ОДЕССКАЯ НАЦИОНАЛЬНАЯ АКАДЕМИЯ СВЯЗИ
им. А.С. ПОПОВА

Физика электрорадиоматериалов

Рекомендовано
Министерством образования и науки Украины
в качестве учебного пособия
для студентов высших технических учебных заведений

Одеса 2011

ББК 31.233
УДК 621.382: 621.315: 681.7.068
В – 43

Рецензенты:

Ш.Д. Курмашов, заведующий научно-исследовательской лабораторией Одесского национального университета им. И.И.Мечникова, доктор физико-математических наук, профессор, лауреат Государственной премии Украины.

И.Е. Марончук, заведующий кафедрой физики Севастопольского национального университета ядерной энергии, доктор технических наук, профессор, лауреат Государственной премии Украины.

Г.А. Сукач, профессор кафедры общей и технической физики Национального университета биоресурсов и природопользования Украины, доктор физико-математических наук, лауреат Государственной премии Украины.

В–43 Викулин И.М. Физика электрорадиоматериалов: учеб. пос. для студентов вузов/ Викулин И.М., Коробицын Б.В., Криськив С.К. – Одеса, 2010. – 313 с.

Гриф учебного пособия присвоен решением
Министерством образования и науки Украины
№ от

В учебном пособии изложены основы строения и физических явлений, происходящих в проводниковых, полупроводниковых, диэлектрических и магнитных материалах. Показаны особенности применения этих материалов для изготовления устройств твердотельной электроники, включая оптоволоконные линии связи.

Рекомендовано для иностранных студентов высших учебных заведений по направлению «Телекоммуникации», а также может быть полезно широкому кругу специалистов в области твердотельной электроники и оптоволоконных линий связи.

ISBN

© Викулин И.М.,
Коробицын Б.В.,
Криськив С.К., 2011

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ 3

ПРЕДИСЛОВИЕ 7

Глава I. ОБЩИЕ СВЕДЕНИЯ О СТРОЕНИИ ВЕЩЕСТВА 9
1.1. СТРОЕНИЕ ВЕЩЕСТВА 9
1.1.1. Симметрия кристаллических веществ 9
1.1.2. Индексы Миллера 18
1.1.3. Элементарная ячейка кубического кристалла 19
1.1.4. Виды химических связей 21
1.1.5. Аморфные и аморфно-кристаллические вещества 25
1.1.6. Анизотропия кристаллов 26
1.1.7. Полиморфизм (аллотропия) 26
1.1.8. Дефекты кристаллов 27
1.2. ЭЛЕМЕНТЫ ЗОННОЙ ТЕОРИИ ТВЕРДЫХ ТЕЛ 32
1.2.1. Зонные диаграммы 32
1.2.2. Ширина запрещенной зоны 35
1.2.3. Эффективная масса 36
1.2.4. Статистика электронов 38

Глава II. ПРОВОДНИКИ 42
2.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ПРОВОДНИКАХ 42
2.1.1. Основы классической электронной теории металлов 42
2.1.2. Элементы квантовой теории металлов 48
2.1.3. Температурная зависимость удельного сопротивления металлов 54
2.1.4. Влияние примесей на удельное сопротивление металлов. Удельное сопротивление сплавов 57
2.1.5. Работа выхода электронов из металла и эффект Шоттки 60
2.1.6. Контакт двух металлов. Термоэлектродвижущая сила (термо-ЭДС) 62
2.1.7. Особенности сопротивления проводников на высоких частотах 65
2.1.8. Сопротивление тонких металлических пленок 67
2.1.9. Явление сверхпроводимости 69
2.2. КЛАСИФИКАЦИЯ ПРОВОДЯЩИХ МАТЕРИАЛОВ 77
2.2.1. Материалы с высокой проводимостью 77
2.2.2. Проводящие материалы с высоким удельным сопротивлением 78
2.2.3. Другие проводящие материалы 80
2.3. ПРИМЕНЕНИЕ ПРОВОДНИКОВ 84
2.3.1. Применение сверхпроводников 84

2.3.2. Провода ... 87
2.3.3. Резисторы ... 88
2.3.4. Датчики температуры, радиаторы охлаждения ... 90

Глава III. ПОЛУПРОВОДНИКИ ... 95
3.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ПОЛУПРОВОДНИКАХ ... 95
3.1.1. Структура и зонные диаграммы собственных полупроводников ... 96
3.1.1.1. Основные параметры собственных полупроводников ... 100
3.1.2. Структура и зонные диаграммы примесных полупроводников ... 103
3.1.2.1. Основные параметры примесных полупроводников ... 109
3.1.3. Полупроводники в сильных электрических полях ... 117
3.1.3.1. Наклон энергетических зон в электрическом поле ... 117
3.1.3.2. Закон Ома в сильных электрических полях ... 117
3.1.4. Физические явления в полупроводниках при внешних воздействиях ... 121
3.1.4.1. Эффект Холла ... 121
3.1.4.2. Термоэлектрические явления ... 123
3.1.4.3. Тензочувствительность ... 124
3.1.5. Оптические свойства полупроводников ... 125
3.1.5.1. Фотопроводимость ... 125
3.1.5.2. Виды поглощения света ... 127
3.5.5.3. Рекомбинация неравновесных носителей заряда ... 131
3.1.5.4. Люминесценция ... 133
3.2. КЛАССИФИКАЦИЯ ПОЛУПРОВОДНИКОВ ... 134
3.2.1. Простые полупроводники ... 134
3.2.2. Полупроводниковые соединения. Общие свойства ... 139
3.2.3. Полупроводниковые соединения типа $A^{IV}B^{IV}$ ... 140
3.2.4. Полупроводниковые соединения типа $A^{III}B^{V}$ ... 142
3.2.5. Полупроводниковые соединения типа $A^{II}B^{VI}$ (халькогениды) ... 144
3.3. ПРИМЕНЕНИЕ ПОЛУПРОВОДНИКОВ ... 147
3.3.1. Приборы на одном типе проводимости ... 148
3.3.1.1. Датчики Холла ... 148
3.3.1.2. Термоэлектрические устройства ... 149
3.3.1.3. Полупроводниковые тензометры ... 150
3.3.2. Биполярные приборы ... 151
3.3.2.1. Электронно-дырочные переходы ... 151
3.3.2.2. Туннельные диоды ... 161
3.3.2.3. Гетеропереходы ... 164
3.3.2.4. Биполярные транзисторы ... 164
3.3.3. Униполярные приборы ... 169
3.3.3.1. Контакты металл-полупроводник. Диоды Шоттки ... 169
3.3.3.2. Полевые транзисторы. Контакты металл-диэлектрик-полупроводник. МДП-транзисторы. Приборы с зарядовой связью ... 173

3.3.4. Светодиоды 177
3.3.4.1. Физические принципы работы светодиодов 177
3.3.4.2. Характеристики светодиодов 178
3.3.4.3. Материалы для светодиодов 180
3.3.5. Полупроводниковые лазеры 185
3.3.6. Фотодетекторы 189
Глава IV. ДИЭЛЕКТРИКИ 199
4.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ДИЭЛЕКТРИКАХ 199
4.1.1. Основные понятия 199
4.1.1.1. Поляризация диэлектриков 199
4.1.1.2. Виды поляризации диэлектриков 201
4.1.1.3. Классификация диэлектриков по механизмам поляризации 203
4.1.1.4. Поляризация сегнетоэлектриков 205
4.1.1.5. Механизм спонтанной поляризации 208
4.1.1.6. Пироэлектричество 210
4.1.1.7. Пьезоэлектрический эффект Электретные состояния в диэлектриках 211
4.1.1.8. Электретные состояния в диэлектриках 214
4.1.2. Физические свойства слоистых и гетерогенных диэлектриков 216
4.1.2.1. Зависимость свойств диэлектриков от температуры 216
4.1.2.2. Напряженность электрического поля в многослойных диэлектриках 217
4.1.2.3. Диэлектрическая проницаемость гетерогенных и слоистых структур 219
4.1.3. Электрические свойства диэлектриков 221
4.1.3.1. Общие сведения об проводимости 221
4.1.3.2. Проводимость твердых диэлектриков 222
4.1.3.3. Диэлектрические потери 224
4.1.3.4. Виды диэлектрических потерь 228
4.1.3.5. Пробой твердых диэлектриков 230
4.1.3. Оптические свойства диэлектриков 235
4.1.3.1. Дисперсия света 235
4.1.4.2. Полное внутреннее отражение 236
4.1.4.3. Поглощение света 237
4.1.4.4. Поляризация света 237
4.1.4.5. Оптическая анизотропия 239
4.1.4.6. Искусственная оптическая анизотропия 241
4.2. КЛАССИФИКАЦИЯ ДИЭЛЕКТРИКОВ 243
4.2.1. Пассивные диэлектрики 243
4.2.1.1. Твердые органические диэлектрики 243
4.2.1.2. Твердые неогранические диэлектрики 249
4.2.2. Активные диэлектрики 257
4.2.2.1. Сегнетоэлектрики, пьезоэлектрики, электреты 257
4.3. ПРИМЕНЕНИЕ ДИЭЛЕКТРИКОВ 260
4.3.1. Световоды 262
4.3.2. Модуляторы 267
4.3.3. Оптические квантовые генераторы 272

4.3.4. Конденсаторы 277

Глава V. ФЕРРОМАГНЕТИКИ..... 282
5.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ФЕРРОМАГНЕТИКАХ..... 282
5.1.1. Вещество в магнитном поле. Намагниченность..... 282
5.1.2. Свойства ферромагнетиков..... 283
5.1.3. Магнитные потери 288
5.2. КЛАССИФИКАЦИЯ ФЕРРОМАГНЕТИКОВ 291
5.2.1. Магнитомягкие ферромагнетики 291
5.2.2. Магнитотвердые ферромагнетики 295
5.2.3. Ферромагнетики для хранения информации 298
5.3. ПРИМЕНЕНИЕ ФЕРРОМАГНЕТИКОВ 299
5.3.1. Трансформаторы 299
5.3.2. Элементы памяти 301
5.3.3. Ферриты для устройств СВЧ 301

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ..... 309

ЛИТЕРАТУРА..... 311

ПРЕДИСЛОВИЕ

Современную повседневную жизнь невозможно представить себе без радиоэлектронной аппаратуры, значение которой не ограничивается только радиовещанием и связью. На базе электроники развивается радиофизика, радионавигация, радиолокация, телемеханика, радиоастрономия, радиоспектрография, автоматика, современная вычислительная и ядерная ускорительная техника, системы связи, в том числе и оптической.

Основой всех этих областей науки и техники является элементная база электроники, состоящая из диодов, транзисторов, микросхем, конденсаторов, индуктивностей и т.д. Для создания этих элементов необходимо знать не только свойства материалов, из которых они создаются, но и возможности их технологического совмещения в интегральных схемах, где все элементы объединены в одном твердотельном блоке. Дальнейшее развитие микроэлектроники в сторону уменьшения размеров элементов до нанометров привело к созданию наноэлектроники, где требования к технологическому уровню материалов очень высоки.

Важнейшую роль в развитии творческих способностей будущих специалистов, в приобретении ими практических навыков играет самостоятельная работа студентов. Данное учебное пособие предназначено в первую очередь для самостоятельной работы. Большое внимание уделено рассмотрению физических свойств материалов, а также возможностям практического применения материалов в радиоэлектронной аппаратуре различного назначения.

Исследование электрофизических явлений, происходящих в различных материалах в оптическом диапазоне частот, и создание квантовых оптических генераторов привело к возникновению и развитию новой отрасли – оптоэлектроники. Это потребовало разработки принципиально новых материалов (активных диэлектриков, полупроводниковых соединений), способных преобразовывать электрическую энергию в световую и наоборот (лазеры, светодиоды, оптроны, оптические модуляторы и др.).

По общей классификации материалов для электронной техники их делят на четыре класса: проводники, полупроводники, диэлектрики и магнитные материалы. В настоящем пособии вначале рассматривается класс проводников электрического тока с последующим переходом к полупроводникам, диэлектрикам и магнитным материалам. При таком рассмотрении лучше видны логические связи в материаловедении и целесообразность введения новых понятий.

В пособии «Физика электрорадиоматериалов» предполагается сделать упор вначале на физические процессы, а затем на способы их реализации в соответствующих устройствах. Подробнее это выражается в следующем.

Во-первых, изучив принципы работы приборов, можно предопределить, какие качества электрорадиоматериалов наиболее важны для эффективной работы приборов, и какие параметры необходимо изменить, чтобы улучшить

их работу. Потребности же в новых приборах задают ряд требований к материалам, из которых они могут быть созданы.

Во-вторых, необходимо было показать, как, изменяя параметры материалов (уровень легирования, ширину запрещенной зоны и т.д.), можно получить целый ряд приборов, отличающихся по принципам роботы и характеристикам. Например, обычный диод при сильном легировании одного и того же полупроводникового материала становится туннельным диодом, имеющим новые характеристики; при контакте двух полупроводников с разной шириной запрещенной зоны получаются гетеропереходы и т.д.

В-третьих, авторы рассматривают работу приборов именно с точки зрения использования в них основных эффектов и характерных физических свойств материалов. В пособии не приводятся сложные расчеты, а лишь те формулы или их выводы, в которых видно влияние физических свойств и параметров материала на рабочие характеристики приборов.

Пособие состоит из пяти глав, первая из которых представляет обобщенные сведения о строении материалов и основы зонных представлений. Таблицы, приведенные в тексте пособия, являются справочным и обобщающим материалом. Каждая из последующих глав разбита на три раздела, в которых рассматриваются физические процессы в данных материалах, классификация материалов и наиболее важные примеры применения материалов для создания элементов и устройств электронной техники. Разделы «Полупроводники» и «Диэлектрики» несколько расширены ввиду особой важности этих материалов для создания элементной базы в области электроники и связи.

Для лучшего усвоения темы каждая глава завершается вопросами и задачами средней сложности, решение которых не выходит за рамки настоящего пособия. Необходимые для решения справочные данные приводятся в условиях задач, а в конце каждой задачи даны ответы. Более подробное изложение рассматриваемых в пособии вопросов можно найти в списке основной литературы.

Пособие предназначено для иностранных студентов в соответствии с программой курса «Физика электрорадиоматериалов», утвержденного министерством образования и науки Украины и на основе одноименного курса, читаемого кафедрой физики оптической связи в Одесской национальной академии связи имени А. С. Попова.

Глава I. ОБЩИЕ СВЕДЕНИЯ О СТРОЕНИИ ВЕЩЕСТВА

1.1. СТРОЕНИЕ ВЕЩЕСТВА

В твердых веществах атомы и молекулы расположены в геометрически правильном порядке либо в условно хаотическом беспорядке. Вещества, обладающие геометрически упорядоченным расположением атомов или молекул, называют *кристаллическими*, а вещества с условно хаотическим расположением атомов или молекул – *аморфными.* Кристаллические вещества характеризуются *дальним порядком*, аморфные – *ближним порядком* расположения атомов.

Твердое состояние возникает при столь сильном взаимодействии между молекулами (атомами и ионами), что тепловое движение молекул не играет в структуре определяющей роли, которую оно играет в жидкостях и особенно в газах. Молекулы располагаются относительно друг друга в некоторых фиксированных положениях, осуществляя небольшие тепловые колебания около положений равновесия. Взаимное расположение точек равновесия определяется условиями равновесия. Когда эти условия равновесия выполнены в некоторой области пространства, то они выполнены и в другой области и обуславливают аналогичное расположение молекул в любой другой области пространства. Это означает, что структура твердых тел периодична. Она реализуется в виде кристаллической решетки, а точки равновесия составляющих кристалл атомов, молекул или ионов называются узлами кристаллической решетки.

При таком подходе возникает вопрос о причинах существования аморфных твердых тел, таких как стекло, пластик и аналогичных веществ без строго периодической структуры. Ответ состоит в том, что они находятся в не вполне равновесном состоянии, и с течением времени изменяют свою структуру, приближаясь к кристаллическому состоянию. Процесс кристаллизации аморфных тел значительно ускоряется при высоких температурах и давлениях. Аморфные твердые тела иногда называют переохлажденными жидкостями.

Изучение степени упорядоченности строения вещества позволяет не только отнести вещество к кристаллическим или аморфным веществам, но и определить зернистость структуры, т.е. отличить *монокристаллический* материал от *поликристаллического* материала.

1.1.1. Симметрия кристаллических веществ

Большинство твердых веществ, обладающих *кристаллической структурой*, образует *кристаллы* определенной формы. Внешняя форма кристалла часто отражает его внутреннее строение и обусловлена геометрически правильным расположением его атомов или молекул.

Элементарная ячейка – это минимальный объем кристалла, полностью сохраняющий его свойства, и при параллельном переносе (трансляции) которого в трех измерениях можно построить весь кристалл. Элементарная ячейка зачастую имеет форму параллелепипеда.

Выбрав начало координат в некотором узле такой решетки, можно радиус-вектор любого другого узла представить в виде:

$$\vec{r} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3, \qquad (1.1)$$

где $\vec{a}_1$, $\vec{a}_2$ и $\vec{a}_3$ – три вектора, не лежащие в одной плоскости, а n_1, n_2 и n_3 – любые целые числа (включая ноль). Векторы $\vec{a}_1$, $\vec{a}_2$ и $\vec{a}_3$ называют *базисными*, а их совокупность – *базисом решетки*. Длины ребер a_1, a_2 и a_3 вместе с атомами в его вершинах называются *основными периодами* (*параметрами*) решетки. Если в формуле (1.1) числа n_1, n_2, n_3 пробегают всевозможные независимые целые значения, то и радиус-вектор $\vec{r}$ проходит все узлы решетки без исключения. Такая решетка называется *примитивной решеткой* или *решеткой Бравэ*, а ее элементарная ячейка – *примитивной ячейкой*.

Все многообразие реальных кристаллических решеток, вообще говоря, не может быть представлено только в виде примитивных решеток (решеток Бравэ), а является совокупностью нескольких решеток Бравэ. Такие решетки называются *сложными решетками*.

Выбор базиса даже примитивной решетки не является единственным. Различные примитивные базисы отличаются друг от друга длиной базисных векторов или основными периодами решетки, хотя объемы элементарных ячеек с различными базисами не изменяются.

Важнейшим геометрическим свойством кристаллических твердых тел является их *симметрия*. Под симметрией понимается способность твердого тела совмещаться самим с собой в результате его движений или воображаемых операций над точками. Чем большим количеством способов производится такое совмещение, тем более симметричным является тело.

Несмотря на кажущееся большое многообразие типов симметрии твердых тел, все они складываются из четырех видов симметрии: *ось симметрии n-го порядка*, *плоскость симметрии*, *центр симметрии*, *зеркально-поворотная ось n-го порядка*. Различные комбинации этих четырех элементов симметрии и составляют всевозможные симметрии твердых тел.

Ось симметрии n-го порядка. Если тело совмещается с самим собой при повороте вокруг некоторой оси на угол $2\pi/n$, то такая ось называется осью симметрии n-го порядка. Например, квадрат совмещается с самим собой при повороте на угол $\pi/2$ вокруг оси, проходящей через точки пересечения его диагоналей. Следовательно, эта ось является осью 4-го порядка. Любая ось тела является осью симметрии 1-го порядка (угол поворота 2π).

Плоскость симметрии. К совмещениям тела с самим собой только с помощью пространственных движений симметрия не сводится. Если тело

совмещается с самим собой в результате зеркального отражения его точек в некоторой плоскости, то эта плоскость называется плоскостью симметрии.

Центр симметрии. Операция симметрии, называемая *инверсией*, состоит из поворота на угол π и последующего отражения в плоскости, перпендикулярной оси поворота. В результате этой операции симметрии радиус-вектор r меняется на радиус-вектор $(-r)$. Если тело совмещается с самим собой при инверсии относительно некоторой точки, то эта точка называется центром симметрии. Все оси, проходящие через центр симметрии, эквивалентны.

Зеркально-поворотная ось n-го порядка. Если тело совмещается между собой при повороте на угол $2\pi/n$ и отражении в плоскости, перпендикулярной этой оси, то такая ось называется зеркально-поворотная ось n-го порядка.

Элементы симметрии решетки. Всякая примитивная решетка имеет центр симметрии, которым может быть любой узел примитивного параллелепипеда, середины его ребер и центры его граней. Плоскость симметрии также является элементом симметрии решеток. Что же касается осей и зеркально-поворотных осей симметрии, то они могут быть лишь осями 2, 3, 4 и 6-го порядков, а оси других порядков невозможны. Для доказательства этого заметим, что при вращении атомы решетки перемещаются в плоскостях, перпендикулярных оси вращения.

Рассмотрим атомы, лежащие в некоторой плоскости. Они создают плоскую кристаллическую решетку, узлы которой образуют систему правильных одинаковых многоугольников, переходящих при вращении друг в друга и, следовательно, плотно, без промежутков, покрывающих всю поверхность.

Выделим точку O (рис. 1.1), в которой сходятся ребра примыкающих в этой точке правильных многоугольников. Если p – число ребер, сходящихся в этой точке, то угол между ребрами равен $2\pi/p$. С другой стороны, угол между сторонами правильного n-угольника равен $\pi(n\text{-}2)/n$. При заполнении плоскости правильными многоугольниками без промежутков эти углы равны:

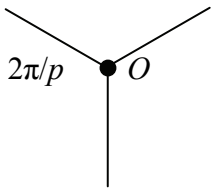


Рис. 1.1 – К определению порядка поворотной оси

$$2\pi/p = \pi(n\text{-}2)/n. \qquad (1.2)$$

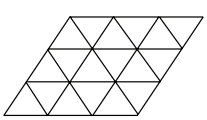
Отсюда следует, что

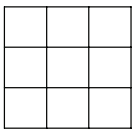
$$p = 2n/(n\text{-}2), \qquad (1.3)$$

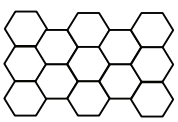
причем, числа p и n должны быть целыми. Решениями этого уравнения в целых числах являются только значения $n = 3$, $n = 4$, $n = 6$.

Таким образом, поверхность без промежутков можно покрыть:

равносторонними треугольниками ($n = 3$); квадратами ($n = 4$); правильными шести-угольниками ($n = 6$).







Других правильных многоугольников, способных без промежутков покрыть поверхность, не существует.

К ним добавляется, очевидно, ось $n = 2$, соответствующая отражению в плоскости, проходящей через ось, а также тривиальная ось $n = 1$, соответствующая углу поворота на 2π. Таким образом, у кристаллической решетки возможны оси симметрии только 2, 3, 4 и 6-го порядков. Аналогичным образом доказывается, что и зеркально-поворотные оси могут быть только тех же порядков.

В результате получается, что число элементов точечных групп симметрии у кристаллических решеток конечно, следовательно, конечно и число возможных симметрий.

Точечные группы симметрии. Совокупность элементов симметрии тела называется его *группой симметрии*. Рассмотренные выше элементы симметрии характерны тем, что они оставляют, по крайней мере, одну точку тела *неподвижной*. Соответствующие им группы симметрии называются *точечными*.

Поскольку сложная кристаллическая решетка состоит из решеток Бравэ, то классифицировать кристаллы в первую очередь целесообразно по точечной симметрии решеток Бравэ. Дадим два эквивалентных определения решетки Бравэ:

а) Решетка Бравэ – это бесконечная периодическая структура, образованная дискретными точками и имеющая абсолютно одинаковый пространственный порядок и ориентацию независимо от того, какую точку мы принимаем за исходную.

