площадь обкладки конденсатора; d – толщина диэлектрика. При переходе к подзатворному диэлектрику из HfO_2 толщина диэлектрика увеличивается в 2,5 раза, следовательно емкость уменьшается относительно традиционного случая с SiO_2 тоже в 2,5 раза, и это ухудшает свойства транзистора. Но у HfO_2 $\epsilon = 20$ и гораздо больше, чем у SiO_2 , $\epsilon = 4$, так что в целом емкость новой структуры возрастает в 2 раза, токи утечки снижаются и получается эффект улучшения. Это был столь существенный технологический скачок, что Мур написал:



«Реализация *high-k*... означает наиболее революционное изменение в технологии со времени изобретения MOS-транзисторов с поликремниевыми затворами в конце 60-х годов» [51].

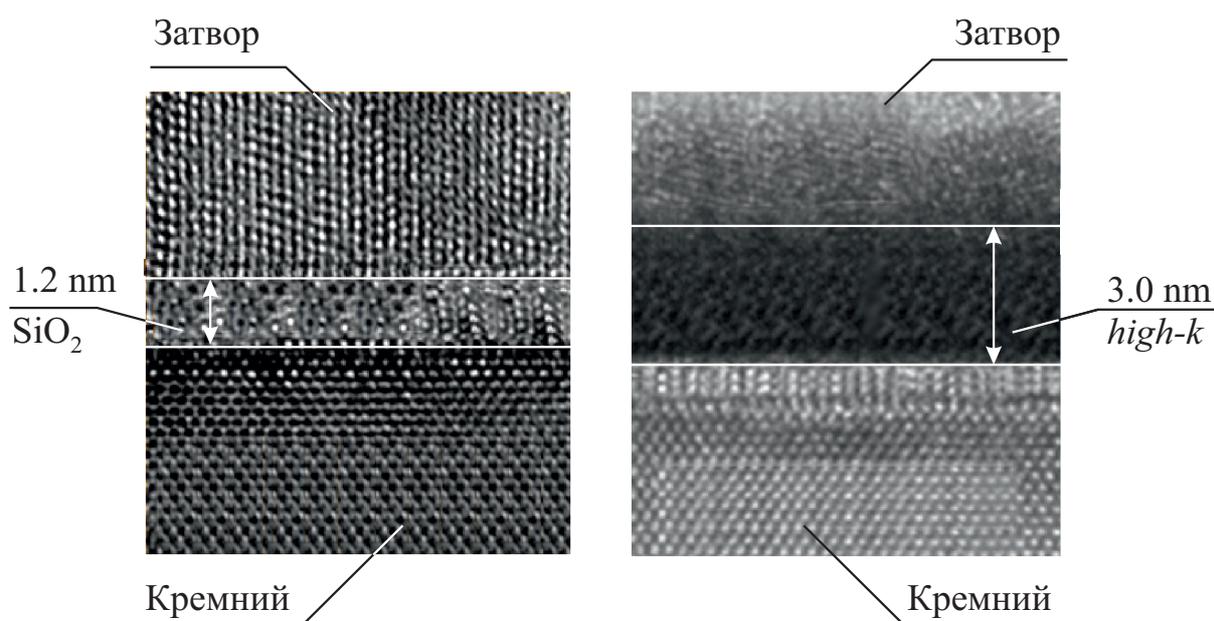


Рис. 1.4 – Структура полевого транзистора с разным подзатворным диэлектриком: диоксид кремния SiO_2 (слева) и оксид гафния HfO_2 (справа)

Проблема № 3. В качестве нового диэлектрика рассматривались карбонитриды В и Si с низкими значениями ϵ , и оказалось возможным синтезировать слои с широкой вариацией свойств, с низкими, по сравнению с SiO_2 , значениями $\epsilon = 3,0$ [50]. Новые вещества, методы и режимы процессов синтеза слоев и структур полезны для продления закона Мура, для дальнейшей миниатюризации кремниевых приборов. Этот процесс будет продолжаться еще 10–15 лет. Дальнейшее увеличение плотности размещения элементов не прекратится и после достижения физического предела кремниевой электроники. Создаются заделы будущих устройств и технологий. Перспективны различные нанобъекты: углеродные и другие нанотрубки, супрамолекулярные соединения.

Вероятней всего приборы следующих поколений будут гибридами, построенными из традиционных фрагментов и новых элементов, преобладающим

подходом будет создание устройств путем синтеза из атомов и молекул (*down up way*), важную роль будут играть различные варианты процессов химического осаждения из газовой фазы (CVD).

1.6 Физические проблемы и новые подходы к созданию электроники наноразмерных элементов и наноструктур

Приблизительно до 2020 г. транзисторы в интегральных наносхемах будут создаваться по современной схеме работы – с двумя электродами и затвором между ними. Пределом этой тенденции было бы *достижение атомного размера элементов электроники*, когда уменьшать их дальше будет невозможно, и уже необходимо искать новые подходы.

Одним из таких новых подходов является организация передачи сигнала с использованием спиновых волн. По мере приближения к нанометровой области, соизмеримой с размерами атомов, проявляются квантовые свойства электронов, и в наноэлектронных приборах поведение электрона определяется его волновыми свойствами, его волновой функцией. Процессы дрейфа и диффузии вовсе отсутствуют в наноэлектронных элементах, в их основе лежат полевые связи, сформированные потенциальные барьеры. «Вход» и «выход» элемента локализованы не в пространстве, а во времени. Их определение возможно только в промежутки времени, когда существует определенный порог внешних воздействий, соответствующий «входу» или «выходу». Этот промежуток времени получил название рефракторного периода и с его помощью обеспечивается распространение сигнала в определенном направлении. Переход от одного устойчивого состояния наноструктуры в другое происходит через возбужденные неустойчивые состояния.

Вторая чисто теоретическая идея заключается в многократном использовании электронов. В современных архитектурах электроны перемещаются от истока к стоку, а затем теряются, но можно производить массу операций, не теряя электронов [51].

Третья альтернатива – углеродные нанотрубки и кремниевые нанопровода (квантовые нити, шнуры), в которых можно достичь более высокой скорости электронов. Транзисторы, изготовленные из таких материалов, имеют сопоставимые с современными размеры. Диаметр углеродных нанотрубок 1–2 нм, но в экспериментальных транзисторах исток и сток расположены по их длине. Это позволяет повысить быстродействие и уменьшить потребляемую энергию, однако размер сильно не сократится. Углеродные нанотрубки могут найти применение в технологии КМОП (комплементарные металл-оксидные полупроводники) не столько для миниатюризации, сколько для повышения производительности или упрощения изготовления. Даже если будет изобретено принципиально иное средство перемещения электронов, повышение плотности и производительности не пойдет дальше пределов технологии КМОП из-за ограничений отвода тепла [51].

Следующий подход – производить чипы больших размеров, наращивая их площадь или строя трехмерные многослойные микросхемы. Технология трехмерной сборки кристаллов (*system in package*) или 3D-сборка является перспективным методом, снижающим размеры микросхем за счет повышения плотности упаковки, увеличивающим пропускную способность соединений внутри кристалла и уменьшающим энергопотребление.

Трехмерная технология производства позволяет наращивать количество транзисторов на кристалле со скоростью, соответствующей закону Мура, причем снижает цены в расчете на устройство. Вертикальное размещение элементов существенно увеличивает плотность и способствует многократному сокращению затрат на производство ИС по сравнению с традиционными ИС при той же технологии производства.

Компании Taiwan Semiconductor Manufacturing Corporation (TSMC, www.tsmc.com) и Matrix Semiconductor (www.matrixsemi.com) в 2001 г. первыми представили трехмерную микросхему памяти «Matrix 3D Memory» – постоянное запоминающее устройство (ПЗУ) с однократной записью. Она изготавливалась по технологии многослойного размещения транзисторов и содержала че-

тыре транзисторных слоя. Для массового производства ряд технологических проблем был преодолен к 2004 г. Потенциальный рынок для них относительно узок, но себестоимость вдвое меньше, чем у флэш-памяти. Основные потребители – заказчики, желающие выпускать кристаллы с записями аудио, видео и другими неизменяемыми файлами. Новые микросхемы энергонезависимой памяти совместимы по контактам с кристаллами флэш-памяти, и карты Matrix 3D Memory можно устанавливать в разъемы для флэш-носителей разнообразных цифровых устройств, применяя их в качестве недорогой альтернативы для однократной записи информации. В 2005 г. Matrix Semiconductor анонсировала новый кристалл трехмерной памяти с 4 уровнями транзисторов емкостью 1 Гб и размером всего 31 кв. мм (рис. 1.5).

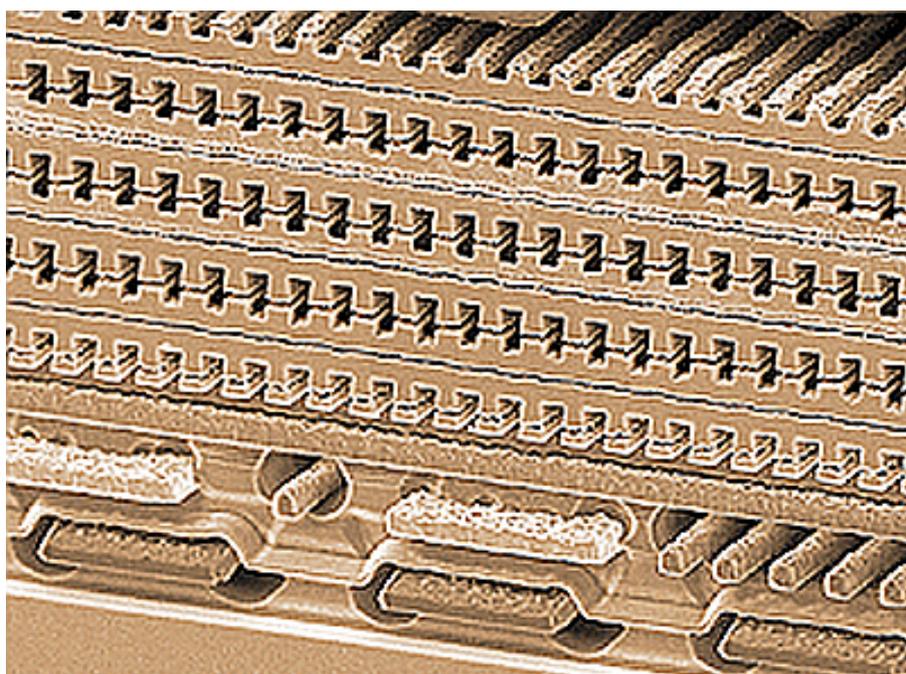


Рис. 1.5 – Кристалл Matrix 3D Memory содержит четыре слоя транзисторов

В результате применения технологии System-in-Package (SiP) возможно совмещать в одном корпусе произведенные по различным технологиям цифровые и аналоговые схемы памяти и микроэлектромеханические системы. 3D-сборка SiP повышает надежность и снижает себестоимость ИС.

В 2008 г. мировой рынок 3D-сборки составил 1,3 млрд долларов, согласно прогнозам в 2015 г. его объем составит 42 млрд долларов.

«Система в корпусе» SiP – очередной этап повышения степени интеграции стандартных модулей электронных устройств. По существу, используется старая идея гибридных микросхем, однако новые технологии принципиально изменяют методы изготовления. «Системы в корпусе» дают преимущества по расширению функциональности и снижению объема, а также возможность сократить цикл проектирования в целом.

Вопросы технологии сборки и тестирования кристаллов многослойной ИС достаточно непросты (рис. 1.6). Кристаллы могут быть расположены как в планарном монтаже, так и в 3D-сборках на общей подложке. Высота таких сборок – весьма критический параметр конструкции SiP. При стыковке до 8–10 кристаллов возможно сделать сборку с высотой 1,2 мм, но требования миниатюризации уже ставят требование к высоте в 1–0,8 мм. Это ставит ряд ограничений по размерам шариков пайки и петель электропроводов для распайки, по толщине подложки и капсулирующего покрытия. Все кристаллы производятся на пластинах стандартной толщины, но чтобы реализовать сборку из 8 кристаллов, пластина изначально утончается в 7 раз. Для этого применяются несколько методов – механическое полирование, химико-механическое полирование, мокрое или сухое травление и плазменная обработка. После обработки механическая прочность пластины сильно уменьшается и при дальнейших обработках весьма вероятно появление трещин и зон с внутренними напряжениями. По этой причине обязательны тестирование и отбраковка после утончения.

В 3D-компановках кристаллы могут быть одинакового или разного размера. Поэтому и применяется непосредственная пирамидальная стыковка (при кристаллах разного размера) или стыковка с прокладками (*interposer*) при кристаллах одинакового размера. Точность стыковки ухудшается из-за разных коэффициентов адгезии отдельных кристаллов. Разварка электрических соединений для кристаллов также отличается существенными ограничениями – изгиб разварки меньше стандартного в 1,5 раза. Технология сборки, хотя достаточно

точно определена, требует внимательной проработки в каждом отдельном случае и является всегда проблемно-ориентированной.

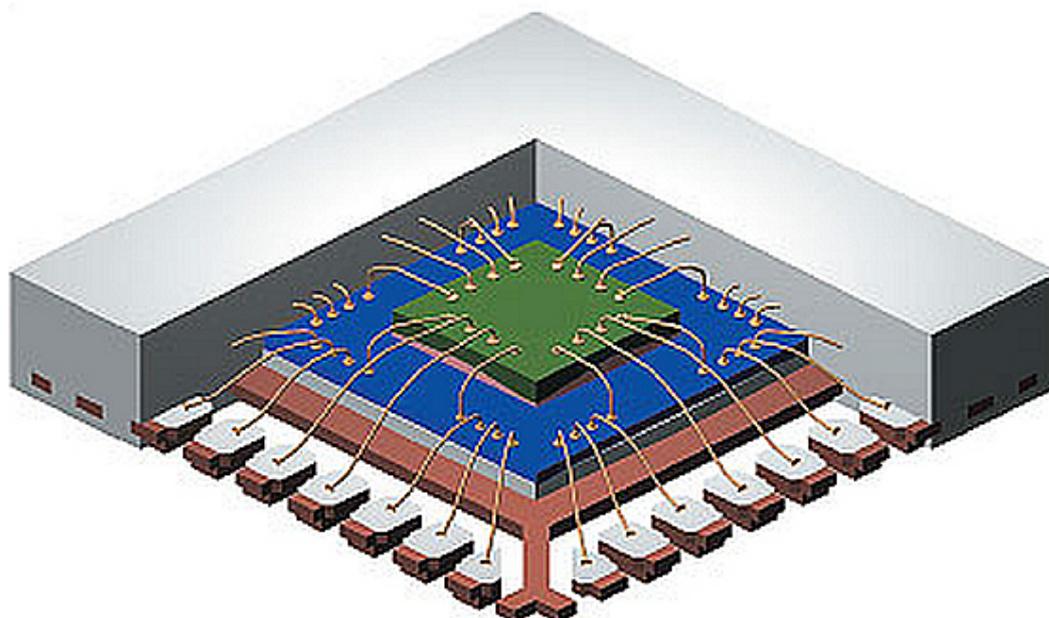


Рис. 1.6 – Сборка многослойной трехмерной микросхемы



Контрольные вопросы по главе 1

1. Опишите послевоенную программу Bell Labs по глобальной связи, созданную для нее научную группу и начало ее работы.
2. Опишите изобретение и устройство точечного германиевого транзистора. Кто и когда его изобрел?
3. Каковы принципиальные качественные изменения, связанные со значительными уменьшениями размеров элементов ИС?
4. Назовите и опишите проблемы перехода от микро- к нанoeлектронике.
5. Назовите новые подходы к созданию электроники наноразмерных элементов и наноструктур.

2 Необходимые сведения из квантовой физики

2.1 Переход от классической физики к квантовой

Все не раз слышали о квантах света, и что немецкий физик М. Планк открыл квант энергии, и что нагретое тело излучает энергию дискретно, а не непрерывно, как считали долгое время. Обычно кое-что слышали о том, что вещество тоже состоит из квантов, но есть ли здесь связь с квантами света? Как были открыты кванты, мало кто знает, а за этим незнанием часто стоит непонимание всей квантовой физики. Поэтому рассмотрим открытие кванта и переход к квантовой физике. Мы будем пользоваться сборником русских переводов работ Планка [52]. Для физиков XIX в. особый интерес представляло тепловое излучение нагретых тел в связи с тем, что оно находится в тепловом равновесии с нагретыми телами, в отличие от всех других видов излучения. При тепловом излучении энергия, уносимая от источника, подводится к нему за счет нагрева. Исследуя закономерности теплового излучения, физики стремились установить связь между оптикой и термодинамикой. Тепловое излучение, находящееся в тепловом равновесии с телами, нагретыми до температуры T , называется равновесным, или черным, излучением, и у него плотность энергии и спектральный состав определяются только температурой и не зависят от природы излучающего вещества. Если тело может поглощать всю падающую на него лучистую энергию любого спектрального состава, то его называют «абсолютно черное тело». При $T = \text{const}$ тепловое излучение абсолютно черного тела должно иметь тот же спектральный состав, что и окружающее это тело равновесное излучение. Поэтому нужно изучить спектральный состав излучения абсолютно черного тела.



В эксперименте абсолютно черное тело моделируется замкнутой металлической полостью с небольшим отверстием, из которого происходит излучение абсолютно черного тела, когда полость

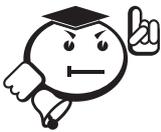
нагрета до температуры T . С ростом T будет расти энергия излучения из отверстия и меняться его спектральный состав. Для абсолютно черного тела распределение энергии по длинам волн при $T = \text{const}$ характеризуется излучательной способностью $r(\lambda, T)$, и это мощность излучения единичной площадки поверхности по всем направлениям в интервале длин волн $\Delta\lambda$.

.....

Тогда полный поток излучения для всех длин волн получится интегралом от $r(\lambda, T)$ по длине волны $R(T) = \int_0^{\infty} r(\lambda, T) d\lambda$. Излучательная способность $r(\lambda, T)$, или через частоту $\nu - r(\nu, T)$, называется также спектральная светимость, а $R(T)$ – интегральная светимость.

Эксперимент показал, что при $T = \text{const}$ функция $r(\lambda, T)$ имеет выраженный максимум, который при росте T смещается в область коротких волн (рис. 2.1).

.....



Немецкий физик Вильгельм Вин в 1893–1896 гг. вывел выражение для спектральной светимости:

$$r(\nu, T) = C_1 \nu^3 e^{-C_2 \nu / T} \quad (2.1)$$

(C_1 и C_2 – константы) – и показал, что максимум спектральной светимости смещается в область коротких волн при нагреве абсолютно черного тела.

.....

Этот закон назвали закон смещения Вина (рис. 2.1), причем произведение длины волны λ_m максимума на температуру T остается постоянным: $\lambda_m \cdot T = b$, где b – константа Вина, $b = 2,898 \cdot 10^{-3}$ м·К. Формула Вина хорошо соответствовала эксперименту в коротковолновой области. Английские физики на основе термодинамики просчитали спектральную светимость $r(\lambda, T)$: Дж. Релей (1900 г.) для длинных волн и Дж. Джинс (1905 г.) для всех длин волн.

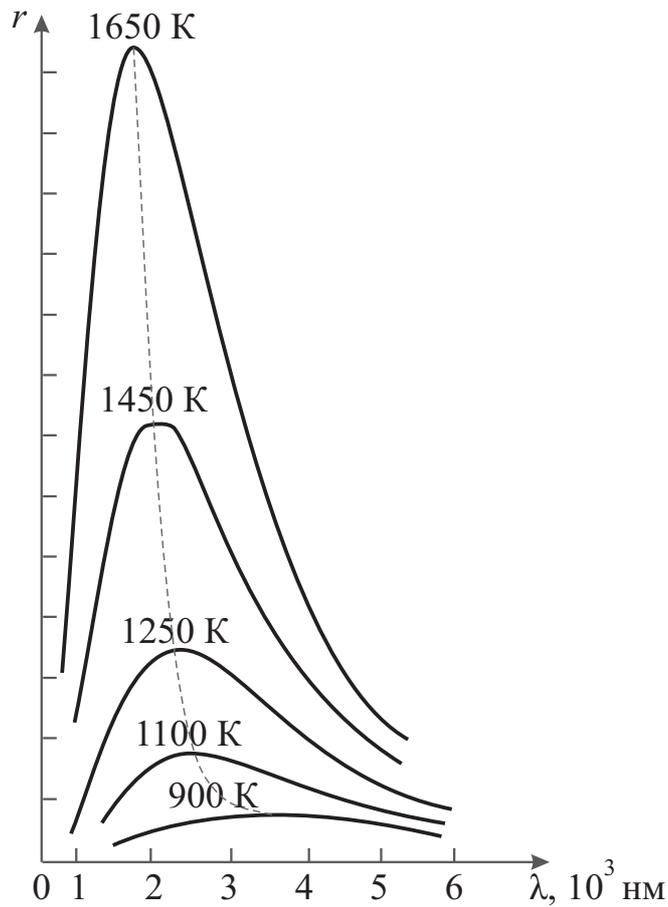


Рис. 2.1 – Спектральное распределение $r(\lambda, T)$ при различных T

Они получили закон излучения абсолютно черного тела соответственно названный их именами, закон Релея – Джинса:

$$r(\lambda, T) = 8\pi k T \lambda^{-4}. \quad (2.2)$$



Однако этот закон хорошо соответствовал эксперименту только в области больших длин волн (рис. 2.2). Оказалось, что классическая физика не способна дать достоверные теории, объясняющие экспериментальные данные по излучению абсолютно черного тела.

Немецкий физик Макс Планк был характерным воспитанником школы классической физики, но именно он совершил переворот в науке и открыл новую физику, способную объяснить многие сложные вещи, в том числе и законы излучения абсолютно черного тела. Классицизм впоследствии еще долго не да-

вал ему покоя в стремлении соединить и согласовать классическую физику с новой физикой. В работе [53] показана логика работы Планка, приведшая его к введению в физику квантования.

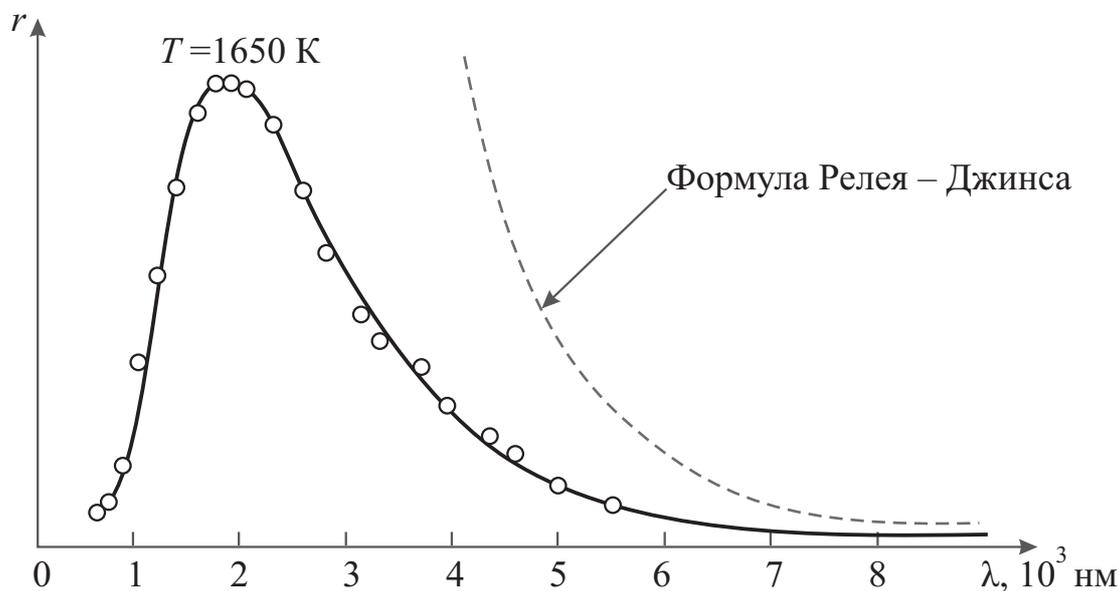


Рис. 2.2 – Сравнение экспериментального распределения $r(\lambda, T)$ с законом Релея – Джинса

Планк, как и его коллеги, долгое время занимался проблемой излучения абсолютно черного тела. В том же 1900 г., когда Релей получил свои результаты, Планк эмпирически получил свою формулу, описывающую закон излучения абсолютно черного тела, которая хорошо согласуется с экспериментальными измерениями как в длинноволновой, так и в коротковолновой областях. Одна из работ Планка так и называлась – «Об одном улучшении закона Вина». Так как $r(\nu, T)$ не зависит от материала стенки полости модели абсолютно черного тела, то можно для этих стенок предположить любое строение, удобное для вычислений. Единственное условие: стенки должны излучать и поглощать радиацию, то есть осуществлять обмен энергией между веществом стенки и средой. Как известно из развития радио, вибратор Герца излучает на своей частоте и может действовать приемником при поглощении излучения как резонатор, что и использовал Герц в своих опытах для приема и передачи радиоволн.

Естественно, Планк выбрал простейшее для расчетов строение стенок полости абсолютно черного тела в виде ансамбля вибраторов Герца (осцилляторов), каждый из которых имел свою частоту и действовал при поглощении излучения как резонатор. Совершенно согласно методам классической электродинамики, не вводя никаких противоречащих или дополняющих ее предположений, Планк получил для интервала частот $d\nu$ связь интенсивности излучения r и энергии вибратора U :

$$r = \frac{8\pi\nu^2}{c^3} U d\nu. \quad (2.3)$$

Планк очень ясно формулировал недостаток понятия температуры и преимущество энтропии: в то время как о температуре можно говорить только в случае состояния равновесия, понятие энтропии в статистической термодинамике может быть определено для всевозможных процессов. Планк подозревал о существовании глубокой связи между энтропией и энергией. Планк обращается к понятию энтропии S и ее связи со статистической вероятностью и вычисляет энтропию резонирующих осцилляторов. Статистическая вероятностная концепция энтропии требует ее логарифмической связи с энергией осциллятора U . Это и получает Планк: так как из эксперимента и закона Релея (2.2) ясно, что интенсивность излучения пропорциональна T , то и энергия осциллятора пропорциональна T , ведь энергия U пропорциональна интенсивности (2.3). Тогда $U = cT$, а $T = U/c$. Известно, что для энтропии $dS = \frac{dU}{T}$, тогда $dS = c \frac{dU}{U}$, отсюда, интегрируя, получается $S = c \ln U$. Следовательно $\frac{d^2S}{dU^2} = -\frac{c}{U^2}$ для больших значений U . А для малых U справедливо другое равенство $\frac{d^2S}{dU^2} = -\frac{1}{\beta U}$. Из этих крайних случаев получаются законы Вина и Релея. Оба крайних следствия реализуются как граничные случаи в общем соотношении, которое и нашел Планк:

$$\frac{d^2S}{dU^2} = \frac{1}{U \left(\beta + \frac{U}{c} \right)}. \quad (2.4)$$

Планк указывает, что это выражение является простейшим из всех, которые дают S как логарифмическую функцию U (что и требуется) и при малых значениях U приводят к указанному выше выражению закона Вина. А если положить $\beta = 0$, то (2.4) увязывается и с законом Релея.



Из общего выражения (2.4) Планк получает свою знаменитую эмпирическую формулу для спектральной плотности (формулу Планка), которая зависит от двух констант и является наипростейшей с точки зрения электромагнитной теории излучения:

$$r(\lambda, T) = \frac{C\lambda^{-5}}{e^{c/\lambda T} - 1}. \quad (2.5)$$

Далее, чтобы теоретически обосновать выражение (2.5) Планк берет выражение для энтропии, несколько отличное от прежнего, приведшего к формуле Вина, и тут ему пришлось воспользоваться статистикой Больцмана. Метод Планка базируется на законах электромагнитного излучения, термодинамики и теории вероятностей. Это новый, совершенно элементарный подход, позволяющий получить без какого-либо знания о виде спектральной формулы и без какой-либо теории распределение некоторой заданной энергии по длинам волн нормального спектра, причем соответствующая связь будет выражена с помощью одной мировой постоянной (h), в то время как другая мировая постоянная (k) свяжет величину температуры и данную энергию излучения.

Планк берет систему из очень большого числа N одинаковых осцилляторов. Затем вычисляет вероятность W того, что эта система обладает заданной энергией $E = U_N$. При этом по статистике Больцмана энтропия связана с вероятностью выражением:

$$S = k \ln W. \quad (2.6)$$

Но вероятность обладания системой энергией U_N можно найти только путем подсчета по выборке *элементов* энергий, составляющих U_N . Поэтому энергию U_N нужно рассматривать как сумму **дискретных**, равных друг другу **элементов энергии** отдельного вибратора (осциллятора) ε с частотой ν . Количество элементов ε , для статистики, должно представляться очень большим числом P . Тогда

$$E = U_N = N \cdot U = P \cdot \varepsilon.$$

Если бы энергия E делилась бы между резонаторами как безгранично делимое количество, то такое распределение возможно провести бесконечным числом способов. Чтобы этого избежать, и это самое важное здесь у Планка, надо чтобы энергия E делилась на точно определенное число равных частей. Чтобы это обеспечить, Планк вводит универсальную мировую постоянную $h = 6,55 \cdot 10^{-27}$ эрг·сек (современное значение $h = 6,63 \cdot 10^{-27}$ эрг·сек), которая, умноженная на частоту осцилляторов ν , дает элемент энергии ε в эргах:

$$\varepsilon = h \cdot \nu. \quad (2.7)$$

Тогда энергия системы разделится на точное число равных частей на каждый осциллятор: $E / \varepsilon = P$, где P – число элементов энергии, распределенных между N резонаторами (осцилляторами). Если указанное отношение не равно целому числу, берем для P ближайшее целое значение.

Далее Планк выводит спектральное распределение энергии в виде:

$$r(\nu, T) = \frac{8\pi\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1}, \quad (2.8)$$

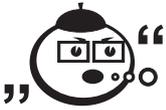
которое в точности соответствует спектральной формуле (2.5), которую можно теперь записать в виде:

$$r(\lambda, T) = \frac{8\pi ch}{\lambda^5} \frac{1}{e^{ch/k\lambda T} - 1}. \quad (2.9)$$

В формулу Планка, написанную в любом виде, входят две универсальные постоянные h и k . Постоянная k обычно называется постоянной Больцмана, так как она является также коэффициентом пропорциональности в формуле Больцмана (2.6), связывающей энтропию и вероятность. Планк, однако, обращает

внимание на то, что сам Больцман никогда этой постоянной не пользовался (он пользовался универсальной газовой постоянной R), и численное значение k впервые вычислил Планк.

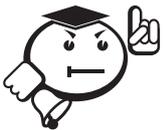
Здесь хочется процитировать Планка.



.....

«Если это соотношение (здесь это (2.6). – П. Д.) действительно должно иметь общее значение, то так как энтропия есть аддитивная величина, а вероятность – мультипликативная, постоянная k должна быть универсальной величиной, зависящей только от выбора единиц измерения. Ее нередко, по понятным причинам, называют постоянной Больцмана. К этому, конечно, надо заметить, что Больцман эту постоянную никогда не вводил и, как мне известно, вообще не думал интересоваться ее численным значением. Ибо тогда ему пришлось бы выяснять число действительных атомов – задача, которую он полностью предоставил своему коллеге Й. Лошмидту, в то время как сам он при своих вычислениях всегда придерживался взгляда, что кинетическая теория газов представляет собой только механическую картину. Поэтому его вполне устраивало иметь дело с грамм-атомами. Буквой k стали пользоваться лишь постепенно. Еще долгие годы после ее введения было принято вести подсчеты, пользуясь вместо нее числом Лошмидта L , обозначающим число атомов, соответствующее одному грамм-атому» [52, с. 440].

.....



В современном виде формулу Планка для спектральной светимости абсолютно черного тела записывают в виде:

$$r(\nu, T) = \frac{2\nu^2}{c^2} \frac{h\nu}{e^{h\nu/kT} - 1}, \quad (2.10)$$

где c – скорость света, h – постоянная Планка, k – постоянная Больцмана, T – абсолютная температура. Формула Планка хорошо опи-

сывает спектральное распределение излучения черного тела при любых частотах. Она прекрасно согласуется с экспериментальными данными. Из формулы Планка можно вывести закон Вина. При $h\nu \ll kT$ формула Планка переходит в формулу Релея – Джинса.

.....

В отличие от своих коллег, считавших, что излучение и поглощение происходит непрерывно, Планк пришел к выводу о дискретности этих процессов, излучение и поглощение энергии происходит не непрерывно, а определенными порциями – квантами – и это не любая величина, а именно $\varepsilon = h\nu$, где h – определенная константа (постоянная Планка), а ν – частота света. Постоянную h еще называют «элементарный квант действия», открытие которого увековечило имя Планка. Его работы привели к признанию, наравне с атомизмом вещества, атомизма энергии или действия, дискретного, квантового характера излучения.

2.2 Работы Эйнштейна

Хотя первые исследователи (А. Беккерель, Г. Герц, А. Г. Столетов) и сумели экспериментально установить, что фотоны могут вызвать эмиссию электронов из некоторых материалов, до М. Планка и А. Эйнштейна не существовало ответа на загадочный вопрос: почему, если число полученных электронов пропорционально интенсивности света, их энергия пропорциональна длине волны света?

Квантовая гипотеза Планка стала основой теории, объясняющей поведение электронов в твердых телах. На ее основе Эйнштейн смог объяснить в 1905 г. фотоэлектрический эффект, впервые открытый Беккерелем [54]: свет, падающий на полупроводник, вызывает появление свободных электронов. В 1905 г. в немецком журнале «Annalen der Physik» появились три статьи 26-летнего Эйнштейна – о теории относительности, квантовой теории фотоэффекта [55] и теории броуновского движения, – которые обозначили поворот в развитии физики. Главным в работе Эйнштейна [55, 56] является гипотеза квантов света – фотонов и ее доказательство. Во введении Эйнштейн пишет:

«Я и в самом деле думаю, что опыты, касающиеся «излучения черного тела»... и других групп явлений, связанных с возникновением и превращением света, лучше объясняются предположением, что энергия света распределяется по пространству дискретно» [56, с. 93]. Сразу отметим, что работы Эйнштейна по теории излучения не явились непосредственным продолжением работ Планка, Эйнштейн только опирается на планковскую гипотезу квантов энергии монохроматического излучения черного тела. Эйнштейн показывает, что



.....
«определение элементарных квантов, данное Планком, является до известной степени независимым от созданной им теории «излучения черного тела» [56, с. 95].

Он пишет:



.....
«В дальнейшем «излучение черного тела» будет рассматриваться в связи с опытом, а не на основе каких-либо представлений о возникновении и распространении излучения» [56, с. 96].

Далее он на работы Планка больше не опирается и не ссылается, проводит все теоретические выкладки самостоятельно. Подход Эйнштейна коренным образом отличался от подхода Планка. В своей работе Эйнштейн оперирует введенными им константами α , β , универсальной постоянной R и другими, связанными с h , ν и k . Мы физические величины и соотношения будем приводить в современном виде, в терминах $h\nu$ и kT . На основе статистического рассмотрения законов равновесного излучения Эйнштейн делает заключение: «Монохроматическое излучение... ведет себя так, как будто оно состоит из независимых друг от друга квантов энергии величиной $h\nu$ » [56, с. 102], подтверждая достижение Планка. А величина $E/h\nu$ представляет число квантов света. Эйнштейн также вычисляет среднюю энергию квантов энергии излучения черного тела $3kT$ и сравнивает ее со средней кинетической тепловой энергией молекул $(3/2)kT$.

Эйнштейн говорит:



.....

«Но если монохроматическое излучение... ведет себя как дискретная среда, состоящая из квантов энергии величиной $h\nu$, то напрашивается вопрос, не являются ли и законы возникновения и превращения света такими, как будто свет состоит из подобных же квантов энергии» [56, с. 102].

.....

Далее Эйнштейн переходит от равновесного излучения к исследованию процессов взаимодействия излучения с веществом, основанному на представлении о дискретности энергии излучения, и соответственно, дискретности самих процессов поглощения и испускания.

Эйнштейн рассматривает фотоэффект на основе того, «что возбуждающий свет состоит из квантов с энергией $h\nu$ » [56, с. 104], предполагает, что «один световой квант отдает всю свою энергию одному электрону» [56, с. 104], и учитывает, «что каждый электрон, покидая тело, должен совершить некоторую работу P (характерную для данного тела)» [56, с. 104], т. е. вводит понятие о работе выхода. Приравнивая максимальную кинетическую энергию вырываемого электрона $h\nu - P$ к величине Πe , где Π – величина задерживающего потенциала, а e – заряд электрона, Эйнштейн получает свое знаменитое уравнение для фотоэффекта в виде (в современной записи)

$$eV = h\nu - P. \quad (2.11)$$

В 1839 г. французский ученый Александр Эдмон Беккерель впервые случайно обнаружил влияние света на увеличение тока в электрической цепи [54]. А. Э. Беккерель экспериментировал с электролитическим элементом из двух электродов, помещенных в электропроводящий водный раствор азотной кислоты. Один был изготовлен из листовой платины и покрыт тонким слоем хлорида серебра AgCl – на этот электрод и падал свет, вызывая фотоэффект. После Беккереля, что известно более широкой аудитории, опыты по фотоэффекту проводил Герц, а российский ученый А. Г. Столетов исследовал фотоэффект и ре-

зультаты опытов обобщил в следующем законе: 1) число электронов, освобожденных светом за единицу времени, пропорционально интенсивности (энергии) светового потока. Дальнейшее исследование позволило установить второй закон: 2) максимальная кинетическая энергия фотоэлектронов пропорциональна частоте поглощаемого света и не зависит от его интенсивности.

Уравнение Эйнштейна (2.11) по существу – это закон сохранения энергии: $h\nu = eV + P$ – энергия фотона тратится на работу выхода электрона P и на передачу электрону кинетической энергии eV . Для начала фотоэффекта нужен свет с минимальной частотой ν_0 и энергией

$$h\nu_0 = P, \quad (2.12)$$

ν_0 – минимальная частота света, называемая *красной границей* фотоэффекта.

Если Планк предложил считать, что свет *излучается* квантованными порциями, то Эйнштейн (идею подал Планк) уже предложил считать, что свет *существует* и *состоит* из квантов – фотонов. Тогда первый закон Столетова объясняется квантовой теорией Эйнштейна следующим образом.

Известно, что интенсивность света:

$$I = WSt, \quad (2.13)$$

где W – энергия падающего света, S – площадь поверхности, на которую падает свет, t – время. Энергия света определяется произведением энергии кванта (фотона) $h\nu$ и числа фотонов, падающих на вещество N_f : $W = N_f h\nu$. Очевидно, что число электронов N_e , вырванных из вещества, пропорционально числу фотонов, падающих на вещество, т. е. $N_e \sim N_f$, а следовательно из (2.13) $N_e \sim I$. Это квантовое доказательство первого закона фотоэффекта: число фотоэлектронов пропорционально интенсивности света.

Второй закон фотоэффекта доказывается уравнение Эйнштейна (2.11): энергия фотоэлектронов пропорциональна частоте (длине волны) света и не зависит от интенсивности света. Ну и третий закон фотоэффекта: красная граница фотоэффекта (2.12) зависит от рода вещества, так как работа выхода P у каждого вещества своя.

Итак, по Эйнштейну, свет не только излучается и поглощается, но и состоит из дискретных, неделимых порций, квантов света. Чем ярче свет, тем больше фотонов и больше число выбитых с поверхности вещества электронов, но не их скорость. Более быстрые фотоэлектроны можно получить, направляя излучение с большей частотой ν , так как фотоны такого излучения содержат больше энергии.



И еще одно важное предположение выдвинул Эйнштейн о том, что свет имеет двойственную природу и одновременно является и волной, и потоком частиц. Действительно, со времен Ньютона известна оптика и ее законы дифракции, интерференции, «угол падения равен углу отражения» и другие волновые законы и свойства света. Но по квантовой теории и экспериментальным законам фотоэффекта свет – это поток частиц, фотонов (квантов света).

2.3 Работы Резерфорда и Бора

Английский физик новозеландского происхождения Эрнест Резерфорд в 1911 г. провел оригинальный опыт, показавший, что положительные α -частицы, создающиеся при излучении радиоактивным материалом, при прохождении их через золотую фольгу отклоняются и на люминесцентном экране после фольги видны вспышки вокруг оси потока α -частиц. Оказалось, что некоторые частицы отклоняются на углы вплоть до 180° (рис. 2.3).

То, что отражало мощную, быструю и тяжелую положительно заряженную α -частицу, должно было иметь невероятно плотную массу, сосредоточенную в центре атома, и весь положительный заряд атома, так как только силами электрического отталкивания возможно отклонение на угол более 90° . Это образование в центре атома Резерфорд назвал ядром, а вокруг ядра – система отрицательных электронов, которые удерживаются притяжением ядра. То, что отражало мощную, быструю и тяжелую положительно заряженную α -частицу,

должно было иметь невероятно плотную массу, сосредоточенную в центре атома, и весь положительный заряд атома, так как только силами электрического отталкивания возможно отклонение на угол более 90° . Это образование в центре атома Резерфорд назвал ядром, а вокруг ядра – система отрицательных электронов, которые удерживаются притяжением ядра.

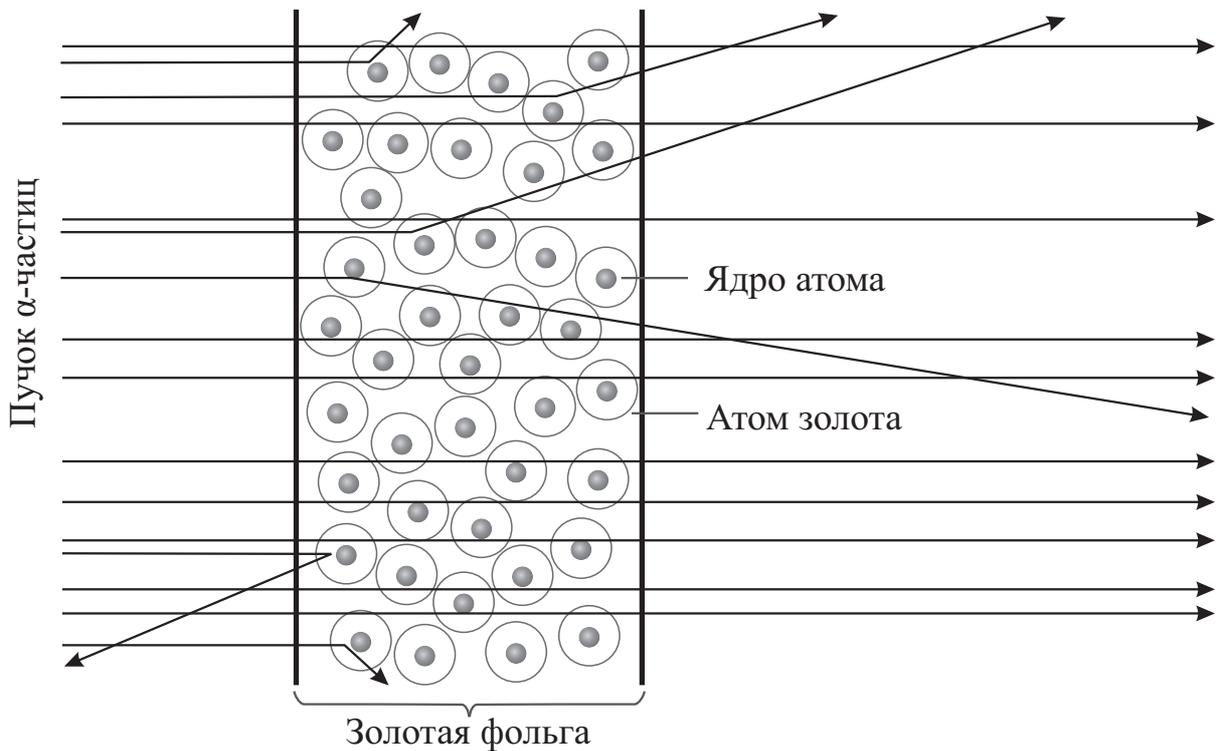


Рис. 2.3 – Иллюстрация опыта Резерфорда

Общий отрицательный заряд электронов равен положительному заряду ядра, линейные размеры ядра исключительно малы по сравнению с линейными размерами всего атома. Число электронов в атоме приблизительно равно половине атомного веса. Эксперименты Г. Гейгера в 1913 г. и Дж. Чэдваика в 1920 г. подтвердили теорию атома Резерфорда и показали, что заряд ядра атома равен Ze , где Z – атомный номер элемента, а e – заряд электрона.

В 1913 г. датский физик Нильс Бор, отдавая должное экспериментальной основе модели атома Резерфорда, указал на ее противоречивость: модель, согласно классическим законам, должна была быть неустойчивой из-за потерь энергии на излучение ускоренно движущимися электронами.



Заряженный электрон вращается вокруг ядра с большим центростремительным ускорением, значит, он непрерывно должен испускать свет, вследствие чего его энергия должна уменьшаться и электрон все ближе должен приближаться к ядру, пока не упадет на него. В действительности ничего такого нет, атомы свет не испускают.

Бор пишет:



«Обсуждение этого вопроса приводит к выводу, что классическая электродинамика, очевидно, неприменима для описания поведения систем атомных размеров. Что касается законов движения электронов, то представляется необходимым ввести в эти законы чуждую классической электродинамике величину, а именно постоянную Планка, или, как ее часто называют, элементарный квант действия. Если ввести эту величину, то вопрос о стабильных конфигурациях электронов в атомах существенно меняется, так как размерность и величина этой постоянной таковы, что вместе с массой и зарядом частиц она позволяет определить длину [радиус орбиты] нужного порядка» [57, с. 85].

Для объяснения поведения атомных систем модели Резерфорда и в основу своей теории атомных систем Бор положил знаменитые два постулата, получившие его имя.

1. Электрон может находиться только на таких орбитах, для которых момент количества движения электрона (то есть произведение количества движения электрона mv на радиус орбиты r , а v – скорость электрона) кратен значению $h/2\pi$. Пока электрон находится на одной из таких орбит, излучения энергии не происходит. Каждой разрешенной орбите электрона соответствует определенная энергия, а атому – определенное энергетическое состояние, назы-

ваемое стационарным. Находясь в стационарном состоянии, атом не излучает света. Математически первый постулат Бора можно записать так:

$$m\omega r = n(h/2\pi), \quad (2.14)$$

где $n = 1, 2, 3, \dots$ – некоторое целое число, называемое *главным квантовым числом*.

2. Поглощение или испускание света атомом происходит только при переходах атома из одного стационарного состояния в другое. При этом энергия поглощается или испускается определенными порциями, квантами, значение которых $h\nu$ определяется разностью энергий, соответствующих начальному и конечному стационарным состояниям атома:

$$h\nu = E_m - E_n,$$

где E_m – энергия начального состояния атома, E_n – энергия конечного состояния атома.

Если $E_m > E_n$, то энергия излучается атомом, если же $E_m < E_n$, то поглощается. Электрон не может изменять свою траекторию постепенно и непрерывно, а может только «перепрыгивать» с одной стационарной орбиты на другую. При переходе со стационарной орбиты, удаленной от ядра, на стационарную орбиту ближе к ядру, происходит излучение света.

Для разрешенных орбит можно найти радиусы из законов классической механики, учитывая, что электроны и ядро – заряженные частицы и сила их взаимодействия определяется электрическим полем и законом Кулона, и применяя первый постулат Бора. Тогда для стационарных состояний системы, состоящей из ядра и одного вращающегося вокруг него электрона:

$$r = \frac{n^2 h^2}{4\pi^2 m e^2}, \quad (2.15)$$

$n = 1$ отвечает наиболее близкой к ядру разрешенной орбите. Из (2.15), применяя экспериментальные значения величин m , e и h , можно вычислить радиус первой орбиты:

$$r_1 = 0,52 \cdot 10^{-10} \text{ м} = 0,052 \cdot 10^{-9} \text{ м} = 0,052 \text{ нм}. \quad (2.16)$$

Радиус первой орбиты электрона в атоме составляет пять сотых долей нанометра ($\frac{5}{100}$ нм). Эта величина берется за радиус атома водорода. Из (2.15)

следует, что любая другая n -я орбита с квантовым числом n имеет радиус:

$$r_n = n^2 r_1. \quad (2.17)$$

Видно, что радиусы последующих орбит электрона увеличиваются как квадрат числа n (рис. 2.4).

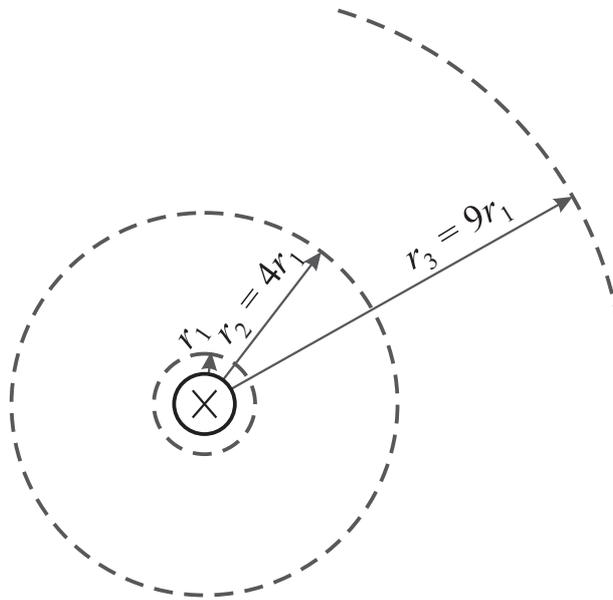


Рис. 2.4 – Радиусы разрешенных орбит в атоме водорода

Разрешенному положению электрона соответствует свое энергетическое состояние атома. В теории Бора полная энергия атома в состоянии, когда электрон находится на n -й орбите, определяется выражением:

$$E_n = -\frac{2\pi^2 m e^4}{n^2 h^2}. \quad (2.18)$$

Эти значения энергий называются *энергетическими уровнями* атома и их можно изобразить в виде графика энергий – это так называемый *энергетический спектр* разрешенных состояний атома (рис. 2.5).

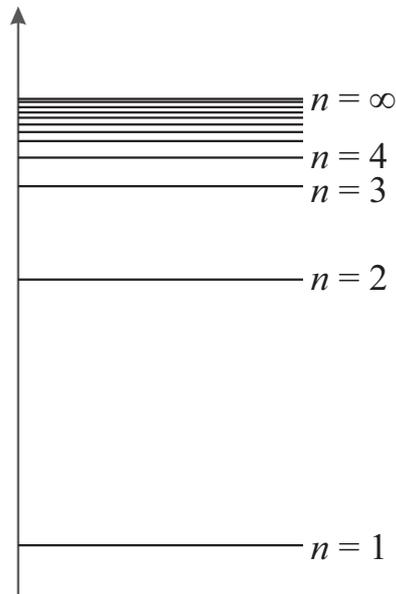


Рис. 2.5 – Энергетический спектр разрешенных состояний

Так же, как увеличиваются по квадрату n радиусы разрешенных орбит, соответствующие им энергии увеличиваются обратно квадрату n . Поэтому расстояние между последовательно расположенными энергетическими уровнями быстро уменьшается: увеличение энергии атома (за счет поглощения атомом энергии извне) сопровождается переходом электрона на все более удаленные орбиты, где взаимодействие ядра и электрона ослабевает. Очень малое изменение энергии нужно для перехода между близкими удаленными орбитами. При этом удаленные уровни так близки, что спектр становится почти непрерывным и вверху заканчивается состоянием $n = \infty$, когда атом ионизуется, а электрон отделяется от ядра и становится свободным. Этому состоянию $n = \infty$ соответствует нулевая энергия атома, а состояниям $n \neq \infty$ – отрицательные энергии (2.18), и чем ближе к ядру, тем больше по модулю выражение (2.18), и тем меньше энергия атома. Уровню $n = 1$ соответствуют наименьшая энергия и наименьший радиус разрешенной орбиты электрона. Это – *основной*, или *невозбужденный*, уровень. Остальные уровни $n = 2, 3, 4, \dots$ называются *уровнями возбуждения*.

Теория Бора – это сочетание классической механики движения электрона по круговой орбите и представлений о квантах излучения. Теория Бора хорошо

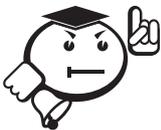
описывала только одноэлектронный атом водорода. В многоэлектронных атомах начиная с двухэлектронного атома гелия, теория Бора не работала. Некоторые улучшения дал следующий шаг, когда орбиту стали представлять не круговой, а эллиптической, когда ядро расположено в одном из фокусов эллипса. Оказалось, что дальнейшее сочетание классики и квантовых представлений себя исчерпало. Нужен кардинальный отказ от понятий и представлений классической механики (невозможно описание движения электронов в атоме в классических образах траектории, орбиты и др.). Стало ясно, что нужна теория, которая оперировала бы величинами, относящимися лишь к начальному и конечному состояниям атома.

2.4 Волны де Бройля и квантово-волновой дуализм

частиц вещества



Одним из еще более радикальных теоретических принципов стала гипотеза французского физика Луи де Бройля (1923 г.) о том, что корпускулярно-волновой дуализм носит всеобщий характер, волновые свойства присущи любым частицам материи, не только кванту энергии – фотону, но и электрону, протону и др. материальным частицам.



Де Бройль предположил [58], что всякая частица связана с каким-то периодическим процессом, а всякая движущаяся частица, будь то квант света или электрон, связана с особыми «фазовыми волнами», частота которых определяется энергией частицы, а длину волны которых удобнее всего определить из условия, что количество движения частицы mv равно количеству движения какого-то «кванта», соответствующего этим волнам.

Энергия кванта равна $h\nu$, скорость света связана с частотой и длиной волны как $c = \lambda\nu$, тогда

$$m\nu = p = \frac{h\nu}{c} = \frac{h}{\lambda}. \quad (2.19)$$

Откуда

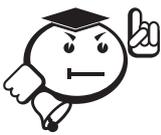
$$\lambda = \frac{h}{m\nu} = \frac{h}{p}. \quad (2.20)$$

Сорт квантов, то есть кванты с данной длиной волны, обычно принято характеризовать не длиной волны, а величиной, которая называется «волновым вектором». Эта величина просто связана с длиной волны $k = \frac{2\pi}{\lambda}$.

Скорость распространения этих фазовых волн определяется формулой:

$$\omega = \frac{c^2}{v} = \frac{c}{\beta}. \quad (2.21)$$

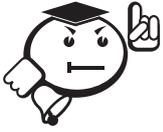
Длина волны связана с периодом T соотношением $\lambda = cT$, а период связан с частотой ω по формуле $T = \frac{2\pi}{\omega}$. Частота ω всегда больше скорости света и только групповая скорость фазовых волн $\omega' = v$ оказывается равной скорости самой частицы. Так де Бройль разрешает противоречие со скоростями: фазовая скорость «волн материи» больше скорости света, но групповая – меньше.



По де Бройлю, фазовые волны, связанные с движущимся электроном, настолько же реальны, насколько реальны обычные световые волны, «связанные» с квантами света, наблюдаемыми нами при всех явлениях поглощения или иного действия света.

Вместо электрона, движущегося по орбите, мы должны представить себе группу волн, движущихся по этой орбите. Чтобы эти волны не уничтожали друг друга вследствие интерференции, они должны двигаться так, чтобы в орбите создалась стационарная, или стоячая, волна. Необходимое условие для этого: длина пути должна равняться целому числу длин волн $2\pi r = n\lambda$, или,

подставив длину волны в виде (2.20), получим $2\pi r = \frac{nh}{m\upsilon}$, или $mrv = \frac{nh}{2\pi}$, что совпадает с выражением (2.14) теории Бора.



.....

Де Бройль, восприняв идею Эйнштейна о двойственной природе света, распространил ее на вещество, предсказав, что поток материальных (вещественных) частиц должен одновременно обладать волновыми свойствами, однозначно связанными с массой и энергией частиц по выражению (2.20).

.....

Эта идея была высказана де Бройлем впервые в 1923 г., а 1924 г. он развил ее в ряде статей, и в знаменитой докторской диссертации «Recherche sur la theorie des quanta», которую защитил в ноябре 1924 г. К новым идеям в этой диссертации с большим вниманием отнеслись П. Ланжевен, А. Эйнштейн, М. Борн и другие крупные физики того времени. В 1925 г. К. Дэвиссон прислал Борну необычные результаты своих измерений по отражению электронов от кристаллических поверхностей: в определенных направлениях наблюдались максимумы отраженных электронов (рис. 2.6). М. Борн и Дж. Франк объяснили кривые Дэвиссона как результат интерференции электронных волн де Бройля. В 1927 г. К. Дэвиссон и Л. Джермер опубликовали результаты систематически проведенных экспериментов, окончательно подтвердивших справедливость идеи де Бройля и правильность выражения (2.20) [59].

Центральным затруднением квантовой теории был конфликт между волнами и частицами. С одной стороны, действуют законы сохранения материи, энергии и т. д., которые говорят нам, что материя сохраняется, и наделяют каждую частицу энергии индивидуальностью, так что мы можем проследить ее историю. С другой стороны, мы имеем теоремы интерференции света, а теперь и материи, которые так же определенно говорят нам, что те вещи, которые мы раньше считали частицами, должны расплываться, теряя при этом свою индивидуальность.

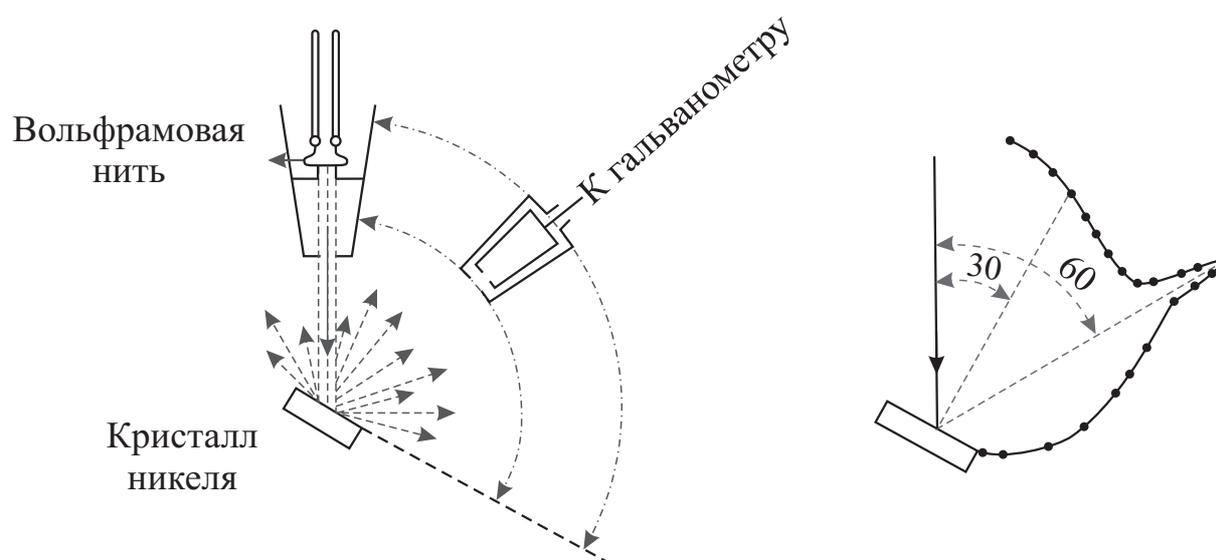


Рис. 2.6 – Экспериментальная установка для исследования рассеяния электронов кристаллом и типичная кривая, показывающая пучок правильно отраженных электронов. Бомбардирующий потенциал 83 вольт; угол падения пучка электронов 30 градусов [59]

2.5 Волновая механика и уравнение Шредингера

В 1925 г. Эрвин Шредингер, развивая идеи де Бройля, построил так называемую волновую механику, в основе которой представление о том, что квантовые процессы следует понимать как некие волновые процессы, характеризующиеся волновой функцией Ψ , которая определяет однозначность и конечность в пространстве и времени. Действительно, по де Бройлю, электрон, движущийся по орбите, мы должны представить как группу волн, движущихся по орбите, и они должны двигаться так, чтобы в орбите создалась стационарная, или стоячая, волна. Там, где укладывается целое число длин волн электрона, волны складываются и образуют боровские разрешенные орбиты. А там, где целое число длин волн не укладывается, гребни волн компенсируют впадины и орбиты не будут разрешены. Это также означает, что образ электрона как материальной точки, занимающей определенное место в пространстве – орбиту, является очень приближенным.

Уместно вспомнить про волновое уравнение: уравнение любой волны есть решение некоторого дифференциального уравнения, называемого волновым, которое в декартовых координатах имеет вид:

$$\frac{d^2U}{dx^2} + \frac{d^2U}{dy^2} + \frac{d^2U}{dz^2} = \frac{1}{\omega^2} \frac{d^2U}{dt^2}, \quad (2.22)$$

где ω – скорость распространения волны, t – время, а U – волновая функция, характеризующая смещение точки, совершающей колебание, от положения равновесия. Например, в случае электромагнитных волн U может представлять собой электрический или магнитный вектор волны, распространяющейся в пространстве.

Отрицательный электрон притягивается к положительному ядру, и эта сила не дает ему уйти от ядра: как только электрон за счет внешней силы начнет отходить от ядра, сила притяжения вернет его назад. В силу инерционности электрон будет совершать колебательное движение, и его можно рассматривать как стоячую волну, аналогичную волне при колебании закрепленной в двух точках струны.

Шредингер использует квантовые соотношения (2.7), (2.19) и (2.20) и, полагая, что $\omega \ll c$, а кинетическая энергия частицы – это разность между полной энергией \mathcal{E} и потенциальной энергией V , записывает $m\upsilon^2 = 2(\mathcal{E} - V)$. Тогда

$$\upsilon = \sqrt{\frac{2(\mathcal{E} - V)}{m}}, \text{ а импульс } p = m\upsilon = m\sqrt{\frac{2(\mathcal{E} - V)}{m}} = \sqrt{2m(\mathcal{E} - V)}, \text{ так что}$$

$$\omega = \frac{h\nu}{p} = \frac{h\nu}{\sqrt{2m(\mathcal{E} - V)}}. \quad (2.23)$$

Вообще-то физический смысл имеют только волновые функции, которые являются гармоническими функциями времени, то есть выражаются через $\sin 2\pi\nu t$ или $\cos 2\pi\nu t$, или через их комбинации в виде экспоненты $\exp(i2\pi\nu t) = e^{i2\pi\nu t}$, где i – мнимая единица, $i = \sqrt{-1}$.

Берем волновую функцию в виде $U = ue^{i2\pi\nu t}$, причем амплитуда u зависит только от координат, а от времени не зависит и подставляем. Тогда

$$\frac{d^2U}{dt^2} = -4\pi^2\nu^2ue^{i2\pi\nu t} = -4\pi^2\nu^2U. \quad (2.24)$$

Подставим (2.23) и (2.24) в волновое уравнение (2.22) – получим уравнение Шредингера для отдельной частицы:

$$\frac{d^2U}{dx^2} + \frac{d^2U}{dy^2} + \frac{d^2U}{dz^2} = -\frac{8\pi^2m}{h^2}(\mathcal{E} - V)U. \quad (2.25)$$

Сократим каждый член на экспоненту $e^{i2\pi\nu t}$ – получим уравнение для амплитуды, куда время уже не входит:

$$\frac{d^2u}{dx^2} + \frac{d^2u}{dy^2} + \frac{d^2u}{dz^2} = -\frac{8\pi^2m}{h^2}(\mathcal{E} - V)u. \quad (2.26)$$

Положительный заряд ядра будет создавать сферически симметричное электростатическое поле, значит, в уравнении Шредингера надо использовать не прямоугольные декартовы координаты, а сферические: r – радиус-вектор полярных координат – расстояние точки до начала координат, θ и φ – зенитный и азимутальный углы соответственно. При этом обозначим волновую функцию $\Psi(r, \theta, \varphi)$, а потенциальная энергия взаимодействия зарядов ядра и электрона равна $V = -\frac{Ze^2}{r}$. В сферических координатах уравнение Шредингера имеет вид [60]:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Psi}{dr} \right) + \frac{1}{r^2 \sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d\Psi}{d\theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{d^2\Psi}{d\varphi^2} + \frac{2m}{\hbar^2} \left(\mathcal{E} + \frac{Ze^2}{r} \right) \Psi = 0, \quad (2.27)$$

где r – расстояние от электрона до ядра; $\hbar = \frac{h}{2\pi} = 1,0546 \cdot 10^{-34}$ Дж·с.

Решение уравнения (2.27) делается разделением переменных, поэтому искомую волновую функцию представляют в виде композиции трех функций $\Psi(r, \theta, \varphi) = R(r) \cdot \Theta(\theta) \cdot \Phi(\varphi)$. В результате получают три уравнения с одной переменной r , θ и φ . Разделение переменных и решение уравнений возможны, только если ввести три целых величины, которые называют *квантовыми числами*: n , l и m , от которых зависят решения: $\Psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) \cdot \Theta_{lm}(\theta) \cdot \Phi_m(\varphi)$.

Квантовые числа n , l и m описывают состояния электрона в атоме, а волновые функции называют *атомными орбиталями* (АО).

Слово «орбиталь» происходит от латинского «орбита» (путь, колея), но имеет другой смысл, не совпадающий с понятием траектории (пути) электрона вокруг атома, предложенным Бором для планетарной модели атома. Контуры атомной орбитали – это графическое отображение волновой функции, полученной при решении волнового уравнения для одного электрона.

Всякая атомная орбиталь имеет строго определенную комбинацию значений n , l и m и соответствующие ей энергию, форму и ориентацию в пространстве [60].

Функции $\Theta(\theta)$ и $\Phi(\varphi)$ от расстояния не зависят, а только от углов, поэтому их объединяют (после нахождения из решения уравнения Шредингера) в угловые функции $Y(\theta, \varphi) = \Theta(\theta) \Phi(\varphi)$ – это сферические гармоники и они полностью определяют угловую зависимость волновой функции электрона. Их еще называют *угловой составляющей волновой функции*. Тогда решения уравнения Шредингера имеют вид $\Psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) \cdot Y_{lm}(\theta, \varphi)$ [60].

Атомные орбитали обозначают, в зависимости от величины квантового числа l , малыми буквами латинского алфавита: s ($l = 0$), p ($l = 1$), d ($l = 2$), f ($l = 3$) по начальным буквам серий линий, присутствующих в оптических спектрах атомов: *sharp* (резкая), *principal* (главная), *diffuse* (диффузная) и *fundamental* (фундаментальная).



.....

Физический смысл имеет не сама волновая функция, т. е. атомная орбиталь, а квадрат волновой функции, он описывает пространственное распределение плотности вероятности нахождения электрона около ядра – плотности отрицательного заряда в околоядерном пространстве.

.....

В такой трактовке квадраты угловых функций Y в сферических координатах описывают угловую зависимость вероятности нахождения электрона в цен-

тральном поле атома, а сам электрон рассматривается как некоторое пространственное образование – электронное облако. Размеры электронных облаков выбирают исходя из того соображения, чтобы их граничная поверхность охватывала от 0,9 до 0,99 вероятности нахождения электрона.

Главное – это понять физический смысл волновой механики [61].

Само понятие орбиты, или траектории движения электрона в атоме, лишено физического смысла. Никакой определенной последовательности появления электрона в различных точках пространства установить нельзя, он оказывается «размазанным» в области, называемой *электронным облаком*. Облако это для невозбужденного атома водорода имеет форму шара, но плотность его не одинакова. Вероятность обнаружения электрона будет наибольшей вблизи сферы с радиусом r_1 , соответствующим радиусу первой боровской орбиты. Под орбитой электрона в атоме будем понимать геометрическое место точек, которые характеризуются наибольшей вероятностью обнаружения электрона, или, другими словами, область пространства с наибольшей плотностью электронного облака.

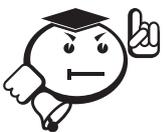
Всегда сферическим электронное облако будет лишь для случая невозбужденного состояния атома водорода, когда главное квантовое число $n = 1$. Если же $n = 2$, то помимо сферического облака, размеры которого будут теперь в четыре раза больше, электрон может создать облако в виде гантели. С появлением несферичности области преимущественной локализации электрона (электронного облака) связано введение второго квантового числа l , называемого *орбитальным квантовым числом*. Каждому значению главного квантового числа n соответствуют положительные целочисленные значения квантового числа l от нуля до $n - 1$: $l = 0, 1, 2, \dots (n - 1)$.

Если $n = 1$, то l имеет единственное значение, равное нулю. Если же $n = 3$, то l может принимать значения 0, 1, 2. При $n = 1$ имеется только сферическая орбита, поэтому и $l = 0$. Когда $n = 2$, возможны как сферическая, так и гантелеобразные орбиты, поэтому и l может быть равным либо нулю, либо единице.

Если $n = 3$, то $l = 0, 1, 2$. Электронное облако, соответствующее значению $l = 2$, приобретает уже довольно сложный характер. Для нас, однако, важна не форма электронного облака, а то, какая ему соответствует энергия атома.

Энергия атома водорода определяется только значением главного квантового числа n и не зависит от значения орбитального числа l . Иначе говоря, если $n = 3$, то атом будет иметь определенную энергию E_3 независимо от того, на какой из возможных орбит, соответствующих данному значению n и различным возможным значениям l , находится электрон. Это означает, что при возвращении с уровня возбуждения на основной уровень атом будет испускать фотоны, энергия которых не зависит от значения l .