б) Трехмерная решетка Бравэ образована всеми точками с радиус-векторами $\vec{r}$ вида (1.1). Поэтому, чтобы достигнуть точки $\sum n_i \vec{a}_i$, необходимо проделать n_i шагов длиной a_i в направлении $\vec{a}_i$, где $i = 1,2,3\ldots$

Рассмотрим классификацию решеток Бравэ. Существует только семь различных точечных групп, которые может иметь решетка Бравэ. Любая кристаллическая структура принадлежит к одной из *семи кристаллических систем* (*сингоний*) в зависимости от того, какая из семи точечных групп является точечной группой данной решетки Бравэ.

Ниже мы рассмотрим все кристаллические системы, а также решетки Бравэ, принадлежащие каждой из них. Число решеток Бравэ в системе укажем в скобках после названия системы.

Кубическая система **(3).** Кубическая система содержит те решетки Бравэ, точечная группа которых совпадает с группой симметрии куба (рис. 1.2, *а*). Существуют три кубические решетки Бравэ: *простая кубическая*, *объемноцентрированная кубическая* (ОЦК), *гранецентрированная кубическая* (ГЦК). Если частицы расположены только в вершинах решетки, то такая решетка называется *простой***.** *Объемноцентрированные* решетки содержат дополнительную частицу в точке пересечения пространственных диагоналей, *гранецентрированные*

решетки содержат дополнительные частицы в точках пересечения диагоналей граней.

***Тетрагональная система* (2).** Чтобы понизить симметрию куба, можно взять его за противоположные грани и вытянуть в прямую призму с квадратным основанием ($a_1 = a_2$), но с высотой a_3, не равной сторонам квадрата. Группа симметрии такого объекта – тетрагональная группа. Растягивая подобным образом простую кубическую решетку, можно получить простую тетрагональную решетку Бравэ. Третий вектор называют a_3-осью. Растягивая подобным образом объемноцентрированную и гранецентрированную кубические решетки, удается получить лишь одну решетку тетрагональной системы – объемноцентрированную тетрагональную (рис. 1.2, *б*).

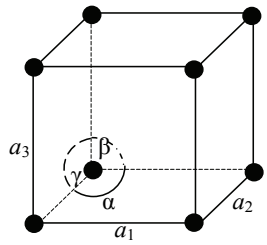
Ромбическая система (**4**). Переходя к менее симметричным деформациям куба, мы можем понизить тетрагональную симметрию, преобразовав в прямоугольники квадратные грани объекта, показанного на рис. 1.2, *б*. Вытягивая простую тетрагональную решетку вдоль одной из осей, например, вдоль a_2-оси (рис. 1.2, *б*), мы приходим к простой ромбической решетке Бравэ. В результате получается объект, имеющий три взаимно перпендикулярные грани разной длины (рис. 1.2, *в*), группу симметрии которого называют ромбической. Если простую тетрагональную решетку растянуть вдоль диагонали квадрата основания, то получится решетка Бравэ с ромбической точечной группой симметрии, называемая *базоцентрированной ромбической*. (*Базоцентрированные* решетки содержат дополнительные частицы в точках пересечения диагоналей только на двух противолежащих друг другу гранях.)

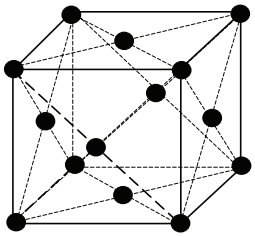
Растягивая подобным образом тетрагональную решетку вдоль других пространственных диагоналей, можно получить объемно-центрированную и гранецентрированную ромбические решетки Бравэ.

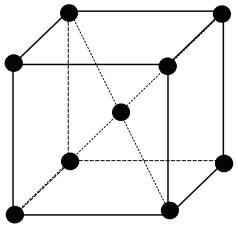
Этими четырьмя решетками Бравэ полностью исчерпывается ромбическая система.

Моноклинная система (**2**). Ромбическую симметрию можно понизить, превратив прямоугольные грани, перпендикулярные a_3-оси (рис 1.2, *в*), в произвольные параллелограммы. Получающийся объект (рис 1.2, *г*) имеет моноклинную группу симметрии. Искажая таким способом простую ромбическую решетку, мы получаем *простую моноклинную* решетку Бравэ. Искажая аналогичным образом базоцентрированную ромбическую решетку Бравэ, мы тоже получаем простую моноклинную пространственную группу. Однако если исказить подобным образом гранецентрированную решетку Бравэ или объемноцентророванную ромбическую решетку, то в результате возникает *базоцентрированная моноклинная* решетка Бравэ (рис. 1.2, *г*).

Триклинная система (**1**). Искажение куба завершится, если наклонить a_3-ось (рис. 1.2, *г*) так, чтобы она не была перпендикулярна двум другим осям. Получающийся в результате объект изображен на рис. 1.2, *д*; он не должен удовлетворять никаким ограничениям, кроме требования параллельности противоположных граней. Искажая таким путем любую из

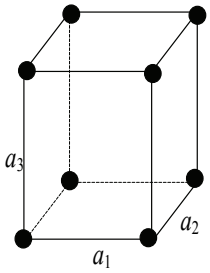


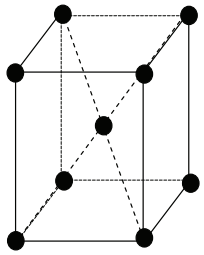




простая гранецентрированная объемноцентрированная

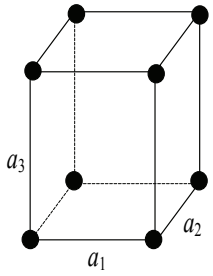
а) кубическая система ($a_1 = a_2 = a_3$; α = β = γ = 90°)

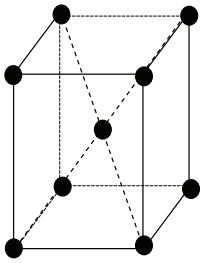


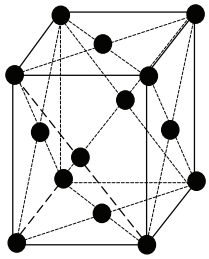


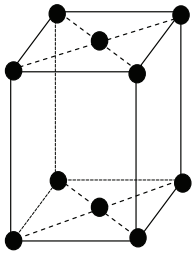
простая объемноцентрированная

б) тетрагональная система ($a_1 = a_2 \neq a_3$; α = β = γ = 90°)



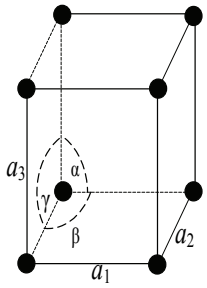


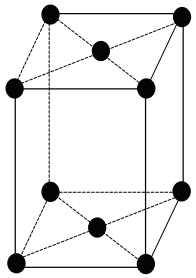




простая объемноцентрированная гранецентрированная базоцентрированная

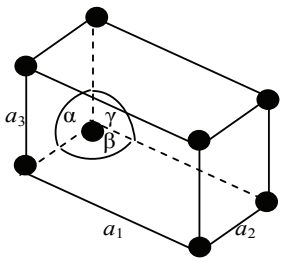
в) ромбическая система ($a_1 \neq a_2 \neq a_3$; α = β = γ = 90°)





простая базоцентрированная

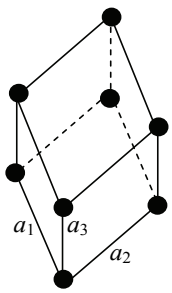
г) моноклинная система ($a_1 \neq a_2 \neq a_3$; α = γ = 90°; β ≠ 90°)



простая

д) триклинная система

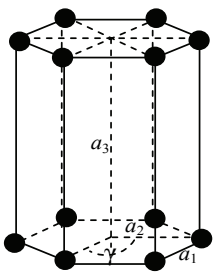
($a_1 \neq a_2 \neq a_3$; α ≠ β ≠ γ; γ = 90°)



простая

е) тригональная система

($a_1 = a_2 = a_3$; α = β = γ ≠ 90°)



простая

ж) гексагональная система

($a_1 = a_2 \neq a_3$; α = β = 90°; γ = 120°)

Рис. 1.2 – Объекты с симметриями точечных групп решеток Бравэ, принадлежащих семи кристаллическим системам: кубической (*а*), тетрагональной (*б*), ромбической (*в*), моноклинной (*г*), триклинной (*д*), тригональной (ромбоэдрическая) (*е*), гексагональной (*ж*).

моноклинных решеток Бравэ, можно построить *триклинную* решетку Бравэ. Эта решетка порождается тройкой основных векторов, не связанных какими-либо соотношениями, следовательно, она представляет собой решетку Бравэ с минимальной симметрией. Все же триклинная группа не является группой объекта без всякой симметрии, так как решетка Бравэ всегда инвариантна относительно *инверсии* с центром в любой точке решетки.

Исказив такими способами куб, мы получили двенадцать из четырнадцати решеток Бравэ, и пять из семи кристаллических систем. Чтобы найти тринадцатую и шестую, вернемся к исходному кубу и исказим его несколько иным способом.

Тригональная система (**1**). Тригональная точечная группа описывает симметрию объекта, который получается, если растянуть куб вдоль объемной диагонали (рис. 1.2, *е*). В результате такого искажения любой из трех кубических решеток Бравэ возникает *ромбоэдрическая* (или *тригональная*) решетка Бравэ. Она порождается тремя основными векторами равной длины, образующими равные углы друг с другом.

И, наконец, последняя из систем, которая не имеет отношения к кубу.

Гексагональная система (**1**). Гексагональная точечная группа есть группа симметрии правильной шестиугольной призмы (рис. 1.2, *ж*). Простая гексагональная решетка Бравэ имеет гексагональную точечную группу и является единственной решеткой Бравэ в гексагональной системе.

Описанные выше семь кристаллических систем и четырнадцать решеток Бравэ исчерпывают все возможные случаи. Бравэ произвел эти расчеты в1845 г. и показал, что, хотя симметрия решетки не обязательно совпадает с симметрией любой примитивной ячейки, можно всегда найти такую примитивную ячейку, которая имеет те же элементы симметрии, что и решетка в целом. Это возможно для всех типов решеток, за исключением гексагональных, где примитивная ячейка не содержит всех элементов симметрии, которые имеются у решетки в целом. Из 7 точечных групп (сингоний) решетки Бравэ можно построить 32 кристаллографические группы, 11 из которых обладают центральной симметрией.

Трансляционная симметрия. Ввиду бесконечной протяженности решетка кроме симметрий, характерных для твердых тел, обладает *трансляционной симметрией*, т.е. способностью совмещаться с собой в результате поступательного перемещения. Это главный тип симметрии кристаллов.

Пространственные группы. Элементы симметрии твердого тела образуют точечные группы симметрии. Если к ним добавить трансляционную симметрию, характерную для периодических бесконечных структур, то совокупность этих симметрий образует *пространственную группу*. Как было показано Е.С. Федоровым, всего может существовать 230 различных пространственных групп.

Исходя из симметрии, применяют различные классификации кристаллов. Наиболее общей является деление на категории, которое связано с понятием *особенного* или *единичного направления*. *Единственное*, *не*

повторяющееся в кристаллическом многограннике направление называют особенным или единичным. Всего существуют три категории кристаллов.

Таблица 1.1 – Деление кристаллов на категории

Категория	Число особенных направлений	Особенности морфологии
Высшая	Несколько осей симметрии порядка больше 2. Обязательно есть 4 оси 3-го порядка	Внешняя форма изометрична
Средняя	Одно особенное направление (одна ось симметрии порядка: 3, 4, 6)	Характерные формы – призмы, пирамиды и пр.
Низшая	Несколько особенных направлений. Нет осей симметрии порядка выше, чем 2	Ярко выражена анизотропия формы

Результаты вышеприведенного кристаллографического анализа можно представить в виде итоговой таблицы 1.2.

Таблица 1.2 – Классификация пространственных решеток кристаллических систем.

Категория	Оптическая классификация	Кристаллическая система	Пространственная решетка	Соотношение между осевыми углами и осевыми единицами
Высшая	Изотропная среда	1. Кубическая система	простая, объемноцентрированная, гранецентрированная	$a_1 = a_2 = a_3$; $\alpha = \beta = \gamma = 90°$
Средняя	Одноосные кристаллы	2. Тетрагональная система	простая, объемноцентрированная	$a_1 = a_2 \neq a_3$; $\alpha = \beta = \gamma = 90°$
		3. Тригональная (ромбоэдрическая) система	простая	$a_1 \neq a_2 \neq a_3$; $\alpha = \beta = \gamma = 90°$
		4. Гексагональная система	простая	$a_1 = a_2 \neq a_3$; $\alpha = \beta = 90°$; $\gamma = 120°$
Низшая	Двуосные кристаллы	5. Триклинная система	простая	$a_1 \neq a_2 \neq a_3$; $\alpha \neq \beta \neq \gamma$; $\gamma = 90°$
		6. Моноклинная система	простая, базоцетрированная	$a_1 \neq a_2 \neq a_3$; $\alpha = \gamma = 90°$; $\beta \neq 90°$
		7. Ромбическая (орторомбическая) система	простая, базоцетрированная, объемноцентрированная, гранецентрированная	$a_1 \neq a_2 \neq a_3$; $\alpha = \beta = \gamma = 90°$

Так как в дальнейшем нас будет интересовать применение кристаллических материалов в оптоэлектронике, дадим краткие сведения о том, как симметрия кристаллов влияет на их оптические свойства. Одной из важнейших оптических характеристик является показатель преломления света n в кристаллах, равный отношению скорости света в пустоте к скорости света в среде $n = c_0/c$.

В кубических кристаллах скорость света во всех направлениях одинакова. Мы встречаемся с принципиально важным случаем: симметрия

оптических явлений в кубических кристаллах гораздо выше, чем симметрия среды.

В кристаллах средних сингоний скорость света зависит от направления. При этом возможны два новых явления: *двойное лучепреломление* прошедших через кристаллы лучей и *вращение плоскости поляризации* лучей. *Двойное лучепреломление* заключается в раздвоении падающего неполяризованного луча в кристалле на два луча с взаимно перпендикулярными поляризациями, распространяющимися с разными скоростями, т.е. характеризующимися двумя различными показателями преломления.

Оптической осью кристалла называется кристаллографическое направление в оптически анизотропном кристалле, вдоль которого свет распространяется не испытывая двойного лучепреломления.

Оптически анизотропные кристаллы бывают, в зависимости от типа их симметрии, *одноосными* либо *двуосными*, т.е. имеют одну или две оптические оси.

В одноосных кристаллах один из лучей, образующихся при двойном лучепреломлении, подчиняется законам преломления света: он лежит в плоскости падения и удовлетворяет закону Снеллиуса. Его называют *обыкновенным лучом*. Второй луч называют *необыкновенным* лучом, так как он, в общем случае, не лежит в плоскости падения и не подчиняется закону Снеллиуса. В двуосных кристаллах оба преломленных луча ведут себя как необыкновенные.

Явление искусственного двойного лучепреломления, широко применяющегося в оптических модуляторах света, подробнее будет рассмотрено в IV главе.

1.1.2. Индексы Миллера

Кристаллографические направления – прямые, выходящие из любой точки отсчета, вдоль которых располагаются атомы. Точками отсчета могут служить, например, вершины параллелепипеда или шестигранной призмы. Примерами кристаллографических направлений являются ребра и диагонали граней призмы, но могут быть и другие направления. Кристаллографические направления и плоскости задаются индексами Миллера. Кристалло-графические направления перпендикулярны кристаллографическим плоскостям с одинаковым набором индексов.

Индексы Миллера – это набор не имеющих общего множителя минимальных целых чисел, обратно пропорциональных длинам отрезков, отсекаемых кристаллографической плоскостью на осях кристалла: $h:k:l=\frac{1}{x}:\frac{1}{y}:\frac{1}{z}$. Для нахождения индексов Миллера нужно сначала найти точки, в которых рассматриваемая плоскость пересекает координатные оси, и записать их в единицах постоянных решетки. Затем взять обратные значения полученных целых чисел и привести их к наименьшему целому, кратному

каждому из значений. Полученный результат записывают в круглых скобках (*hkl*). Это и есть индексы Миллера отдельной плоскости или системы параллельных плоскостей. На рис. 1.3 показаны основные плоскости кубического кристалла и соответствующие индексы Миллера.

Приведем некоторые принятые обозначения:

($\bar{h}\,kl$) – плоскость, пересекающая ось *x* при отрицательных значениях этой координаты.

{*hkl*} – для плоскостей эквивалентной симметрии. Например, для кубического кристалла {100} соответствует плоскостям (100), (010), (001), ($\bar{1}\,00$), ($0\,\bar{1}\,0$) и ($00\,\bar{1}$).

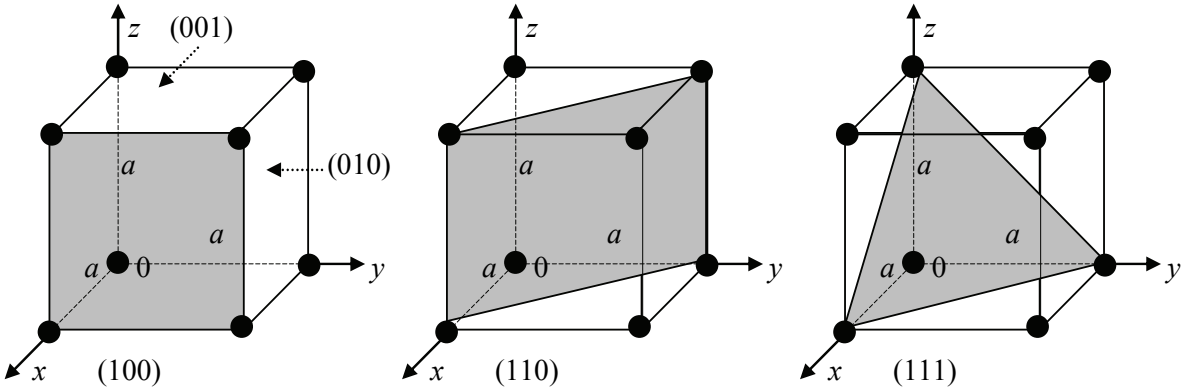


Рис. 1.3 – Индексы Миллера основных плоскостей кубического кристалла

[*hkl*] – для кристаллографического направления; например, [100] – направление вдоль оси *x*.

$\langle hkl \rangle$ – для совокупности всех эквивалентных направлений.

[$a_1a_2a_3c$] – для гексагональных решеток, где обычно используется четыре оси, причем ось *c* соответствует направлению [0001].

Для двух элементарных полупроводников (германия и кремния) атомные плоскости {111} являются плоскостями наиболее легкого разрушения или скола. В противоположность этому арсенид галлия с такой же кристаллической структурой и сравнительно небольшой ионной компонентой в атомных связях скалывается по плоскости {110}.

1.1.3. Элементарная ячейка кубического кристалла

Рассмотрим более подробно высшую форму кристаллической симметрии – кубическую сингонию. Известно, что большинство металлов, полупроводников и ионных кристаллов, являющихся диэлектриками, кристаллизуются именно в этой системе.

Для кубических кристаллических решеток удобно в качестве элементарной ячейки выбрать объем, заключенный внутри граней куба. При этом атомы, центры которых расположены в вершинах куба, входят каждый одновременно в восемь ближайших элементарных ячеек, давая вклад в одну

ячейку, равный одному атому. Таким образом, долю участия конкретного атома в элементарной ячейке следует точно учитывать. Элементарная ячейка кубического кристалла имеет вид простой, гранецентрированной или объемноцентрированной ячейки. Если частицы расположены только в вершинах ячейки, то такая ячейка называется *простой* (см. рис. 1.4). На

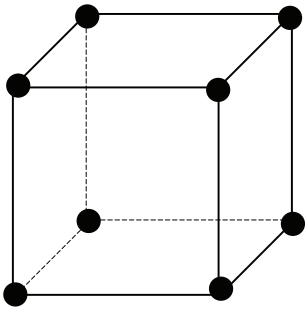


Рис. 1.4 – Ячейка простой кубической решетки

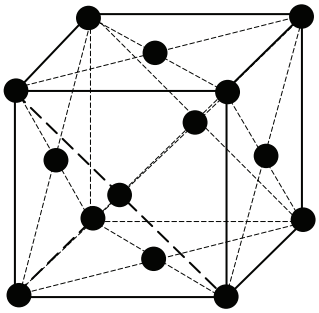


Рис. 1.5 – Ячейка гранецентрированной кубической решетки

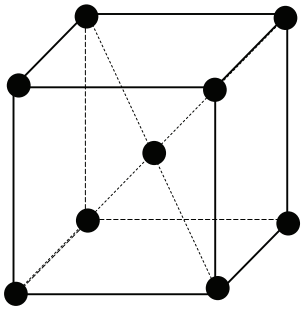


Рис. 1.6 – Ячейка объемноцентрированной кубической решетки

объем такой ячейки приходится один узел решетки, т.е. одна частица (N = 1). *Объемноцентрированные* кубические решетки (ОЦК) содержат дополнительную частицу в точке пересечения пространственных диагоналей (см. рис. 1.5), *гранецентрированные* кубические решетки (ГЦК) содержат дополнительные частицы в точках пересечения диагоналей граней (см. рис. 1.6). На объем одной элементарной ячейки объемноцентрированной решетки приходятся две частицы (N = 2), гранецентрированной – четыре (N = 4).

Важнейшими параметрами кристалла являются размеры элементарной ячейки; эти размеры определяют как равновесные расстояния в направлении осей между центрами частиц, занимающих соседние узлы решетки, и называют *периодами* (*параметрами*) *решетки* a_1, a_2, a_3.

Расстояние d между ближайшими соседними атомами в кубической решетке:

простой $d = a$;

гранецентрированной $d = \dfrac{a\sqrt{2}}{2}$;

объемноцентрированной $d = \dfrac{a\sqrt{3}}{2}$, где a – параметр (период) решетки.

Параметр кубической решетки a можно найти, зная некоторые макроскопические величины, характеризующие данный кристалл, такие как молярная масса кристалла M и плотность кристалла D. Молярный объем кристалла можно определить по формуле:

$$V_M = \frac{M}{D}. \qquad (1.4)$$

Объем $V_Я$ элементарной ячейки кристалла кубической структуры

$$V_Я = a^3, \qquad (1.5)$$

где a – параметр (период) решетки.

Объем атома рассчитаем по формуле

$$V_{\text{а}} = \frac{4}{3}\pi R_{\text{а}}^3, \tag{1.6}$$

где $R_{\text{а}}$ – радиус атома.

Число Z_M элементарных ячеек в одном моле кристалла

$$Z_M = \frac{N_A}{N}, \tag{1.7}$$

где N_A – число Авогадро; N – число атомов; приходящихся на одну элементарную ячейку.

Найдем объем элементарной ячейки кубической решетки, разделив молярный объем кристалла V_M (см. формулу (1.4)) на число элементарных ячеек в одном моле кристалла Z_M (см. формулу (1.7)). Тогда параметр кристаллической решетки определим по формуле:

$$a = \sqrt[3]{\frac{NM}{DN_A}}. \tag{1.8}$$

Параметр кубической решетки порядка 10^{-10} м.

Концентрация атомов в кристалле равна:

$$n = \frac{N}{V_{\text{я}}} = \frac{DN_A}{M}, \tag{1.9}$$

эта величина порядка 10^{28} м$^{-3}$.