Электронные облака имеют строго определенную ориентацию, которая задается магнитным квантовым числом m , которое может принимать целочисленные значения от $-l$ до $+l$, включая 0. При данной форме (данном значении l) электронное облако может иметь несколько различных ориентаций в пространстве. При $l = 1$ их будет три, соответствующих значениям магнитного квантового числа m , равным $-1, 0$ и $+1$. Если $l = 2$, то различных ориентаций электронного облака будет 5, соответствующих значениям $m = -2, -1, 0, +1$ и $+2$. Естественно, что если уж форма электронного облака в свободном атоме водорода не влияет на энергию атома, то тем более не влияет на энергию атома ориентация этого облака в пространстве.



.....

Экспериментально выяснено, что сами электроны могут находиться на орбитах в двух возможных состояниях, определяемых направлением так называемого *спина электрона*.

.....

В 1925 г. английские физики Дж. Уленбек и С. Гоудсмит для объяснения результатов спектральных экспериментов предложили гипотезу, согласно которой каждый электрон вращается вокруг своей собственной оси подобно волчку или веретену. При таком вращении электрон приобретает некоторый момент импульса, который и получил название *спин* (в переводе с английского *спин*

означает «вращение, веретено»). Поскольку вращение может происходить по часовой стрелке или против, то спин (иначе говоря, вектор момента импульса) может иметь два направления. В единицах \hbar спин равен $1/2$, а благодаря различным направлениям имеет знак «+» или «-». Таким образом, ориентация электрона на орбите определяется *спиновым квантовым числом* σ , равным $\pm 1/2$. Отметим, что и ориентация спина, как и ориентация орбиты электрона, не влияет на энергию атома водорода, находящегося в свободном состоянии.

Но и со спином природа оказалась еще сложнее. Более поздние исследования и расчеты показали, что объяснить спин электрона простым вращением его вокруг оси нельзя. При подсчете угловой скорости вращения электрона выяснилось, что линейная скорость точек, лежащих на экваторе электрона (если считать, что электрон имеет шарообразную форму), должна быть больше скорости света, чего не может быть. Полагать, что электрон имеет форму шара нельзя. В любом случае, спин является некоторой неотъемлемой характеристикой электрона, такой, например, как его масса или заряд.



.....

Квантовые числа задают адрес электрона в атоме. Для определения состояния электрона в атоме, то есть для описания движения электрона в атоме, необходимо задать набор из четырех квантовых чисел: n , l , m и σ .

.....

Главное квантовое число n определяет размеры электронной орбиты. Чем больше n , тем большее пространство охватывает соответствующее электронное облако, n может принимать любые целочисленные значения от 1 до ∞ : $n = 1, 2, 3, \dots$.

Орбитальное число l определяет форму электронного облака и выделяет орбиты, имеющие одинаковую форму, из всей совокупности орбит, относящихся к одному и тому же значению n . Каждому значению l соответствует своя подболочка. Число подболочек равно n , так как l может принимать значения от 0 до $(n-1)$: $l = 0, 1, 2, \dots, (n-1)$.

Магнитное квантовое число m определяет пространственную ориентацию орбиты в группе орбит, имеющих одинаковую форму, то есть относящихся к одной подоболочке. В каждой подоболочке насчитывается $(2l + 1)$ различно ориентированных орбит, поскольку m может принимать значения от 0 до $\pm l$:

$$m = -l, -(l-1), \dots, -1, 0, +1, \dots, +(l-1), +l.$$

Наконец, спиновое квантовое число σ определяет ориентацию спина электрона на заданной орбите. Значений у σ всего два: $\sigma = \pm 1/2$.

Рассмотренную взаимосвязь допустимых значений квантовых чисел удобно представить в виде таблицы 2.1 [60].

Таблица 2.1 – Наборы значений квантовых чисел для различных атомных орбиталей

n	l	m	Число АО	Обозначение АО
1	0	0	Одна	1s
2	0	0	Одна	2s
	1	-1, 0, 1	Три	2p
3	0	0	Одна	3s
	1	-1, 0, 1	Три	3p
	2	-2, -1, 0, 1, 2	Пять	3d
4	0	0	Одна	4s
	1	-1, 0, 1	Одна	4p
	2	-2, -1, 0, 1, 2	Три	4d
	3	-3, -2, -1, 0, 1, 2, 3	Пять	4f
	0	0	Одна	4s

Графическое изображение полной волновой функции:

$$\Psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) \cdot Y_{lm}(\theta, \varphi),$$

точнее говоря, ее модуля или квадрата модуля в трехмерном пространстве невозможно. Ее же угловую зависимость визуализируют различными способа-

ми [60]. В одном из них используют трехмерное изображение (аналогичное полярному графику), полагая радиальную функцию постоянной, то есть $R(r) = \text{const}$, и умножая ее на значения $Y(\theta, \varphi)$, получаемые при различных значениях соответствующих углов. Вид этих функций для s -, p - и d -состояний представлен на рисунке 2.7 [60].

Рассматривая атом водорода и оперируя понятиями «оболочка», «подоболочка», «орбита», мы говорили не столько о строении атома, сколько о возможностях, открывающихся перед единственным электроном, содержащимся в этом атоме. Электрон в атоме водорода может переходить с оболочки на оболочку и с орбиты на орбиту в пределах одной оболочки. Гораздо сложнее оказывается картина распределения электронов и возможностей их переходов в многоэлектронных атомах.

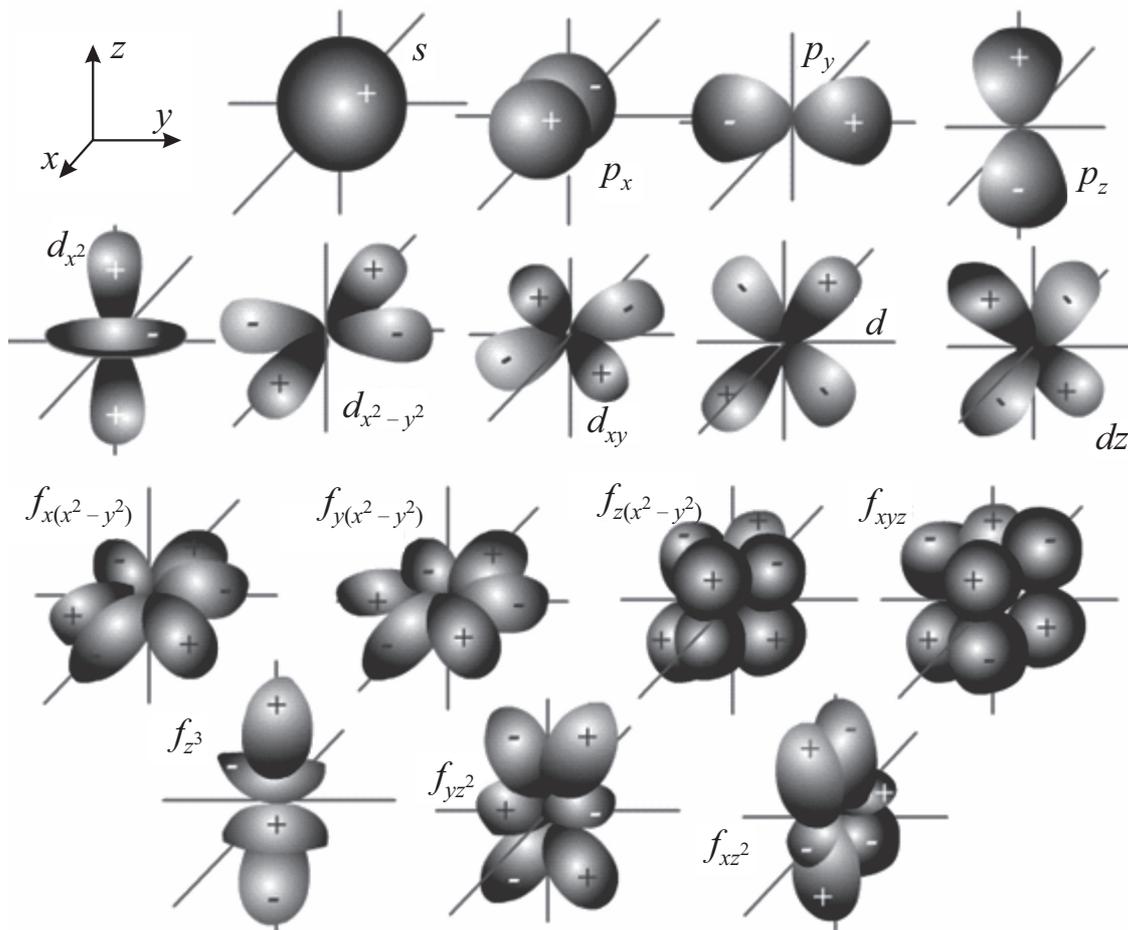


Рис. 2.7 – Изображение s -, p - и d -орбиталей (в сферических координатах, соотнесенных с декартовыми координатами)

Водородоподобный атом содержит только один электрон, поэтому на него действуют только силы притяжения со стороны положительного заряда ядра. В многоэлектронных атомах дело обстоит намного сложнее, поскольку в них присутствует более одного электрона, а значит необходимо учитывать также и силы отталкивания между электронами. Эта задача не поддается точному решению, поэтому используется так называемое приближение *центральносимметричного* поля: движение каждого электрона приближенно рассматривается в суммарном электростатическом поле, которое создает заряд ядра. В некоторой степени заряд ядра «погашен» остальными электронами, поскольку их заряд имеет противоположный знак. Степень такого «экранирования» ядра остальными электронами рассчитывается с помощью соответствующих приемов. Но самое главное – делается допущение, что это суммарное электростатическое поле, действующее на отдельный электрон в многоэлектронном атоме, является по-прежнему сферически симметричным. Это означает, что присутствие многих электронов изменяет только радиальную часть волновой функции, а сферические гармоники остаются по своей форме такими же, как в водородоподобном атоме. Именно это обстоятельство делает их изображения универсальными и весьма полезными при рассмотрении возможных вариантов взаимодействия атомов при образовании химических связей.

Шредингер вскоре показал, что его волновая механика оказывается идентичной матричной форме квантовой механики, разработанной В. Гейзенбергом, М. Борном и П. Йорданом.

2.6 Принцип запрета Паули, вырождение состояний и снятие вырождения под внешним воздействием

В многоэлектронных атомах очень важным является квантовый принцип запрета, сформулированный в 1925 г. известным швейцарским физиком В. Паули. Согласно этому принципу в атоме в любом состоянии, описываемом набором четырех квантовых чисел (n, l, m, σ) , не может находиться более одного электрона. Иначе говоря, на любой атомной стационарной орбите одновре-

менно может находиться только один или два электрона, причем в последнем случае спины электронов должны иметь противоположные направления, то есть для одного электрона $\sigma = +1/2$ а для другого $\sigma = -1/2$. Тогда для известного количества орбиталей (орбит), определяемых разными квантовыми числами, можно точно определить возможное число электронов в каждой оболочке и подоболочке атома.

Первая оболочка $n = 1$ не делится на подоболочки, так как для нее $l = 0$. А так как и $m = 0$, то первая оболочка состоит из одной орбиты с двумя электронами с разными спинами.

Для второй оболочки $n = 2$, $l = 0$ или $l = 1$ и она состоит из s - и p -подоболочек. На s -подоболочке с круговой орбитой размещаются 2 электрона, на p -подоболочке из трех орбит p_x, p_y, p_z (рис. 2.7) могут находиться 6 электронов, а всего на второй оболочке могут находиться 8 электронов.

И так далее, например, на $3d$ -подоболочке ($n = 3, l = 2$) может быть 10 электронов: по два электрона на каждой из пяти орбит, характеризующихся различными значениями квантового числа. Максимальное число электронов в любой подоболочке равно $2(2l + 1)$. В спектроскопии различным оболочкам присваивают буквенные обозначения – термы: первая оболочка обозначается буквой K , вторая – L , третья – M и т. д.

В атоме водорода единственный электрон находится в центрально-симметричном поле атомного ядра, и энергия электрона определяется только значением главного квантового числа n и не зависит от значений остальных квантовых чисел (рис. 2.5).

В многоэлектронном атоме энергия электрона зависит и от главного квантового числа n , и от орбитального квантового числа l и не зависит от m и σ . Например, все 6 электронов, находящихся на $3p$ -подоболочке, хотя и характеризуются различными n и l , имеют одну и ту же энергию. Состояния с разными n, l, m, σ , но с одинаковой энергией, называются *вырожденными*. И уровни, к которым относятся эти состояния, называются вырожденными. Поэтому многоэлектронных атомах энергетический спектр существенно отличается от спек-

тра атома водорода. На рисунке 2.8 показана часть спектра многоэлектронного атома для трех первых оболочек. Кружочки на уровнях обозначают максимальное число электронов, которым разрешено расположиться в соответствующей подоболочке.

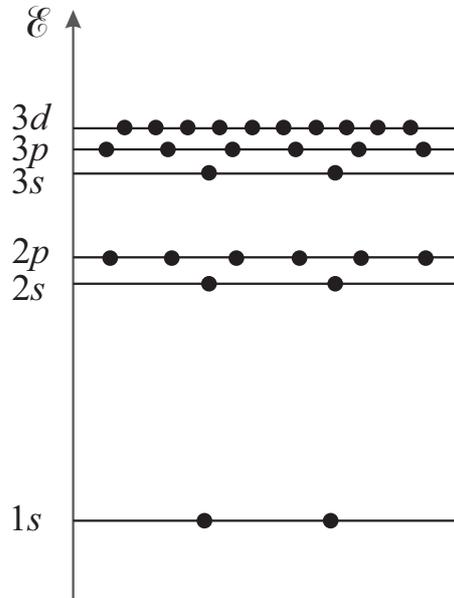


Рис. 2.8 – Энергетический спектр разрешенных состояний в многоэлектронном атоме

По мере увеличения числа электронов в атомах происходит заполнение лежащих выше подоболочек и оболочек. При этом верхняя оболочка может быть заполнена только одним электроном. У лития в атоме всего три электрона. Два располагаются в первой оболочке и один – во второй, ему на первой оболочке по принципу запрета Паули места нет. Поэтому третий электрон расположен во второй оболочке с $n = 2$, и она частично заполнена. Это означает, что электрон на второй оболочке слабо связан с атомным остатком и легко от него отделяется.

Вырожденность энергетических уровней сохраняется, пока атомы свободные. Если же подвергнуть атомы воздействию, например сильным магнитным или электрическим полем, то вырождение частично или полностью снимается.

Разные числа m отражают разную ориентацию в пространстве одинаковых электронных орбит. По отношению к внешнему полю орбиты ориентированы по-разному (рис. 2.7), под действием поля изменение энергии электрона будет разным и по величине и по знаку: у одних электронов энергия возрастет, у других – уменьшится, соответственно уровни изменят свое положение. Причем вместо одного уровня появятся подуровни, соответствующие одинаковым орбитам с разной ориентацией, и число подуровней соответствует разным возможным значениям квантового числа m .

Пример воздействия внешнего магнитного поля на уровень $3d$, для которого $n = 3$ и $l = 2$, приведен на рисунке 2.9, *а*: видно, что одновременно происходит и расщепление уровня на подуровни, и смещение подуровней. Это процесс внесения различия в ранее неразличимые по энергии вырожденные уровни называется *снятие вырождения*.

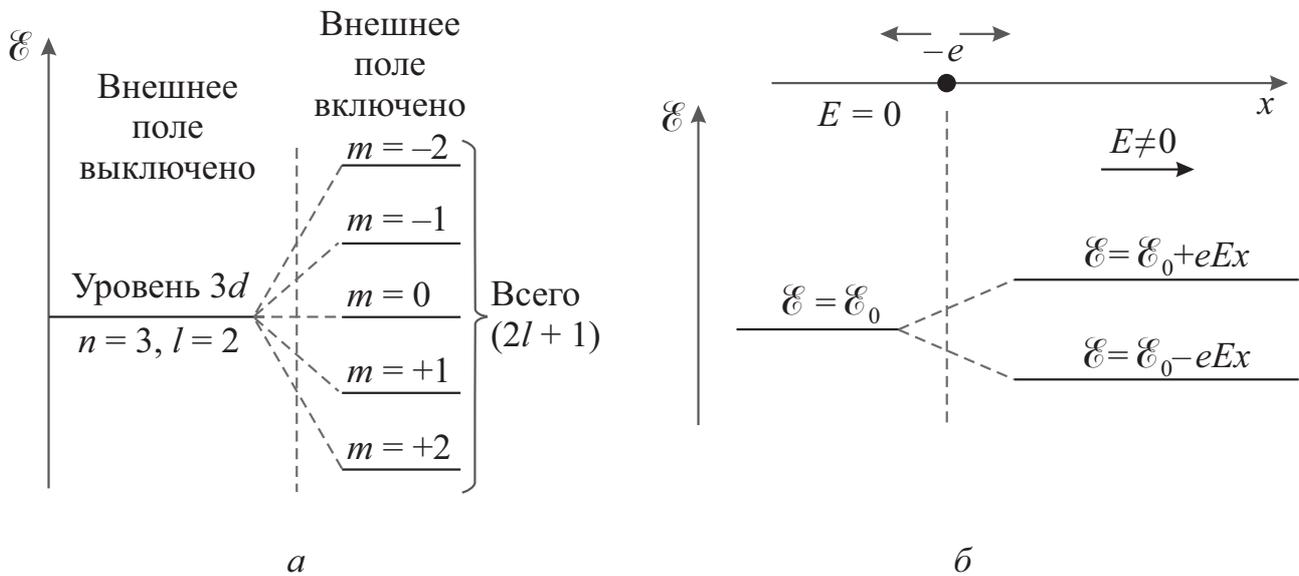
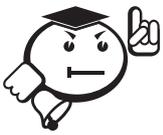


Рис. 2.9 – Снятие вырождения: *а*) под действием магнитного поля;
б) под действием электрического поля E

На рисунке 2.9, *б* показано снятие вырождения под действием электрического поля. Без поля ($E = 0$) электрон занимает состояние \mathcal{E}_0 , то есть у него двукратное вырождение в соответствии с двумя значениями спина. При включении электрического поля вдоль оси x (рис. 2.9, *б*) электрон, движущийся по полю,

будет тормозиться и его энергия будет $\mathcal{E}_0 - eEx$, а движущийся в обратном направлении электрон будет иметь энергию $\mathcal{E}_0 + eEx$ (x – это пройденное расстояние). Соответственно двум энергетическим состояниям на спектре появляются два энергетических уровня – происходит расщепление вырожденного уровня \mathcal{E}_0 на два невырожденных: $\mathcal{E}_0 - eEx$ и $\mathcal{E}_0 + eEx$.

2.7 Образование энергетических зон и их заполнение в кристаллах



.....

Можно представить себе как образуется кристалл из свободных атомов. При их сближении между ними происходит взаимодействие, аналогичное тому воздействию, которое происходит от электрического поля. Оно приводит к снятию вырождения с энергетических уровней: каждый вырожденный уровень расщепляется на $(2l + 1)$ невырожденных подуровней. Подуровни, получившиеся из одинаковых атомных уровней не сливаются в линию, а отличаются по энергии: одни – несколько выше исходных уровней свободных атомов, другие – несколько ниже. Так работает принцип Паули, обобщенный на весь кристалл: в кристалле не может быть двух одинаковых по энергии невырожденных подуровней.

.....

Поэтому каждый энергетический уровень при образовании кристалла из N атомов расщепляется в энергетическую зону, состоящую из $N(2l + 1)$, отличающихся по энергии невырожденных подуровней. Так, уровень $1s$ расщепляется в зону $1s$, состоящую из N подуровней, на которых могут расположиться $2N$ электронов, уровень $2p$ размывается в зону $2p$, состоящую из $3N$ подуровней и способную вместить $6N$ электронов и т. д.

На рисунке 2.10 показано расщепление уровней при сближении атомов. Чем меньше между атомами расстояние r , тем сильнее воздействие на атомы со

стороны соседних атомов и тем больше размываются уровни. При формировании кристалла устанавливается межатомное расстояние a_0 , характерное для данного кристалла. Этим a_0 определяется энергетический спектр кристалла.

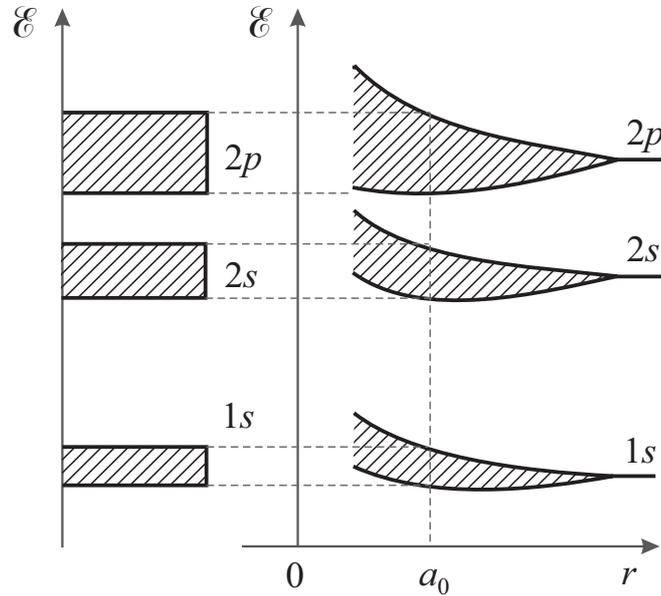


Рис. 2.10 – Схема расщепления уровней при образовании кристалла

Как сильно размываются уровни зависит от глубины их расположения в атоме. Внутренние электроны сильнее связаны с ядром и сильно экранируются от внешнего воздействия наружными электронными оболочками, поэтому соответствующие им уровни испытывают слабое размытие. Внешние электроны подвергаются наибольшему воздействию со стороны поля кристаллической решетки, и соответствующие им уровни сильнее размываются. Размытие уровней в зоны не зависит от того, имеются на соответствующих атомных уровнях электроны или они пусты. Если уровни пусты, то размытие уровней отражает расширение диапазона возможных энергий, которые может приобрести электрон в кристалле.

Каждому разрешенному атомному уровню в кристалле соответствует целая зона разрешенных значений энергии – *разрешенная зона*. Разрешенные зоны чередуются с зонами запрещенных энергий, или *запрещенными зонами*. Чем выше по шкале энергии расположен разрешенный атомный уровень, тем более размытой оказывается соответствующая зона. В чистом кристалле электроны

не могут обладать энергией, лежащей в диапазоне запрещенных зон. Ширина же запрещенных зон по мере увеличения энергии становится все более узкой.

Расстояние между подуровнями в разрешенной зоне очень мало и по порядку величины лежит в пределах от 10^{-22} до 10^{-24} эВ. Это очень малое различие в энергии и обычно зоны рассматриваются как непрерывные. Тем не менее сам факт дискретности подуровней в зонах и то обстоятельство, что число подуровней в зоне всегда конечно, играют принципиальную роль во всей физике кристаллических сред, так как именно в зависимости от степени заполнения зон электронами все твердые тела делятся на проводники, полупроводники и диэлектрики.

Заполненные уровни создают заполненные зоны, а пустые уровни – пустые зоны. Заполнение уровней электронами будет определяться в первую очередь заполненностью электронами соответствующих атомных уровней.

Для полупроводников с кристаллической решеткой типа алмаза, кремния и германия перекрытие зон приводит к интересному результату.

В атомах кремния на уровне $3p$ (в подоболочке $3p$) находятся всего два электрона, хотя на этом уровне могут располагаться шесть электронов. При образовании кристалла кремния верхняя энергетическая зона (зона $3p$) окажется заполненной лишь частично, в то время как предшествующая ей зона (зона $3s$) будет заполнена целиком (она образуется из полностью заполненного уровня $3s$). Однако в действительности эффект перекрытия приводит не только к возникновению гибридной зоны, составленной из подуровней $3s$ и $3p$, но и к дальнейшему расщеплению создавшейся гибридной зоны на две подзоны, разделенные между собой промежутком запрещенных энергий E_g (рис. 2.11). Всего в гибридной зоне $3s + 3p$ должно быть по 8 вакантных электронных мест на каждый атом (2 места от подоболочки $3s$ и 6 мест от подоболочки $3p$). При расщеплении же гибридной зоны в каждой из подзон оказывается по 4 места на атом. Стремясь занять нижние энергетические уровни, электроны третьих оболочек атомов кремния (а их всего четыре – два на подоболочке $3s$ и два

на подболочке $3p$) как раз и заполняют нижнюю подзону. Верхняя подзона при этом остается пустой.

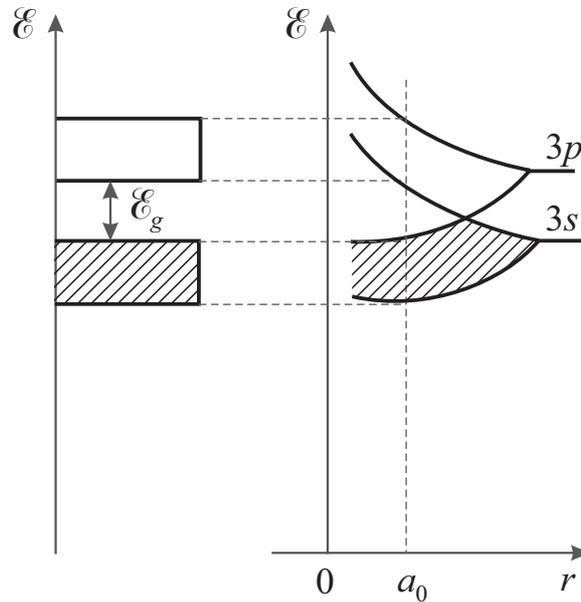


Рис. 2.11 – Образование запрещенной зоны и разрешенной зоны в кристаллах с кристаллической решеткой типа алмаза

2.8 Соотношение неточностей Гейзенберга

В 1934 г. на русском языке были опубликованы Нобелевские лекции В. Гейзенберга, Э. Шредингера и П. Дирака [61], а в 1935 г. была опубликована на русском языке работа Макса фон Лауэ о соотношениях неточностей Гейзенберга и их теоретико-познавательном значении [62]. В современной физике соотношения неточностей называют соотношением неопределенностей.

Гейзенберг является главным представителем Боровской «копенгагенской» школы, защищающей положение о построении теории на основе «принципиально наблюдаемых» величин. Между точностью, с которой может быть установлено положение частицы, и точностью, с которой одновременно может быть измерен ее импульс, существует определенное соотношение. Именно, произведение вероятных ошибок при измерении положения q и импульса p всегда будет по меньшей мере равняться постоянной Планка, деленной на 4π [61]. Таким образом, всегда должно иметь место

$$\Delta p \Delta q \geq \frac{h}{4\pi}. \quad (2.28)$$

Волновые свойства частиц и возможность задать для частицы лишь вероятность *ее пребывания в данной точке пространства* приводят к тому, что сами понятия *координаты частицы и ее скорости (или импульса) могут применяться в квантовой механике в ограниченной мере.*

К пересмотру понятия орбиты электрона вынуждает, например, анализ возможных экспериментов, необходимых для наблюдения этой орбиты. Наиболее непосредственным способом будет наблюдение электрона, скажем, в микроскоп исключительно большой разрешающей силы. Но так как освещение объекта наблюдения в этом микроскопе должно быть произведено светом весьма малой длины волны, то уже первый световой импульс, который доходит от источника света к электрону и к нашему глазу выбросит электрон из его орбиты. Таким образом, из всей орбиты возможно будет наблюдать только одну точку [61].

Итак, нельзя одновременно характеризовать частицу ее положением в пространстве (координатами) и скоростью (или импульсом). Электрон не может иметь одновременно точных значений координаты q и импульса p .

Из (2.28) ясно, что чем меньше неопределенность импульса (Δp), тем больше неопределенность координаты (Δq). Если координата имеет точное значение ($\Delta q = 0$), то импульс (скорость) совершенно неопределен ($\Delta p \rightarrow \infty$), и наоборот. Если бы нам удалось абсолютно точно установить координаты квантовой частицы, о ее скорости мы не имели бы ни малейшего представления. То есть из (2.28) следует, что нет таких энергетических состояний электрона, в которых его координаты и импульс одновременно имели бы точные значения. Поэтому невозможно измерять с любой наперед заданной точностью одновременно и координаты электрона, и его скорость (импульс).

Само измерение вносит неопределенность. Действительно, нужна некоторая твердая система (например, координатная система, в которой покоится наблюдатель), относительно которой и производятся все измерения. Допуская,

что эта система является «твердой», мы тем самым отказываемся заранее от знания ее импульса, так как понятие «твердый» как раз и означает, что какие бы то ни было изменения импульса системы при взаимодействии с ней не должны иметь места. Принципиально необходимая неопределенность, допущенная в этом месте, распространяется затем дальше через измерительный аппарат на атомные процессы.



.....

Соотношение (2.28) имеет место и для других пар величин, как их называют в классической механике, – канонически сопряженных. Так что, в общем виде, утверждение о том, что произведение неопределенностей значений двух сопряженных переменных не может быть по порядку меньше постоянной Планка, называется соотношением неопределенностей Гейзенберга.

.....

Для энергии и времени, которые тоже канонически сопряженные величины, также справедливо соотношение:

$$\Delta t \cong \frac{h}{4\pi\Delta\mathcal{E}}, \quad \Delta\mathcal{E}\Delta t \geq \frac{h}{4\pi} = \frac{\hbar}{2}. \quad (2.29)$$

То есть, определение энергии с точностью $\Delta\mathcal{E}$ должно занять интервал времени, равный, по меньшей мере, $\Delta\mathcal{E}\Delta t \geq \frac{h}{4\pi}$, $\Delta t \geq \frac{h}{4\pi\Delta\mathcal{E}}$.

В классической механике измерение координаты и импульса может быть произведено с любой точностью, поэтому соотношение неопределенностей является, таким образом, *квантовым ограничением применимости классической механики к микрообъектам*.

Соотношение неопределенностей указывает с какой степенью точности можно говорить о траекториях микрочастиц. Движение по траектории характеризуется вполне определенными значениями координат и скорости в каждый момент времени. Подставив в (2.28) вместо p произведение mv , получим соотношение:

$$\Delta v \Delta q \geq \frac{h}{4\pi m}, \quad (2.30)$$

из которого следует, что чем больше масса частицы, тем меньше неопределенности координат и скорости, значит, тем с большей точностью можно применять к частице понятие траектории.

Рассмотрим работу соотношения (2.28) на примерах. Возьмем макрочастицу – пылинку с массой 10^{-12} кг и размером 10^{-6} м. Точность ее координаты определена до 0,01 ее размера ($\Delta q = 10^{-8}$ м), тогда неопределенность ее скорости по соотношению (2.30) $\Delta v \geq \frac{h}{\Delta q \cdot 4\pi \cdot m} = \frac{6,26 \cdot 10^{-34}}{10^{-8} \cdot 4\pi \cdot 10^{-12}} = 0,5 \cdot 10^{-14}$ м/с. Это такая малая величина, что частица может двигаться с любой возможной для нее скоростью. Получается, что для макротел их волновые свойства не играют никакой роли, а их координаты и скорости измеряются с высокой точностью. И для описания движения макротел можно пользоваться законами классической механики.

Теперь возьмем поток микрочастиц – электронов, движущихся со скоростью $v = 10^8$ м/с, масса электрона $9,11 \cdot 10^{-31}$ кг, точность определения скорости 0,01%, то есть $\Delta v = 10^4$ м/с. Тогда точность определения координаты $\Delta q \geq \frac{h}{\Delta v \cdot 4\pi \cdot m} = \frac{6,26 \cdot 10^{-34}}{10^4 \cdot 4\pi \cdot 9,11 \cdot 10^{-31}} = 0,5 \cdot 10^{-6}$ м. Такая точность в тысячные доли миллиметра для определения положения электрона достаточна для рассмотрения его движения по определенной траектории, то есть для описания движения таких электронов классической механикой.

Теперь рассмотрим электрон, движущийся в атоме водорода. Неопределенность его координаты порядка размеров атома. $\Delta q \cong 0,1 \cdot 10^{-10}$ м. Тогда неопределенность в скорости:

$$\Delta v \geq \frac{h}{\Delta q \cdot 4\pi \cdot m} = \frac{6,26 \cdot 10^{-34}}{0,1 \cdot 10^{-10} \cdot 4\pi \cdot 9,11 \cdot 10^{-31}} = 0,6 \cdot 10^7 \text{ м/с.}$$

По классической оценке при движении электрона по круговой орбите радиуса $\cong 0,5 \cdot 10^{-10}$ м его скорость $v \cong 2 \cdot 10^6$ м/с, так что рассчитанная неопреде-

ленность скорости в несколько раз больше самой скорости. Ясно, что в таком случае нельзя говорить о движении электронов в атоме по определенной траектории и для описания движения электронов в атоме нельзя пользоваться законами классической физики.



.....
Контрольные вопросы по главе 2
.....

1. В чем состоит гипотеза Планка? Поясните формулу Планка для излучения.
2. В чем состоит гипотеза де Бройля? Поясните формулу де Бройля.
3. В чем смысл функции, относительно которой записывается уравнение Шредингера?
4. Приведите и поясните соотношение неопределенностей Гейзенберга.
5. Поясните принцип запрета Паули и заполнение энергетических зон электронами.

3 Квантовое поведение электронов в кристалле

3.1 Уравнение Шредингера для кристалла

Ядра атомов в полупроводниковом кристалле образуют кристаллическую решетку, которая обладает свойством пространственной периодичности. При наложении внешнего электрического поля решетка практически не деформируется, хотя ядра атомов и заряжены. Силы, удерживающие ядра атомов в узлах кристаллической решетки, обычно значительно больше тех сил, которые создаются внешними электрическими полями. Те из электронов, которые не находятся близко к ядру атома и поэтому к нему не очень сильно притягиваются, могут передвигаться по твердому телу, создавая электрический ток. Однако при количественном описании этого явления возникают серьезные трудности. Они связаны с тем, что электроны являются заряженными частицами и при своем движении по твердому телу встречаются с другими электронами. Но так как между электронами действуют электрические силы отталкивания, движение электрона оказывается зависимым от движения окружающих его электронов.

Иными словами, в рассматриваемом случае надо решать не одноэлектронную, а многоэлектронную задачу. Поэтому для определения стационарных состояний и энергетического спектра совокупности большого числа атомных ядер и электронов в кристалле нужно решить уравнение Шредингера, в котором число независимых переменных определяется полным числом частиц в кристалле. Число частиц в кубическом сантиметре кристалла порядка 10^{23} и такая задача не может быть решена в общем виде. Возможно только приближенное решение задачи. Среди возможных приближений известны адиабатическое приближение и одноэлектронное приближение, например, кратко представленные в [63].

Адиабатическое приближение. Скорости электронов в кристаллах приблизительно на два порядка больше, чем скорости ядер, поэтому для любой, даже неравновесной конфигурации ядер, всегда будет устанавливаться соответ-

ствующее ей электронное равновесие. По этой причине движением ядер можно пренебречь и считать их неподвижными. Тогда исключается обмен энергией между электронной и ядерной системами, поэтому приближение называется адиабатическим. Естественно, что в адиабатическом приближении нельзя рассматривать такие явления, как диффузия, ионная проводимость и др., связанные с движением атомов или ионов.

Одноэлектронное приближение. Вместо взаимодействия данного электрона с остальными электронами и ядрами по отдельности рассматривают его движение в некотором результирующем усредненном поле остальных электронов и ядер. Такое поле называют самосогласованным. В одноэлектронном приближении задача сводится к независимому описанию каждого электрона в среднем внешнем поле с потенциальной энергией $V(r)$. Вид функции $V(r)$ определяется свойствами симметрии кристалла. Основное свойство самосогласованного поля заключается в том, что оно имеет тот же период, что и поле ядер.



.....

Аналитическое решение одноэлектронного уравнения Шредингера также невозможно получить в общем виде для произвольного потенциала $V(r)$. Здесь необходимо сделать дополнительные предположения.

.....

Часто используют следующие приближения.

1. *Приближение слабой связи.* Электроны в кристалле рассматривают как почти свободные частицы, на движение которых оказывает слабое возмущение поле кристаллической решетки. Данное допущение применимо, когда потенциальная энергия взаимодействия электрона с решеткой много меньше его кинетической энергии. Такой подход иногда называют «приближением почти свободных электронов».
2. *Приближение сильной связи.* В этом приближении состояние электрона в кристалле мало отличается от его состояния в изолированном атоме. Приближение сильной связи применимо, когда потенциальная

энергия электрона значительно больше его кинетической энергии. Этот метод применим чаще всего в полупроводниках для анализа их физических свойств.

Одноэлектронное уравнение Шредингера записывается в виде:

$$\left[-\frac{\hbar^2}{2m} \Delta + V(r) \right] \Psi(r) = E \Psi(r), \quad (3.1)$$

и нам нужно изучить те общие свойства уравнения (3.1), которые обусловлены периодичностью потенциала $V(r)$.



.....

Феликс Блох в 1928 г. [64] доказал теорему, на которой основана одноэлектронная теория твердых тел: *волновые функции, являющиеся решениями одноэлектронного уравнения Шредингера (3.1) с периодическим потенциалом, имеющим период решетки, представляют собой плоские волны, модулированные периодической функцией*

$$\Psi_k(r) = u_k(r) e^{ikr}, \quad (3.2)$$

здесь $u_k(r)$ – периодическая функция с периодом решетки, зависящая от квазиволнового вектора k , характеризующего квантовое состояние электрона в кристалле.

.....

Независимые электроны, каждый из которых подчиняется одноэлектронному уравнению Шредингера с периодическим потенциалом, называют блоховскими (в отличие от «свободных», к которым блоховские электроны сводятся, если периодический потенциал тождественно равен нулю). Из периодичности потенциала $V(r)$ следует очень важное свойство стационарных состояний блоховских электронов.

3.2 Модель Кронига – Пенни

Для движения электронов в кристалле к настоящему времени удалось решить только очень упрощенные задачи такого движения электрона, в основ-

ном не трех-, а одномерные. Ниже рассмотрим результаты решения одной из них (модель Кронига – Пенни) об одномерном движении электрона в периодическом потенциале. Потенциальная энергия электронов в кристалле является функцией координат с периодом, равным расстоянию между атомами. Электрон в периодическом поле имеет ряд общих свойств, не зависящих от конкретной формы периодического потенциала. Большинство свойств можно выяснить при рассмотрении упрощенной модели кристалла, представляя его в виде линейной одномерной цепочки атомов. При этом периодический потенциал становится одномерным. Форма периодического потенциала выбирается такой, чтобы расчет оказался наиболее простым. В этом случае возможно проведение строгого решения уравнения Шредингера. В такой модели почти свободных электронов, которую предложили Крониг и Пенни в 1931 г. [65], рассматривается движение электрона в линейной цепочке прямоугольных потенциальных ям. Одномерный потенциал имеет вид прямоугольных ям с шириной a , и они отделены друг от друга потенциальными барьерами толщиной b и высотой U_0 (рис. 3.1), которые чередуются с прямоугольными потенциальными ямами:

$$\begin{aligned} U(z) &= 0, \text{ если; } nc < z < nc + a \\ U(z) &= U_0, \text{ если } (nc + a) < z < (n+1)c, \end{aligned} \quad (3.3)$$

постоянная кристаллической решетки $c = a + b$, $n = 0, \pm 1, \pm 2, \dots$

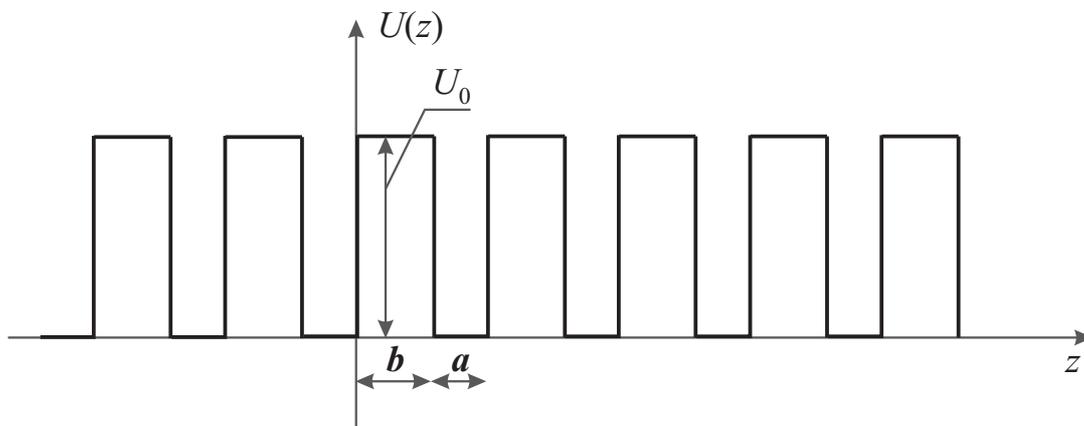


Рис. 3.1 – Рельеф потенциала электрона в кристалле в модели Кронига – Пенни

В кристалле внешнее электрическое поле E действует на электрон так же, как на свободный электрон: с кулоновской силой $F = -eE$, направленной против поля. Однако кроме силы $-eE$ на электрон действуют значительные внутренние силы, создаваемые периодическим полем решетки. Это означает, что ускорение электрона в решетке в общем случае может быть не направлено параллельно внешней силе F . Следовательно, движение электрона в кристалле будет более сложным, чем движение свободного электрона. Именно поэтому энергия электрона, движущегося в периодическом поле кристалла, не имеет квадратичной зависимости от волнового вектора $\mathcal{E} \neq p^2/2m$. Это также следует из модели Кронига и Пенни. Однако для практических целей удобно сохранить зависимость энергии электрона от квазиимпульса в классическом виде, а все различия, вызванные влиянием периодического поля, включить в массу электрона. Тогда формулу $\mathcal{E} \neq p^2/2m$ можно переписать в виде $\mathcal{E} \neq p^2/2m^*$, где вместо m появляется некоторая величина m^* , которая является функцией энергии и называется *эффективной массой*. В этом проявляется действие поля решетки – движение электрона определяется не его обычной массой, а эффективной массой.

Введенный нами волновой вектор k при обсуждении функций Блоха играет важную роль в движении электрона в периодическом поле кристалла, такую же роль, какую играет волновой вектор в задаче о движении электрона. Состояние свободно движущегося электрона характеризуется энергией \mathcal{E} и импульсом p , при этом $\mathcal{E} = p^2 / 2m$. Этому электрону соответствует волна де Бройля с длиной $\lambda = \frac{h}{p} = \frac{h}{m\nu}$, ν – скорость электрона. Учитывая, что $|\vec{k}| = \frac{2\pi}{\lambda}$ получим, что $\vec{p} = \hbar\vec{k}$, откуда видно, что волновой вектор пропорционален импульсу электрона.

Энергия свободного электрона связана с \vec{k} соотношением

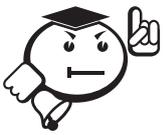
$$\mathcal{E} = \frac{\hbar^2 k^2}{2m}. \quad (3.4)$$

Если на электрон никакие силы не действуют, то его энергия остается постоянной $\mathcal{E}(\vec{k}) = \text{const}$. Это означает, что не меняется \vec{k} и остается постоянным импульс \vec{p} . По существу, это закон сохранения энергии и импульса.



На электрон, движущийся в кристалле, всегда действует периодическое поле решетки. Энергия этого взаимодействия является периодической функцией координат. Следовательно, энергия и импульс электрона в кристалле изменяются со временем под действием этого поля, т. е. не сохраняются. Однако, пользуясь понятием волнового вектора \vec{k} , введенного для электрона в кристалле, т. е. входящего в функцию Блоха (3.2), можно ввести характеристику, аналогичную импульсу, но сохраняющуюся во времени:

$$\vec{P} = \hbar \vec{k}. \quad (3.5)$$



Чтобы подчеркнуть сходство и одновременно отметить отличие в (3.5) величины $\hbar \vec{k}$ от истинного импульса, эту величину называют квазиимпульсом электрона.

Уравнение Шредингера необходимо рассматривать в каждой области с последующим сшиванием полученных решений: для $U(z) = 0$ уравнение

$$-\frac{\hbar^2}{2m_n} \frac{d^2}{dz^2} \Psi = \mathcal{E} \Psi, \quad (3.5.1)$$

а для $U(z) = U_0$ уравнение

$$\left[-\frac{\hbar^2}{2m_n} \frac{d^2}{dz^2} + U_0 \right] \Psi = \mathcal{E} \Psi, \quad (3.5.2)$$

m_n – эффективная масса электрона в кристалле.

Решение уравнения Шредингера дает уравнение, определяющее значения энергии электрона \mathcal{E} , для которых существуют нетривиальные решения в виде:

$$\cos(\alpha a) \cosh(\beta b) + \frac{\beta^2 - \alpha^2}{2\alpha\beta} \sin(\alpha a) \sin(\beta b) = \cos(\lambda), \quad (3.6)$$

где $\alpha = \frac{1}{\hbar} \sqrt{2m_n \mathcal{E}}$, $\beta = \frac{1}{\hbar} \sqrt{2m_n (U_0 - \mathcal{E})}$. Анализ уравнения (3.6) осложнен тем, что энергия входит в α и в β . Это уравнение можно упростить, если еще более идеализированно представить потенциал. Произведение βb определяет степень прозрачности потенциального барьера для электронов, в него входит ширина барьера b и высота барьера U_0 . Рассмотрим случай, когда ширина барьера устремляется к нулю, а высота – к бесконечности, причем так, что произведение $U_0 b$ остается конечным ($U_0 b = \text{const}$) (то есть, мы рассматриваем тонкие высокие барьеры). Так как $U_0 \sim \beta^2$, то при $b \rightarrow 0$ и $U_0 \rightarrow \infty$ произведение $\beta b \rightarrow 0$.

Обозначим

$$P = \lim_{\substack{b \rightarrow 0 \\ U_0 \rightarrow \infty}} \frac{\beta^2 a b}{2}. \quad (3.7)$$

Тогда можно упростить уравнение (3.6):

$$\cos(\alpha a) + P \frac{\sin(\alpha a)}{\alpha a} = \cos(\lambda). \quad (3.8)$$



.....

Уравнение (3.8) называется уравнением Кронига – Пенни. Параметр P пропорционален площади потенциального барьера; он характеризует степень прозрачности барьера для электронов или степень связанности электрона внутри потенциальной ямы.

.....

На рисунке 3.2 показан графический анализ уравнения Кронига – Пенни. Изобразим графически его левую часть как функцию аргумента αa (рис. 3.2, сплошная линия). Действительные корни этого уравнения существуют только при тех значениях αa , при которых левая часть уравнения принимает значения в интервале $[-1; 1]$. На рисунке 3.2 заштрихованы области допустимых значений αa . Их ширина зависит от параметра P : чем он меньше, тем они шире,

и при фиксированном значении P становятся шире с увеличением αa , а значит и энергии.

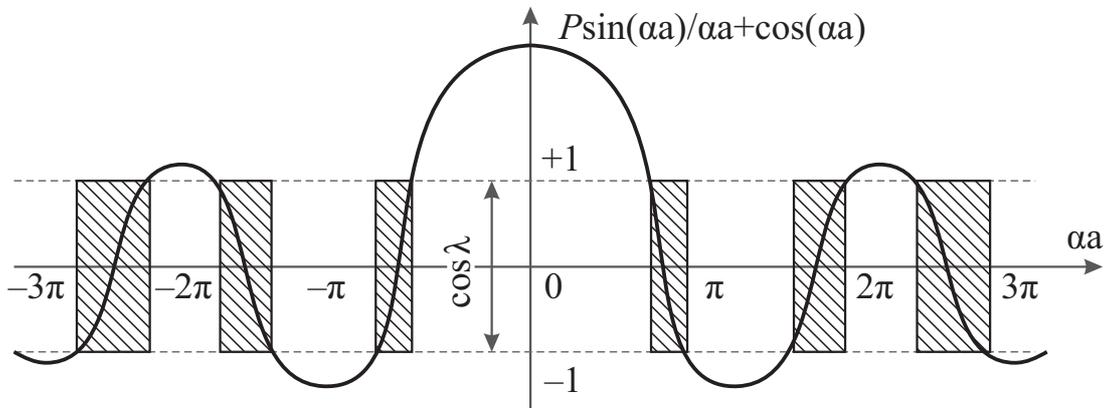
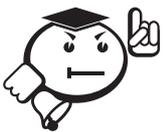


Рис. 3.2 – Разрешенные (заштрихованные) и запрещенные (светлые) зоны



Таким образом, энергия электрона в периодическом поле не может принимать любое значение, как для свободного электрона. Она ограничена рядом полос (зон) разрешенных значений, отделенных друг от друга запрещенными зонами – энергетический спектр электрона в периодическом поле имеет зонную структуру. Ширина разрешенных зон определяется степенью связанности электрона внутри потенциальной ямы.

3.3 Переход к нанофизике полупроводников

Диапазон линейных размеров элементов нанoeлектроники 100...10 нм – это фундаментальный физический барьер, за которым резко меняются все свойства твердого тела, включая электропроводность. В полную силу и в полной мере начинают проявляться квантовые эффекты, и физика проводимости определяется квантово-механической интерференцией электронных волн. Существуют два способа перехода к наноструктурам: 1) путем уменьшения линейных размеров элементов в чипе, то есть традиционный; 2) путь изготовления

искусственных периодических слоистых структур со слоями нанометровой толщины. В таких слоях открывается возможность формирования заданного энергетического спектра электронов. А с появлением технологической возможности изготавливать такие структуры появился даже термин «зонная инженерия».



.....

2000 год – весьма символический рубеж перехода в XXI в.: это и переход к 100-нанометровой технологии в построении чипов (см. рис. 1.2 в главе 1), и переход к новому состоянию физики полупроводников, когда она стала главным образом физикой полупроводниковых низкоразмерных структур (наноструктур). Современные технологии и методы позволяют создавать монокристаллические слои и многослойные гетероструктуры с толщиной слоев 1–10 нм, сравнимой с длиной волны де Бройля носителей заряда

$$\lambda = \frac{h}{m_n v}.$$

.....

Современная физика и технологии уже научились создавать приборы, работающие на квантовых принципах. Сейчас можно создавать искусственные кристаллы, размещая атомы с точностью до одного-двух слоев. Такие полупроводниковые наноструктуры имеют размеры в несколько нанометров. Основные типы наноструктур: *квантовые ямы, нити, точки, сверхрешетки*. Это открывает принципиальную возможность наблюдения в полупроводниках и использования явлений, обусловленных волновой природой электрона, например интерференцию электронных волн и вызванные ею размерные квантовые эффекты, такие как квантование энергии и импульса носителей заряда в наноструктурах, резонансный характер прохождения электронов через них и другие. Размерное квантование кардинально меняет энергетический спектр носителей заряда, фононов, квазичастиц и ведет к возникновению целого ряда новых физических явлений и свойств полупроводниковых наноструктур.



Важное достоинство наноструктур связано с тем, что, изменяя геометрические размеры и конфигурацию нанообъектов, можно управлять свойствами системы. Открывается широкая возможность конструирования параметров структур и, прежде всего, энергетического спектра носителей заряда. От простой однобарьерной структуры переходят к двухбарьерной и трехбарьерной структуре, то есть рассматриваются одиночные или двойные квантовые ямы, отделенные барьерами конечной толщины. А периодическая структура с квантовыми ямами, разделенными не очень широкими потенциальными барьерами, образует сверхрешетку, в которой носители заряда могут туннелировать из ямы в яму, и длина свободного пробега этого носителя вдоль оси, перпендикулярной плоскостям нанослоев, превышает период структуры.



Обобщая материал главы 2, можно выделить и подчеркнуть необходимые положения квантовой теории. В силу корпускулярно-волнового дуализма носители заряда, электроны и дырки могут интерферировать, проникать через узкие барьеры и щели, при этом обладая признаками материальной частицы: масса, заряд, импульс и энергия. Но однозначно и достоверно предсказать, рассчитать местоположение носителя заряда невозможно. Можно описать распределение их в пространстве и оперировать с вероятностью распределения. Это распределение описывает волновая Ψ -функция, она дает вероятностное описание отдельного носителя заряда, так что вероятность его обнаружить в интервале пространства Δx равна $|\Psi|^2 \Delta x$. Волновая функция содержит полную информацию об элек-

тронах, дырках и других квантовых частицах в атоме или в кристалле.

.....

3.4 Поведение носителей заряда в прямоугольной потенциальной яме. Квантовое ограничение

При движении электрона в неограниченном объеме его энергия определяется квадратом импульса или, в волновом представлении, квадратом волнового вектора (выражение (3.4)). Плотность электронных состояний при этом является

непрерывной функцией энергии $n_{3D}(\mathcal{E}) = \frac{m_n \sqrt{2m_n \mathcal{E}}}{\pi^2 \hbar^3}$.



.....

Длина волны де Бройля для металлов составляет 1–10 нм, для полупроводников около 100 нм [66]. Поэтому в структурах очень малых размеров движение электрона будет ограничено в том направлении (или направлениях), в котором размер наноструктуры сравним с длиной волны де Бройля.

.....

При движении электрона в ограниченной области спектр его энергий оказывается квантован и его энергия имеет строго определенные дискретные значения [66]. Если электрон заперт в атоме или в молекуле, или в любой другой потенциальной яме, то его волновая функция представляет собой стоячую волну. Потенциальные ямы прямоугольной формы могут быть двух видов: симметричная и несимметричная (рис. 3.3).

В направлении ограничения движения электрона (все равно, какое направление, пусть в направлении z) потенциальная энергия его может быть представлена квантовым колодцем с бесконечно высокими краями, то есть бесконечно глубокой потенциальной ямой симметричного вида (рис. 3.4).

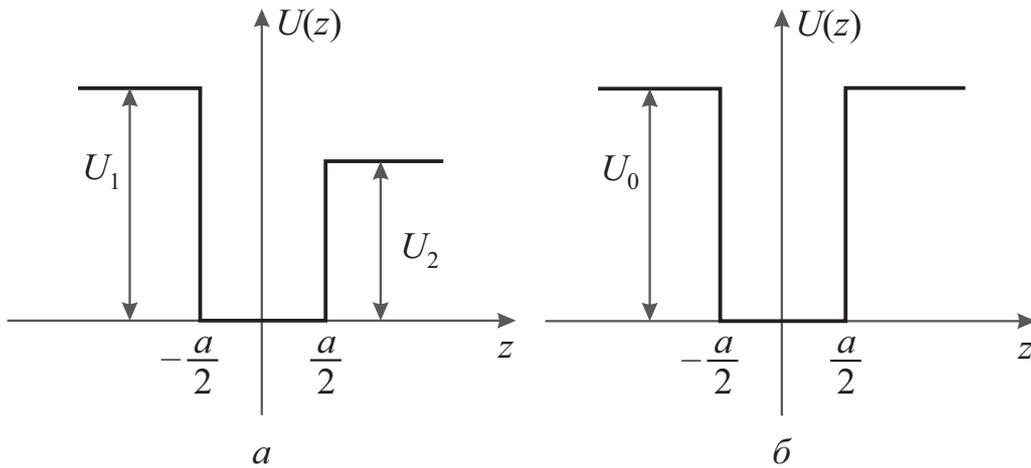


Рис. 3.3 – Потенциальный рельеф прямоугольной потенциальной ямы: а) несимметричной; б) симметричной

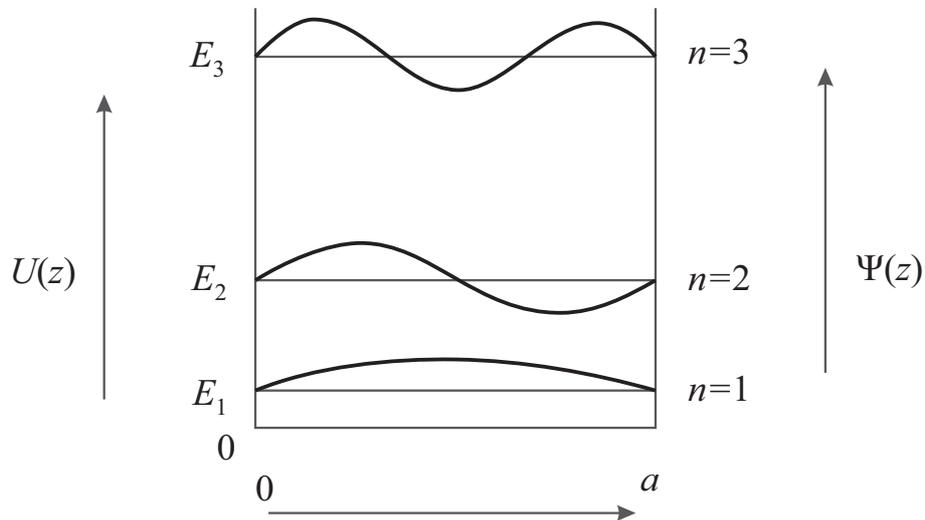


Рис. 3.4 – Прямоугольная симметричная потенциальная яма и волновые функции и уровни энергии частицы в ней

В такой потенциальной яме по своей форме волна будет такой же, как и в случае натянутой струны (такой подход использовался ранее в п. 2.5), только природа волны другая относительно струны, а дискретным будет не спектр частот, а спектр энергий. Бесконечно высокий потенциальный барьер делает невозможным нахождение электрона за границами области $0 < z < a$.

Одноэлектронное уравнение Шредингера (см. п. 3.1 и 3.2) имеет вид:

$$-\frac{\hbar^2}{2m_n} \frac{d^2}{dz^2} \Psi(z, \mathcal{E}) = \mathcal{E} \Psi(z, \mathcal{E}), \quad (3.9)$$

и ему удовлетворяет волновая функция $\Psi(z, \mathcal{E})$ внутри потенциальной прямоугольной ямы, а за пределами ямы $\Psi(z, \mathcal{E}) = 0$, что определяет граничные условия для модели рисунка 3.3 в виде:

$$\Psi\left(\pm \frac{a}{2}, \mathcal{E}\right) = 0. \quad (3.10)$$

Собственными значениями уравнения (3.9) являются дискретные значения энергии \mathcal{E} , при которых оно имеет решение:

$$\mathcal{E} = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a} \right)^2. \quad (3.11)$$

Условию бесконечно высокого потенциального барьера соответствует только ограниченный набор волновых функций – стоячие волны, описывающие электронные состояния в яме, это синусоиды, обращающиеся в точках $z = 0$ и $z = a$ в нуль.

$$\Psi_n(z) = \sqrt{\frac{2}{a}} \sin \frac{\pi z}{a} n, \quad (3.12)$$

где n – номер квантового состояния, a – размер ямы.

На рисунке 3.4 изображены три такие функции, соответствующие главному квантовому числу $n = 1, 2, 3, \dots$. Из рисунка видно, что есть максимумы и минимумы плотности вероятности и электронная плотность в яме распределяется неравномерно. Также следует, что длины волн волновых функций, описывающих электронные состояния с различными n , удовлетворяют условиям:

$$\frac{\lambda_n}{2} = \frac{a}{n}. \quad (3.13)$$

Для первого состояния $n = 1$: $\frac{\lambda_1}{2} = a$; для второго состояния $n = 2$: $\frac{\lambda_2}{2} = \frac{a}{2}$,

то есть $\lambda_2 = a$; для третьего состояния $n = 3$: $\frac{\lambda_3}{2} = \frac{a}{3}$. Таким образом, в прямоугольной потенциальной яме укладывается целое число полуволн. Конечно, для второго состояния таких полуволн укладывается две, то есть на ширине потенциальной ямы укладывается целая волна.

Соответствующие разрешенные значения волнового вектора дискретны и равны

$$k_n = \frac{2\pi}{\lambda_n} = \frac{n\pi}{a}. \quad (3.14)$$

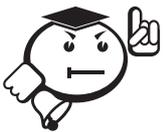
С учетом (3.14) дискретный спектр разрешенных энергетических состояний электрона в бесконечно глубокой яме можно переписать в виде:

$$\mathcal{E} = \frac{\hbar^2}{2m_n} \left(\frac{n\pi}{a} \right)^2 = \frac{\hbar^2 k_n^2}{2m_n} \cong \frac{h^2 n^2}{8m_n a^2}. \quad (3.15)$$

Из (3.14) и (3.15) следует, что электрон в ограниченной области пространства, размеры которой сравнимы с его длиной волны (длиной волны де Бройля) может занимать только дискретные разрешенные уровни. Наиболее низкое энергетическое состояние электрона:

$$\mathcal{E}_1 = \frac{\hbar^2 \pi^2}{2m_n}, \quad (3.16)$$

которая больше нуля всегда, в отличие от классической системы, в которой частица на дне потенциальной ямы имеет нулевую энергию. К тому же разрешенные энергии для электрона в потенциальной яме оказались квантованными и пропорциональными квадрату главного квантового числа n .



.....

Принцип неопределенности тоже приводит к ненулевому значению минимальной энергии электрона в потенциальной яме. В соотношении неопределенностей Гейзенберга $\Delta p \Delta z \geq \hbar/2$ в нашем случае потенциальной ямы в качестве координатной неопределенности надо подставлять $\Delta z = a$. Тогда неопределенность $\Delta p \geq \hbar/2a$, что определяет минимальное изменение энергии $\Delta \mathcal{E} = (\Delta p)^2 / (2m_n) = \hbar^2 / 8m_n a^2$, которое с учетом $\hbar = h/2\pi$, с точностью до $\pi^2/4$, соответствует приведенному выше выражению для \mathcal{E}_1 .

.....



Установленные нами квантованность разрешенных дискретных состояний носителей заряда (электронов, дырок) и ненулевая минимальная их энергия в потенциальной яме называются *квантовым ограничением* – все это следствия ограниченности движения носителей заряда в низкоразмерной структуре и яркое проявление их корпускулярно-волновой природы.

Таким образом, квантовое ограничение сопровождается как увеличением минимальной энергии запертого электрона, так и дополнительным квантованием энергетических уровней, соответствующих его возбужденному состоянию. Это приводит к тому, что электронные свойства наноразмерных структур отличаются от известных объемных свойств материала, из которого они сделаны. Квантовое ограничение в полупроводниках, как и в любом твердом теле, может реализоваться в любом из трех пространственных направлений [66].



Количество направлений, в которых нет квантового ограничения и носитель заряда движется свободно, используется в качестве критерия для классификации элементарных низкоразмерных структур по трем группам: квантовые пленки (ямы), квантовые шнуры (проволоки, нити) и квантовые точки.

3.4 Квантовые структуры с двумерным электронным газом



Одним из примеров таких структур являются тонкие пленки, соответственно для них одно из направлений (z) очень мало, следовательно, энергетический спектр в этом направлении можно описать формулой (3.15), в которой величина a соответствует толщине пленки, так как в данном направлении образуется

потенциальная яма. В оставшихся двух направлениях x и y электроны могут передвигаться свободно. Ансамбль электронов в квантовых пленках называют **двумерным электронным газом (two-dimensional electron gas, 2D EG)**.

.....

Считая образованную потенциальную яму бесконечно глубокой, получаем на практике толщину пленки порядка нанометров, что соответствует нескольким межатомным расстояниям. Полная энергия носителей в квантово-размерной пленке носит смешанный дискретно-непрерывный спектр, представляя собой сумму дискретных уровней, связанных с движением в направлении квантового ограничения и непрерывной компоненты, описывающей движение в плоскости слоя:

$$\mathcal{E} = \mathcal{E}_n + \frac{p_x^2 + p_y^2}{2m}, \quad (3.17)$$

где p_x и p_y – компоненты импульса в плоскости пленки.

С учетом выражения (3.15) можно записать (3.17) в следующем виде:

$$\mathcal{E} = \frac{\hbar^2}{2m_n} \left(\frac{n\pi}{a} \right)^2 + \frac{\hbar^2 k_x^2}{2m_n} + \frac{\hbar^2 k_y^2}{2m_n}$$

или, вместо a обозначив толщину пленки как l_z :

$$\mathcal{E} = \frac{\hbar^2}{2m_n} \left(\frac{n\pi}{l_z} \right)^2 + \frac{\hbar^2 k_x^2}{2m_n} + \frac{\hbar^2 k_y^2}{2m_n}. \quad (3.18)$$

В пространстве волновых векторов зонная диаграмма квантовой пленки – это семейство перекрывающихся параболических зон, образующих подзоны. Электрон с минимальной энергией, которая определяется выражением (3.16), неподвижен в плоскости пленки.

Поскольку волновой вектор имеет только k_x и k_y компоненты, то вместо параболической зависимости, как в объемных трехмерных структурах, в квантовой пленке зависимость плотности электронных состояний от энергии имеет ступенчатый вид ($\Theta(\mathcal{E} - \mathcal{E}_i)$ – ступенчатая функция):

$$n_{2D}(\mathcal{E}) = \frac{m_n}{\pi \hbar^2 l_z} \sum_i \Theta(\mathcal{E} - \mathcal{E}_i); \quad i = 1, 2. \quad (3.19)$$

Установленные особенности энергетического спектра низкоразмерной двумерной структуры в сравнении с объемной структурой представлены на рисунке 3.5 в графической форме [66].

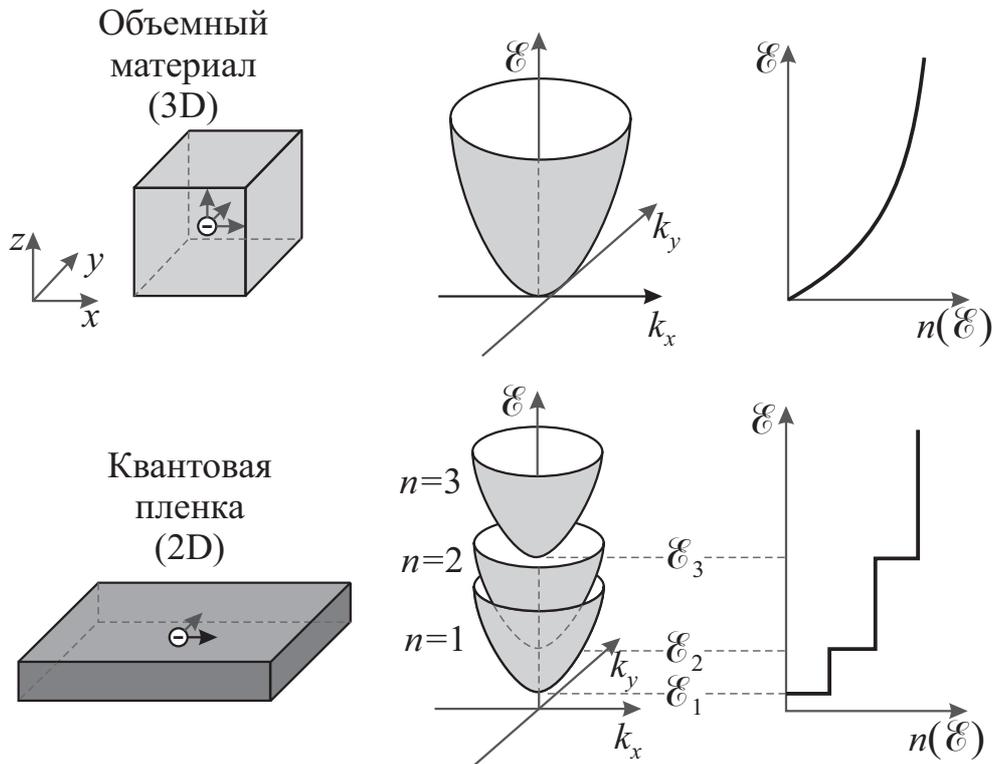


Рис. 3.5 – Двумерная низкоразмерная структура, ее энергетическая диаграмма и плотность состояний $N(\mathcal{E})$ в сравнении с трехмерной структурой

3.5 Квантовые структуры с одномерным электронным газом



В квантовых структурах, в которых движение носителей ограничено не в одном, а в двух направлениях (y и z), носители заряда могут свободно двигаться лишь в одном направлении (ось x). Такая структура представляется в виде **квантовой нити (quantum wires)** – это одномерная структура $1D$, ее еще называют **квантовый шнур**, или **квантовая проволока**.

Она имеет два нанометровых размера, в направлении которых действует эффект квантового ограничения (а в квантовой пленке был один такой размер – толщина пленки вдоль z). В поперечном сечении (плоскость yz) двумерное движение описывается двумя квантовыми числами, n и m , при этом энергия квантуется и принимает дискретные значения \mathcal{E}_{nm} . Полная энергия носителей в квантово-размерной нити, аналогично тонким пленкам, носит смешанный дискретно-непрерывный спектр, но лишь с одной непрерывной степенью свободы:

$$\mathcal{E} = \mathcal{E}_{nm} + \frac{p_x^2}{2m}, \quad (3.20)$$

где p_x – компонента импульса в направлении нити (x).

Выражение (3.20) можно переписать с учетом выражения (3.15) в следующем виде:

$$\mathcal{E} = \frac{\hbar^2}{2m_n} \left(\frac{n\pi}{l_y} \right)^2 + \frac{\hbar^2}{2m_n} \left(\frac{m\pi}{l_z} \right)^2 + \frac{\hbar^2 k_x^2}{2m_n}; \quad n, m = 1, 2, \dots, \quad (3.21)$$

так что полная энергия носителя заряда состоит из кинетической энергии вдоль одного направления (в данном случае вдоль x) и из квантованных дискретных значений в двух других направлениях.