Коэффициентом упаковки называют долю объема кристалла, занятую атомами. Очевидно, он больше нуля, но меньше единицы. Коэффициент упаковки для кубической решетки:

$$k_{\text{уп}} = \frac{V_{\text{а}} N}{V_{\text{я}}} = \frac{4\pi R_{\text{а}}^3 N}{3a^3}, \tag{1.10}$$

где $V_{\text{а}}$ – объем атома; $V_{\text{я}}$ – объем элементарной кубической ячейки; N – число атомов в ячейке.

В кристаллографии важным параметром является число ближайших соседей – *координационное число z*.

Полупроводниковые материалы обычно кристаллизуются в структуре алмаза, цинковой обманки кубического типа (сфалерит) и в алмазоподобной структуре гексагонального типа (вюрцит). Более подробно эти структуры и их свойства будут рассмотрены в III главе.

1.1.4. Виды химических связей

Химическая связь – это совокупность сил, действующих между атомами или группой атомов, в результате чего образуются молекулы и кристаллы. При обсуждении химической связи обычно пользуются термином валентность. Химическая связь осуществляется в основном за счет валентных электронов. Описать химическую связь – значит, выяснить, как распределяется электронная плотность в молекуле или кристалле.

Рассмотрим основные виды связей, за счет которых происходит объединение атомов в молекулы и кристаллы.

Ковалентная связь. *Ковалентная связь*, которую иногда называют *гомеополярной*, образуется за счет взаимодействия между двумя электронами, ранее принадлежавшими отдельным атомам, с образованием устойчивой электронной пары. Электроны в такой паре должны иметь противоположное направление спинов. *Спин* – собственный момент элементарной частицы, в данном случае, электрона. Спин является квантовой характеристикой элементарных частиц и не имеет аналогов в классической физике.

Электрон в атоме на самом деле не является локализованной частицей, а представляет собой облако с определенным распределением плотности. Орбите электрона в квантовой механике соответствует линия максимальной плотности электронного облака. Но многие явления в атомах удобно объяснять в виде движения электрона как частицы по орбите, не слишком искажая сущности явлений.

Электроны при движении по молекулярной орбите чаще всего находятся между ядрами, где создается избыток отрицательного заряда, что способствует сближению атомов (рис 1.7). Если двухатомная молекула состоит из атомов одного элемента (H_2, N_2, Cl_2), то электронная пара в одинаковой степени принадлежит обоим атомам. В подобном случае молекулу и существующую в ней ковалентную связь называют *неполярной* или *нейтральной*. В неполярных молекулах центры положительных или отрицательных зарядов совпадают, т. е. дипольный момент равен нулю.

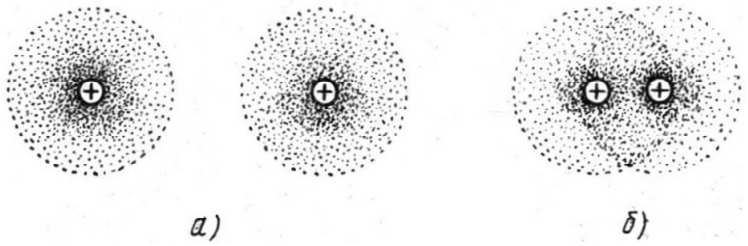


Рис. 1.7 – Схема ковалентной связи в молекуле водорода: *а*) распределение электронного облака в атоме водорода; *б*) распределение электронного облака в молекуле водорода

Если же двухатомная молекула состоит из атомов различных элементов, то электронная пара может быть смещена к одному из атомов. В таких случаях ковалентную связь называют *полярной*, а молекулы с полярной связью, у которых центры положительных и отрицательных зарядов не совпадают, – *полярными* или *дипольными*. Дипольная молекула характеризуется величиной электрического дипольного момента p, измеряемого в Кл·м

$$p = ql, \qquad (1.11)$$

где q – абсолютное значение заряда; l – расстояние между центрами положительного и отрицательного зарядов.

Многоатомные молекулы также могут быть неполярными – при симметричном расположении зарядов или полярными – при асимметричном

расположении. В качестве примера веществ с полярной ковалентной связью можно привести молекулы типа H_2O, H_2S, HCl, N_2O_5 и всех кислот.

Ковалентная связь является сильной связью. Например, энергия связи между двумя атомами углерода в кристалле алмаза составляет величину 7,3 эВ, что сравнимо с энергией связи в ионных кристаллах, хотя ковалентная связь образуется между нейтральными атомами. Вещества, образованные молекулами с ковалентной связью, характеризуются высокой твердостью. Они тугоплавки и химически устойчивы, что объясняется прочностью ковалентной связи.

Ковалентная связь является *пространственно ориентированной*. В кристаллах алмаза, кремния и германия со структурой алмаза каждый атом расположен в центре тетраэдра, образованными четырьмя ближайшими атомами. Углероду, кремнию и германию не хватает четырех электронов до заполнения их электронных оболочек для образования устойчивой восьмиэлектронной конфигурации, и поэтому эти атомы будут притягиваться друг к другу при обобществлении электронных оболочек.

В соединениях с ковалентной связью типа $A^{III}B^{V}$ и $A^{II}B^{VI}$ для различных атомов присутствует доля «ионности» в химических связях (на ковалентную связь приходится порядка 75%, на ионную – 25% общей энергии связи).

Ионная связь. *Ионная связь* вызывается силами электростатического притяжения между положительными и отрицательными ионами, когда один атом отдает электрон, а другой его принимает. В качестве примеров ионных соединений можно привести элементы, очень отличающиеся по свойствам, например, типичные *металлы* и типичные *неметаллы*. На рис. 1.8 показана ионная связь на примере ионного кристалла хлористого цезия CsCl. Молекулы вещества с ионной связью полярны. Ионные кристаллы состоят из положительных и отрицательных ионов. Эти ионы образуют кристаллическую решетку за счет того, что кулоновское притяжение разноименно заряженных ионов сильнее, чем кулоновское отталкивание одноименных ионов. Электронные оболочки образовавшихся ионов соответствуют стабильным электронным оболочкам, характерным для инертных газов. Например, нейтральные атомы лития и фтора имеют электронные конфигурации: Li – $1s^2 2s$, F – $1s^2 2s^2 2p^5$, в то же время в кристалле фтористого лития однократно заряженные ионы имеют электронные конфигурации, характерные соответственно для атомов гелия и неона: Li^+ – $1s^2$, F^- – $1s^2 2s^2 2p^6$.

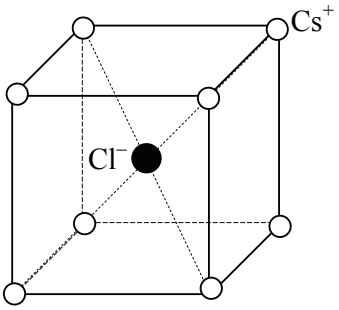


Рис. 1.8 – Кристалл хлорида цезия с ионной

Рассмотрим вклады в потенциальную энергию ионного кристалла по мере уменьшения их значимости:

- кулоновское притяжение ионов разного знака;
- кулоновское отталкивание ионов одного знака;

- квантовомеханическое отталкивание при перекрытии электронных облаков;

- ван-дер-ваальсовое притяжение между ионами.

Основной вклад в энергию связи ионных кристаллов дает электростатическая энергия (первые две строки в списке), наименьшее (~ 1…2%) – притяжение Ван-дер-Ваальса.

К веществам с ионной связью относится большинство солей и некоторые оксиды (Cu_2O, ZnO, CdO, Fe_2O_3, NiO и др.). Ионные соединения по механической прочности и химической стойкости уступают веществам, образованным с помощью ковалентной связи, так как эта связь обычно менее прочна, чем ковалентная. Это связано с тем, что ионная связь не имеет столь выраженной направленности, как ковалентная. С другой стороны, ионная связь и не столь «подвижна» как металлическая. Это приводит к тому, что ионные кристаллы не обладают пластичностью металлов.

Металлическая связь. Металлы характеризуются высокой электропроводностью, так как значительная часть электронов в них свободна. Обычно на атом приходится 1-2 свободных валентных электронов. Такие электроны называют *электронами проводимости*.

В металлах реализуется химическая связь, основанная на обобществлении валентных электронов, принадлежащих не двум, а практически всем атомам в кристалле. Атомы металлов отдают внешние электроны, превращаясь в положительные ионы. Таким образом, металл можно рассматривать как систему, построенную из положительных ионов, находящихся в среде свободных, коллективизированных электронов. Кристаллы металлов можно представить себе в виде правильно расположенных положительных ионов, погруженных в однородную электронную «жидкость».

Благодаря наличию свободных электронов металлы обладают высокой электро- и теплопроводностью. Высокая пластичность металлов объясняется отсутствием направленности металлической связи.

Размещение частиц в пространственной решетке осуществляется таким образом, чтобы обеспечить максимально возможную для данного типа кристалла энергию их связи. Так как энергия связи всегда отрицательна, то более сильная связь – большая по модулю. Для частиц сферической формы очень часто наиболее благоприятным оказывается такое их размещение, при котором каждая сфера находится в соприкосновении с наибольшим числом ближайших соседей. Подобные пространственные образования называются *структурами плотной упаковки*.

Металлы кристаллизуются в относительно плотно упакованные структуры с большими координационными числами (числом ближайших соседей): гранецентрированную кубическую, объемноцентрированную кубическую, гексагональную плотноупакованную. Следует отметить, что для модели плотноупакованных шаров коэффициент упаковки одинаков для ГЦК и ГПУ структур и равен максимальному значению 0,74. При этом их координационное число также одинаково и равно максимально возможному

значению – 12. Равенство этих параметров в способах расположения этих атомов говорит и о близости энергии связи. Действительно, ряд металлов может при определенных условиях менять свою структуру от ГЦК к ГПУ и наоборот.

Многие вещества, являющиеся при нормальных условиях полупроводниками или диэлектриками, при повышении давления испытывают фазовые переходы особого рода и приобретают металлические свойства; при этом и связь в них приобретает металлический характер. Действительно, принудительное сближение атомов усиливает перекрытие электронных оболочек, что способствует обобществлению электронов. Так, полупроводник Te (теллур) становиться металлом при давлениях порядка 4 ГПа, Ge (германий) – при 16 ГПа, InSb – при 2 ГПа.

Связи Ван-дер-Ваальса. Такие связи наблюдаются у ряда веществ между молекулами с ковалентным характером внутримолекулярного взаимодействия. Наличие межмолекулярного притяжения в этом случае возможно при согласованном движении валентных электронов в соседних молекулах. Тогда силы притяжения валентных электронов ядром соседней молекулы оказываются сильнее сил взаимного отталкивания электронных оболочек этих молекул. Вещества с межмолекулярным характером связи обладают низкими температурами плавления (например, у парафина $t_{пл}$ = 50…52 ^{0}C). Это свидетельствует о том, что связь Ван-дер-Ваальса очень слабая.

Виды кристаллических решеток. В зависимости от вида химической связи и частиц, образующих кристаллическую решетку, различают *атомные*, *ионные*, *металлические* и *молекулярные* кристаллические решетки.

В узлах *атомных решеток* находятся нейтральные атомы, которые связаны друг с другом ковалентной связью. Примерами кристаллов этого типа могут служить алмаз, кремний, германий, серое олово, а также кристаллы сложных веществ с высокой долей ковалентной связи, таких как кварц, карбид кремния, сульфид цинка, нитрид алюминия.

В узлах *ионных решеток* располагаются, чередуясь, положительные и отрицательные ионы. Примерами таких кристаллов могут быть NaCl, CaF_2, KF. Многие ионные кристаллы относятся к классу диэлектриков.

В узлах *металлических решеток* расположены положительные ионы, в промежутках между которыми находятся свободные электроны. Устойчивость кристалла обеспечивается силами притяжения между положительными ионами и электронным газом. Примерами таких кристаллов являются металлы и их сплавы. По энергии кристаллической решетки металлы занимают промежуточное положение между кристаллами с атомными решетками и молекулярными решетками. Плотность большинства металлов велика, что свидетельствует о плотной структуре металлических кристаллов. Высокая пластичность указывает на отсутствие жесткости в металлических решетках: их атомные плоскости относительно легко сдвигаются одна относительно другой.

В узлах *молекулярных* решеток находятся молекулы. Такие решетки образуются за счет ковалентных и ионных связей. Молекулы органических соединений (полимеров) могут иметь вытянутую форму в виде зигзагообразных скрученных цепочек, сетчатую или другую сложную пространственную форму.

1.1.5. Аморфные вещества

Аморфные вещества. Внутреннее строение аморфных веществ характеризуется беспорядочным расположением атомов и молекул, т.е. в них отсутствует кристаллическое строение. Они делятся на две большие группы: а) простые аморфные вещества – низкомолекулярные жидкости, неорганические стекла, плавленый кварц; б) высокополимерные соединения – каучуки, резины, органические стекла, смолы.

Свойства аморфных тел не зависят от направления, т.е. аморфные тела *изотропны*. Например, кристалл кварца, если его расплавить, после охлаждения образует так называемый плавленый кварц, обладающий свойствами, одинаковыми по всем направлениям, и при этом сильно отличающимися от свойств кристаллического кварца.

Некоторые вещества могут находиться как в кристаллическом состоянии, так и в аморфном. Так, оксид кремния SiO_2 (кремнезем) встречается в природе в виде кристаллов кварца, а также в аморфном состоянии в виде минерала – опала. При исследовании опала под электронным микроскопом выяснилось, что кремнезем в основном образует мелкие сферы (шарики), обычно имеющие около 0,2 мкм в диаметре. Эти сферы плотно упакованы и уложены в строгом геометрическом порядке, а промежутки между ними заполнены воздухом, водяным паром или водой.

Как говорилось выше, аморфное состояние вещества является менее устойчивым, чем кристаллическое, поэтому при определенных условиях возможен переход вещества из аморфного состояния в кристаллическое. Примером такого превращения служит расстеклование – самопроизвольная кристаллизация стекол при повышенных температурах или давлении. В связи с образованием мелких кристалликов стекла меняют свои свойства и превращаются, например, в *ситаллы*.

1.1.6. Анизотропия кристаллов

Свойства кристаллов (электрические, магнитные, механические, химические) зачастую зависят от кристаллографического направления. Такую неодинаковость физических свойств кристалла в разных направлениях называют *анизотропией*.

Степень анизотропии может быть значительной. Например, исследования монокристалла меди показали, что предел прочности на растяжение изменяется от 120 до 360 МПа, а относительное удлинение при растяжении – от 10 до 55% в зависимости от кристаллографического направления. Анизотропия физических свойств кристаллических тел проявляется тем сильнее, чем класс симметрии кристаллической решетки.

Твердые тела могут находиться и в *поликристаллическом* состоянии: они состоят из большого количества мелких кристаллитов (зерен), кристаллографические оси которых ориентированы в пространстве хаотично. Размеры зерен колеблются в диапазоне 1 … 1000 мкм. Размеры зерен в металлах влияют на их механические свойства: чем мельче зерна, тем пластичнее металл.

Свойства поликристаллического тела за счет усреднения свойств отдельных зерен, в целом, не зависят от направления. Такое состояние называют *квазиизотропией.* Однако с помощью специальной обработки могут быть получены текстурованные материалы с ориентированным расположением кристаллитов, в которых проявляется анизотропия. Эту особенность можно использовать, например, для изготовления текстурованной стали крупнокристаллического строения, ребра кристаллитов которой расположены параллельно направлению прокатки. Магнитные свойства стали в направлении прокатки выше, чем по другим направлениям.

1.1.7. Полиморфизм (аллотропия)

Многие вещества в зависимости от условий кристаллизации могут иметь различные формы кристаллической решетки. Способность вещества иметь различные формы кристаллической решетки называют *полиморфизмом* или *аллотропией*. Полиморфные превращения происходят главным образом под влиянием температуры, хотя замечено влияние и других факторов – изменения давления, наличия посторонних примесей. Сущность полиморфизма состоит в том, что при определенных температурах в твердом веществе возникают новые центры кристаллизации, в процессе роста которых образуется новая решетка.

Различные полиморфные формы вещества называют *модификациями* и обозначают начальными буквами греческого алфавита (α, β, χ, δ и т. д.). Полиморфные превращения происходят при определенных условиях в кобальте, железе, олове, титане, селене, углероде, фосфоре и др.

Классическим примером полиморфизма является низкотемпературное превращение олова. Известно, что обычное белое олово (β-Sn) при низких температурах может переходить в серое олово (α-Sn). Распространенная прежде оловянная посуда, хранившаяся зимой в плохо отапливаемых помещениях, внезапно рассыпàлась, превращаясь в серый порошок. Следов этого порошка в таких условиях достаточно, чтобы вызвать «заражение» и гибель остальной посуды. Это явление получило название «оловянной чумы».

Другой пример вещества, имеющего разные аллотропические модификации – углерод, существующий в виде графита и алмаза. Графит, имеющий гексагональную кристаллическую решетку, мягкий и хрупкий материал, хорошо проводит электрический ток. В обычных условиях графит является устойчивой модификацией. При определенных значениях температуры и давления графит можно перевести в алмаз. На этом принципе основано получение искусственных алмазов. Алмаз, кристаллическая

решетка которого состоит из двух гранецентрированных кристаллических решеток, сдвинутых относительно друг друга на 1/4 постоянной кристаллической решетки, является самым твердым веществом в природе.

1.1.8. Дефекты кристаллов

Приведенные на рис. 1.2 модели кристаллических решеток соответствуют идеальным кристаллам. Теория идеальных кристаллов позволяет объяснить многие свойства кристаллической решетки – плотность, удельную теплопроводность, упругие свойства. Все эти свойства называются *объемными* или *структурно-нечувствительными*. Однако реальные кристаллические тела не идеальны: в них в некотором количестве существуют нарушения структуры, называемые *несовершенствами* или *дефектами*. Некоторые свойства реальных кристаллических тел слабо зависят от этих несовершенств. Другие же свойства кристаллов, такие как прочность, величина и тип электропроводности в полупроводниках, гистерезисные потери в ферромагнитных материалах и т. д. сильно зависят от степени совершенства кристаллов и называются *структурно-чувствительными*.

Приведем классификацию дефектов в кристаллах.

Если дефекты рассматривать как n-мерные нарушения в кристаллической решетке, то они могут быть: ***точечными*** (n = 0); ***линейными*** (n = 1); ***плоскими*** (поверхностными) (n = 2); ***объемными*** (n = 3).

Кроме того, различные n-мерные дефекты при благоприятных условиях могут взаимодействовать друг с другом с образованием *сложных ассоциированных дефектов*.

Точечные (нульмерные) дефекты имеют размеры порядка диаметра атома. Основной причиной их возникновения является переход атомов за счет теплового движения из узлов кристаллической решетки в междоузлие. Образуются так называемые вакансии – пустые узлы основной решетки (матрицы) и междоузельные атомы самой матрицы. К точечным дефектам относятся также атомы примеси, которые могут располагаться как в узлах, так и в междоузлиях матрицы, а также посторонние атомы, адсорбированные на поверхности кристалла.

Точечные дефекты всегда присутствуют в кристаллах, и при росте температуры количество их увеличивается.

Линейные (одномерные) дефекты имеют малые размеры в двух измерениях и большую протяженность в третьем измерении. Такие несовершенства, представляющие собой нарушение кристаллической структуры вдоль некоторой линии, называют *дислокациями*.

К ***плоским (поверхностным) дефектам*** принадлежат границы зерен кристаллов (двойников), границы самого кристалла.

Объемные (трехмерные) дефекты кристалла имеют существенные размеры во всех трех измерениях. К объемным дефектам относятся пустоты, включения отдельных кристаллических зерен или иной кристаллической

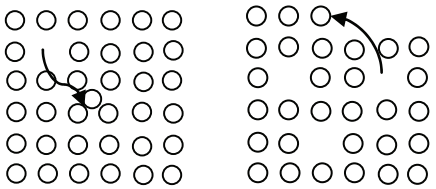
модификации. Такие дефекты возникают при нарушении условий в процессе роста кристалла.

Сложные дефекты менее изучены. Они могут возникать за счет взаимодействия атомов или ионов примеси друг с другом и с вакансиями. Поэтому такие дефекты еще называют *ассоциированными*.

Рассмотрим более подробно приведенные дефекты.

1) ***Точечные дефекты*** **(*n* = 0).** ***Вакансии.***

Различают два вида вакансий: вакансии по Френелю и вакансии по Шоттки. При образовании вакансий по Френелю, оторвавшийся атом занимает в решетке положение в междоузлии, иногда его также называют дислоцированным атомом (рис. 1.9, *а*). В плотноупакованных кристаллах размеры междоузлий обычно невелики и образование френкелевских вакансий маловероятно. Они могут образовываться в достаточно рыхлых кристаллах, таких, например, как кристаллы щелочных металлов, ионы которых очень малы по сравнению с размером междоузлия.



а) *б*)

Рис. 1.9 – Вакансии в кристалле:
а – вакансии по Френелю; *б* – вакансии по Шоттки

В полупроводниках более вероятным является образование вакансий по Шоттки (рис. 1.9, *б*), при котором атомы, покидающие свое место в решетке, уходят на какие-либо *стоки*. Стоками вакансий могут служить дислокации, границы блоков, поверхность кристалла.

Физическая природа образования вакансий связана с тепловыми колебаниями атомов. Амплитуда колебаний атомов при умеренных температурах может достигать нескольких процентов от межатомного расстояния, а в процессе роста температуры – намного больше.

Из общих соображений, очевидно, что концентрация вакансий обоих типов зависит от температуры и от энергии образования вакансий, которую называют *энергией дефектообразования W*. Обычно значения энергии дефектообразования велики и составляют 1…2 эВ для различных кристаллов. При комнатной температуре (300 К) концентрация образующихся вакансий обычно ничтожно мала: порядка 10^{-17} от концентрации атомов.

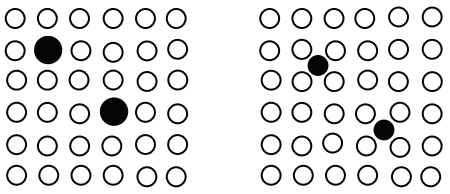
В процессе кристаллизации твердого тела с уменьшением температуры количество вакансий в кристалле уменьшается, но происходит это не мгновенно, поэтому для уменьшения количества вакансий охлаждение кристалла должно происходить достаточно медленно, чтобы не возник эффект закалки.

2) ***Точечные дефекты*** **(*n* = 0).** ***Примесные атомы.***

Из всех вышеперечисленных дефектов наиболее важным типом дефектов для управления свойствами полупроводникового материала являются атомы примесей.

Чужеродные примесные атомы с матрицей основного кристалла могут образовывать либо твердые растворы замещения, либо твердые растворы

внедрения. (Обычно это очень разбавленные растворы). В первом случае примесные атомы находятся в узлах кристаллической решетки, во втором – в междоузлиях (рис. 1.10). Та или другая ситуация зависит в основном от двух факторов: *геометрического* и *электрохимического*.



а) *б*)

Рис. 1.10 – Примеси в кристалле:
а – примеси замещения; *б* – примеси внедрения

Формулировка эмпирического правила для геометрического фактора принадлежит Юм-Розери, установившему, что *растворы замещения образуют лишь те атомы*, *радиусы которых отличаются от радиусов матричных атомов не более чем на* 15%. Чтобы применять это правило, необходимо выбрать определенный способ оценки атомных радиусов. Анализ большинства металлических соединений показывает, что наиболее удобно в качестве атомного радиуса брать половину наименьшего межатомного расстояния кристалла из чистого элемента. Преимущество такого определения очевидно: нет необходимости знать структуру соединений.