Плотность электронных состояний в квантовой нити для каждой пары дискретных уровней зависит от энергии по закону $\mathcal{E}^{-1/2}$:

$$n_{1D}(\mathcal{E}) = \frac{(2m_n)^{1/2}}{\pi \hbar l_y l_z} \sum_{i,j} (\mathcal{E} - \mathcal{E}_{i,j})^{-1/2}; \quad i, j = 1, 2, \dots \quad (3.22)$$

Энергетический спектр квантовых нитей, изображенный на рисунке 3.6, также представляет собой совокупность подзон размерного квантования, но уже не двумерных, а одномерных.

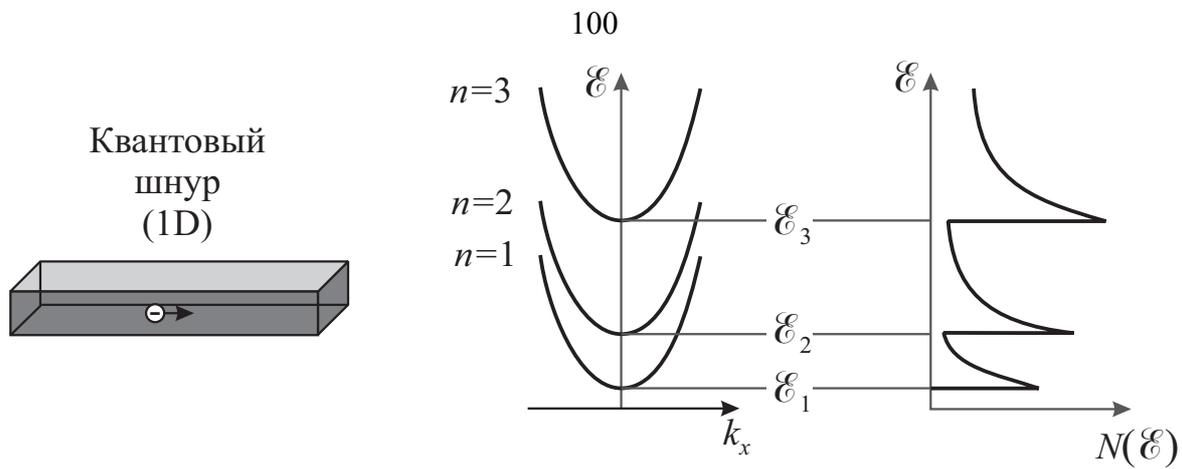


Рис. 3.6 – Одномерная низкоразмерная структура, ее энергетическая диаграмма и плотность состояний $N(\mathcal{E})$

3.6 Квантовые структуры с нульмерным электронным газом



Нульмерные квантовые структуры представляют собой квантовые точки, движение электрона в которых квантово ограничено во всех трех направлениях x , y , z . Квантовые точки особенно интересны тем, что их свойства подобны свойствам дискретного атома. Соответственно квантовую точку так и называют – искусственный атом. В квантовых точках энергетический спектр не содержит непрерывной компоненты, то есть не состоит из непрерывных подзон и является совершенно дискретным. В точности как в атоме, энергетический спектр квантовой точки описывается тремя квантовыми числами n , l , m (не учитывая спиновое квантовое число σ). Энергию можно записать в виде $\mathcal{E} = \mathcal{E}_{nlm}$, энергетические состояния, как и в атоме, могут быть вырождены и зависеть лишь от одного или двух квантовых чисел (рис. 3.7).

Квантовая
точка
(0D)

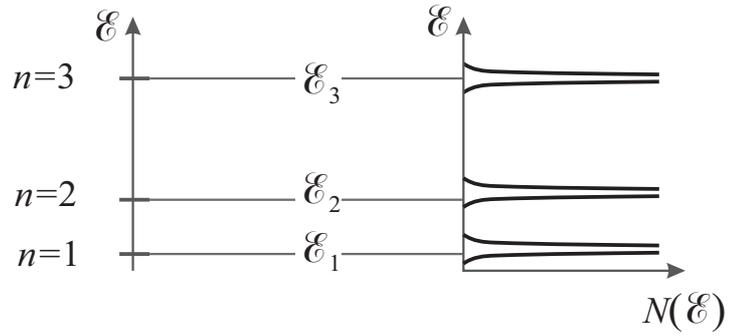
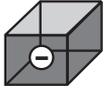


Рис. 3.7 – Нульмерная низкоразмерная структура, ее энергетическая диаграмма и плотность состояний $N(\mathcal{E})$

В соответствии с (3.5) можно записать минимальную энергию электрона в квантовой точке, имеющей форму куба с ребром a или сферы с радиусом a .

$$\mathcal{E}_{1,\text{куб}} = \frac{3\hbar^2\pi^2}{2m_n a^2} = \frac{3h^2}{8m_n a^2}. \quad (3.23)$$

Плотность состояний описывается δ -функциями и представляет собой набор острых пиков:

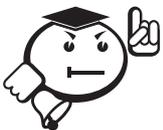
$$n_{0D}(\mathcal{E}) = \frac{2}{l_x l_y l_z} \sum_{i,j,k} \delta(\mathcal{E} - \mathcal{E}_{i,j,k}); \quad i, j, k = 1, 2, \dots, \quad (3.24)$$

$\mathcal{E}_{i,j,k}$ – разрешенные значения энергии, определяемые квантовым ограничением, соответственно, в направлениях x, y, z .

Для квантовой точки в виде сферы самое низкое разрешенное состояние для электронов в ней:

$$\mathcal{E}_{1,\text{сфера}} = \frac{h^2}{2m_n a^2}. \quad (3.25)$$

Из сравнения выражений (3.23) и (3.25) видно, что при одном и том же характерном размере a в квантовой сфере минимальная энергия больше, чем в кубе. Это объясняется меньшим объемом сферы по сравнению с объемом куба $\pi a^3 / 6 < a^3$. Поэтому условия квантового ограничения в сфере более жесткие, чем в кубе.



Три основных составляющих энергии определяют минимальную энергию для образования пар носителей заряда электрон –

дырка в квантовой точке. Первая компонента – это ширина запрещенной зоны полупроводника, из которого изготовлена квантовая точка E_g , вторая – это минимальная энергия, задаваемая квантовым ограничением (quantum confinement) для электронов и для дырок, третья – это энергия кулоновского взаимодействия электронов и дырок.

.....

Вторая составляющая записывается в виде:

$$\mathcal{E}_{conf} = \mathcal{E}_{conf}^{electr} + \mathcal{E}_{conf}^{hole} = \frac{\hbar^2}{2a^2} \left(\frac{1}{m_n} + \frac{1}{m_p} \right), \quad (3.26)$$

m_n и m_p – эффективные массы электронов и дырок.

Энергия кулоновского взаимодействия для сферической квантовой точки, изготовленной из материала с диэлектрической проницаемостью ϵ :

$$\mathcal{E}_{coule} = -\frac{1,8e^2}{2\pi\epsilon\epsilon_0 a}. \quad (3.27)$$

В итоге

$$\mathcal{E}_{min} = E_g + \mathcal{E}_{conf} + \mathcal{E}_{coule} = E_g + \frac{\hbar^2}{2a^2} \left(\frac{1}{m_n} + \frac{1}{m_p} \right) - \frac{1,8e^2}{2\pi\epsilon\epsilon_0 a}. \quad (3.28)$$

В рассмотренных элементарных низкоразмерных структурах (квантовые пленки, нити и точки) квантовое ограничение возникает, когда движение электронов в одном из направлений оказывается ограниченным потенциальными барьерами. Квантовое ограничение изменяет спектр разрешенных энергетических состояний и влияет на перенос носителей заряда через наноструктуры.

.....



Транспорт носителей заряда может осуществляться как параллельно, так и перпендикулярно потенциальным барьерам. При движении носителей заряда вдоль потенциальных барьеров преобладающими эффектами оказываются баллистический транспорт и квантовая интерференция. Прохождение носителей заряда через

потенциальные барьеры происходит исключительно посредством их туннелирования.

.....

3.7 Баллистический транспорт

Транспорт носителей заряда в низкоразмерных структурах показывает массу интересных явлений, совершенно новых в сравнении с поведением носителей заряда в макроскопических системах.

.....



*В наноструктурах с размером меньше длины свободного пробега перенос носителей заряда происходит без их рассеяния. Такой перенос называют **баллистическим транспортом** (ballistic transport). Здесь важно соотношение между размерами наноструктуры и характерными длинами свободного пробега для упругого и неупругого рассеяния носителей заряда.*

.....

Важную роль также играет соотношение размера наноструктуры и *длины фазовой когерентности* электрона. Это расстояние между двумя ближайшими точками, в которых происходит рассеяние электрона на примесных атомах, фоновых или других неоднородностях в материале, вследствие чего изменяется фаза электронной волны.

Средняя длина свободного пробега при неупругом рассеянии (inelastic mean free path) – это расстояние, на котором электронная волна изменяет свою фазу вследствие рассеяния. Она равна $l_{in} = v_F \tau_\varphi$, где v_F – скорость Ферми $v_F = \sqrt{2F / m_n}$, F – энергия Ферми; τ_φ – время релаксации фазы (или энергии). Также неупругое рассеяние носителей заряда характеризует *длина фазовой когерентности* (phase coherence length) $l_\varphi = \sqrt{D\tau_\varphi}$ – расстояние, на котором электронная волна сохраняет свою фазу (D – коэффициент диффузии).

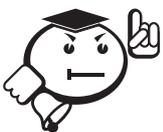
Следует иметь в виду, что l_{in} и l_{ϕ} представляют собой различные характеристики. При неупругом рассеянии $l_{in} > l_{\phi}$. Оба параметра важны при анализе условий интерференции электронных волн.

Средняя длина свободного пробега при упругом рассеянии (elastic mean free path) – это среднее расстояние l , которое проходит носитель заряда между двумя последовательными актами упругого рассеяния: $l = v_F \tau_{sc}$, где τ_{sc} – время рассеяния (scattering). $\tau_{sc} = Dd / v_F^2$, где d – размерность структуры.

Важной характеристикой наноструктур является *длина волны Ферми* (Fermi wavelength) $\lambda_F = 2\pi / k_F$, где k_F – волновой вектор Ферми. При $T = 0$ электроны находятся в состояниях, когда их волновой вектор по модулю меньше волнового вектора Ферми: $|\mathbf{k}| < k_F$, что эквивалентно электронным волнам с длиной волны $\lambda > \lambda_F$.

Сравнивая размер наноструктуры со средней длиной свободного пробега электронов и с длиной волны Ферми (являющихся характеристиками материала, из которого данная структура изготовлена), можно предсказать основные особенности движения носителей заряда в этой наноструктуре.

Транспорт носителей заряда в полупроводниках характеризуется большой, до нескольких микрометров, *средней* длиной свободного пробега электронов. При комнатной температуре l_{in} достигает 50–100 нм в Si и около 120 нм в GaAs. Очевидно, что в наноструктурах из полупроводников баллистический транспорт легко реализуется. Более того, длина волны Ферми в полупроводниках достигает 30–50 нм. Когда размер структуры сравнивается с этими длинами, квантование энергии вследствие квантового ограничения становится существенным фактором, определяющим ее электронные свойства и транспорт носителей заряда.



Идеальный баллистический транспорт носителей заряда в наноструктурах характеризуется универсальной баллистической проводимостью (universal ballistic conductance), которая не зависит

от типа материала и определяется только фундаментальными константами.

.....

Проводимость идеального одномерного проводника в баллистическом режиме описывается выражением:

$$G = I / (V_1 - V_2) = \frac{e^2}{\pi\hbar} = \frac{2e^2}{h}. \quad (3.29)$$

Она определяется только фундаментальными константами: зарядом электрона и постоянной Планка.



.....

Величину $e^2/h = 38,740$ мкСм называют квантовой единицей проводимости (quantum conductance unit), или квантом проводимости. Соответствующее сопротивление равно $h/e^2 = 25812,807$ Ом – это квант сопротивления.

.....

Приведенные рассуждения можно обобщить на случай, когда проводящий канал имеет более одного энергетического состояния ниже уровня Ферми. Для канала с N разрешенными состояниями получаем:

$$G = N \frac{2e^2}{h}, \quad (3.30)$$

то есть проводимость канала квантуется в единицах $\frac{2e^2}{h}$.

Как только движение электрона становится когерентным, то есть электрон проходит через всю структуру без рассеяния, его волновая функция не меняется и сохраняет свою фазу. При этом возникают разнообразные интерференционные эффекты, представляющие особенности транспорта носителей заряда вдоль потенциальных барьеров.

3.8 Туннелирование

Термин «туннелирование» (tunneling) означает перенос частицы через область, ограниченную потенциальным барьером, высота которого больше полной энергии данной частицы (или проникновение частицы в эту область). Такой эффект невозможен с точки зрения классической механики, однако имеет место для квантовых частиц, которым, как известно, присущ корпускулярно-волновой дуализм. Волновые свойства квантовых частиц приводят и к другому, аномальному с точки зрения классической механики, эффекту – надбарьерному отражению. Взаимодействие квантовых частиц с различными потенциальными барьерами показано на рисунке 3.8.

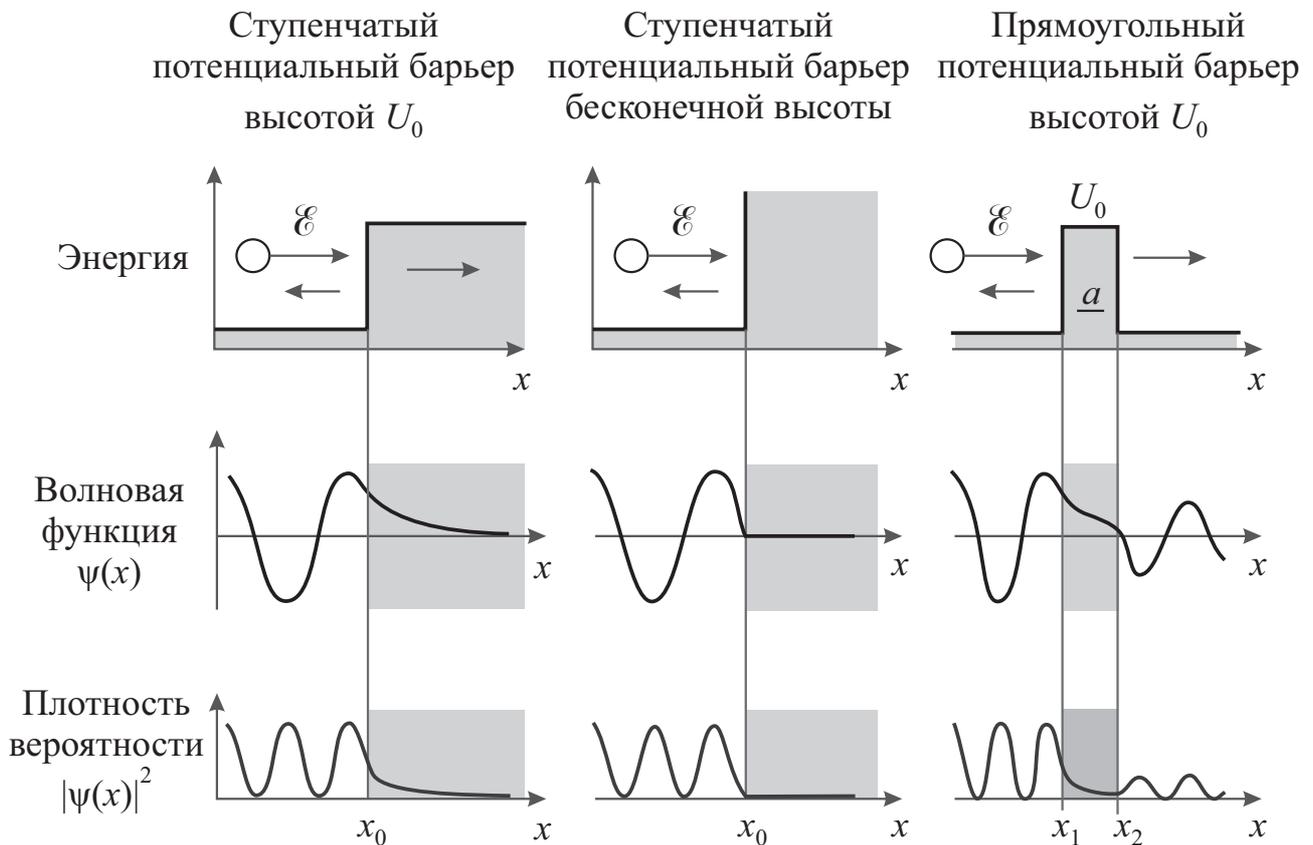


Рис. 3.8 – Взаимодействие квантовой частицы с полной энергией \mathcal{E} со ступенчатым барьером конечной высоты U_0 , с бесконечно высоким потенциальным барьером и с прямоугольным барьером высотой U_0 и шириной a

Движение частицы в одномерном потенциальном поле $U(x)$ описывается уравнением Шредингера вида (3.5.2) из модели Кронига – Пенни (п. 3.2). Рас-

смотрим случай, когда энергия квантовой частицы больше высоты барьера $\mathcal{E} > U_0$.

В области перед потенциальным барьером, где $U(x) = 0$, решение уравнения Шредингера для частицы с импульсом $k_1 = \frac{1}{\hbar} \sqrt{2m_n \mathcal{E}}$ имеет вид суперпозиции двух волн:

$$\psi = A' \exp(ik_1 x) + B \exp(-ik_1 x), \quad (3.31)$$

где A' и B – константы. Первый член в (3.31) соответствует падающей на барьер волне, движущейся вдоль оси x слева направо. Второй член описывает отраженную волну, движущуюся вдоль оси x в противоположном направлении, справа налево.

При $x \rightarrow \infty$ волновая функция прошедшей над потенциальным барьером частицы имеет вид:

$$\psi = A \exp(ik_2 x), \quad (3.32)$$

где $k_2 = \frac{1}{\hbar} \sqrt{2m_n (\mathcal{E} - U_0)}$. Плотность потока в падающей волне пропорциональна k_1 , в отраженной – $k_1 |B|^2$, а в прошедшей – $k_2 |A|^2$.

Коэффициент прохождения частицы через границу потенциального барьера, определяемый как отношение плотности потока в прошедшей волне к плотности потока в падающей волне, равен

$$T(\mathcal{E}) = \frac{k_2}{k_1} |A|^2. \quad (3.33)$$

Его также называют коэффициентом туннельной прозрачности барьера. Коэффициент отражения частицы от потенциального барьера как отношение плотности отраженного потока к плотности падающего потока равен

$$R(\mathcal{E}) = |B|^2 = 1 - T(\mathcal{E}) = 1 - \frac{k_2}{k_1} |A|^2. \quad (3.34)$$

Постоянные A и B , определенные из условия непрерывности волновой функции и ее первой производной при $x = x_0$, равны

$$A = \frac{2k_1}{k_1 + k_2}, \quad B = \frac{k_1 - k_2}{k_1 + k_2}. \quad (3.35)$$

Тогда коэффициенты отражения и прохождения имеют вид:

$$R(\mathcal{E}) = \left(\frac{k_1 - k_2}{k_1 + k_2} \right)^2, \quad T(\mathcal{E}) = \frac{4k_1 k_2}{(k_1 + k_2)^2}, \quad (3.36)$$

откуда следует, что при $\mathcal{E} = U_0$ ($k_2 = 0$) коэффициент отражения $R = 1$, а $T = 0$.

С ростом энергии частицы коэффициент отражения уменьшается и стремится

к нулю при $\mathcal{E} \rightarrow \infty$, так как $R \approx \left(\frac{U_0}{4\mathcal{E}} \right)^2$ при $\mathcal{E} \gg U_0$.

Если квантовая частица движется над прямоугольным потенциальным барьером высотой U_0 и конечной шириной $a = x_2 - x_1$, то максимум коэффициента прохождения, $T(\mathcal{E}) = 1$ достигается для частиц с энергиями, соответствующими решению уравнения Шредингера для частиц в прямоугольной потенциальной яме (п. 3.4, выражение (3.11)) плюс высота U_0 :

$$\mathcal{E} = U_0 + \frac{\pi^2 \hbar^2}{2m_n a^2} n^2, \quad (3.37)$$

где $n = 0, 1, 2, \dots$.

При других значениях энергии наблюдается частичное отражение падающих на барьер частиц. Таким образом, из квантовой теории следует, что даже когда энергия падающей на потенциальный барьер частицы больше высоты этого барьера, коэффициент ее отражения может быть отличен от нуля. Этим квантовая частица отличается от классической, для которой никакого отражения в подобной ситуации быть не может.

Существование прошедшей через барьер волны, отвечающей квантовой частице с энергией меньше высоты барьера, называют *туннельным эффектом*.

Туннелирование электронов характеризуется временами порядка $10^{-13} \div 10^{-15}$ с. Теоретически Т. Хартман нашел выражение, отражающее взаимосвязь времени туннелирования электронов с энергией \mathcal{E} через прямоугольный потенциальный барьер высотой U_0 :

$$\tau = \frac{\hbar}{\sqrt{\mathcal{E}(U_0 - \mathcal{E})}}. \quad (3.38)$$

Время туннелирования зависит только от энергии частицы и высоты барьера. Минимальное значение времени, равное $\frac{\hbar}{\mathcal{E}}$, достигается при $\mathcal{E} = 0,5U_0$. Время туннелирования не зависит ни от массы частицы, ни от ширины потенциального барьера, и, туннелируя через «толстые» барьеры, квантовые частицы могли бы достигать сверхсветовых значений скорости движения. Это следствие называют «парадоксом Хартмана», пока оно не получило всестороннего понимания [67].



Контрольные вопросы по главе 3

1. Что такое наноразмерная структура?
2. Какой критерий используется для классификации элементарных низкоразмерных структур?
3. Что такое баллистический транспорт носителей заряда?
4. Какими параметрами характеризуют транспорт носителей заряда в твердотельных структурах?
5. Что такое туннелирование квантовой частицы?

4 Полупроводниковые гетероструктуры и сверхрешетки



.....

Наш замечательный соотечественник, ученый с мировым именем, Ж. И. Алферов внес выдающийся вклад мирового уровня в развитие физики и техники полупроводниковых гетероструктур. Мировым признанием работ Алферова стало присуждение ему в 2000 г. Нобелевской премии «за исследование полупроводниковых гетероструктур, которые нашли применение в высокочастотной и оптической электронике». В результате работы Алферова появилось новое направление – физика гетероструктур, а также оптоэлектроника и электроника на их основе. Обобщение этого пути Алферов сделал в работе [68], материалы которой послужили основой для Нобелевской лекции Алферова [69].

.....

В настоящее время очень трудно представить современную физику твердого тела без полупроводниковых гетероструктур. Полупроводниковые гетероструктуры и особенно двойные гетероструктуры (ДГС), включая квантовые ямы, нити и точки, являются сегодня предметом изучения 2/3 исследовательских групп в области физики полупроводников.

Гетероструктуры позволяют решить проблему управления фундаментальными параметрами в полупроводниковых кристаллах и приборах: шириной запрещенной зоны, эффективными массами носителей и их подвижностями, показателем преломления, электронным энергетическим спектром и так далее [68].

4.1 Гетеропереходы и гетероструктуры

Основным элементом гетероструктур различного типа является гетеропереход.



.....

Гетеропереходы представляют собой контакт двух разнородных полупроводников с различной шириной запрещенной зоны и различной степенью легирования. Это контакт двух различных по химическому составу или (и) фазовому состоянию полупроводников, при котором кристаллическая решетка одного материала без нарушения периодичности переходит в решетку другого материала.

.....

Различают изотипные и анизотипные гетеропереходы. Изотипный гетеропереход образован двумя полупроводниками одного типа проводимости. Анизотипный гетеропереход образуется полупроводниками с разным типом проводимости. Примеры контактов, используемых для образования гетероперехода: Ge – Si, GaAlAs – GaAs, GaAs – Ge, InGaAs – InP.

Существуют три модели гетероперехода: 1) идеальный гетеропереход; 2) неидеальный гетеропереход; 3) гетеропереход с промежуточным слоем. В идеальном гетеропереходе на границе раздела материалов отсутствуют локальные энергетические состояния для электронов, это и отличает его от неидеального. Гетеропереход с промежуточным слоем формируется через слой конечной толщины и локальные энергетические состояния могут существовать как в самом промежуточном слое, так и на границах его раздела.



.....

Главная особенность гетероперехода по сравнению с гомопереходом (обычным *p-n*-переходом) состоит в скачкообразном изменении каких-либо свойств на границе раздела полупроводников: ширины запрещенной зоны, подвижности носителей заряда, их эффективной массы, энергии сродства к электрону и других.

.....

Наибольший интерес представляют гетеропереходы, свойства которых обусловлены различиями в значениях ширины запрещенной зоны и энергии сродства к электрону контактирующих полупроводников. Энергетическая диаграмма таких гетеропереходов содержит скачки энергии в зоне проводимости и в валентной зоне, так что высота потенциального барьера для электронов и дырок оказывается различной. Эта особенность гетеропереходов позволяет изменять соотношения между потоками носителей заряда над барьерами (или сквозь барьеры), например создавать практически одностороннюю инжекцию носителей заряда, в отличие от *p-n*-гомоперехода, где основную роль играет различие в концентрациях примеси.



.....

Необходимо, чтобы у материалов гетеропереходов с высокой точностью совпадали температурный коэффициент расширения и постоянная решетки. Поэтому количество материалов для гетеропереходов ограничено. Наиболее распространенными из этих пар являются германий (Ge), арсенид галлия (GaAs), фосфид индия (InP), трехкомпонентный раствор (AlGaAs), четырехкомпонентный раствор (InGaAsP).

.....

Во избежание дефектов в кристаллической решетке двух материалов, составляющих гетеропереход, необходимо, как минимум, чтобы два материала имели одну и ту же кристаллическую структуру и близкие периоды решеток. В этом случае структура получается без напряжений. Только некоторые материалы могут удовлетворять этим требованиям для создания гетеропереходов. На рисунке 4.1 приведены данные из книги [70] для полупроводников с близкими значениями постоянных кристаллических решеток, отрицательное значение, приведенное для энергии запрещенной зоны HgSe, является спорным, что обсуждается в работах [71–73].

Используя данные рисунка 4.1 можно изготавливать требуемые гетеропереходы с заданными свойствами.

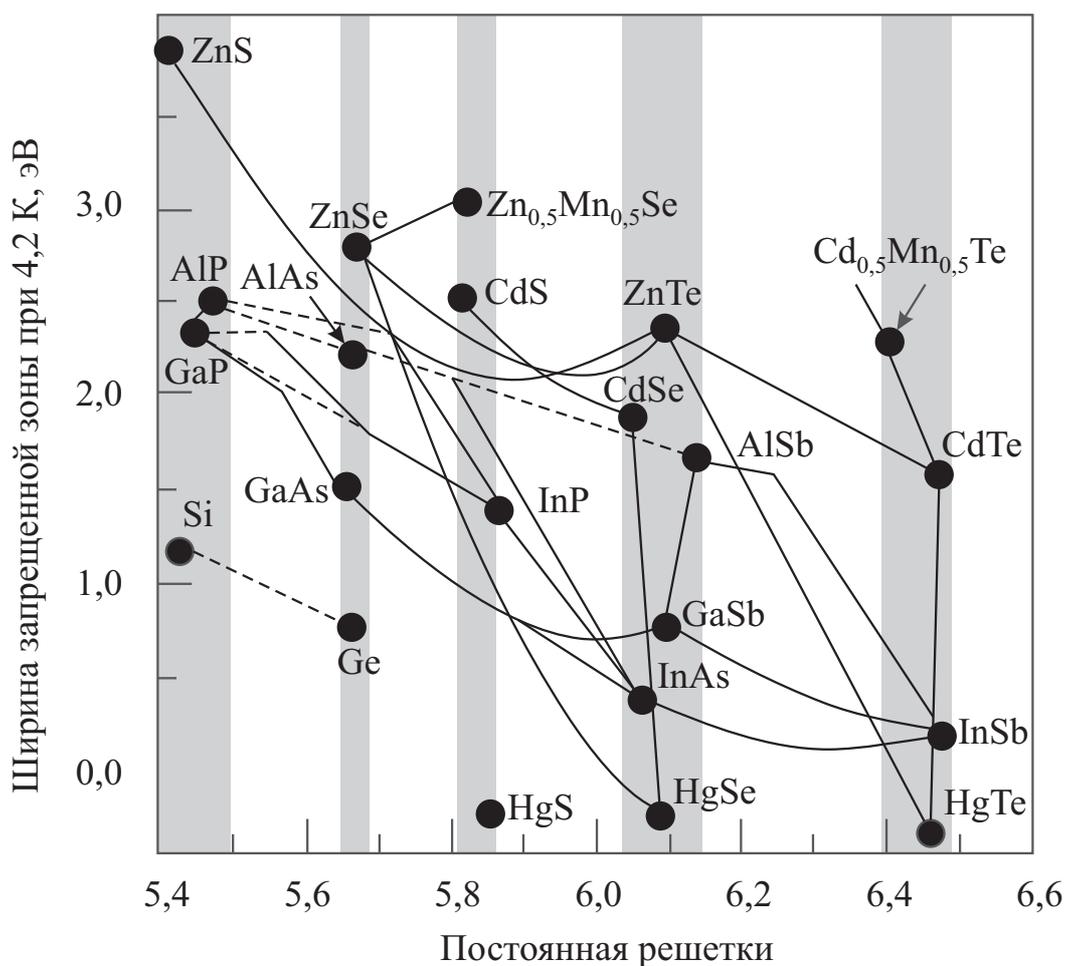


Рис. 4.1 – График зависимости энергии запрещенной зоны при низкой температуре от постоянной решетки для ряда полупроводников со структурой алмаза и цинковой обманки. Затененные области объединяют группы полупроводников с близкими значениями постоянных кристаллических решеток. Полупроводники, соединенные сплошными линиями, образуют между собой стабильные твердые растворы. Штриховые линии указывают на непрямые запрещенные зоны



Основным методом создания гетеропереходов является эпитаксия. При условии высокой степени совпадения типа, ориентации и периода кристаллических решеток сопрягаемых полупроводниковых материалов метод эпитаксии позволяет получать гетеропереходы с границей раздела, практически свободной от структурных и других дефектов (дислокаций, точечных дефектов и т. п.).

Обозначим дно зоны проводимости первого и второго полупроводника E_{c1} и E_{c2} соответственно. Из-за различия электронного сродства χ_1 и χ_2 (расстояние от дна зоны проводимости до уровня вакуума) в контактирующих полупроводниках E_{c1} выходит на плоскость контакта в точке, не совпадающей с точкой выхода на эту плоскость E_{c2} , и формируется разрыв зоны проводимости ΔE_c (рис. 4.2).

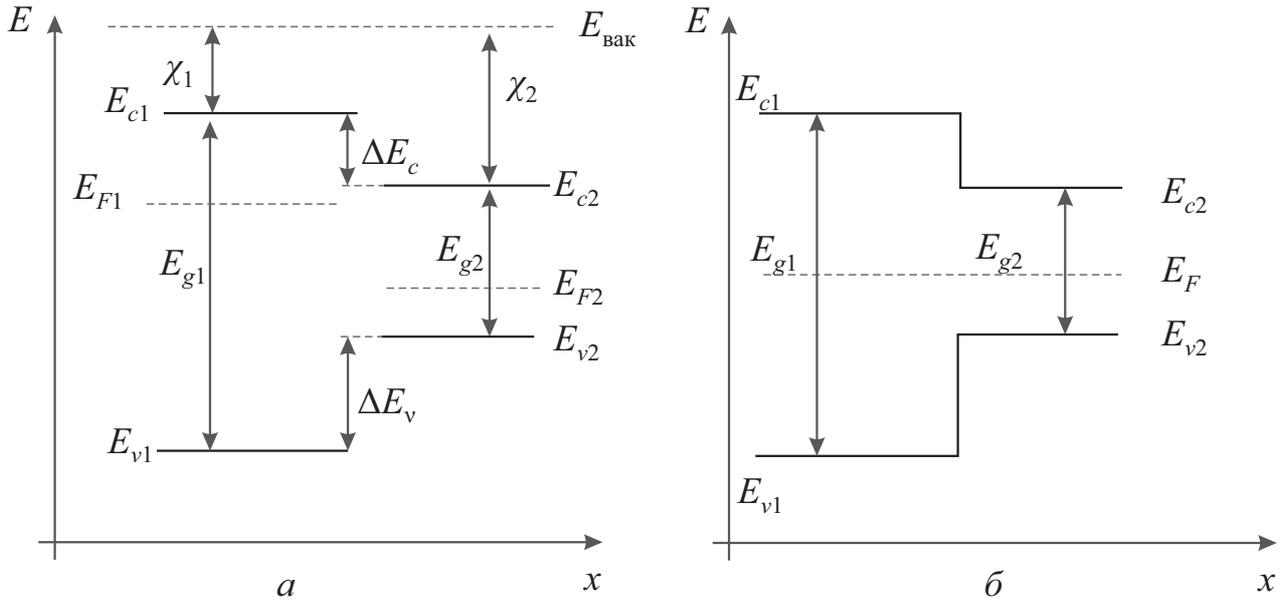


Рис. 4.2 – энергетическая диаграмма при формировании гетероперехода:
а) до контакта и б) после контакта

Также формируется разрыв валентной зоны ΔE_v (рис. 4.2). Величины этих разрывов обычно различны:

$$\Delta E_c = \chi_1 - \chi_2, \quad (4.1)$$

$$\Delta E_v = E_{g2} - E_{g1} - \Delta E_c. \quad (4.2)$$

Разрывы зон могут быть положительными или отрицательными. Можно выделить следующие разновидности гетеропереходов в соответствии со знаками разрывов зон:

- 1) гетеропереход I типа (стандартный) или охватывающий переход: $\Delta E_c > 0$ и $\Delta E_v > 0$. Пример: состав GaAs – AlGaAs, показан на рисунке 4.2, б;

- 2) гетеропереход II типа (ступенчатый): $\Delta E_c > 0$, $\Delta E_v < 0$ или наоборот $\Delta E_c < 0$, $\Delta E_v > 0$. Пример: InP – In_{0,52}Al_{0,48}As, показан на рисунке 4.3;

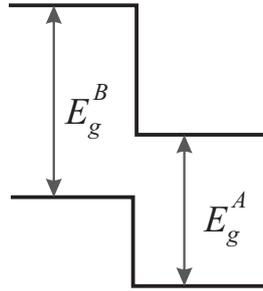


Рис. 4.3 – Гетеропереход II типа

- 3) гетеропереход III типа (разрывный), когда запрещенные зоны вообще не перекрываются по энергии. Пример: InAs – GaSb.

Разрыв зоны проводимости G также применяют как параметр, описывающий гетеропереход. Это доля разрыва, приходящаяся на зону проводимости:

$$G = \frac{\Delta E_c}{\Delta E_g} = \frac{\Delta E_c}{E_{g2} - E_{g1}}. \quad (4.3)$$



.....

Свойства тройных и четверных соединений можно моделировать с помощью эмпирического закона (правила) Вегарда, установленного в 1921 г. [74]. Правило говорит, что между концентрацией отдельных элементов и свойствами кристаллической решетки сплава при постоянной температуре существует линейная связь. Это правило распространяется также и на определение запрещенной зоны.

.....

Пользуясь правилом Вегарда можно подобрать состав тройного соединения, при котором будут совпадать параметры решеток, но зонные параметры будут различны. Таким образом, можно получать гетеропереходы с заданными зонными параметрами. Так, на подложке GaAs может быть выращен без значительных напряжений AlAs, а также твердый раствор Al_xGa_{1-x}As практически любого состава. У твердого раствора Al_xGa_{1-x}As период решетки меняется ме-

нее чем на 0,15% при изменении x от 0 до 1. Существует очень мало материалов, которые образуют гетеропереход с GaAs без напряжений. Альтернативой подложке из арсенида галлия является фосфид индия InP. Но на подложке из InP без напряжений можно вырастить только два твердых раствора: $\text{Al}_{0,48}\text{In}_{0,52}\text{As}$ и $\text{Ga}_{0,47}\text{In}_{0,53}\text{As}$ [75, 76].

Полупроводниковая система с различными гетеропереходами называется *гетероструктурой*. На рисунке 4.4 показана энергетическая диаграмма двойного гетероперехода. Двойной гетеропереход первого рода $B/A/B$ представляет гетероструктуру с одиночной квантовой ямой, если ширина запрещенной зоны в материале A меньше, чем в материале B , то есть $E_g^A < E_g^B$ или структуру с одиночным барьером, если $E_g^A > E_g^B$. В первом случае внутренний слой A образует потенциальную яму (рис. 4.4, *а*), в которой происходит размерное квантование электронных и дырочных состояний. Во втором случае слой A образует потенциальный барьер для электронов и дырок (рис. 4.4, *б*). На рисунке 4.4, *а* схематически показана структура с прямоугольной ямой. Используя в качестве композиционного материала A твердый раствор и изменяя его состав в процессе роста, можно создавать ямы другой формы – параболические, треугольные и т. п.

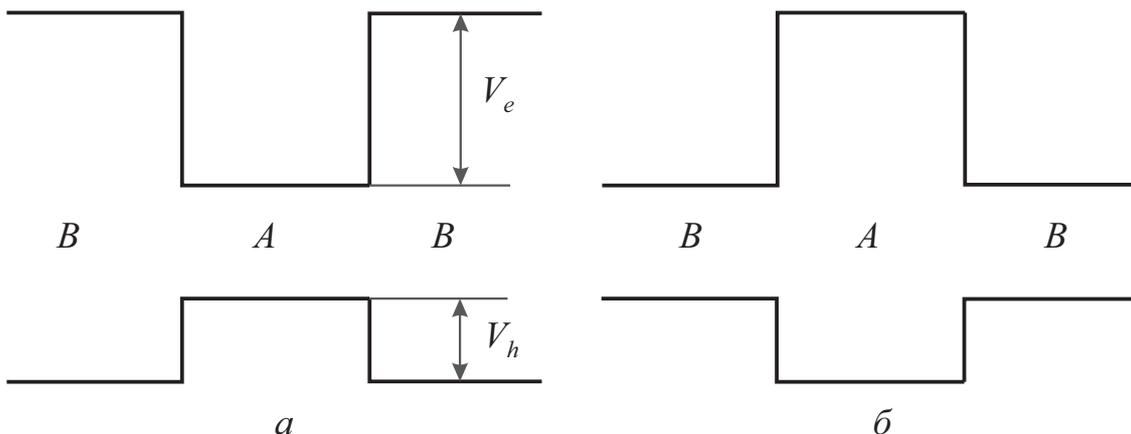


Рис. 4.4 – Зонная схема двойного гетероперехода первого рода с квантовой ямой *а* или барьером *б*; $V_{e,h}$ – высота потенциального барьера (глубина квантовой ямы) или разрыв зон на интерфейсе в зоне проводимости и валентной зоне соответственно [76]

В соответствии с материалом главы 3, п. 3.4, в прямоугольной потенциальной яме электрон в ограниченной области пространства, размеры которой сравнимы с его длиной волны (длиной волны де Бройля) может занимать только дискретные разрешенные уровни, энергия которых определяется выражением (3.11) $\mathcal{E} = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a} \right)^2$. Меняя ширину потенциальной ямы a можно управлять энергетическим расстоянием между разрешенными уровнями и, соответственно, менять длину волны поглощаемого или испускаемого излучения. Оценка показывает, что эффекты квантового ограничения проявляются в различных материалах при значениях a от нескольких до 100 параметров решетки или от 1 до 50 нм. Ширина потенциальной ямы a должна быть не больше волны де Бройля, что для арсенида галлия составляет величину приблизительно 14 нм.

Скачкообразное изменение свойств гетеропереходов и возможность целенаправленного управления этими свойствами подбором сопрягаемых полупроводниковых материалов открывает широкие возможности для совершенствования на основе гетеропереходов традиционных полупроводниковых приборов (диодов, транзисторов, фотоэлементов, светоизлучающих диодов и т. д.), создания оригинальных гетеропереходных полупроводниковых приборов, а также многослойных гетероструктур типа *сверхрешеток* с толщиной слоев менее 10 нм.

4.2 Сверхрешетки

Когда в практических структурах содержится большое количество одинаковых или почти одинаковых квантовых объектов, то обычно это не меняет физической картины, так как вклады от всех объектов просто суммируются. Ситуация резко меняется, если отдельные объекты находятся так близко друг к другу, что носители заряда могут туннелировать между ними.



.....

Периодическая структура с квантовыми ямами, разделенными не очень широкими потенциальными барьерами, образует

сверхрешетку, в которой носители заряда могут туннелировать из ямы в яму и длина свободного пробега таких носителей заряда вдоль оси, перпендикулярной плоскостям слоев, превышает период структуры.

.....

Полупроводниковые сверхрешетки состоят из чередующихся слоев двух полупроводников, различающихся или составом, или типом проводимости. Только с развитием новых методов выращивания гетероструктур в 1970-е гг. (см. главу 1) стала возможной реализация высококачественных двойных гетероструктур со сверхтонкими слоями. Получают сверхрешетки, например, с помощью технологии МПЭ, позволяющей наращивать чередующиеся слои любого состава и толщины. Период повторения слоев составляет от нескольких нанометров до десятков нанометров (для сравнения – постоянные решетки кристаллов Si и GaAs равны примерно 0,5 нм). Широко используют два типа полупроводниковых сверхрешеток: композиционные и легированные [77, 78]. Композиционные сверхрешетки – это гетероструктуры из чередующихся слоев различного состава и ширины запрещенной зоны, но с близкими значениями постоянной решетки. Например, $\text{Al}_x\text{Ga}_{x-1}\text{As} - \text{GaAs}$; $\text{In}_x\text{Ga}_{x-1}\text{As} - \text{GaAs}$; $\text{ZnS} - \text{ZnSe}$; $\text{In}_x\text{Ga}_{x-1}\text{As} - \text{InP}$ и др. Здесь дополнительный периодический потенциал создается за счет периодического изменения ширины запрещенной зоны. Легированные сверхрешетки – это периодическая последовательность слоев n - и p -типа одного и того же полупроводника. Донорные атомы в n -слоях отдают электроны, которые связываются акцепторными атомами в p -слоях. Дополнительный периодический потенциал создают чередующиеся заряды ионизированных акцепторов и доноров. Существуют также сверхрешетки из металлов, сверхпроводников и диэлектриков. Дополнительный периодический потенциал сверхрешетки изменяет зонную структуру исходных полупроводников. Поэтому сверхрешетку можно рассматривать как новый, синтезированный полупроводник, не существующий в природе и обладающий необычными свойствами.

Подбором материала и состава чередующихся слоев можно в широких пределах варьировать зонную структуру сверхрешетки.

На рисунке 4.5, *a* показана энергетическая диаграмма (потенциальный профиль) композиционной сверхрешетки $\text{Al}_x\text{Ga}_{x-1}\text{As} - \text{GaAs}$ в направлении, перпендикулярном слоям. Вследствие периодического изменения ширины запрещенной зоны $\Delta E_g = E_{c1,2} - E_{v1,2}$ создается последовательность прямоугольных квантовых ям, разделенных барьерами. Ямы образуются в узкозонном полупроводнике: для электронов – в зоне проводимости, для дырок – в валентной зоне [77, 78].

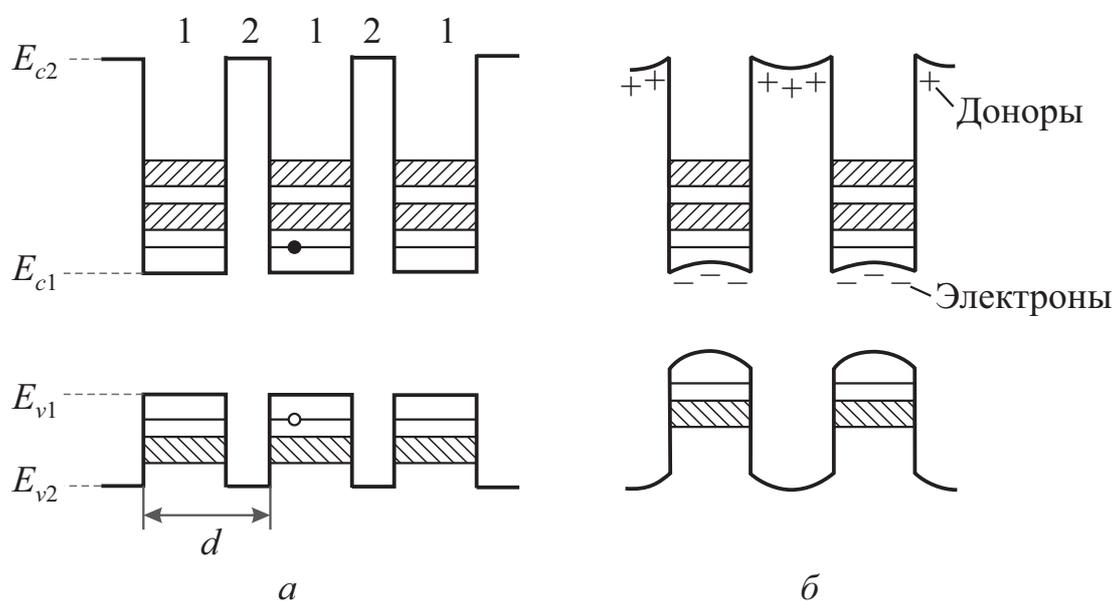


Рис. 4.5 – Энергетические диаграммы простой композиционной (*a*) и модулированно-легированной (*б*) сверхрешеток; d – период решетки

На рисунке 4.5, *б* показан потенциальный профиль модулированно-легированной композиционной сверхрешетки. Здесь легируется донорной примесью только широкозонный материал. Электроны с донорных уровней переходят в квантовые ямы, пространственно разделяясь с ионизованными донорами. Чередование зарядов вызывает периодические изгибы краев зон. На рисунке 4.5, *a*, *б* штриховкой показаны мини-зоны, на которые разбивается валентная зона и зона проводимости. На рисунке 4.6 показан потенциальный профиль легированной сверхрешетки [77, 78].

Чередующиеся заряды ионизированных доноров и акцепторов создают последовательность потенциальных ям для электронов и дырок. Электроны и дырки оказываются пространственно разделенными: дырки находятся в потенциальных ямах валентной зоны p -слоя, электроны – в потенциальных ямах зоны проводимости n -слоя. Штриховкой показаны мини-зоны; ΔE_g – ширина запрещенной зоны исходного полупроводника, $\Delta E_{g\text{эфф}}$ – эффективная ширина запрещенной зоны сверхрешетки. Для получения легированных сверхрешеток часто используют GaAs.

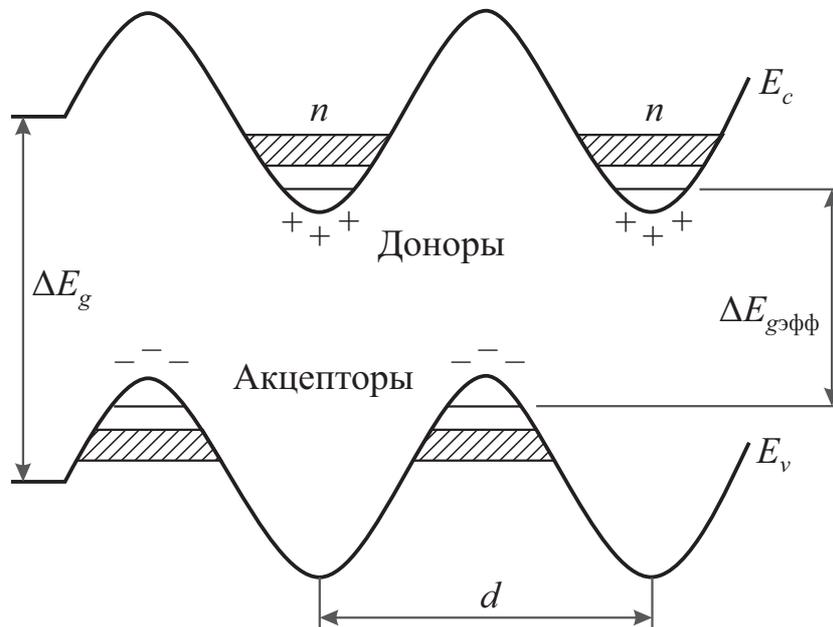
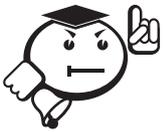


Рис. 4.6 – Энергетические диаграммы легированной сверхрешетки: $\Delta E_{g\text{эфф}}$ – эффективная ширина запрещенной зоны сверхрешетки; d – ее период

Энергетический спектр электронов в сверхрешетках находится из решения уравнение Шредингера в модели Кронига – Пенни, как и в случае массивного кристалла, но с учетом дополнительного периодического потенциала. Используя результаты расчета зонной структуры кристалла, можно сделать 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. Так как период сверхрешетки d значительно больше постоянной

решетки, то получающиеся сверхрешеточные зоны представляют собой более мелкое дробление энергетических зон исходных полупроводниковых кристаллов и называются *мини-зонами*. Расщепление зоны проводимости и валентной зоны на мини-зоны показано штриховкой на рисунках 4.5 и 4.6. Чем меньше ширина ямы, тем больше расстояние между мини-зонами и больше эффективная ширина запрещенной зоны $\Delta E_{g\text{эфф}}$; чем меньше ширина барьера, тем шире мини-зоны. Таким образом, можно перестраивать энергетический спектр сверхрешетки простым изменением толщины слоев, что легко осуществить в методе МПЭ. График плотности состояний $N(E)$ имеет ступенчатый вид, как и $N(E)$ для квантовой ямы, но с иной формой ступеней.



.....

Электронный газ в сверхрешетках приобретает оригинальные свойства. В легированных сверхрешетках электроны и дырки пространственно разделены (рис. 4.5). Неравновесные носители, например генерируемые светом пары электрон – дырка, также пространственно разделяются, что подавляет рекомбинацию и увеличивает их время жизни до $\sim 10^3$ с.

.....

В модулированно-легированных композиционных сверхрешетках широкозонный полупроводник (например, AlGaAs) легируется донорной примесью. Электроны с донорных уровней барьера переходят в ямы зоны проводимости узкозонного полупроводника, например GaAs (см. рис. 4.4, б). В барьерах остаются ионизованные доноры (примесные центры), в ямах образуется 2D электронный газ с высокой плотностью и подвижностью электронов. Высокая подвижность обусловлена тем, что плотность электронов в узкозонном слое больше плотности центров рассеяния, а донорные примесные центры заключены в широкозонных слоях. Эффект увеличения подвижности особенно значителен при низких температурах, когда главный вклад в рассеяние движущихся электронов вносит их рассеяние на примесях. Высокая подвижность электронов позволяет создавать на сверхрешетках быстродействующие приборы, например

транзисторы с проводящими каналами, параллельными слоям. Время переключения таких транзисторов может составлять пикосекунды.



.....
Контрольные вопросы по главе 4
.....

1. Что такое гетеропереход? Назовите три модели гетероперехода.
2. Какая технология в основном применяется для создания гетеропереходов?
3. Что такое гетероструктура?
4. Как формулируется и для чего используется правило Вегарда?
5. Что такое сверхрешетка, композиционная сверхрешетка и легированная сверхрешетка?

5 Технологии формирования квантоворазмерных наноструктур

Прогресс в экспериментальном изучении физических свойств квантово-размерных структур тесно связан с развитием технологии их получения и в значительной мере им и определяется. Изготовление квантово-размерных структур в большинстве случаев требует создания полупроводниковых гетеропереходов с необходимыми свойствами. Для этого прежде всего необходимо подобрать подходящую пару полупроводниковых материалов. Выбор этих материалов зависит от типа структуры, которую необходимо получить.

Для создания структур с двумерным электронным газом на основе гетеропереходов основным условием является требование равенства постоянных решетки у обоих полупроводников. Нарушение этого условия может привести к образованию высокой плотности дислокаций несоответствия вблизи гетерограницы, что резко ухудшает свойства переходов и делает невозможным наблюдение эффектов размерного квантования. Гетеропереход GaAs – Al_xGa_{1-x}As наиболее часто используют для изготовления квантово-размерных структур [77, 78]. Эти материалы обладают хорошим согласием решеток при любом составе твердого раствора и поэтому позволяют создавать гетеропереходы с различными разрывами зон на границе. Среди полупроводников класса A^{III}B^V есть еще несколько идеальных гетеропар, например InP – In_{0.53}Ga_{0.47}As, но они обладают согласием решеток лишь при одном фиксированном составе и потому не позволяют варьировать ΔE_c и ΔE_v . Другие классы полупроводников реже используются для изготовления квантовых гетероструктур.

Одно из главных требований к технологии изготовления квантовых гетероструктур связано с необходимостью получения очень резких гетеропереходов с переходным слоем промежуточного состава, имеющим толщину всего в несколько постоянных решетки. Как видно из рисунка 5.1, для плавного перехода образуется потенциальная яма большой ширины, что уменьшает расстоя-

ние между уровнями E_n и затрудняет наблюдение квантовых размерных эффектов [77, 78].

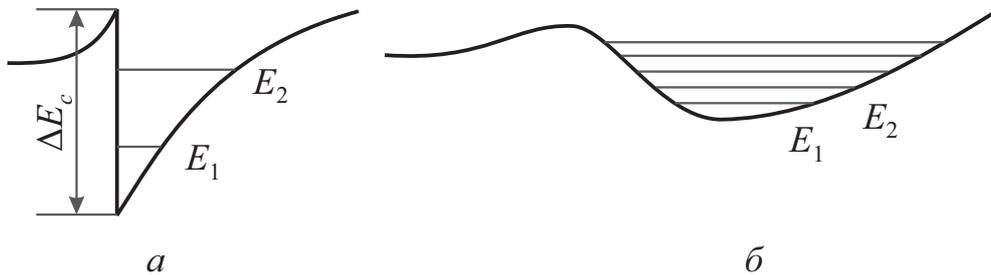


Рис. 5.1 – Потенциальный профиль края зоны проводимости для резкого (*a*) и плавного (*б*) гетеропереходов

Другим важным требованием к технологии является возможность получения сложных профилей состава и легирования. При изготовлении сверхрешеток необходимо строго периодически менять состав растущего слоя, причем период изменения может составлять лишь несколько десятков ангстрем. При выращивании 5 слоев и структур с модулированным легированием необходимо со столь же высокой точностью менять концентрацию легирующих примесей.

Далеко не всякая эпитаксиальная технология удовлетворяет приведенным требованиям. Плохо пригодными, в частности, оказываются наиболее распространенные методики газотранспортной и жидкостной эпитаксии. Получение сложного профиля состава и легирования наталкивается на ряд технических трудностей, а сравнительно высокая температура роста вызывает диффузионное размывание гетерограницы, не позволяя получать очень резкие границы. В отличие от МПЭ газовая эпитаксия из металлоорганических соединений не требует столь сложного и дорогого оборудования, обладает меньшими возможностями контроля и управления, но, тем не менее, позволяет растить гетеропереходы требуемой резкости и структуры достаточно сложного профиля, включая сверхрешетки, квантовые нити и точки. Получение одно- и нульмерных структур возможно путем выделения их с помощью литографии. Однако это требует применения особых методов литографии – нанолитографии, поскольку столь малые структуры трудно изготовить при помощи стандартных методов.

Значительные успехи были достигнуты в изготовлении квантовых точек и квантовых нитей с использованием эффектов самоорганизации упорядоченных наноструктур на полупроводниковых подложках. Этот метод предъявляет специфические требования к материалам гетеропары [77, 78].

5.1 Формирование квантовых ям

Для формирования квантовых ям используют молекулярно-пучковую (лучевую) эпитаксию (МЛЭ, от англ. *MBE – Molecular Beam Epitaxy*). Эпитаксия – это закономерное нарастание одного кристаллического материала на другом (от гр. *επι* – «на» и *ταξις* – «упорядоченность»), то есть ориентированный рост одного кристалла на поверхности другого (подложки). Материал подложки в этом процессе играет роль затравочного кристалла.



.....

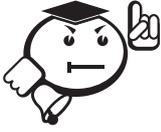
*Если материалы получаемого слоя и подложки идентичны, например когда GaAs выращивают на GaAs, то процесс называют **автоэпитаксиальным**, или **гомозэпитаксиальным**. Если же материалы слоя и подложки различаются (хотя их кристаллическая структура должна быть сходной для обеспечения роста монокристаллического слоя), то процесс называют **гетерозэпитаксиальным**.*

.....

С помощью эпитаксиальных процессов получают пленки металлов, оксидов, органических соединений и полупроводников. В основе метода лежит осаждение испаренного в молекулярном источнике вещества на кристаллическую подложку. Несмотря на достаточно простую идею, реализация данной технологии требует чрезвычайно сложных технических решений. Основные требования к установке эпитаксии:

- В рабочей камере установки необходимо поддерживать сверхвысокий вакуум (около 10^{-8} Па).
- Чистота испаряемых материалов должна достигать 99,999999%.

- Необходим молекулярный источник, способный испарять тугоплавкие вещества (такие как металлы) с возможностью регулировки плотности потока вещества.



.....

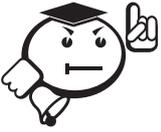
Особенностью эпитаксии является невысокая скорость роста пленки (обычно менее 1000 нм в час).

.....

5.2 Формирование квантовых точек

Упорядоченные самоорганизованные массивы квантовых точек были получены при отработке лазерных структур. Квантовые точки имеют дискретный энергетический спектр (глава 3, п. 3.6) и высокую плотность состояний, что должно давать (при использовании квантовых точек в качестве активной среды лазеров) большой коэффициент усиления, низкий пороговый ток и его температурную стабильность. В течение долгого времени предпринимались попытки изготовления квантовых точек и приборов на их основе традиционными методами, например литографией, селективным травлением структур с квантовыми ямами. На этом пути отработывалась литография с предельно высоким разрешением, но квантовых точек с дискретным атомоподобным спектром получить не удалось [77].

«Настоящие» квантовые точки были получены после смены технологии по типу «сверху вниз» на технологию «снизу вверх», когда стали использовать эффекты самоорганизации при эпитаксиальном наращивании. При определенных условиях в эпитаксиальном слое образуются упорядоченные массивы квантовых точек. На квантовых точках, полученных посредством самоорганизации, были впервые продемонстрированы физические свойства, предсказанные теоретически, в частности специфическая дискретность их энергетического спектра.



.....

Процесс самоорганизации при эпитаксиальном росте есть следствие стремления системы к уменьшению энергии. Экспериментально установлено, что при осаждении атомов на подложку из газовой фазы наблюдаются три типа начальной стадии роста: механизм Франка – Ван дер Мерве (слоевый рост), Фольмера – Вебера (островковый рост) и Странского – Крастанова (промежуточный тип).

.....

1. *Механизм Франка – Ван дер Мерве.* Осаждаемый материал смачивает подложку, постоянные решеток практически совпадают. Происходит послойный двумерный рост (рис. 5.2, а).

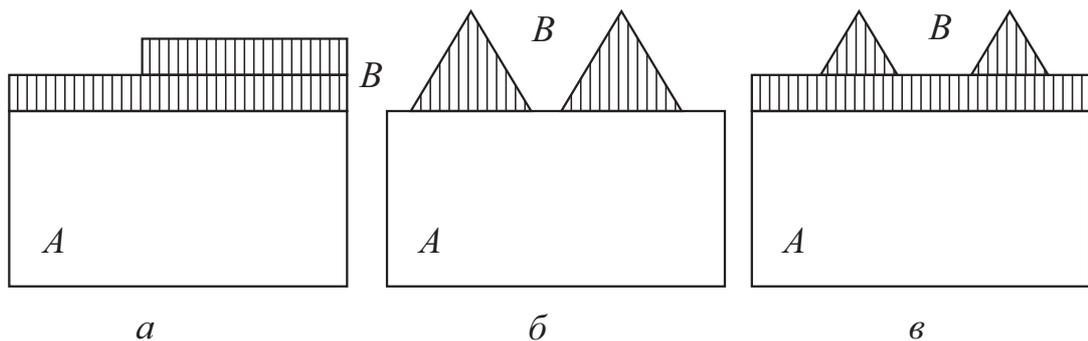


Рис. 5.2 – Схемы трех типов начальной стадии гетероэпитаксиального роста. Материал *B* осаждается на подложку *A*

2. *Механизм Фольмера – Вебера.* Осаждаемый материал не смачивает подложку (это материалы, различные по свойствам, или с большим различием постоянных решеток). Происходит островковый (трехмерный) рост. Материал *B* стягивается в наноструктуры на поверхности подложки *A* (рис. 5.2, б).
3. *Механизм Странского – Крастанова* (рис. 5.2, в). Осаждаемый материал смачивает подложку, но имеется рассогласование постоянных решеток (порядка нескольких процентов). Именно этот механизм роста используется для получения массивов квантовых точек, например

квантовых точек InAs в матрице GaAs (рассогласование решеток 7%) или квантовых точек германия в кремниевой матрице (рассогласование 4%). На начальном этапе идет послойный рост материала B на подложке A с образованием смачивающего слоя, а затем происходит переход к формированию трехмерных островков из материала B на покрытой подложке. Каждая вертикальная атомная плоскость подложки продолжается в объеме островка, но из-за различия постоянных решетки материалов A и B островок становится напряженным (используется термин «когерентно напряженный») [77].

Постоянная решетки InAs больше, чем у GaAs. При осаждении InAs на подложку GaAs сначала формируется слой InAs. Этот слой из-за различия решеток напряженный (ячейки как бы стремятся выгнуться). По мере увеличения толщины слоя упругая энергия растет, связи между атомами слоя начинают рваться, некоторые атомы частично освобождаются. Происходит перераспределение материала, и образуются трехмерные островки (см. рис. 5.2, в). Когда образуется островок, решетка InAs частично распрямляется и получается выигрыш в энергии. Образование трехмерных островков начинается после осаждения 1,6–1,7 слоя InAs.

После осаждения четырех монослоев получается плотный массив островков правильной формы. Если на островки InAs снова нарастить GaAs, получатся квантовые точки InAs (узкозонный полупроводник) в матрице GaAs (широкозонный).

Формирование массивов островков наблюдается также в системах Si – Ge, InAs – InP, AlInAs – AlGaAs, PbSe – PbTe и в ряде других.

Размер и форма островков определены для заданных условий роста и соответствуют минимуму энергии системы. Увеличение или уменьшение размеров островков приводит к увеличению энергии. Большинство островков имеют одинаковый размер, который зависит от условий роста.

Островки ограняются поверхностями с малой поверхностной энергией (так называемые кристаллографические грани с низкими индексами Милле-

ра). Например, на подложке GaAs с ориентацией (100) равновесная форма островков InAs – это пирамидки с квадратным основанием (рис. 5.3, а).

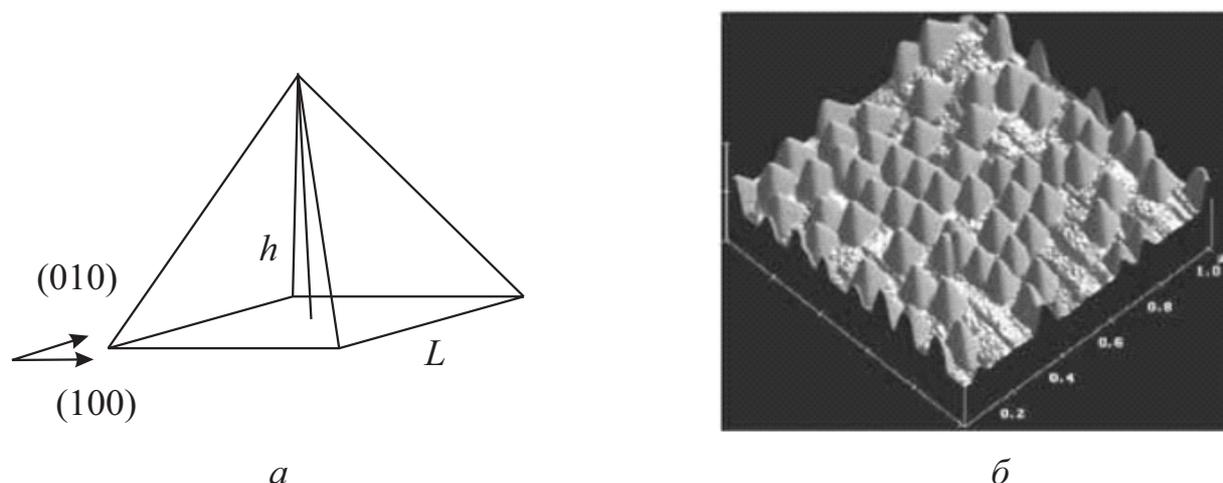


Рис. 5.3 – Островки PbSe на поверхности PbTe: а) схема островка; б) схематическое изображение островков, полученное с помощью атомно-силового микроскопа

Когда расстояние между островками становится сравнимым с размерами островков, начинается их упругое взаимодействие, обусловленное проникновением в подложку напряжений, создаваемых островками. Взаимодействие островков на поверхности – всегда отталкивание. Это обеспечивает устойчивость массива островков. На упругоанизотропных поверхностях {100} островки образуют квадратную сетку.

На рисунке 5.3, б показаны пирамидальные островки PbSe на поверхности PbTe (111).



Для приборных приложений массив островков должен обладать определенными качествами: отсутствие дефектов, однородность массива по форме, взаимному расположению и размерам островков, достаточная поверхностная плотность островков. Например, для лазерных структур необходима плотность островков $\sim 10^{11} \text{ см}^{-2}$.

Регулированием условий роста можно изменять размеры островков и их поверхностную плотность. Например, с ростом температуры подложки увеличивается размер основания L , уменьшается высота h и уменьшается поверхностная плотность островков. Эти параметры зависят также от давления паров мышьяка и других условий роста.

Подбор оптимальных условий получения островков производится экспериментально. Существуют минимальный и максимальный размеры островков, когда их можно рассматривать как квантовые точки. Для системы InAs – GaAs это 4 и 20 нм соответственно. Минимальный размер определяется тем, что квантовая точка должна иметь хотя бы один энергетический уровень; максимальный размер определяется необходимостью выполнения условия $E_2 - E_1 \gg kT$.

Образование вертикальных массивов квантовых точек. Если первый слой квантовых точек – островков закрыть несколькими монослоями матричного материала и снова осаждают материал островков, то новые островки – квантовые точки – будут формироваться строго над островками предыдущего слоя (рис. 5.4).

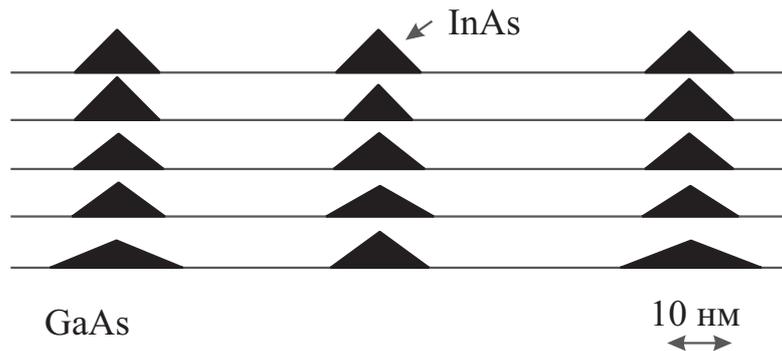


Рис. 5.4 – Схематическое изображение связанных массивов квантовых точек

Это объясняется тем, что из-за различия решеток осаждаемый материал матрицы над островками оказывается упругонапряженным. Например, в матрице GaAs над островком InAs создается максимальное растяжение. Так как атомы индия имеют больший радиус, чем атомы галлия, то при осаждении ато-

мы индия притягиваются к участкам поверхности GaAs с максимальным растяжением, то есть к участкам над зарощенными островками предыдущего слоя.

Так получают вертикальные массивы квантовых точек для приборных структур. Регулируя толщину осажденных слоев GaAs, можно получить туннельно-связанные и электронно-связанные квантовые точки в вертикальных массивах (см. рис. 5.4).

5.3 Формирование квантовых проволок (нитей)

В настоящее время получение квантовых нитей методом МПЭ стало достаточно рутинной процедурой. Большинство способов изготовления квантовых нитей основывается на том, что в системе с двумерным электронным газом (как правило, на основе гетероструктур) тем или иным способом ограничивается движение электронов еще в одном из направлений. Для этого есть несколько способов. Эти структуры можно сформировать, например, на границе раздела двух полупроводников, где находится двумерный электронный газ. Это можно сделать, если нанести дополнительные барьеры, ограничивающие движение электронов еще в одном или двух направлениях. Квантовые нити формируются в нижней точке *V*-образной канавки, созданной на полупроводниковой подложке. Если в основание канавки осадить полупроводник с меньшей шириной запрещенной зоны, то электроны этого полупроводника будут заперты в двух направлениях. Наиболее перспективным методом сейчас выглядит непосредственное «вырезание» узкой полоски с помощью литографической техники (рис. 5.5).



Один из способов изготовления квантовых проволок заключается в использовании электронно-лучевой литографии. Данный метод позволяет получать квантовые проволоки самого высокого качества. Изготавливаются квантовые проволоки $\text{In}_{0,53}\text{Ga}_{0,47}\text{As}$ при помощи пучков электронного микроскопа (200 кВ).

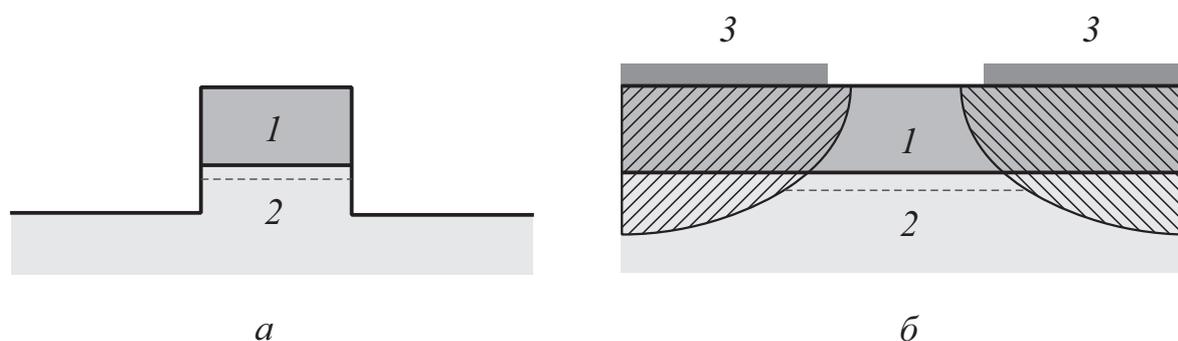


Рис. 5.5 – Полупроводниковые гетероструктуры с квантовыми нитями, полученные с помощью субмикронной литографии за счет вытравливания узкой полоски из самой структуры (*а*) или щели в затворе Шоттки (*б*):

1 – полупроводник с широкой запрещенной зоной (например, AlGaAs);

2 – полупроводник с узкой запрещенной зоной (например, GaAs);

3 – металлический затвор. Образующийся вблизи гетерограницы узкий электронный канал показан штриховой линией. Заштрихованы области обеднения электронами

Для таких высоких энергий электронов эффект близости значительно ослабляется, что позволяет получать низкоразмерные структуры с малым периодом [77].

На буферный слой InP толщиной 200 нм эпитаксиальным методом наращивается слой InGaAs толщиной 5 нм, а затем слой InP толщиной 8 нм. Квантовая яма образуется слоем узкозонного InGaAs между двумя широкозонными слоями InP. На верхний слой InP наносится слой резиста ПММА толщиной 100 нм и экспонируется система параллельных полос шириной 15 нм с расстоянием между ними 60 нм. После проявления резиста сверху наносится слой золота толщиной 20 нм. Затем резист, оставшийся после проявления, покрытый золотом, удаляется, и на верхнем слое InP остаются полоски золота шириной 15 нм, служащие масками при последующем травлении. В результате из квантовой ямы вырезаются квантовые проволоки шириной 16–18 нм. Слой InGaAs толщиной 5 нм расположен на 8 нм ниже верха структуры [77].

При ориентации проволок в направлении [011] стенки проволок получают практически вертикальными.

На рисунке 5.6 приведена микрофотография рассматриваемой структуры, полученная с помощью сканирующего электронного микроскопа.

В другом способе поверхность полупроводниковой структуры покрывают металлическим электродом, создающим с полупроводником контакт Шоттки и имеющим узкую щель (см. рис. 5.5). Если гетерограница находится достаточно близко от поверхности, в слое обеднения, то двумерные электроны на границе будут отсутствовать всюду, кроме узкой области под щелью. Такой тип одномерной структуры обладает дополнительным преимуществом: меняя напряжение на затворе, мы можем управлять эффективной шириной квантовой нити и концентрацией носителей в ней.



Рис. 5.6 – Микрофотография квантовых проволок $\text{In}_{0,53}\text{Ga}_{0,47}\text{As}$



Контрольные вопросы по главе 5

1. Почему необходимо равенство постоянных решетки у полупроводников, составляющих гетеропереход?
2. Какая технология может в ряде случаев заменить МЛЭ без потери резкости и качества гетероперехода?
3. Расскажите об эффектах самоорганизации при эпитаксиальном наращивании.

4. Назовите три типа начальной стадии эпитаксиального роста.
5. Какая технология используется для изготовления квантовых проволок?

6 Квантовые эффекты

6.1 Двумерный электронный газ в магнитном поле

Магнитное поле действует на свободное движение электрически заряженного электрона так, что электрон начинает вращаться в плоскости, перпендикулярной вектору магнитной индукции \vec{B} и движется по окружности радиуса $r = m v / B$, вращаясь с круговой частотой $\omega = Be / m$, где $B = |\vec{B}|$.

В твердом теле электрон испытывает столкновения с атомами решетки. Частота столкновений $\nu_{ст} = \tau^{-1}$, где τ – время свободного пробега. Поэтому в твердом теле движение электрона по окружности возможно только при включении достаточно сильного магнитного поля, когда $\omega / 2\pi \gg \nu_{ст}$ или $Be / 2\pi m_n \gg \tau^{-1}$. Достаточно сильные магнитные поля, для которых выполняются два условия: $\omega / 2\pi \gg \nu_{ст}$ и $\frac{\hbar Be}{2\pi m_n} \gg kT$ – называются квантующими. Для описания движения в таких полях необходим квантово-механический подход.

Характер движения частиц в квантующих полях значительно отличается от классического, особенно в 2D-системах. Для квантовой частицы понятие траектории, вообще говоря, неприменимо, когда она движется в достаточно ограниченной области. Как показывает решение уравнения Шредингера, движение частицы в магнитном поле можно считать в определенной степени ограниченным (как, например, в потенциальной яме конечной глубины) [77]. Движение в плоскости, перпендикулярной вектору магнитной индукции \vec{B} , можно считать ограниченным площадкой радиуса $r = \sqrt{\frac{\hbar}{Be}}$. Энергия такого движения, как всякого ограниченного движения, квантуется. Согласно решению уравнения Шредингера возможные значения энергии $E_m = \hbar \left(m + \frac{1}{2} \right)$, где $m = 0, 1, 2 \dots$. Эти значения энергии называются *уровнями Ландау*.

Если вектор \vec{B} направлен вдоль оси z , перпендикулярной плоскости 2D-газа, то возможные значения энергии движения вдоль оси z – это размерные уровни E_n , а возможные значения энергии движения в плоскости (xy) – это уровни Ландау E_m .

Полная энергия электронов 2D-газа: $E = E_n + E_m$. Энергия 2D-газа в магнитном поле полностью квантована, так же, как для атомов, квантовых точек и других микрообъектов.

6.2 Целочисленный и дробный квантовый эффект Холла



.....

*Существованием уровней Ландау объясняется квантование так называемого холловского сопротивления – **квантовый эффект Холла**. Квантовый эффект Холла является макроскопическим квантовым эффектом, наблюдаемым в низкоразмерных твердотельных структурах. Этот эффект носит фундаментальный характер, поскольку представляет собой новое макроскопическое проявление квантовых свойств вещества, и в то же время имеет важное прикладное значение как метод точного измерения универсальных физических констант. Наблюдается эффект при очень низких температурах ($T \sim 1$ К) в достаточно сильных полях (~ 15 Тл).*

.....

Экспериментально проявление квантового эффекта Холла в сильных магнитных полях установлено в работе [79] К. Клитцингом с соавторами. За открытие квантового эффекта Холла Клитцинг был награжден в 1985 г. Нобелевской премией. Условие сильного магнитного поля выполнялось при низких гелиевых температурах ($T \sim 1$ К), когда подвижности носителей заряда значительно возрастают, при $B \sim 15$ Тл [79]. Экспериментальным образцом являлся кремниевый транзистор MOSFET (metal-oxide-semiconductor field effect transistor), то есть полевой МОП-транзистор, в инверсном слое которого изме-

рялся эффект Холла. На рисунке 6.1 показана схематичная диаграмма МОП-транзистора, приведенная в работе [79] и адаптированная к русскоязычным терминам.

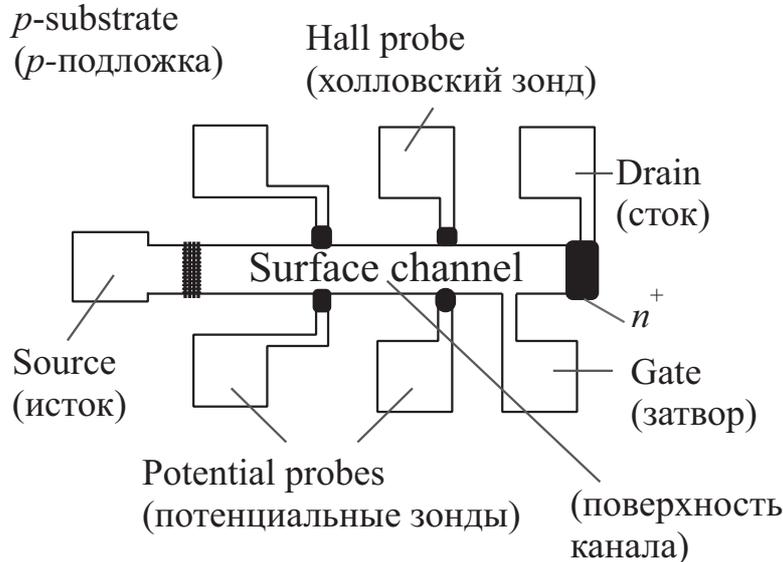


Рис. 6.1 – Схематичная диаграмма типичного для [79] МОП-транзистора

Было установлено, что почти линейная зависимость холловского сопротивления сменяется плато (рис. 6.2). Величина сопротивления Холла на плато рисунка 6.2 определяется комбинацией фундаментальных физических констант, деленной на целое число i :

$$R_H = h / ie^2. \quad (6.1)$$

Среднее значение сопротивления Холла для всех исследованных образцов было $6453,22 \pm 0,10$ Ом для измерений в энергетической щели между уровнями Ландау $n = 0$ и $n = 1$ (что соответствует $i = 4$ в формуле (6.1)), $3226,62 \pm 0,10$ Ом для измерения в энергетической щели между уровнями Ландау $n = 1$ и $n = 2$ ($i = 8$), и $12\,906,6 \pm 1,012$ для измерений в энергетической щели между уровнями с разными спиновыми состояниями $n = 0$ ($i = 2$) [79]. Эти сопротивления хорошо согласуются с расчетами.

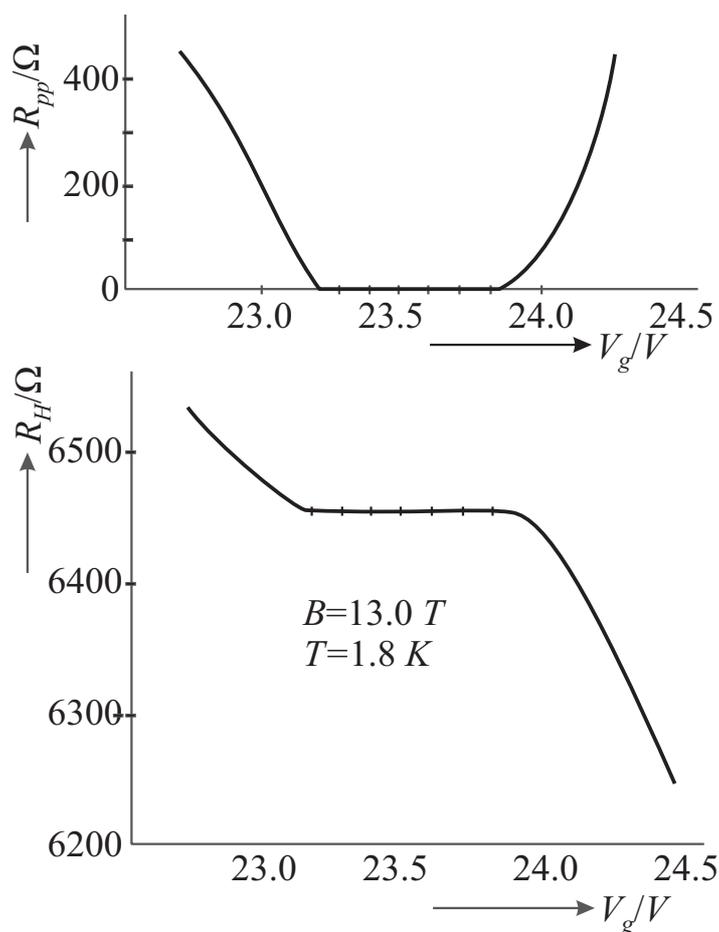


Рис. 6.2 – Сопротивление Холла R_H и сопротивление прибора R_{pp} между потенциальными зондами в зависимости от напряжения на затворе V_g , величине которого соответствуют полностью занятые минимальные ($n = 0$) состояния уровня Ландау. Плато R_H имеет значение $6453,3 \pm 0.1$ Ом. Геометрия прибора $L = 400$ мкм; $W = 50$ мкм; расстояние между зондами $L_{pp} = 130$ мкм, $B = 13$ Тл

В своей Нобелевской лекции [80] Клитцинг привел дальнейшее развитие своих экспериментов, уже на квантовых гетероструктурах $\text{Al}_x\text{Ga}_{1-x}\text{As}$ (рис. 6.3).

В таких гетероструктурах при нулевом напряжении на затворе $V_g = 0$ при изменении магнитного поля от 0 до 5 Тл хорошо наблюдаются сменяющие себя участки плато в сопротивлении Холла $R_H(B)$ (рис. 6.4).

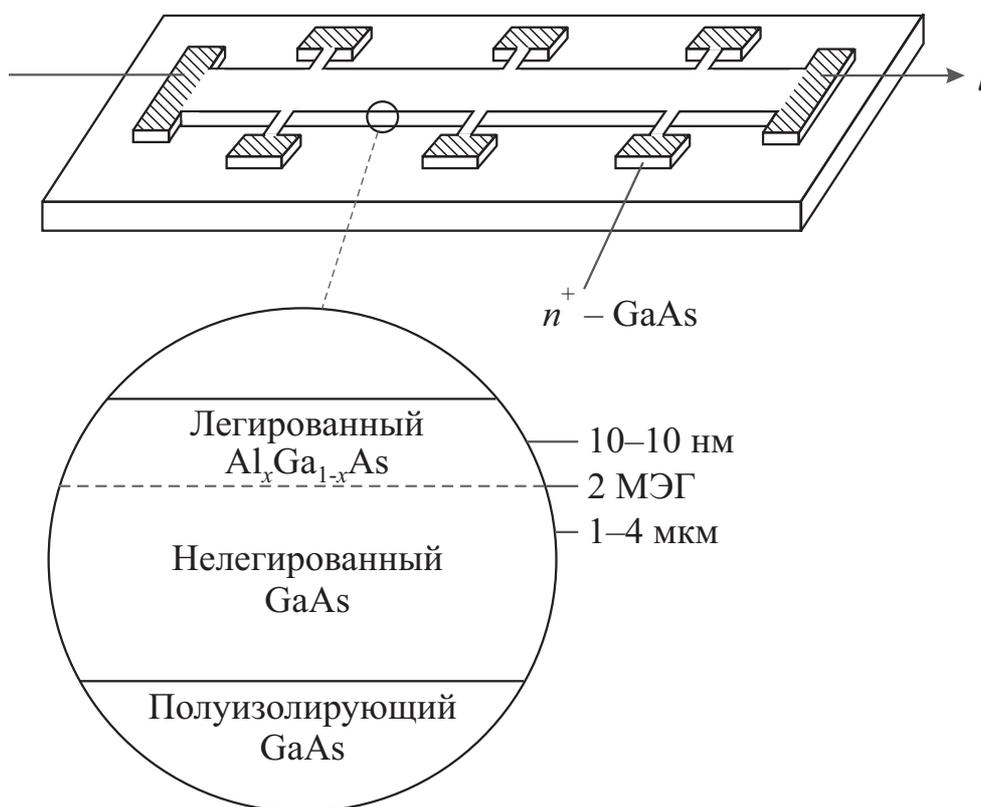


Рис. 6.3 – Типичные форма и сечение гетероструктуры $\text{Al}_x\text{Ga}_{1-x}\text{As}$, используемой при измерениях эффекта Холла



.....

Таким образом, в эксперименте в сильных магнитных полях наблюдаются два эффекта: 1) R_H выходит на плато и 2) в области плато R_H продольное сопротивление прибора между потенциальными зондами, в пределах ошибки высокоточных измерений, становится равным нулю (рис. 6.2). Оба эти эффекта получили название **квантового целочисленного эффекта Холла**.

В 1982 г. в работе Д. Тсу, Г. Штермера и А. Госсарда [81] было отмечено, что в более сильных полях плато на холловском сопротивлении наблюдается не только при целочисленных значениях параметра i , но и при $i = 1/3$. А в последствии плато сопротивления были обнаружены и при других дробных значениях $i = 2/5$, $i = 3/7$ и т. д. Интересно отметить, что знаменатель в дробном i всегда нечетный. Этот эффект получил название **квантового дробного эффекта Холла**.

.....

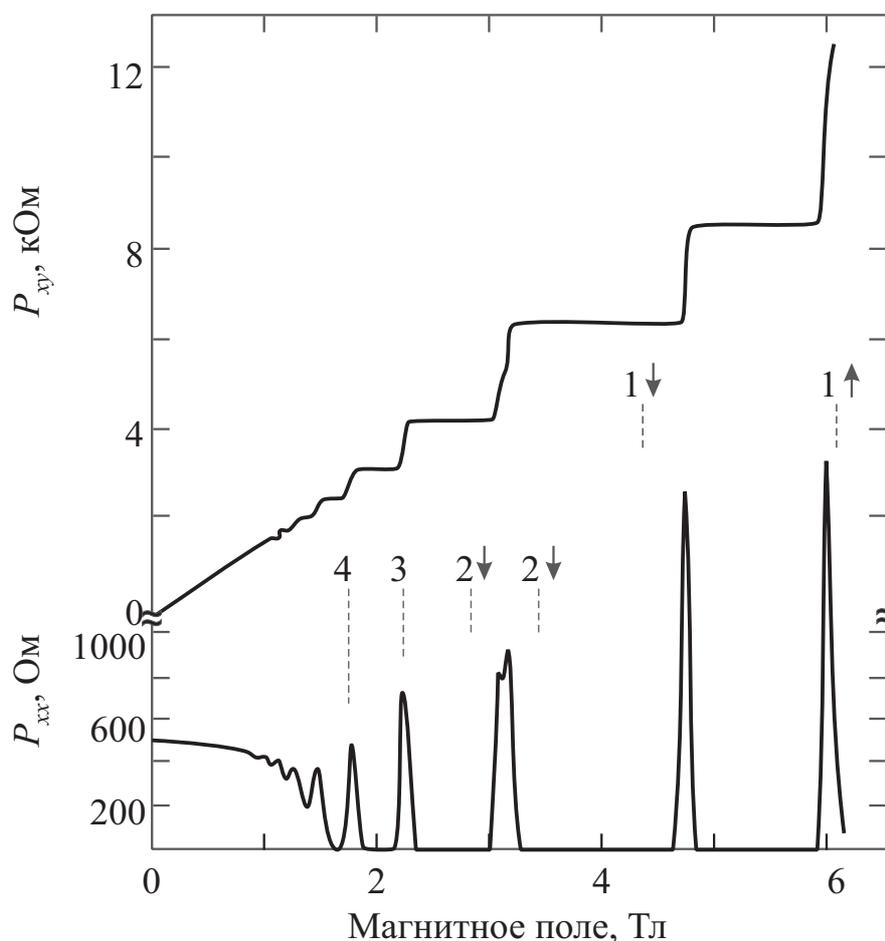


Рис. 6.4 – Экспериментальные зависимости холловского сопротивления $R_H = \rho_{xy}$ и удельного сопротивления $\rho_{xx} \sim R_x$ гетероструктуры от магнитного поля при фиксированной плотности носителей, соответствующей напряжению на затворе $V_g = 0$. Температура около 8 мК

6.3 Эффект Ааронова – Бома



.....

Эффект Ааронова – Бома – квантово-механический эффект, характеризующий влияние внешнего электромагнитного поля, сосредоточенного в области, недоступной для заряженной частицы, на ее квантовое состояние. Наличие такого нелокального воздействия электромагнитного поля на заряженную частицу, исчезающего в классическом пределе, подчеркивает, что при квантовом рассмотрении взаимодействие заряженной частицы с электромагнитным полем не сводится к локальному действию на нее силы Лоренца. Подробное теоретическое изучение эффекта

проведено в 1959 г. Я. Аароном и Д. Бомом, отметившими его тесную связь с фундаментальными положениями квантовой теории. Их исследования привлекли внимание к особой роли электромагнитных потенциалов в квантовой теории.

.....

В 1959 г. Я. Ааронов и Д. Бом показали, что в квантовой механике электромагнитные потенциалы A (векторный) и φ (скалярный) приобретают непосредственный физический смысл и могут быть измерены.

В этой связи необходимо сказать, что в классической физике (классической электродинамике) имеют смысл только электрическое E и магнитное B поля, определяющие силу F , действующую на электрический заряд e : $F = eE$ или на проводник с током (т. е. на движущийся со скоростью v заряд): $F = evB$.

Формально напряженность электрического поля E и индукция магнитного поля B могут быть выражены через векторный A и скалярный φ потенциалы следующим образом:

$$\vec{B} = \text{rot} \vec{A}, \quad (6.2)$$

$$E = -\text{grad} \varphi - \frac{\partial \vec{A}}{\partial t}. \quad (6.3)$$

Оказалось, что не напряженности полей, а именно электромагнитные потенциалы определяют фазу волновой функции электрона.

.....



Эффект Ааронова – Бома как раз и состоит в том, что электромагнитные потенциалы непосредственно влияют на интерференционную картину, образуемую электронами, движущимися в той области пространства, где присутствует векторный и/или скалярный потенциал, однако напряженности электрического и/или магнитного полей равны нулю. Ааронов и Бом обратили внимание на то, что электромагнитный вектор-потенциал должен сдвигать фазу волновой функции электрона ψ (даже в том

случае, когда путь электрона лежит в области, где нет никаких электрических или магнитных полей) на величину

$$\Delta\varphi = \frac{e}{\hbar} \int (V dt - \vec{A} d\vec{S}), \quad (6.4)$$

где $d\vec{S}$ и dt – элементы пути и времени на траектории электрона; V – напряжение электрического поля; A – вектор-потенциал магнитного поля.

.....

В 1959 г. Ааронов и Бом [82] нашли решение уравнения Шредингера, описывающее рассеяние электронов на бесконечно длинном и бесконечно тонком соленоиде. Полученная ими формула для сечения рассеяния электронов с импульсом p на угол θ имеет вид:

$$d\sigma = \frac{\hbar}{p} \sin^2 \left(\pi \frac{\Phi}{\Phi_0} \right) \frac{d\theta}{2\pi \cos^2(\theta/2)}, \quad (6.5)$$

где Φ – поток, захваченный соленоидом, $\Phi_0 = hc/e$ – квант потока.

Эффект Ааронова – Бома заключается в том, что электрон рассеивается, двигаясь по области, где нет никакого поля – об этом говорит формула (6.5), которая интересна множителем $\sin^2 \left(\pi \frac{\Phi}{\Phi_0} \right)$. Этот множитель говорит о том, что электроны не ощущают существования соленоида, если

$$\Phi = n\Phi_0, \quad n = 1, 2, 3, \dots \quad (6.6)$$

Если магнитный поток вычислять как интеграл через поверхность, ограниченную окружностью радиуса r с центром в начале координат, то результат совпадает с выражением (6.6). Поскольку вычисленный поток не зависит от радиуса окружности, то ясно, что весь он сосредоточен в начале координат. Это приводит в итоге к появлению бесконечно длинного и бесконечно тонкого соленоида с целым числом квантов потока в нем, и мы приходим к заключению, что бесконечно длинный и бесконечно тонкий соленоид с целым числом квантов потока в нем ненаблюдаем.

6.4 Квантово-размерный эффект Штарка

В этом разделе рассматриваются сдвиг и расщепление энергетических уровней в атоме под действием внешнего электрического поля. Начальные представления об этом эффекте даны в п. 2.6, что облегчает дальнейшее рассмотрение материала. Хороший обзор о штарковском сдвиге уровней в атоме и в многоэлектронных системах представлен в работе [83].



.....

*Сдвиг уровня в постоянном электрическом поле называется **статическим штарковским сдвигом**; сдвиг уровня в переменном монохроматическом электромагнитном поле называется **динамическим штарковским сдвигом**.*

.....

Статический штарковский сдвиг атомных уровней был открыт Й. Штарком в 1913 г. До середины XX в. как экспериментальные, так и теоретические исследования этого эффекта ограничивались случаем постоянного электрического поля. Лишь в 1960-х гг. создание лазеров стимулировало исследователей обратиться к возмущению (сдвигу и расщеплению) атомных уровней под действием переменного электромагнитного поля и, в первую очередь, под действием светового поля лазерного излучения. Эти исследования вошли в качестве составляющей в общую задачу изучения взаимодействия интенсивного света с веществом [83].

Эффект Штарка в слабом постоянном электрическом поле детально исследован экспериментально и описан теоретически. Электрон, связанный в атоме и имеющий (отрицательную) энергию E_n , приобретает во внешнем постоянном электрическом поле с напряженностью F дополнительную энергию $\delta E_n(F)$, так как атом в поле поляризуется [83].

Если указанное состояние не вырождено, то его энергия во внешнем поле имеет вид:

$$E_n(F) = E_n + \delta E_n(F). \quad (6.7)$$

Вырожденное состояние с энергией E_m , где m – магнитное квантовое число, испытывает в электрическом поле расщепление на $2l + 1$ компоненту, так как подуровни, соответствующие различным значениям m , приобретают в поле различную дополнительную энергию $\delta E_{nm}(F)$. В частности, открытие Й. Штарка состояло в наблюдении расщепления бальмеровской линии H_α в спектре атома водорода под действием постоянного электрического поля.

Атом водорода (кроме основного состояния), а также высоковозбужденные водородоподобные (ридберговские) состояния многоэлектронных атомов имеют дополнительное вырождение по орбитальному квантовому числу l . Оно снимается в электрическом поле. В отсутствие поля такой атом имеет постоянный дипольный момент d . Дополнительная энергия, приобретаемая атомом с постоянным дипольным моментом во внешнем электрическом поле, равна

$$\delta E_{n_1, m}(F) = -d_{n_1, m} F. \quad (6.8)$$

Здесь n_1 – параболическое квантовое число, характеризующее расщепленный подуровень. Из (6.7) и (6.8) следует, что в этом случае изменение энергии пропорционально первой степени напряженности поля.



.....

Это так называемый линейный эффект Штарка. В этом случае получается симметричная относительно первичной спектральной линии картина расщепления.

.....

Многоэлектронные атомы, находящиеся в основном и низколежащих возбужденных состояниях, не имеют постоянного дипольного момента. Под действием внешнего поля в них возникает индуцированный дипольный момент:

$$d_{nlm} = \alpha_{nlm} F. \quad (6.9)$$

Величина α_{nlm} представляет собой статическую поляризуемость атома в состоянии с указанными квантовыми числами. Из (6.7)–(6.9) следует, что дополнительная энергия в этом случае определяется соотношением:

$$\delta E_{n_1, m}(F) = -(1/2) \alpha_{n_1, m} F^2. \quad (6.10)$$



.....

Это квадратичный эффект Штарка, характерный для многоэлектронных атомов. В этом случае получается несимметричная относительно первичной спектральной линии картина расщепления.

.....

На основе эффекта Штарка разработаны методы изучения свойств частиц и электрических полей в веществе. В частности, эффект Штарка в переменном электрическом поле используется для изменения частоты квантового перехода (штарковская модуляция), например в устройствах микроволновой спектроскопии, для оценки концентрации заряженных частиц в плазме; расщепление электронных уровней энергии ионов в кристаллической решетке под действием внутрискристаллического поля учитывается в спектроскопии кристаллов, в квантовой электронике.

Например, эффект Штарка с успехом применяется для исследования квантово-размерных полупроводниковых гетеронаноструктур (ГНС) [84, 85]. Влияние электрического поля на энергетический спектр квантово-размерных слоев называют квантово-размерным эффектом Штарка (КЭШ, quantum-confined Stark effect). Исследование КЭШ позволяет определить дипольный момент неравновесных электронно-дырочных пар, в частности связанных пар – экситонов, который чувствителен к распределению химического состава и поля упругих напряжений в этих слоях. Экситон состоит из электрона и дырки, но его считают самостоятельной элементарной квазичастицей, перемещающейся по кристаллу, причем это движение не связано с переносом электрического заряда и массы.

Применение методик фотоэлектрической спектроскопии обеспечивает возможность определения энергетического спектра и некоторых других оптоэлектронных характеристик квантово-размерных ГНС (КГНС) In(Ga)As/GaAs с практически любым типом квантово-размерных слоев – квантовые точки (КТ), квантовые ямы (КЯ) и комбинированный слой КЯ/КТ – и местом их рас-

положения в структурах на полуизолирующей или проводящей подложке. С применением комплекса этих методик установлен ряд новых закономерностей, касающихся особенностей энергетического спектра и возможностей его управления в КГНС с КТ и КЯ In(Ga)As/GaAs [84–86].

КЭШ в ГНС с КЯ проявляется в красном смещении линии экситонного поглощения КЯ. Это смещение связано с изменением формы потенциальной ямы в электрическом поле. На рисунке 6.5 показаны энергетические диаграммы и огибающие волновые функции фотовозбужденных электронов и дырок для гетероструктуры с КЯ InGaAs/GaAs в отсутствие электрического поля и в электрическом поле с напряженностью F . В достаточно совершенных КЯ обычно образуются связанные (экситонные) состояния электронов и дырок, которые проявляются в наличии пика на краю межзонного оптического поглощения КЯ. Энергия основного оптического перехода в КЯ:

$$E_0 = E_{e_1} - E_{hh_1}, \quad (6.11)$$

где E_{e_1} и E_{hh_1} – энергии основного состояния электрона и тяжелой дырки соответственно.

При отсутствии постоянного дипольного момента электронно-дырочной пары в КЯ красное смещение экситонного пика основного оптического перехода в электрическом поле:

$$\Delta E_0 = E_0(0) - E_0(F) = -(\mathbf{p} \cdot \mathbf{F}) = -\alpha F^2, \quad (6.12)$$

определяется индуцированным дипольным моментом $\mathbf{p} = \alpha \mathbf{F}$ и квадратично зависит от напряженности приложенного электрического поля F .

Кроме смещения линии экситонного поглощения в достаточно сильном электрическом поле происходит уширение линии, связанное с уменьшением времени жизни экситона, и уменьшение коэффициента поглощения, связанное с уменьшением интеграла перекрытия огибающих волновых функций электронов и дырок (рис. 6.5, б).

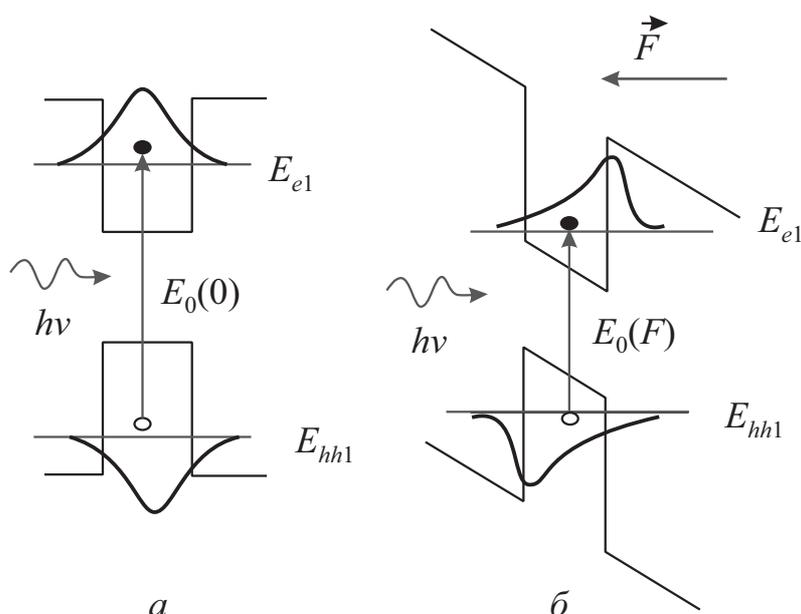


Рис. 6.5 – Энергетические диаграммы гетероструктуры с КЯ: *а* – в отсутствие электрического поля, *б* – в электрическом поле с напряженностью F

Эффект Штарка на линиях экситонного поглощения КЯ Al(In)GaAs/GaAs [84–86] представляет практический интерес в связи с созданием быстродействующих оптических затворов и модуляторов для области спектра 0,85–1,1 мкм. Они представляют собой многоямную структуру, встроенную в *p-i-n*-диод (рис. 6.6) и обеспечивают глубину оптической модуляции больше 50% для света, падающего нормально на модулятор. В оптических модуляторах используются гетероструктуры с КЯ GaAs/AlGaAs и с КЯ InGaAs/GaAs для модуляции излучения с длиной волны $\lambda \approx 0,85$ мкм и $\lambda \approx 1$ мкм соответственно. Последние структуры могут быть использованы для модуляции излучения твердотельного лазера на основе алюмо-иттриевого граната («YAG», $Y_3Al_5O_{12}$), легированного ионами неодима (Nd), обычно называемого Nd:YAG лазером ($\lambda = 1,064$ мкм) [84–86].



С помощью КЭШ можно управлять спектром оптического поглощения и излучения ГНС. На этом эффекте основана работа быстродействующих электрооптических модуляторов, которые обычно представляют собой многоямную ГНС, встроенную

в $p-i-n$ -диод (диод с высокоомной базой из полупроводника, близкого к собственному (intrinsic)).

.....

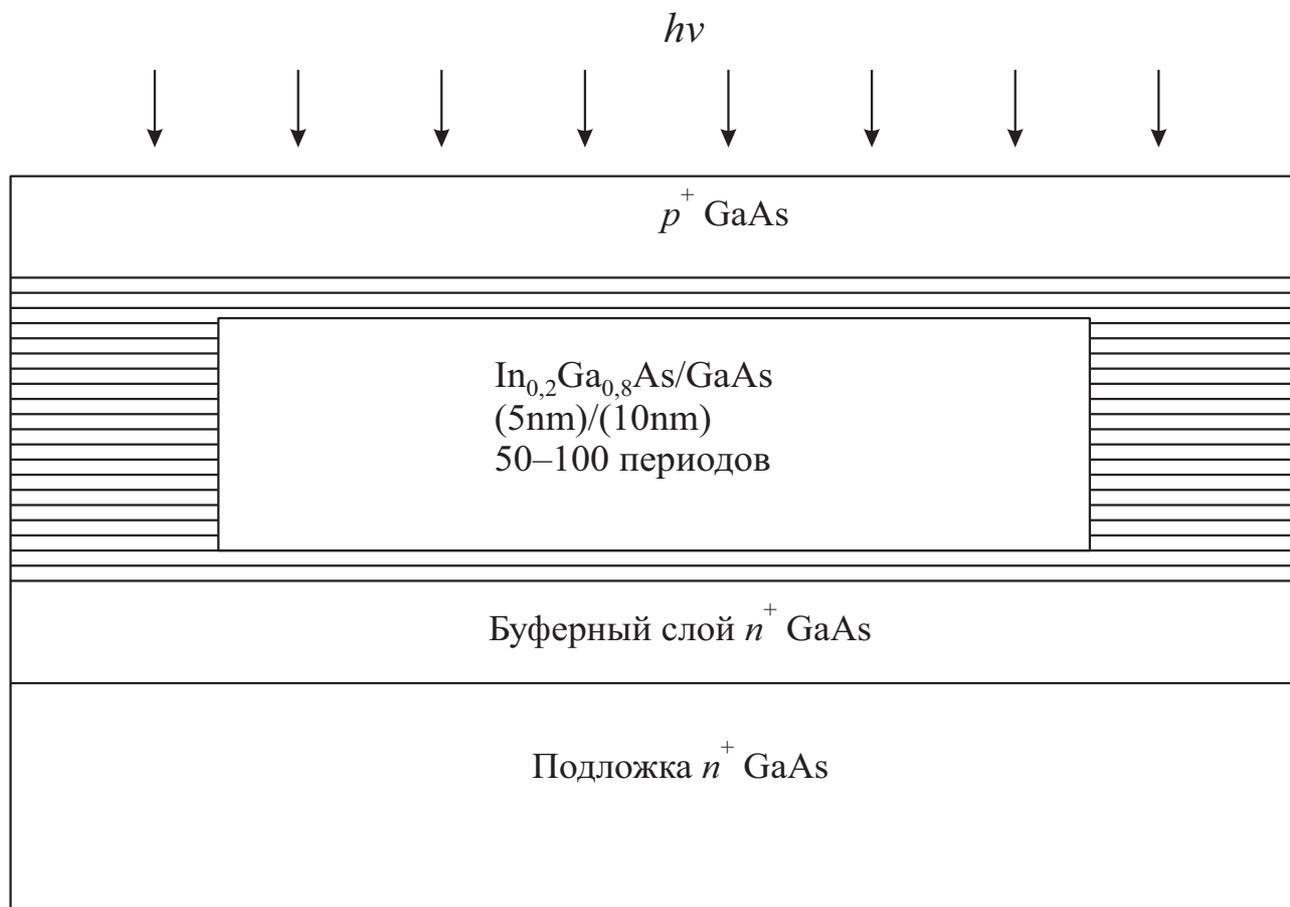


Рис. 6.6 – Схема многоямной гетеронаноструктуры InGaAs/GaAs, встроенной в $p^+ - i - n^+$ -диод

При приложении напряжения к такой структуре происходит смещение экситонных пиков поглощения (основного и возбужденных переходов) (рис. 6.7). Электрическое поле в низколегированной i -области однородно и определяется следующей формулой:

$$F = \frac{V + \phi_0}{L_i}, \quad (6.13)$$

где V – приложенное обратное смещение на диоде, ϕ_0 – высота барьера, L_i – ширина i -области ($L_i = (L_{QW} + L_b) N_{QW}$, где N_{QW} – число КЯ в многоямной структу-

ре). При $L_{QW} = 5$ нм, $L_b = 10$ нм, $N_{QW} = 50$, $L_i \approx 750$ нм, $F = \varphi_0 / W \approx 20$ кВ/см при $V = 0$.

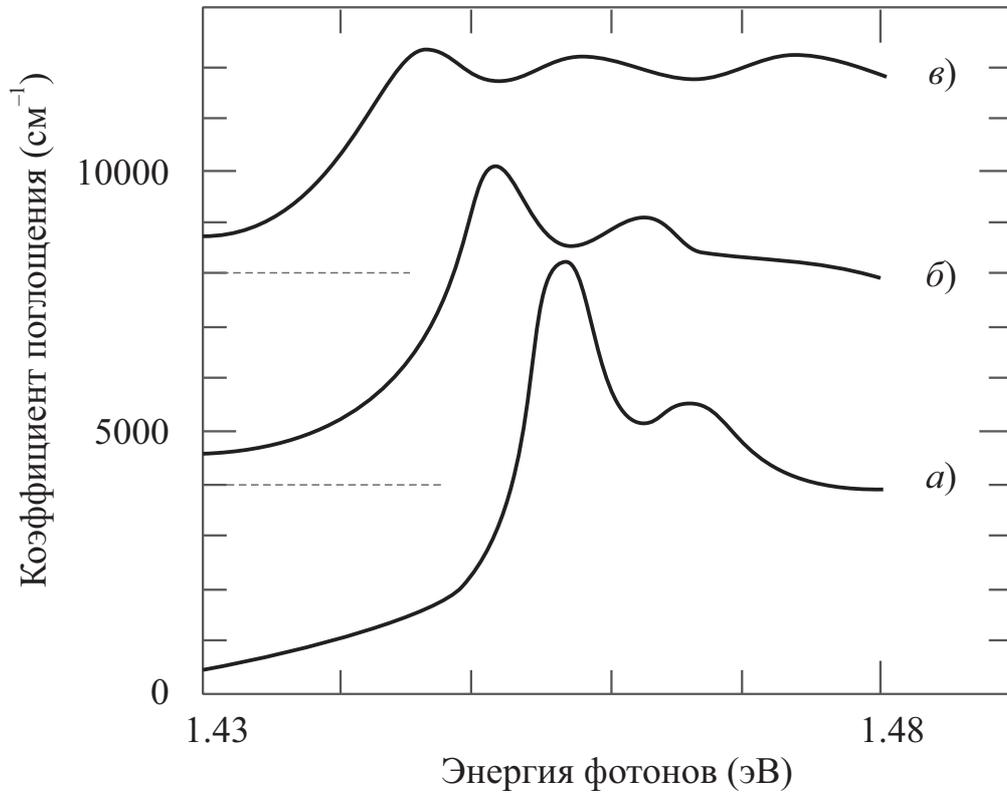


Рис. 6.7 – Спектры поглощения многоямной структуры GaAs/AlGaAs при различных значениях электрического поля F в i -области: *a*) $F = 10$ кВ/см, *б*) $F = 47$ кВ/см, *в*) $F = 73$ кВ/см

Такие модуляторы демонстрируют хорошие электрические характеристики и глубину модуляции (контраст поглощения) до 57% на экситонных пиках [84–86].

Коэффициент экситонного поглощения близок по величине к коэффициенту межзонного поглощения одиночной КЯ:

$$\beta^{QW} = \frac{\Delta I}{I_0}, \quad (6.14)$$

где $\Delta I = I_0 - I_1$, I_0 , I_1 – интенсивности падающего и прошедшего через плоскость КЯ излучения.

Для кейновского закона дисперсии (длина волны электрона велика по сравнению с постоянной решетки, а энергетический зазор между разрешенными

ми зонами больше энергии электрона) при нормальном падении излучения на плоскость КЯ коэффициент поглощения для электронных переходов в пределах первых подзон ($h\nu > E_{el} - E_{hh1}$) [84–86]:

$$\beta^{QW} = \frac{6\gamma\pi}{\sqrt{\varepsilon}} \frac{E_g(E_g + \Delta)|X|^2}{(E_{e1} - E_{hh1})(3E_g + \Delta)} \frac{m^*}{m_0} \left(\frac{m_0}{m_e} - 1 \right), \quad (6.15)$$

где $\gamma = e^2 / c\hbar \cong 1/137$ – постоянная тонкой структуры, E_g – ширина запрещенной зоны в твердом растворе, Δ – энергия спин-орбитального расщепления зон,

$$|X|^2 = \int \chi_s^*(z) \chi_h(z) dz, \quad (6.16)$$

$|X|^2$ – интеграл перекрытия огибающих волновых функций электронов и дырок, ε – относительная диэлектрическая проницаемость, $m^* = m_e m_{hh} (m_e + m_{hh})$.

Из (6.14) следует, что для структуры GaAs/In_xGa_{1-x}As с $x < 0,3$, когда $\Delta \ll E_g$, $E_{e1} - E_{hh1} \cong E_g$, $m_e \ll m_0$, m_{hh} и $|X|^2 \cong 1$, коэффициент поглощения КЯ слабо зависит от состава твердого раствора и ширины ямы. Он также не зависит от $h\nu$ вблизи края и приближенно может быть найден из простого выражения:

$$\beta^{QW} \approx \frac{2\gamma\pi}{\sqrt{\varepsilon}}. \quad (6.17)$$

При $\varepsilon = 11$, $\beta \cong 1,4 \cdot 10^{-2}$, т. е. КЯ поглощает около 1% падающего на нее излучения. Заметим, что точный расчет по формуле (6.15) дает $\beta = 7,1 \cdot 10^{-3}$.

6.5 Квантово-размерный туннельный эффект

Туннелирование и его закономерности было подробно рассмотрено в главе 3, п. 3.8.



Туннельный квантово-механический эффект впервые был открыт в 1927 г. советским физиком Георгием Гамовым, эмигрировавшим в США в 1933 г. Гамов изучал вопросы распада ядра атома на α -частицу и ядерный остаток (α -распад). Он обратил внимание на установленные закономерности α -распада, свидетельствующие

о том, что квантовая частица может проходить через классически запрещенные области.

Туннельный эффект является принципиально квантово-механическим эффектом, не имеющим аналога в классической механике. Своим существованием туннельный эффект подтверждает основополагающее положение квантовой механики – корпускулярно-волновой дуализм свойств элементарных частиц.

.....

В п. 3.8 были даны определения коэффициента прохождения частицы через границу потенциального барьера (коэффициент туннельной прозрачности барьера) и коэффициент отражения частицы от потенциального барьера.

Коэффициентом прозрачности D в случае прямоугольного барьера определяется выражением [77]:

$$D = D_0 e^{-\frac{2}{\hbar} \sqrt{2m(U^0 - E)}d}, \quad (6.18)$$

D_0 – предэкспоненциальный множитель, U^0 – высота барьера, d – ширина барьера, $(U^0 - E)$ – дефицит энергии частицы внутри барьера.

Главная особенность выражения (6.18) заключается в том, что очень малая величина \hbar стоит в знаменателе экспоненты, вследствие чего коэффициент туннелирования через барьер классической частицы большой массы очень мал. Чем меньше масса частицы, тем больше и вероятность туннельного эффекта. Например, при $U^0 = 2$ эВ и $d = 10^{-8}$ см вероятность прохождения сквозь барьер для электрона с энергией 1 эВ равна 0,78, а для более тяжелого протона с той же энергией лишь $3,6 \cdot 10^{-19}$. Если взять вообще макроскопическое тело – шарик массой в 1 г, движущийся по горизонтальной поверхности с очень малой скоростью (кинетическая энергия близка к нулю), то вероятность преодоления им препятствия – лезвия бритвы толщиной 0,1 мм, выступающего над горизонтальной поверхностью на 0,1 мм, равна 10^{-26} .

Прозрачность барьера тем больше, чем меньше ширина барьера d и разность между его высотой и энергией частицы.

Потенциальный барьер прямоугольной формы – это упрощающее теорию приближение, а в реальности форма барьера может быть весьма произвольной. Но такой барьер (рис. 6.8, *а*) можно представить в виде последовательности большого числа узких прямоугольных потенциальных барьеров, расположенных один за другим (рис. 6.8, *б*). Будем считать, что барьер имеет достаточно плавную форму, то есть полагать, что его высота на расстоянии, сравнимом с длиной волны де Бройля, изменяется незначительно. В этом случае отражением волны на выступающих участках прямоугольных барьеров можно пренебречь и считать, что ослабление волны происходит в основном за счет поглощения.

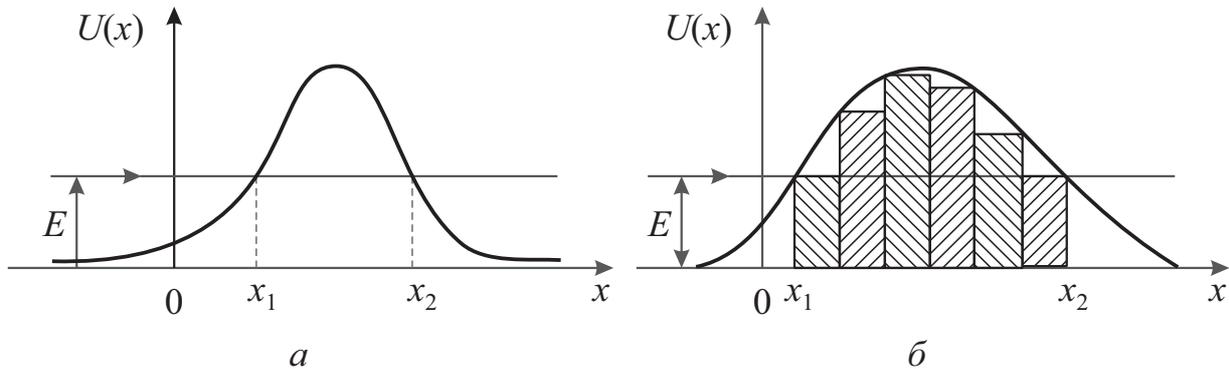
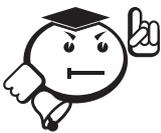


Рис. 6.8 – Барьер произвольной формы (*а*) и его представление в виде последовательности большого числа узких прямоугольных потенциальных барьеров, расположенных один за другим (*б*)



.....

Волна де Бройля, прошедшая через i -й прямоугольный барьер, представляет собой волну, падающую на $(i + 1)$ -й барьер и т. д. Вероятность прохождения частицы через цепочку последовательно расположенных потенциальных барьеров равна произведению вероятностей прохождения через каждый из барьеров. Таким образом, коэффициент прохождения (прозрачности) D равен произведению коэффициентов прохождения D_i для каждого барьера:

$$D = \prod_i D_i \approx \prod_i \exp \left\{ -\frac{2\Delta x_i}{\hbar} \sqrt{2m_0 [U(x_i) - E]} \right\} =$$

$$= \exp \left\{ -\sum_i \frac{2\Delta x_i}{\hbar} \sqrt{2m_0 [U(x_i) - E]} \right\}, \quad (6.19)$$

где Δx_i – ширина, а $U(x_i)$ – высота i -го барьера. Переходя в (6.19) от суммирования к интегрированию, получаем:

$$D \approx \exp \left\{ -\frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m_0 [U(x) - E]} dx \right\}, \quad (6.20)$$

где x_1 и x_2 – значения координат, при которых $U(x) = E$ (рис. 6.8).

.....

Эффект туннелирования в реальных наноструктурах проявляется, например, при прохождении электрона сквозь слои диэлектрика. При этом время туннелирования τ очень мало, порядка $\hbar/(U^0 - E) \sim 10^{-15}$ с, эта оценка возможна на основе соотношения неопределенностей Гейзенберга (2.29).



Туннельный эффект имеет заметную вероятность, если d соизмеримо с длиной волны де Бройля. По существу туннельный эффект определяет пределы функционирования элементов ИМС на основе традиционных принципов.

.....

Но если положить туннельный эффект в основу принципа действия прибора, то это может повысить его быстродействие (до сотен терагерц). Например, на этом принципе работают одноэлектронные устройства [77].

6.6 Эффект Джозефсона



В 1962 г. Брайн Д. Джозефсон, будучи дипломником Кембриджского университета, предсказал замечательное явление. Опираясь на чисто теоретический анализ сверхпроводимости (резкого исчезновения электрического сопротивления некоторых веществ

при температурах, близких к абсолютному нулю), Джозефсон пришел к выводу о том, что сверхпроводящий ток, определяемый коррелированными парами электронов, может протекать через пленку изолятора, разделяющую два сверхпроводника, если ее толщина достаточно мала. Он показал далее, что это «туннелирование» электронных пар через изолятор может принимать две формы, которые и называются теперь эффектами Джозефсона. Оба эти эффекта были найдены экспериментально.

.....

Коррелированными парами электронов являются куперовские пары (по имени американского ученого Л. Купера), предсказавшего в 1956 г. объединение в связанное состояние двух электронов с противоположными спинами и импульсами движения.

Для того чтобы понять природу эффектов Джозефсона, рассмотрим, что происходит при протекании тока через сверхпроводник. Если пропускать ток через брусок из сверхпроводящего вещества, то вольтметр, присоединенный к концам этого бруска, покажет, что падение напряжения на нем равно нулю (рис. 6.9, *а*). Таким образом, брусок не имеет сопротивления и, как говорят, находится в сверхпроводящем состоянии. Если разрезать теперь этот брусок на две части и разнести их на некоторое расстояние, равное, скажем, одному сантиметру, то ток через такую открытую цепь прекратится, а вольтметр покажет напряжение, равное напряжению источника тока (например, батареи) в открытой цепи (рис. 6.9, *б*) [87].

Если, однако, уменьшить расстояние между двумя кусками примерно до десяти ангстрем, может возникнуть одно из двух удивительных явлений. Возможно, что вольтметр по-прежнему не будет показывать падения напряжения между этими кусками сверхпроводника, но, несмотря на пространственное их разделение, через цепь будет проходить ток (рис. 6.9, *в*).

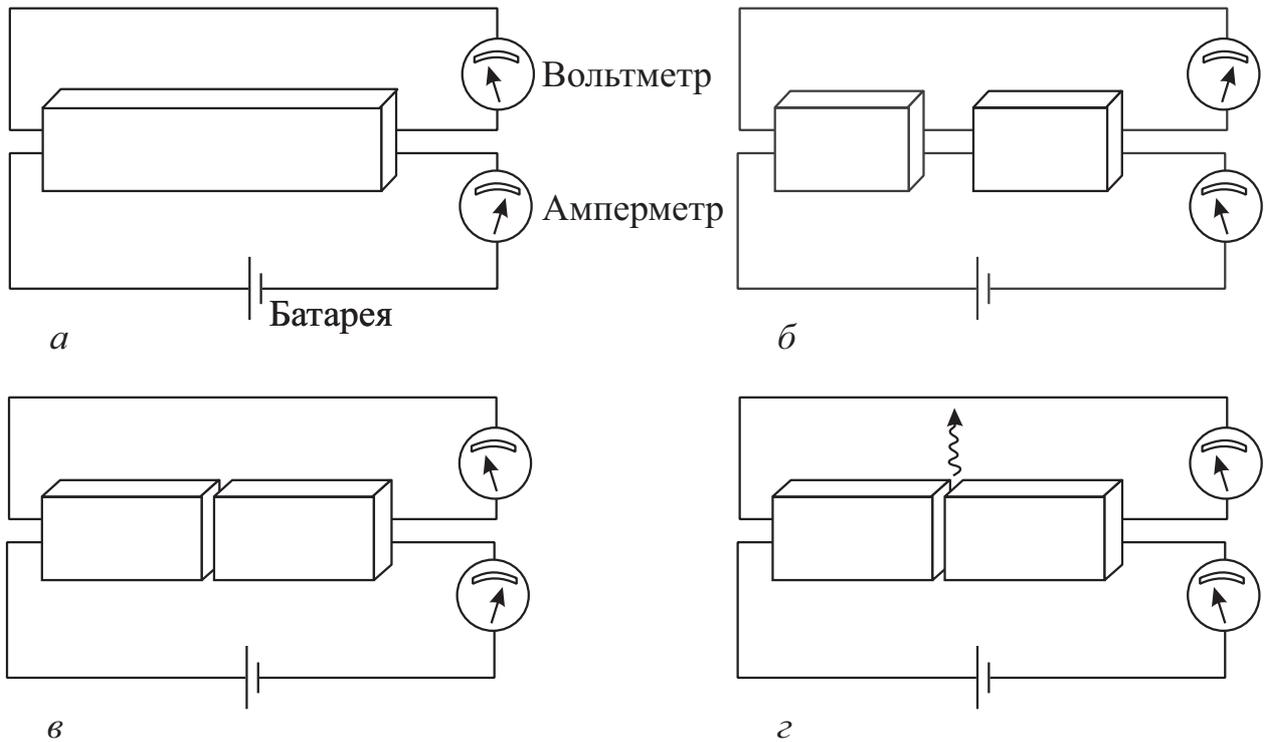


Рис. 6.9 – Токи и напряжения в системе сверхпроводящих брусьев



Другими словами, ток может протекать без сопротивления не только через два куска сверхпроводника, но и через щель между ними. Это явление называется **стационарным эффектом Джозефсона**.

Может быть и так, что при наличии тока в цепи вольтметр показывает падение напряжения. При этом из щели между сверхпроводниками излучаются высокочастотные электромагнитные волны, что указывает на присутствие там переменного тока (рис. 6.9, г). Это явление называется **нестационарным эффектом Джозефсона**.

Оба эти эффекта являются следствием особой природы сверхпроводящего состояния.

Исследование этих двух эффектов сделало важный вклад в понимание сверхпроводимости. Но они могут иметь и технические приложения, такие как

генерация коротковолнового электромагнитного излучения, измерение очень малых магнитных полей и создание эталонов напряжения высокой точности. Возможно, более важно то, что эффекты Джозефсона позволяют физикам глубже понять явление сверхпроводимости и измерять некоторые фундаментальные физические постоянные с недостижимой ранее точностью [87].

Эффект Джозефсона экспериментально обнаружен американскими физиками П. Андерсоном и Дж. Роуэллом в 1963 г. Электроны проводимости, объединенные в куперовские пары, проходят через тонкий слой диэлектрика (обычно пленку оксида металла, исходного металлического сверхпроводника толщиной ~ 1 нм) благодаря туннельному эффекту. Если ток через этот слой не превышает определенного *критического* значения, то падение напряжения на контакте Джозефсона отсутствует (стационарный эффект).

При токах больше критического на контакте возникает падение напряжения U и он излучает электромагнитные волны (нестационарный эффект). Частота излучения ν связана с U соотношением $\nu = 2eU / h$. Излучение обусловлено тем, что объединенные в куперовские пары электроны при переходе через контакт приобретают избыточную по отношению к энергии их основного состояния энергию $2eU$. Единственная возможность для такой пары вернуться в основное состояние – это излучить квант энергии $h\nu = 2eU$. Существует и обратный эффект: если поместить контакт Джозефсона в СВЧ электромагнитное поле, то при достижении на контакте напряжения с амплитудой $U = h\nu / 2e$ между двумя сверхпроводниками потечет ток, проявляющийся в виде скачков на вольт-амперной характеристике такого контакта.

Аналогичные эффекты наблюдаются также в так называемых слабосвязанных сверхпроводниках, то есть сверхпроводниках, соединенных тонкой перемычкой (микромостиком или точечным контактом). На основе контактов Джозефсона и слабосвязанных сверхпроводников созданы сверхчувствительные измерительные приборы – гальванометры, магнитометры, СВЧ-детекторы и т. п.; контакты Джозефсона используются для создания быстродействующих элементов логики и памяти ЭВМ, перестраиваемых микромощных

СВЧ-генераторов и др. устройств. Использование контактов Джозефсона в цифровых устройствах (в ячейках логики и памяти) основано на представлении единицы (бита) информации электрическим током, циркулирующим в сверхпроводящем контуре. Контакты Джозефсона служат для управления этим током (его генерации, уничтожения или переключения направления).

Примером простейшей логической ячейки с управляемым контактом Джозефсона является туннельный криотрон. Конструктивно такая ячейка выполнена в виде двух шин (рис. 6.10): управляющей и шины вентиля, содержащей контакт Джозефсона. Действие ячейки основано на изменении критического значения силы тока через контакт магнитным полем, создаваемым током, протекающим через управляющую шину (управляющим током). В отсутствие управляющего тока критический ток превышает силу тока I_B в шине вентиля; в этом случае напряжение на контакте и, следовательно, сигнал на выходе ячейки отсутствуют. При подаче управляющего тока критический ток понижается до значения, меньшего I_B , что приводит к появлению напряжения на контакте и выходного сигнала.

Через согласованную линию сигнал с ячейки поступает в нагрузку, при этом согласованная линия может служить управляющей шиной для других логических элементов. Ячейка памяти с неразрушающим считыванием содержит два контакта Джозефсона в запоминающем контуре (рис. 6.10, в) и один в цепи считывания. Запись информации в такую ячейку осуществляется последовательной подачей токов в адресную и разрядную шины. При этом в зависимости от направления тока в адресной шине сверхпроводящий ток, возникающий в контуре, циркулирует либо по часовой стрелке (информация записана как «1»), либо против часовой стрелки (записана как «0»). При считывании информации токи подаются в разрядную шину и шину считывания; сигнал в цепи считывания появляется только в том случае, когда ток в контуре циркулирует по часовой стрелке.

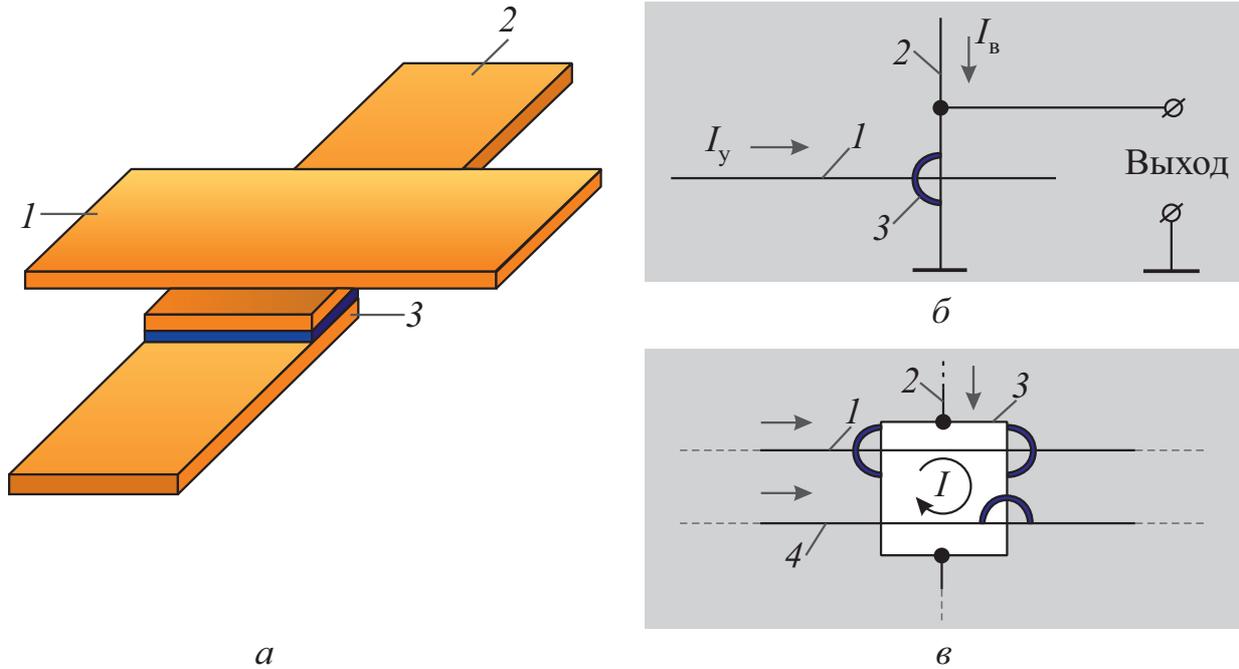


Рис. 6.10 – Конструкция логической ячейки на основе управляемого эффекта Джозефсона (а) и ее принципиальная схема (б): 1 – управляющая шина; 2 – шина вентиля; 3 – контакт Джозефсона; I_y – управляющий ток; I_B – ток в шине вентиля; (в) принципиальная электрическая схема ячейки памяти с неразрушающим считыванием на основе трех контактов Джозефсона: 1 — адресная шина; 2 – разрядная шина; 3 – запоминающий контур; 4 – шина считывания. Стрелками указаны направления токов в шинах и запоминающем контуре для случая, когда информация в ячейке записана как «1»



Основными сверхпроводящими материалами для создания контактов Джозефсона и устройств на их основе являются свинец, его сплавы (PbInAu, PbBi и др.), а также ниобий. Последний отличается высокой механической прочностью, химической стабильностью и служит для создания высококачественных и надежных контактов. В качестве разделительных материалов используются оксидные пленки сверхпроводников либо сверхтонкие прослойки полупроводниковых материалов или металлов в нормальном состоянии. Весьма перспективны для создания контактов Джозефсона синтезированные сверхпроводящие керамические материалы на основе соединений La-Sr-Cu-O, Y-Ba-Cu-O, обладающие критической

температурой 100 К и выше, что значительно расширит область технического применения эффекта Джозефсона.

.....

Первый переход Джозефсона представлял собой два свинцовых сверхпроводника, разделенных слоем диэлектрика. В настоящее время используется трехслойная структура из алюминия и ниобия. Электроны могут совершать туннельный переход даже при нулевом напряжении, приложенном к сверхпроводнику.

Если же к переходу приложено напряжение, то ток будет снижаться и колебаться на высокой частоте, пропорциональной напряжению. Отношение между приложенным напряжением и частотой настолько точно, что стандартный вольт определяется в данном случае в терминах частоты колебаний джозефсоновского перехода. Джозефсоновский переход может служить в качестве сверхчувствительного детектора слабых магнитных полей. Он также очень чувствителен к электромагнитным излучениям, например от микроволновых печей.

.....



Транзистор на эффекте Джозефсона: электрод, расположенный рядом с окисью свинца джозефсоновского перехода может влиять на переход посредством емкостной связи. Прибор, изображенный на рисунке 6.11, б, называется транзистором на эффекте Джозефсона. Главной особенностью транзистора на эффекте Джозефсона является низкая рассеиваемая мощность, что очень важно в таких устройствах, как компьютеры. Подобный транзистор является, как правило, частью более сложных сверхпроводящих приборов, таких как сверхпроводящий квантовый интерферометр, или устройств на базе сверхпроводной быстрой одноквантовой логики (БОКЛ).

.....

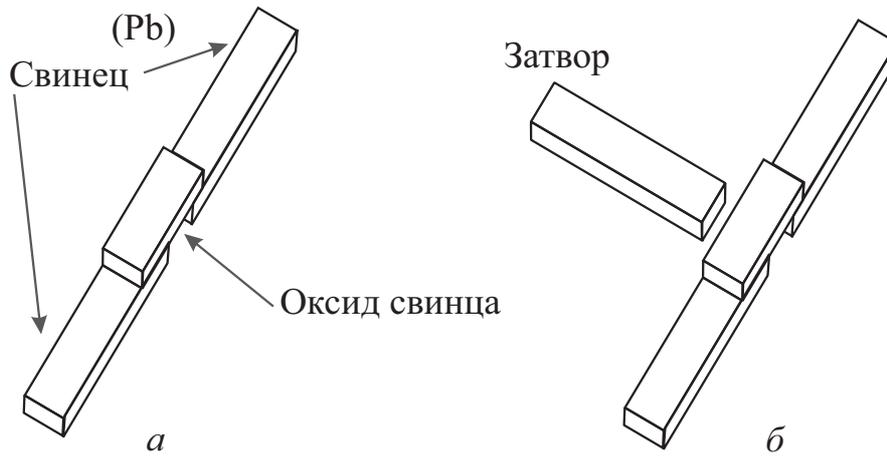


Рис. 6.11 – а) джозефсоновский переход, б) транзистор на эффекте Джозефсона

Наиболее распространенное практическое применение эффекта Джозефсона вытекает из другого прогноза, даваемого квантовой механикой. Если сделать небольшой сверхпроводящий контур с двумя встроенными переходами Джозефсона на каждом конце, а затем пропустить по нему ток, мы получим прибор под названием «сверхпроводниковый квантовый интерферометр», или СКВИД (от англ. *SQUID – Superconducting QUantum Interference Device*). В основе работы СКВИД лежат два физических явления: стационарный эффект Джозефсона и эффект квантования магнитного потока. В зависимости от интенсивности внешнего электромагнитного поля ток в его цепи может изменяться от нуля (когда токи, идущие от двух переходов, взаимно гасятся) до максимума (когда они однонаправлены и усиливают друг друга).

СКВИД, состоящий из двух переходов, включенных параллельно и работающих при постоянном токе смещения, называется СКВИД постоянного тока, такие приборы в настоящее время изготавливаются по тонкопленочной технологии.

Схема СКВИДа представляет собой замкнутый контур из сверхпроводника с четырьмя выводами, служащими для подачи тока и снятия напряжения, в который включены, как показано на рисунке 6.12, два джозефсоновских перехода.

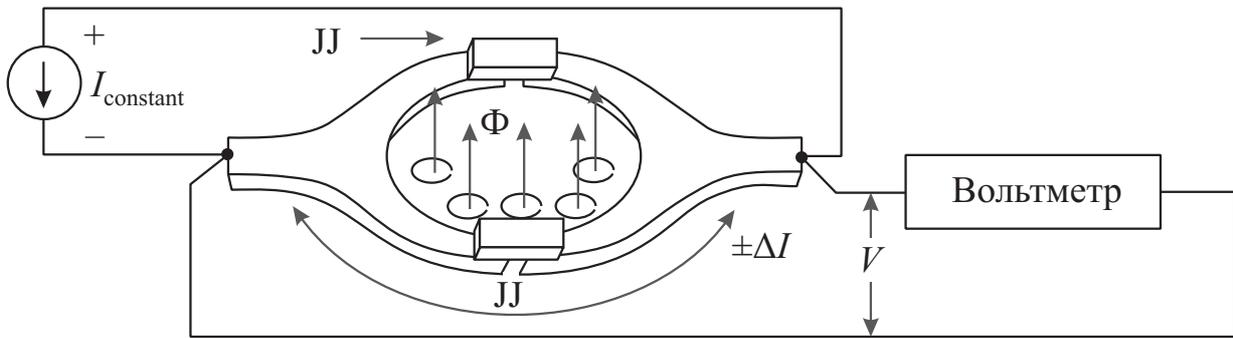


Рис. 6.12 – Два джозефсоновских перехода внутри сверхпроводникового кольца. Изменение потока вызывает изменение напряжения на паре джозефсоновских переходов. JJ – Josephson Junction – с англ. «джозефсоновский переход»

К кольцу прикладывается постоянное токовое смещение параллельно переходу Джозефсона (см. рис. 6.12). В отсутствие магнитных полей ток распределяется равномерно между двумя переходами и присоединенный к СКВИДу вольтметр не покажет падения напряжения. В работающем СКВИДе при наличии небольшого постоянного смещения в непосредственной близости от тех точек, в которых критический ток равен нулю, сверхпроводимость туннельных контактов будет разрушена и на присоединенном к интерферометру вольтметре будет показано падение напряжения. В этом случае при изменении магнитного потока даже на тысячные доли кванта (Φ_0) рядом с такой точкой на интерферометре будет достаточно напряжение, которое пропорционально величине магнитного поля. Таким образом, чувствительность интерферометра к магнитному потоку может достигать порядка $(10^{-5} - 10^{-6})\Phi_0$. С помощью преобразования различных физических величин в поток магнитного поля можно получить сверхчувствительные приборы для измерения тока, температуры или напряжения.

Сверхпроводниковый квантовый интерферометр – самый точный на сегодняшний день прибор для измерения магнитных полей, и при этом весьма компактный. Он находит самое широкое практическое применение в самых разных областях, начиная с предсказания землетрясений и заканчивая медицинской диагностикой.