Однако при таком определении атомного радиуса часто возникают затруднения из-за того, что в кристалле из чистого элемента атомы могут быть в одном состоянии, а в соединении – в другом, например, за счет ионизации. Из-за того, что в полупроводниковых кристаллах преобладает ковалентный характер химической связи, в качестве геометрического фактора принято считать разность ковалентных радиусов атомов примеси и матрицы.

В растворах замещения важную роль играет электрохимический фактор. Для образования такого вида растворов необходимо, чтобы разность электроотрицательностей примесного и матричного атомов была мала. В противном случае один из компонентов будет слишком электроотрицателен, а это приведет скорее к образованию химического соединения, чем к твердому раствору. (Электроотрицательность характеризует способность атома в молекуле притягивать валентные электроны).

Исследование твердых растворов внедрения было проведено американским ученым Хаггом. Им было показано, что твердые растворы внедрения подчиняются правилу: *отношение радиуса внедряющегося атома к радиусу атома матрицы должно быть меньше*, *чем* 0,59, т.е. разница в диаметрах должна быть больше, чем 41%, от диаметра большего из атомов. Значение 0,59 соответствует наибольшей сфере, которая может уместиться в промежутках плотно упакованной решетки.

Таким образом, если отношение $r_п/r_м$ ($r_п$ – радиус примесного атома; $r_м$ – радиус матричного атома) заключено в пределе $0< r_п/r_м< 0{,}59$, то примесь образует раствор внедрения, а если $0{,}85< r_п/r_м< 1{,}15$ – раствор замещения,

область же $0{,}59 < r_п/r_м < 0{,}85$ вообще неблагоприятна для образования твердых растворов.

3) ***Точечные дефекты* (*n* = 0). *Поверхностные уровни*.**

На поверхности кристалла граничные атомы имеют незавершенные химические связи, это означает, что поверхностные атомы основного кристалла ведут себя иначе, чем атомы в его объеме. Существование самой поверхности, как было указано И.Е. Таммом, уже является нарушением внутреннего периодического поля. И как при всяком нарушении периодичности, в этом случае должны возникать локальные уровни внутри запрещенной зоны полупроводника. Характер этих уровней может быть как акцепторным, так и донорным. Отличительной чертой этих уровней является их локализация на поверхности раздела полупроводника с вакуумом или газом. Такие поверхностные уровни называют *таммовскими уровнями*.

Плотность таммовских уровней порядка плотности атомов на 1 $м^2$ поверхности. Так как параметр кристаллической решетки составляет примерно $3 \cdot 10^{-10}$ м, то плотность таммовских уровней составляет около 10^{19} $м^{-2}$. Роль этих уровней велика, когда их количество станет сравнимо с количеством объемных уровней. Это возможно в тех случаях, когда объемные процессы, определяющие работу того или иного полупроводникового устройства, происходят в приповерхностных слоях толщиной порядка 10^{-8} м.

4) ***Точечные дефекты* (*n* = 0). *Отклонение от стехиометрического состава*.**

Своеобразным видом точечных дефектов является отклонение состава вещества от строго *стехиометрического* состава, характерное для многих химических соединений. Его можно представит как частичную потерю соединением одной из своих более летучих составляющих, например испарением при повышенной температуре, что и имеет место в сульфиде кадмия. Отклонение от точно стехиометрического состава резко меняет электропроводность кристаллов, осложняет технологию их получения.

5) ***Линейные дефекты* (*n* = 1). *Дислокации*.**

Дислокации представляют собой нарушение правильной структуры вдоль некоторой линии в кристалле. Эта линия не обязательно является прямой и не обязательно лежит в одной плоскости, хотя наиболее часто встречаются дислокации, линии которых расположены в одной плоскости. Линия дислокации либо выходит на поверхность кристалла, либо разветвляется на другие дислокации, либо образует замкнутую петлю.

Основная геометрическая характеристика дислокации – вектор сдвига $\vec{b}$, называемый *вектором Бюргерса*. Вектор Бюргерса характеризует величину и направление сдвига атомных плоскостей, приводящих к образованию дислокаций. В зависимости от угла сдвига φ различают *краевые* (φ = 90º), *винтовые* (φ = 0º) и *смешанные* (φ произвольно) дислокации.

Краевые дислокации возникают за счет удаления из кристалла одной полуплоскости или образования в кристалле дополнительной полуплоскости,

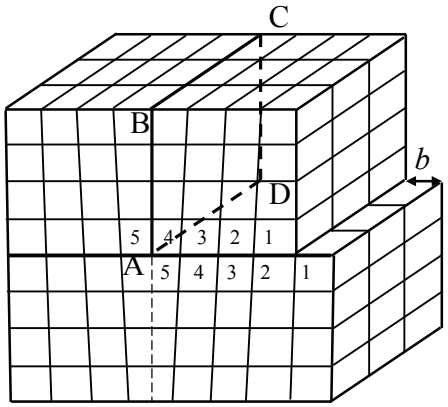


Рис. 1.11 – Краевые дислокации в кристаллах

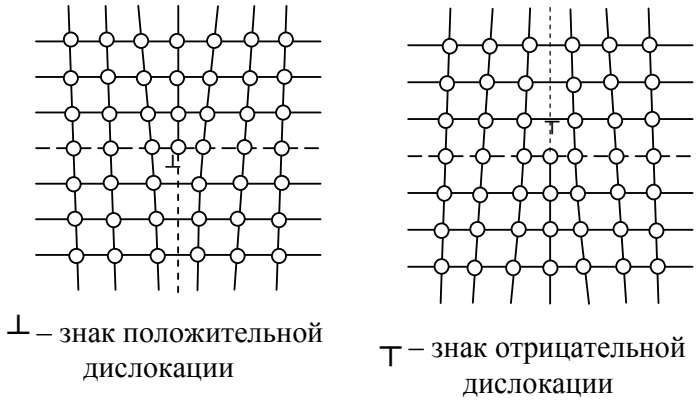


Рис. 1.12 – Краевые дислокации в кристаллах: а) – положительные; *б*) - отрицательные

когда линия дислокации перпендикулярна направлению вектора сдвига $\vec{b}$. Для получения краевой дислокации сделаем мысленно неполный надрез, перпендикулярный плоскости ABCD, и сдвинем друг относительно друга по надрезу верхнюю и нижнюю половины кристалла на период решетки. Во всех частях кристалла все атомы решетки не изменяют своих положений, кроме области, примыкающей к линии AD. Теперь атомы верхней полуплоскости 1 расположатся над атомами нижней полуплоскости 2; атомы верхней полуплоскости 2 – над атомами нижней 3 и т.д. (рис. 1.11). В области же линии AD образуется своеобразная складка – это и есть линия дислокации. Краевые дислокации условно подразделяются на *положительные* и *отрицательные* дислокации. Положительная дислокация соответствует случаю, когда сверху есть лишняя атомная полуплоскость (рис. 1.12, *а*), соответственно в верхней половине кристалла действует сжимающее напряжение, а в нижней – растягивающее. Отрицательные дислокации соответствуют случаю, когда верхняя половина кристалла растянута, а нижняя сжата (рис. 1.12, *б*). Эти дислокации различаются лишь поворотом на 180^{o}. Знак дислокации существенен, если рядом есть другая дислокация. Силы упругого взаимодействия между дислокациями зависят от знака дислокации: одноименные – отталкиваются, разноименные – притягиваются. Под действием внешнего напряжения дислокации в кристаллах могут двигаться. Разноименные дислокации, двигающиеся в одной плоскости, при встрече уничтожают друг друга, в результате чего восстанавливается

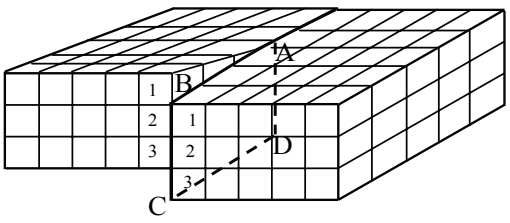


Рис. 1.13 – Винтовые дислокации в кристаллах

целостность решетки.

Причиной возникновения винтовых дислокаций является смещение части кристаллической решетки не менее чем на один период параллельно вектору сдвига $\vec{b}$ (рис. 1.13).

В реальных кристаллах присутствуют дислокации смешанных форм, которые могут быть разложены на краевую и винтовую компоненты.

Образование дислокаций происходит за счет термических и механических напряжений в процессе кристаллизации, при пластической деформации, термической обработке и др.

Дислокации существенно ухудшают физические свойства веществ. Например, в металлах наличие дислокаций резко снижает механическую прочность, в полупроводниках значительно увеличивает неконтролируемую проводимость. В местах выхода дислокаций на поверхность кристалла изменяется химическая активность вещества.

6) ***Поверхностные* (*двухмерные*) *дефекты* (*n* = 2).**

Поверхностные (*двухмерные*) *дефекты* малы только в одном измерении, они представляют собой поверхности раздела между отдельными зернами или блоками вещества (рис. 1.14). Отдельные зерна обычно разориентированы относительно друг друга на углы, достигающие нескольких десятков градусов, что приводит к образованию по границам зерен дислокаций и вакансий. Поверхностные дефекты бывают двух типов: *наружные* и *внутренние* (поверхности раздела между кристаллитами).

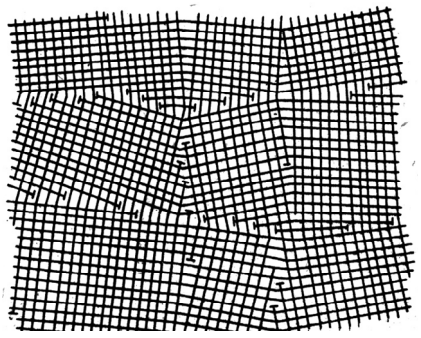


Рис. 1.14 – Схема блочной структуры кристалла

Важным видом поверхностных дефектов является сама поверхность кристалла, представляющая собой естественное нарушение его кристаллической структуры. Свойства поверхностных слоев кристалла существенно отличаются от его объемных свойств.

7) ***Сложные дефекты***.

К *сложным дефектам* относятся различные ассоциации точечных дефектов. Примерами могут служить комплексы вакансия-примесь, которые образуются при облучении полупроводников различным корпускулярным излучением (электроны, протоны, нейтроны). При облучении происходит выбивание атомов из узлов решетки в междоузлия с образованием вакансий.

В большинстве случаев после облучения со временем происходит обратный переход (*отжиг дефектов*). Однако при определенных условиях дефекты живут долго и способны объединяться с примесными атомами в ассоциаты. Подобные комплексы образуют локальные уровни в запрещенной зоне полупроводника, как правило, глубокие.

Для получения совершенных монокристаллических тел с минимальным количеством дефектов необходимо тщательно контролировать процесс кристаллизации и избегать попадания посторонних примесей. Однако, как

показывает развитие науки, многие отрицательные физические явления рано или поздно приносят пользу. Не исключено, что образование сложных и плоских дефектов может быть использовано для придания материалам новых необходимых свойств.

Поиск материалов с новыми свойствами зачастую основан на модификации либо радикальном изменении расположения атомов в веществе. Это связано со значительными изменениями в структуре и симметрии по сравнению с идеальными кристаллами. Последние достижения технологии привели к возможности построения твердого тела буквально по атому, так что можно создавать вещества в твердой фазе, которые состоят из моноатомных слоев, чередующихся в заданном исследователями порядке – квантовые сверхструктуры (сверхрешетки).

1.2. ЭЛЕМЕНТЫ ЗОННОЙ ТЕОРИИ ТВЕРДЫХ ТЕЛ

1.2.1. Зонные диаграммы твердых тел

Все тела в зависимости от их электрических свойств условно подразделяют на проводники, полупроводники и диэлектрики.

Численно решить задачу о нахождении возможных значений энергии в кристалле можно только с помощью квантового уравнения Шредингера. Однако в квантовой механике, так же как и в классической, отсутствуют методы точного решения динамической задачи для системы многих частиц. Поэтому эта задача решается приближенно сведением задачи многих частиц к одноэлектронной задаче – задаче об одном электроне, движущимся в периодическом потенциальном поле, результатом решения которой является появление зонного энергетического спектра.

Мы рассмотрим упрощенную качественную модель энергетических зон, опуская сложные теоретические вопросы квантовой механики и физики твердого тела. Тем не менее, такое рассмотрение позволяет с единой точки зрения объяснить: а) деление твердых кристаллических тел на металлы, полупроводники и диэлектрики; б) их основные свойства.

Любая система спонтанно принимает состояние с минимальной энергией, т.е. все электроны стремятся занять наиболее низкие энергетические свободные состояния. В основе заполнения атомных уровней и энергетических зон лежит *принцип Паули, согласно которому в одном энергетическом состоянии* (*на одном уровне*) *может находиться не более двух электронов с противоположными спинами*.

Электроны в несвязанных друг с другом атомах имеют только дискретные значения энергии, отделенные друг от друга широкими запрещенными интервалами. Рассмотрим теперь, какие изменения претерпевает энергетический спектр электрона при образовании кристалла из отдельных атомов. При сближении N одинаковых атомов, энергетические уровни отдельных атомов, согласно принципу Паули расплываются в энергетическую зону, состоящую из N близко расположенных уровней. Из каждого дискретного уровня обычно образуется одна энергетическая зона шириной в единицы электрон-вольт, включающая столько состояний,

сколько атомов в кристалле. Так, например, в кристалле размерами 0,1·0,1·0,1 мм содержится $\sim 10^{17}$ атомов, т.е. фактически зона представляет собой сплошную совокупность состояний – континуум.

В нейтральном изолированном атоме низшие энергетические уровни полностью заполнены электронами. Следовательно, энергетические зоны в кристаллах, образованные в результате расщепления этих атомных уровней, также будут полностью заполнены электронами. Во всех глубоко лежащих зонах процесс переноса заряда невозможен, т.к. в этих зонах все состояния заняты. Интерес с точки зрения проводимости и многих других физических процессов представляют лишь верхние зоны, а рассмотрение глубоко лежащих зон можно исключить.

Для изображения энергетической структуры кристалла обычно пользуются упрощенной энергетической схемой. Так как физические процессы в кристалле (электрические, магнитные, оптические и др.) обусловлены состоянием валентных электронов, то на схеме изображают только две разрешенные зоны: *валентную зону* и *зону проводимости*, которые могут быть разделены зоной запрещенных энергий (*запрещенной зоной*). Самая верхняя зона, полностью заполненная электронами при температуре $T = 0$ К, называется *валентной зоной*. Она возникает из атомарного уровня, на котором находятся валентные электроны в основном состоянии атома. Более высокие разрешенные зоны будут свободны. Ближайшая пустая или частично заполненная электронами зона при температуре $T = 0$ К называется *зоной проводимости*.

В зависимости от степени заполнения валентной зоны электронами и ширины запрещенной зоны возможны три случая (см. рис. 15, *а*, *б*, *в*).

Если валентная зона заполнена электронами не полностью (рис. 15, *а*), то электрон, получив сколь угодно малую энергию (например, за счет теплового движения или электрического поля), сможет перейти на более высокий энергетический уровень той же зоны, т. е. стать свободным и участвовать в проводимости. Внутризонный переход вполне возможен, так как, например, при 1К энергия теплового движения $kT \approx 10^{-4}$ эВ, т.е. гораздо больше разности энергий между соседними состояниями. Кристалл, имеющий такую зонную диаграмму, является металлом. Типичными представителями веществ, у которых наблюдается частичное заполнение валентной зоны (в случае металла ее называют также зоной проводимости) являются одновалентные щелочные металлы и благородные металлы, такие как Cu, Ag, Au, Pt и др. Таким образом, у одновалентных металлов валентная зона (зона проводимости) при $T = 0$ К заполнена только в нижней части до некоторого энергетического уровня, называемого уровнем Ферми E_F (рис. 15, *а*). Энергия электронов, находящихся на этом уровне, называется энергией Ферми.

Твердое тело является проводником электрического тока и в том случае, когда валентная зона полностью заполнена электронами, но она перекрывается со свободной зоной (зоной проводимости), не занятой электронами (рис. 15, *б*). В данном случае образуется так называемая

«гибридная» зона, которая заполняется валентными электронами лишь частично. Это имеет место для двухвалентных щелочноземельных элементов (Be, Mg, Ca и т.д.).

Рассмотрим случай, когда отдельная валентная зона полностью занята электронами (рис. 15, *в*). Для того, чтобы перебросить электрон из валентной зоны в зону проводимости, необходимо сообщить ему энергию, не меньшую, чем ширина запрещенной зоны E_g. Электрическое поле, не приводящее к пробою кристалла, сообщить электрону такую энергию не может. Если ширина запрещенной зоны невелика ~ 1эВ, то энергии теплового движения при реальных температурах оказывается достаточно, чтобы перевести часть электронов из валентной зоны в зону проводимости. Такие вещества называются полупроводниками. К ним относятся такие элементы как германий (Ge), кремний (Si), селен (Se) и др.

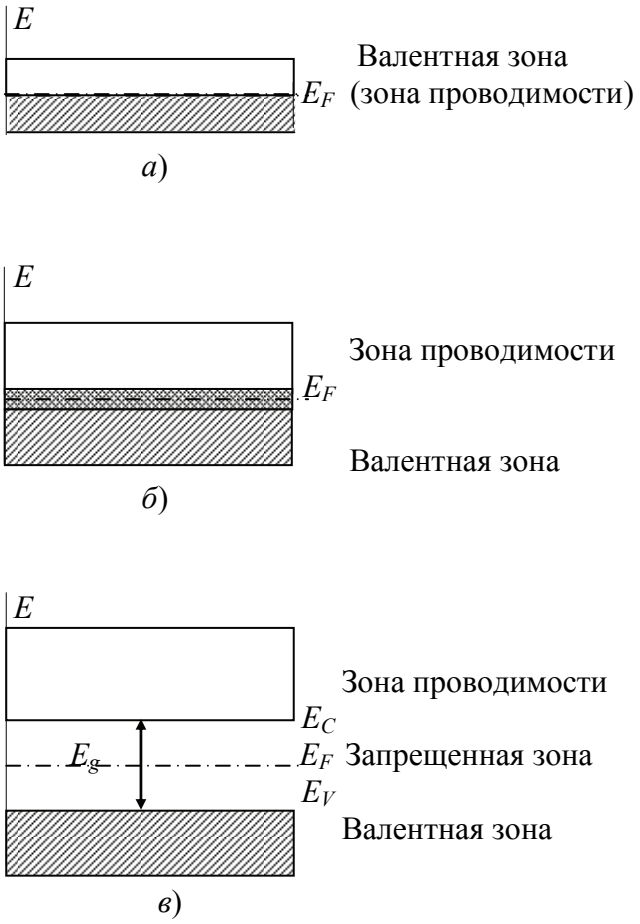


Рис. 1.15 – Зонные диаграммы одновалентных металлов (*а*), двухвалентных металлов (*б*) и полупроводников (*в*); E_V – потолок валентной зоны, E_C – дно зоны проводимости, E_g – ширина запрещенной зоны, E_F – уровень Ферми

1.2.2. Ширина запрещенной зоны

Шириной запрещенной зоны называется ширина энергетической щели, разделяющая зону проводимости и валентную зону. Это очень важный параметр полупроводников и диэлектриков. Запрещенная зона E_g равна энергии, которую должен получить валентный электрон, чтобы перейти из валентной зоны в зону проводимости. Электроны не могут обладать энергиями, заключенными в интервалах запрещенных зон.

Различие между диэлектриками и полупроводниками условно определяется шириной запрещенной зоны: если $0 < E_g \le 3$ эВ, то такие вещества являются полупроводниками, если $E_g > 3$ эВ – диэлектриками (изоляторами).

Ширина запрещенной зоны в полупроводниках зависит от температуры. В первом приближении функцию $E_g(T)$ можно представить в виде

$$E_g = E_{g0} - \alpha T, \qquad (1.12)$$

где E_g – ширина запрещенной зоны при температуре T, К; E_{g0} – ширина запрещенной зоны при температуре 0 К; α – температурный коэффициент. Величина α у полупроводников приблизительно $10^{-5}…10^{-4}$ эВ/К. Таким образом, ширина запрещенной зоны E_g с ростом температуры уменьшатся.

Такая температурная зависимость объясняется тем, что с повышением температуры усиливаются тепловые колебания узлов решетки; соответственно, сильнее перекрываются электронные оболочки соседних атомов. Это приводит к уширению разрешенных энергетических зон, а ширина запрещенной зоны уменьшается. Исключение из отмеченной закономерности составляют халькогениды свинца (PbS, PbSe, PbTe – полупроводниковые соединения типа $A^{IV}B^{VI}$), у которых с повышением температуры наблюдается увеличение ширины запрещенной зоны.

1.2.3. Эффективная масса

Как отмечалось, проводимость в полупроводнике возникает в результате того, что часть электронов перебрасывается из валентной зоны в зону проводимости. Валентная зона оказывается не полностью заполненной, и в ней становится возможным определенное перемещение электронов, т.е. повышение их энергии при приложении электрического поля.

Привычным является описание движения тел в рамках классической механики, которая устанавливает связь между внешней силой F, приложенной к телу, и его ускорением $a = \frac{d\upsilon}{dt}$:

$$F = m\frac{d\upsilon}{dt}. \qquad (1.13)$$

Прохождение тока через рассматриваемый кристалл обусловлено одновременным перемещением электронов и в зоне проводимости, и в валентной зоне. Такое перемещение должно отличаться от движения действительно свободных электронов в вакууме, так как в кристалле на электроны действует еще и внутреннее периодическое потенциальное поле, явившееся следствием действия на электрон многих внутренних сил. Эти силы также должны учитываться величиной F. Рассчитать истинную траекторию электрона в кристалле и решить поставленную задачу в рамках классической теории не представляется возможным. Используя же квантовые представления, можно получить зависимость силы от ускорения. Запишем это выражение в форме, аналогичной закону Ньютона:

$$F = \frac{h^2}{d^2E/dk^2}\frac{d\upsilon}{dt}. \qquad (1.14)$$

Сравнивая (1.13) с законом Ньютона для обычных тел (1.14), можно увидеть аналогию, если величину

$$m^* = \frac{h^2}{d^2E/dk^2} \qquad (1.15)$$

интерпретировать как аналог массы электрона. Величину m^* называют *эффективной массой*. Смысл введения эффективной массы заключается в том, что с ее помощью учитывают совместное действие периодического потенциального поля и внешней силы на электрон в кристалле. Посредством введения эффективной массы удалось сложные законы движения электронов в кристалле свести к законам, которые по форме совпадают с законами классической механики.

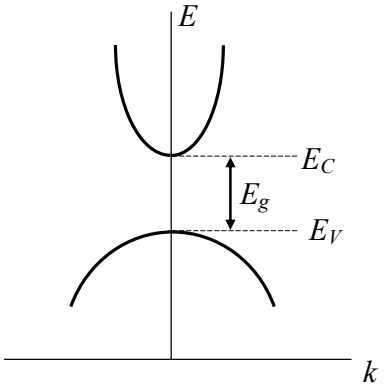


Рис.1.16 – Энергетический спектр электронов и дырок в координатах E–k ($m_p^* > m_n^*$)

Следует понимать, что эффективная масса имеет мало общего с обычной массой, т.е. она не является мерой инерции и не связана с силами тяготения. Более того, из рис. 1.16 видно, что энергия E зависит от волнового числа k ($k = \frac{p}{\hbar}$, где p – квазиимпульс электрона), а значит, d^2E/dk^2 может быть переменной величиной, т. е. эффективная масса может быть переменной в отличие от обычной истинной массы электрона $m_0 = 9{,}11 \cdot 10^{-31}$ кг.