6.7 Кулоновская блокада

В МДМ-структуре (металл – диэлектрик – металл) с тонким слоем диэлектрика электрический заряд переносится непрерывно в проводнике и дискретно через диэлектрик. В начальный момент времени МД-граница нейтральна, но после подачи напряжения на МДМ-структуру происходит непрерывное изменение заряда в проводниках и накопление заряда на МД-границе. Когда заряд на границе накопится достаточный для отрыва от границы одного электрона, происходит его туннелирование через диэлектрик, после чего система возвращается в первоначальное состояние. Этот процесс повторяется неоднократно, если приложенное напряжение сохраняется. Электрон туннелирует в прямом направлении, когда накопится заряд больше $+e/2$, и туннелирует в обратном направлении, когда накопится заряд меньше $-e/2$, поскольку только при этом условии электростатическая энергия системы уменьшается [66]. При промежуточных значениях накопленного заряда туннелирование невозможно из-за кулоновского взаимодействия электрона с другими подвижными и неподвижными зарядами в проводнике. Данное явление называют кулоновской блокадой (Coulomb Blockade).



.....

*Можно дать следующее определение: если кулоновское взаимодействие препятствует туннелированию электрона, то отсутствие тока через туннельный переход при наличии напряжения называется **кулоновской блокадой**.*

.....

Эффект туннелирования одного электрона используется на практике в так называемых приборах одноэлектроники (*single-electronics*). Приятно отметить, что это новое направление электроники было основано на работах отечественных ученых [88–90].

6.7.1 Кулоновская блокада с одним туннельным переходом

Рассмотрим туннельный переход «металл (M1) – диэлектрик (D) – металл (M2)» (рис. 6.13) [77]. Пусть первоначально система не заряжена. Систему можно рассматривать как конденсатор с некоторой емкостью C . Если перенести электрон e с пластины M1 на M2, то конденсатор окажется заряженным. Перенос заряда требует энергии ΔE , так как происходит против сил взаимодействия с положительным зарядом, возникающим на пластине M1 (см. рис. 6.13, а) [77].

Эта энергия равна энергии заряженного конденсатора $\Delta E = \frac{e^2}{2C}$. Величина ΔE называется *энергией одноэлектронной зарядки*. Величину $\frac{e^2}{2C}$ называют кулоновским зазором (*coulomb gap*).

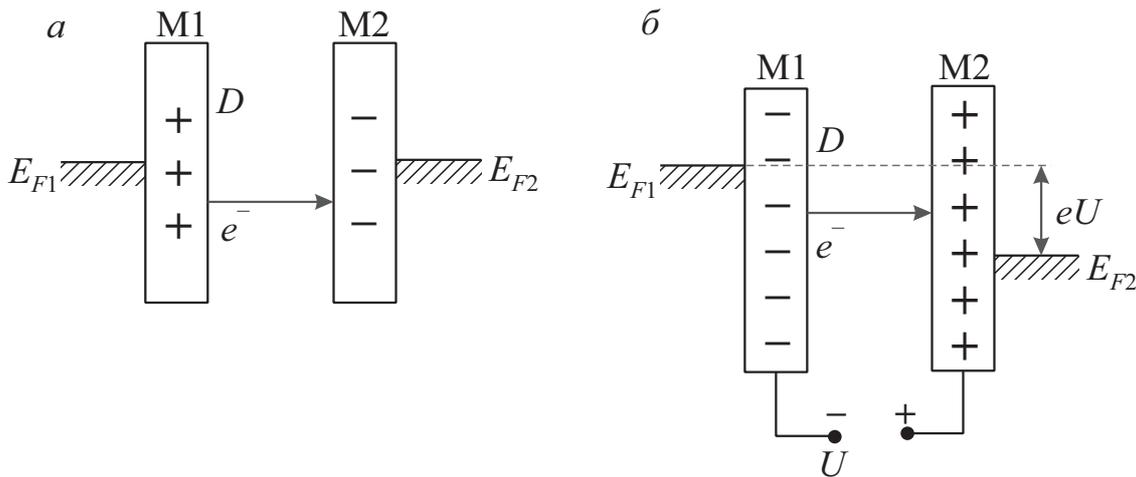


Рис. 6.13 – Схема, поясняющая механизм кулоновской блокады при туннелировании электронов

Если приложить к пластинам напряжение U (как это показано на рис. 6.13) и постепенно увеличивать его, то на пластинах станут накапливаться заряды.

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 пластины M1 на пластину M2 по величине остается меньше работы

против кулоновских сил отталкивания. Иными словами, протеканию туннельного тока препятствует кулоновская блокада.

Блокада будет «прорвана», когда напряжение U достигнет значения U_k , определяемого соотношением $eU_k = \frac{e^2}{2C}$ или $U_k = \frac{e}{2C}$.

Напряжение U_k называется напряжением кулоновской блокады.



После «прорыва блокады» туннелирует один электрон, затем снова накапливается заряд на пластинах и т. д. Электроны туннелируют через переход по одному. Наблюдение эффекта кулоновской блокады возможно в условиях, когда энергия теплового движения электронов недостаточна для преодоления блокады, т. е. $\Delta E \gg kT$ или $C \ll \frac{e^2}{2kT}$.

Подставив в последнее неравенство значения e и k , получим, что для наблюдения эффекта кулоновской блокады необходима емкость $C \ll 9 \cdot 10^{-16}$ Ф при температуре 1 К и $C \ll 3 \cdot 10^{-18}$ Ф при температуре 300 К. Для наглядности оценим радиусы сфер с соответствующими значениями C . Как известно, $C_{сф} = 4\pi\epsilon\epsilon_0 r$.

Полагая $\epsilon = 10$, получим $r_{1К} = 0,8$ мкм, $r_{300К} = 2,7$ нм. Таким образом, блокаду можно наблюдать или при температурах ниже гелиевых, или при очень малых емкостях. В обычных условиях ΔE пренебрежимо мало, так как C велико. Если $T = 300$ К, то устройство типа изображенного на рисунке 6.13, б должно для проявления кулоновской блокады иметь пластины (1) и (2) размером порядка нескольких нанометров, если толщина диэлектрика (D) составляет примерно 10 нм [77].

Второе условие наблюдения кулоновской блокады – энергия одноэлектронного заряда ΔE должна превышать квантовые флуктуации энергии $\Delta E = \frac{\hbar}{\tau}$,

где $\tau = R_T C$ – время зарядки конденсатора, а R_T – сопротивление туннельного перехода: $\Delta E \gg \Delta E'$. Последнее условие можно переписать в виде: $R_T C \gg 4R_C$, где R_C – квант сопротивления ($R_C \sim h/4e^2 \approx 6,45$ кОм) [77].

Туннельную структуру можно проанализировать на языке эквивалентных схем.

Для однобарьерной структуры с одним туннельным барьером эквивалентная схема и вольт-амперная характеристики представлены на рисунке 6.14, ее основными элементами, связанными переходом, являются R_T и C_T – емкость туннельного перехода, R_T учитывает состояние перехода до накопления заряда. C_E – эквивалентная емкость подводящей ток цепи, включая источник тока, она учитывает изменение заряда в цепях при протекании тока.

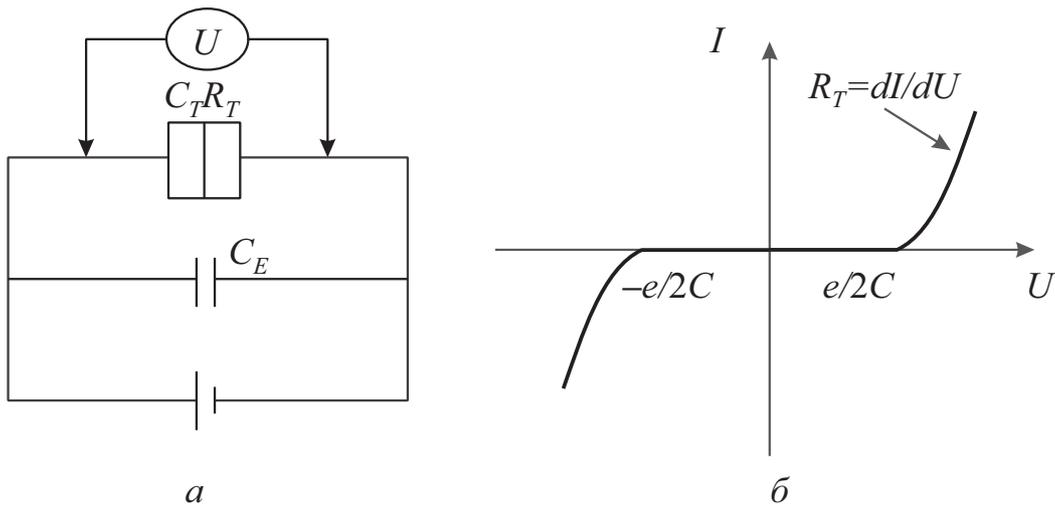


Рис. 6.14 – Эквивалентная схема (а) и вольт-амперная характеристика однобарьерной структуры (б)

Общая емкость в цепи $C = C_E + C_T$. При туннелировании одного электрона изменение электростатической энергии определяется разностью между накопленной энергией в конденсаторе и энергией, затраченной на работу источника напряжения $\Delta E = \frac{e^2}{2C} - eU$. С точки зрения минимума энергозатрат туннелирование выгодно, когда $\Delta E \leq 0$, что приводит к выражению для порогового напряжения $U_k = \frac{e}{2C}$ как при прямом, так и при обратном смещении.

Вольт-амперная характеристика однобарьерной структуры показана на рисунке 6.14, б. Электрический ток в такой структуре появляется только при ее смещении напряжением выше порогового. На «открытом» участке вольт-амперной характеристики ток определяется величиной туннельного сопротивления R_T . Связанные с одноэлектронным туннелированием осцилляции напряжения изменяют его величину на $\pm e/2C$ и происходят с частотой $f = I/e$, где I – протекающий через структуру электрический ток [66].

6.7.2 Кулоновская блокада с двумя туннельными переходами

В приборных структурах одноэлектроники используют два туннельных перехода, чтобы ослабить шунтирующее действие подводящих проводов.

На рисунке 6.15, а показан конденсатор с двумя туннельными переходами. Металлическая гранула (остров) I помещена в диэлектрическую среду (заштрихована) между двумя соединенными металлическими пластинами 2. Остров и пластины образуют две обкладки сложного конденсатора емкостью C , перенос электрона с пластин на остров (или наоборот) требует энергии [77].

На рисунке 6.15, б приведена схема для наблюдения кулоновской блокады. Металлические пластины исток (Ис) и сток (Ст), металлический островок (О). Заштрихованы туннельные диэлектрические барьеры.

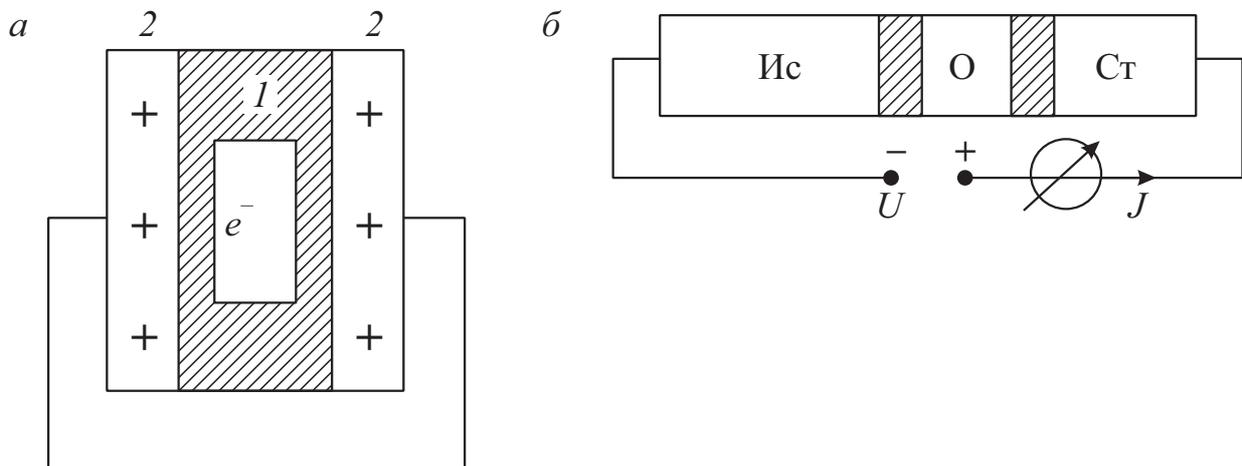


Рис. 6.15 – Структура, включающая два электрода и расположенный между ними в диэлектрической среде наноструктуры в отсутствие (а) и при наличии (б) напряжения между электродами Ис и Ст [77]

На рисунке 6.16 приведены энергетические диаграммы, поясняющие эффект кулоновской блокады. В состоянии равновесия ($U = 0$, уровни Ферми истока $E_{F_{Ис}}$, острова E_{F_0} и стока $E_{F_{Ст}}$ находятся на одной высоте; тока в структуре нет. Символами $E_{F_{0+}}$ и $E_{F_{0-}}$ обозначены уровни Ферми острова при удалении с него электрона и добавлении электрона соответственно, $E_{F_{0-}} - E_{F_{0+}} = 2\Delta E = \frac{e^2}{C}$ (рис. 6.16, а).

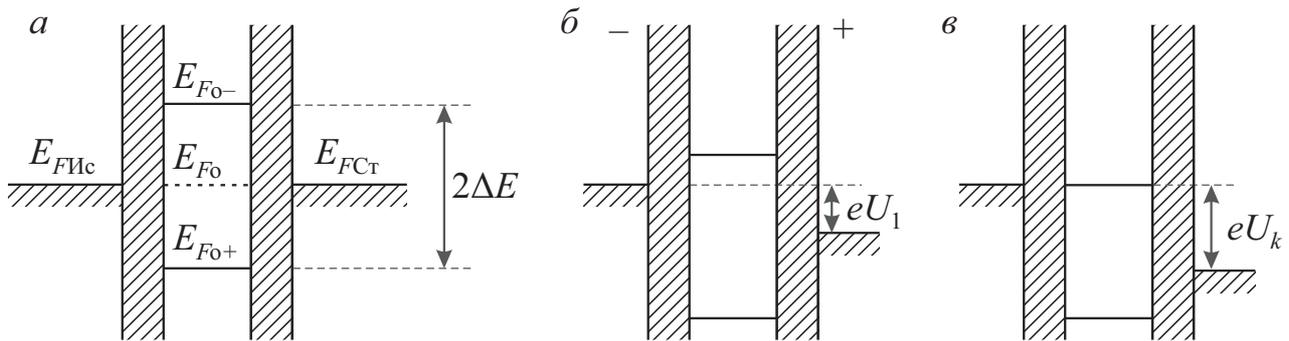


Рис. 6.16 – Энергетические диаграммы, поясняющие принцип работы элемента на основе эффекта кулоновской блокады

Рисунок 6.16, б иллюстрирует случай, когда между истоком и стоком приложено напряжение U_1 , которое не может вызвать туннелирования, ток в цепи равен нулю. Электроны туннелируют с истока на остров, а затем с острова на сток при некотором критическом значении напряжения U_k (рис. 6.16, в). В цепи появляется ток J , ток растет при дальнейшем увеличении U .

Вольт-амперная характеристика структуры показана на рисунке 6.17, а. Подаваемое напряжение измеряется единицами милливольт, одноэлектронный туннельный ток составляет десятки наноампер.

Когда напряжение U соответствует рисунку 6.16, в электроны туннелируют и проходят через остров по одному. Пока на острове есть один добавочный электрон, другой не может туда проникнуть из-за кулоновского отталкивания. Если при увеличении U между $E_{F_{Ис}}$ и $E_{F_{Ст}}$ окажется несколько зарядовых состояний острова, то возможны мультиэлектронные переходы. Но если один

из барьеров будет значительно более толстым, то электроны снова будут проходить через остров по одному.

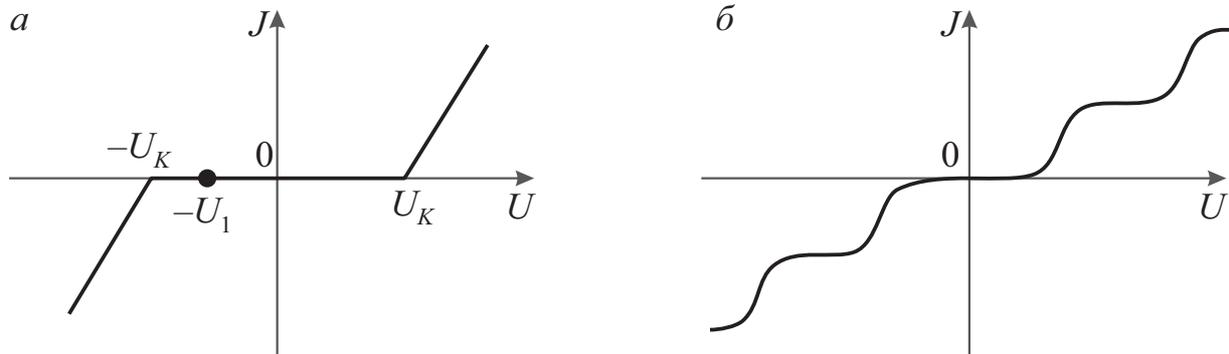


Рис. 6.17 – Вольт-амперные характеристики элемента, работающего на основе эффекта кулоновской блокады



.....

Когда один из барьеров значительно толще другого, либо один из них имеет более высокую прозрачность, вольт-амперная характеристика структуры имеет специфический, ступенчатый вид из-за различия скоростей туннелирования через первый и второй барьер (рис. 6.17, б) и называется **кулоновской лестницей** (Coulomb staircase).

.....

6.7.3 Сотуннелирование

Описанное выше одноэлектронное туннелирование, контролируемое кулоновской блокадой, может иметь место только при температуре $T = 0$ К в структуре с электрическим сопротивлением барьеров, значительно превышающим квант сопротивления ($R_T > h/e^2$). Но в условиях реального эксперимента действуют факторы, приводящие к определенным отклонениям от идеальной картины [66].

Одним из таких факторов являются квантовые флуктуации числа электронов в островке (квантовой точке). Они приводят к тому, что в островке воз-

никают виртуальные состояния, и в эти состояния из электрода-истока туннелируют электроны, энергия которых меньше энергии, необходимой для преодоления кулоновской блокады в разделяющем их барьере. Благодаря внешнему источнику напряжения уровень Ферми в электроде-стоке расположен ниже, чем в истоке. Поэтому электрон в виртуальном состоянии имеет энергию, уже достаточную для преодоления кулоновской блокады в барьере, отделяющем островок от стока, и быстро покидает островок. Такой перенос электронов происходит параллельно с одноэлектронными процессами, контролируемые кулоновской блокадой [66]. Он получил название «сотуннелирование» (*co-tunneling*), или «макроскопическое квантовое туннелирование» (*macroscopic quantum tunneling*). Термин «макроскопическое» отражает факт изменения в данном процессе такой макроскопической характеристики, как суммарный перенесенный через структуру заряд, хотя при этом очевидно нарушение ограничений, связанных с кулоновской блокадой. Различают упругое (elastic) и неупругое (inelastic) сотуннелирование.



.....

В случае упругого сотуннелирования электрон туннелирует в определенное энергетическое состояние островка и затем уходит из этого же состояния, которое в результате остается неизменным.

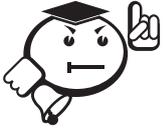
.....

Соответствующая такому переносу электронов плотность туннельного тока равна:

$$J = \frac{h\sigma_1\sigma_2\Delta}{8\pi^2e^2} \left(\frac{1}{E_1} + \frac{1}{E_2} \right) V, \quad (6.21)$$

где σ_1 и σ_2 – проводимости барьеров в отсутствие туннельных процессов; Δ – средний энергетический зазор между состояниями в островке; E_1 – энергия зарядки, связанная с добавлением одного электрона в островок; E_2 – энергия зарядки, связанная с уходом одного электрона из островка. Для упругого сотуннелирования характерно линейное соотношение между током и приложенным напряжением и отсутствие в явном виде зависимости от температуры. Эквива-

лентная проводимость в данном случае изменяется пропорционально отношению величины энергетического зазора Δ к кулоновскому зазору $e^2/2C$.



При неупругом сотуннелировании электрон, вошедший в островок, занимает одно энергетическое состояние, а покидает островок электрон с другого уровня.

Энергетическое состояние островка при этом изменяется вследствие соответствующего электронно-дырочного возбуждения, а плотность туннельного тока определяется следующим соотношением:

$$J = \frac{h\sigma_1\sigma_2\Delta}{6e^2} \left(\frac{1}{E_1} + \frac{1}{E_2} \right)^2 \left[(k_B T)^2 + \left(\frac{eV}{2\pi} \right)^2 \right] V. \quad (6.22)$$

Плотность тока J при неупругом сотуннелировании нелинейно зависит от приложенного напряжения V и температуры T и имеет две составляющие: первая линейна по V и квадратична по T , а вторая не зависит от T и пропорциональна V^3 . Обе составляющие связаны с возрастанием электронно-дырочного возбуждения в процессе туннелирования. Соотношение между упругим и неупругим сотуннелированием определяется плотностью состояний в островке. В металлах, обычно имеющих высокую плотность состояний, неупругие процессы преобладают над упругими. В полупроводниках же, где плотность состояний ниже, и те и другие вносят свой вклад в туннельный ток. Сотуннелирование проявляется на участках, контролируемых кулоновской блокадой, в виде дополнительного тока, зависящего от приложенного напряжения. Оно приводит к отличиям между экспериментальными и теоретическими вольт-амперными характеристиками одноэлектронных структур [66].

6.8 Квантовое магнетосопротивление. Спинтроника



Спинтроника (spintronics – SPIN TRansport electrONICS) – направление нанoeлектроники, «электроника на основе переноса

спина», в котором для представления и обработки информации наряду с зарядом используется спин электрона.

.....

О спине подробно говорилось в главе 2, п. 2.5–2.6, спин – словно стрелка компаса – очень чутко реагирует на переключение магнитного поля.

В современной электронике уже есть устройства, работающие на спиновых явлениях. Традиционно развитие магнетизма вообще и исследование новых магнитных материалов в особенности в значительной мере стимулируют изобретение и разработку новых устройств магнитной памяти. Это, например, выпускаемые фирмой IBM считывающие головки для магнитных дисков и новый тип магнитной памяти – MRAM (*magnetic random access memory* – магнитная память с произвольной выборкой). Устройства работают на эффектах квантового магнетосопротивления. Развитие технологии в последнее время позволило создавать ультратонкие пленки с практически совершенной кристаллической структурой и конструировать на их основе принципиально новые магнитные материалы: магнитные мультислойные структуры и сверхрешетки. Открытие эффекта гигантского магнетосопротивления в таких системах дало, прежде всего, мощный импульс работам по созданию сверхплотной памяти на магнитных носителях ($10 \text{ Гбит} \times \text{дюйм}^{-2}$) [77, 91].



***Магнетосопротивлением** (*magnetoresistance*) называется эффект изменения электропроводности материала при помещении его в магнитное поле.*

.....

Это явление в металлах известно давно и объясняется тем, что электроны проводимости в магнитном поле под действием силы Лоренца должны двигаться по спиральным траекториям, что приводит к уменьшению подвижности носителей заряда. Магнетосопротивление в металлах наблюдается в сильных полях при низких температурах.

В 1988 г. был обнаружен аналогичный, но гораздо более сильный эффект в гетероструктурах, состоящих из чередующихся наноразмерных слоев магнит-

ного и немагнитного металлов, например Fe/Cr, Co/Cu, NiFe/Cu и др. В различных гетероструктурах сопротивление может изменяться в пределах от 6 до 100%, что значительно больше, чем для обычного магнитосопротивления в тех же условиях. Так, сопротивление чередующихся слоев Fe (3 нм) / Cr (0,9 нм) изменяется в два раза при переключениях магнитного поля, поэтому эффект назван «гигантским магнитосопротивлением» (ГМС) [77]. За открытие эффекта гигантского магнетосопротивления в 2007 г. присуждена Нобелевская премия по физике проф. А. Ферту (Франция) и П. Грюнбергу (Германия).

Сущность эффекта ГМС поясняется на рисунке 6.18, где изображены чередующиеся нанометровые слои магнитного и немагнитного материалов, например Co – Cu – Co.

Есть два варианта проявления эффекта. Первый вариант – ток перпендикулярен слоям структуры (рис. 6.18, *а*, *б*). Когда ферромагнитный материал намагничен, магнитные моменты его атомов имеют одинаковое направление (вертикальные стрелки на рисунке 6.18, *а*, *б*). Неполаризованные по спину электроны из внешней цепи (ток J), проходя через первый ферромагнетик, приобретают такое же направление спина (\uparrow). Ток становится спинполяризованным (J_{\uparrow}). Если второй ферромагнетик намагничен в том же направлении (рис. 6.18, *а*), ток J_{\uparrow} свободно входит в слой этого ферромагнетика и сопротивление структуры мало. Если намагниченности ферромагнитных слоев ориентированы противоположно (рис. 6.18, *б*), то спинполяризованные электроны отражаются от второй границы, сопротивление структуры велико.

Тип магнитных материалов первого и третьего слоев не имеет принципиального значения для проявления эффекта. В практических применениях используют магнитомягкий и магнито жесткий материалы для того, чтобы под действием внешнего магнитного поля изменял направление намагниченности только один из них.

Второй вариант эффекта – ток проходит вдоль слоев структуры (рис. 6.18, *в*, *г*). Если намагниченности ферромагнитных слоев антипараллельны, происходит сильное рассеяние носителей заряда на поверхностях раздела и,

кроме того, сужается канал прохождения тока (рис. 6.18, в). В результате сопротивление структуры увеличивается. Если намагниченности слоев параллельны (рис. 6.18, з), то оба эти фактора не проявляются и сопротивление значительно уменьшается. На практике в ГМС-элементах чаще используется второй вариант эффекта, так как сопротивление нанометровых слоев в направлении, перпендикулярном слоям, очень мало.

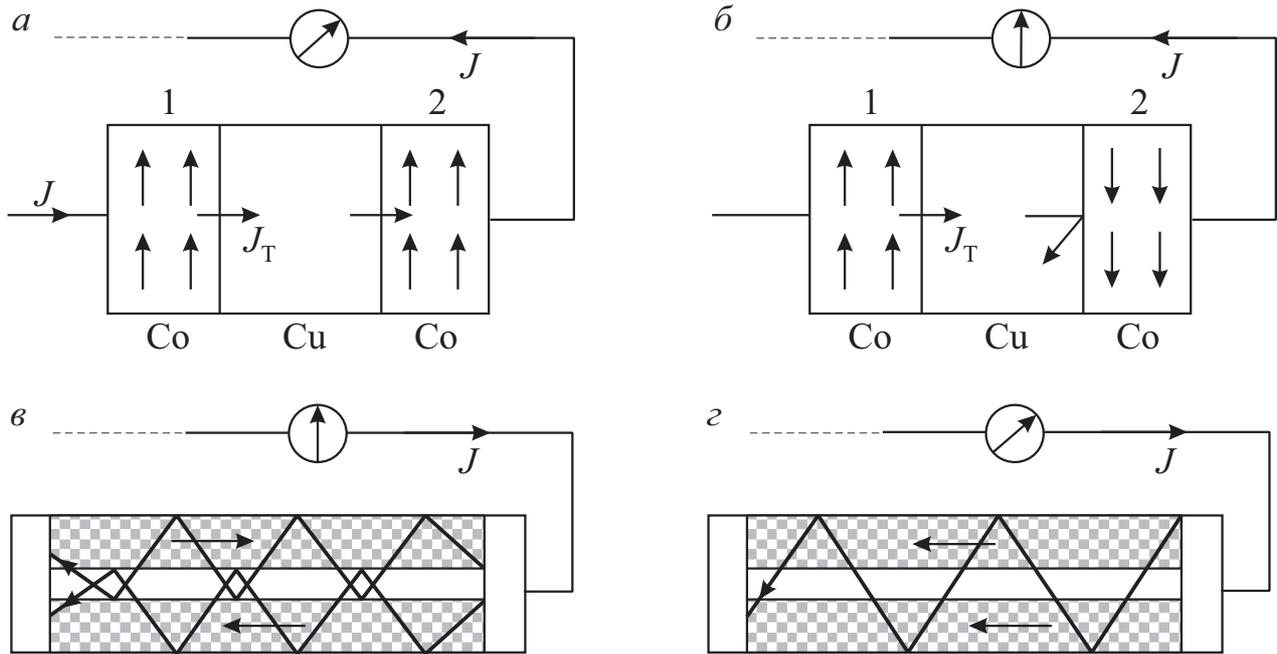


Рис. 6.18 – Схемы, поясняющие механизм ГМС

Таким образом, сопротивление структуры можно существенно изменять переключением внешнего магнитного поля. В ферромагнетике существуют небольшие области, самопроизвольно намагниченные до насыщения, их называют домены. ГМС-структуры реагируют на очень малые изменения внешних полей, например при переходе от магнитного домена одной ориентации к домену противоположной ориентации. На этом основано действие считывающих ГМС-головок. Принцип их считывания следующий. Цифровая информация (в виде битов) хранится на магнитных дисках или лентах в виде доменов. При записи намагниченность каждого домена принимает одно из двух направлений, например \rightarrow или \leftarrow . Если первое направление отождествить с нулем «0», то второе будет отражать единицу «1». Магнитное поле домена практиче-

ски не выходит за пределы внутренней области домена, за исключением участков вблизи полюсов домена.

На рисунке 6.19, *а* показаны линии вектора индукции магнитного поля \vec{B} одного домена; на рисунке 6.19, *б*, *в* приведены направления результирующего поля $\vec{B}_{\text{рез}}$ на стыках двух доменов.

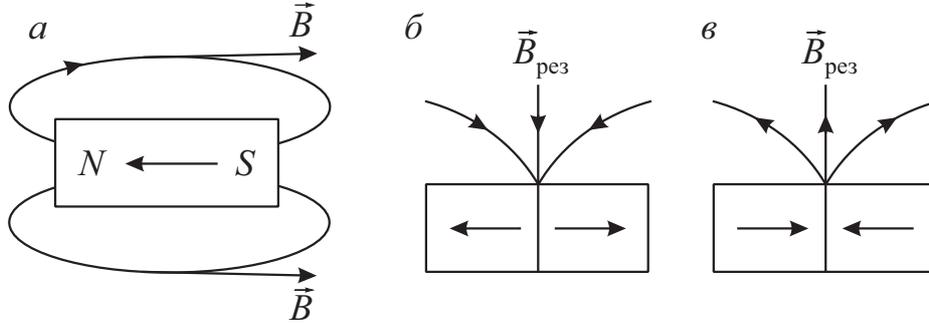


Рис. 6.19 – Направления линий вектора магнитной индукции \vec{B} одного домена (*а*) и на границе двух доменов $\vec{B}_{\text{рез}}$ (*б*) и (*в*)

Именно $\vec{B}_{\text{рез}}$ считается ГМС-головкой. Для различных материалов область междоменной стенки составляет от десяти до нескольких десятков нанометров.

На рисунке 6.20 приведена схема ГМС-головки на магнитной ленте: 1, 2, 3 – слои NiFe, Cu (2 нм), Co (2,5 нм) соответственно, 4 – магнитная лента. Стрелки на ленте обозначают направление намагниченности доменов. Намагниченность слоя Co (магнито жесткого материала) направлена стабильно перпендикулярно плоскости ленты, например вверх. Намагниченность магнитомягкого слоя NiFe принимает направление $\vec{B}_{\text{рез}}$ на каждом стыке $\uparrow\uparrow$ или $\downarrow\downarrow$ относительно слоя Co. Соответственно, ток J возрастает или уменьшается.

До открытия ГМС для намагничивания малых областей носителя (режим записи) и последующего определения направления намагниченности (режим считывания) использовались индукционные катушки. ГМС-головки много чувствительнее индукционных, поэтому емкость магнитных дисков выросла на два порядка. ГМС-элемент после переключения сохраняет намагниченность до нового переключения, поэтому он может использоваться как элемент памяти.

ГМС-элементы используются для ячеек в магнитной памяти MRAM. Развитие спинтроники должно существенно повысить быстродействие компьютеров и плотность записи информации.

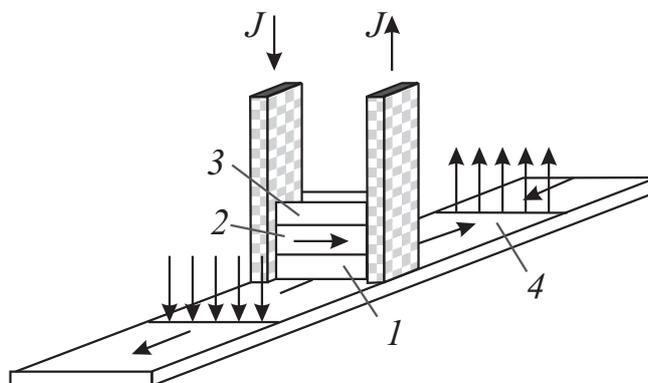


Рис. 6.20 – Схематическое изображение ГМС-головки на магнитной ленте



Туннельное магнитосопротивление (ТМС) – это эффект спин-зависимого туннелирования электронов через нанометровый слой диэлектрика или полупроводника, помещенный между двумя ферромагнетиками [77, 91].

Используемая структура имеет такой же вид, как на рисунке 6.18, а, б, но вместо слоя меди применяется слой диэлектрика (Al_2O_3) или полупроводника. Как и в случае ГМС, электроны туннелируют, создавая ток из левого ферромагнетика в правый, если их намагниченности параллельны. Если намагниченности ферромагнетиков антипараллельны, то вероятность туннелирования резко уменьшается, следовательно, падает ток через структуру, т. е. существенно увеличивается ее сопротивление. Изменение сопротивления составляет $\sim 30\%$ при комнатной температуре, что позволяет использовать рассматриваемый эффект в приборных структурах. Как в случае ГМС, применяются мягкий и жесткий ферромагнетики.

Состояние намагниченности сохраняется до нового переключения, поэтому магнитный переход может использоваться как носитель одного бита информации в электронной памяти.



.....

Ведутся разработки памяти на эффекте ТМС. В приборах на ТМС используются весьма слабые токи, поэтому такие приборы имеют низкое энергопотребление. Недостатком является то, что в них токи направлены перпендикулярно слоям. Поэтому при уменьшении площади слоев электрическое сопротивление прибора возрастает.

.....

Хотя уже созданы устройства на основе эффектов гигантского магнитного сопротивления и туннельного магнитного сопротивления, но эти устройства работают на металлах, а современные технологии ориентированы на полупроводники. Разработан ряд других спинтронных приборов, требующих спинполяризованного тока, но получение эффективной инжекции из ферромагнитного металла в полупроводник встречает большие трудности. Поэтому важное значение имеет разработка полупроводниковой спинтроники, совместимой с современной технологией чипов [77].

Для полупроводниковой спинтроники необходимы материалы нового типа – магнитные полупроводники. Как уже отмечалось в предыдущей главе, обнаружено, что ферромагнитные свойства можно придать соединениям A_3B_5 и A_2B_6 посредством легирования их ионами железа, кобальта, марганца. Однако технология легирования встречает большие трудности из-за низких пределов растворимости указанных примесей в кристаллах A_3B_5 и A_2B_6 . Кроме того, температуры Кюри (T_K) оказываются много ниже комнатной (например, $\sim 120\text{--}150$ К для $Ga_{1-x}Mn_xAs$). Ведутся разработки и исследования свойств новых магнитных полупроводников. Получены материалы с T_K выше комнатной температуры, например GaMnN, GaCrN и другие.

В современных информационных технологиях обработка информации и вычисления производятся в ИМС, а хранится информация на магнитных дисках. Применение магнитных полупроводников дало бы возможность размещать процессор и память на одном чипе, убрав медленные каналы ввода-вывода, что значительно повысило бы быстродействие.

Еще одно преимущество – магнитные полупроводники, способные воспринимать и усиливать оптические сигналы, дали бы возможность прямого преобразования информации из оптического представления в электронное, без процесса детектирования.

6.9 Эффект Кондо

Согласно классическим представлениям электрическое сопротивление металлов уменьшается при понижении температуры, так как электроны все слабее и слабее рассеиваются на тепловых колебаниях кристаллической решетки. Но в области очень низких температур возможны три различных сценария поведения сопротивления. Сопротивление большинства чистых металлов (например, Cu, Au и Al) уменьшается при понижении температуры до тех пор, пока не достигнет определенной остаточной величины (рис. 6.21, пунктирная линия). Некоторые металлы (Pb и Nb и др.) при охлаждении становятся сверхпроводящими: их сопротивление падает фактически до нуля при критической температуре T_c (рис. 6.21).



.....

*Еще в 30-х гг. XX в. было обнаружено, что сопротивление целого ряда металлов при низких температурах **начинает увеличиваться с понижением температуры** (рис. 6.21). Это могло означать только существование нового механизма рассеяния электронов, эффективность которого растет с уменьшением температуры или энергии электрона. Дальнейшие эксперименты показали, что этот эффект связан с наличием в металле магнитных примесей, то есть примесей атомов с незаполненными внут-*

ренными оболочками, сохранивших свой спин при помещении в основную немагнитный металл. Эффект наблюдался лишь при очень малой концентрации примесей ($<0,1\%$) [92].

.....

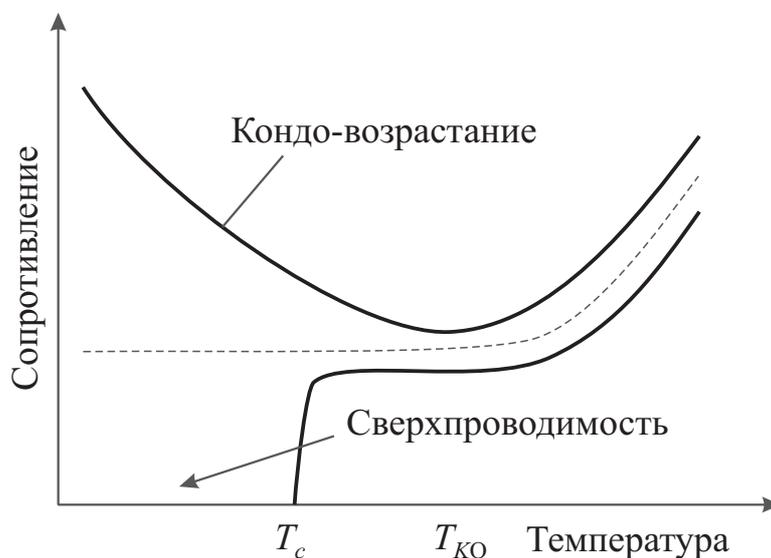


Рис. 6.21 – Типы температурных зависимостей сопротивления металлов в области низких температур (0–10 К)

В настоящее время известно большое количество сочетаний основного металла и магнитных примесей (таких как Fe, Co, Ni), в которых имеется этот эффект. В таких металлах, содержащих небольшое количество магнитных примесей, сопротивление при очень низких температурах увеличивается, что связано с ненулевым полным спином всех электронов в образце.

Количественное измерение температурной зависимости сопротивления было проделано в середине 50-х гг. XX в. Крофтом с соавторами на меди и Н. Е. Алексеевским и Ю. П. Гайдуковым на золоте с малой примесью железа. Оказалось, что при низких температурах сопротивление описывается формулой $\rho = \rho_1 + \rho_2 \ln(1/T)$. Теория этого явления была дана японским теоретиком Д. Кондо в 1964 г., после чего эффект стали называть его именем.



.....
 Температуру $T_{ко}$, при которой начинается рост сопротивления, называют **температурой Кондо** (*Kondo temperature*).

Согласно Кондо, взаимодействие электрона с примесью содержит помимо обычной потенциальной части обменный член, пропорциональный σs , где σ и s – операторы спина электрона и примеси. При расчете амплитуды рассеяния по теории возмущений в первом (борновском) приближении оба взаимодействия дают константу. Для обычного взаимодействия это справедливо и при точном расчете. Однако для обменной части второе приближение даст логарифмический член в амплитуде рассеяния и в соответствии с этим также в сопротивлении. В результате оказывается [92]:

$$\rho = \rho_V + \rho_{ex}^{(0)} \left[1 - (3J / zE_F) \ln(E_F / T) \right], \quad (6.23)$$

где ρ_V – потенциальная часть, $\rho_{ex}^{(0)}$ – обменная часть сопротивления в борновском приближении, E_F – энергия Ферми, J – энергия обменного взаимодействия, z – число валентных электронов на один атом основного металла.

Есть целый ряд явлений, в которых проявляется эффект Кондо: характеристика туннельного контакта, теплопроводность, магнитное упорядочение и другие. Эффект Кондо наблюдается не только в металлах с магнитными примесями, но и в квантовых точках.

Квантовые точки представляют собой еще один класс систем, транспортные свойства которых можно регулировать, используя эффект Кондо. По сравнению с магнитными примесями в объемных металлах для квантовых точек возможен более высокий уровень контроля параметров системы. Квантовая точка, содержащая строго определенное число электронов, может вести себя как магнитная примесь.

Полный спин квантовой точки равен нулю или целому числу при четном числе электронов в ней и полуцелому числу при нечетном числе электронов. Последний случай – классический пример для наблюдения эффекта Кондо. При

этом все электроны, кроме одного (с самой высокой энергией), можно игнорировать, что эквивалентно наличию в квантовой точке единственного носителя спина $\sigma = 1/2$.

Квантовую точку с затворами можно электрически переключать из состояния «кондо-система» в состояние «не-кондо-система», поскольку число электронов в ней может быть изменено с нечетного на четное. Схематически такая структура и ее проводимость показаны на рисунке 6.22, *а*. Квантовая точка соединена с двумя источниками электронов туннельными каналами, пропускная способность которых управляется напряжением на входном и выходном затворах. Связь точки со стоком и истоком приводит к появлению у энергетического уровня электрона в ней дополнительной составляющей ЛЕ. Число электронов в точке и их уровни энергии регулируются напряжением, прикладываемым к центральному затвору. Температуру Кондо можно изменить, меняя напряжение на затворе, поскольку при этом энергия одноэлектронного состояния приближается к уровню Ферми.

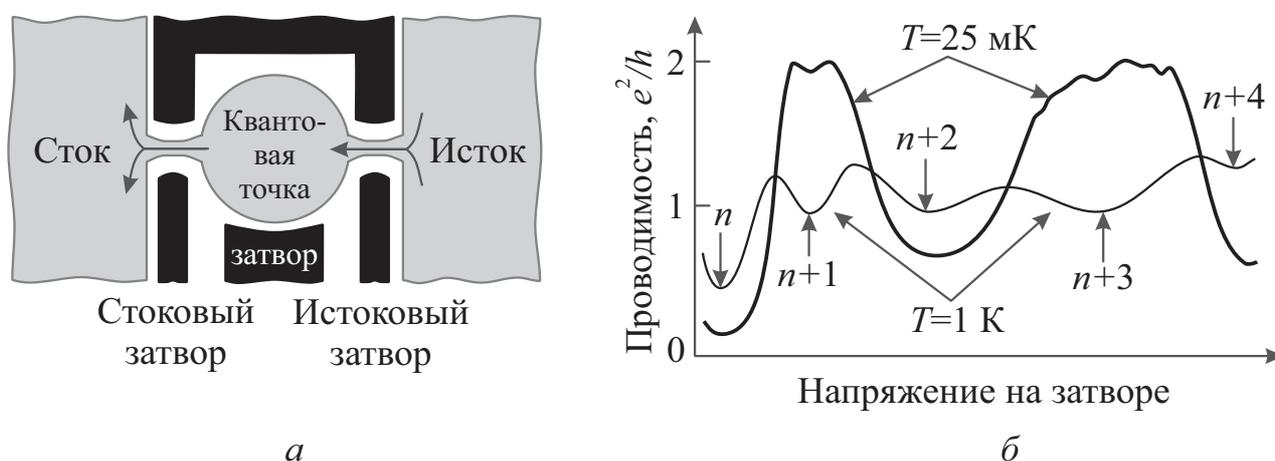


Рис. 6.22 – Управляемая квантовая точка (*а*) и ее проводимость как функция напряжения на затворе при двух различных низких температурах (*б*).

n – четное число

Отличие квантовой точки от объемного металла с магнитной примесью обусловлено, главным образом, разной природой электронных состояний в них. В металлах эти состояния представляют собой плоские электронные волны,

а энергетический спектр является квазинепрерывным. В квантовой же точке могут существовать только стоячие электронные волны, а энергетические уровни квантуются вследствие эффекта квантового ограничения. Это приводит к диаметрально противоположному характеру температурных зависимостей проводимости квантовой точки и объемного металла с магнитной примесью.

На рисунке 6.22, б показано изменение проводимости квантовой точки как функции напряжения на затворе, регулирующем число электронов, заключенных в точке при двух различных низких температурах. Когда в квантовой точке заключено четное число электронов ($n, n + 2, n + 4, \dots$), ее проводимость при понижении температуры уменьшается, что указывает на отсутствие эффекта Кондо. При нечетном числе электронов ($n + 1, n + 3, \dots$), когда эффект Кондо имеет место, наблюдается рост проводимости.

Подобно сопротивлению массивного образца в кондо-состоянии проводимость квантовой точки зависит только от отношения T/T_{KO} . При очень низкой температуре проводимость приближается к своему квантовому пределу $2e^2/h$. Это указывает на то, что электроны проходят через точку насквозь, то есть эффект Кондо делает квантовую точку полностью прозрачной. Эффект Кондо может также наблюдаться и в квантовых точках с четным числом электронов, но для этого образец нужно поместить в магнитное поле, чтобы добиться расщепления электронных состояний по спинам и заполнения соответствующих уровней энергии.



.....
Контрольные вопросы по главе 6
.....

1. Назовите особенности двумерного электронного газа. Что такое уровни Ландау?
2. Назовите и опишите пять основных квантовых эффектов (Холла, Штарка и т. д.).
3. Что такое кулоновская блокада и кулоновская лестница?
4. Что называется эффектом гигантского магнитосопротивления?
5. Какие материалы необходимы для полупроводниковой спинтроники?

7 Устройства наноэлектроники

7.1 Приборы на одноэлектронном туннелировании



.....

Структуры на основе эффекта одноэлектронного туннелирования (кулоновской блокады) являются перспективными для создания широкого спектра твердотельных приборов, в том числе интегральных схем нового поколения сверхвысокой степени интеграции. Известно большое количество структур рассматриваемого типа различной конфигурации и назначения, и число публикаций в данном направлении продолжает возрастать. Становится достаточно сложно ориентироваться в этой области, так как в настоящее время полная классификация приборов одноэлектроники отсутствует. Попытка классификации приборных одноэлектронных структур предпринята в работе [93]. Достоинством классификации является то, что известные в настоящее время структуры рассматриваемого типа могут быть описаны с ее использованием, а также то, что на ее основе могут быть предложены новые приборы одноэлектроники.

.....



.....

Одноэлектронными называются приборы, в которых контролируется перемещение одного электрона или малого их количества.

.....

Эти устройства представляют особый интерес в связи с приближением технологии к пределу миниатюризации элементов классических интегральных микросхем. Одноэлектронные приборы обеспечивают дальнейшее увеличение быстродействия, снижение габаритов и энергопотребления электронной аппа-

ратуры. В этих приборах перемещение электронов происходит посредством туннельного эффекта. Так как время туннелирования мало, то предел быстродействия высок (10^{14} Гц). Работа, необходимая для перемещения одного электрона, мала – энергопотребление и тепловыделение должны быть низкими (теоретически – 10^{-8} Вт для одного элемента) [77].

7.1.1 Одноэлектронный транзистор

По предложенной классификации к классу «Цепочки туннельных переходов» относятся структуры, содержащие два и более туннельных переходов в активной области, соединенные последовательно.



.....

Один из наиболее изученных приборов из этого класса – одноэлектронный транзистор. Он содержит два туннельных перехода, отделяющих малый «островок» полупроводника от областей истока и стока.

.....

В настоящее время созданы транзисторы, в которых переключение происходит под действием одного электрона, и запоминающие устройства, в которых носителем бита информации является один электрон. Для сравнения: в современных устройствах переключение (переход между состояниями «0» и «1») требует не менее 10^5 электронов.

Принцип действия *одноэлектронного транзистора* основан на явлении кулоновской блокады.

Как и в обычном транзисторе, имеются три электрода: исток («Ис»), сток («Ст») и затвор («З») (рис. 7.1). Между истоком и стоком приложено небольшое напряжение U . Напряжение между истоком и затвором равно U_3 , на затвор подается «+».

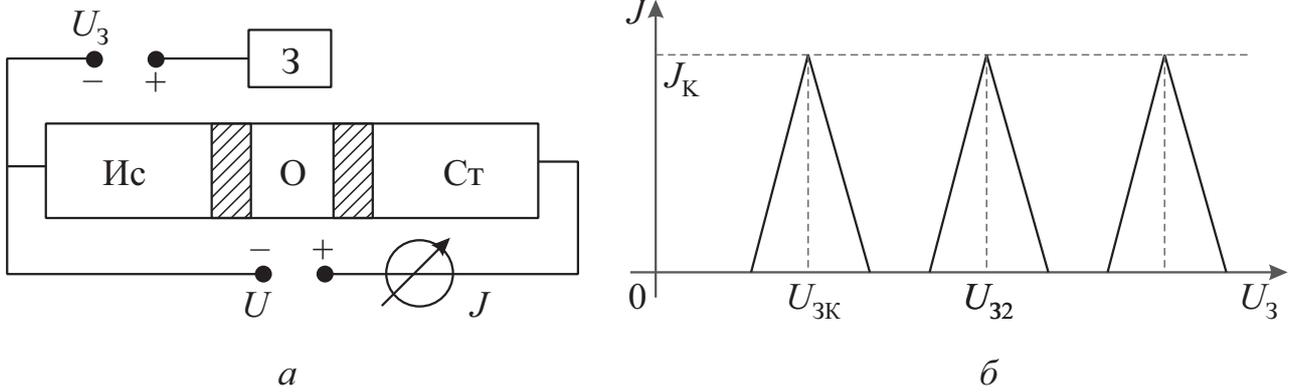


Рис. 7.1 – а) принципиальная схема транзистора, работающего на эффекте кулоновской блокады; б) зависимость тока от напряжения на затворе

На рисунке 7.2 приведены энергетические диаграммы, поясняющие работу транзистора. Обозначения такие же, как на рисунке 6.16. Цифры над уровнями Ферми острова ($-1, 1, 2, \dots$) означают число электронов, перешедших на остров с истока. При этом заряд острова составляет $-e, -2e, -3e, \dots$ соответственно; (-1) означает, что электрон удален с острова. Заряд острова в таком случае равен $+e$.

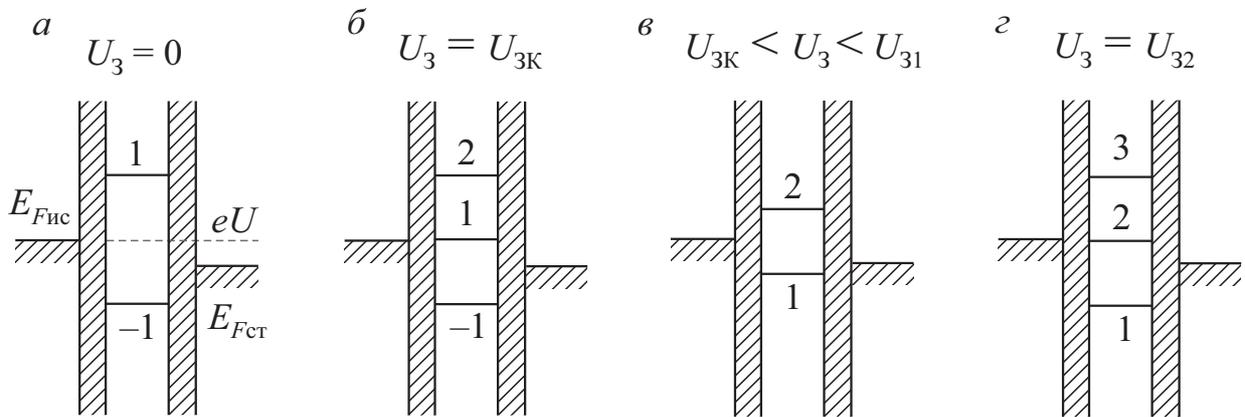


Рис. 7.2 – Энергетические диаграммы, поясняющие работу транзистора, схема которого представлена на рисунке 7.1

Поле положительно заряженного затвора смещает уровни острова вниз (уменьшает энергию электронов острова). Если напряжение $U_3 = 0$, то хотя $E_{F_{Ст}} < E_{F_{Ис}}$ (рис. 7.2, а), кулоновская блокада подавляет туннелирование электронов с истока и поэтому $J = 0$. При $U_3 = U_{3К}$ (рис. 7.2, б) кулоновская

блокада прорвана, электроны туннелируют с истока на остров, а затем на сток – в цепи появляется ток J . При дальнейшем увеличении напряжения U_3 и смещении уровня (1) вниз снова возникает блокада (рис. 7.2, в), ток отсутствует, но на острове находится один добавочный электрон. При $U_3 = U_{32}$ ток возникает снова.

По мере увеличения U_3 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 рисунке 7.1, б показан вид зависимости тока от напряжения на затворе транзистора. Возникающие скачки тока называются *осцилляциями кулоновской блокады*. Напряжение U_3 имеет порядок единиц милливольт, ток составляет десятки пикоампер.

Таким образом, транзистор может находиться в состояниях «включено» и «выключено» в зависимости от напряжения на затворе. Поэтому одноэлектронный транзистор может использоваться как переключатель. При прохождении электрона сменяются состояния «включено» и «выключено». Теоретически время туннелирования равно $\sim 10^{-15}$ с, этим определяется верхний предел быстродействия прибора (ТГц). Так как одноэлектронный транзистор может находиться в двух состояниях, то на его основе могут быть созданы ячейки памяти с малыми размерами, высоким быстродействием и низким энергопотреблением. Проекты таких ячеек уже имеются [77].

В этом разделе рассмотрен металлический транзистор. Разработано и реализовано много вариантов полупроводниковых одноэлектронных транзисторов. В таких транзисторах островом служит квантовая точка, отделенная обедненными областями – барьерами от истока и стока. Подобный транзистор можно реализовать в едином кристалле. В настоящее время в ряде лабораторий получены образцы одноэлектронных полупроводниковых и металлических транзисторов, работающих при комнатной температуре.

7.1.2 Одноэлектронный насос (стандарт силы тока).

Одноэлектронная память

Большинство других известных в настоящее время одноэлектронных приборов относятся также к классу «Цепочки туннельных переходов»: «одноэлектронный насос», одноэлектронная память и другие.

С помощью одноэлектронных эффектов удастся определить единицу силы тока со значительно большей точностью, чем в донаноэлектронную эпоху, что с успехом можно использовать для высокоточного стандарта силы тока. В качестве стандартов единицы сопротивления (Ом) и единицы напряжения (вольт) могут служить соответственно квантовый эффект Холла и эффект Джозефсона [78].

Для стандарта силы тока можно использовать структуру из квантовой точки, связанной туннельными переходами с истоком и стоком.

Планарную квантовую точку выделяют из области двумерного электронного газа с помощью системы затворов (рис. 7.3). Затвор 3 аналогичен затвору одноэлектронного транзистора и служит для изменения электростатической энергии квантовой точки. Затворы 2 и 4 совместно с затвором 1 определяют высоту, а значит, и прозрачность туннельных барьеров, через которые происходит перенос электронов. Принципиальное отличие такой структуры от обычного одноэлектронного транзистора состоит в том, что с помощью напряжения на затворах 2 и 4 высоту левого и правого туннельных барьеров можно изменять независимо.

Между истоком и стоком структуры прикладывают небольшое напряжение смещения V_C , такое, чтобы в квантовую точку можно было поместить только один электрон: $eV_C > E(N+1) - E(N)$ (рис. 7.4, а). На затворы 2 и 4 подают в противофазе переменное напряжение с частотой f . Таким образом, когда повышается один барьер, другой в этот момент понижается.

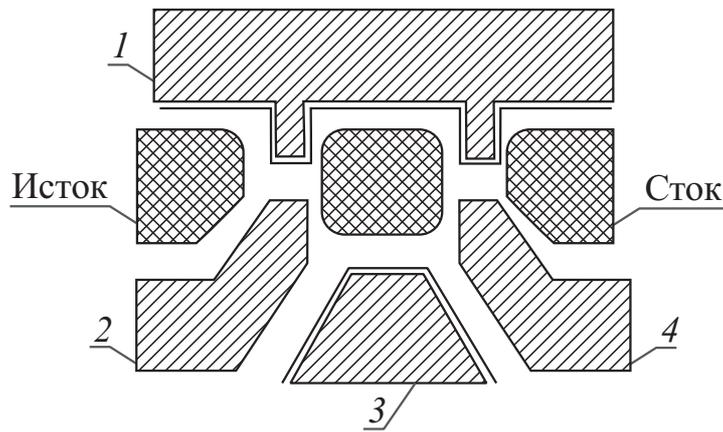


Рис. 7.3 – Расположение поверхностных затворов (1–4) на структуре с двумерным электронным газом для создания стандарта силы тока [78]

Прохождение электрона от левого электрода к правому через квантовую точку в течение периода изменения напряжения на затворах 2 и 4 продемонстрировано на рисунке 7.4, а–г. Через пониженный левый барьер электрон проникает в квантовую точку, но выйти из нее в правый электрод не может из-за низкой вероятности туннелирования через высокий правый барьер (рис. 7.4, б). Только лишь в следующий полупериод переменного напряжения на затворах электрон покидает квантовую точку и переходит в правый электрод (рис. 7.4, г).



.....

Таким образом, за один период изменения напряжения на затворах 2 и 4 через квантовую точку проходит только один электрон. По этой причине такую структуру называют *одноэлектронным насосом*.

.....

Ток, протекающий через точку, будет $I = ef$. Величина заряда электрона известна с очень высокой степенью точности, измерение частоты f сводится к измерению интервалов времени, что можно сделать тоже достаточно точно. В результате оказывается возможным измерить силу тока с точностью, значительно превосходящей точность других стандартов тока.

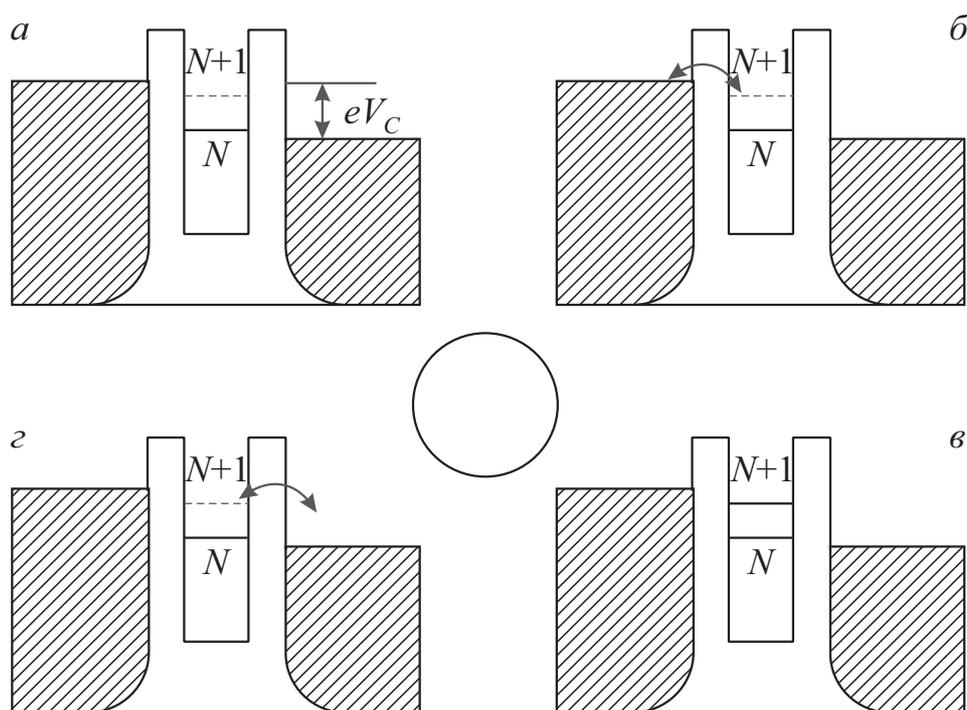


Рис. 7.4 – Схема прохождения одного электрона через квантовую точку при периодической модуляции высоты туннельных барьеров [78]

Ток, протекающий через точку, будет $I = ef$. Величина заряда электрона известна с очень высокой степенью точности, измерение частоты f сводится к измерению интервалов времени, что можно сделать тоже достаточно точно. В результате оказывается возможным измерить силу тока с точностью, значительно превосходящей точность других стандартов тока.



.....

Ячейка памяти также может быть реализована на основе одноэлектронных транзисторов. Одноэлектронный транзистор может находиться в двух состояниях – либо в состоянии кулоновской блокады, либо в проводящем состоянии. Этот факт позволяет создавать на его основе устройства, которые могут служить ячейкой компьютерной памяти.

.....

Очень малые размеры одноэлектронных транзисторов при высоком быстродействии и низкой потребляемой мощности делают такое их применение очень перспективным. Рассмотрим одну из возможных конструкций ячейки

памяти на основе одноэлектронных эффектов (рис. 7.5). В качестве активных элементов в ней используются три одноэлектронных транзистора на основе множественных туннельных переходов (*MTJ* – *Multi-Tunnel Junctions*), представляющих собой структуры с двумя и более центральными электродами, отделенными друг от друга туннельно-прозрачными барьерами. В таких структурах подавляется сотуннелирование и вероятность мгновенного переноса заряда из истока в сток практически равна нулю [78].

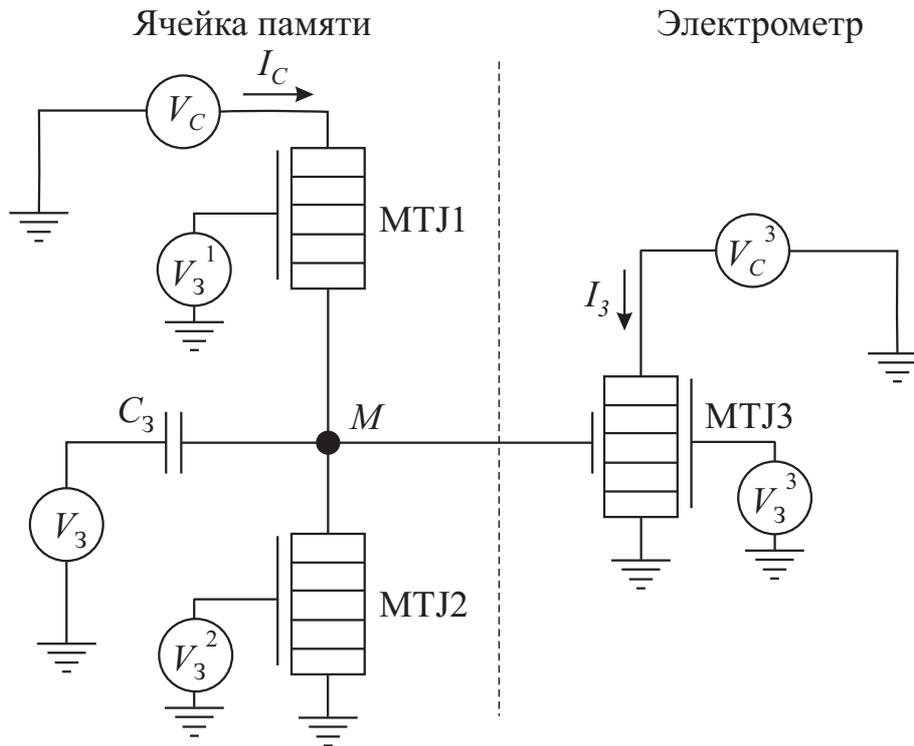


Рис. 7.5 – Структурная схема ячейки памяти на одноэлектронных транзисторах

Основными действующими элементами ячейки являются транзисторы MTJ1 и MTJ3. Транзистор MTJ2 играет вспомогательную роль и при работе ячейки находится в выключенном состоянии (в режиме кулоновской блокады). Транзистор MTJ3 имеет два затвора. С помощью напряжения V_3^3 на правом затворе рабочую точку транзистора устанавливают на пороге его открывания, поэтому очень малые изменения потенциала второго затвора, соединенного с узлом запоминания M , приводят к увеличению или уменьшению тока I_C через

МТJ3. Таким образом, этот транзистор выполняет функции электрометра – с его помощью можно регистрировать малые изменения заряда в точке M .

На рисунке 7.6 представлена зависимость тока электрометра I_3 от напряжения на основном затворе ячейки V_3 . Гистерезис этой зависимости (направления изменения V_3 указаны стрелками) связан с различным состоянием транзистора МТJ1 на разных участках его петли. На участке AB он закрыт – находится в режиме кулоновской блокады [78].

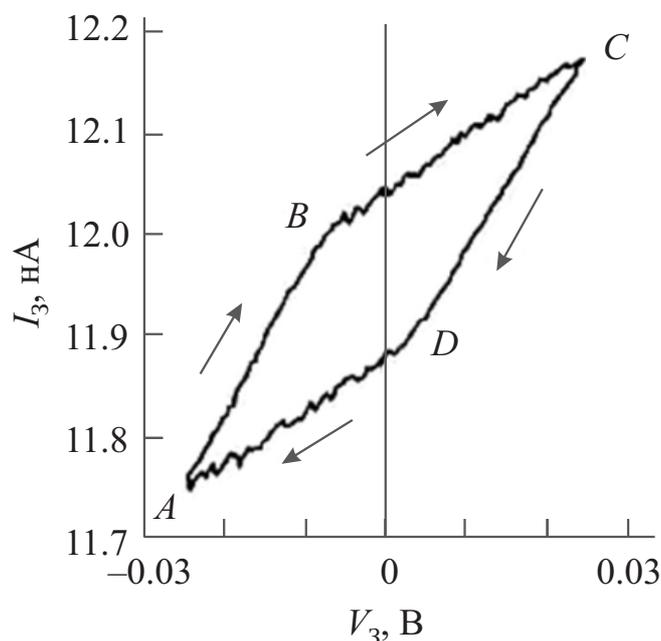


Рис. 7.6 – Гистерезис в зависимости тока электрометра при различных направлениях изменения напряжения V_3 на затворе ячейки памяти

Скорость накопления заряда в точке M большая, она определяется параметрами емкостного делителя, образованного емкостью C_3 и емкостями транзисторов МТJ1 и МТJ2. Однако при потенциале в узле M , соответствующем точке B кривой, транзистор МТJ1 открывается – выходит из режима кулоновской блокады. В результате скорость накопления заряда в узле замедляется – заряд стекает через открытый транзистор. При смене направления изменения V_3 в точке C кривой транзистор МТJ1 закрывается, и скорость изменения заряда снова будет определяться только емкостным делителем. В точке D кривой транзистор МТJ1 снова открывается – при другой полярности снятия кулонов-

ской блокады. Изменение потенциала в точке M снова замедляется. Таким образом, при $V_3 = 0$ ячейка имеет два устойчивых состояния, которые могут быть использованы для записи бита информации [78].

7.2 Приборы на резонансном туннелировании

7.2.1 Диоды на резонансном туннелировании

Полезными для нанoeлектроники особенностями обладает *резонансный туннельный эффект*. Он 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 (рис. 7.7, в). 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 электроном структуры включает, помимо времен туннелирования, время пребывания электрона в яме, разделяющей барьеры, то есть время его жизни τ на резонансном уровне. Например, оценка показывает для двойной гетероструктуры, состоящей из слоев $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ (5 нм) – GaAs (7 нм) – $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ (5 нм) при высоте барьеров 0,2 эВ, время $\tau \sim 8 \cdot 10^{-12}$ с, то есть все же достигается терагерцевый диапазон.

Величина τ уменьшается при дальнейшем уменьшении размеров структуры. Двухбарьерные структуры представляют большой интерес для электроники, на их основе могут работать СВЧ-приборы в диапазоне сотен ГГц и переключатели с задержкой менее 1 пикосекунды. Созданы приборы на основе двухбарьерной структуры – резонансно-туннельные диод и транзистор. Разработаны и находят применение многобарьерные структуры (сверхрешетки).

Вариант структуры резонансно-туннельного диода изображен на рисунке 7.7, а. На рисунке 7.7, б, в, г приведены энергетические диаграммы, поясняющие работу диода. Основные части диода: 1 и 5 – сильнолегированные слои n^+ -GaAs, (эмиттер и коллектор), толщина этих слоев лежит за пределами нано-

метрового диапазона; 3 – квантовая яма толщиной $d_3 = 3\text{--}10$ нм, слой (из слаболегированного $n\text{-GaAs}$): 2, 4 – барьеры из $\text{Al}_x\text{Ga}_{1-x}\text{As}$ толщиной $d_2 = d_4 = 2\text{--}5$ нм.