Таким образом, эффективная масса есть коэффициент пропорциональности в законе, связывающем внешнюю силу, действующую на электрон в кристалле, с его ускорением.

Согласно квантовой теории твердого тела поведение электронов, которые находятся в нижней части зоны проводимости, представляется как движение отрицательно заряженных частиц с *положительной* массой ($d^2E/dk^2 > 0$, следовательно, и $m^* > 0$), которая отличается от массы свободного электрона. Движение электронов, которые перемещаются в верхней части валентной зоны, ускоряется электрическим полем в направлении обратном ускорению электронов, которые движутся в нижней части зоны проводимости; то есть следует считать, что масса этих электронов *отрицательна* ($d^2E/dk^2 < 0$, следовательно, и $m^* < 0$). Для более наглядного описания такого движения было введено понятие положительной квазичастицы – дырки, перемещение которой в электрическом поле происходит в противоположном направлении по сравнению с электроном. Эффективная масса такой положительной дырки будет положительна.

Конечно, истинными носителями тока в металлах и полупроводниках являются реальные электроны, а не формально введенные дырки. Никаких дырок, как реально существующих частиц, в действительности нет. Однако представление о дырках оказалось весьма плодотворным по следующим соображениям. Классическими законами для движения электронов можно пользоваться только в тех случаях, когда концентрация этих частиц в соответствующей зоне мала. Этому условию удовлетворяют электроны в зоне проводимости полупроводника. Но в валентной зоне мы имеем дело с

противоположным случаем. Там почти все состояния заполнены электронами, зато мала концентрация дырок. Здесь классические представления о движении электронов не применимы, описывать движение таких электронов необходимо, пользуясь квантовыми законами. Это значительно усложняет описание движения таких частиц. Однако введение понятия дырок приводит к поразительному результату. Оказывается, что в электрическом поле дырки движутся так, как двигались бы при классическом рассмотрении положительно заряженные частицы, обладающие определенной массой.

Эффективная масса носителей заряда обратно пропорциональна кривизне той зоны в *k*-пространстве ($m^* = \frac{h^2}{d^2E/dk^2}$), где они находятся.

Эффективная масса носителей заряда чаще всего выражается в долях массы покоя электрона m_0. Например: $m_p^* = 1{,}01m_0$ или $m_n^* = 0{,}85m_0$.

Электроны располагаются у дна зоны проводимости, и увеличению их энергии соответствует перемещение вверх. Дырки же располагаются у потолка валентной зоны, и увеличению их энергии соответствует перемещение вниз по энергетической шкале. Иными словами, для дырок и электронов отсчет энергий идет в противоположных направлениях.

1.2.4. Статистика электронов

При рассмотрении системы многих частиц предполагается, что они обладают какими-то признаками, которые позволяют отличать их друг от друга, хотя частицы и принимались совершенно одинаковыми. В связи с этим при подсчете микросостояний два микросостояния, которые отличаются тем, что две частицы поменялись местами, рассматривались как различные. Такая модель *различимых* частиц называется моделью Максвелла-Больцмана, а получающаяся при этом статистическая теория – статистикой Максвелла-Больцмана

$$f(E) = e^{-\frac{E-E_F}{kT}}, \qquad (1.15)$$

где k – постоянная Больцмана; T – абсолютная температура; E_F – уровень Ферми (химический потенциал).

Неразличимость частиц. Признаки, по которым можно было бы отличить одну частицу от другой, отсутствуют, если частицы совершенно идентичны. Представим себе две совершенно одинаковые частицы в некоторых состояниях. Тогда ясно, что в физической ситуации ничего не изменится, если частицы поменять местами. Если взять два электрона, то их неразличимость очевидна, потому что отпадают соображения о возможных различиях в их внутреннем состоянии. Если принимать частицы *неразличимыми*, то получаются другие правила подсчета микроскопических состояний, чем в модели Максвелла-Больцмана.

Модели Бозе-Эйнштейна и Ферми-Дирака. Модели, в которых частицы рассматриваются как неразличимые, называются моделями Бозе-

Эйнштейна и Ферми-Дирака, а статистические теории – статистиками Бозе-Эйнштейна и Ферми-Дирака.

Между собой эти модели различаются по поведению частиц в отношении микросостояний. Если в данном состоянии может находиться не более одной частицы, то такая модель называется моделью Ферми-Дирака, а если может находиться сколько угодно частиц – то моделью Бозе-Эйнштейна. Подчеркнем, что состояние характеризуется не только своей энергией, но и другими параметрами. Например, состояния с одинаковой энергией, но отличающиеся направлением импульса частицы, являются разными. Более точная формулировка гласит: в модели Бозе-Эйнштейна в каждом квантовом состоянии может находиться любое число частиц, а в модели Ферми-Дирака – не больше одной. Обобщая вышесказанное, можно сказать, что в статистике Ферми-Дирака имеется сильная «конкуренция» при занятии состояния: если состояние занято какой-то частицей, то другая занять это состояние не может. В определенном смысле, частица, занимающая некоторое состояние, «отталкивает» от этого состояния другие частицы. В статистике Бозе-Эйнштейна такая «конкуренция» полностью отсутствует: частица может занять некоторое состояние независимо от того, занято оно другими частицами или свободно.

Как было показано Паули, частицы с целым спином (фотоны) подчиняются статистике Бозе-Эйнштейна

$$f(E)=\frac{1}{e^{\frac{E-E_F}{kT}}-1}, \tag{1.16}$$

а с полуцелым спином (электроны) – статистике Ферми-Дирака

$$f(E)=\frac{1}{e^{\frac{E-E_F}{kT}}+1}. \tag{1.17}$$

Здесь $f(E)$ – вероятность заполнения данного состояния электронами. Нет микрочастиц, подчиняющихся статистике Максвелла-Больцмана, и, тем не менее, эта статистика правильно описывает поведение частиц в большом числе практически важных случаев, с которыми наиболее часто приходится встречаться. Это обусловлено тем, что статистики Бозе-Эйнштейна и Ферми-Дирака переходят в статистику Максвелла-Больцмана при условии, что число доступных для частиц состояний значительно больше, чем число частиц, которые могли бы занять эти состояния.

Особо отметим, что в предельных случаях речь идет о совпадении формул, а отнюдь не о том, что изменяется квантовое поведение частиц.

Вид распределения Ферми-Дирака для $T = 0$ К показан на рис. 1.16. Такое распределение обусловливается необходимостью двух требований: минимальности полной энергии и соблюдения принципа Паули. Поэтому электроны начинают заполнять квантовые состояния с самого нижнего энергетического уровня, последовательно занимая квантовые состояния, причем каждое из них лишь одним электроном. После того, как электроны оказались исчерпанными, заполнение прекращается. Последний электрон

занимает энергетический уровень с максимальной энергией. Этот энергетический уровень называется *уровнем Ферми* (*химическим потенциалом*), а энергия уровня – *энергией Ферми*. Однако такое наглядное

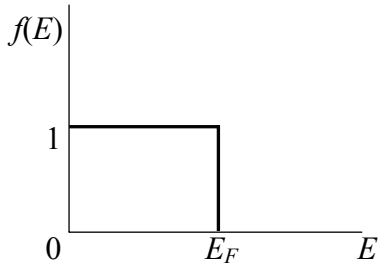


Рис. 1.16 – Распределение Ферми-Дирака при температуре $T = 0$ К

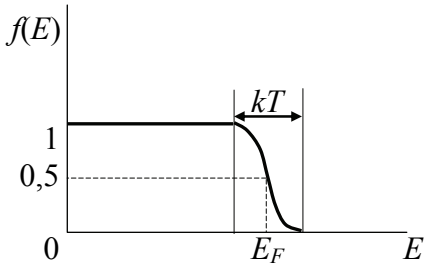


Рис. 1.17 – Распределение Ферми-Дирака при температуре $0 < T < T_F$, К

определение имеет смысл лишь в применении к свободным электронам в металле. В общем случае такое определение не совсем точно. Например, в полупроводниках энергия Ферми приходится примерно в средине запрещенной зоны, где заведомо нет электронов, которые обладали бы такой энергией. Поэтому более общее определение гласит: *энергией Ферми называется такая энергия, при которой функция распределения f*(*E*) *Ферми-Дирака принимает значение* ½.

При $T > 0$ распределение Ферми-Дирака перестает быть ступенчатым и размывается в окрестности уровня Ферми (рис. 1.17). Причина размывания – взаимодействие электронов с тепловым движением атомов. Поскольку средняя энергия теплового движения атомов имеет порядок kT, то и область размывания энергий электронов вблизи уровня Ферми также имеет порядок kT. Вместо скачкообразного уменьшения $f(E)$ при $T = 0$ К происходит плавное уменьшение $f(E)$. Например, переход от $f(E) = 0{,}9$ к $f(E) = 0{,}1$ происходит при увеличении энергии на $3{,}9\ kT$.

ВОПРОСЫ.

1. Что такое симметрия твердого тела?
2. Какие порядки осей вращения возможны у кристаллической решетки? Каким фигурам, покрывающим плоскость без промежутков, эти оси соответствуют?
3. Что такое точечные и пространственные группы симметрии?
4. Сколько типов кристаллических систем (сингоний) имеется? Сколько имеется типов решеток Бравэ?
5. Что такое индексы Миллера?
6. Что такое коэффициент упаковки и координационное число?
7. Каковы основные виды химической связи в материалах и чем они обусловлены?
8. В чем различие между монокристаллами, поликристаллами и аморфными веществами?

9. Охарактеризуйте явление полиморфизма. Приведите примеры полиморфных веществ.

10. Приведите примеры точечных и протяженных дефектов структуры в реальных кристаллах.

11. В чем сущность геометрического и электрохимического факторов, определяющих образование растворов замещения или внедрения примесных атомов в кристалл?

12. Что такое уровни Тамма? Какова их природа?

13. Почему при образовании твердого тела энергетические уровни атомов расщепляются в энергетические зоны? Чем различаются зонные структуры проводника (металла), полупроводника и диэлектрика?

14. От чего зависит ширина запрещенной зоны и число энергетических уровней в ней?

15. Как температура влияет на ширину запрещенной зоны?

16. В чем заключается физический смысл понятия эффективной массы?

17. Чем отличаются друг от друга статистики Бозе-Эйнштейна и Ферми-Дирака? При каких условиях эти статистики переходят в статистику Максвелла-Больмана?

18. Могут ли электроны находиться на уровне Ферми? В каких веществах?

ЗАДАЧИ.

1. Определить, сколько атомов приходится на одну элементарную ячейку в кристаллах с простой, объемноцентрированной и гранецентрированной кубической решеткой. (Ответ: 1; 2; 4).

2. Найти коэффициент упаковки в простой, объемноцентрированной и гранецентрированной кубических решетках, состоящих из атомов одного вида. (Ответ: 0,52; 0,68; 0,74).

3*. Вычислить относительное изменение объема материала при переходе железа из гранецентрированной кубической решетки в объемноцентрированную кубическую решетку, если межатомные расстояния в этих структурах соответственно равны 0,254 и 0,248 нм. (Ответ: $\Delta V/V = 1{,}6\%$).

4. Известно, что алюминий кристаллизуется в решетке гранецентрированного куба с параметром решетки a = 0,4041 нм. Вычислить концентрацию свободных электронов, считая, что на каждый атом кристаллической решетки приходятся три электрона. (Ответ: $n = 1{,}818 \cdot 10^{28}$ м$^{-3}$)

5. Вычислить параметр кристаллической решетки меди, если ее плотность D = 8920 кг/м^3, а элементарная ячейка представляет собой гранецентрированный куб. Определить объем, приходящийся на один атом. (Ответ: a = 0,372 нм; $V_a = 5{,}15 \cdot 10^{-29}$ м^3).

6*. Определить углы между следующими кристаллографическими направлениями в кубическом кристалле: а) [001] и [111]; б) [011] и [101]; в) [011] и [111]. (Ответ: а) $\theta = 54{,}74°$; б) $\theta = 60°$; в) $\theta = 35{,}26°$)

ГЛАВА II. ПРОВОДНИКИ

2.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ПРОВОДНИКАХ

К проводникам относятся вещества, основным свойством которых является высокая *электрическая проводимость* (в дальнейшем, *проводимость*). Проводниками электрического тока могут служить твердые тела, жидкости, а при соответствующих условиях, и газы.

Рассмотрим более подробно физические свойства твердых проводников, а именно, металлов и сплавов металлов. Это связано с тем, что металлы и их сплавы являются важнейшими проводниковыми материалами, широко применяемыми в современной радиоэлектронике и электротехнике.

2.1.1. Основы классической электронной теории металлов

Металлы занимают особое положение в физике твердого тела, обнаруживая ряд особых свойств, отсутствующих у других твердых тел. Все они прекрасные проводники электричества и тепла, обладают ковкостью и пластичностью. Химические элементы явно предпочитают металлическое состояние: более двух третей из них – металлы. Даже для объяснения свойств неметаллов необходимо понять свойства металла: лишь объяснив, почему медь такой хороший проводник, мы узнаем, почему им не является обычная поваренная соль.

Существование свободных электронов в металлах можно объяснить следующим образом. При образовании кристаллической решетки металла в результате сближения изолированных атомов, валентные электроны, сравнительно слабо связанные с атомными ядрами, отрываются от атомов металла и становятся «свободными», т.е. могут перемещаться по всему объему. Таким образом, в узлах кристаллической решетки располагаются ионы металла, между которыми хаотически движутся свободные электроны, образуя своеобразный «электронный газ». Плотность «газа» электронов примерно в 1000 раз больше плотности классического газа при нормальных температурах и давлении. К такому «газу», состоящему из электронов массой m, которые (в отличие от молекул обычного газа) движутся на фоне тяжелых неподвижных ионов, Друде применил кинетическую теорию идеального газа.

В теории Друде, как и в кинематической теории, столкновения – это мгновенные события, внезапно меняющие скорость электрона. Он связывал эти столкновения с тем, что электроны отскакивают от непроницаемых сердцевин ионов, а не считал их электрон-электронными столкновениями по аналогии с механизмом рассеяния в обычном газе. Такая простая механистическая модель, согласно которой электрон отскакивает от иона к иону, весьма далека от действительности. Однако во многих задачах это и не важно: для качественного, а иногда даже и количественного понимания проводимости металлов достаточно предположить существование какого-либо механизма рассеяния, не вдаваясь в подробности того, каков этот механизм.

Удельная проводимость. Электроны проводимости при своем движении сталкиваются с ионами решетки, в результате чего устанавливается термодинамическое равновесие между электронным газом и решеткой. По теории Друде-Лорентца, электроны обладают такой же энергией теплового движения, как и молекулы одноатомного газа. В соответствии с молекулярно-кинетической теорией идеальных газов средняя кинетическая энергия электронов, находящихся в состоянии непрерывного хаотического движения, линейно возрастает с температурой:

$$\frac{m_0\langle\upsilon\rangle^2}{2}=\frac{3}{2}kT, \qquad (2.1)$$

где $\langle\upsilon\rangle$ – средняя скорость теплового движения; k – постоянная Больцмана; m_0 – масса электрона. Среднюю скорость теплового движения можно найти из формулы (2.1):

$$\langle\upsilon\rangle=\sqrt{\frac{3kT}{m_0}}. \qquad (2.2)$$

При температуре T = 300 К средняя скорость $\langle\upsilon\rangle$ достаточно велика и достигает значений 10^5 м/с, но так как тепловое движение электронов хаотично, то оно не приводит к появлению тока.

При наложении внешнего электрического поля на металлический проводник кроме уже существующего хаотического теплового движения электронов, возникает их упорядоченное движение, т.е. электрический ток. По плотности тока (например, для медных проводов допустимая плотность тока $j = 10^7$ А/м^2) можно найти значение средней скорости упорядоченного движения, т.е. скорость дрейфа $\langle\upsilon_{др}\rangle=\frac{j}{en}$. Эта скорость порядка 10^{-5} м/с, что значительно меньше средней скорости теплового движения электронов (примерно в 10^{10} раз). Это и позволяет при вычислениях суммарную скорость ($\langle\upsilon\rangle+\langle\upsilon_{др}\rangle$) заменять только средней скоростью теплового движения $\langle\upsilon\rangle$.

Казалось бы, незначительная дрейфовая скорость электронов противоречит факту практически мгновенной передачи электрических сигналов на большие расстояния. Но дело в том, что замыкание электрической цепи влечет за собой распространение электрического поля со скоростью света ($c = 3\cdot10^8$ м/с). Через время $t = l/c$ (l – длина цепи) вдоль цепи установится стационарное электрическое поле и в ней начнется упорядоченное движение электронов. Поэтому электрический ток возникает в цепи практически одновременно с ее замыканием.

Согласно теории Друде удельную электрическую проводимость металла можно найти из следующих рассуждений.

Пусть в металлическом проводнике существует электрическое поле напряженностью E = const. Со стороны поля заряд e испытывает действие силы $F = eE$ и приобретает ускорение $a=\frac{F}{m}=\frac{eE}{m}$. Таким образом, во время

свободного пробега электроны движутся равноускоренно, приобретая к концу свободного пробега скорость

$$\upsilon_{max} = a \cdot \langle t \rangle = \frac{e \cdot E \cdot \langle t \rangle}{m}, \quad (2.3)$$

где $\langle t \rangle$ – среднее время между двумя последовательными соударениями электрона с ионами решетки.

Согласно теории Друде, в конце свободного пробега электрон, сталкиваясь с ионами решетки, отдает им накопленную кинетическую энергию, поэтому скорость его упорядоченного движения станет равной нулю. Следовательно, средняя скорость направленного движения электрона

$$\langle \upsilon_{др} \rangle = \frac{\upsilon_{max} + 0}{2} = \frac{e \cdot E \cdot \langle t \rangle}{2m}. \quad (2.4)$$

Классическая теория металлов не учитывает распределения электронов по скоростям, поэтому среднее время $\langle t \rangle$ свободного пробега определяется средней длиной свободного пробега $\langle l \rangle$ и средней скоростью движения электронов относительно кристаллической решетки проводника, равной $(\langle \upsilon \rangle + \langle \upsilon_{др} \rangle)$, где $\langle \upsilon \rangle$ – средняя скорость теплового движения электронов. Как было показано, $\langle \upsilon_{др} \rangle << \langle \upsilon \rangle$, поэтому среднее время

$$\langle t \rangle = \frac{\langle l \rangle}{\langle \upsilon \rangle}. \quad (2.5)$$

Подставив значения $\langle t \rangle$ в формулу (2.4), получим

$$\langle \upsilon_{др} \rangle = \frac{e \cdot E \cdot \langle l \rangle}{2m \cdot \langle \upsilon \rangle}. \quad (2.6)$$

Плотность тока в металлических проводниках

$$j = ne \langle \upsilon_{др} \rangle = \frac{n \cdot e^2 \cdot \langle l \rangle}{2m \cdot \langle \upsilon \rangle} E. \quad (2.7)$$

Отсюда видно, что плотность тока пропорциональна напряженности поля, т. е. можно записать закон Ома в дифференциальной форме $j = \sigma E$. Коэффициент пропорциональности между j и E – это удельная электрическая проводимость металла

$$\sigma = \frac{n \cdot e^2 \cdot \langle l \rangle}{2m \cdot \langle \upsilon \rangle}, \quad (2.8)$$

где n – концентрация электронов в металле; m – масса электрона; $\langle l \rangle$ – средняя длина свободного пробега; $\langle \upsilon \rangle$ – средняя скорость теплового движения электронов.

Вывод формулы (2.8) нельзя считать строгим, так как принималось во внимание движение лишь одного электрона, а выводы распространялись на все свободные электроны. Более правильным было бы рассмотреть действие

электрического поля на всю совокупность свободных электронов, у которых суммарный импульс изменяется как под действием поля, так и под действием соударений с узлами кристаллической решетки. Такой анализ приводит к тому, что средняя дрейфовая скорость электронов оказывается вдвое больше. С учетом этой поправки выражение для удельной проводимости примет вид:

$$\sigma = \frac{n \cdot e^2 \cdot \langle l \rangle}{m \cdot \langle \upsilon \rangle}. \qquad (2.9)$$

Так как в (2.9) n и $\langle l \rangle$ от температуры не зависят, а средняя скорость теплового движения $\langle \upsilon \rangle \sim \sqrt{T}$, то $\sigma \sim \frac{1}{\langle \upsilon \rangle} \sim \frac{1}{\sqrt{T}}$. Тогда удельное сопротивление металла ρ, равное $\rho = \frac{1}{\sigma}$, должно возрастать пропорционально $\sqrt{T}$. Этот вывод классической электронной теории противоречит опытным данным, согласно которым $\rho \sim T$ ($\rho = \alpha \rho_0 T$; ρ_0 – удельное сопротивление при 0 °С, α – температурный коэффициент).

Удельная теплопроводность. Представления о свободных электронах позволяют легко перейти к экспериментальному закону *Видемана-Франца*, устанавливающему связь между электропроводностью и теплопроводностью металла: *для всех металлов при одной и той же температуре отношение удельной теплопроводности* λ *к удельной электрической проводимости* σ *одинаково*, т.е.

$$\frac{\lambda}{\sigma} = \beta T, \qquad (2.9)$$

где β – константа, не зависящая от рода металла.

Из опыта известно, что металлы обладают как высокой электропроводностью, так и высокой теплопроводностью. Способностью проводить тепло обладают и неметаллические кристаллы, например, диэлектрики. Однако теплопроводность металлов значительно превосходит теплопроводность диэлектриков. Так, медь имеет удельную теплопроводность 406 Вт/К·м, серебро 453 Вт/К·м, алюминий 218 Вт/К·м, что значительно выше чем у большинства диэлектриков. Например, слюда имеет удельную теплопроводность 0,44 Вт/К·м, кварцевое стекло ~ 1,4 Вт/К·м. Из этого следует, что теплопередача в металлах осуществляется в основном не кристаллической решеткой, а свободными электронами, т.е. носителями тока и теплоты в металлах являются одни и те же частицы – свободные электроны.

Элементарная классическая теория электропроводности металлов позволила найти значение константы β. В соответствии с молекулярно-кинетической теорией идеального газа удельная электронная теплопроводность может быть записана в виде:

$$\lambda = \frac{1}{2} k n \langle \upsilon \rangle \langle l \rangle. \qquad (2.10)$$

где $\langle l \rangle$ – средняя длина свободного пробега; $\langle \upsilon \rangle$ – средняя скорость теплового движения электронов; n – концентрация электронов в металле.

Поделив (2.10) на удельную электропроводность (2.9), и с учетом формулы (2.2) получим, что

$$\frac{\lambda}{\sigma} = \frac{\frac{1}{2} kn \langle \upsilon \rangle \langle l \rangle}{\frac{e^2 n \langle l \rangle}{2m \langle \upsilon \rangle}} = \frac{km \langle \upsilon \rangle^2}{e^2} = \frac{3mk^2 T}{me^2} = 3k^2 e^{-2} T .$$

$$\frac{\lambda}{\sigma} = 3k^2 e^{-2} T = LT , \qquad (2.11)$$

где $L = \frac{3k^2}{e^2} = 2,23 \cdot 10^{-8}$ Дж /(Кл·К)^2 – число Лоренца. При T = 300 К для отношения $\frac{\lambda}{\sigma}$ получается значение 6,7·10^{-6} Дж·Ом/(с·К), хорошо согласующееся с экспериментальными данными (для алюминия это отношение равно 5,8·10^{-6}, для меди – 6,4·10^{-6}, для свинца – 7,0·10^{-6} Дж·Ом/(с·К)). Однако впоследствии выяснилось, что столь хорошее совпадение оказалось случайным. Когда Лоренц учел распределение электронов по скоростям, для отношения $\frac{\lambda}{\sigma}$ получилось значение $2k^2 e^{-2} T$, которое хуже согласовалось с опытными данными.

Теплоемкость. Рассчитаем теплоемкость твердого тела. В качестве модели твердого тела рассмотрим правильно построенную кристаллическую решетку, в узлах которой находятся частицы (атомы, ионы, молекулы), принимаемые за материальные точки, колеблющиеся около положения равновесия – узлов решетки – в трех взаимно перпендикулярных направлениях. Электроны учитывать не будем. Таким образом, каждой составляющей кристаллическую решетку частице приписывается три колебательных степени свободы, каждая из которых обладает энергией kT. Внутренняя энергия одного моля твердого тела

$$E_M = 3N_A kT = 3RT , \qquad (2.12)$$

где N_A – число Авогадро; $N_A k = R$ (R – универсальная газовая постоянная).

Молярную теплоемкость твердого тела можно найти, продифференцировав энергию моля вещества

$$C = \frac{dE_M}{dT} = 3R \approx 25 \text{ Дж/(моль·К)}. \qquad (2.13)$$

Из (2.13) видно, что молярная (атомная) теплоемкость химически простых тел в кристаллическом состоянии одинакова (равна 3R) и не зависит от температуры. Этот закон был эмпирически получен французскими учеными П. Дюлонгом и Л. Пти и носит название *закона Дюлонга* и *Пти*. Если твердое тело является химическим соединением (например, NaCl), то число частиц в

моле равно $n_a N_A$, где n_a – число атомов в молекуле. Таким образом, молярная теплоемкость твердых химических соединений

$$C = 3n_a R \approx 25n_a \text{ Дж/(моль·К)}, \qquad (2.14)$$

т. е. она равна сумме атомных теплоемкостей элементов, составляющих эти соединения.

В металлах в процессе поглощения энергии должны принимать участие не только колеблющиеся атомы, теплоемкость которых можно найти по формуле (2.13), но и свободные электроны. С этой точки зрения теплоемкость металлов должна быть существенно больше, чем теплоемкость диэлектриков, которые не содержат свободных электронов. Принимая, что число свободных электронов в металле равно числу атомов, и учитывая, что электроны как частицы идеального газа имеют среднюю тепловую энергию $\frac{3kT}{2}$, для молярной теплоемкости металлов получим

$$C_V = C + C_e = 3kN_A + \frac{3}{2}kN_A = 3R + \frac{3}{2}R = \frac{9}{2}R, \qquad (2.15)$$

где C – теплоемкость атомов моля вещества; C_e – теплоемкость электронов.

Однако в действительности теплоемкость металлов при высоких температурах не равна $\frac{9}{2}R$ и мало отличается от теплоемкости кристаллических диэлектриков. Это свидетельствует о том, что электронный газ практически не поглощает теплоты при нагревании металлического проводника, что противоречит классической электронной теории.

Успехи классической теории Друде-Лоренца неоспоримы, так как она объяснила и обосновала законы Ома и Джоуля-Ленца, а также дала качественное объяснение закона Видемана-Франца. Однако при дальнейшем своем развитии классическая теория зашла в тупик, потому что, несмотря на многочисленные усилия, так и не смогла объяснить целый ряд экспериментально установленных закономерностей.

1. Одним из главных противоречий теории Друде-Лоренца является несоответствие теоретически получаемых значений теплоемкости металлов с опытными данными.

2. Имеются экспериментальные доказательства, что средняя длина свободного пробега электронов в металлах высокой проводимости при комнатных температурах составляет $10^{-8}…10^{-5}$ м, а при низких температурах – существенно больше. К примеру, экспериментально установлено, что для чистой меди при температуре 300 К длина свободного пробега $\langle l \rangle = 3\cdot10^{-4}$ м, при температуре жидкого гелия (4,2 К) – $\langle l \rangle = 3\cdot10^{-2}$ м, а у сверхчистой меди при этой же температуре длина свободного пробега достигает значений порядка 0,1 м. Таким образом, приходится предположить, что электрон проходит без соударений с ионами решетки сотни, а то и сотни тысяч межузельных расстояний. Однако это предположение невозможно объяснить в рамках теории Друде-Лоренца.

3. Экспериментально установлено, что в широком интервале температур удельное электрическое сопротивление пропорционально температуре $\rho \sim T$. А из формулы удельной проводимости (2.9) следует, что $\rho \sim \sqrt{T}$. Это положение классической электронной теории также противоречит экспериментальным данным.

Перечисленные выше трудности удалось преодолеть с помощью квантовой теории металлов, основы которой разработаны немецким физиком А. Зоммерфельдом.

2.1.2. Элементы квантовой теории металлов

Основные недостатки классической теории исходят не столько из представлений о существовании в металлах свободных электронов, сколько от применения к ним законов классической статистики Максвелла-Больцмана (см. гл. I, ф-ла (1.15)). Зоммерфельд сохранил модель свободных электронов, но применил к ней вместо классической статистики квантовую статистику Ферми-Дирака (см. гл. I, ф-ла (1.17)).

Квантовая статистика базируется на принципе Паули, откуда вытекает различие классического и квантового распределений электронов по энергиям. С классической точки зрения энергия всех электронов при температуре абсолютного нуля должна равняться нулю. По принципу Паули даже при абсолютном нуле число электронов в каждом энергетическом состоянии не может превышать двух. И если общее число свободных электронов в кристалле равно n, то при температуре 0 К они займут $n/2$ наиболее низких энергетических состояний. Максимальная энергия, которой может обладать электрон в металле при температуре абсолютного нуля, называется энергией Ферми E_F. Энергию Ферми можно найти по формуле:

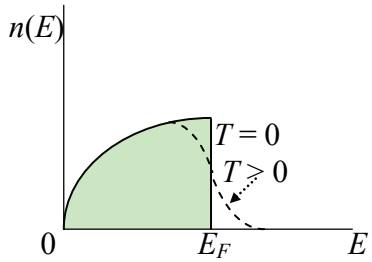


Рис. 2.1 – Распределение электронов по энергиям в металлах при температурах $T = 0$ и $T > 0$

$$E_F = \frac{h^2}{2m_n^*} \cdot \left(\frac{3n}{8\pi}\right)^{\frac{2}{3}}, \qquad (2.16)$$

где h – постоянная Планка; E_F – энергия Ферми; m_n^* – эффективная масса электрона; n – концентрация электронов. Из (2.16) видно, что энергия Ферми E_F определяется концентрацией свободных электронов. Так как концентрация свободных электронов в металле велика ($\sim 10^{28}$ м$^{-3}$), то энергия Ферми также оказывается большой и в типичных случаях составляет 3…15 эВ.

При нагревании кристалла ему сообщается тепловая энергия порядка kT *на каждую частицу*. За счет этого некоторые электроны, находящиеся вблизи уровня Ферми, начинают заполнять состояния с более высокой энергией и при $T > 0$ график распределения электронов по энергиям $n(E)$ становится более пологим (рис. 2.1). Из распределения Ферми-Дирака (ф-ла

(1.17)) видно, что при любой температуре для уровня с энергией Ферми вероятность заполнения электронами равна 0,5 (рис. 1.16). Все уровни с энергией меньше энергии Ферми, заполнены электронами с вероятностью большей 0,5; а уровни с энергией большей энергии Ферми свободны от электронов с вероятностью более 0,5 (рис. 1.17).

При обычных температурах электроны в металлах следует рассматривать как «вырожденный газ», поведение которого описывается статистикой Ферми-Дирака. В состоянии вырождения средняя энергия электронного газа, а также средняя тепловая скорость электронов практически не зависят от температуры, так как тепловая энергия расходуется на увеличение амплитуды колебаний узлов кристаллической решетки.

Электронный газ в металле остается вырожденным до тех пор, пока любой из электронов не сможет обмениваться энергией с кристаллической решеткой. А это возможно лишь тогда, когда средняя энергия тепловых колебаний станет близкой энергии Ферми:

$$kT_F = E_F . \qquad (2.17)$$

Для металлов температура снятия вырождения T_F (температура Ферми) составляет $10^4 \ldots 10^5$ К, т. е. превышает не только температуру плавления, но и температуру испарения металлов.

Итак, электронный газ в металлах вырожден практически при всех температурах. Этим объясняются неудачи теории Друде-Лорентца, оперировавшей с невырожденным газом.

Квантовая теория электропроводности металлов сумела устранить те противоречия, которые были заложены в классической электронной теории, и сумела объяснить:

1. Малую теплоемкость электронного газа.

2. Уточненное значение числа Лоренца.

3. Аномально большие величины (порядка сотен периодов решетки) средней длины свободного пробега электронов в металле.

4. Обратно пропорциональную зависимость удельной проводимости от температуры $\sigma \sim \frac{1}{T}$, которая соответствует экспериментальным данным (классическая теория дает, что $\sigma \sim \frac{1}{\sqrt{T}}$).

Выводы квантовой теории по отношению к вышеперечисленным вопросам таковы:

1. Так как в любом состоянии может находиться только одна частица (принцип Паули), электрон с малой энергией оказывается не способным принять предоставляемую ему порцию энергии. Все более высокие состояния уже заняты, и электрону приходится оставаться в том энергетическом состоянии, в котором он и находился. Только вблизи уровня Ферми электроны имеют какую-то вероятность перейти в состояние с большей энергией. И только эти электроны дают вклад в теплоемкость, а так

как их количество незначительно, то теплоемкость электронного газа практически очень мала. Была рассчитана молярная теплоемкость электронного газа

$$C_V = \frac{\pi^2 N_A}{2}\frac{k^2}{E_F}T, \qquad (2.18)$$

где E_F – энергия Ферми, T – абсолютная температура. Сравним теперь эту теплоемкость с молярной теплоемкостью классического одноатомного газа $C_V^{\text{кл}} = \frac{3N_A k}{2}$ (при комнатных температурах): $\frac{C_V}{C_V^{\text{кл}}} = \frac{\pi^2}{3}\frac{kT}{E_F} \sim 0{,}015$. При обычных температурах теплоемкость в 50…70 раз меньше, чем в классической теории. Этим устраняются трудности, на которые натолкнулась классическая теория в вопросе о теплоемкости электронов в металлах.

2. Более строгий анализ, основанный на квантовой статистике электронов, дает несколько иное выражение и для числа Лоренца, хотя численное значение его существенно не изменяется:

$$L = \frac{\lambda}{\sigma T} = \frac{\pi k^2}{3e^2} = 2{,}45\cdot 10^{-8}\ \text{Дж}/(\text{Кл}\cdot\text{К})^2. \qquad (2.19)$$

3. Квантовая теория рассматривает движение электронов с учетом их взаимодействия с кристаллической решеткой. Согласно корпускулярно-волновому дуализму, движение электрона сопоставляют с волновым процессом. Идеальная кристаллическая решетка (в ее узлах находятся неподвижные частицы, и в ней отсутствует нарушение периодичности) ведет себя подобно оптически однородной среде – она «электронные волны» не рассеивает. Это соответствует тому, что металл не оказывает электрическому току никакого сопротивления. «Электронные волны», распространяясь в кристаллической решетке, как бы огибают узлы решетки и проходят значительные расстояния. Это означает, что длина свободного пробега электронов может быть бесконечной, тогда и электрическая проводимость металлов должна стремиться к бесконечности.

Но в реальных кристаллических решетках всегда имеются дефекты решетки: динамические (тепловые колебания узлов кристаллической решетки) и статические (одномерные, линейные дефекты решетки, границы кристалла). На этих дефектах и происходит рассеяние «электронных волн», что и является причиной электрического сопротивления металлов. Рассеяние «электронных волн» на неоднородностях, связанных с тепловыми колебаниями, можно рассматривать как столкновения свободных электронов с фононами. Таким образом, для большинства чистых металлов в широком температурном интервале длина свободного пробега обратно пропорциональна температуре: $\langle l \rangle \sim \frac{1}{T}$ (см. п. 2.1.3). При очень низких температурах ситуация меняется и появляется сверхпроводимость (см. п. 2.1.9).

4. Вследствие вырождения в процессе электропроводности могут принимать участие не все свободные электроны, а только небольшая их

часть, имеющая энергию, близкую к энергии Ферми. Только эти электроны способны менять свои состояния под действием поля. При воздействии электрического поля эти электроны ускоряются на длине свободного пробега и приобретают добавочную скорость направленного движения $\upsilon_{\text{др}}^{F}$:

$$\upsilon_{\text{др}}^{F} = \frac{eE}{m_n^*}\tau_F = \frac{eE}{m_n^*}\frac{\langle l_F \rangle}{\upsilon_F}, \qquad (2.20)$$

где τ_F и $\upsilon_{\text{др}}^{F}$ – соответственно время свободного пробега и дрейфовая скорость электронов, обладающих энергией, близкой к энергии Ферми.

Если значение дрейфовой скорости (2.20) подставить в выражение для плотности тока $j = ne\upsilon_{\text{др}}^{F}$, то расчет удельной проводимости металлов, приводит к следующему выражению

$$\sigma = \frac{ne^2 \langle l_F \rangle}{m_n^* \upsilon_F}\ , \qquad (2.21)$$

которое по внешнему виду напоминает классическую формулу (2.9) для удельной проводимости σ, но имеет совершенно другое физическое содержание. Здесь n – концентрация электронов проводимости в металле, $\langle l_F \rangle$ – средняя длина пробега электронов, имеющих энергию Ферми, υ_F – скорость электронов, имеющих энергию Ферми.

Согласно классической теории, $\langle \upsilon \rangle \sim \sqrt{T}$ (см. формулу (2.2)), поэтому эта теория так и не смогла объяснить истинную зависимость проводимости σ от температуры T. В квантовой механике скорость υ_F зависит только от энергии Ферми, а энергия Ферми с изменением температуры остается практически неизменной. Например, при нагревании серебра от 0 до 1000 К энергия Ферми в нем уменьшается лишь на 0,2 %. Таким образом, скорость υ_F не является скоростью теплового движения и от температуры не зависит. Оценить значение скорости υ_F быстрых электронов, имеющих энергию Ферми, можно исходя из следующих соображений. Энергия Ферми равна кинетической энергии движения быстрых электронов

$$E_F = \frac{m_n^* \upsilon_F^2}{2}, \qquad (2.22)$$

отсюда

$$\upsilon_F = \sqrt{\frac{2E_F}{m_n^*}}\ . \qquad (2.23)$$

Если учесть, что концентрация электронов в металлах порядка 10^{29} м$^{-3}$, эффективная масса электрона в металлах не отличается от его реальной массы, то энергию Ферми E_F можно найти по формуле (2.16). Так как $E_F \approx 6$эВ $\approx 10^{-18}$ Дж, то $\upsilon_F \approx 10^6$ м/с.

С повышением температуры возрастает рассеяние «электронных волн» на тепловых колебаниях решетки (фононах), что соответствует уменьшению средней длины свободного пробега электронов $\langle l_F \rangle$. В широком

температурном интервале $\langle l_F \rangle \sim \frac{1}{T}$, поэтому, учитывая независимость скорости υ_F от температуры, получим, что удельная проводимость металлов обратно пропорциональна температуре $\sigma \sim \frac{1}{T}$, тогда удельное сопротивление металлов в соответствие с данными опытов растет пропорционально температуре $\rho \sim T$. Таким образом, удельное сопротивление металлических проводников зависит в первую очередь от средней длины свободного пробега электронов $\langle l_F \rangle$.

Как было показано выше, основное количество электронов не изменяет своего энергетического состояния при наложении поля, поэтому в процессе электропроводности могут принимать участие не все свободные электроны, а только небольшая их часть, имеющая энергию, близкую к энергии Ферми. Как в этом случае объяснить высокую проводимость металлов, если концентрация электронов n в (2.21) мала?

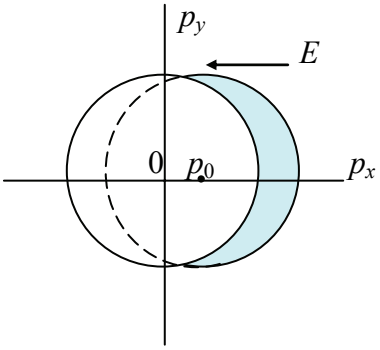


Рис. 2.2 – Смещение центра распределения Ферми под действием внешнего электрического поля. Радиус сферы равен импульсу Ферми $R = p_F$

На рис. 2.2 показано распределение электронов в пространстве импульсов при наличии внешнего электрического поля. Поверхность, ограничивающая это распределение, является поверхностью Ферми, соответствующей энергии Ферми. В отсутствие поля ее центр находится в начале координат. При воздействии поля наблюдается смещение центра распределения в направлении против поля на величину импульса p_0, который является средним импульсом для электронного ансамбля в данном электрическом поле. Энергию Ферми (см. формулу (2.16)) можно представит через импульс p_F электронов

$$E_F = \frac{p_F^2}{2m^*}, \tag{2.24}$$

где m^* – эффективная масса электронов; p_F – импульс электронов, обладающих энергией Ферми (максимальный импульс электронов). Из (2.16) и (2.24) получим импульс

$$p_F = \sqrt{h^2\left(\frac{3n}{8\pi}\right)^{\frac{2}{3}}} = h\left(\frac{3n}{8\pi}\right)^{\frac{1}{3}} \tag{2.25}$$

и объем сферы Ферми

$$V_F = \frac{4}{3}\pi R^3 = \frac{4}{3}\pi p_F^3 = \frac{4\pi h^3 3n}{3\cdot 8\pi} = \frac{h^3}{2}n\,. \tag{2.26}$$

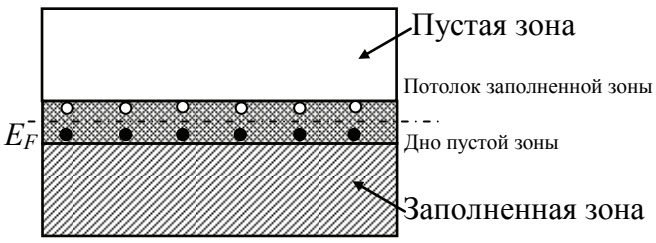
Объем сферы Ферми определяется полной концентрацией электронов в металле. Поэтому движение электронов в металле во внешнем поле можно представить как движение центра сферы Ферми со скоростью

$$\upsilon_{др} = \upsilon_0 = \frac{p_0}{m_0 n}, \qquad (2.27)$$

где υ_0 – скорость дрейфа центра тяжести ансамбля электронов, n – концентрация электронов.

Таким образом, скоростью дрейфа можно считать скорость направленного движения электронного распределения как целого, не учитывая интенсивного хаотического движения электронов, которое не переносит ток. Скорости дрейфа соответствует эффективная подвижность ансамбля электронов, которую можно измерить на опыте. С другой стороны, большую проводимость металлов следует объяснять скорее большой величиной тока, переносимого небольшой группой быстрых электронов на вершине распределения Ферми (затемненная область на рис. 2.2), чем высоким значением полной концентрации свободных электронов, которым можно придать небольшую дрейфовую скорость. Тем не менее, в формулу (2.21) входит полная концентрация электронов n. Это связано с тем, что большие скорости на вершине распределения Ферми определяются именно высокой полной концентрацией электронов.

Для того чтобы описывать движение электрона в твердом теле, как свободное перемещение заряженной частицы без учета периодического поля кристаллической решетки, вводится понятие об эффективной массе m^* (см. гл. I), которая может быть как больше, так и меньше массы свободного электрона и даже иметь отрицательный знак. Носителями заряда в металлах, как правило, являются электроны. Эффективные массы таких носителей заряда положительны, а сами металлы называются электронными (например, алюминий, медь). Однако существуют металлы и с отрицательными массами носителей заряда (например, цинк, молибден, магний). Носителями заряда в таких металлах являются дырки, а сами металлы называются дырочными. Появление дырок в качестве носителей заряда в металлах можно объяснить следующим образом. В случае перекрытия полностью заполненной энергетической зоны пустой зоной (характерно для металлов с валентностью бóльшей, чем единица), часть состояний у дна изначально пустой зоны заполняется электронами из области потолка изначально заполненной зоны (рис. 2.3). Такое поведение



● – электроны; ○ – дырки

Рис. 2.3 – Зонная диаграмма металла с перекрытием зон. Объяснение механизма дырочной проводимости в металлах

частиц является энергетически выгодным, так как электроны стремятся занять более низкие по энергиям свободные состояния. Электроны, перешедшие с потолка заполненной зоны на дно пустой зоны, становятся электронами проводимости, а опустевшие состояния у потолка нижней зоны проявляют себя как дырки. Уровень Ферми находится между потолком нижней и дном верхней перекрывшихся зон. В таких металлах проводимость носит смешанный характер, и роль дырок в некоторых из них может преобладать.

2.1.3. Температурная зависимость удельного сопротивления металлических проводников.

Удельное сопротивление проводников. Величину, обратную удельной проводимости σ называют удельным сопротивлением ρ и для проводника с постоянным поперечным сечением S определяют по формуле:

$$\rho = R\frac{S}{l}, \tag{2.28}$$

где l – длина проводника. Единицей удельного сопротивления в СИ является Ом·м, однако в практике чаще пользуются внесистемной единицей мкОм·м.

Следует отметить, что в отличие от полупроводников диапазон удельных сопротивлений металлических проводников достаточно мал – от 0,016 мкОм·м для серебра и примерно до 10 мкОм·м для железо-хромо-кобальто-алюминиевых сплавов, т.е. всего три порядка.

Температурная зависимость удельного сопротивления металлических проводников. Как было показано ранее (см. предыдущий параграф) в идеально чистых металлах единственной причиной, которая ограничивает длину свободного пробега, являются тепловые колебания узлов кристаллической решетки (фононы). Удельное сопротивление металла, обусловленное этим фактором, обозначим как ρ_T. С ростом температуры возрастают амплитуды фононов и связанные с этим флуктуации периодического поля решетки. Это повышает рассеивание электронов, уменьшает длину свободного пробега и вызывает возрастание удельного сопротивления. Для упрощенной одномерной модели решетки длина свободного пробега электронов $\langle l_T \rangle$ определяется как:

$$\langle l_T \rangle = \frac{1}{\pi(\Delta a)^2 n}, \tag{2.29}$$

где Δa – амплитуда фононов; $\pi(\Delta a)^2$ – эффективное сечение рассеяния; n – концентрация атомов в металле.

Потенциальная энергия атома, отклоненного на Δa от узла решетки:

$$E_n = \frac{k_{\text{упр}}(\Delta a)^2}{2}, \tag{2.30}$$

где $k_{\text{упр}}$ – коэффициент упругости связи.