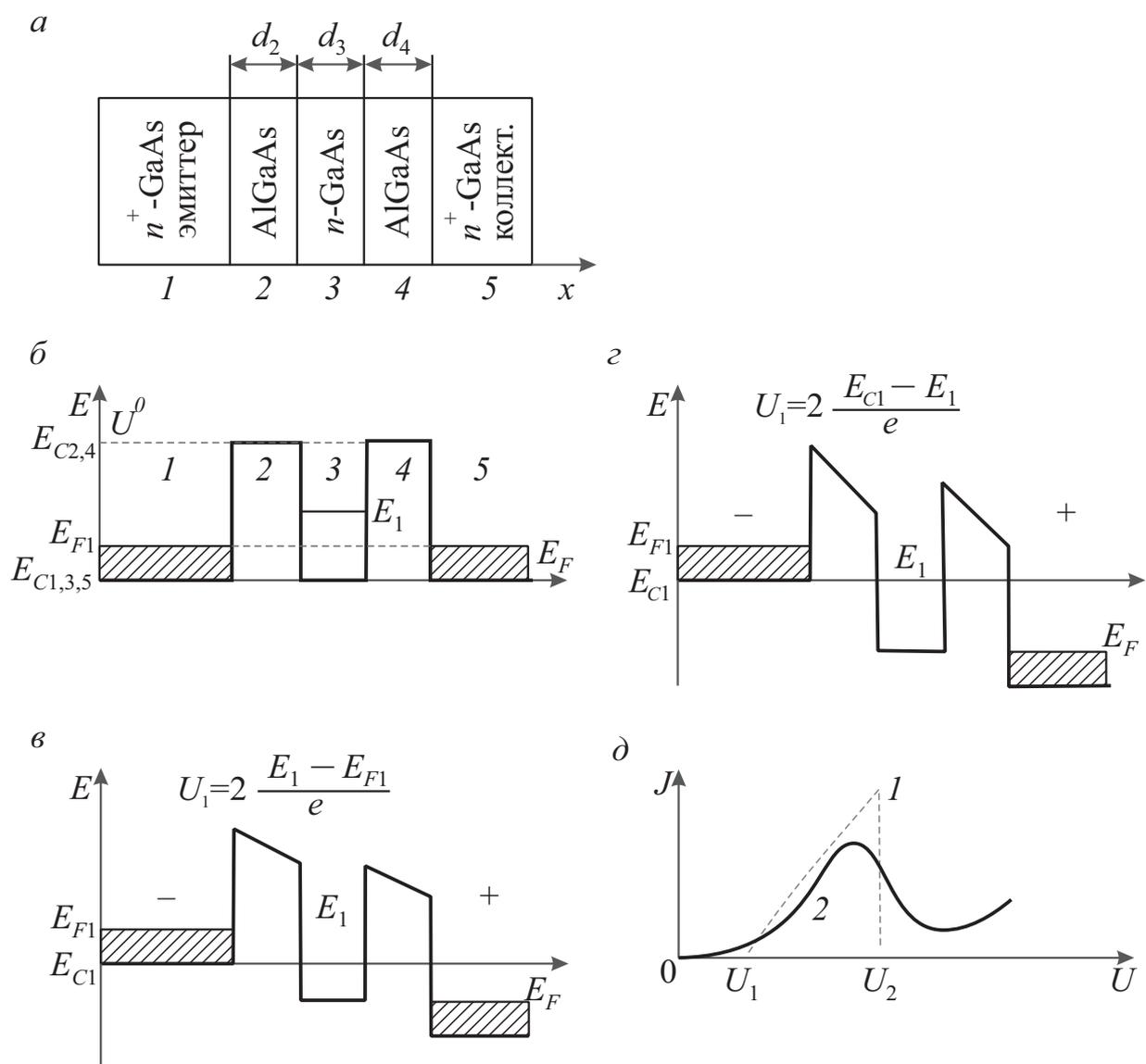


Рис. 7.7 – Структура (а), энергетические диаграммы (б, в, г) и вольт-амперные характеристики (д) резонансно-туннельного диода: пунктир на рисунке 7.7, д – идеализированная характеристика, сплошная кривая – реальная характеристика

Высота барьеров U^0 зависит от концентрации алюминия x . Величина U^0 возрастает от 0,2 эВ при $x = 0,3$ до 0,35 эВ при $x = 1$. Предполагается, что высота барьера U^0 и ширина потенциальной ямы d_3 подобраны так, что в яме образуется только один размерный квантовый уровень E_1 . Величина $E_{C1,5}$ соответствует дну зоны проводимости арсенида галлия GaAs ; величина $E_{C2,4}$ – дну

зоны проводимости соединения AlGaAs; E_{F1} , E_F – уровни Ферми в сильнолегированных слоях GaAs. В сильнолегированных n -полупроводниках уровень Ферми располагается в зоне проводимости, и все уровни от дна зоны проводимости до уровня Ферми при низких температурах заняты электронами. На рисунке 7.7, б, в, г штриховкой показаны заполненные уровни.

Энергетическая диаграмма, представленная на рисунке 7.7, б соответствует состоянию равновесия структуры. Напряжение между эмиттерами и коллектором $U = 0$, ток в структуре отсутствует.

Если приложить небольшое напряжение U и постепенно повышать его, то через структуру пойдет слабый ток, обусловленный обычным туннелированием электронов через потенциальный барьер 2 (см. участок $0-U_1$ на рисунке 7.7, д). При этом энергетические уровни эмиттера 1 будут подниматься относительно уровней коллектора 5 (полагаем, что напряжение падает только на барьерах). Значительно больший ток через структуру, связанный с резонансным туннелированием, появится при напряжении U_1 когда уровень E_{F1} (последний уровень в эмиттере, заполненный электронами) сравняется с размерным уровнем E_1 (см. рис. 7.7, в). На рисунке видно, что U_1 определяется положениями уровней E_1 и E_{F1} , и можно записать $U_1 = 2(E_1 - E_{F1})/e$.

С увеличением U ток I резко возрастает до тех пор, пока с размерным уровнем E_1 не сравняется дно зоны проводимости E_{C1} (см. рис. 7.7, г). Это произойдет, если $U_2 = 2(E_{C1} - E_1)/e$.

В этом выражении E_{C1} – исходное положение дна зоны проводимости в слое 1 (см. рис. 7.7, в). При дальнейшем увеличении U уровень E_1 опускается ниже дна зоны проводимости, в запрещенную зону, где электронов нет, и ток резко падает (вблизи $U = U_2$ см. рис. 7.7, б).

На рисунке 7.7, д показаны вольт-амперные характеристики (ВАХ) диода: 1 – идеализированная, 2 – реальная. На ВАХ имеется падающий участок при $U = U_2$, то есть участок с отрицательным дифференциальным сопротивлением. Это позволяет использовать резонансный туннельный прибор для генерации

СВЧ-колебаний с частотами выше (на порядок и более), чем в обычных туннельных диодах. В экспериментах достигнута частота генерации 700 ГГц. Величинами U_1 и U_2 можно управлять, изменяя ширину ямы, что приводит к изменению положения размерного уровня.

На основе резонансно-туннельных диодов разработаны аналого-цифровые преобразователи на несколько ГГц, логические элементы, запоминающие устройства, которые можно использовать в процессорах, и другие цифровые устройства для быстродействующей электроники [77].

При обсуждении преимуществ резонансно-туннельных диодов заслуживает внимания их сравнение с диодами Есаки (туннельными диодами) [77]. Диоды Есаки состоят из сильно легированного p - n -перехода и имеют подобные вольт-амперные характеристики. Одно из наиболее важных преимуществ резонансно-туннельных диодов – это возможность получить высокую плотность максимального тока при относительно низкой емкости. Так, для резонансно-туннельных диодов достигнута экстремально высокая плотность тока $6,8 \cdot 10^5$ А/см² при удельной емкости около $1,5 \cdot 10^{-7}$ Ф/см². Эти величины указывают на то, что показатель скорости, который определяется отношением удельной емкости к пиковой плотности тока C/J_p , меньше чем 0,22 пс/В.

Показатель скорости соответствует скорости изменения напряжения, когда емкость диода заряжается его максимальным током. Эта величина намного меньше, чем для диодов Есаки, которая для последних больше чем 10 пс/В. Возможность такого различия связана с тем, что плотность тока в резонансно-туннельных диодах можно увеличить изменением толщины барьера и ямы, а это может быть достигнуто без уменьшения толщины обедненного слоя. С другой стороны, для того чтобы увеличить плотность тока в диодах Есаки, должна быть увеличена концентрация примеси для увеличения толщины туннельного барьера (равно как и обедненного слоя). Следовательно, максимальное быстродействие резонансно-туннельных диодов может быть намного больше, чем у диодов Есаки. Более того, в резонансно-туннельных диодах

можно избежать деградации, наблюдаемой в диодах Есаки из-за диффузионного перераспределения примесей вблизи сильнолегированного p - n -перехода.

7.2.2 Транзисторы на резонансном туннелировании



Резонансно-туннельный диод преобразуется в резонансно-туннельный транзистор (*resonant tunneling transistor*), если к центральной части структуры (см. рис. 7.7, а) подвести электрод, регулирующий положение размерного уровня E_1 с помощью электрического поля.

Условное обозначение резонансно-туннельного транзистора в электрических схемах и его вольт-амперная характеристика изображены на рисунке 7.8.

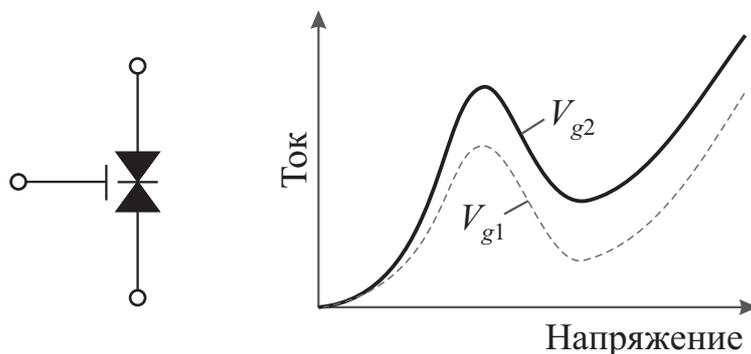


Рис. 7.8 – Условное обозначение резонансно-туннельного транзистора в электрических схемах и его вольт-амперные характеристики

Потенциал, подаваемый на дополнительный электрод, смещает вольт-амперную характеристику диода вдоль оси тока. Возможна также комбинация резонансно-туннельного диода с обычным транзистором. Этот вариант использован для создания резонансно-туннельного биполярного транзистора (*resonant tunneling bipolar transistor*) и резонансно-туннельного транзистора на горячих электронах (*resonant tunneling hot electron transistor*). Резонансно-туннельный биполярный транзистор представляет собой биполярный транзистор с резонансно-туннельной структурой, встроенной в область перехода эмиттер–база

или в саму базу. У резонансно-туннельного транзистора на горячих электронах эта структура встраивается в эмиттер. Такие приборы имеют отрицательную крутизну вольт-амперной характеристики в схеме включения с заземленным эмиттером.



Кроме перечисленных приборов существуют также транзисторные структуры, представляющие собой управляемые затвором резонансно-туннельные диоды (*gated resonant tunneling diodes*). С целью управления условиями резонанса тока затворы в них изготавливают в виде барьеров Шоттки или p - n -переходов вокруг эмиттера для внешнего управления условиями резонанса тока в диоде.

На рисунке 7.9 показан пример такого прибора. В нем область эмиттера (а следовательно, и эмиттерный ток) можно модулировать с помощью потенциала на затворе, выполненного в виде окружающего эмиттер p - n -перехода. Такое управление эмиттерным током позволяет управлять максимальным током, протекающим через структуру в резонансных условиях.

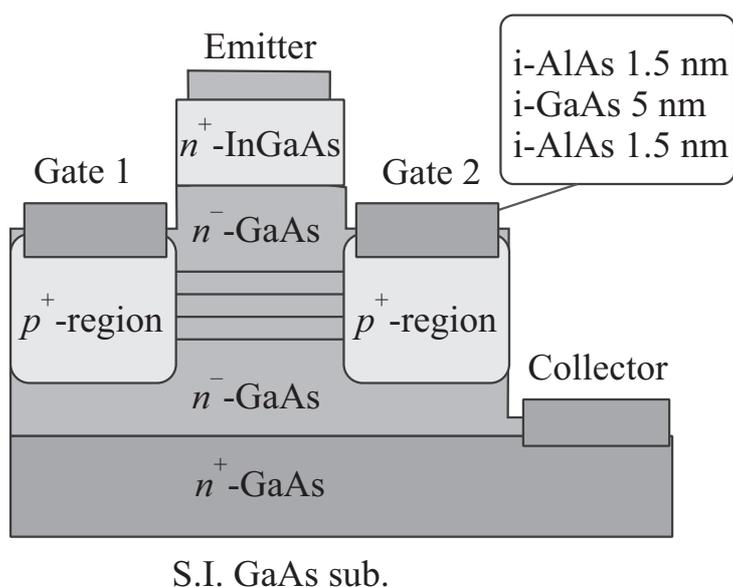


Рис. 7.9 – Управляемый затвором резонансно-туннельный диод на основе GaAs – AlAs

7.2.3 Логические элементы на резонансно-туннельных приборах



Принципы построения оригинальных сверхбыстродействующих логических вентилях были предложены в работе [95]. В основе работы логического вентиля использован переход из моностабильного в бистабильное состояние (*monostable-bistable transition logic elements – MOBILE*) схемы, состоящей из двух последовательно соединенных приборов с отрицательным дифференциальным сопротивлением. При этом по крайней мере один из приборов должен иметь третий вывод для модуляции пикового тока.

Устройство MOBILE состоит из двух резонансно-туннельных транзисторов (РТТ), соединенных последовательно и управляемых переключением напряжения смещения V_{bias} [95] (будем использовать обозначения, принятые в оригинальной работе [95]).

Схематическая структура поперечного сечения РТТ с затвором из $p^+ - n$ -перехода показана на рисунке 7.10, а [95].

Напряжение затвора изменяет ширину обедненного слоя и изменяет сечение для прохождения тока резонансного туннелирования. Количество входных терминалов поэтому может быть увеличено делением электрода затвора. Эффективность модуляции пикового тока будет пропорциональна ширине каждого затвора. Эпитаксиальная структура выращена методом МЛЭ и имела три входных затвора с отношением ширин затвора 1 : 2 : 4 (3, 6, 12 мкм) [95]. Все измерения были сделаны при комнатной температуре. На рисунке 7.10, б показана схема конфигурации многовходового MOBILE. При измерениях взвешенной суммы операций постоянное напряжение V_C прикладывалось к нижнему РТТ для контроля логического порога (L_{th}). Импульсное напряжение смещение V_{bias} и напряжения входных затворов V_{G1} , V_{G2} и V_{G3} задавались $V_{bias} = 2$ В, $V_{G1} = 1$ В, $V_{G2} = 1$ В и $V_{G3} = 1$ В, соответственно. Полное изменение пикового

тока верхнего РТТ пропорционально взвешенной сумме входных сигналов. Поэтому выходной сигнал O такой структуры MOBILE можно записать как

$$O = u(y), \quad y = S - L_{th},$$

где $S = \sum w_i x_i$ – взвешенная сумма; $x_i = (0,1)$ и $w_i = (1,2,4)$ – входной сигнал и вес i -го затвора соответственно.

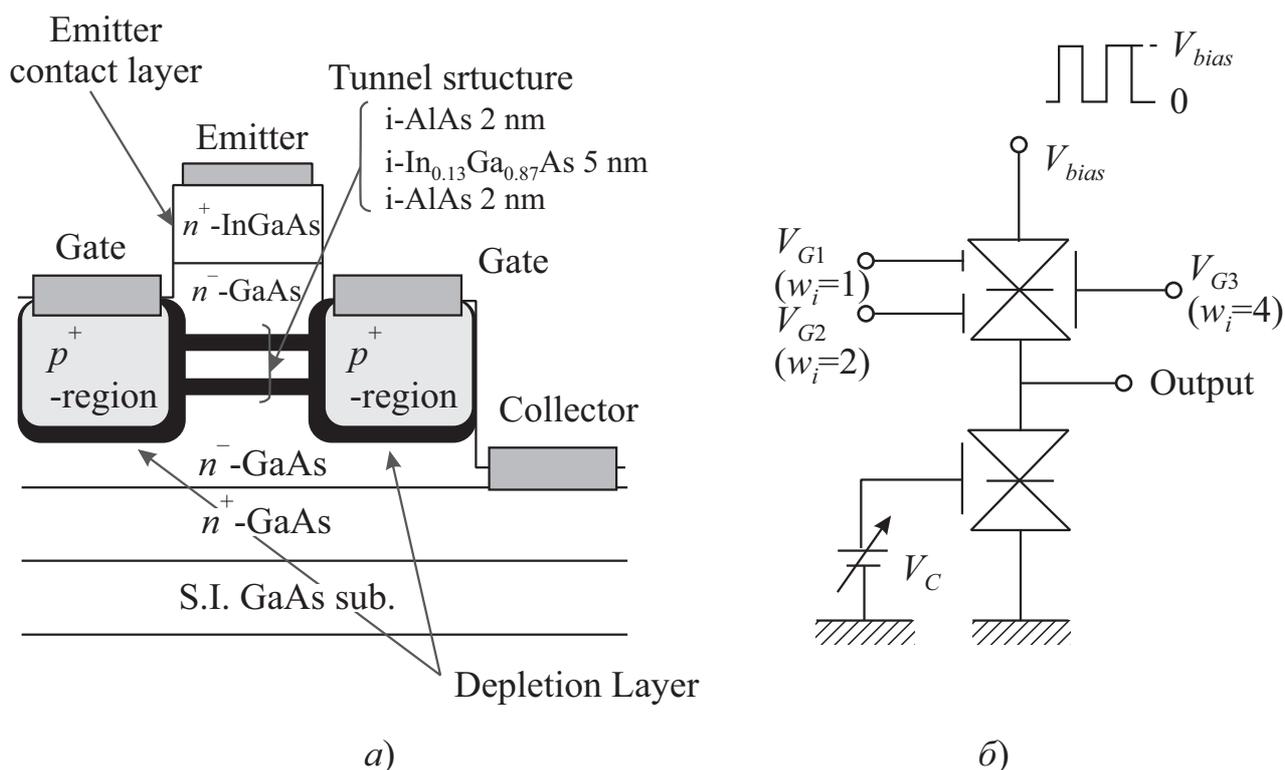


Рис. 7.10 – а) схематическая структура РТТ, имеющего затвор в виде $p^+ - n$ -перехода; б) конфигурация схемы трехвходового устройства MOBILE

Шаг функции $u(y)$ равен 1 при $y > 0$ и 0, при $y < 0$. L_{th} может изменяться от $-0,5$ до $7,5$ за счет изменения V_C . Это указывает на то, что все входные модели ($2^3 = 8$) можно различить с помощью многовходового устройства MOBILE.

Пример операции для $L_{th} = 2,5$ показан на рисунке 7.11. Благодаря большому весу в затворе 3, выход для $(V_{G1}, V_{G2}$ и $V_{G3}) = (0,0,1)$ «выше», в противоположность «низкому» выходу $(1, 0, 0)$ или $(0, 1, 0)$. Эти результаты показывают, что устройство MOBILE может выполнять пороговую логическую функцию для взвешенной суммы входных сигналов. Десять или больше транзисторов

необходимы для этой функции, когда мы используем обычную схему MOS. Эта функция имеет широкий спектр применения в новых вычислительных архитектурах, таких как клеточные нейронные сети и многослойные персептроны (нейронные сети прямого распространения сигнала, от слоя к слою) [95].

По описанной выше технологии были успешно изготовлены устройства MOBILE с многоходовыми логическими затворами. Устройство MOBILE, имеющее трехходовый затвор с отношением ширины 1 : 2 : 4, может отличить все 8 (2^3) входных моделей, соответствующих каждой взвешенной сумме, в зависимости от порогового значения, выбранного на управляющем затворе. Это показывает, что пороговая логическая операция для взвешенной суммы входных сигналов может быть реализована устройством MOBILE.

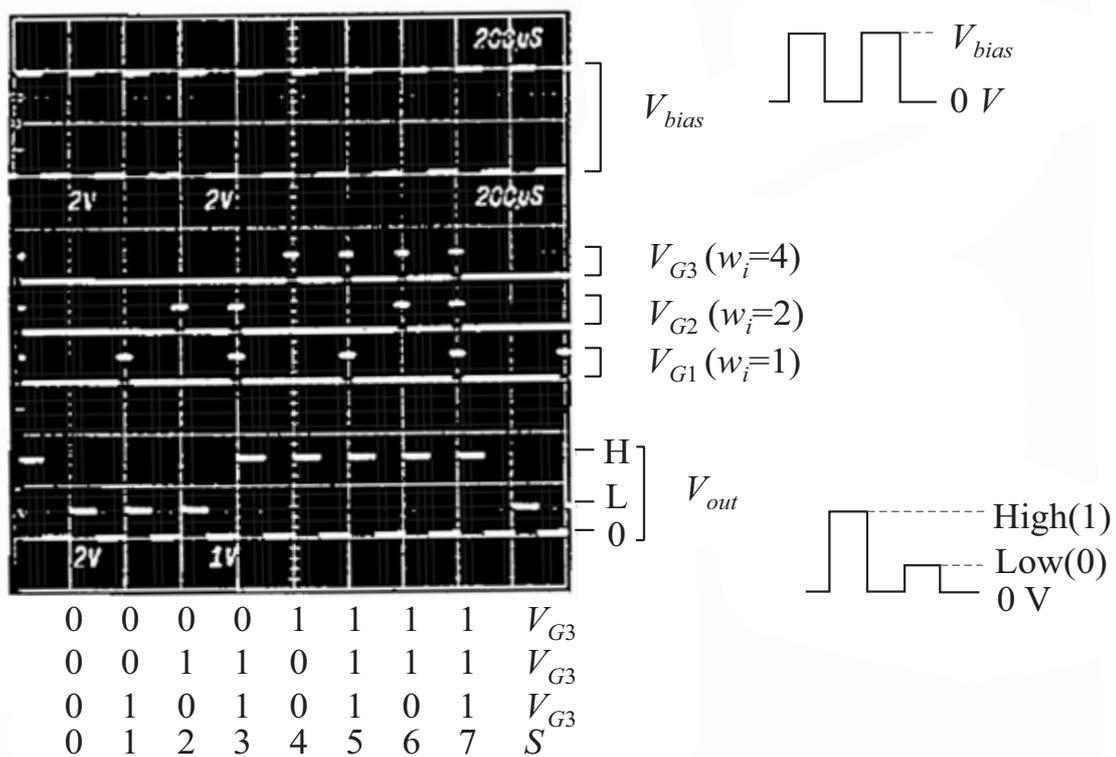


Рис. 7.11 – Пример пороговой логической операции для взвешенной суммы входных сигналов, когда $L_{th} = 2,5$

На рисунке 7.12 показаны нагрузочные кривые и соответствующие диаграммы потенциальной энергии для такой схемы.

Количество устойчивых точек равно единице (рис. 7.12, *a*), когда напряжение смещения меньше, чем удвоенное пиковое напряжение ($2V_p$). Эта устойчивая точка расщепляется на две ветви, S_1 и S_2 (рис. 7.12, *в*), когда напряжение смещения увеличивается за пределы $2V_p$.

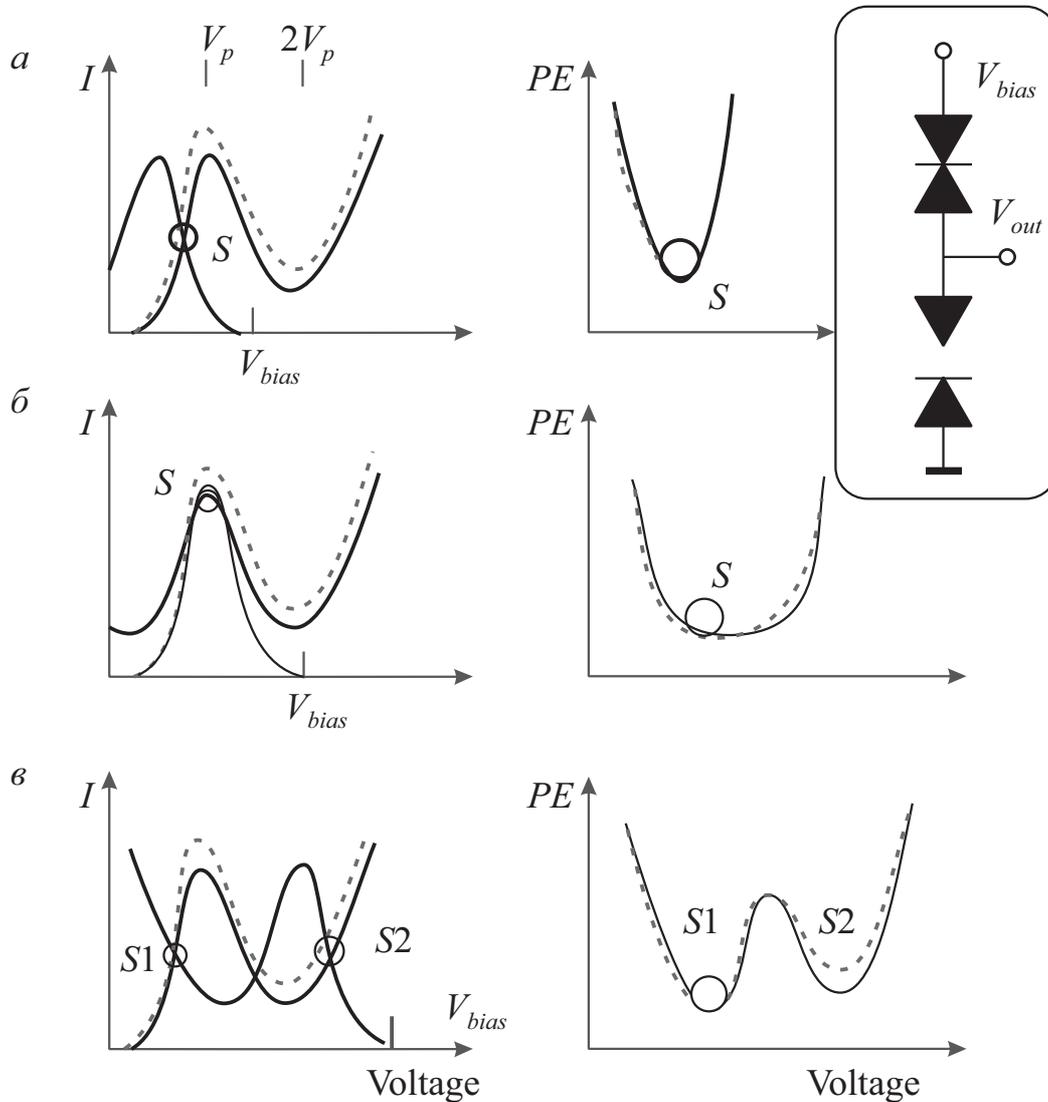
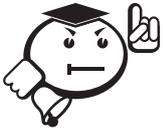


Рис. 7.12 – Нагрузочные кривые и соответствующие диаграммы потенциальной энергии для базового элемента, обеспечивающего переход из моностабильного в бистабильное состояние

Небольшое различие пикового тока у двух приборов определяет состояние схемы после перехода. Например, больший пиковый ток в задающем приборе (нижний резонансно-туннельный диод или транзистор) дает стабильную точку S_1 (штриховая линия). Различие пиковых токов может быть слишком ма-

лым для переключения, так как в точке перехода (рис. 7.12, б) система чувствительна к разнице. При импульсном изменении напряжения смещения V_{bias} схема работает как логический вентиль.



В заключение следует отметить, что наиболее важными особенностями рассмотренных наноэлектронных приборов на туннельных эффектах являются их расширенные, по сравнению с традиционными приборами, функциональные возможности, связанные с уникальностью вольт-амперных характеристик и высоким быстродействием. Благодаря этим выдающимся параметрам эти приборы играют важную роль в разработке сверхбыстродействующих интегральных микросхем со сверхвысокой степенью интеграции [66].

7.3 Приборы на основе сверхрешеток

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. Рассмотрим некоторые применения сверхрешеток в оптоэлектронике. К ним относятся, в частности, ИК-фотоприемники и лазеры на двойной гетероструктуре (ДГС) [77].

7.3.1 Инфракрасные фотоприемники

Фотоприемники предназначены для регистрации и измерения оптических сигналов и для получения изображений во всех областях спектра, в том числе в ИК- и УФ-диапазонах. Некоторые типы фотоприемников были созданы еще в начале XX в., но наиболее быстрыми темпами развивается разработка и выпуск фотоприемников в последние десятилетия. Объем рынка приемников излучения всех классов достигает \$1 млрд/год. Потребителями являются отрасли научного, гражданского и оборонного приборостроения. Наибольшие средства вкладываются в разработки ИК-фотоприемников, используемых в аппаратуре и системах наблюдения, ночного видения, самонаведения, тепловидения, а также в волоконно-оптических линиях связи. Для ИК-фотоприемников используются как легированные сверхрешетки, так и многослойные гетероструктуры. Рабочий диапазон легированных сверхрешеток на основе GaAs 0,8–1,4 мкм. Слои легируются так, что число доноров в *n*-слое равно числу акцепторов в *p*-слое. При этом в равновесном состоянии практически отсутствуют свободные носители, и решетка представляет собой полностью обедненный, высокоомный материал. Свет с длиной волны, соответствующей условию

$$h\nu = \frac{hc}{\lambda} = \Delta E_{g\text{эф}}$$

генерирует пары электрон – дырка. Речь идет об электронах в нижней мини-зоне зоны проводимости *n*-слоя и дырках в верхней мини-зоне валентной зоны *p*-слоев (см. рис. 7.13).

Так как $\Delta E_{g\text{эф}} < \Delta E_g$, то рабочий диапазон λ лежит за краем поглощения GaAs в ИК-области; λ можно изменять с помощью изменения толщины слоев. Генерируемые светом электроны и дырки сразу оказываются пространственно разделенными, поэтому рекомбинация подавлена и время жизни τ велико. К внешним слоям структуры прикладывается напряжение, которое «вытягивает» на электроды неравновесные носители, регистрируя сигнал. Темновой ток мал, так как практически нет равновесных носителей. Генерируемые электроны

при переходе должны туннелировать сквозь барьер высотой $(\Delta E_g - \Delta E_{g\text{эф}})$, что уменьшает вероятность поглощения, но компенсируется большим значением τ .

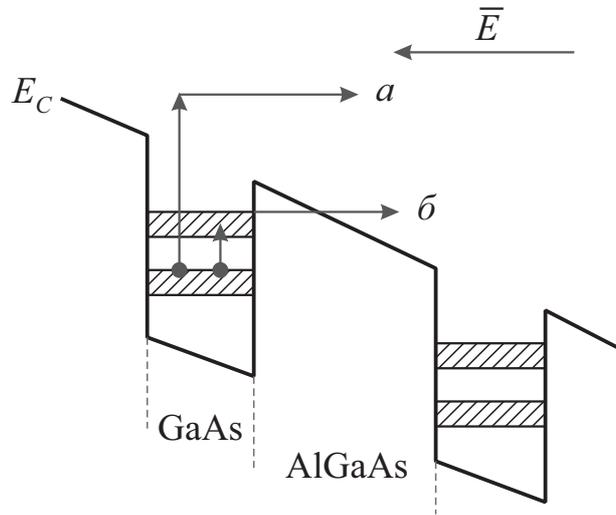


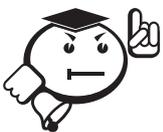
Рис. 7.13 – Энергетическая диаграмма, иллюстрирующая принцип работы ИК-приемника на основе гетероструктур с квантовыми ямами

На рисунке 7.13 приведена энергетическая диаграмма зоны проводимости структуры с квантовыми ямами. Приложенное внешнее электрическое поле с напряженностью E создает наклон зоны проводимости многослойной структуры. Ширина ям составляет $\sim 4\text{--}5$ нм, ширина барьеров – порядка $40\text{--}50$ нм. Штриховкой показаны размерные подзоны в ямах. В равновесии электроны занимают нижние подзоны. Барьеры относительно широкие, туннельных переходов между размерными уровнями ям нет, поэтому ток через структуру в отсутствие освещения (темновой ток) очень мал. Освещение вызывает переходы электронов типа 1 или 2 (см. рис. 7.13), при этом в цепи появляется ток (явление фотопроводимости). Переход 1 – это переход с основного уровня в свободное состояние над барьером в зоне проводимости. Переход 2 – с основного уровня на возбужденный, с последующим туннелированием в состояния над барьером (стрелка b на рис. 7.13) [77].

Изменением толщины слоев можно в широких пределах регулировать спектральные характеристики фотоприемника. Для тепловизионных систем, работающих в диапазоне $8\text{--}20$ мкм, в настоящее время наряду с фотоприемни-

ками на основе CdHgTe и других материалов используются также детекторы ИК-излучения на основе многослойных гетероструктур GaAs – AlGaAs с квантовыми ямами. Тепловидение – это визуализация слабо нагретых объектов по их собственному ИК-излучению. Для тел с температурой, близкой к комнатной, максимум интенсивности излучения приходится на $\lambda = 10$ мкм. Например, в полной темноте человека можно увидеть на расстоянии ~ 30 м в ручной тепловизор. Объекты военной техники просматриваются на расстоянии 2–3 км. Тепловидение может быть широко использовано в авиации, машиностроении, строительстве, микроэлектронике, медицине, геологии. Перспективность гетероструктур GaAs – AlGaAs связана с разработанностью методов формирования тонких слоев, возможностью обеспечить высокий уровень интеграции фотоприемных элементов и элементов обработки фотосигнала. В настоящее время создана технология гибридных фотоматриц. Отдельно изготавливаются фотоматрица из фотоприемных элементов (форматом, например, 128×128 или 320×256) и матрица СБИС на основе Si для считывания и обработки сигналов, регистрируемых фотоэлементами. Затем эти матрицы стыкуются при помощи индиевых микроконтактов. Размеры фоточувствительных элементов составляют ~ 20 – 50 мкм. Такие гибридные фотоматрицы – продукт самых высоких технологий [77].

7.3.2 Сверхрешетки в лазерных структурах



Самым распространенным типом полупроводникового лазера является лазер на двойной гетероструктуре, где активная область представляет собой тонкий слой узкозонного полупроводника между двумя широкозонными полупроводниками. Двойные гетероструктуры предложены в 1963 г. Ж. И. Алферовым [96], выдающийся вклад которого в развитие гетероструктур отмечался в главе 4.

На рисунке 7.14 представлен первый лазер на двойной гетероструктуре (ДГС), работающий в непрерывном режиме, эта работа Алферова была послана в печать в мае 1970 г. [68]. Непрерывный режим лазерной генерации был реализован в лазерах с полосковой геометрией, для формирования которых была использована фотолитография, и установленных на медных теплоотводах, покрытых серебром (рис. 7.14).

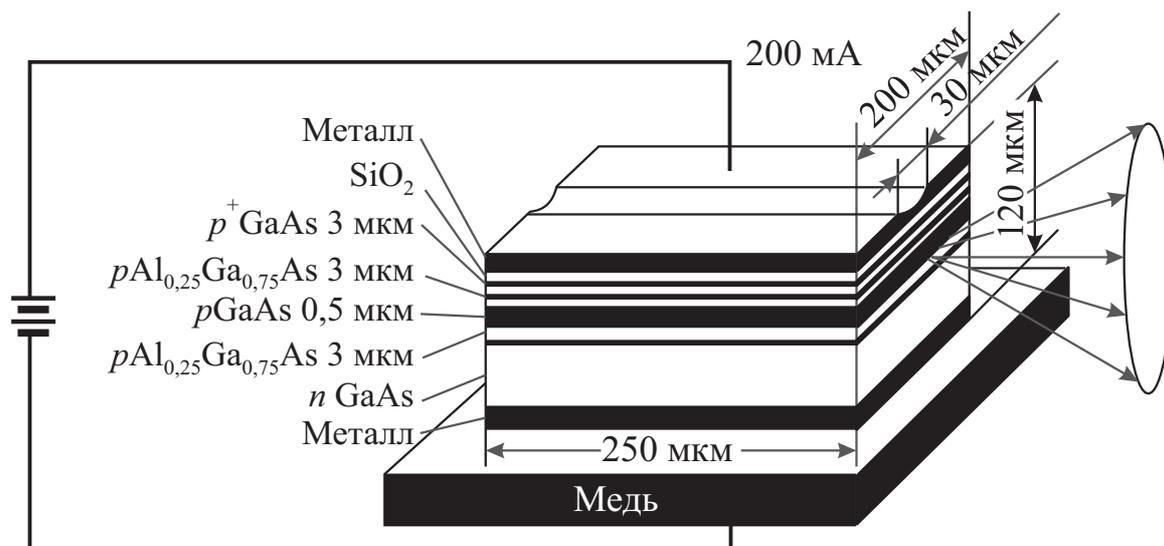


Рис. 7.14 – Схематическое изображение структуры первого инжекционного ДГС лазера, работающего в непрерывном режиме при комнатной температуре [69]

Основное влияние на свойства лазеров оказывает изменение плотности состояний, происходящее под влиянием размерного квантования. Если в массивном полупроводнике в непосредственной близости от края зоны плотность состояний мала, то в квантово-размерной системе она не убывает вблизи края, оставаясь равной $m/\pi\hbar^2$. Благодаря этому факту условия создания инверсной населенности в двумерных системах оказываются более благоприятными, чем в трехмерных. Это привело к конкретным практическим результатам. Создание лазеров с квантово-размерной активной областью позволило получить непрерывную генерацию при комнатной температуре и в дальнейшем снизить пороговый ток инжекционного лазера до рекордно низких значений (50 А/см^2).

Благодаря иной энергетической зависимости плотности состояний меняется не только величина порогового тока, но и его температурная зависимость.

Она становится более слабой, в силу чего непрерывную генерацию удастся получить не только при комнатной температуре, но и при температурах на много десятков градусов выше.

Другой важной особенностью лазеров на квантовых ямах является возможность их частотной перестройки. Минимальная энергия излучаемых световых квантов равна $\hbar\omega = E_g + E_1^e + E_1^h$. Она меняется при изменении ширины квантовой ямы a , то есть изменяя a можно осуществлять перестройку частоты генерации, сдвигая ее в коротковолновую сторону по сравнению с лазерами с широкой (классической) активной областью.

В квантовых точках энергетический спектр меняется еще более радикально, чем в квантовых ямах. Плотность состояний имеет δ -образный вид, и в результате отсутствуют состояния, которые не принимают участия в усилении оптического излучения, но содержат электроны. Это уменьшает потери энергии и как следствие уменьшает пороговый ток. Лазеры могут содержать одну или (для увеличения оптического усиления) несколько плоскостей, заполненных квантовыми точками. Согласно теоретическим оценкам диодные лазеры с активной средой из квантовых точек должны обладать значительно лучшими свойствами по сравнению с лазерами на квантовых ямах: существенно большим коэффициентом усиления, меньшей пороговой плотностью тока, полной невосприимчивостью к температуре решетки, лучшими динамическими характеристиками и большими возможностями контроля за энергией кванта излучения [77]. Показано, что в лазерах на основе вертикально связанных квантовых точек пороговый ток при комнатной температуре может быть снижен до 15 A/cm^2 .

Квантовые каскадные лазеры на сверхрешетках – один из наиболее ярких примеров достижений зонной инженерии, они работают на переходах между подзонами одной и той же зоны проводимости.

Излучения различных каскадных лазеров лежат в диапазоне $\lambda = 2,5\text{--}13 \text{ мкм}$, т. е. перекрывают окна прозрачности атмосферы. Лазеры работают при температурах от криогенных до комнатных, как в импульсном, так и в

непрерывном режиме. На рисунке 7.15 представлена энергетическая диаграмма зоны проводимости, поясняющая принцип работы квантового каскадного лазера (один из вариантов каскадных лазеров).

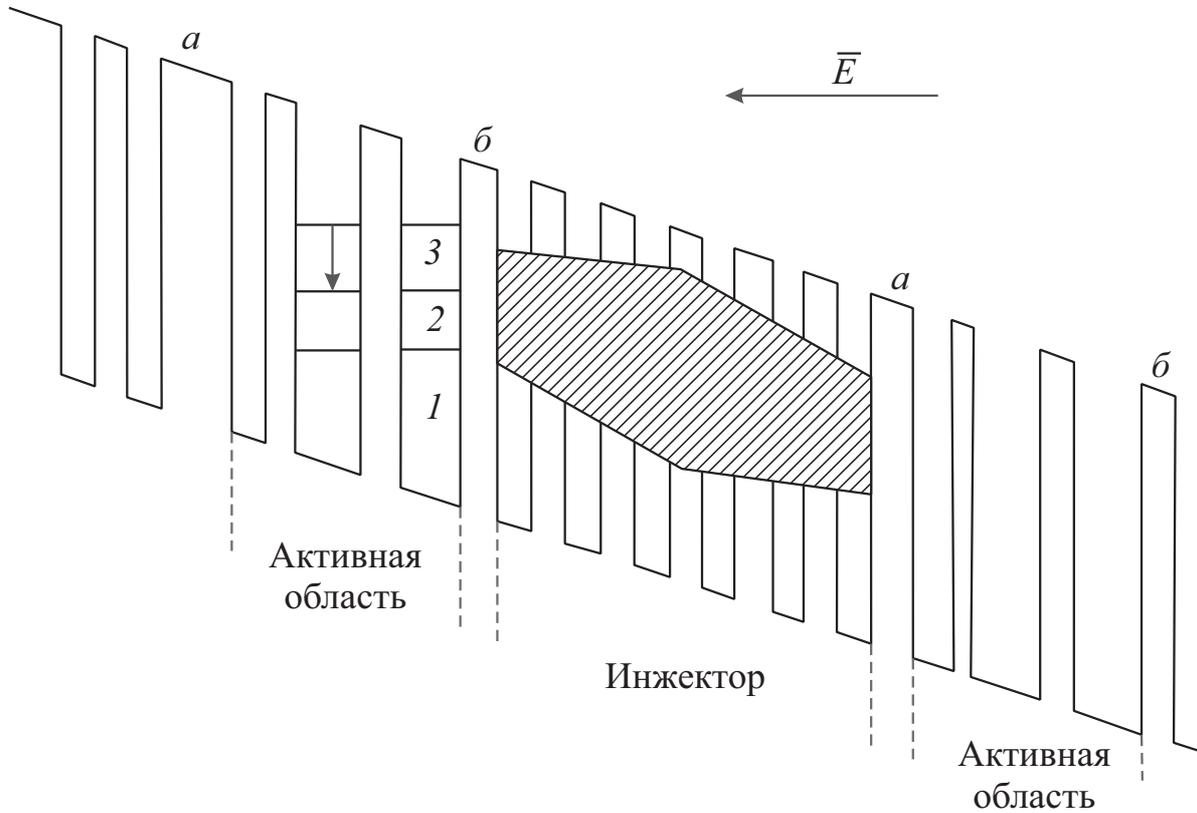


Рис. 7.15 – Схема, поясняющая работу квантового каскадного лазера

В рассматриваемом примере активные области состоят из трех туннельно связанных квантовых ям (InGaAs), разделенных тонкими AlInAs-барьерами. Инжекторы представляют собой сверхрешетки из тех же материалов. Активные области и инжекторы чередуются, образуя каскад из нескольких десятков ступеней (обычно в каскад входит 25–30 ступеней). В квантовых ямах активных областей уровень 2 (квантовое число $n = 2$) и уровень 3 ($n = 3$) – первый и второй возбужденные уровни, уровень 1 ($n = 1$) – основной. Уровни уширены вследствие наличия электрического поля E (величина E достигает 10^5 В/см). Это поле создает также наклон энергетического профиля зоны проводимости (ее нижняя граница представлена на рисунке 7.14).

Лазерное излучение испускается в активных областях при переходах $3 \rightarrow 2$ (вертикальная стрелка на рисунке 7.15), для чего необходима инверсная населенность уровня 3. Инверсия населенности соответствует большему числу электронов на уровне 3, чем на уровне 2. Для этого толщины слоев выбирают так, чтобы: 1) переход $2 \rightarrow 1$ был быстрым и безызлучательным, то есть время жизни электронов на уровне 2 было малым и он быстро опустошался; 2) с уровня 3 наиболее вероятным был излучательный переход на уровень 2 и время жизни электрона на уровне 3 было много больше, чем на уровне 2. При этих условиях создается инверсная населенность на уровне 3 и возможна генерация.

Для непрерывной работы лазера необходимо постоянно инжектировать электроны на уровни 3 и отводить их с уровнями 1. Инжекция и отвод осуществляются через инжекторные области посредством резонансного туннелирования через входной и выходной барьеры (см. рис. 7.15, а, б). Инжекторные области – это сверхрешетки. В сверхрешетках размерные уровни отдельных ям расширяются в мини-зоны.

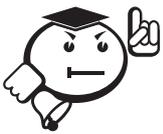
В электрическом поле возможен наклон мини-зон или даже их исчезновение из-за расхождения уровней в соседних ямах. Чтобы это не происходило, в инжекторе ширина ям последовательно уменьшается и получаются горизонтальные мини-зоны (см. рис. 7.12). Толщина слоев в сверхрешетке подбирается так, чтобы на выходе из активной области электроны с уровня 1 или 2 туннелировали в мини-зону, а против уровня 3 была запрещенная мини-зона и утечка электронов с уровня 3 становилась невозможной. Величина электрического поля выбирается таким образом, чтобы дно мини-зоны совпадало с уровнем 3 следующей активной области. Кроме того, мини-зоны сужаются к барьерам инжекции; электроны идут к очередной активной области как по воронке, что условно отражено формой заштрихованной части рисунка 4.15. Электроны, испустившие излучение в активной области, туннелируют в сверхрешетку через выходной барьер, пролетают по сверхрешетке посредством резонансных туннельных переходов, передавая избыточную энергию ее атомам, переходят на уровень у дна мини-зоны и туннелируют на уровень 3 следующей активной

зоны. Выбором ширины ям и барьеров в активной области обеспечивается наибольшая вероятность нахождения электронов на уровне 3 в первой и второй ямах, а на уровнях 2 и 1 – во второй и третьей ямах, что повышает эффективность инжекции и отвода носителей заряда.

Приведем пример ширины ям и барьеров для одного из действующих лазеров (его энергетическая диаграмма дана на рисунке 4.15). Для активных областей крайние барьеры – входной и выходной. При этом входной барьер имеет ширину 5 нм, выходной – 3 нм, а потенциальные ямы – $1,0$; $4,7$; $4,0$ нм, промежуточные барьеры – $5,0$; $1,5$; $2,2$; $3,0$ нм соответственно. В инжекционной области ямы имеют размеры $2,3$; $2,2$; $2,0$; $2,0$; $1,9$; $1,9$ нм, а барьеры – $2,3$; $2,2$; $2,0$; $2,3$; $2,28$ нм. Если в каскадном лазере N ступеней, то каждый электрон, пройдя весь каскад, создает N фотонов (по одному на каждой ступени). Приведенный пример иллюстрирует, сколь высоки требования, предъявляемые к методам выращивания приборных многослойных наноструктур. Лазерная структура наращивается послойно методом молекулярно-лучевой эпитаксии на подложке InP. Толщина всех слоев должна выдерживаться с атомной точностью. Разработаны также инфракрасные лазеры на межзонных переходах в сверхрешетках с более сложным энергетическим профилем на основе соединений In, Ga, Al, Sb.

7.3.3 Лавинные фотодиоды

Лавинные фотодиоды представляют собой фоточувствительные приборы с внутренним усилением, позволяющие получать высокую чувствительность. Основным их недостатком является то, что с лавинным умножением связан дополнительный шум, ограничивающий возможность детектирования слабых сигналов.



Уже давно установлено, что для получения низкого уровня шума при большом внутреннем усилении необходимо, чтобы коэффициенты ударной ионизации электронов α и дырок β резко различались между собой. К сожалению, в большинстве соединений A_3B_5

$\alpha/\beta = 1$, что приводит к возрастанию шума при умножении [78]. Поэтому большое практическое значение имеют методы, позволяющие в указанных материалах увеличить отношение α/β . Один из способов сделать это – создание структур типа сверхрешеток, использующих явление ударной ионизации на разрыве энергетических зон [78].

.....

Рассмотрим зонную диаграмму сверхрешеточной структуры в сильном электрическом поле обратносмещенного p - i - n -диода (рис. 7.16). Пусть мы имеем горячий электрон, ускоряющийся в барьерном слое широкозонного полупроводника. Влетая в узкозонный слой, он резко увеличивает энергию на величину разрыва зоны проводимости ΔE_C .

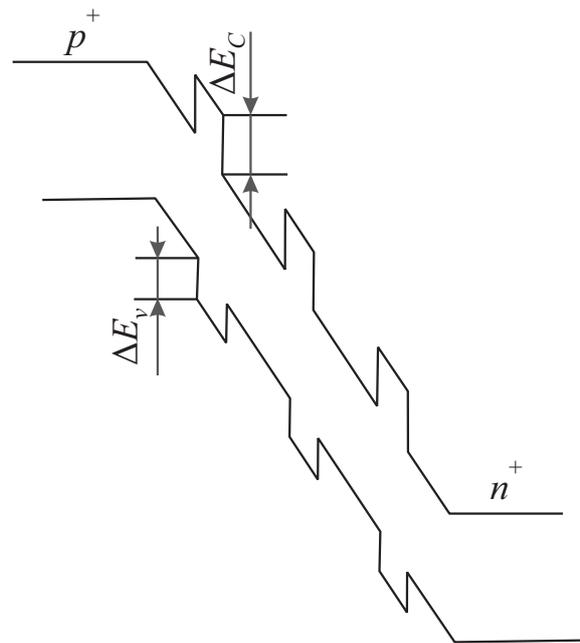


Рис. 7.16 – Схема лавинного фотодиода с системой квантовых ям

Это эквивалентно тому, что он «видит» энергию ионизации уменьшенной на ΔE_C по сравнению с пороговой энергией в массивном узкозонном полупроводнике. Поскольку коэффициент ударной ионизации α с уменьшением пороговой энергии экспоненциально растет, следует ожидать резкого увеличения эффективного значения α . В следующем барьерном слое пороговая энергия

увеличивается на ΔE_C , уменьшая тем самым значение α в этом слое. Но поскольку $\alpha_1 \ll \alpha_2$ (индексы 1 и 2 относятся соответственно к широкозонному и узкозонному материалу), то экспоненциальный рост α_2 приводит к тому, что и среднее значение:

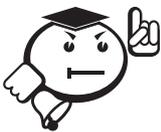
$$\bar{\alpha} = \frac{\alpha_1 b + \alpha_2}{a + b},$$

значительно увеличивается [78].

Если разрывы в валентной зоне ΔE_V значительно меньше разрывов в зоне проводимости (как это, в частности, имеет место в системе GaAs – AlGaAs), то подобный эффект для дырочного коэффициента β оказывается значительно меньше. Окончательным результатом будет сильное увеличение отношения α/β .

7.3.4 Оптические модуляторы

Структуры с квантовыми ямами могут использоваться не только для генерации, но и для модуляции светового излучения из области межзонного оптического поглощения.



.....

Физической основой эффекта модуляции являются два экспериментальных факта: а) в структурах с квантовыми ямами существуют резкие линии экситонного поглощения, имеющие значительно большую интенсивность, чем в однородных полупроводниках, и наблюдаемые не только при низких, но и при комнатных температурах; б) спектральное положение линий заметно сдвигается под воздействием сильного электрического поля, перпендикулярного слоям гетероструктуры.

.....

В однородных полупроводниках электрическое поле разрушает экситоны, растаскивая электрон и дырку в разные стороны. Наличие же квантовой ямы позволяет без вреда для экситона приложить в поперечном направлении достаточно сильное поле, поскольку стенки ямы не дают электрону и дырке возмож-

ности разойтись. Поэтому экситонная линия поглощения в поле не исчезает, но ее положение достаточно сильно сдвигается. Это обусловлено двумя факторами. Во-первых, поле искажает форму самой квантовой ямы, превращая ее из прямоугольной в трапециевидную (рис. 7.17). При этом меняется энергия квантовых уровней как в зоне проводимости \tilde{E}_1^e , так и в валентной зоне \tilde{E}_1^h , а следовательно, и эффективная ширина запрещенной зоны $E_g^{\text{эф}} = E_g + E_1^e + E_1^h$.

Во-вторых, энергия связи экситона E_{ex} , хотя и не обращается в нуль, как только что было отмечено, но, тем не менее, зависит от приложенного поля. В результате спектральное положение экситонной линии $\hbar\omega_{ex} = E_g^{\text{эф}} - E_{ex}$ оказывается сильно зависящим от напряженности электрического поля F .

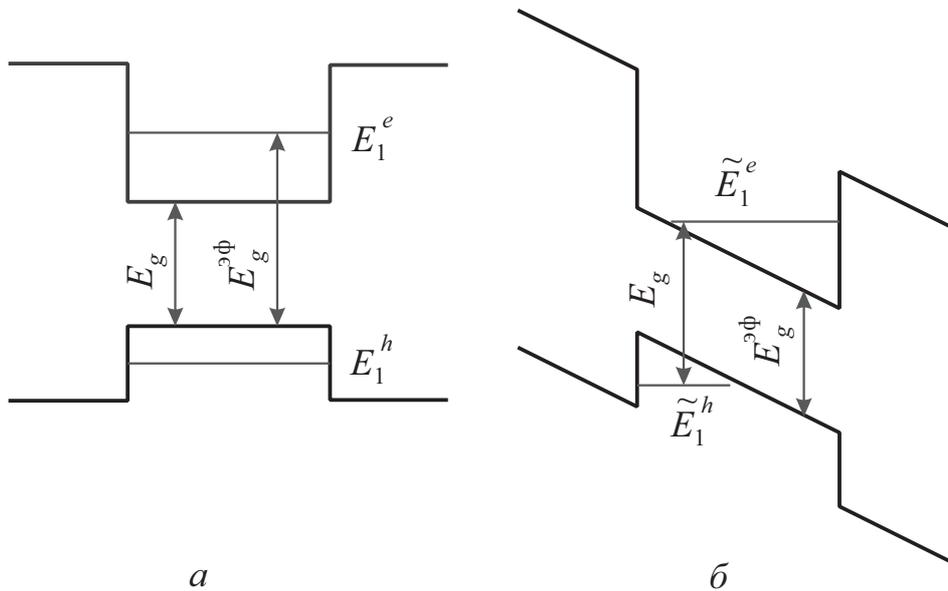


Рис. 7.17 – Энергетический спектр квантовой ямы в отсутствие (а) и при наличии (б) приложенного электрического поля

Поскольку экситонная линия в спектре поглощения весьма резка, коэффициент поглощения α для света с частотой вблизи этой линии очень сильно меняется при изменении положения линии. Это делает весьма эффективной модуляцию света с помощью приложенного электрического поля. Подобные модуляторы находят применение в системах оптической передачи информации.

7.4 Транзисторы для СВЧ-электроники

В России (как и ранее в СССР) основным рынком сбыта изделий СВЧ-электроники является рынок радиоэлектронной аппаратуры военного назначения. Именно он формирует требуемые совокупности технических и эксплуатационных характеристик для конкретных изделий СВЧ-электроники, а также является главным двигателем научно-технического прогресса в этой области. Изделия СВЧ-электроники, предназначенные, прежде всего, для генерации, усиления и преобразования энергии СВЧ-колебаний в приемных и передающих трактах радиоэлектронной аппаратуры различных частотных диапазонов, в значительной степени определяют облик, архитектуру построения и основные тактико-технические характеристики современных и перспективных систем и комплексов радиолокации, радионавигации, всех видов связи, радиоэлектронной разведки и радиоэлектронного противодействия, систем управления вооружением, систем высокоточного оружия и др.

Но военное применение СВЧ-приборов – это не более 10% всего их рынка. Целые направления современной мировой индустрии строятся на достижениях твердотельной СВЧ-электроники, например сотовая и спутниковая связь, телевидение, радиолокационные станции (РЛС) для управления воздушным движением, спутниковые трансиверы и прочие сферы применения.

По диапазонам частот основной потребитель СВЧ-компонентов (пассивных, транзисторов и СВЧ монолитных ИС (МИС)) – системы сотовой связи для диапазонов 0,8–1,9 ГГц и беспроводные локальные и персональные сети передачи информации 2,4 и 5–6 ГГц. Для стандарта передачи информации региональных и беспроводных сетей WiMAX – диапазон от 3 до 11 ГГц. Для систем сотового телевидения – LMDS/MVDS (Local Multipoint Distribution Service/Multipoint Video Distribution Systems) – американская/европейская системы, работающие в диапазонах 27–31 и 40,5–42,5 ГГц соответственно. Лицензируемые диапазоны (США) 71–76 и 81–86 ГГц для создания сверхбыстрых систем передачи данных с радиусом действия 1–2 км.

Системы спутниковой связи в Ku- (13,75–14,5 ГГц) и Ka-диапазоне (29,5–30 ГГц), цифровое радиовещание (DAB) в диапазонах 6–42 ГГц. Автомобильные радары различного назначения (радары дальнего действия для систем автоматического контроля курса, системы предупреждения/предотвращения столкновений, а также радары ближнего действия для бокового и заднего обзора). Рабочие диапазоны подобных систем различны – 19, 24, 60 и 77 ГГц.

В области диапазонов частот от единиц до сотен гигагерц и мощностей от милливатт до сотен ватт используются разнообразные полупроводниковые технологии и материалы. Основные типы транзисторов в этой области – это традиционные полевые и биполярные приборы, полевые транзисторы с затвором Шоттки (MESFET), полевые транзисторы с гетеропереходом (HEMT), а также биполярные транзисторы с гетеропереходом (HBT).

С точки зрения полупроводниковых материалов на рынке СВЧ-электроники безраздельно доминируют GaAs-приборы: в самой емкой рыночной нише – усилители мощности сотовых телефонов – им принадлежит 80% всего рынка. Причем когда сегодня говорят об GaAs, имеют в виду прежде всего гетероструктуры – AlGaAs/GaAs, InGaAs/GaAs и т. д. – и созданные на их базе транзисторы, биполярные (HBT) и полевые с барьером Шоттки (HEMT). Арсенид галлия традиционно рассматривался как основной материал для СВЧ-приборов. И тому есть очевидные предпосылки: подвижность электронов – $8\,500\text{ см}^2/\text{Вс}$, ширина запрещенной зоны и напряженность поля пробоя – 1,42 эВ и $4\cdot 10^5\text{ В/см}$ соответственно – делают GaAs предпочтительнее кремния в области СВЧ. Это был первый освоенный промышленностью материал из группы полупроводников A^3B^5 , с которыми и сегодня связаны многие перспективы СВЧ-электроники. Не сошел с арены и «старый добрый» кремний. В нише сотовых телефонов приборы на основе этого материала занимают порядка 20% рынка. Обладая приемлемыми характеристиками в области диапазонов частот до 3 ГГц (с тенденцией расширения до 10 ГГц), кремниевые транзисторы (биполярные, МОП, Би-КМОП) развиваются весьма интенсивно. Порядка 10 лет назад начали интенсивно прогрессировать и сегодня весьма

распространены SiGe-технологии, прежде всего SiGe HBT и Би-КМОП-транзисторов, которые можно рассматривать как развитие Si-технологий.

На фоне этих доминирующих сегодня технологий развиваются направления, которые, возможно, будут играть важнейшую роль в СВЧ-электронике завтра. Прежде всего, речь идет о фосфиде индия и твердых растворах на его основе (AlInP). В последние годы наблюдается необычайно бурное развитие широкозонных материалов – группы III-нитридов, прежде всего GaN, а также приборов на основе карбида кремния SiC. Появляется все больше сообщений о транзисторах на основе экзотических пока материалов InAs, AlSb и InSb, сулящих сверхбыстродействие при минимальном энергопотреблении. Исследуются возможности создания полупроводниковых приборов на основе алмазов. Все эти технологии переживают различные фазы становления, их время еще впереди.

7.4.1 Транзисторы с высокой подвижностью (HEMT)

Одним из самых распространенных видов полупроводниковых приборов являются полевые транзисторы, в том числе МДП-структуры. Если к инверсионному слою создать два контакта, называемых истоком и стоком и расположенных на расстоянии l друг от друга, то при напряжении между ними, равном V , в канале будет течь ток, определяемый, кроме размеров канала, подвижностью μ и концентрацией n_S носителей заряда в канале:

$$I = e\mu n_S V (b/l), \quad (7.1)$$

где b – ширина канала. Изменяя с помощью затвора концентрацию в канале n_S , мы можем осуществлять управление током исток – сток. Двумерная плотность электронов в слое n_S (плотность состояний электронов на единицу площади двумерного электронного газа) пропорциональна напряжению на затворе [78]:

$$n_S = \frac{k_d}{4\pi e d} (V_3 - V_0), \quad (7.2)$$

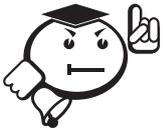
где k_d – диэлектрическая проницаемость диэлектрика; V_0 – пороговое напряжение, соответствующее открытию инверсионного канала, т. е. появлению

в нем электронов; d – толщина диэлектрика, отделяющая металлический затвор от объема полупроводника. Тогда, с учетом (7.1), важнейшая характеристика транзистора – крутизна G – в нашем случае определяется выражением:

$$G = \frac{dI}{dV_3} = \frac{\varepsilon_d \mu b V}{4\pi d l}. \quad (7.3)$$

Видно, что крутизна полевого транзистора пропорциональна подвижности носителей. Поэтому сразу возникает желание использовать в полевых транзисторах структуры с модулированным легированием, обладающие гигантской подвижностью носителей. Для этого требуется решить задачу об управлении концентрацией двумерных носителей в этих структурах. Она решается путем создания контакта Шоттки на широкозонном легированном слое. Если толщина этого слоя достаточно мала, то области обеднения вблизи гетероперехода и контакта Шоттки перекрываются и все доноры широкозонного слоя оказываются ионизованными. В результате он играет роль диэлектрика, гетероструктура становится аналогом МДП-структуры и концентрация двумерных электронов линейно зависит от напряжения на затворе (контакте Шоттки) [78]. Созданные таким образом полевые транзисторы часто обозначают английской аббревиатурой НЕМТ (*High Electron Mobility Transistor* – транзистор с высокой подвижностью электронов). На самом деле полностью использовать преимущества высокой подвижности двумерного газа и получить транзисторы с крутизной, во много раз большей, чем у обычных МДП-транзисторов, не удастся. Причина в том, что в реальных приборах для получения высокого быстродействия и высокой плотности интеграции в БИС расстояние между истоком и стоком l делается весьма малым (~ 1 мкм). При этом напряженность поля в канале настолько велика, что дрейфовая скорость электронов уже не пропорциональна полю (с коэффициентом, равным подвижности), а выходит на насыщение. В результате крутизна G не описывается формулой (7.3), а слабее зависит от подвижности носителей. Тем не менее, использование структур с модулированным легированием приводит к некоторому повышению крутиз-

ны полевых транзисторов. При температуре 77 К она может быть доведена до значений порядка 500 мСм/мм.



Полевые GaAs-транзисторы с гетеропереходом и управляющим затвором Шоттки (ГПТШ) сейчас уже вышли на уровень массового производства. Хотя они еще и не потеснили Si- и GaAs-биполярные транзисторы с гетеропереходом (*HBT – hetero bipolar transistor*), однако такая тенденция прослеживается.

Типичная структура GaAs ГПТШ приведена на рисунке 7.18. Его конструкция отличается от обычных полевых транзисторов с управляющим барьером Шоттки тем, что между барьерным (донорным) и нелегированным канальным слоями (в данном случае между слоями GaAs и AlGaAs) имеется гетеропереход. Поскольку ширина запрещенной зоны материала канального слоя больше, чем барьерного, в канале у границы слоев формируется потенциальная яма (квантовый колодец) – тонкий слой, в котором накапливаются свободные носители, образуя двумерный электронный газ (2DEG). Двумерный электронный газ (2D-газ) представляет собой систему электронов, энергетические состояния которых соответствуют свободному движению только вдоль определенной плоскости. Важным свойством двумерного электронного газа является то, что возможно регулирование в широких пределах плотности электронов под воздействием поперечного электрического поля. Электроны в ДЭГ имеют повышенную эффективную концентрацию и подвижность: так как канальный слой не легирован, в нем рассеяние на примесных центрах и дислокациях решетки минимальны, а подвижность носителей, соответственно, высока. Поэтому данный класс приборов называют транзисторами с высокой подвижностью электронов (HEMT). Задача буферного слоя – обеспечить структурный переход от полуизолирующей подложки к совершенной структуре канального слоя.

На рисунке 7.18, б приведена зонная диаграмма гетероперехода между арсенидом галлия и арсенидом галлия алюминия $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$. Величина x характеризует содержание алюминия, и с ростом x увеличивается ширина запрещенной зоны данного твердого раствора. Для типичного значения $x = 0,3$ ширина запрещенной зоны твердого раствора $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ равна 1,8 эВ.

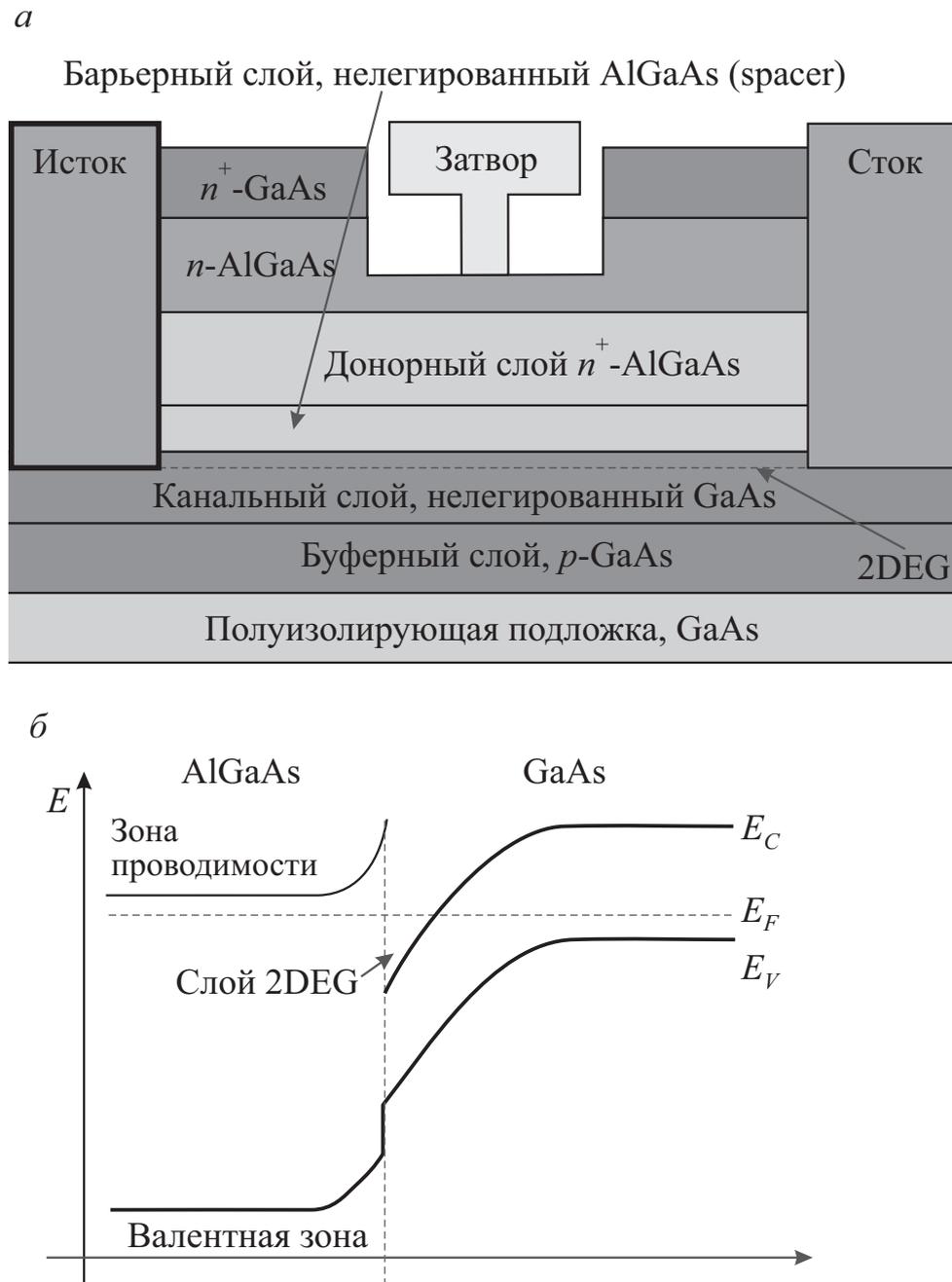


Рис. 7.18 – Типичная структура GaAs ГПТШ (*a*); его зонная диаграмма (*б*)

Конструкция НЕМТ многослойна. Между металлическим затвором и легированным слоем на основе AlGaAs формируется управляющий переход металл – полупроводник. Обедненная область этого перехода располагается в слоях арсенида галлия алюминия.

Постоянные кристаллических решеток AlGaAs и GaAs достаточно близки. НЕМТ-транзисторы, гетеропереход которых образован материалами с существенно разными константами решетки (AlGaAs/InGaAs, InGaAs/InAlAs, InGaP/InGaAs), получили название псевдоморфных транзисторов (pHEMT). Приборы этого типа за счет увеличения разрыва между границами зон проводимости и значениями подвижности электронов обладают более высоким пробивным напряжением (выше 12 В, рабочее напряжение – 5–6 В) и высокими рабочими частотами, КПД достигает 60%. Сейчас производятся транзисторы pHEMT с минимальным размером элементов 150 нм. Базовая структура транзистора формируется на полуизолирующей GaAs-подложке, на ней создают буферный слой в виде сверхрешетки AlAs/GaAs, InGaAs-канальный слой и AlGaAs-барьерный слой *n*-типа.