Согласно классической статистике средняя энергия одномерного гармоничного осциллятора равняется kT. Тогда

$$\frac{k_{\text{упр}}(\Delta a)^2}{2} = kT\,, \tag{2.31}$$

где k – постоянная Больцмана. Тогда из (2.29) и (2.31) получим выражение:

$$\langle l_T \rangle = \frac{k_{\text{упр}}}{2\pi kn} \cdot \frac{1}{T}\,. \tag{2.32}$$

Необходимо отметить, что полученное отношение не выполняется при достаточно низких температурах. Дело в том, что с понижением температуры могут уменьшаться не только амплитуды тепловых колебаний атомов, но и частоты колебаний. Поэтому в области низких температур рассеяние электронов тепловыми колебаниями становится не эффективным. Взаимодействие электронов с колеблющимися атомами лишь незначительно меняют импульс электронов. В теории колебаний атомов решетки температуру оценивают относительно некоторой характеристической температуры, которую называют *температурой Дебая* θ_D. Температура Дебая определяет максимальную частоту тепловых колебаний, которые могут возбуждаться в кристалле: $\theta_D = \frac{h\nu_{\max}}{k}$. Эта температура зависит от сил связи между узлами кристаллической решетки и является важным параметром твердого тела. При температуре выше этой, теплоемкость вещества постоянна.

Если подставить (2.32) в (2.21), то при $T > \theta_D$ получим формулу для расчета удельного сопротивления:

$$\rho_T = \frac{1}{\sigma_T} = \frac{2\pi m^* k \langle \upsilon_F \rangle}{e^2 k_{\text{упр}}} T\,. \tag{2.33}$$

Из (2.33) видно, что с ростом температуры удельное сопротивление чистых металлов возрастает прямо пропорционально температуре. В действительности же эта зависимость является более сложной (рис. 2.4).

На участке I (рис. 2.4) при комнатных температурах зависимость $\rho = f(T)$ является линейной, что соответствует (2.33). Т.е. с ростом температуры возрастает амплитуда тепловых колебаний узлов кристаллической решетки, что уменьшает длину свободного пробега электронов.

На участке II (рис. 2.4), ниже температуры Дебая ($T << \theta_D$) изменяется частота тепловых колебаний узлов кристаллической решетки, поэтому зависимость $\rho = f(T)$ является нелинейной и подчиняется закону:

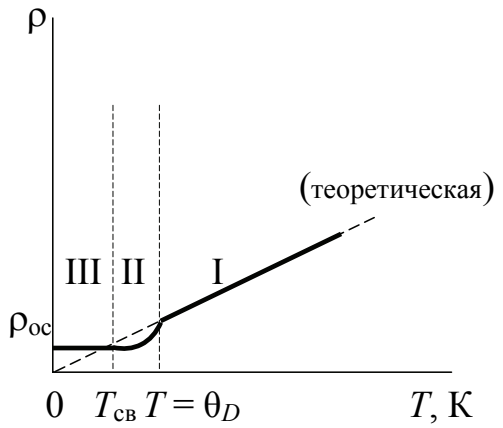


Рис. 2.4 – Зависимость удельного сопротивления проводника от температуры

$$\rho = A \cdot T^n , \qquad (2.34)$$

где n – число, меняющееся от 4 до 6, как правило, в небольшом температурном интервале.

На участке III (рис. 2.4) некоторые металлы имеют конечное сопротивление ($\rho_{ост}$) даже при температуре $T = 0$ К. Это объясняется наличием в металле *статических дефектов решетки*, прежде всего примесей, и позволяет оценить чистоту металлов на основании отношения:

$$\rho_{300K} / \rho_{4K} , \qquad (2.35)$$

где ρ_{300K}, ρ_{4K} – удельное сопротивление металла при 300 К и 4,2 К соответственно (температура кипения жидкого гелия). Чем меньше это отношение, тем чище металл.

У некоторых металлов при температуре ниже $T_{св}$ наблюдается резкое уменьшение удельного сопротивления до нуля. Такое явление называют сверхпроводимостью (смотри подробнее в п. 2.1.9).

Температурный коэффициент удельного сопротивления. В диапазоне температуры, где зависимость ρ от t °С близка к линейной (рис. 2.3, участок III) допустима линейная аппроксимация этой зависимости, и величина удельного сопротивления в конце диапазона температуры t может быть рассчитана по формуле:

$$\rho_t = \rho_0(1+\alpha_\rho \Delta t), \qquad (2.36)$$

где ρ_0 – удельное сопротивление в начале диапазона.

Величину α_ρ из выражения (2.36) называют средним *температурным коэффициентом удельного сопротивления* в данном диапазоне температуры:

$$\alpha_\rho = \frac{\rho_t - \rho_0}{\rho_0 \Delta t} , \qquad (2.37)$$

α_ρ измеряется в град$^{-1}$.

Дифференциальное выражение для температурного коэффициента α_ρ имеет вид:

$$\alpha_\rho = \frac{1}{\rho}\frac{d\rho}{dt}. \qquad (2.38)$$

Значения α_ρ чистых металлов в твердом состоянии близки друг к другу, и поэтому приближенно можно считать $\alpha_\rho \approx 0{,}004$, град$^{-1}$.

Исключение составляют элементы, относящиеся к ферромагнетикам – железо, никель, кобальт, гадолиний, а также натрий, калий, хром и др., однако и для них α_ρ отличается от приведенной величины только в 1,5…2 раза.

Наличие примесей уменьшает значение температурного коэффициента α_ρ. У некоторых сплавов температурный коэффициент α_ρ может быть равен нулю и даже может иметь небольшие отрицательные значения. Это объясняют тем, что при более сложных составах по сравнению с чистыми металлами, сплавы нельзя рассматривать как классические металлы, т. е.

изменение их проводимости обуславливается не только уменьшением подвижности носителей заряда, но в некоторых случаях, и частичным возрастанием концентрации носителей при повышении температуры. Сплав, у которого уменьшение подвижности с увеличением температуры полностью компенсируется возрастанием концентрации носителей заряда, имеет нулевой температурный коэффициент удельного сопротивления. Это явление используется для изготовления термостабильных сплавов, например, константана и манганина**.**

2.1.4. Влияние примесей на удельное сопротивление металлов. Удельное сопротивление сплавов

Влияние примесей на удельное сопротивление металлов. Как уже говорилось, причинами рассеяния электронов в металлах являются не только тепловые колебания узлов кристаллической решетки, но и наличие статических дефектов, которые, прежде всего, связанны с примесями. Рассеивание на статических дефектах не зависит от температуры. Поэтому при абсолютном нуле сопротивление реальных металлов остается конечным. Из этого следует *правило Маттиссена* об аддитивности удельного сопротивления:

$$\rho = \rho_T + \rho_{ост}, \tag{2.39}$$

где ρ – полное сопротивление металла с примесью; ρ_T – сопротивление, обусловленное рассеянием электронов на фононах; $\rho_{ост}$ – остаточное сопротивление, обусловленное рассеиванием электронов на статических дефектах решетки.

Наибольший вклад в остаточное сопротивление вносит рассеяние на примесных атомах, которые практически всегда имеются в металлах. Поэтому длина свободного пробега электронов в металлах с примесью имеет вид:

$$\frac{1}{\langle l \rangle} = \frac{1}{\langle l_T \rangle} + \frac{1}{\langle l_{пр} \rangle}, \tag{2.40}$$

где $\langle l_T \rangle$, $\langle l_{пр} \rangle$ – длины свободного пробега электронов, ограниченные фононами и примесями, соответственно.

Определим длину свободного пробега $\langle l_{пр} \rangle$:

$$\frac{1}{\langle l_{пр} \rangle} \approx N_{пр} \cdot S_{пр}, \tag{2.41}$$

где $N_{пр}$ – концентрация атомов примеси; $S_{пр}$ – эффективное сечение рассеивания электронов атомами примеси. Тогда удельное сопротивление проводника с примесью

$$\rho = \frac{m^* \upsilon_F}{e^2 n \langle l \rangle} = \rho_T + \frac{m^* \upsilon_F N_{пр} S_{пр}}{e^2 n}. \tag{2.42}$$

Таким образом, наличие примеси увеличивает удельное сопротивление металла, но его зависимость от температуры остается линейной (рис. 2.5). Различные примеси по-разному влияют на сопротивление металла. Эффективность примесного рассеяния определяется возмущающим потенциалом в решетке, значение которого тем выше, чем сильнее различаются валентности примесных атомов металла основы (матрицы). Это зависит от деформаций кристаллической решетки атомами примеси. Чем большая разность в размерах собственных и примесных атомов, тем больше остаточное сопротивление металла, т.е. выполняется правило Линде:

$$\Delta\rho_{ост} = a + b(\Delta Z)^2, \qquad (2.43)$$

где $\Delta\rho_{ост}$ – изменение остаточного сопротивления при изменении примеси; ΔZ – разность валентностей собственного атома и атома примеси; a, b – константы.

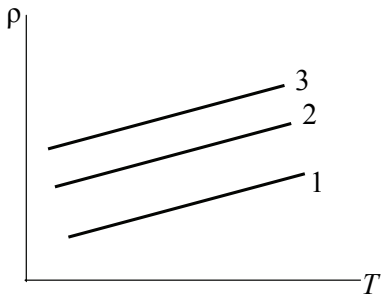


Рис. 2.5 – Температурная зависимость металлов с примесями: 1 – чистый металл; 2 – легирование металлом; 3 – легирование металлоидом

Таким образом, на сопротивление металлов меньшее влияние оказывают примесные атомы металла, а большее – атомы металлоидов.

Изменение удельного сопротивления металла при упругих деформациях объясняется изменением амплитуды колебания узлов кристаллической решетки. Увеличение амплитуды колебания узлов решетки приводит к уменьшению длины свободного пробега носителей заряда и удельное сопротивление возрастает. Пластическая деформация, как правило, повышает удельное сопротивление металлов вследствие искажения кристаллической решетки. При рекристаллизации путем термической обработки (отжига) удельное сопротивление может быть опять снижено до первоначальных значений.

Удельное сопротивление сплавов. В технике очень широко используют металлические сплавы, имеющие значительную концентрацию атомов примеси, со структурой неупорядоченного твердого раствора. Статистическое распределение атомов разного вида в узлах кристаллической решетки вызывает значительные флуктуации периодического поля кристалла, рассеивающего электроны. Как и в случае металла (см. ф-лу (2.39)), полное сопротивление сплава можно выразить в виде суммы двух сопротивлений: сопротивления ρ_T, обусловленного рассеянием электронов на тепловых колебаниях решетки, и остаточного сопротивления $\rho_{ост}$, связанного с рассеянием электронов нам неоднородностях структуры сплава.

Для многих двухкомпонентных сплавов изменение $\rho_{ост}$ от состава сплава хорошо описывается законом Нордгейма:

$$\rho_{ост} = C x_A x_B = C\, x_B\,(1\text{-}x_B) = C x_A(1\text{-}x_A), \qquad (2.44)$$

где C – константа; x_A, x_B – атомные доли компонентов в сплаве. В бинарных твердых растворах A-B остаточное сопротивление возрастает как при

добавлении атомов металла B к металлу A, так и при добавленные атомов металла A к металлу B и достигает максимума при $x_A = x_B = 0{,}5$. Закон Нордгейма описывает изменение остаточного сопротивления для непрерывного ряда твердых растворов. Если сплав отжечь, то он может стать упорядоченным и, если при этом возникают интерметаллические соединения, которые имеют собственную кристаллическую решетку, зависимость остаточного сопротивления разделяется на части, соответственно числу интерметаллических соединений.

Обычно, удельное сопротивление металлических сплавов выше сопротивления чистых металлов. Это физическое свойство сплавов используется для получения высокоомных проводниковых материалов. Однако, если отсутствует взаимная растворимость компонентов, образующих сплав, то возникает структура в виде смеси зерен практически чистых металлов. Удельное сопротивление такого сплава линейно возрастает пропорционально доле металла с бóльшим значением удельного сопротивления.

Следует отметить, что для ферромагнитных проводников (например, для железа) и других ферромагнитных металлов и их сплавов характерна нелинейная зависимость удельного сопротивления от температуры. Эта особенность обусловлена изменением спонтанной намагниченности по мере приближения к температуре Кюри (T_K), выше которой ферромагнитные свойства отсутствуют (см. гл.V). При очень низких температурах магнитные моменты в ферромагнитных материалах ориентированы параллельно друг другу. Благодаря такому упорядоченному, периодическому расположению они не вызывают рассеяния электронов, движущихся под действием электрического поля. При повышении температуры спиновая упорядоченность нарушается, что вызывает дополнительное рассеяние электронов проводимости. Согласно правилу Матиссена, различные механизмы рассеяния дают аддитивный (суммарный) вклад в полное сопротивление:

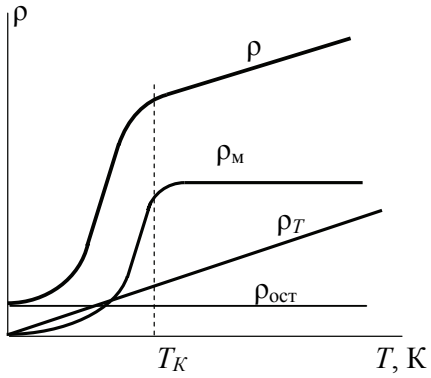


Рис. 2.6 – Температурная зависимость удельного сопротивления ферромагнитного материала

$$\rho = \rho_T + \rho_{ост} + \rho_м, \qquad (2.45)$$

где ρ – полное сопротивление металла с примесью; ρ_T – сопротивление, обусловленное рассеянием электронов на фононах; $\rho_{ост}$ – остаточное сопротивление, обусловленное рассеиванием электронов на статических дефектах решетки; $\rho_м$ – магнитный вклад в электрическое сопротивление, обусловленный беспорядком в системе спинов. Отдельные составляющие удельного сопротивления ферромагнитного материала и итоговое удельное сопротивление схематически показаны на рис. 2.6. До температуры Кюри зависимость удельного сопротивления от температуры носит сложный

характер. Выше температуры Кюри T_K магнитная составляющая остается постоянной, благодаря чему зависимость удельного сопротивления от температуры принимает линейный характер.

2.1.5. Работа выхода электронов из металла и эффект Шоттки.

При нагревании или освещении светом электроны могут покидать металл. Детальные экспериментальные исследования показали, что электрон покидает металл, только если его энергия выше некоторого порогового значения. Пороговое значение называют работой выхода χ. Это – константа, различная для разных металлов, может служить характеристикой металла. *Работа выхода представляет собой ту минимальную энергию, которую должен затратить электрон, чтобы выйти из металла при температуре абсолютного нуля*. При температуре абсолютного нуля все состояния ниже уровня Ферми заняты и существует поверхностный потенциальный барьер χ (рис. 2.7). Поскольку барьер достаточно высок, лишь незначительная часть электронов может покинуть металл.

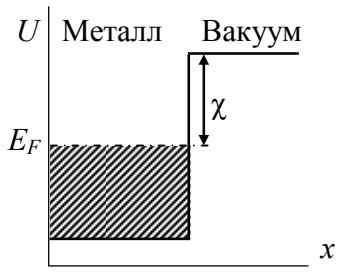


Рис. 2.7 – Энергетическая диаграмма металл- вакуум

Существует простое следствие законов электростатики: сила, действующая на электрон со стороны плоскости с бесконечной проводимостью, не изменится, если эту плоскость заменить зеркальным изображением истинных зарядов, т.е. положительными зарядами, находящимися на том же расстоянии. Сила, действующая между этими зарядами, равна

$$F = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{(2x)^2}. \qquad (2.46)$$

Потенциальную энергию можно найти, интегрируя выражение (2.46) от x до бесконечности, т. е.

$$U(x) = \int_0^{\infty} F(y)dy = -\frac{e^2}{16\pi\varepsilon_0 x}. \qquad (3.47)$$

Здесь, следуя принятой в электростатике договоренности, мы приняли за нуль потенциальную энергию при $x = \infty$. Однако, ранее за нулевую точку потенциальной энергии принималась энергия покоя валентного электрона. Это противоречие легко устранимо. Продемонстрируем это на энергетических диаграммах. На рис. 2.8, *а* показана энергетическая диаграмма на границе раздела металл-вакуум. Зеркальные заряды изменяют профиль потенциального барьера (рис. 2.8, *б*).

Когда электрическое поле отсутствует, изменение крутизны потенциального барьера ничем себя не проявляет. Однако стоит ему появиться, и незначительные поправки, обусловленные силами зеркальных зарядов, становятся существенными.

Для простоты рассмотрим случай, когда электрическое поле однородно в пространстве. Тогда (рис. 2.8, *в*)

$$U(x) = -eEx. \qquad (2.48)$$

Если принять во внимание и электрическое поле, и зеркальные заряды, то потенциальный барьер примет вид, указанный на рис. 2.8, *г*. В точке,

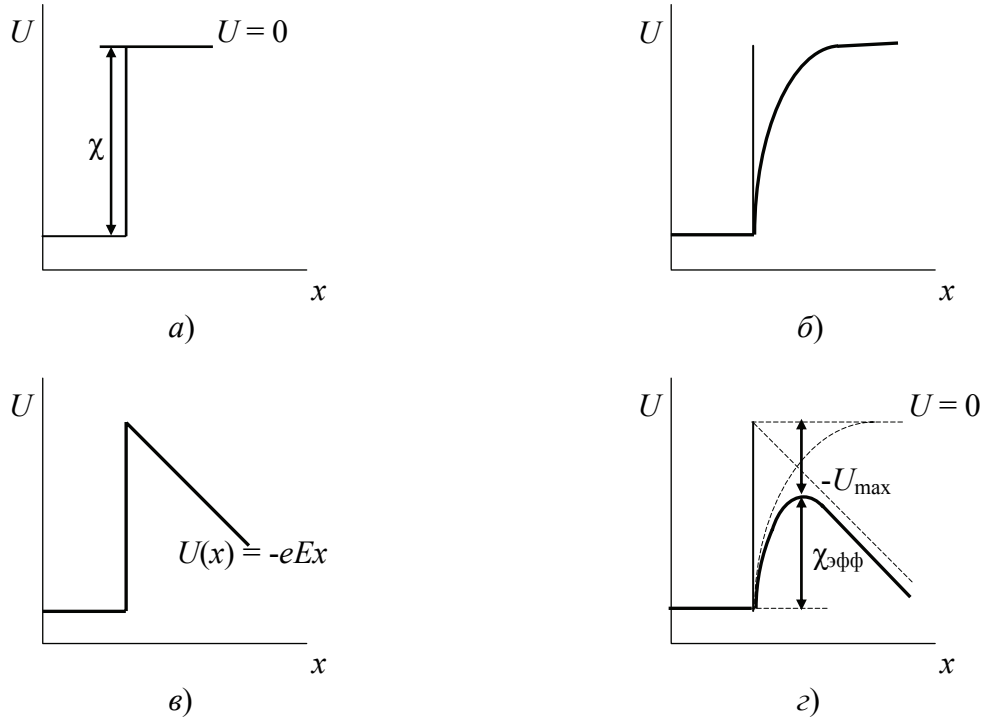


Рис. 2.8 – Эффект Шоттки: *а* – потенциальный барьер на границе раздела металл-вакуум; *б* – потенциал с учетом зеркальных зарядов; *в* – распределение потенциалов на границе раздела металл-вакуум при наличии разности потенциалов; *г* – суммарное распределение потенциала с учетом зеркального изображения и разности потенциалов

являющейся корнем уравнения

$$\frac{d}{dx}\left(-\frac{e^2}{16\pi\varepsilon_0 x} - eEx\right) = 0, \qquad (2.49)$$

существует максимум

$$U_{\max} = -e\left(\frac{eE}{4\pi\varepsilon_0}\right)^{\frac{1}{2}}. \qquad (2.50)$$

Следовательно, высота потенциального барьера, который должен преодолеть электрон, понижается на величину $-U_{\max}$. Эффективная работа выхода уменьшается от χ до $\chi_{эфф}$

$$\chi_{эфф} = \chi - e\left(\frac{eE}{4\pi\varepsilon_0}\right)^{\frac{1}{2}}. \qquad (2.51)$$

Подобное уменьшение работы выхода известно как *эффект Шоттки*.

2.1.6. Контакт двух металлов. Термоэлектродвижущая сила (термо-ЭДС)

При соприкосновении двух различных металлов между ними возникает разность потенциалов, которая называется *контактной разностью*

потенциалов. Это явление было открыто Вольтой в 1797 г. Исследуя различные металлы, Вольта расположил их в следующей ряд: Al, Zn, Sn, Pb, Sb, Bi, Hg, Fe, Cu, Ag, Au, Pt, Pd. Если привести их в контакт в указанной последовательности, то каждый предыдущий при соприкосновении с одним из следующих зарядится положительно. Этот ряд называется рядом Вольта.

Вольта установил два экспериментальных закона:

1. Контактная разность потенциалов зависит лишь от химического состава и температуры соприкасающихся металлов.

2. Контактная разность потенциалов последовательно соединенных различных металлов, находящихся при одинаковой температуре, не зависит от химического состава промежуточных проводников, и равна контактной разности потенциалов, возникающей при непосредственном соединении крайних проводников.

Контактная разность потенциалов для различных металлов составляет от десятых до целых вольт.

Для объяснения возникновения контактной разности потенциалов воспользуемся представлениями зонной теории. Рассмотрим контакт двух металлов с различными работами выхода, т.е. различными положениями уровней Ферми.

Термодинамическая работа выхода электронов из металлов A и B равняется, соответственно, χ_A и χ_B. Пусть, для определенности, $\chi_A > \chi_B$, тогда Уровень Ферми в металле B выше уровня Ферми в металле A (рис. 2.9, *а*). Следовательно, при контакте металлов электроны с более

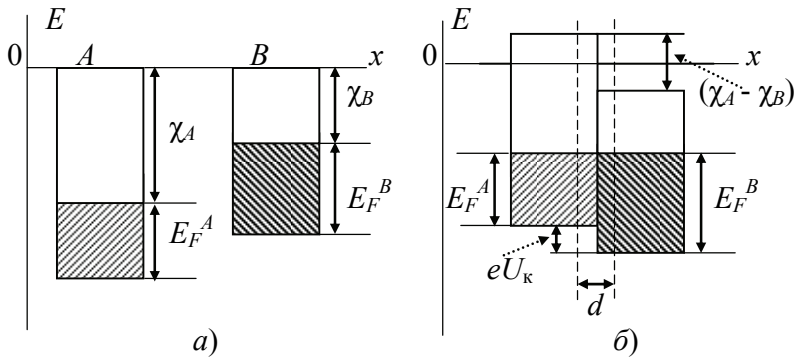


Рис. 2.9 – Энергетическая диаграмма контакта двух металлов: *а* – диаграмма двух изолированных металлов; *б* – диаграмма контакта двух металлов

высоких уровней в металле B будут переходить на более низкие уровни металла A, что приведет к тому, что металл B зарядится положительно, а металл A – отрицательно. Одновременно произойдет относительное смещение энергетических уровней: в металле, заряженном положительно, все уровни будут смещаться вниз, а в металле, заряженном отрицательно – вверх. Этот процесс будет происходить до тех пор, пока между соприкасающимися металлами не установится равновесие, которое характеризуется совпадением уровней Ферми в обоих металлах (рис. 2.9, *б*).

Итак, если уровни Ферми двух металлов не одинаковы, то между ними в случае их контакта возникает *внутренняя контактная разность потенциалов* $U_к$, которая блокирует дальнейший переход носителей заряда. Равновесие наступает при условии:

$$eU_к = E_F^B - E_F^A. \qquad (2.52)$$

Наличие контактного поля обеспечивает равновесие потоков электронов из одного металла в другой. Равновесие вследствие большой скорости теплового движения устанавливается очень быстро (приблизительно за 10^{-16} с). Двойной слой d, который возникает при этом в области контакта, будет очень тонким ~ 10^{-10} м, приблизительно равным периоду решетки, поэтому он не влияет на прохождение электрического тока через контакт.