Сверху и снизу к каналному слою примыкает так называемый спейсер (*spacer*) – тонкий слой нелегированного AlGaAs. Концентрация электронов проводимости в канале достигает $3,2 \cdot 10^{12} \text{ см}^{-2}$, а их подвижность – $6500 \text{ см}^2/\text{Вс}$. В результате максимальный ток канала транзистора может составлять 680 мА/мм (при напряжении канала 1,5 В), напряжение пробоя –13 В, граничная частота $f_T = 52 \text{ ГГц}$, максимальная частота усиления мощности $f_{\text{max}} = 150 \text{ ГГц}$. Оптимальное рабочее напряжение транзистора – 6 В. Выходная мощность может достигать 815 мВт/мм, а КПД – 40%.

По этой технологии производятся усилители с мощностью 2 Вт, рассчитанные на диапазоны частот 32–38 ГГц и 45 ГГц соответственно. Такой трехкаскадный усилитель размещается на кристалле размером 2,8×2,3 мм, общая ширина затвора транзисторов в выходном каскаде – 4,16 мм. При пиковой выходной мощности 2,5 Вт в диапазоне 32–38 ГГц КПД прибора составил 25%, номинальное усиление – 18 дБ, рабочее напряжение – 6 В, напряжение пробоя

сток–затвор – 11 В. Усилитель с аналогичными электрическими параметрами в диапазоне 44–46 ГГц выполнен по балансной трехкаскадной схеме. Общая площадь кристалла 3,4×4,3 мм. КПД усилителя при входной мощности 20 дБм (мощность сигнала по отношению к 1 мВт) составляет 14%.

С целью увеличения рабочих частот рНЕМТ-транзисторов на основе InAlAs/InGaAs-структур разработчики стремятся повысить концентрацию индия в $\text{In}_x\text{Ga}_{1-x}\text{As}$ -канальном слое. Но это ведет к увеличению несоответствия постоянных кристаллических решеток материала канального слоя и GaAs-подложки, что, в конечном счете, ухудшает характеристики прибора. Одно из возможных решений – применение InP-подложек, постоянная решетки которых близка к InAlAs. Для подобных структур получены рекордные значения частоты. Так, их граничная частота может превышать 400 ГГц. В частности, о создании рНЕМТ с InGaAs/InAlAs/InP-структурой на 100-мм пластинах по 0,1-мкм технологии одной из первых объявила компания Northrop Grumman Space Technology (NGST). Эта технология – результат совершенствования уже освоенного процесса изготовления рНЕМТ на 75-мм пластинах, на базе которого производятся серийные МИС. В транзисторах, создаваемых в рамках данного процесса, концентрация индия в InGaAs-канальном слое составляет 60%, подвижность носителей в канале достигает $10\,000\text{ см}^2/\text{Вс}$ при плотности электронов в 2DEG-слое $\sim 3,5 \cdot 10^{12}\text{ см}^{-2}$. При этом крутизна транзистора G_m – примерно 800 мСм/мм, плотность тока – 540 мА/мм, граничная частота f_T – более 190 ГГц. На базе рНЕМТ с такой структурой компанией создан ряд МИС, в частности – малошумящий двухкаскадный балансный усилитель с коэффициентом усиления свыше 17 дБ и уровнем шумов менее 2,4 дБ в диапазоне 27–39 ГГц.

Однако, несмотря на усилия производителей, даже с учетом перехода к 100-мм InP-подложкам, стоимость приборов на основе InP продолжает оставаться высокой. В качестве возможного решения этой проблемы рассматриваются так называемые метаморфные НЕМТ-структуры (МНЕМТ), т. е. приборы со сложным многослойным буфером со ступенчатым изменением concentra-

ции примеси для обеспечения плавного перехода от GaAs-подложки к InGaAs-каналу. Тем самым решается проблема применения более дешевых GaAs-подложек. Разработана технология изготовления InAlAs/InGaAs/GaAs МНЕМТ-структур с концентрацией индия в InGaAs-канале 60%. Им удалось создать транзисторы с крутизной порядка 850 мСм/мм, напряжением пробоя 8 В и током канала до 700 мА/мм (в постоянном режиме). В импульсном режиме эти показатели составили 6 В и 200 мА/мм соответственно. При суммарной ширине затвора 480 мкм усиление на частоте 94 ГГц при напряжении 2,5 В составило 12 дБ, выходная мощность 360 мВт/мм, КПД – свыше 30%. На базе этих МНЕМТ создан ряд однокаскадных усилителей мощности W -диапазона (56–100 ГГц). Их усиление при напряжении 3,3 В составило 6,5 дБ, КПД – 17% при максимальной выходной мощности 226 мВт (коэффициент усиления снижается на 1 дБ).

7.4.2 Транзисторы на основе SiGe-технологии



В последние годы растет популярность СВЧ SiGe-устройств, которые уже сопоставимы по частотным характеристикам с GaAs-приборами. SiGe-транзисторы можно считать результатом развития кремниевой технологии. Первой промышленный технологический процесс создания SiGe-транзисторных структур разработала компания IBM в 1996 г.

С тех пор эту технологию поддержали многие ведущие производители. Основные достоинства Si-Ge-структур – при незначительном (на 10–20%) увеличении стоимости процессов по сравнению со стандартной Si-технологией существенно возрастают рабочие частоты и линейность характеристик. Так, компания IBM в рамках своего Би-КМОП-техпроцесса с технологическими нормами 0,18 мкм производит МИС с граничными частотами 70/45 ГГц для n -МОП/ p -МОП транзисторов и 120 ГГц – для НВТ (Би-КМОП-технология яв-

ляется объединением КМОП- и биполярных транзисторов в единой приборной структуре). По технологии 8НР с 0,13-мкм топологическими нормами производятся биполярные *n-p-n*-транзисторы с граничной частотой 210 ГГц, максимальная частота усиления – 185 ГГц.

Процесс изготовления SiGe-приборов, по сравнению с кремниевыми, требует нескольких дополнительных операций. Так, в биполярных кремниевых *n-p-n*-транзисторах в базе формируется эпитаксиальный SiGe-слой, в результате чего образуется гетеропереход Si/SiGe, позволяющий увеличить уровень легирования базы и тем самым снизить ее сопротивление. Это сокращает время переключения. Снижается и уровень шумов. Уменьшение ширины запрещенной зоны базы позволяет увеличить коэффициент усиления, поскольку из-за разрыва границ валентной зоны в переходе база – эмиттер сокращается инжекция дырок в эмиттер (что приводит к увеличению коэффициента эффективности эмиттера).

В МОП-транзисторах в буферном слое используется SiGe для формирования в канале области двумерного электронного газа (рис. 7.19).

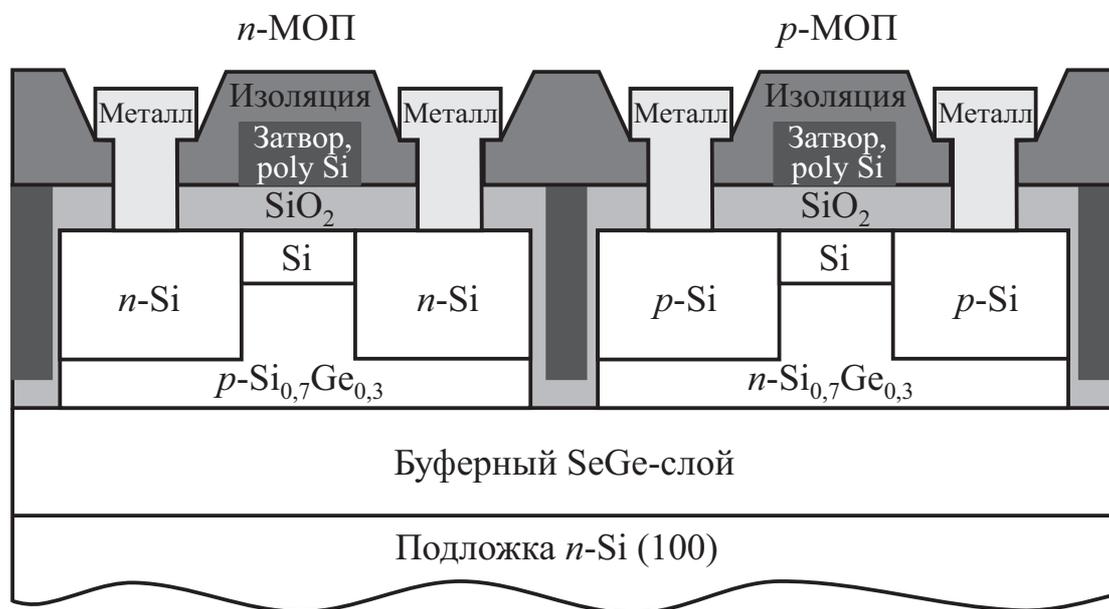


Рис. 7.19 – Структура SiGe КМОП-ячейки ИС

В результате подвижность носителей в канале *n*- и *p*-типа составляет 2 800 и 1 400 см²/Вс соответственно, что обуславливает высокие частотные характеристики: граничные частоты SiGe МОП-транзисторов могут достигать

100 ГГц и даже выше. И это не предел, учитывая тенденцию уменьшения геометрических размеров элементов транзистора.



Однако сдерживающий фактор для SiGe-приборов – малая ширина запрещенной зоны. У германия она равна 0,68 эВ, у кремния – 1,1 эВ, у их композиции зависит от концентрации германия и составляет ~1 эВ.

Это существенно ограничивает (на уровне порядка 2 В) пробивное напряжение и мощностные показатели. Поэтому SiGe-приборы сегодня находят широкое применение в маломощных приложениях – сотовых телефонах, мобильных устройствах и т. п., то есть там, где важно минимизировать энергопотребление и не требуется высокая мощность. Отметим, что благодаря конструктивным особенностям пробивное напряжение перехода эмиттер – коллектор SiGe HBT может составлять 3,5–6 В. Однако, возможно, наиболее существенная особенность SiGe-транзисторов – простота их формирования на одном кристалле с цифровыми кремниевыми схемами. Это важно, прежде всего, для создания разнообразных СБИС смешанной обработки сигнала, включая «системы на кристалле». Неслучайно SiGe-технология – одна из основных при изготовлении СВЧ-трансиверов, в том числе для сотовых телефонов. Этому способствует и современная тенденция применения трансиверов прямого преобразования, т. е. устройств с нулевой промежуточной частотой. Например, характеристики SiGe-усилителя: диапазон частот 0–6 ГГц с усилением 9–15,5 ГГц (в зависимости от частоты), выходная мощность достигает 10 дБм. Усилитель работает от однополярного источника питания с напряжением 3–10 В. Другой пример – СВЧ SiGe-трансивер, работающий в диапазоне частот 4–7 ГГц с шириной полосы до 250 МГц. Прибор обеспечивает подавление несущей до 34 дБс (относительно несущей) и способен работать с уровнем сигнала внешнего тактового генератора от –3 до 6 дБм. Поставляется в миниатюрном пластиковом корпусе QFN габаритами 3×3 мм.

7.4.3 Транзисторы на основе технологии нитрида галлия (GaN) на подложке из карбида кремния (SiC)



.....

Одно из наиболее активно осваиваемых сегодня направлений СВЧ-электроники – мощные приборы на основе полупроводниковых материалов с широкой запрещенной зоной. Широкая запрещенная зона – это большие значения пробивного и рабочего напряжения, высокая рабочая температура перехода. В гетероструктурах она обеспечивает значительный разрыв границ зоны проводимости. Однако сегодня активно развиваются всего два широкозонных полупроводниковых материала – карбид кремния и нитрид галлия.

.....

Нитрид галлия представляется наиболее перспективным широкозонным полупроводниковым материалом, как и вся группа III-нитридов. Совокупность свойств широкозонного материала и возможности формирования на основе GaN гетеропереходов позволяют добиться уникальных характеристик транзисторов – рабочего напряжения свыше 60 В, частот до 70 ГГц, мощности до 200 Вт. Уже продемонстрированы GaN HEMT с выходной мощностью свыше 176 Вт, рабочим напряжением 63 В, КПД 54% при усилении 12,9 дБ на частоте 2,1 ГГц. Рекордное на сегодня значение удельной мощности GaN HEMT – 32,2 Вт/мм при КПД 54,8%, рабочем напряжении 120 В и частоте 4 ГГц.

Карбид кремния как материал для СВЧ-приборов активно осваивается последние десять лет. Его главный недостаток – невозможность создания на нем гетероструктур. Вследствие низкой подвижности носителей ($370 \text{ см}^2/\text{Вс}$ для политипа 6Н) рабочие частоты SiC-приборов не превышают 4–7 ГГц. Однако остальные достоинства, прежде всего высокая плотность мощности и большая теплопроводность, заставляют активно развивать их промышленные технологии. Зарубежной промышленностью выпускаются мощные дискретные транзисторы SiC MESFET, изготовленные на 4H-SiC-подложке: номинальная выходная мощность – 60 Вт; КПД прибора при выходной мощности около

80 Вт и линейном усилении 12 дБ в полосе частот 0,9–1,6 ГГц превышает 40% (50% на частоте 1 ГГц). Кроме дискретных освоены процесс изготовления SiC МИС на 75-мм пластинах. Их базовыми элементами служат MESFET с шириной затвора 0,5 мкм. Плотность мощности MESFET – 4 Вт/мм, КПД – 60% (при комнатной температуре), рабочее напряжение – 50 В, рабочая частота достигает 6 ГГц. При выходной мощности 3,8 Вт на частоте 3,5 ГГц транзисторы демонстрируют КПД 68%.

Но все достоинства SiC-структур, даже в частотном диапазоне менее 6 ГГц, разбиваются о чрезвычайно высокую стоимость SiC-пластин. Проблема эта актуальна не только для SiC-транзисторов, но и для всех мощных приборов на основе широкозонных полупроводников. Действительно, большая плотность мощности означает малую площадь теплоотвода. Следовательно, необходимы подложки с высокой теплопроводностью. Один из лучших с этой точки зрения материалов – карбид кремния. Практически все GaN-приборы с рекордными характеристиками выполнены именно на SiC-подложках. С ценой на GaN-подложки ситуация еще тяжелее. Этот материал хоть и уступает по теплопроводности SiC, но в качестве подложки для GaN-приборов предпочтителен благодаря структурному соответствию. GaN-подложки, причем диаметром всего 50 мм, имеют высокие цены, число пластин ограничено, а их качество (плотность дефектов) не отвечает требованиям, предъявляемым к пластинам для крупномасштабного производства.

Если изготовление полуизолирующих GaN-подложек осваивается, то гораздо сложнее с AlN-подложками, их диаметры малы – 25 мм. Правда, есть надежда, что в скором времени это положение изменится. Кроме того, ведутся работы по созданию подложек на основе искусственных алмазов. В то же время значимые успехи достигнуты в области создания GaN-приборов на недорогих Si-подложках с кристаллографической ориентацией (111). На GaN-эпитаксиальных структурах на таких подложках создана МИС двухкаскадного усилителя X-диапазона мощностью 25 Вт. Ширина затвора транзистора в выходном каскаде составляет 11,4 мм, размер кристалла – 3×4,5 мм. Мощ-

ность 25 Вт в импульсном режиме была достигнута на частоте 10 ГГц при напряжении стока 30 В и усилении 15 дБ, КПД – 21%.

7.5 Перспективные материалы и приборы

7.5.1 Молибденит MoS_2 – перспективный материал нанoeлектроники

Когда академика РАН К. А. Валиева спросили о том, что будет в развитии электроники после кремния, он ответил: «Кремний – это фундамент микроэлектроники. Я думаю, что кремний будет всегда, по крайней мере, до тех пор, пока совершенно новое не превратит кремниевые СБИС (сверхбольшие интегральные схемы) в изделия «каменного века» [97].

Среди современных альтернативных кремнию материалов полупроводниковой электроники можно особо выделить молибденит – дисульфид молибдена MoS_2 , английское название: Molybdenite (рис. 7.20). В природе встречается в виде минерала. В полупроводниковой электронике, в ранней радиотехнике природные кристаллы молибденита, наряду с кристаллами галенита, сульфида свинца PbS , использовались в кристаллических детекторах в радиоприемных устройствах.



Рис. 7.20 – Природный кристалл молибденита



.....

Древние греки минерал сульфида свинца (галенит) называли «молибдена». Молибденит очень похож на галенит. В результате два разных минерала не различали, принимая за один и тот же, не только древние греки, но и химики XVIII в. Поэтому, когда в 1778 г. шведский химик Карл Вильгельм Шееле обнаружил в минерале «молибдена» новый элемент, он ошибочно назвал его «молибдос» (Mo), что в переводе означает «свинец». Хотя на самом деле это был не свинец (Pb). Впоследствии в названии минерала сульфида свинца перешли с греческого («молибдена») на латынь («галена»), а за минералом состава MoS закрепилось название молибденит.

.....

Кристаллическая сингония молибденита гексагональная, в его химическом составе возможны рений Re и селен Se, поэтому молибденит служит рудной основой для добычи этих элементов. По гексагональной форме и низкой твердости молибденит можно спутать с графитом. Различают их так: при растирании черты молибденита она приобретает зеленоватый цвет, а черта графита остается серой. У галенита кубическая форма кристалла, а удельный вес и твердость выше, чем у молибденита.

В 2011 г. ученые из Швейцарии, из Федеральной политехнической школы города Лозанны (Ecole Polytechnique Federale de Lausanne (EPFL)) объявили, что в их Лаборатории наноэлектроники и наноструктур создана интегральная электронная схема, в которой вместо обычного для современной электроники кремния применен дисульфид молибдена.



.....

Молибденит имеет ширину запрещенной зоны, характерную для полупроводников, 1,8 эВ [98]. Это полупроводник, идеально подходящий для изготовления транзисторов: по некоторым своим характеристикам он потенциально превосходит кремний, а по нескольким параметрам – даже такую «экзотику», как графен.

.....

Директор Лаборатории наноэлектроники и наноструктур Андрас Кис рассказал о новой разработке:



.....
«Мы сделали первый прототип, он состоит всего из нескольких типовых транзисторов, но при этом способен производить бинарные логические операции, а это значит, что мы сможем сделать большие, намного более сложные, чипы» [99].

Важным достоинством дисульфида молибдена по сравнению с кремнием является его двумерная структура, тогда как кремний образует объемные кристаллы. Двумерная структура позволяет легко формировать тонкие пленки толщиной $6,5 \text{ \AA}$ ($0,65 \text{ нм}$), в которых подвижность электронов при комнатной температуре $200 \text{ см}^2 \text{ В}^{-1} \text{ с}^{-1}$ соответствует подвижности электронов в слое кремния толщиной 2 нм [98].



.....
 Преимущества молибденита проявляются в связи с дальнейшей миниатюризацией кремниевых чипов и отдельных их элементов, таких как транзисторы. Слой кремния невозможно сделать тоньше двух нанометров – иначе он начинает окисляться, что резко снижает его электронные качества. Интегральная схема из молибденита стабильно работает даже при толщине в три атома, что потенциально позволяет делать гораздо более миниатюрные чипы. Транзисторные MoS_2 -ключи могут переключаться быстрее, чем кремниевые. По механическим качествам молибденит представляется весьма привлекательным материалом для использования в гибкой электронике, из него можно создавать целые «простыни» из микрочипов, которые впоследствии будут использовать, например, для производства компьютеров, сгибаемых в трубку, или для приборов, наносимых непосредственно на человеческую кожу.

Испытания молибденитового микрочипа подтвердили прежние предположения сотрудников Лаборатории наноэлектроники и наноструктур о том, что молибденитовая электроника способна преодолеть физические ограничения, наложенные на кремниевую в отношении таких характеристик, как миниатюризация, электропотребление и механическая гибкость.

На рисунке 7.21 показан полевой транзистор со сверхмалым энергопотреблением – транзисторы, произведенные из молибденита, в неактивном состоянии потребляют в 100 тысяч раз меньше энергии, чем транзисторы из традиционного кремния. Канал транзистора из дисульфида молибдена сформирован на подложке из оксида кремния и отделен от затвора слоем материала с высокой диэлектрической проницаемостью ϵ – диоксида гафния. Об использовании диоксида гафния в качестве диэлектрика вместо традиционного диоксида кремния мы говорили в главе 1, п. 1.5.

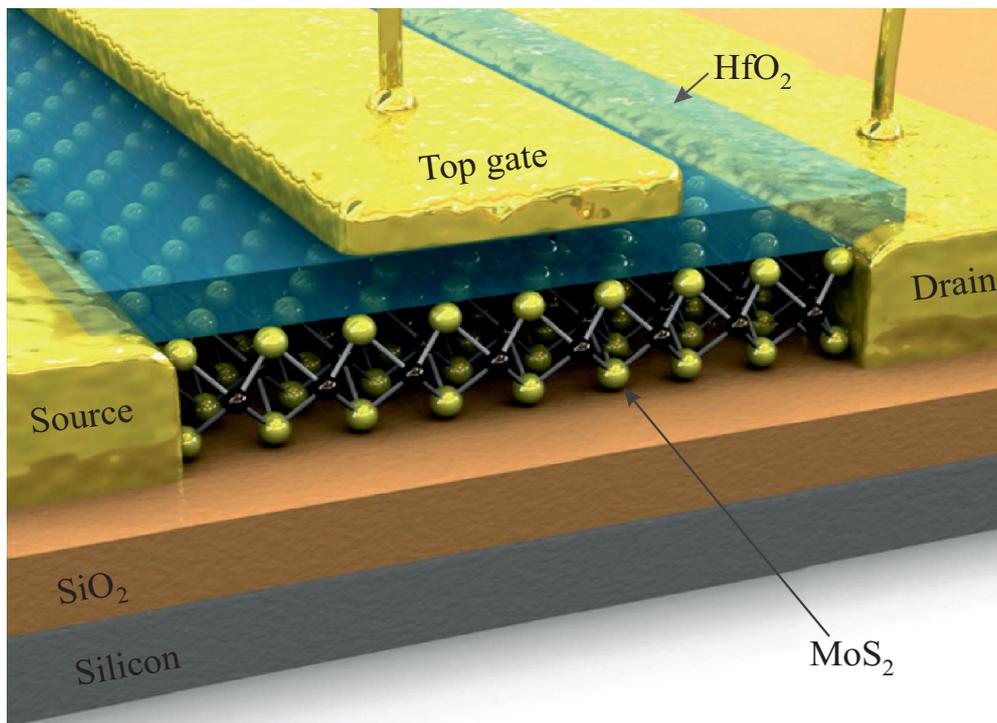


Рис. 7.21 – Молибденитовый полевой транзистор: source – исток, drain – сток, Top gate – затвор, HfO_2 – диэлектрик оксид гафния

7.5.2 Алмаз как материал для СВЧ-приборов



Алмаз как полупроводниковый материал давно привлекает внимание исследователей благодаря своим выдающимся электрофизическим характеристикам. При комнатной температуре ширина его запрещенной зоны – 5,47 эВ, подвижность электронов и дырок – 1 800 и 1 200 см²/Вс соответственно (может достигать 4 500 и 3 800 см²/Вс соответственно). Как показывают расчеты, на основе алмазов можно создавать МДП-структуры с удельной плотностью мощности до 27 Вт/мм, плотностью носителей заряда в тонком слое до 10¹³ см⁻², напряжением пробоя порядка 10 МВ/см и рабочими частотами 100 ГГц. Не менее важное свойство алмазов – их высокая теплопроводность, около 20 Вт/К см, что более чем в 30 раз выше, чем у GaAs.

Поэтому первые попытки применения алмазов в СВЧ-приборах были направлены на создание теплопроводящих подложек, как правило, методом химического осаждения тонких пленок из газовой фазы (CVD-алмазы) [100, 101]. Это направление развивается и сегодня, в том числе в России. С формированием активных приборных структур на основе алмазов трудностей гораздо больше. Технические проблемы главным образом связаны с легированием. Легирующие примеси алмаза – бор (акцептор), азот и фосфор (доноры). Но в структуре алмаза энергия активации (ионизации) этих элементов весьма высокая – у бора $E_A = 0,37$ эВ, фосфора и азота – 0,62 и 1,7 эВ соответственно. Это означает, что при комнатной температуре донорная примесь не будет активироваться, и единственный легирующий элемент – это бор. Но энергия ионизации бора также велика, и полная активация акцепторных связей при комнатной температуре наступает при концентрации свыше 10²⁰ см⁻³.

Путем легирования тонкого слоя алмаза бором ученые факультета электронных приборов и схем университета г. Ульм (Германия) совместно со спе-

циалистами компании Element Six (Великобритания) создали в начале 2005 г. действующую MESFET-структуру (рис. 7.22) [100, 101]. На подложку искусственного алмаза, выращенного при высоких значениях давления и температуры (high pressure high temperature – HPHT) (ориентация 100), наносился слой CVD-алмаза толщиной 250 нм. На нем вырастили δ -слой CVD-алмаза, легированного бором с максимальной концентрацией $5 \cdot 10^{20} \text{ см}^{-3}$. Толщина этого δ -слоя на уровне половины максимальной концентрации не превышала 10 нм. Поверх δ -слоя осаждался тонкий ($< 20 \text{ нм}$) слой CVD-алмаза. Переход Шоттки создавался путем нанесения на этот CVD-слой алюминиевого затвора. Канал формировался непосредственно над легированным δ -слоем.

Отметим, что чем острее профиль легирования δ -слоя, тем выше подвижность дырок в канале, которая может достигать $3\,800 \text{ см}^2/\text{Вс}$. Ток сток – исток первого «алмазного» транзистора не превысил 1 мкА при размере затвора $100 \times 0,5 \text{ мкм}$. В целом эксперимент подтвердил правильность расчетов. Таким образом, первый транзистор на CVD-алмазах может стать родоначальником семейства перспективных приборов.

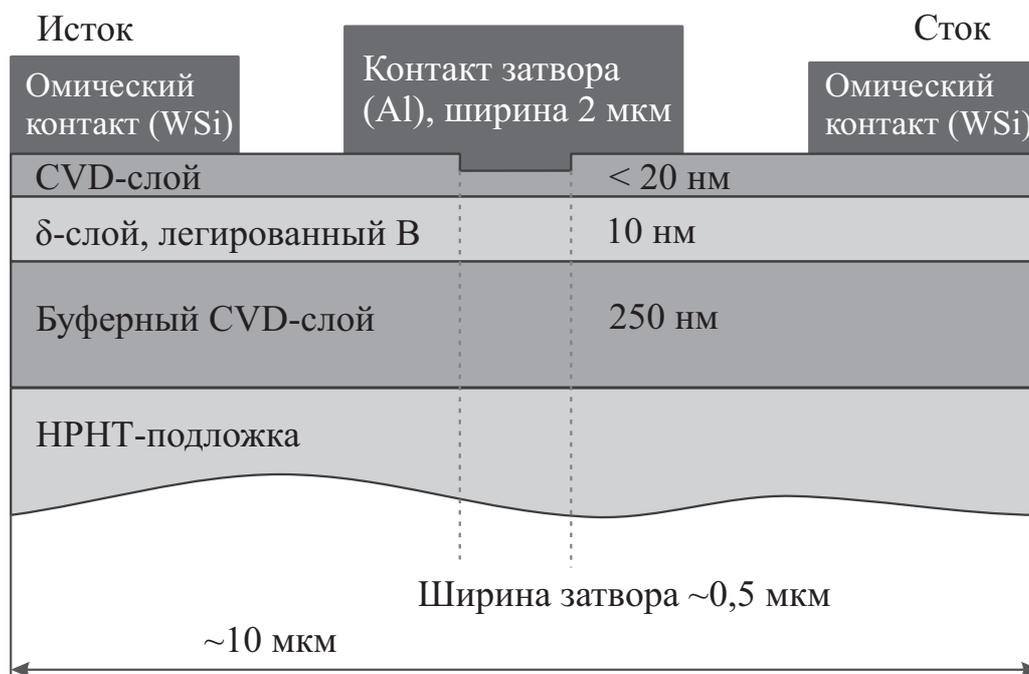
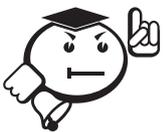


Рис. 7.22 – Структура первого MESFET-транзистора на основе алмаза с δ -слоем, легированным бором [101]

Есть сообщения и о другом способе создания транзисторных структур на основе алмазов. Так, обработка поверхности алмазной пластины водородом (hydrogen termination) приводит к появлению тонкого приповерхностного слоя с проводимостью p -типа, в результате формируется слой двумерного дырочного газа с концентрацией носителей $\sim 10^{13} \text{ см}^{-2}$. Однако пока физика этого эффекта до конца не объяснена. Тем не менее, уже созданы опытные МДП-структуры на основе поверхностной проводимости алмазных пленок. Успехов в этой области добились упомянутые исследователи университета г. Ульм, создав как планарный МДП-транзистор с использованием CaF_2 в качестве затворного диэлектрика, так и MESFET с самосовмещенным Т-образным затвором. MESFET-транзисторы изготавливались на различных подложках – искусственных НРНТ-алмазах с ориентацией (111) и (100), CVD-монокристаллических подложках и квазиподложках, выращенных на Ir/SrTiO_3 . Во всех случаях были сформированы транзисторные структуры с нормально обогащенным каналом и малыми токами утечки. При длине канала 0,2 мкм плотность тока достигала 360 мА/мм. Измеренная подвижность дырок на поверхности оказалась на порядок ниже расчетной ($150 \text{ см}^2/\text{Вс}$), что объясняется рассеянием на поверхностных состояниях. Тем не менее, при низких уровнях тока удалось добиться превосходных частотных характеристик: $f_T = 25 \text{ ГГц}$ и f_{max} до 81 ГГц. На частоте 3 ГГц уровень собственных шумов составил 0,72 дБ.

7.5.3 Транзисторы на антимонидах и арсенидах индия



Первые сообщения о AlSb/InAs/AlSb НЕМТ появились в конце 1980-х гг. в работе известного нам лауреата Нобелевской премии за развитие гетероструктур Г. Кремера [102]. Однако вплоть до недавнего времени ощутимого прогресса в этой области не наблюдалось. Сегодня исследования приборов на основе этих материалов поддерживаны программой DARPA «Сложные полупроводники на основе антимонидов» (Antimonide Based Compound Semicon-

ductor – ABCS). DARPA (Defense Advanced Research Projects Agency) – агентство передовых оборонных исследовательских проектов, США. Речь идет о полупроводниковых соединениях GaSb, InAs и AlSb с близкими значениями постоянных кристаллической решетки – примерно 6,1 ангстрем. У таких материалов весьма узкая запрещенная зона, но чрезвычайно высокие значения подвижности электронов (до 33 000 см²/Вс) и скорости насыщения носителей (свыше 4·10⁷ см/с) при комнатной температуре. Это позволяет создавать на их основе НЕМТ с чрезвычайно высоким быстродействием и рабочими напряжениями порядка 0,5–1 В.

.....

В этой области уже достигнуты значимые практические результаты. Так, сотрудники компании NGST совместно со специалистами Исследовательской лаборатории ВМС США (NRL) сообщили о создании AlSb/InAs МНЕМТ-структур с длиной затвора 0,1 мкм. Разрыв границ зоны проводимости в области гетероперехода AlSb/InAs достигает 1,35 эВ, что приводит к формированию глубокой квантовой ямы и обуславливает высокую концентрацию носителей в 2DEG-области (до 1,28·10¹² см⁻² при подвижности электронов 26 300 см²/Вс). Крутизна транзисторов достигает 1,3 См/мм при рабочем напряжении до 0,3 В, значения f_T и f_{max} превысили 220 и 270 ГГц соответственно, напряжение пробоя сток – затвор – 1,42 В.

При таких высоких частотах удельная рассеиваемая мощность приборов оказалась чрезвычайно низкой. В частности, при напряжениях сток – исток 0,2 и 0,5 В граничная частота в среднем составила 160 и 220 ГГц, а ток канала – 135 и 340 мА/мм, соответственно. Столь небольшая рассеиваемая мощность (27 и 170 мВт/мм) по меньшей мере в 5–10 раз ниже, чем у InP НЕМТ на аналогичных частотах. В диапазоне 2–26 ГГц удельная рассеиваемая мощность не превысила 6 мВт/мм при коэффициенте шума 0,85 дБ и усилении 11,5 дБ. На базе AlSb/InAs НЕМТ уже создана МИС усилителя мощности, коэффициент шума которого в диапазоне 80–100 ГГц составил 5,4 дБ, усиление – 11,1 дБ, а общая

рассеиваемая мощность – 1,8 мВт. При усилении 16 дБ в полосе 77–103 ГГц рассеиваемая мощность была равна 4,41 мВт.

Еще более многообещающим по сравнению с InAs-приборами выглядят транзисторы с каналным слоем из антимонида индия. InSb характеризуется самой высокой подвижностью носителей из известных полупроводников. Его свойства были продемонстрированы более восьми лет назад в МДП-транзисторе с InSb-каналом, сформированном на InSb-подложке с SiO₂ в качестве подзатворного диэлектрика. Первый МНЕМТ-транзистор на основе Al_xIn_{1-x}Sb/InSb-структуры с 0,1-мкм затвором создали сотрудники технологического центра QinetiQ и компании Intel (рис. 7.23).

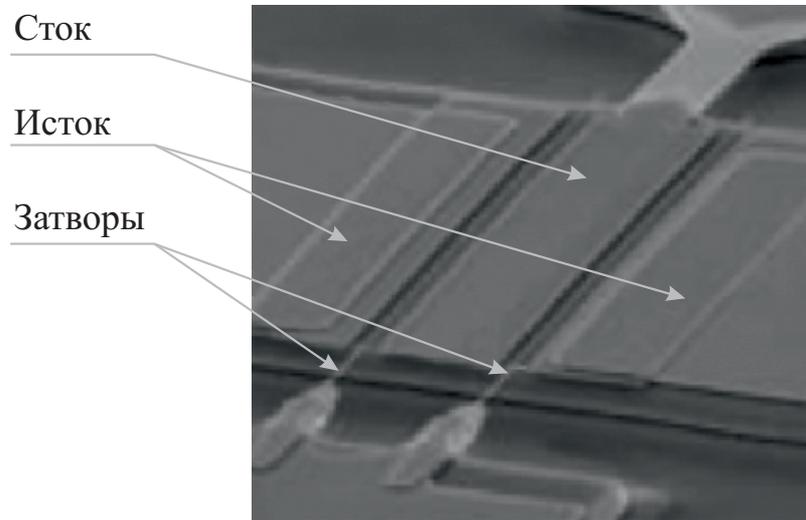


Рис. 7.23 – «Двухпальцевый» Al_xIn_{1-x}Sb/InSb МНЕМТ-транзистор с 0,1-мкм затвором

Прибор был изготовлен на GaAs-подложке, на ней располагались Al_yIn_{1-y}Sb-буферный слой толщиной 3 мкм, нижний барьерный слой Al_xIn_{1-x}Sb, InSb-каналный слой толщиной 20 нм и Al_xIn_{1-x}Sb-спейсер толщиной 5 нм. Выше следовали *d*-легированный теллуром тонкий слой Al_xIn_{1-x}Sb (с концентрацией донорной примеси 10¹² см⁻²) и верхний барьерный слой Al_xIn_{1-x}Sb толщиной 15–45 нм. При доле алюминия в барьерном слое 30% подвижность электронов в канале составила 30 000·см²/Вс при концентрации носителей 1,3·10¹² см⁻². В испытываемых InSb-транзисторах с затвором длиной 0,1 мкм и шириной 40 мкм доля Al в Al_xIn_{1-x}Sb составляла 20%, что обеспечивало концентрацию носителей

в канале 10^{12} см^{-2} и их подвижность $25\,000 \text{ см}^2/\text{Вс}$. Максимальная крутизна прибора при напряжении сток – исток $0,5 \text{ В} – 800 \text{ мСм/мм}$, ток в канале – 250 мА/мм , экстраполированные значения f_T и f_{\max} – 210 и 270 ГГц соответственно. Рассеиваемая мощность InSb-прибора примерно в 10 раз меньше, чем у кремниевого n -канального МОП-транзистора с 80-нм каналом на такой же граничной частоте. Задержка включения (вентильная задержка) по крайней мере второе меньше, чем в современных или перспективных Si-транзисторах при равной длине затвора.

7.5.4 Транзисторы на углеродных нанотрубках

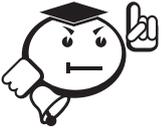


.....

Приблизительно до 2020 г. транзисторы в интегральных наносхемах будут создаваться по современной схеме работы – с двумя электродами и затвором между ними. Ясно, что определенным пределом этой тенденции было бы достижение атомного размера элементов электроники, когда уменьшать их дальше будет невозможно. Следовательно, уже сейчас необходимо искать новые подходы.

.....

Одним из таких новых подходов являются углеродные нанотрубки и кремниевые нанопровода, в которых можно достичь более высокой скорости электронов. Углеродные нанотрубки – своеобразные цилиндрические молекулы диаметром примерно от половины нанометра и длиной до нескольких микрометров. Эти системы впервые обнаружили в 1991 г. как побочные продукты синтеза фуллерена C_{60} [103], причем автор работы, используя просвечивающий электронный микроскоп с высоким разрешением, обнаружил эти структуры и описал их образование, геометрию, кристаллическую структуру. Основным элементом таких структур является графитовый слой – поверхность, выложенная правильными шестиугольниками с атомами углерода, расположенными в вершинах.



Углеродные нанотрубки – цилиндрические кристаллы, состоящие из одних лишь атомов углерода. Внешне выглядят как свернутая в цилиндр графитовая плоскость.

Нанотрубки отличаются широким разнообразием физико-химических свойств и привлекают значительный интерес исследователей и технологов. В работе [104] представлена информация о методах получения, структуре и основных физико-химических характеристиках углеродных нанотрубок. Анализируются возможные механизмы образования нанотрубок; рассматриваются вопросы прикладного использования этого нового материала.

Как показали наблюдения, выполненные с помощью электронных микроскопов, большинство нанотрубок состоит из нескольких графитовых слоев, либо вложенных один в другой, либо навитых на общую ось. Расстояние между слоями практически всегда составляет 0,34 нм, что соответствует расстоянию между слоями в кристаллическом графите. Нанотрубки обычно заканчиваются полусферической головкой, структура которой включает в себя наряду с шестиугольниками также правильные пятиугольники.

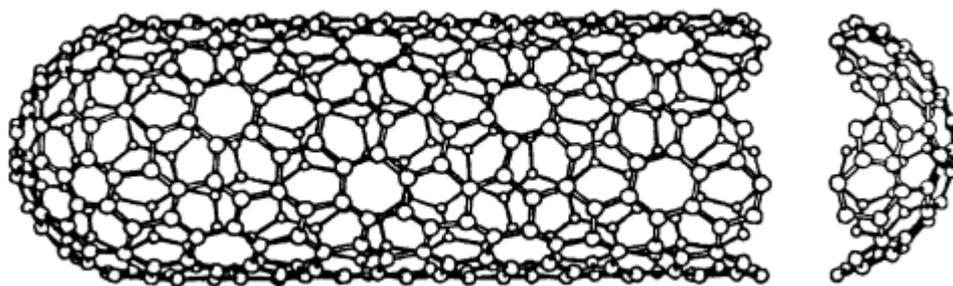


Рис. 7.24 – Идеализированная модель однослойной нанотрубки [104]

На рисунке 7.24 показана свернутая трубка, результат операции сворачивания зависит от угла ориентации графитовой плоскости относительно оси нанотрубки. Угол ориентации задает *хиральность* нанотрубки, которая определяет, в частности, ее электрические характеристики.

Хиральность нанотрубок обозначается набором символов (m, n) , указывающим координаты шестиугольника, который в результате сворачивания плоскости должен совпасть с шестиугольником, находящимся в начале координат. Некоторые из таких шестиугольников вместе с соответствующими обозначениями отмечены на рисунке 7.24. Другой способ обозначения хиральности состоит в указании угла α между направлением сворачивания нанотрубки и направлением, в котором соседние шестиугольники имеют общую сторону. Среди различных возможных направлений сворачивания нанотрубок выделяются направления, для которых совмещение шестиугольника (m, n) с началом координат не требует искажения в его структуре. Этим направлениям соответствуют угол $\alpha = 0$ (armchair-конфигурация) и $\alpha = 30^\circ$ (zigzag-конфигурация). Указанные конфигурации отвечают хиральностям $(m, 0)$ и $(2n, n)$ соответственно.

Индексы хиральности однослойной нанотрубки (m, n) однозначным образом определяют ее диаметр D . Эта связь очевидна и имеет следующий вид:

$$D = \sqrt{m^2 + n^2 - mn} \frac{\sqrt{2}d_0}{\pi}, \quad (7.4)$$

где $d_0 = 0,142$ нм – расстояние между соседними атомами углерода в графитовой плоскости. Связь между индексами хиральности (m, n) и углом α дается соотношением:

$$\alpha = \tan^{-1} \frac{\sqrt{3}n}{2m - n}. \quad (7.5)$$

Разрешающая способность современных электронных микроскопов недостаточна для непосредственного различения нанотрубок с разной хиральностью, поэтому основной способ определения данного параметра связан с измерением их диаметра [104].

С помощью углеродной нанотрубки (УНТ) можно создать полевой транзистор, принцип действия которого полностью эквивалентен работе традиционного полевого транзистора, за исключением того, что каналом переноса но-

сителей заряда является углеродная трубка. Простейшая схема такого транзистора изображена на рисунке 7.25.

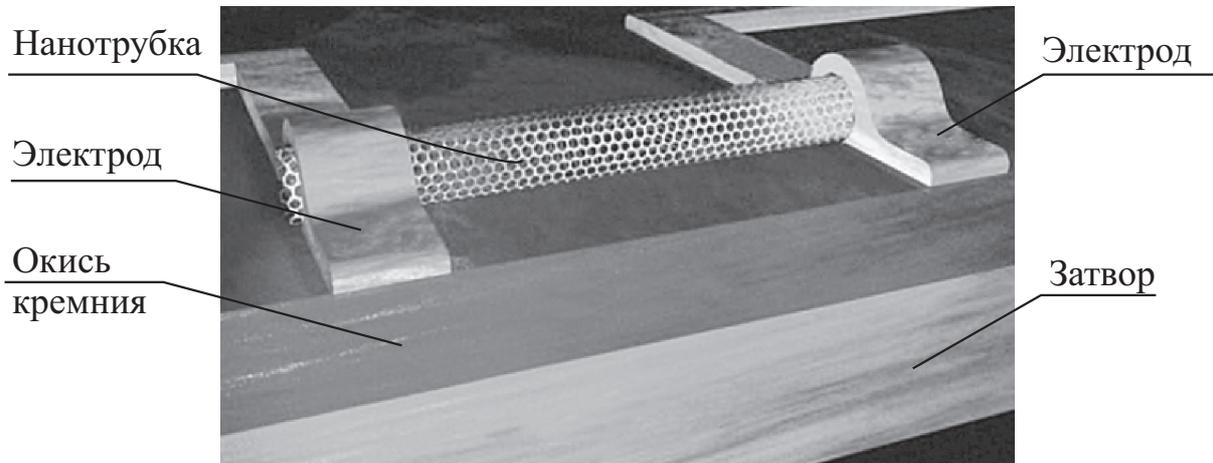


Рис. 7.25 – Схема полевого транзистора на УНТ

Транзистор изготавливают следующим образом. На кремниевую пластину наносят пару электродов – сток и исток, между которыми располагают нанотрубку. Сама пластина является затвором. В обычном состоянии канал закрыт, так как имеется потенциальный барьер для дырок. Зона проводимости и валентная зона разделены запрещенной зоной с шириной в несколько эВ (рис. 7.26, *а*).

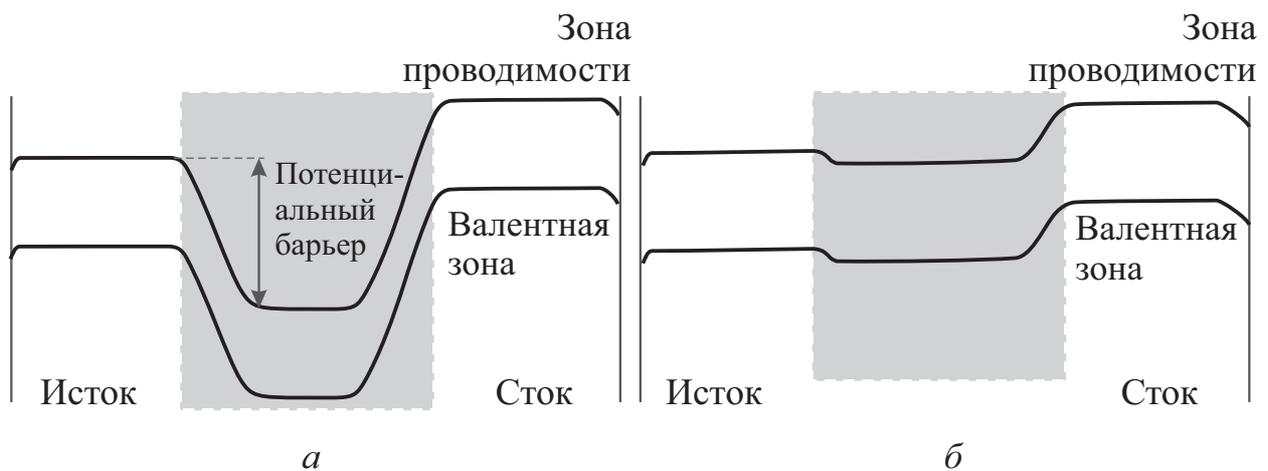


Рис. 7.26 – Зонные диаграммы транзистора без внешнего электрического поля (*а*) и в его присутствии (*б*)

Но если на затвор подать напряжение, которое приведет к возникновению электрического поля там, где находится УНТ, то ее зонная диаграмма перестраивается, она становится хорошим проводником (рис. 7.26, б).

Таким образом, меняя напряжение на затворе, можно управлять проводимостью нанотрубки и соответственно открывать или запирают транзистор.



.....

Чем УНТ-транзистор может быть лучше кремниевого? Во-первых, скорость работы УНТ-транзистора намного превосходит быстродействие кремниевых транзисторов. По некоторым оценкам, нанотрубка может работать на частоте в 1 ТГц, что в сотни раз быстрее, чем скорости современных компьютеров. В настоящее время уже созданы устройства на основе нанотрубок, работающие на частотах до 30 ТГц, что на порядок больше тактовой частоты хорошего современного процессора. Это достигается за счет высокой подвижности электронов в нанотрубках (в кремнии этот параметр составляет $1\,400\text{ см}^2/\text{Вс}$, а в нанотрубках – около $100\,000\text{ см}^2/\text{Вс}$).

Во-вторых, теоретический предел для миниатюризации кремниевых элементов составляет 12 нм. Для УНТ такого предела нет, размеры элементов на их основе могут достигать размеров молекулы. Уже созданы транзисторы размером $18\times 1\text{ нм}$, которые, даже без существенной оптимизации технологии их изготовления, по многим параметрам работают не хуже кремниевых, гораздо больших по размерам. И это не предел миниатюризации.

В-третьих, процесс производства транзисторов на основе УНТ может быть сделан значительно более простым, чем производство кремниевых элементов. Это возможно благодаря технологии печати транзисторов краской, состоящей из углеродных нанотрубок. Возможна печать транзистора целиком, включая электроды, слои изоляции и каналы из УНТ. Несомненным преимуществом создаваемой технологии станет резкое снижение количества вредных

веществ, поступающих в окружающую среду; выбросы углекислого газа, по оценкам, могут быть уменьшены более чем на 90% [105].

.....

Важным для применения в электронике свойством нанотрубок является то, что согласно расчетам электронные свойства, а также хиральность («скрученность» атомарной решетки) идеальной структуры нанотрубок меняются при внедрении в однослойную нанотрубку в качестве дефекта пары пятиугольник – семиугольник. При рассмотрении структуры (8,0)/(7,1) расчеты показали, что трубка с хиральностью (8,0) – это полупроводник с шириной запрещенной зоны 1,2 эВ, а трубка с хиральностью (7,1) является полуметаллом с нулевой шириной запрещенной зоны. Это позволяет создавать на базе нанотрубок электронные приборы (диоды, транзисторы, резисторы), подобные традиционным кремниевым.

Гетеропереходы полупроводник – полупроводник с различными значениями ширины запрещенной зоны получают таким же образом посредством внедрения дефекта. Поэтому нанотрубку с внедренным в нее дефектом можно рассматривать в качестве гетероперехода металл – полупроводник. На основе этого гетероперехода можно реализовать полупроводниковый элемент очень малых размеров, меньше нынешних кремниевых.

Уменьшение размеров кремниевых транзисторов за последние десятилетия привело к существенному удешевлению электроники, снижению энергопотребления и росту производительности компьютеров. Однако технологические нормы 10 нм считаются тем рубежом, где вступают в силу непреодолимые физические ограничения. Дальнейшее уменьшение размеров затрудняет управление потоком электронов в кремниевом канале. Уже при переходе к нормам 22 нм компании Intel пришлось изменить структуру транзистора, отказавшись от использовавшейся в течение десятилетий планарной схемы в пользу объемной компоновки. Однако кремний остается кремнием, говорят ученые, и дальнейшее уменьшение размеров элементов невозможно даже при использовании новой структуры. Поскольку возможности кремния безграничны, а токи

утечки транзисторов растут по мере уменьшения геометрических размеров элементов, исследователи активно ищут замену кремнию.

Недавно ученым IBM удалось создать транзистор размером 9 нм на основе цилиндрических углеродных структур, конструкция и характеристики которого показаны на рисунке 7.27 [106]. На рисунке 7.27, б показана нанофотография структуры сверху, а сечение ее показано на рисунке 7.27, в, на котором различима длина канала 9 нм. По своим характеристикам новый прибор превосходит кремниевые транзисторы сопоставимого размера.

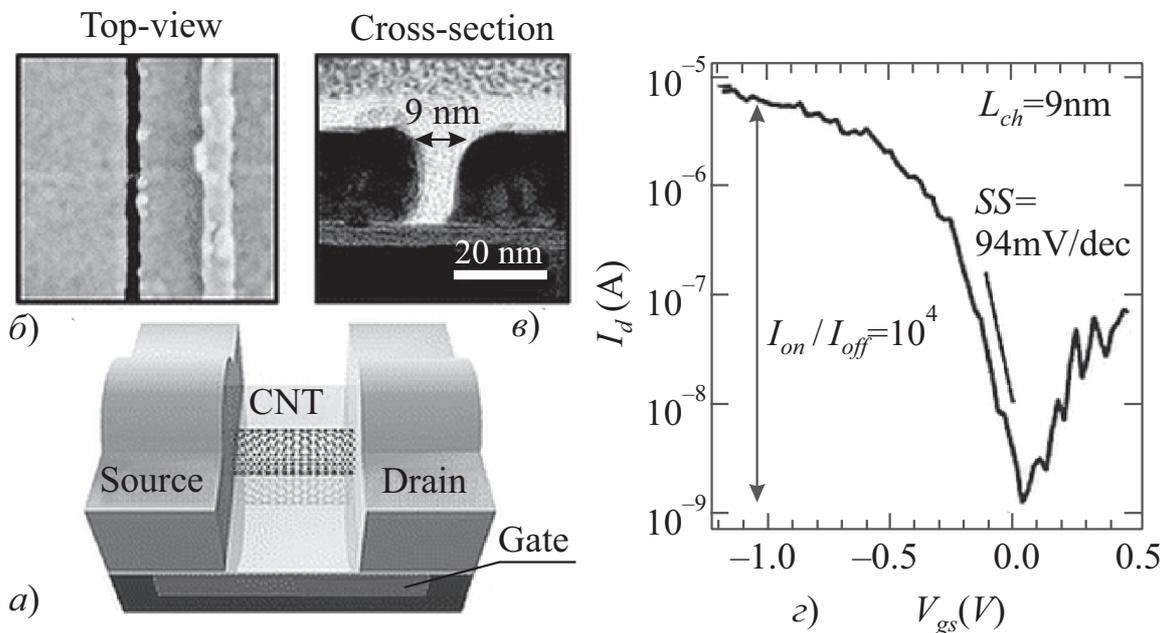


Рис. 7.27 – а) модель транзистора: затвор (gate), исток (source), сток (drain), УНТ (CNT); б) вид сверху; в) поперечное сечение (Cross section); г) электрические характеристики [106]

Очень важно было убедиться, что нанотрубки имеет смысл применять при размерах транзисторов меньших 10 нм, поскольку иначе у них нет преимуществ над кремнием.

Чтобы выяснить, как размер влияет на работу транзистора, исследователи сформировали несколько транзисторов разного размера на одной нанотрубке. Это позволило сравнить их работу, исключив влияние, связанное с вариациями параметров нанотрубок.

Эксперимент показал, что транзистор размером 9 нм потребляет гораздо меньше электроэнергии, чем другие транзисторы такого же размера.

Пока речь идет об отдельных транзисторах. Технология находится на лабораторном этапе и не подходит для серийного производства. Для практического применения разработки предстоит создать технологию изготовления чистых углеродных нанотрубок (без примеси металлических нанотрубок, которые могут стать причиной замыканий) и их точного позиционирования на подложке.



Контрольные вопросы по главе 7

1. Опишите принцип действия одноэлектронного и резонансно-туннельного транзистора.
2. Кто впервые в мире изготовил полупроводниковый лазер непрерывного действия? На какой структуре и чем обусловлено его непрерывное действие?
3. Почему реализуется высокая подвижность носителей заряда в НЕМТ-структурах?
4. Опишите перспективный полупроводниковый материал молибденит.
5. Что такое углеродные нанотрубки? Поясните смысл параметра хиральность.

8 Графеновая электроника

8.1 Получение и открытие свойств графена



.....

В сентябре 2004 г. редакция престижного научного журнала Science, после долгих девяти месяцев доработки, приняла к публикации статью А. Гейма и К. Новоселова (Department of Physics, University of Manchester) и их соавторов из Манчестерского университета, а также тех соавторов, кто работает в России (Институт проблем технологии микроэлектроники АН СССР, Черноголовка) [107]. Наиболее существенной частью этой работы были электрические измерения по влиянию электрического поля на свойства двумерного слоя углерода, то есть слоя толщиной в один атом, но имеющего длину и ширину. Такой атомарный слой углерода в свое время получил название графен, а результаты работы 2004 г. [107] послужили основой для Нобелевской премии по физике 2010 г. для Андрея Гейма совместно с Константином Новоселовым [108].

В лабораторных условиях графен можно получить простым способом. Высокоориентированный пиролитический графит (ВОПГ) – это стандартно используемый материал для сканирующей туннельной микроскопии (СТМ), где образец со свежей поверхностью обычно готовится путем отслаивания верхнего слоя графита с помощью липкой ленты [108]. А. Гейм отметил: «Мы годами использовали эту технику, но никогда внимательно не смотрели на то, что же мы выбрасываем вместе с лентой» [108, с. 1290]. Когда остатки графита на скотче были рассмотрены в микроскоп, то были обнаружены фрагменты очень малой толщины (рис. 8.1), некоторые из чешуек были толщиной всего в несколько нанометров [108].

.....

На рисунке 8.1, *з* показано одно из первых лабораторных устройств. В центральной части находится кристалл графита толщиной ~ 20 нм, а его поперечный размер соизмерим с диаметром человеческого волоса, ~ 50 мкм. К кристаллу изготовлены четыре близко расположенных контакта с помощью серебряной пасты. В самом первом образце, сделанном вручную на стекле, ясно проявился эффект электрического поля (ЭЭП), который заключался в том, что сопротивление образца можно было изменять на несколько процентов.

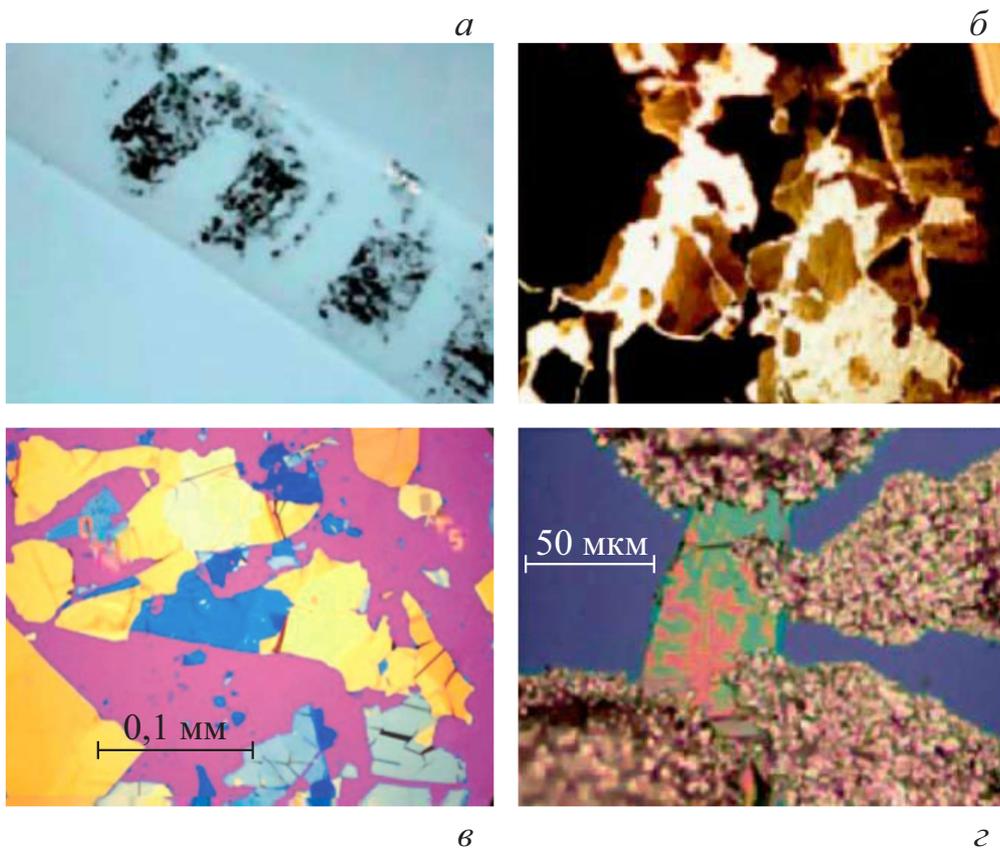
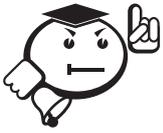


Рис. 8.1 – *а*) следы ВОПГ, оставшиеся на скотче; *б*) некоторые кристаллики оптически прозрачны, если взглянуть на них через оптический микроскоп или даже через увеличительное стекло; *в*) на подложке из оксида кремния светопрозрачные кристаллики дают различные оттенки голубого цвета; *з*) одно из самых первых устройств, изготовленное вручную в лабораторных условиях с помощью пинцета, зубочистки и серебряной пасты для электрических контактов [108]

Это было первое существенное достижение. Если эти устройства примитивного вида, сделанные вручную из относительно больших и толстых пласти-

нок, уже демонстрируют некоторое влияние внешнего поля, то, если использовать самые тонкие кристаллы и применить весь арсенал технологического оборудования, результаты будут еще лучше [108].



.....

В работе 2004 г. [107] впервые был показан амбиполярный эффект электрического поля, при котором сопротивление графенового образца меняется приблизительно в 100 раз. Свойства графена можно изменить простым изменением напряжения на затворе. Графен можно перестраивать из состояния, близкого к нормальному металлу с концентрацией электронов $\sim 10^{21} \text{ см}^{-3}$, до металла с такой же концентрацией дырок, т. е. можно проделать весь путь, минуя «полупроводниковое» состояние с невысокой концентрацией носителей заряда [108].

.....

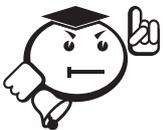
За период с 2005 по 2010 г. в лаборатории Гейма выполнен цикл приблизительно десятка научных работ [108], из которых установлены важные свойства графена как материала, пригодного для современной полупроводниковой электроники. В физике полупроводников качество электронных свойств определяется подвижностью носителей заряда μ . В работе [107] для графена при комнатной температуре приводится значение $\mu = 10\,000 \text{ см}^2/\text{В}\cdot\text{с}$. Такая большая подвижность носителей заряда (а для электронов это самая большая подвижность среди всех известных материалов) делает его перспективным материалом для самых разнообразных применений, например как основу наноэлектроники [109] или замену кремния в микросхемах в будущем.

Элементарная частица вещества в общем случае называется фермион. Примеры фермионов – кварки, лептоны и дырки в полупроводнике. Из интересных для электроники фермионов можно выделить электроны, которые относятся к лептонам.

Выяснилось, что электроны в графене ведут себя как релятивистские частицы без массы, так называемые безмассовые фермионы Дирака, передвигаясь

с максимально возможной для среды скоростью, что для графена приблизительно 10^6 м/с, поэтому их движение нужно описывать не стандартным уравнением Шредингера, а уравнением типа уравнения Дирака. Электрическое напряжение, приложенное к подложке, на которой находится графен, создает носители заряда – электроны либо дырки, что зависит лишь от знака напряжения.

У графена отсутствует запрещенная зона – энергетический зазор запрещенных энергий между валентной зоной и зоной проводимости. Установлено, что графен остается металлическим в пределе отсутствия носителей заряда, даже тогда, когда в устройстве микронных размеров остается всего лишь несколько электронов. Предложено использовать графеновые устройства для регистрации туннелирования Клейна – эффекта, известного в квантовой электродинамике на протяжении многих десятилетий, но считавшегося ненаблюдаемым, пока несколько групп не продемонстрировали эффект экспериментально.



В двухслойном графене электроны приобретают новый облик, становясь теперь *массивными* дираковскими фермионами, то есть имеют конечную массу. Эти свойства раскрылись при обнаружении двух новых типов целочисленного квантового эффекта Холла, соответствующих двум типам дираковских фермионов. Установлено, что двуслойный графен – это полупроводник, причем с управляемой запрещенной зоной, и из него можно точно вырезать нанометровые устройства [108].

8.2 В направлении к графеновой электронике

Известно, что современная микросхема изготавливается на пластинке кремния, причем, чем больше таких пластинок умещается на одной пластине, вырезанной из выращенного слитка кремния, тем лучше для массовости произ-

водства микросхем, снижения их себестоимости. Поэтому требовалось и требуется получение кремниевых слитков и, соответственно, пластин из них, большего диаметра, чем больше, тем лучше. Это же относится и к другим перспективным материалам полупроводниковой электроники, например к арсениду галлия и, само собой, к графену.

Изготовление графеновых пластин, на базе которых становится возможным массовое производство графеновых полупроводниковых изделий, связано с определенными сложностями, прежде всего с его уникальной одноатомной толщиной. Методы создания пластинок небольшой площади, достаточные для лабораторных опытов, совсем не подходят для промышленных масштабов.

Одно дело, фигурально выражаясь, послойно отдирать скотчем крохотные пластинки хаотично напластованного графена от графита, и совсем другое дело – получать пластины диаметром в десятки и сотни миллиметров с четко ориентированными слоями графена.

На рисунке 8.2 показана структура графена. Массовое производство интересно поточный и, по возможности, недорогой метод получения исходного материала высокой чистоты. К настоящему времени известно несколько методов получения графеновых пластин, в определенной мере удовлетворяющих потребностям электронной промышленности.

1. Классический метод химического осаждения углерода из распыленного газообразного состояния. Разработан в Массачусетском технологическом институте (Massachusetts Institute of Technology) под руководством профессора Джинг Конга (Jing Kong). Используется оборудование, схожее с применяемым в обычном полупроводниковом производстве [110]. Метод недорог и вполне пригоден для коммерческого внедрения, возможно, что для специфических приложений, например для производства чипов, установку придется доработать для повышения качества и однородности получаемых на выходе листов графена. Тем не менее, уже в нынешнем виде установка вполне годится для получения графена для таких приложений, как электроды солнечных батарей [111].

2. Химическое осаждение из газовой фазы в парах метана и водорода на медную фольгу. Метод разработан коллективом ученых из Техасского университета в Остине (The University of Texas at Austin) и из компании Texas Instruments [113].

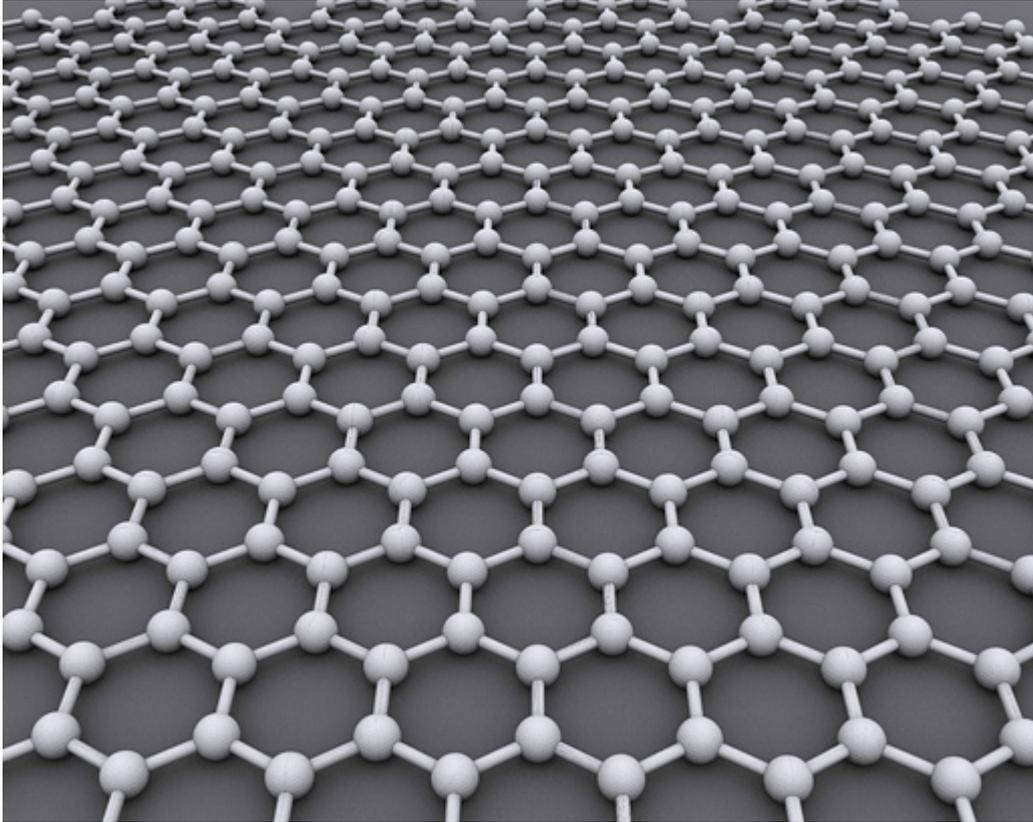
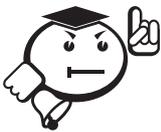


Рис. 8.2 – Графен представляет собой гексагональную решетку из атомов углерода толщиной в один атом



Успешно получены графеновые пленки, перенесенные на обычную кремниевую пластину с верхним слоем диоксида кремния. Именно синтез графена на кремниевых подложках методами, совместимыми с традиционными технологическими процессами полупроводниковой индустрии, считается основой будущего нанoeлектроники. Выращены графеновые пластины площадью 10×10 мм на медной фольге, однако размер графеновых пластин теоретически

ограничен лишь возможностями использовавшегося оборудования, а именно обжиговой печи.

.....

Полученные в результате экспериментов образцы площадью 1 квадратный сантиметр почти полностью покрыты слоем мономолекулярного графена; лишь малая площадь – менее 5% – покрыта двуслойными и трехслойными «напластованиями». На основе полученных графеновых пластин были созданы 2-затворные полевые транзисторы с находящимся сверху затвором, электрически изолированным от графена тончайшим слоем окиси алюминия.

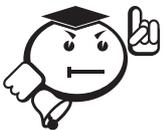
Определение подвижности носителей полученных таким образом транзисторов показало результаты, значительно более высокие, нежели у традиционных кремниевых полупроводников – более $4\ 050\ \text{см}^2/\text{Вс}$.

3. Процесс кремниевой сублимации карбида кремния. Разработан в Электрооптическом Центре (Electro-Optics Center, EOC) при кафедре материаловедения Пенсильванского университета (Pennsylvania State University).

Впервые изготовлена графеновая пластина диаметром 100 мм (4 дюйма) (рис. 8.3). При термической обработке пластины карбида кремния в высокотемпературной печи после полного удаления кремния на поверхности остается слой пленки из углерода толщиной в один-два атома, что, по сути, и является пластиной самого настоящего графена [113].

Достижения ученых из EOC можно назвать серьезной вехой на пути развития графеновой электроники в направлении массового производства. Даже 4-дюймовые графеновые пластины могут со временем стать основой для производства различных электронных компонентов. Так, исследователи из EOC уже в самое ближайшее время планируют начать исследования высокочастотных полевых транзисторов, производимых на 100-мм графеновых пластинах таким «полупромышленным» способом.

.....



Благодаря тому что электроны могут двигаться в графене с очень высокой скоростью (значительно большей, чем в кремнии),

графеновая электроника изначально сулит появление терагерцовых процессоров, скорость которых теоретически сможет превышать современные кремниевые полупроводники на пару порядков. Однако технологам придется еще поработать над чистотой и однородностью получаемых графеновых пластин, поскольку именно от этого в первую очередь зависит возможность достижения предельных тактовых частот получаемых полупроводников.

.....

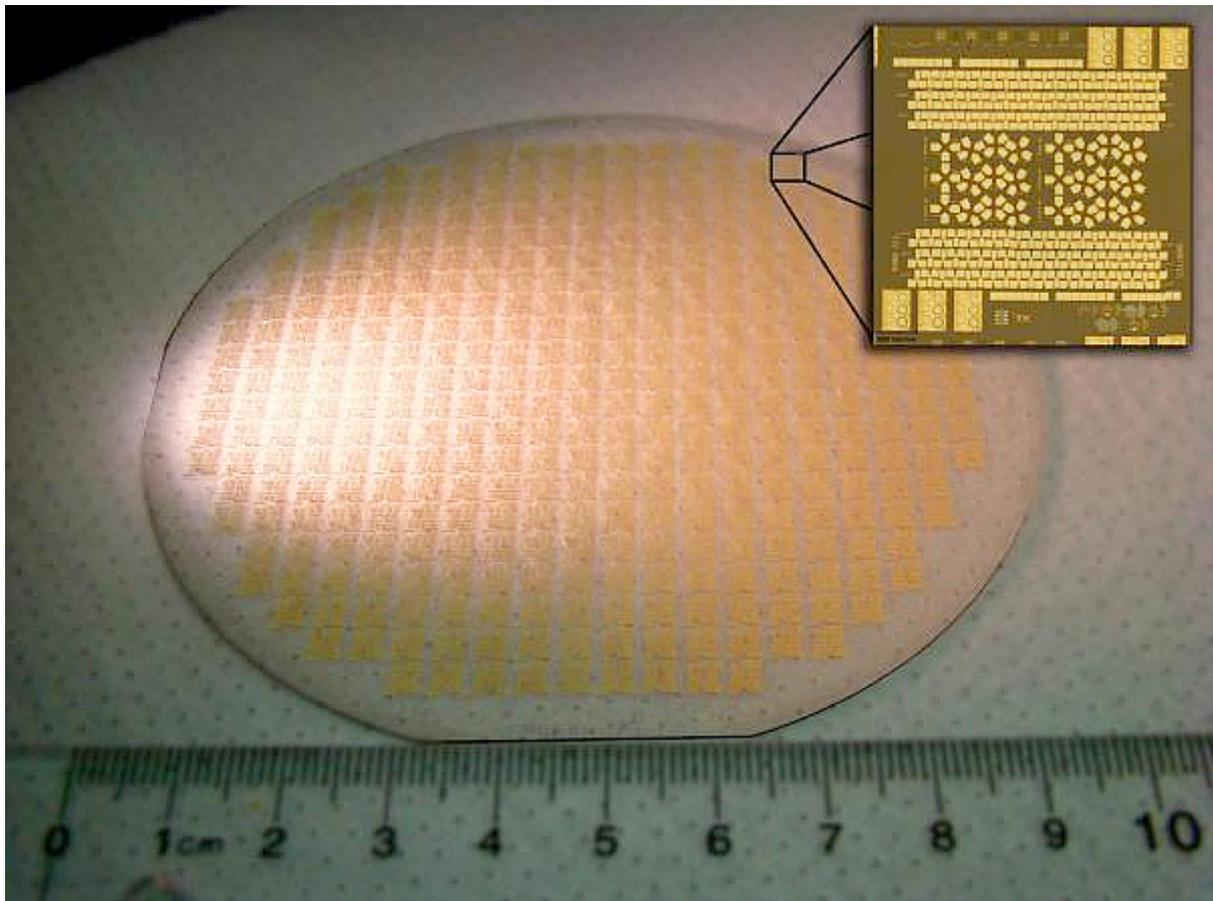


Рис. 8.3 – 100-мм пластина графена, содержащая приблизительно 75 000 устройств и тестовых структур. На вставке – оптическое изображение одного чипа. Каждый маленький квадрат на чипе составляет всего 100 мкм – толщина человеческого волоса. Устройства были изготовлены на Penn State Nanofab, учреждение Научно-исследовательского института материалов (фото: Джошуа Робинсон) [113]

4. Метод синтезирования графена на кремнии. Находится в разработке в Электрооптическом центре (Electro-Optics Center, EOC) при кафедре материаловедения Пенсильванского университета (Pennsylvania State University).

В случае успеха это позволит производить пластины диаметром 200 мм и более, а значит, плавно интегрировать процессы производства графеновых полупроводников в уже существующую инфраструктуру полупроводниковой индустрии.

8.3 Первые транзисторы на графене

Так как графен впервые был получен всего несколько лет назад, то вполне естественно, что в настоящее время пока еще нет массового производства работающих устройств на его основе, хотя список перспективных технологий довольно обширный.

На рисунке 8.4 приведен пример возможной реализации одноэлектронного транзистора на базе графена. Слева показана схема транзистора, состоящего из графеновых истока (source) и стока (drain), соединенных островком (island) из проводящего материала или из квантовой точки, шириной около 100 нм. Справа показан тестовый транзистор, изображение которого увеличено в 40 000 раз. Островок транзистора настолько мал, что способен уместить только один электрон в один момент времени.

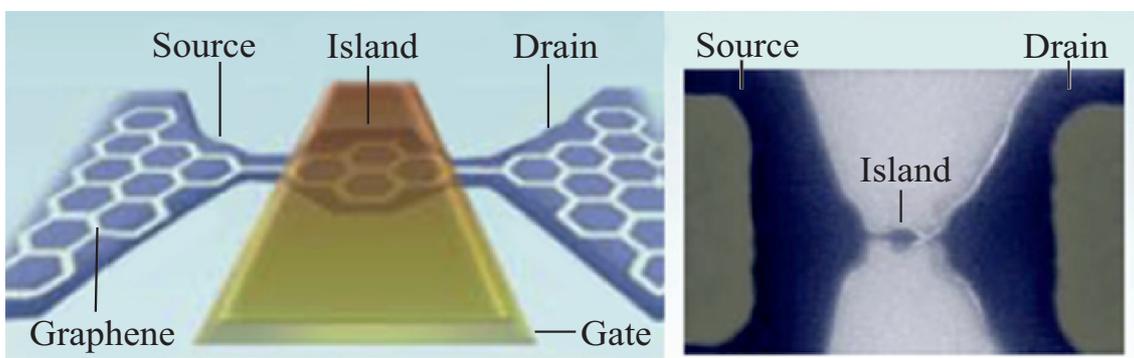


Рис. 8.4 – Первый графеновый полевой транзистор

Если к островку подходят новые электроны, то они отбрасываются электростатической силой. Любой электрон из истока квантово-механически туннелирует (есть отличная от нуля вероятность его прохождения через энергетический барьер) к островку, после чего «исчезает», просачиваясь к истоку. Напряжение, приложенное к третьему электроду – затвору (не показан на микроснимке), – управляет входом и выходом электрона с островка, таким образом регистрируя либо логический 0 (на островке нет электрона), либо 1 (электрон на островке).

8.4 Создание графеновой транзисторной технологии GNRFETs

В работе [114] при комнатной температуре исследованы лабораторные образцы полевых транзисторов с каналом из полоски (ленты) графена длиной 200 нм и различной ширины: 50, 30, 20, 10 и менее 10 нм. Исток и сток транзистора имели электроды из палладия Pd. Оказалось, что свойства графена удивительным образом меняются от чисто металлических, которые проявляются при больших значениях ширины полоски графена ω от 50 до 10 нм включительно, до полупроводниковых, когда ширина полоски графена становится менее 10 нм. При больших значениях ω графен имеет ничтожно малую ширину запрещенной зоны, а при суб-10-нм ширине полоски проявляется заметная ширина запрещенной зоны, что полезно для полевого транзистора, работающего при комнатной температуре и обладающего великолепной скоростью переключения и высокой подвижностью носителей заряда. Такое поведение графена предсказывалось теоретически в работах, выполненных в 2006–2007 гг. [114].

Таким образом, транзисторная структура имела вид, приведенный на рисунке 8.5, а [115]. На подложку из сильно легированного p^{++} -Si с высокой проводимостью, который служил тыловым (нижним) затвором (G), наносился слой подзатворного диэлектрика SiO₂ толщиной 300 нм. Исток (S) и сток (D) формировались металлическим палладием Pd. Было установлено, что при комнатной температуре отношение токов включения-выключения I_{on}/I_{off} увеличивалось экспоненциально при уменьшении ширины полоски графена: $I_{on}/I_{off} \sim 1, \sim 5,$

~ 100 , and $> 10^5$ для $\omega \sim 50$ нм, $\omega \sim 20$ нм, $\omega \sim 10$ нм, and $\omega \sim$ суб-10-нм соответственно. Важно, что все структуры с шириной $\omega \leq 10$ нм (а их было более 30 штук) проявляли значение $I_{on}/I_{off} > 10^5$, даже если смещение сток – исток V_{ds} достигало 1 В. Это означало, что такие структуры были полупроводниковыми и имели значительную ширину запрещенной зоны.

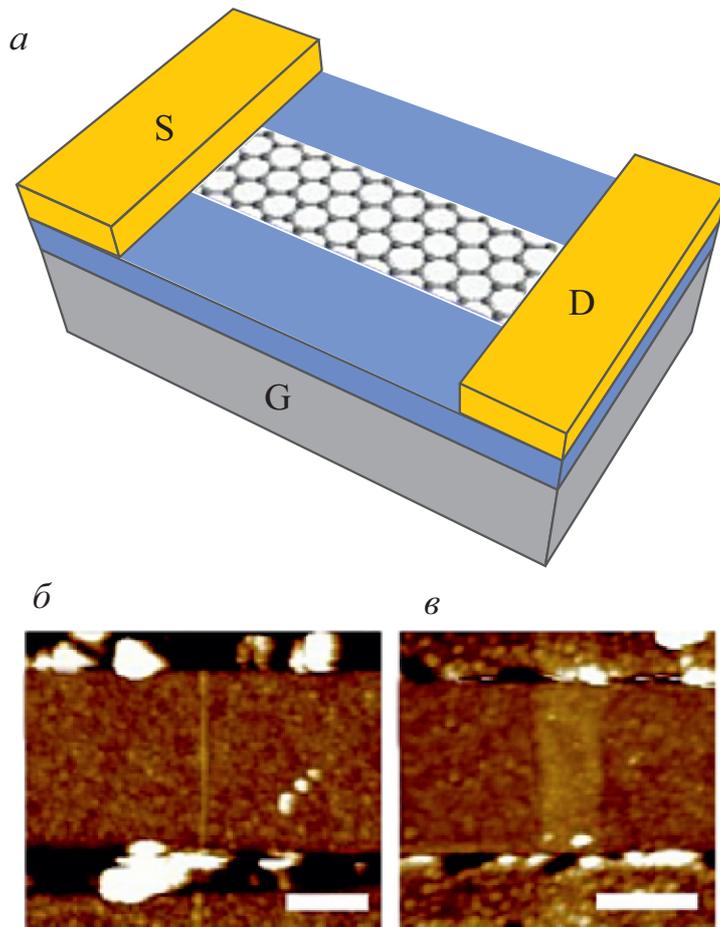
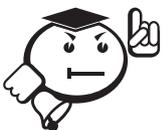


Рис. 8.5 – а) структура и внешний вид графенового полевого транзистора. Схема GNRFETs с использованием 10 нм SiO_2 и Pd стоком-исток, p^{++} -Si использовался как нижний затвор; б) AFM изображение структуры с шириной полосы графена $\omega \sim 2 \pm 0,5$ нм, длиной канала $L = 236$ нм. Метка измерительной линейки (белая прямоугольная полоска) – 100 нм; в) AFM изображение структуры с $\omega \sim 60 \pm 5$ нм, длиной канала $L = 190$ нм. Измерительная линейка – 100 нм



Такие графеновые структуры по-английски назвали *graphene nanoribbon field-effect transistors* (GNRFETs), а сами полосы графе-

на – *graphene nanoribbon* (GNR) (*ribbon* – англ., «полоска, лента»). Один из методов их изучения – атомно-силовая микроскопия, или *atomic force microscopy* (AFM). Для структур с $\omega \sim$ суб-10-нм из емкостных измерений была сделана оценка подвижности дырок 100–200 см²/Вс.

.....

В последующей работе [114] была отработана технология создания полевых транзисторов на основе графена в планарной конфигурации, исследована зависимость основных характеристик таких устройств при различной ширине затвора. При этом ширина полоски графена, как показано в подписи к рисунку 8.5, менялась в широких пределах от 2 до 60 нм. Авторы измерили подвижность и длину свободного пробега носителей заряда в транзисторе с $\omega \leq 10$ нм: подвижность оказалось 200 см²/Вс, длина свободного пробега 10 нм.

Оказалось, что полевые транзисторы GNRFETs обладают схожими характеристиками и способны в будущем составить конкуренцию транзисторам на углеродных нанотрубках. Успешные исследования в области построения такого рода устройств, а также понимание роли граней графеного слоя и применение *high-k* материалов позволят создать новое поколение полевых транзисторов с отличными характеристиками [114].

8.5 Высокоскоростные графеновые транзисторы

В работе [115] усовершенствовали технологию создания полевых транзисторов на основе графена в планарной конфигурации и исследовали зависимость основных характеристик таких устройств при различной ширине затвора. После измерений выходных характеристик устройства с шириной затвора 360 нм (зависимости тока стока от напряжения сток – исток V_{DS} , крутизны характеристики прямой передачи от V_{TG} и S -параметров от частоты) была построена зависимость коэффициента усиления по току для полевого транзистора, при этом критическая частота составила всего 4 ГГц. В подтверждение правильности полученных результатов ученые построили зависимость критической ча-

стоты от крутизны характеристики прямой передачи, которая оказалась линейной. Это характерно для полевых транзисторов. Измерения, проведенные на полевых графеновых транзисторах с различной шириной затвора, показали, что для ширины затвора $L_G = 150$ нм критическая частота составляет 26 ГГц, а сама зависимость f_T от ширины затвора представляет собой квадратичную гиперболу.

Ученые считают, что полученные результаты – не предел для электроники, основанной на графеновых транзисторах, и при соблюдении некоторых технических тонкостей (например, сохранении высокой подвижности носителей заряда при производстве таких транзисторов) можно достичь терагерцовых частот.

При этом из графена можно собирать микрочипы по обычной, отработанной годами планарной технологии, ныне применяемой при создании микросхем на основе кремния. А благодаря двумерной структуре графена, управляющий ток может быть легко увеличен за счет изменения ширины проводящего канала [116].



.....

В заключение отметим результаты работы [117], в которой испытан скоростной графеновый транзистор с рекордно высоким для транзисторов на базе графена быстродействием и расширенным диапазоном рабочих температур.

.....

Благодаря хорошей подвижности носителей заряда графен заслужил репутацию перспективного материала для электроники. Однако когда графеновый лист размещают на диэлектрической подложке (обычно на диоксиде кремния), подложка влияет на поведение одноатомного листа углерода, создает эффект рассеивания носителей заряда. Ученые из исследовательского центра Уотсона (IBM Thomas J. Watson Research Center) решили эту проблему. При помощи химического осаждения из пара они сначала создали на медной пленке лист графена, а затем перенесли его на подложку из алмазоподобного углерода,

в свою очередь расположенную на традиционной для электроники кремниевой пластине (рис. 8.6).

На этой базе исследователи построили полевой транзистор с затвором длиной всего 40 нанометров (рис. 4.6). На испытаниях прибор показал очень высокую частоту среза – 155 гигагерц.

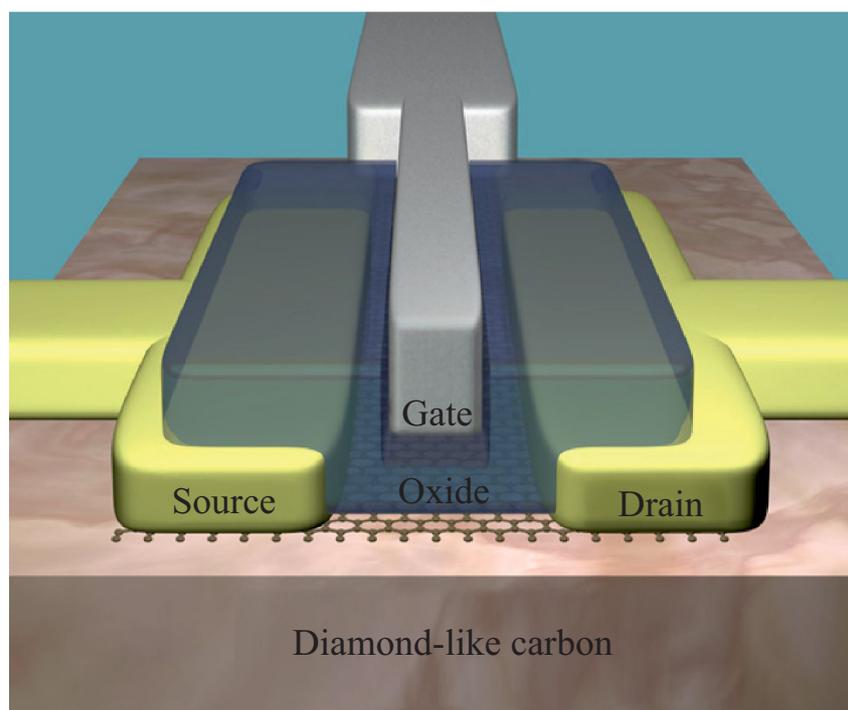


Рис. 8.6 – Принципиальная схема графенового транзистора на алмазоподобной подложке (фото: P. Avouris) [117]

Кроме того, авторы прибора проверили его функционирование при криогенных температурах, и выяснилось, что новый графеновый транзистор хорошо себя чувствует вплоть до температуры в 4,3 К. При этом отмечается, что качество использованного графена было не самым высоким, так что достигнутые показатели не предел для новой технологии.



Контрольные вопросы по главе 8

1. В каком году был получен графен? Опишите простой способ получения графена в лабораторных условиях. Что такое ВОПГ?
2. Какой графен является полупроводником?
3. Какие простые инструменты были использованы для изготовления первого графенового устройства в лаборатории?
4. Чем обусловлено в графеновой электронике повышение рабочих частот до терагерц?
5. Опишите технологию GNR-FETs и скоростной графеновый транзистор.

Заключение

Вы изучили весь материал настоящего учебного пособия. Во-первых, разобрались с историей развития электроники, переходом к микроэлектронике и, далее, к наноэлектронике – к электронике наноразмеров составляющих ее элементов: нанотранзисторов, нанодиодов и других. Стали понятны физические и технологические проблемы, возникающие на пути последовательного уменьшения элементов интегральных схем до наноразмеров. Во-вторых, Вы последовательно освоили основы квантовой физики и понимаете сущность квантов энергии и квантов вещества, смысл волновой функции, принципы запрета и неопределенностей и их предназначение в наномире. Это дает Вам, в-третьих, понимание квантового поведения электронов в кристалле, принципов квантового ограничения, движения электронов и дырок в потенциальных ямах и барьерах. Далее, на этой основе, в-четвертых, Вы хорошо понимаете сверхрешетки на базе гетероструктур и гетеропереходов и технологии их формирования, что дает Вам, в-пятых, основу для изучения и понимания всех основных квантовых эффектов в полупроводниках. И, наконец, пройденный материал позволяет Вам хорошо разобраться в устройстве и работе основных современных приборов наноэлектроники и материалов для них, в том числе в графеновой электронике.

Таким образом, предложенное Вам учебное пособие представляет собой целостный материал, все разделы которого логично взаимосвязаны и дают единую, цельную картину о наноэлектронике. При этом несомненной положительной чертой пособия является обширный список первоисточников, в основном представляющих собой оригинальные научные работы. Там, где это возможно, автор указал соответствующие ссылки на оригиналы электронных документов в сети Интернет, чтобы читатель сам мог изучить эти источники и глубже разобраться в материале. Это настоятельно рекомендуется делать, так как даже прикосновение к первоисточнику и его просмотр значительно упрощают восприятие самого сложного материала, делают его осязаемым и простым.

Выбранная Вами специальность чрезвычайно наукоемка, она основана на фундаментальных достижениях и физики, и химии, и материаловедения. Элементы нанотехнологий и наноэлектроники строятся на понимании фундаментальных, базисных законов природы и строения вещества на уровне отдельных атомов и межатомных расстояний. Осваивая профессию, Вы постигнете фундаментальные законы и устройство вещества так глубоко, что это позволит Вам легко ориентироваться в различных аспектах своей профессиональной деятельности.

Литература

1. Дробот П. Н. Промышленные технологии и инновации : учеб. пособие / Томский государственный университет систем управления и радиоэлектроники, Институт инноватики. – Томск : ТУСУР, 2012. – 145 с.
2. Lojek B. History of Semiconductor Engineering. – Berlin, Heidelberg : Springer, 2007. – 387 P.
3. Wilson A. H. The Theory of Electronic Semi-conductors // Proceedings Royal Society. – 1931. – Vol. 133. – P. 458–491 [Электронный ресурс]. – URL: <http://rspa.royalsocietypublishing.org/content/133/822/458.full.pdf> (дата обращения: 16.11.2015).
4. Давыдов Б. И. О выпрямляющем действии полупроводников // ЖЭТФ. – 1938. – Т. 8. – С. 3.
5. Давыдов Б. И. О выпрямлении тока на границе между двумя полупроводниками // ДАН СССР. – 1938. – Т. 20. – С. 279–282.
6. Давыдов Б. И. К теории твердых выпрямителей // ДАН СССР. – 1938. – Т. 20. – С. 283.
7. Mott N. F. Note on the contact between a metal and an insulator or semiconductor // Proceedings of the Cambridge Philosophical Society. – 1938. – Vol. 34, iss. 04. – P. 568.
8. Schottky W. Halbleitertheorie der Sperrschicht // Naturwissenschaften. – Vol. 26, iss. 52. – P. 843 [Electronic resource]. – URL: http://link.springer.com/chapter/10.1007/978-94-009-0657-0_4 (дата обращения: 26.04.2015).
9. Ohl R. S. Light-sensitive electric device. U. S. Patent № 2402662 (Filed May 27, 1941. Issued June 25, 1946) // Google Patent [Electronic resource]. – URL: <http://www.google.com/patents/US2402662> (дата обращения: 16.11.2015).

10. Scaff J. H., Ohl R. S. Development of silicon crystal rectifiers for microwave radar receivers // Bell System Technical Journal. – 1947. – Vol. 26. – P. 1–30 [Electronic resource]. – URL: <https://archive.org/details/bstj26-1-1> (дата обращения: 16.11.2015).
11. Лашкарев В. Е. Исследование запирающих слоев методом термозонда // Известия АН СССР. Сер. физ. – 1941. – Т. 5. – С. 442.
12. Seitz F. Research on Silicon and Germanium in World War II // Physics Today. – 1995. – P. 22–27 [Electronic resource]. – URL: http://www.physicstoday.org/resource/1/phtoad/v48/i1/p22_s1?isAuthorized=no (дата обращения: 16.11.2015).
13. Riordan M. The road to silicon was paved with germanium // Silicon Materials Science and Technology : Proceedings of the Eighth International Symposium on Silicon Materials Science and Technology «Semiconductor Silicon 1998» commemorating the 50th Anniversary of the Commercialization of the Point-Contact Transistor. – 1998. – Vol. 98-1. – P. 134–142 [Electronic resource]. – URL: http://books.google.ru/books?id=cr8FPGkiRS0C&dq=Research+on+Silicon+and+Germanium+in+World+War+II&hl=ru&source=gbs_navlinks_s (дата обращения: 16.11.2015).
14. Bray R. The Origin of Semiconductor Research at Purdue University. – Purdue University College of Science. Department of Physics [Electronic resource]. – URL: http://www.physics.purdue.edu/about_us/history/semiconductor_research.shtml (дата обращения: 16.11.2015).
15. Bray R., Gartenhaus S., Tubis A., Cassidy D. A History of Physics at Purdue: The Post-War Years (1945–1958). – Purdue University. College of Science. Department of Physics [Electronic resource]. – URL: http://tesla.physics.purdue.edu/about_us/history/post_war.shtml (дата обращения: 16.11.2015).
16. Torrey H. C., Whitmer C. A. Crystal Rectifiers // MIT Radiation Laboratory Series. – McGraw-Hill, New York, 1948. – Vol. 15 [Electronic re-

- source]. – URL: https://www.febo.com/pages/docs/RadLab/VOL_15_Crystal_Rectifiers.pdf (Имеется русский перевод: Торрей Г. С., Уитмер С. А. Кристаллические детекторы / пер. с англ. под ред. Е. Я. Пумпера. – Т. 1, 2. – М. : Советское радио, 1950. – 664 с.) (дата обращения: 16.11.2015).
17. Shockley W. The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors // Bell System Technical Journal. – 1949. – Vol. 28, iss. 3. – P. 435–489 [Electronic resource]. – URL: <https://archive.org/details/bellsystemtechni28amerrich> (дата обращения: 16.11.2015).
 18. Matare H. How Europe missed the transistor // IEEE Spectrum. – 2005 [Electronic resource]. – URL: www.pbs.org/transistor/materials/how-europe-missed-transistor.pdf (дата обращения: 16.11.2015).
 19. Riordan M., Hoddeson L. The Moses of Silicon Valley // Physics Today. – 1997. – P. 42–47 [Electronic resource]. – URL: <http://link.aip.org/link/phtoad/v50/i12/p42/s1> (дата обращения: 16.11.2015).
 20. Riordan M. From Bell labs to Silicon Valley: A saga of semiconductor technology transfer, 1955–61 [Electronic resource]. – URL : http://www.electrochem.org/dl/interface/fal/fal07/fall07_p36-41.pdf (дата обращения: 16.11.2015).
 21. Holonyak N. The Origins of Diffused-Silicon Technology at Bell Labs, 1954–55 [Electronic resource]. – URL: http://www.electrochem.org/dl/interface/fal/fal07/fall07_p30-34.pdf (дата обращения: 16.11.2015).
 22. Aschner J. F., Bittmann C. A., Hare W. F. J., Kleimack J. J. A Double Diffused Silicon High Frequency Switching Transistor Produced by Oxide Masking Techniques // Journal of the Electrochemical Society. – 1959. – Vol. 106, iss. 5. – P. 415–417 [Electronic resource]. – URL: <http://jes.ecsdl.org/content/106/5/415.full.pdf> (дата обращения: 16.11.2015).
 23. Tanenbaum M., Thomas D. E. Diffused emitter and base silicon transistors // Bell System Technical Journal. – 1956. – Vol. 35, № 1. – P. 1–15

- [Electronic resource]. – URL: <https://archive.org/details/bstj35-1-1> (дата обращения: 16.11.2015).
24. Moore G., Noyce R. Method for Fabricating Transistors. U. S. Patent 3108359 (Filed June 30, 1959. Issued October 29, 1963) [Electronic resource]. – URL: <http://google.com/patents/US3060123> (дата обращения: 16.11.2015).
 25. Dummer G. W. A. Electronic components in Great Britain // *Electrical Engineering*. – 1953. – Vol. 72, iss. 2. – P. 167–169 [Electronic resource]. – URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6438507&queryText%3DElectronic+components+in+Great+Britain> (дата обращения: 16.11.2015)
 26. Килби Дж. С. Возможное становится реальным: изобретение интегральных схем. Нобелевская лекция. Стокгольм, 8 декабря 2000 г. // *УФН*. – 2002. – Т. 172. – С. 1102–1109.
 27. Kilby J. S. Miniaturized Electronic Circuits. U. S. Patent 3138743 (Filed February 6, 1959. Issued June 23, 1964) [Electronic resource]. – URL: <http://google.com/patents/US3138743> (дата обращения: 27.04.2015).
 28. Hoerni J. A. Planar Silicon Diodes and Transistors // *Electron Devices Meeting*, 1960, Washington, D. C. – October 1960 reprinted as Fairchild Semiconductor Technical Paper TP-14. – 1961. – P. 1–9 [Electronic resource]. – URL: <http://corphist.computerhistory.org/corphist/documents/doc-472a58a7b0f46.pdf> (дата обращения: 16.11.2015).
 29. Hoerni J. A. Method of Manufacturing Semiconductor Devices. U. S. Patent 3025589 (Filed May 1, 1959. Issued March 20, 1962) [Electronic resource]. – URL: <http://google.com/patents/US3025589> (дата обращения: 16.11.2015).
 30. Hoerni J. A. Semiconductor Device. U. S. Patent 3064167 (Filed May 1, 1959. Issued May 19, 1960) [Electronic resource]. – URL: <http://google.com/patents/US3064167> (дата обращения: 16.11.2015).

31. Riordan M. The Silicon Dioxide Solution: How physicist Jean Hoerni built the bridge from the transistor to the integrated circuit // IEEE Spectrum. – 2007. – P. 44–50 [Electronic resource]. – URL: <http://spectrum.ieee.org/semiconductors/design/the-silicon-dioxide-solution> (дата обращения: 16.11.2015).
32. Noyce R. Semiconductor device-and-lead structure. U. S. Patent 2981877 (Filed July 30, 1959. Issued April 25, 1961) [Electronic resource]. – URL: <http://google.com/patents/US2981877> (дата обращения: 16.11.2015).
33. Lehovec Kurt. Multiple Semiconductor Assembly. U. S. Patent 3029366 (Filed April 22, 1959. Issued April 10, 1962) [Electronic resource]. – URL: <http://google.com/patents/US3029366> (дата обращения: 16.11.2015).
34. Kahng D. Electric Field Controlled Semiconductor Device. U. S. Patent No. 3102230 (Filed 31 May 31, 1960, issued August 27, 1963) [Electronic resource]. – URL: <http://google.com/patents/US3102230> (дата обращения: 16.11.2015).
35. Atalla M. M., Tannenbaum E., Scheibner E. J. Stabilization of Silicon Surfaces by Thermally Grown Oxides // Bell System Technical Journal. – 1959. – Vol. 38, iss. 3. – P. 749–783 [Electronic resource]. – URL: <https://archive.org/details/bstj38-3-749> (дата обращения: 16.11.2015).
36. Sah R. L-Y. Evolution of the MOS transistor-from conception to VLSI // Proceedings of the IEEE. – Vol. 76, iss. 10. – P. 1280–1326 [Electronic resource]. – URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=16328&queryText%3DEvolution+of+the+MOS+Transistor> (дата обращения: 16.11.2015).
37. Hofstein S. R., Heiman F. P. The silicon insulated-gate field-effect transistor // Proceedings of the IEEE. – Vol. 51, iss. 9. – P. 1190–1202 [Electronic resource]. – URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=1444418&queryText%3DThe+Silicon+Insulated+Gate+Field+Effect+Transistor> (дата обращения: 16.11.2015).

38. Moore E. Cramming More Components onto Integrated Circuits // Electronics. – 1965. – P. 114–117 [Electronic resource]. – URL: http://web.eng.fiu.edu/npala/EEE5425/Gordon_Moore_1965_Article.pdf (дата обращения: 16.11.2015).
39. Christensen H., Teal G. Method of Fabricating Germanium Bodies. U. S. Patent No. 2692839 (Filed 7 April 1951. Issued 26 October 1954) [Electronic resource]. – URL: <http://google.com/patents/US2692839> (дата обращения: 16.11.2015).
40. Кокориш Н. П., Красилов А. В., Шефталь Н. Н. Способ получения монокристаллов германия и кремния с заданным содержанием примесей. Описание изобретения к авторскому свидетельству, 1955 г. // База патентов СССР [Электронный ресурс]. – URL: <http://patentdb.su/1-107450-sposob-polucheniya-monokristallov-germaniya-i-kremniya-s-zadannym-soderzhaniem-primesejj.html> (дата обращения: 16.11.2015).
41. Hall Eldon C. A Case History of the AGC Integrated Logic Circuits. Apollo Guidance and Navigation E-1880, MIT Instrumentation Laboratory (December 1965) // NASA Office of Logic Design. A scientific study of the problems of digital engineering for space flight systems, with a view to their practical solution [Electronic resource]. – URL: http://klabs.org/history/history_docs/mit_docs/1716.pdf (дата обращения: 16.11.2015).
42. Hall Eldon C. Reliability History of the Apollo Guidance Computer // NASA Office of Logic Design. A scientific study of the problems of digital engineering for space flight systems, with a view to their practical solution [Electronic resource]. – URL: http://klabs.org/history/history_docs/mit_docs/1698.pdf (дата обращения: 16.11.2015).
43. Безбородов А. Б. История СССР/РФ в контексте современного Руссиеведения / И. А. Агакишиев, А. Н. Бачинин, А. Б. Безбородов и др.;

- отв. ред. Е. И. Пивовар, А. Б. Безбородов ; Министерство образования и науки РФ, РГГУ. – М. : Проспект, 2011. – 400 с.
44. Чертопруд С. В. Научно-техническая разведка от Ленина до Горбачева. – М. : ОЛМА-ПРЕСС, 2002. – 447 с.
45. Малин Б. В. Создание первой отечественной микросхемы // Виртуальный компьютерный музей [Электронный ресурс]. – URL: http://www.computer-museum.ru/technlgy/su_chip.htm (дата обращения: 16.11.2015).
46. Малашевич Б. М. Зеленоградский Центр микроэлектроники: Создание, расцвет, закат... // Виртуальный компьютерный музей [Электронный ресурс]. – URL: http://www.computer-museum.ru/technlgy/su_chip.htm (дата обращения: 16.11.2015).
47. Cho A. Y. Growth of Periodic Structures by the Molecular Beam Method // Appl. Phys. Lett. – 1971. – Vol. 19. – P. 467 [Electronic resource]. – URL: <http://scitation.aip.org/content/aip/journal/apl/19/11/10.1063/1.1653775> (дата обращения: 16.11.2015).
48. Прошло 50 лет, а закон Мура продолжает удваивать ставку // Новости высоких технологий. 21.04.2015 [Электронный ресурс]. – URL: <http://hi-news.ru/hardware/proshlo-50-let-a-zakon-mura-prodolzhaet-udvaivat-stavku.html> (дата обращения: 16.11.2015).
49. Выполнение закона Мура на протяжении 50 лет стало сюрпризом для его автора // Наука и техника. Lenta-Ru. 25.03.2015 [Электронный ресурс]. – URL: <http://lenta.ru/news/2015/03/25/50moore/> (дата обращения: 16.11.2015)
50. Кузнецов Ф. А. Новые материалы и технологии для приборов нанoeлектроники // Учреждение Российской академии наук Институт неорганической химии им. А. В. Николаева СО РАН [Электронный ресурс]. – URL: <http://www.abercade.ru/research/analysis/941.html> (дата обращения: 16.11.2015).

51. Закон Мура: каким путем пойдет дальнейшее развитие полупроводников [Электронный ресурс]. – URL: <http://www.nanonewsnet.ru/articles/2009/zakon-mura-kakim-putem-poidet-dalneishee-razvitie-poluprovodnikov> (дата обращения: 16.11.2015).
52. Макс Планк: Избранные труды. Термодинамика. Теория излучения и квантовая теория. Теория относительности. Статьи и речи. – М. : Наука, 1975 [Электронный ресурс]. – URL: <http://eqworld.ipmnet.ru/ru/library/books/Plank1975ru.djvu> (дата обращения: 16.11.2015).
53. Шпольский Э. В. Творческий путь М. Планка // УФН. – 1958. – Т. 66, вып. 3. – С. 535–542 [Электронный ресурс]. – URL: http://ufn.ru/ufn58/ufn58_11/Russian/r5811g.pdf (дата обращения: 16.11.2015).
54. Becquerel E. Memoires sur les effets electriques produits sous l'influence des rayons // Comptes Rendues. – 1839. – Vol. 9. – P. 561–567.
55. Einstein A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt // Annalen der Physik. – 1905. – Bd. 17, hbd. 1. – S. 132–148. (Имеется русский перевод: Эйнштейн А. Об одной эвристической точке зрения, касающейся возникновения и превращения света // Эйнштейн А. Собрание научных трудов. – М. : Наука, 1966. – Т. 3. – 632 с. –с. 93).
56. Эйнштейн А. Об одной эвристической точке зрения, касающейся возникновения и превращения света // Эйнштейн А. Собрание научных трудов. – М. : Наука, 1966. – Т. 3. – 632 с.
57. Бор Н. Избранные научные труды : в 2 т. ; под ред. Е. И. Тамма, В. А. Фока, Б. Г. Кузнецова ; сост. У. И. Франкфурт. – М. : Наука, 1971. – Сер. «Классики науки». – Т. 1.– С. 84–148.
58. Louis de Broglie. Ondes et quanta // Comptes Rendues. – 1923. – Vol. 177. – P. 507–510. (Имеется русский перевод Де-Бройль Л. Волны и кванты // УФН. – 1967. – Т. 93. – С. 178–180.).
59. Дэвиссон К. Дж. Волны ли электроны? //УФН. – 1928. – Т. 8, вып. 4. – С. 483–509.

60. Мычко Д. И., Рябцев А. Н. Графическое изображение атомных орбиталей // *Хімія: праблемы выкладання*. – 2011. – № 6. – С. 25–29 [Электронный ресурс]. – URL: <http://elib.bsu.by/handle/123456789/26818> (дата обращения: 16.11.2015).
61. Гейзенберг В., Шредингер Э., Дирак П. А. М. Современная квантовая механика. Три нобелевских доклада. Л. ; М. : Гостехиздат, 1934 [Электронный ресурс]. – URL: <http://eqworld.ipmnet.ru/ru/library/books/GeizenbergShredingerDirak1934ru.djvu> (дата обращения: 16.11.2015).
62. Лауэ М. О соотношениях неточностей Гейзенберга и их теоретико-познавательном значении // *УФН*. – 1935. – Т. 15, вып. 3. – С. 343–346.
63. Денисенко М. В. и др. Вычисление блоховских функций электрона в одномерном периодическом потенциале / М. В. Денисенко, А. С. Деробенко, С. М. Кашин, А. М. Сатанин. – Н. Новгород : Изд-во ННГУ, 2010. – 33 с.
64. Bloch F. Über die Quantenmechanik der Elektronen in Kristallgittern // *Zeitschrift für Physik*. – 1928. – Vol. 52. – P. 555.
65. Kronig R. de L., Penney W. G. Quantum mechanics of electrons in crystal lattices // *Proceedings of the Royal Society of London*. – Vol. 130A. – P. 499–513.
66. Борисенко В. Е. Нанoeлектроника / В. Е. Борисенко, А. И. Воробьева, Е. А. Уткина. – 2-е изд., перераб. и доп. – М. : БИНОМ. Лаборатория знаний, 2013. – 366 с.
67. Hartman T. E. Tunneling of a wave packet // *Journal of Applied Physics*. – 1962. – Vol. 33, iss. 12. – P. 3427–3433.
68. Алфёров Ж. И. История и будущее полупроводниковых гетероструктур // *ФТП*. – 1998. – Т. 32, № 1. – С. 3–18. [Электронный ресурс]. – URL: <http://journals.ioffe.ru/ftp/1998/01/p3-18.pdf> (дата обращения: 25.05.2015).

69. Alferov Zh. I. Double Heterostructure Concept and its Applications in Physics, Electronics and Technology // Nobel Lectures, Physics 1996–2000 ; Ed. Gösta Ekspong. – Singapore : World Scientific Publishing Co., 2002. – P. 413–441 [Electronic resource]. – URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2000/alferov-lecture.pdf (Имеется русский перевод. Алфёров Ж. И. Двойные гетероструктуры: концепция и применения в физике, электронике и технологии // УФН. – 2002. – Т. 172, № 9. – С. 1068–1086. [Электронный ресурс]. – URL: http://ufn.ru/ufn02/ufn02_9/Russian/r029d.pdf (дата обращения: 16.11.2015).
70. Chen An-Ben, Sher A. Semiconductor Alloys: Physics and Materials Engineering. – New York : Plenum Press, 1995 [Electronic resource]. – URL: <http://link.springer.com/book/10.1007/978-1-4613-0317-6> (дата обращения: 16.11.2015).
71. Gawlik K.-U., Kipp L., Skibowski M., Orlowski N., Manzke R. HgSe: Metal or Semiconductor? // Phys. Rev. Lett. – 1997. – Vol. 78, № 16. – P. 3165–3168 [Electronic resource]. – URL: <http://journals.aps.org/prl/pdf/10.1103/PhysRevLett.78.3165> (дата обращения: 16.11.2015).
72. Dietl T., Dobrowolski W., Kossut J., Kowalski B. J., Szuszkiewicz W., Wilamowski Z. Comment on «HgSe: Metal or Semiconductor?» [Electronic resource]. – URL: <http://journals.aps.org/prl/pdf/10.1103/PhysRevLett.81.1535> (дата обращения: 16.11.2015).
73. Eich D., Hübner D., Fink R., Umbach E., Ortner K., Becker C. R., Landwehr G., Fleszar A. Electronic structure of HgSe(001) investigated by direct and inverse photoemission // Phys. Rev. B. – 2000. – Vol. 61, № 19. – P. 12666–12669 [Electronic resource]. – URL: <http://journals.aps.org/prb/pdf/10.1103/PhysRevB.61.12666> (дата обращения: 01.06.2015).
74. Vegard L. Die Konstitution der Mischkristalle und die Raumbfüllung der Atome // Zeitschrift für Physik. – 1921. – Vol. 5, iss. 1. – S. 17–26.

75. Шиляев П. А. Полупроводниковые гетероструктуры: гетеропереход / сост. П. А. Шиляев, Д. А. Павлов. – Н. Новгород : Нижегородский госуниверситет, 2009. – 18 с.
76. Усанов Д. А. и др. Компьютерное моделирование наноструктур. – Саратов, 2013. – 100 с.
77. Лозовский В. Н., Константинова Г. С., Лозовский С. В. Нанотехнология в электронике. – СПб. : Лань, 2008. – 336 с.
78. Шик А. Я., Бакуева Л. Г., Мусихин С. Ф., Рыков С. А. Физика низкоразмерных структур. – СПб. : Наука, 2001. – 160 с.
79. Klitzing K. V., Dorda G., Pepper M. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance // *Physical Review Letters*. – 1980. – Vol. 45. – P. 494–497 [Electronic resource]. – URL: http://www.physics.oregonstate.edu/~ostroveo/COURSES/ph673/Notes/PRL1980_finestructure_quantumHall.pdf (дата обращения: 16.11.2015)
80. КЛИТЦИНГ К. Квантованный эффект Холла // *УФН*. – 1986. – Т. 150, вып. 1. – С. 107–126 [Электронный ресурс]. – URL: http://ufn.ru/ufn86/ufn86_9/Russian/r869c.pdf (дата обращения: 16.11.2015).
81. Tsui D. C., Störmer H. L., Gossard A. C. Two-Dimensional Magnetotransport in the Extreme Quantum Limit // *Phys. Rev. Lett.* – 1982. – Vol. 48. – P. 1559–1562 [Electronic resource]. – URL: <http://journals.aps.org/prl/pdf/10.1103/PhysRevLett.48.1559> (дата обращения: 16.11.2015).
82. Aharonov Y., Bohm D. Significance of Electromagnetic Potentials in the Quantum Theory // *Phys. Rev.* – 1959. – Vol. 115. – P.485–491.
83. Делоне Н. Б., Крайнов В. П. Динамический штарковский сдвиг атомных уровней // *УФН*. – 1999. – Т. 169, № 7. – С. 753–772.
84. Горшков А. П., Карпович И. А., Кудрин А. В. Исследование эффекта Штарка в гетеронаноструктурах с квантовыми точками и ямами In(Ga)As/GaAs методом фотоэлектрической спектроскопии // *Поверхность*. – 2006. – № 5. – С. 25–29.

85. Карпович И. А. Фотоэлектрическая спектроскопия квантово-размерных гетеронаноструктур InGaAs/GaAs, выращенных газофазной эпитаксией // Вестник нижегородского университета им. Н. И. Лобачевского. – 2010. – № 5-2. – С. 233–242.
86. Волкова Н. С. Влияние поверхностного дефектообразования на фотоэлектрические спектры квантово-размерных гетеронаноструктур In(Ga)As/GaAs, выращенных газофазной эпитаксией // Вестник нижегородского университета им. Н. И. Лобачевского. – 2012. – № 2(1). – С. 34–38.
87. Лангенберг Д. Н., Скалапино Д. Дж., Тейлор Б. Н. Эффекты Джозефсона // УФН. – 1967. – Т. 91, вып. 2. – С. 317–330.
88. Кулик И. О., Шехтер Р. И. Кинетические явления и эффекты дискретности заряда в гранулированных средах // ЖЭТФ. – 1975. – Т. 68, вып. 2. – С. 623–640.
89. Likharev K. K., Zorin A. B. Theory of Bloch-wave oscillations in small Josephson junctions // Journal of Low Temperature Physics. – 1985. – Vol. 59 (3/4). – P. 347–382.
90. Averin D. V., Likharev K. K. Coulomb Blockade of tunneling and coherent oscillations in small tunnel junctions // Journal of Low Temperature Physics. – 1986. – Vol. 62(2). – P. 345–372.
91. Добровицкий В. В., Звездин А. К., Попков А. Ф. Гигантское магнетосопротивление, спин-переориентационные переходы и макроскопические квантовые явления в магнитных наноструктурах // УФН. – 1996. – Т. 166, вып. 4. – С. 439–447.
92. Абрикосов А. А. Некоторые свойства металлов с магнитными примесями (эффект Кондо) // УФН. – 1972. – Т. 107, вып. 5. – С. 160–161.
93. Абрамов И. И., Новик Е. Г. Классификация приборных структур одноэлектроники // ФТП. – 1999. – Т. 33, вып. 11. – С. 1388–1394.
94. Силин А. П. Полупроводниковые сверхрешетки // УФН. – 1985. – Т. 147, вып. 3. – С. 485–521.

95. Akeyoshi T., Meazawa K., Mitzutani T. Weighted sum threshold logic operation of MOBILE (monostable-bistable transition logic element) using resonant-tunneling transistors // IEEE Electron Device Letters. – 1993. – Vol. 14, iss. 10. – P. 475–477.
96. Алфёров Ж. И., Казаринов Р. Ф. А. с. № 181737, заявка № 950840 с присоединением заявки с приоритетом от 30 марта 1963 г.
97. Двуреченский А., Придачин Н. После кремния будет кремний // Наука в Сибири. – 2000. – 10 марта.
98. Radisavljevic B., Radenovic A., Brivio J., Giacometti V., Kis A. Single-layer MoS₂ transistors // Nature Nanotechnology. – 2011. – Vol. 6. – P. 147–150.
99. MoS₂ Integrated Circuits // École polytechnique fédérale de Lausanne (EPFL). Laboratory of Nanoscale Electronics and Structures (LANES) [Electronic resource]. – URL: <http://lanes.epfl.ch/cms/site/lanes/lang/en/mos2ic-news> (дата обращения: 16.11.2015).
100. Chris J. H. Wort, Balmer R. S. Diamond as an electronic material // Materials Today. – 2008. – Vol. 11, iss.1–2. – P. 22–28.
101. Schwitiers M. et al. Diamond-MESFETs – synthesis and integration // Radar Conference (6–7 Oct. 2005). – Paris : EURAD, 2005. – P. 1–4 // IEEE Xplore Digital Library [Electronic resource]. – URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1605549> (дата обращения: 16.11.2015).
102. Tuttle G., Kroemer H. An AlSb/InAs/AlSb quantum well HFT // IEEE Xplore Digital Library [Electronic resource]. – URL: <http://ieeexplore.ieee.org/xpl/abstractAuthors.jsp?arnumber=1486963> (дата обращения: 16.11.2015).
103. Iijima S. Helical microtubules of graphitic carbon // Nature. – 1991. – Vol. 354. – P. 56–58.
104. Елецкий А. В. Углеродные нанотрубки // УФН. – 1997. – Т. 167, № 9. – С. 945–972.

105. Швец П. Транзистор на УНТ: лабораторная реальность // Техника – молодежи. – 2010. – № 5. – С. 22.
106. Franklin A. D., Mathieu Luisier et al. Sub-10-nm Carbon Nanotube Transistor // Nano Lett. – 2012. – № 12 (2). – P. 758–762.
107. Novoselov K. S., Geim A. K., Morozov S. V., Jiang D., Zhang Y., Dubonos S. V., Grigorieva I. V., Firsov A. A. Electric Field Effect in Atomically Thin Carbon Films // Science. – 2004. – Vol. 306. – P. 666–669.
108. Гейм А. К. Случайные блуждания: непредсказуемый путь к графену. Нобелевская лекция. Стокгольм, 8 декабря 2010 г. // УФН. – Т. 181, № 12. – С. 1284–1298.
109. Zhihong Chen, Yu-Ming Lin, Michael J. Rooks, Phaedon Avouris. Graphene nano-ribbon electronics // Physica E: Low-dimensional Systems and Nanostructures. – 2010. – Vol. 40, iss. 2. – P. 228–232.
110. Reina Alfonso, Jia Xiaoting, Ho John, Nezich Daniel, Son Hyungbin, Bulovic Vladimir, Dresselhaus Mildred S., Kong Jing. Large Area, Few-Layer Graphene Films on Arbitrary Substrates by Chemical Vapor Deposition // Nano Letters. – 2009. – Vol. 9 (1). – P. 30–35.
111. Hyesung Park, Jill A. Rowehl, Ki Kang Kim, Vladimir Bulovic, Jing Kong. Doped graphene electrodes for organic solar cells // Nanotechnology. – 2010. – Vol. 21. – P. 1–6.
112. Li X., Cai W., An J., Kim S., Nah J., Yang D., Piner R., Velamakanni A., Jung I., Tutuc E., Banerjee S. K., Colombo L., Ruoff R. S. Large-Area Synthesis of High-Quality and Uniform Graphene Films on Copper Foils // Science. – 2009. – Vol. 324, № 5932. – P. 1312–1314.
113. Suzanne Deffree. University produces 100-mm graphene wafers // Electronics Design News. – 2010. – February 3 [Electronic resource]. – URL: <http://www.edn.com/electronics-news/4312123/100-mm-graphene-wafers-target-next-generation-electronics> (дата обращения: 23.06.2015).
114. Li H. et al. Chemically Derived, Ultrasoother Graphene Nanoribbon Semiconductors // Science. – Vol. 319, № 5867. – P. 1229–1232.

115. Xinran Wang, Yijian Ouyang, Xiaolin Li, Hailiang Wang, Jing Guo, Hongjie Dai. Room-Temperature All-Semiconducting Sub-10-nm Graphene Nanoribbon Field-Effect Transistors // *Phys. Rev. Lett.* – 2008. – Vol. 100, iss. 20. – P. 206803.
116. Yu-Ming Lin, Keith A. Jenkins, Alberto Valdes-Garcia, Joshua P. Small, Damon B. Farmer, Phaedon Avouris. Operation of Graphene Transistors at Gigahertz Frequencies // *Nano Lett.* – 2009. – Vol. 9 (1). – P. 422–426.
117. Yanqing Wu, Yu-ming Lin, Ageeth A. Bol, Keith A. Jenkins, Fengnian Xia, Damon B. Farmer, Yu Zhu & Phaedon Avouris. High-frequency, scaled graphene transistors on diamond-like carbon // *Nature.* – 2011. – Vol. 472. – P. 74–78.

Глоссарий

Ааронова – Бома эффект (Aharonov – Bohm effect) – интерференция электронных волн на выходе кольцевого проводника, на который наложено магнитное поле, проходящее через центр кольца перпендикулярно его плоскости.

Баллистический транспорт (ballistic transport) носителей заряда – перенос носителей заряда без их рассеяния.

Боровский радиус (Bohr radius) – радиус самой низкой орбиты электрона массой m в атоме водорода.

Вегарда правило (Vegard's law) – постоянная решетки промежуточного соединения $a(x)$ линейно изменяется в интервале между постоянными решеток образующих их материалов a_1 и a_2 : $a(x) = xa_1 + (1-x)a_2$, где x – мольная доля материала 1 в материале 2.

Гейзенберга соотношение неопределенностей (Heisenberg's uncertainty principle) утверждает, что невозможно одновременно определить с произвольной точностью импульс (p) и координаты (r) квантовой частицы. Количественно это выражается как $\Delta p \Delta r \geq h / 2\pi$ или, эквивалентно для времени (t) и энергии частицы (E), $\Delta E \Delta t \geq h / 2\pi$, где h – постоянная Планка.

Гетероструктура (heterostructure) – структура, состоящая из нескольких (более одного) материалов, отличающихся по своим свойствам.

Гигантского магнитосопротивления эффект (giant magnetoresistance effect) – существенное изменение электрического сопротивления многослойной тонкопленочной структуры, составленной из чередующихся слоев немагнитного и магнитного проводящих материалов, при изменении направлений намагниченности магнитных слоев с параллельных на антипараллельные.

Графен (graphene) – лист из одного слоя графита *моноатомной толщины*.

Двумерный электронный газ (two-dimensional electron gas, 2DEG) – электроны в квантовых пленках.

Де Бройля волна (de Broglie wavelength) – волна, соответствующая квантовой частице с импульсом p и характеризуемая длиной $\lambda = h/p$, где h – постоянная Планка.

Джозефсона эффекты (Josephson effects) возникают при туннелировании электронов из сверхпроводника через диэлектрик в другой сверхпроводник.

Джозефсона эффекты на переменном токе (alternating current Josephson effect): если к туннельному переходу сверхпроводник/диэлектрик/сверхпроводник приложено напряжение V , через него протекает переменный ток с частотой $2eV/h$, где e – заряд электрона, h – постоянная Планка.

Джозефсона эффекты на постоянном токе (direct current Josephson effect): через туннельный переход между сверхпроводниками протекает суперток при нулевом напряжении на контактах перехода.

Диффузионный транспорт (diffusion transport) носителей заряда – перенос носителей заряда путем их диффузии.

Диэлектрик (dielectric) – вещество, в котором внешнее электрическое поле индуцирует заряды, которые достаточно долго сохраняют в нем электрическое поле даже после снятия внешнего поля.

Интерференция (interference) – взаимодействие двух или более одинаковых или близких по частоте волн, приводящее вследствие их алгебраического или векторного сложения к изменению с расстоянием или временем амплитуды результирующей волны.

Квантовая пленка (quantum film) – двумерная (2D) структура, в которой квантовое ограничение действует только в одном направлении – по толщине пленки.

Квантовая точка (quantum dot) – нульмерная (0D) структура, в которой квантовое ограничение действует во всех трех направлениях.

Квантового сопротивления единица (quantum resistivity unit) – $h/e^2 = 25\,812,807$ Ом. Соответствующая проводимость $e^2/h = 38,740$ мкСм.

Квантовое ограничение (quantum confinement) – специфическое поведение электронов (дырок) в замкнутом пространстве, возникающие как следствие

их квантово-волновой природы и ограниченных возможностей для движения, приводящих к конечному (ненулевому) минимальному значению их энергии и дискретности разрешенных энергетических состояний в низкоразмерной структуре.

Квантовой проводимости единица (quantum conductance unit) – $e^2/h=38,740$ мкСм. Соответствующее сопротивление $h/e^2=25\,812,807$ Ом.

Квантовый колодец (quantum well) – ограниченная потенциальными барьерами конечная область пространства или материала, в которой проявляются квантовые эффекты.

Квантовый шнур (quantum wire) – одномерная (1D) структура, в которой квантовое ограничение действует в двух направлениях – по толщине и ширине шнура.

Кондо температура (Kondo temperature) – температура, начиная с которой электрическое сопротивление металлов с магнитными примесями начинает возрастать при понижении температуры.

Кондо эффект (Kondo effect) – увеличение электрического сопротивления металлов, содержащих магнитные примеси, при температурах, приближающихся к абсолютному нулю; связано с ненулевым полным спином всех электронов в образце.

Кулоновская блокада (Coulomb blockade) – блокирование движения заряженных частиц вследствие их кулоновского взаимодействия с другими подвижными и неподвижными зарядами.

Кулоновская лестница (Coulomb staircase) – вольт-амперная характеристика ступенчатого вида, типичная для двухбарьерной несимметричной одноэлектронной туннельной структуры.

Кулоновские алмазы (Coulomb diamonds) – ромбообразные области электростатической устойчивости одноэлектронного транзистора на диаграмме, описывающей соотношение между напряжением на стоке и напряжением на затворе этого транзистора. В них число электронов в квантовой точке определяется только напряжением на затворе и емкостями затвора и туннельных барье-

ров. При устойчивом состоянии транзистора вследствие кулоновской блокады ток через него не течет.

Куперовская пара (Cooper pair) – пара, образованная двумя электронами с противоположной ориентацией спинов. Результирующий спин таких пар равен нулю, и их распределение по энергиям подчиняется статистике Бозе – Эйнштейна.

Лазер (laser) – аббревиатура от английского выражения *light amplification by stimulated emission of radiation*. Это источник излучения оптического диапазона, в основе работы которого лежит усиление света стимулированной эмиссией излучения.

Ландау уровни (Landau levels) – разрешенные энергетические состояния для электронов, движущихся перпендикулярно магнитному полю и вынужденных изменять траектории движения на круговые орбиты, следуя силе Лоренца. В этом случае электроны выполняют циклотронное движение с угловой частотой $\omega_c = eB/m$, называемой циклотронной частотой, где B – индукция магнитного поля, e – заряд, а m – масса электрона. Это приводит к тому, что разрешенные энергетические состояния для таких электронов становятся квантованными.

Магнитосопротивление (magnetoresistance) – изменение электрического сопротивления материала или структуры, обусловленное воздействием на них внешнего магнитного поля. Величина магнитосопротивления выражается как отношение $\Delta R/R_0$ в процентах, где ΔR – изменение сопротивления, а R_0 – сопротивление при нулевом магнитном поле.

Мазер (maser) – аббревиатура от английского выражения *microwave amplification by stimulated emission of radiation*. Это источник излучения микроволнового диапазона, в основе работы которого лежит усиление излучения за счет стимулированной эмиссии.

Металлоорганика (metal-organics) – органические соединения, в состав которых входят атомы металлов, образующих химические связи металл – угле-

род и металл – кислород – углерод, а также координационные связи между металлами и органическими молекулами.

Нано- (*nano-*) – десятичная приставка, обозначающая 10^{-9} следующей за ней единицы измерения, сокращается – н. Например, 1 нм = 10^{-9} м.

Нанодиагностика – совокупность специализированных методов, исследований, направленных на изучение структурных, морфолого-топологических, механических, электрофизических, оптических, биологических характеристик наноматериалов и наносистем, анализ нанокolicеств вещества, измерение метрических параметров с наноточностью.

Наноматериалы – вещества (и композиции веществ), представляющие собой искусственно или естественно упорядоченную (или неупорядоченную) систему базовых элементов с нанометрическими характеристическими размерами и особым проявлением физического и (или) химического взаимодействий при кооперации наноразмерных элементов, обеспечивающих возникновение у материалов и систем совокупности ранее неизвестных механических, химических, электрофизических, оптических, теплофизических и других свойств, определяемых проявлением наномасштабных факторов.

Нанонаука – система знаний, основанная на описании, объяснении и предсказании свойств материальных объектов с нанометрическими характеристическими размерами или систем более высокого метрического уровня, упорядоченных или самоупорядоченных на основе наноразмерных элементов.

Нанопечать (*nanoimprinting*) – литографическая технология создания наноразмерных элементов интегральных микросхем методом печати. Различают чернильную нанопечать (*inking*) и нанопечать тиснением (*embossing*).

Наносистемотехника – совокупность методов моделирования, проектирования и конструирования изделий различного функционального назначения, в том числе наноматериалов, микро- и наносистем с широким использованием квантово-размерных, кооперативно-синергетических, гигантских эффектов и других явлений и процессов, проявляющихся в условиях материальных объектов с нанометрическими характеристическими размерами элементов.

Наноструктура (nanostructure) – ансамбль связанных атомов, имеющий по крайней мере в одном направлении размер от одного до сотен нанометров.

Нанотехника – машины, механизмы, приборы, устройства, материалы, созданные с использованием новых свойств и функциональных возможностей систем при переходе к наномасштабам и обладающие ранее недостижимыми массогабаритными и энергетическими показателями, технико-экономическими параметрами и функциональными возможностями.

Нанотехнология (nanotechnology) – совокупность методов и способов синтеза сборки, структуро- и формообразования, нанесения, удаления и модифицирования материалов, включая систему знаний, навыков, умений, аппаратное, материаловедческое, метрологическое, информационное обеспечение процессов и технологических операций, направленных на создание материалов и систем с новыми свойствами, обусловленными проявлением наномасштабных факторов.

Нанотрубки – это полые внутри молекулы, состоящие примерно из миллиона атомов углерода и представляющие собой однослойные трубки диаметром около нанометра и длиной в несколько десятков микрон. На поверхности нанотрубки атомы углерода расположены в вершинах правильных шестиугольников.

Наночастица (nanoparticle) – частица с нанометровыми размерами (обычно от 1 до 100 нм). Может содержать от десятка до 10^6 атомов, связанных вместе.

Наноэлектроника (nanoelectronics) – область науки и техники, занимающаяся созданием, исследованием и применением электронных приборов с нанометровыми размерами элементов, в основе функционирования которых лежат квантовые эффекты.

Низкоразмерная структура (low-dimensional structure) – структура, свойства которой по крайней мере в одном из направлений существенно отличаются от свойств составляющего ее материала, присущих ему в объемном (3D) состо-

янии. Различают двумерные (2D), одномерные (1D) и нульмерные (0D) низкоразмерные структуры.

Одноэлектроника (single-electronics) – область электроники, в которой исследуется и используется в устройствах электронной обработки информации явление одноэлектронного туннелирования.

Одноэлектронное туннелирование (single-electron tunneling) – туннелирование электронов «поштучно», контролируемое кулоновской блокадой.

Окисление (oxidation) – класс химических реакций, которые протекают при присоединении к атомам исходного вещества кислорода или отнятия у них водорода, сопровождающиеся увеличением положительной валентности (заряда) или уменьшением отрицательной валентности (заряда) ионов данного вещества.

Оксидирование (oxidation) – группа технологических методов получения оксидных пленок на поверхности материала (обычно на металлах и полупроводниках).

Позитрон (positron) – античастица для электрона, имеющая такую же массу, но противоположный заряд и магнитный момент.

Полупроводник (semiconductor) – вещество с электрическим сопротивлением в диапазоне значений $10^{-4} \div 10^9$ Ом·см и отрицательным температурным коэффициентом сопротивления, промежуточным между проводниками и изоляторами. Энергетический зазор между валентной зоной и зоной проводимости у полупроводников не превышает 5 эВ.

Принцип Паули – две тождественные частицы не могут одновременно находиться в одном и том же квантовом состоянии, например, два электрона могут находиться в одном квантовом состоянии, если они неодинаковы, а это возможно, только если у них разные спины.

Работа выхода (work function) – энергия, необходимая для переноса электрона с уровня Ферми в вакуум.

Рашибы эффект (Rashba effect) – спин-зависимое расщепление энергетических уровней для электронов в отсутствие внешнего магнитного поля, вы-

званное отсутствием инверсной симметрии потенциала кристаллической решетки полупроводника.

Резонансное туннелирование (resonant tunneling) – резкое возрастание туннельного тока, протекающего через квантовый колодец (или колодцы), отделенный потенциальными барьерами от инжектирующего и принимающего электродов, при совпадении уровня Ферми в инжектирующем электроде и дискретного разрешенного уровня в колодце (колодцах).

Самоорганизация (self-organization) – упорядочение расположения взаимодействующих атомов в твердом теле, соответствующее минимуму потенциальной энергии данной атомарной системы.

Самосборка (self-assembling) – хемосорбция и формирование специфического расположения молекул на твердой поверхности.

Сверхпроводимость (superconductivity) – падение до нуля электрического сопротивления материала при его охлаждении. Это явление характеризуется критической температурой T_c , критическим магнитным полем H_c , критической плотностью тока J_c , ограничивающих сверху условия проявления материалами сверхпроводящих свойств.

Сверхрешетка (superlattice) – монокристаллическая пленка из одного материала, воспроизводящая постоянную решетки монокристаллической подложки из другого материала.

Сверхрешетка (устаревшее) – твердотельная периодическая структура, в которой на электроны, помимо обычного потенциала кристаллической решетки, действует дополнительный, обычно одномерный, потенциал с *периодом меньше* длины свободного пробега электронов, но значительно больше периода основной решетки. Это определение больше говорит об электронных свойствах многопериодных сверхрешеток, и меньше – о структурном устройстве, и не дает толкования, почему сверхрешетки бывают псевдоморфными, напряженными и релаксированными, также не охватывает случая однопериодных сверхрешеток.

Сотуннелирование (co-tunneling) – туннелирование подвижных носителей заряда через виртуальные состояния, возникающие в промежуточном островке, вследствие квантовых флуктуаций. Различают **упругое** сотуннелирование – электрон туннелирует в определенное энергетическое состояние в островке и из этого же состояния уходит из него, и **неупругое** сотуннелирование – электрон туннелирует в определенное энергетическое состояние в островке, а покидает островок электрон из другого энергетического состояния.

Спин (spin) – собственный угловой момент элементарной частицы, связанный с ее вращением вокруг собственной оси.

Спинтроника (spintronics) – область науки и техники, занимающаяся созданием, исследованием и применением электронных приборов, в которых спин электрона наравне с его зарядом используется для обработки информации.

Сродство к электрону (electron affinity) – энергия, необходимая для переноса электрона со дна зоны проводимости в вакуум.

Тугоплавкий металл – металл с температурой плавления выше 1850°C (Nb, Mo, Ta, W, Re, Ti, V, Cr, Zr, Hf, Ru, Os, Ir).

Туннелирование (tunneling) – перенос частицы через или *внутри* области, ограниченной потенциальным барьером, высота которого больше полной энергии данной частицы. Это невозможно с точки зрения классической механики, однако допустимо для квантовых частиц, для которых, как известно, характерен корпускулярно-волновой дуализм.

Туннельного магнетосопротивления эффект (tunneling magnetoresistance effect) – существенное изменение электрического сопротивления туннельной структуры с электродами из магнитных материалов при изменении направлений намагниченности магнитных электродов с параллельных на антипараллельные.

Углеродная нанотрубка (carbon nanotube) – естественная самоорганизованная наноструктура в виде трубки, состоящей из атомов углерода с замкнутыми друг на друга связями. Бывают одностенные углеродные нанотрубки (*sin-*

gle-wall carbon nanotubes) и многостенные углеродные нанотрубки (*multiwall carbon nanotubes*).

Ферми-скорость (Fermi velocity) – $v_F = \sqrt{2F / m_n}$, где F – энергия Ферми, m_n – эффективная масса ферми-частицы.

Ферми-энергия, уровень (Fermi energy, level) – энергия наивысшего заполненного состояния в невозбужденной электронной системе при температуре абсолютного нуля.

Ферми-частицы, фермионы (Fermi particles, fermions) – квантовые частицы, подчиняющиеся распределению Ферми – Дирака.

Фонон (phonon) – квант колебаний кристаллической решетки.

Фотолюминесценция (photoluminescence) – излучение света веществом при облучении его потоком фотонов. Фотолюминесценция полупроводников и диэлектриков имеет место только при облучении потоком фотонов с энергией, превышающей ширину их запрещенной зоны.

Фотон (photon) – квант света с энергией $E = h\nu$, где h – постоянная Планка, ν – частота колебаний соответствующей волны. Масса покоя фотона равна нулю.

Фуллерен (fullerene) – сфероподобная молекулярная структура, образованная углеродными атомами в форме футбольного мяча. Может быть **экзоэдральным** (*exohedral fullerene*) – образующимся при присоединении к нему примесных атомов и/или молекул – либо **эндоэдральным** (*endohedral fullerene*) – образующимся при внедрении в его полость примесных атомов.

Фуллерит (fullerite) – твердотельный композиционный материал, образованный из фуллеренов.

Холла эффект (Hall effect) – возникновение поперечной разности потенциалов в проводнике при протекании вдоль него тока, когда на проводник действует внешнее магнитное поле, ориентированное перпендикулярно направлению протекания тока.

Шоттки барьер (Schottky barrier) – барьер на границе металла и полупроводника, для которого характерно образование обедненной подвижными носителями заряда области пространственного заряда в полупроводнике.

Шредингера уравнение (Schrödinger equation) – фундаментальное волновое уравнение движения в квантовой механике. Для квантовой частицы массой m , движущейся в потенциальном поле $V(r)$, оно имеет вид:

$$i\hbar \frac{\partial \Psi(r,t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi(r,t) + V(r) \Psi(r,t),$$

где $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$; i – мнимая единица; \hbar – редуцированная постоянная

Планка, $\hbar = h / 2\pi$. $\Psi(r,t)$ – волновая функция, представляющая квантовую частицу. Уравнение Шредингера часто записывают в короткой форме с использо-

ванием гамильтониана (оператора Гамильтона) $H = -\frac{\hbar^2}{2m} \nabla^2 + V(r)$ в виде

$i\hbar \frac{\partial \Psi(r,t)}{\partial t} = H \Psi(r,t)$. Для расчета энергетических состояний квантовой системы

в стационарных условиях (то есть не меняющихся во времени) используют независящую от времени форму уравнения Шредингера:

$H = -\frac{\hbar^2}{2m} \nabla^2 \psi_n(r) + V(r) \psi_n(r) = E_n \psi_n(r)$ или равноценное этой форме уравнение

в краткой форме $H \psi_n(r) = E_n \psi_n(r)$.

Здесь $\psi_n(r)$ представляет волновую функцию квантовой частицы в энергетическом состоянии E_n , где n – целое число, называемое квантовым числом, обозначающее номер квантового состояния.

Экситон (exciton) – электронно-дырочная пара в связанном состоянии.

Электрон (electron) – стабильная отрицательно заряженная элементарная частица, имеющая массу $9,10939 \cdot 10^{-31}$ кг, неделимый заряд $1,602177 \cdot 10^{-19}$ К и спин $1/2$.