Термоэлектрическими явлениями принято называть эффекты возникновения в проводниках электродвижущих сил и электрических токов под действием тепловых потоков и эффекты возникновения теплоты, дополнительной к джоулевой, при протекании электрического тока. В физике твердого тела существуют 12 термоэлектрических явлений, однако в упрощенном представлении обычно ограничиваются тремя: *эффектом Зеебека, эффектом Пельтье* и *эффектом Томсона.*

Эффект Зеебека (1826 г.) проявляется в том, что в электрической цепи, состоящей из последовательно соединенных различных проводников, контакты (спаи) которых имеют различную температуру, протекает электрический ток. Возможность возникновения термоэлектрического тока видно из формулы (2.52). Дело в том, что энергии Ферми в металлах, хотя и незначительно, но все же зависят от температуры. Поэтому, если температуры спаев будут неодинаковыми, то будут неодинаковыми и внутренние контактные разности потенциалов. Это проводит к нарушению электрического равновесия и возникновению термоэлектрического тока. Кроме того, так как средняя энергия носителей заряда с ростом температуры увеличивается, то градиент температуры вызывает диффузионный ток электронов от горячего конца к холодному.

В разомкнутой цепи в стационарном состоянии плотность тока в любой точке проводника равна нулю. Это происходит потому, что перераспределение носителей заряда в проводнике приводит к появлению электрического поля, которое компенсирует поток носителей, пропорциональный градиенту температуры. При этом в цепи возникает термо-ЭДС. В простейшем случае, когда такая цепь состоит из двух различных проводников, она называется *термопарой.*

Как показывает опыт, в относительно небольшом температурном интервале термо-ЭДС прямо пропорциональна разности температур контактов:

$$U \approx \alpha_T (T_2 - T_1), \qquad (2.53)$$

где α_T – относительная удельная термо-ЭДС. Абсолютные удельные коэффициенты термо-ЭДС α_A и α_B определяются только косвенно, по измеренной теплоте Пельтье и Томсона. При измерении разности

потенциалов на концах проволоки путем подсоединения вольтметра, надо учесть, что подводящие провода вольтметра быстро придут в термодинамическое равновесие с концами проволоки. Поэтому результатом измерения будет относительная термо-ЭДС, так как возникает встречная ЭДС соединительных проводов.

Эффект Пельтье. В 1834 г. Пельтье обнаружил, что в месте контакта разных проводников *A* и *B* при прохождении через него тока, в дополнение к теплоте Джоуля выделяется или поглощается (в зависимости от направления тока) некоторое количество теплоты, называемое теплотой Пельтье. В отличие от теплоты Джоуля, которая пропорциональна квадрату силы тока, теплота Пельтье, выделяющаяся (поглощающаяся) в единицу времени, пропорциональна первой степени силы протекающего тока:

$$Q_{П} = \pi_{AB} I \,, \tag{2.54}$$

где π_{AB} коэффициент, имеющий размерность $Вт \cdot А^{-1} = В$; I – сила тока. Таким образом, явление Пельтье является обратным по отношению к явлению Зеебека.

Объяснить эффект Пельтье можно следующим образом. Электроны по разные стороны спая обладают различной средней энергией. Следует отметить, что имеется в виду именно полная средняя энергия электронов, равная сумме их потенциальной и кинетической энергий. Если электроны пройдут через один спай и попадут в область низких энергий, то избыток своей энергии они отдадут кристаллической решетке, и спай будет нагреваться. В другом спае электроны переходят в область больших энергий и забирают недостающую энергию у кристаллической решетки – спай будет охлаждаться.

Обратимое выделение или поглощение теплоты Пельтье зависит только от свойств каждого проводника термопары, т.е. от абсолютных коэффициентов Пельтье каждого из проводников. В температурных измерениях это явление может играть существенную роль только при достаточно больших плотностях тока в измерительных цепях.

Эффект Томсона (экспериментально установленный Леру в 1867 г.) состоит в следующем. Если вдоль проводника, по которому протекает электрический ток, существует градиент температуры $\dfrac{dT}{dx}$ (неравномерно нагретый проводник), то в дополнение к теплоте Джоуля в объеме проводника выделяется (или поглощается) в зависимости от направления тока некоторое количество теплоты (теплота Томсона). Найдем теплоту Томсона, выделяющуюся в единицу времени на участке провода длиной *dx*:

$$dQ = \tau I \frac{dT}{dx} dx \,, \tag{2.55}$$

где τ – коэффициент Томсона, размерность которого $\dfrac{\text{Вт}}{\text{А} \cdot \text{К}} = \dfrac{\text{В}}{\text{К}}$.

Это явление можно объяснить так. В более нагретой части проводника электроны имеют бóльшую среднюю энергию, чем в менее нагретой части

проводника. Двигаясь в направлении убывания температуры, они отдают часть своей энергии решетке, в результате чего и происходит выделение теплоты Томсона. Если же электроны движутся в сторону возрастания температуры, то они, наоборот, пополняют свою энергию за счет энергии решетки, в результате чего происходит поглощение теплоты Томсона.

Применив к трем указанным термоэлектрическим явлениям первое и второе начала термодинамики, Томсон (1856 г.) вывел следующие соотношения между коэффициентами α_T, π и τ:

$$\tau = T\frac{d\alpha_T}{dT}; \qquad \pi = T\alpha_T . \tag{2.56}$$

2.1.7. Особенности сопротивления проводников на высоких частотах

На высоких частотах имеет место неравномерное распределение электрического тока по сечению проводника: плотность тока максимальна на поверхности и уменьшается в глубине проводника. Такое явление называют *поверхностным эффектом* (скин-эффектом). Неравномерное распределение тока связано с влиянием магнитного поля, создаваемого самим проводником. Сцепленный с проводником магнитный поток пропорционален току:

$$\Phi = L\cdot I , \tag{2.57}$$

где L – индуктивность проводника; I – сила тока.

Изменение магнитного потока вызывает появление вихревого электрического поля $\vec{E}_{\text{вихр}}$ и ЭДС самоиндукции.

$$\varepsilon_i = -L\cdot dI/dt. \tag{2.58}$$

Если ток изменяется по синусоидальному закону $I = I_m\cdot\sin\omega t$, то индуцированная электродвижущая сила пропорциональна частоте ω:

$$\varepsilon_i = -\omega\cdot L\cdot I_m\cdot\cos\omega t. \tag{2.59}$$

Рассмотрим появление скин-эффекта в момент времени, когда переменный ток через проводник увеличивается ($\Delta I > 0$). Увеличение тока приводит к возрастанию создаваемого им магнитного поля с индукцией $\vec{B}$. Рост индукции $\vec{B}$ приводит к появлению ЭДС самоиндукции ε_i и вихревого электрического поля $\vec{E}_{\text{вихр}}$, магнитное поле $\vec{B}'$ которого согласно правилу Ленца направлено навстречу $\vec{B}$. Вихревое поле $\vec{E}_{\text{вихр}}$ вызывает появление тока, имеющего направление, указанное на рис. 2.10. Это приводит к увеличению тока I у поверхности и препятствует увеличению тока на оси проводника.

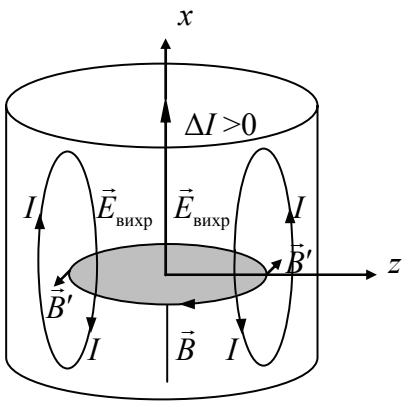


Рис. 2.10 – Иллюстрация скин-эффекта в случае увеличения тока через проводник

Если ток уменьшается ($\Delta I < 0$), то

поле $\vec{E}_{\text{вихр}}$ у поверхности еще больше уменьшает ток I, а на оси препятствует уменьшению I. В обоих случаях вихревое поле $\vec{E}_{\text{вихр}}$ усиливает изменение тока на поверхности и ослабляет изменение тока на оси проводника, следовательно, переменный ток на поверхности увеличивается, а на оси уменьшается. Роль вихревых токов существенно возрастает с увеличением частоты тока, протекающего через проводник (см. ф-лу (2.59)). На высоких частотах ток течет в основном в приповерхностном слое проводника, что эквивалентно уменьшению его поперечного сечения и увеличению сопротивления, а это нежелательно из-за дополнительных энергетических потерь. Для уменьшения сопротивления поверхность высокочастотных проводов обычно покрывают тонким слоем высокопроводящего металла, например, серебром.

Если в проводнике принять направление тока за ось x, а нормаль к поверхности за ось z, тогда распределение тока по сечению проводника будет иметь вид:

$$j_x(z) = j_0 e^{-\frac{z}{\delta}}, \tag{2.60}$$

где j_0 – плотность тока на поверхности; δ – глубина проникновение тока в проводник (толщина скин-слоя). Толщина скин-слоя δ – это поверхностная толщина проводника, за пределами которой плотность тока уменьшается в e раз

$$\delta = \sqrt{\frac{2}{\omega\sigma\mu_0\mu}} = \sqrt{\frac{1}{\pi f\sigma\mu_0\mu}} = \sqrt{\frac{\rho}{\pi f\mu_0\mu}}, \tag{2.61}$$

где μ_0 – магнитная постоянная; μ – магнитная проницаемость; σ – удельная электропроводность; ω – циклическая частота переменного тока; f – частота переменного тока; ρ – удельное сопротивление проводника.

В радиотехнике для плоских проводников используют специальную характеристику – сопротивление квадрата поверхности R_S, которая определяется из выражения:

$$R_S = \frac{\rho}{\delta}, \tag{2.62}$$

где ρ – удельное сопротивление металла. Выражение (2.62) показывает, что активное сопротивление R_S плоского проводника бесконечной толщины при скин-эффекте равно сопротивлению плоского проводника толщиной δ для постоянного тока.

2.1.8. Сопротивление тонких металлических пленок

Металлические пленки широко используются в микроэлектронике, как межэлементные соединения, контактные плоскости, обкладки конденсаторов, магнитных и резистивных элементов интегральных микросхем (ИМС). Электрические свойства *тонких пленок металлов* значительно отличаются от свойств объемных образцов. Это зависит в

первую очередь от структуры пленки, которая может изменяться от неупорядоченного мелкодисперсного состояния (амфотерный конденсат) к достаточно совершенному состоянию (эпитаксийные пленки). Второй причиной является появление размерных эффектов, связанных с поверхностью, если толщина пленки соизмерима со средней длиной свободного пробега электронов. В этом случае удельное сопротивление будет зависеть от геометрических размеров образца. Структура пленки имеет значительные изменения на разных этапах их конденсации. Для большинства пленок, в зависимости удельного сопротивления от толщины пленки, наблюдаются три участка (рис. 2.11). Участок 1 соответствует толщине пленки больше 0,1 мкм. На этом участке удельное сопротивление пленки приблизительно равняется объемному сопротивлению образца. Участок 2 охватывает толщины пленки от 0,1 до 0,01 мкм. На этом участке удельное сопротивление пленки выше объемного, а температурный коэффициент удельного сопротивления α_ρ приближается к нулю. На участке 3 толщиной меньше 10^{-3} мкм удельное сопротивление очень значительно, а температурный коэффициент α_ρ отрицателен.

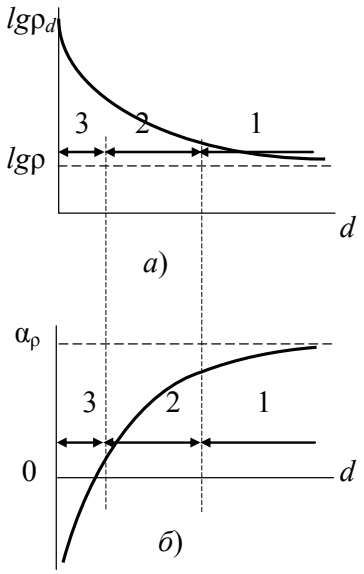


Рис. 2.11 – Зависимости удельного сопротивления (*а*) и температурного коэффициента удельного сопротивления (*б*) тонкой металлической пленки от ее толщины

Для объяснения этого необходимо учесть, что тонкие пленки на первых стадиях конденсации имеют *островковую структуру*, т.е. при малом количестве осажденного металла его частицы располагаются на диэлектрической подложке в виде разрозненных зерен – островков. Электропроводность пленки возникает при некотором минимальном количестве осажденного металла, однако, еще до образования соединительных мостиков между островками металла. При приложении электрического поля в плоскости пленки происходит переход электронов через узкие диэлектрические зазоры между соседними островками за счет термоэлектронной эмиссии и туннелирования (в частности, туннелировать могут электроны, расположенные выше уровня Ферми). Переход электронов облегчается с увеличением температуры. Кроме того, удельное сопротивление островковой структуры определяется поверхностным сопротивлением участков диэлектрической подложки, где нет зерен металла, а поверхностное сопротивление диэлектриков с увеличением температуры падает. Эти причины и обусловливают отрицательный знак температурного коэффициента α_ρ пленок малой толщины.

По мере роста количества осажденного металла, зазоры между островками уменьшаются, и проводимость пленки возрастает, отрицательный α_ρ становится меньше по модулю, а затем меняет знак. Значение толщины пленки, при которой происходит смена знака α_ρ, зависит от рода металла, условий формирования пленки, состояния поверхности подложки и в реальных случаях составляет несколько нанометров.

В процессе дальнейшей конденсации вещества на подложке происходит слияние островков и образование проводящего канала, а затем – сплошного однородного слоя. Но и в сплошной пленке сопротивление больше, чем сопротивление исходного проводника, что является следствием высокой концентрации дефектов структуры: дислокаций, межзерновых границ, вакансий и др.

Возрастанию удельного сопротивления способствует и *размерный эффект*, т.е. уменьшение длины свободного пробега электронов вследствие их отражения от поверхности образца. При комнатной температуре поверхностное рассеяние электронов оказывает значительное влияние на бóльшую часть пленок из чистых металлов, если их толщина меньше 200…300 $\mathring{A}$. Однако, при низких температурах, когда длина свободного пробега электронов в объеме проводящего материала существенно возрастает, влияние размерных эффектов проявляется при гораздо бóльших толщинах пленки.

2.1.9. Явление сверхпроводимости

Сверхпроводимость является одним из фундаментальных свойств вещества. Сверхпроводимость проявляется во многих металлах и сплавах при достаточно низких температурах. Температуры перехода в сверхпроводящее состояние для разных чистых металлов лежат в пределах от 0,01 К (вольфрам) до 9,2 К (ниобий). Изменение сопротивления проходит скачком в узком температурном интервале (10^{-3}…10^{-4} К), может падать более чем на 15 порядков и составлять $5\cdot10^{-18}$ мкОм·м.

Впервые сверхпроводимость была обнаружена у ртути ($T_{св}$ = 4,2 К) голландским физиком Х. Каммерлинг-Оннесом в 1911 г. Обычная методика состоит в том, что в кольце, выполненном из сверхпроводящего материала, возбуждают ток, а затем измеряют магнитное поле этого тока. В нормальных металлах ток исчезает через 10^{-12} с. В сверхпроводнике ток может циркулировать в течение времени, измеряемого не пикосекундами, а годами. Один из самых длительных экспериментов был проведен в США; ток циркулировал без видимого уменьшения в течение трех лет. По скорости уменьшения магнитного поля в кольце была произведена оценка удельного сопротивления материалов в сверхпроводящем состоянии. Его значение составило около 10^{-25} Ом, что в 10^{17} раз меньше сопротивления меди при комнатной температуре.

Физическая природа сверхпроводимости была понята только в 1957 г. Еще задолго до этого советским физиком-теоретиком Л.Д. Ландау была создана теория сверхтекучести гелия-2. Оказалось, что сверхтекучесть – это макроскопический квантовый эффект. Однако перенести теорию Ландау на явление сверхпроводимости мешало то обстоятельство, что атомы гелия, обладая нулевым спином, подчиняются статистике Бозе-Эйнштейна. Электроны же, обладая полуцелым спином, подчиняются принципу Паули и статистике Ферми-Дирака. Для таких частиц невозможна бозе-эйнштейновская конденсация, необходимая для возникновения сверхтекучести. Решающий шаг был сделан американскими учеными Д. Бардином, Л. Купером, Д. Шриффером. В 1957 г. они создали теорию (называемую сегодня теорией БКШ), которая сумела объяснить все главные особенности сверхпроводника. Сущность теории заключается в том, что сверхпроводимость вызывается электронно-решетчатым взаимодействием, а сверхпроводящие электроны являются обычными электронами, сгруппированными в пары.

Таким образом, явление сверхпроводимости возникает тогда, когда два электрона могут притягиваться друг к другу. Это возможно только в среде, имеющей положительно заряженные ионы, электрическое поле которых ослабляет силы кулоновского отталкивания между электронами. Притягиваться могут лишь те электроны, которые принимают участие в проводимости, т.е. расположенные вблизи уровня Ферми. Если такое притяжение имеет место, то электроны с противоположным направлением спинов объединяются в кулоновские пары, которые носят название *куперовских пар*. В образовании куперовских пар решающую роль играет взаимодействие электронов с тепловыми колебаниями решетки – фононами.

Один из электронов, взаимодействующий с решеткой, переводит ее в возбужденное состояние и изменяет свой импульс; второй электрон, также взаимодействующий с решеткой, переводит ее в нормальное состояние и тоже изменяет свой импульс. То есть, состояние решетки не изменяется, а электроны обмениваются квантами тепловой энергии – фононами.

Обменное фононное взаимодействие и вызывает силы притяжения между электронами, которые превышают силы кулоновского отталкивания. Электрон, который движется среди положительных ионов, поляризует решетку, то есть притягивает к себе соседние ионы. Тогда в окрестности траектории электрона локально возрастает плотность положительного заряда. Второй электрон, который движется следом за первым, будет притягиваться положительным

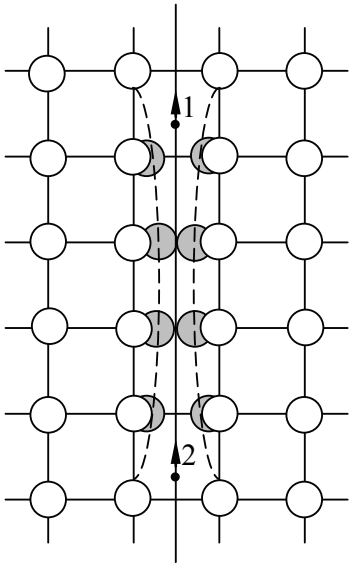


Рис. 2.12 – Возникновение куперовских пар в сверх-проводящем металле

зарядом, поэтому косвенно между электронами 1 и 2 возникнут силы притяжения (рис. 2.12) и эти электроны образуют так называемую куперовскую пару.

Все сверхпроводящие электроны ведут себя совершенно одинаково. До сих пор все электроны, подчиняясь принципу Паули, должны были отличаться друг от друга энергетическими состояниями. В этом случае применялась статистика Ферми-Дирака. В сверхпроводящем состоянии электроны обретают право занимать одинаковые энергетические состояния. Таким образом, имеется огромное число идентичных частиц, подчиняющихся статистике Бозе-Эйнштейна.

Куперовские пары слабо локализованы в пространстве и перекрывают тысячи электрических ячеек. Они беспрерывно возникают и распадаются. Это приводит к возникновению в энергетическом спектре сверхпроводника энергетической щели 2δ (рис. 2.13). Спаренные электроны располагаются на дне энергетической щели. Их количество составляет около 10^{-4} от общего числа электронов. Размер энергетической щели зависит от температуры, которая достигает максимума при абсолютном нуле и исчезает при температуре сверхпроводящего перехода ($T = T_c$). Теория БКШ дает следующую связь ширины щели с критической температурой перехода T_c:

$$2\delta_{0K} = 3{,}5kT_c . \qquad (2.63)$$

Формула (2.63) достаточно хорошо подтверждается экспериментально. Для большинства сверхпроводников энергетическая щель составляет $10^{-4}\ldots10^{-5}$ эВ.

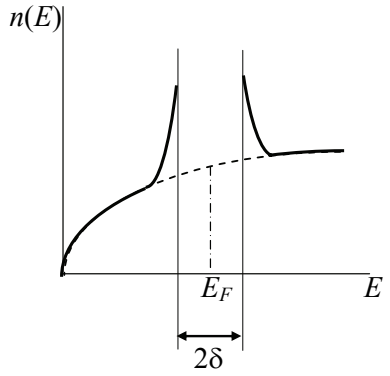


Рис. 2.13 – Распределение электронов по энергиям в металле в состоянии сверхпроводимости

В обычных условиях электрическое сопротивление металла обусловлено рассеиванием электронов на тепловых колебаниях решетки и примесях. Но при наличии энергетической щели, для перехода электронов из обычного состояния в возбужденное состояние, нужна достаточная порция тепловой энергии, которую при низких температурах электроны не могут получить от решетки, поскольку энергия тепловых колебаний меньше ширины щели. Поэтому куперовские пары не рассеиваются на дефектах структуры. Особенностью куперовских пар является их импульсная упорядоченность, состоящая в том, что все пары имеют одинаковый импульс и не могут изменять свои состояния независимо друг от друга. Электронные волны, которые описывают движение пар, имеют одинаковую длину и фазу. Фактически движение всех куперовских пар можно рассматривать как распространение одной электронной волны, которая не рассеивается решеткой, а “обтекает” дефекты решетки. Такая согласованность в поведении пар обусловлена высокой мобильностью

электронного конденсата: непрерывно меняются наборы пар, происходит постоянная смена партнеров.

При абсолютном нуле, все электроны, расположенные вблизи уровня Ферми, связаны в пары. С повышением температуры за счет тепловой энергии происходит разрыв некоторой части электронных пар, вследствие чего уменьшается ширина щели. Движение неспаренных электронов, переходящих с основных уровней на возбужденные, затрудняется рассеянием на дефектах решетки. При температуре $T = T_c$ происходит полный разрыв всех пар, ширина щели обращается в нуль, сверхпроводимость исчезает.

Переход вещества в сверхпроводящее состояние происходит в узком интервале температур (сотые доли градуса). Наличие примесей и дефектов решетки не уничтожает явление сверхпроводимости, а вызовет только расширение температурного интервала перехода из обычного состояния в сверхпроводящее. Электроны, ответственные за состояние сверхпроводимости, не обмениваются энергией с решеткой. Поэтому при температуре ниже критической наблюдается существенное уменьшение теплопроводности.

Магнитные свойства сверхпроводников. Важной особенностью сверхпроводников является то, что внешнее магнитное поле не проникает в толщу образца и затухает в тонком слое (глубина проникновения поля оценивается в десятые и сотые доли микрона). Силовые линии магнитного поля огибают сверхпроводник (эффект Мейсснера). Эффект Мейсснера связан с тем, что в поверхностном слое сверхпроводника, при его внесении в магнитное поле, возникает круговой незатухающий ток, который полностью компенсирует внешнее магнитное поле в толще образца. Таким образом, сверхпроводники являются идеальными диамагнетиками с магнитной проницаемостью, равной нулю. Как всякие диамагнетики сверхпроводники выталкиваются из магнитного поля. При этом эффект выталкивания выражен столь сильно, что открывается возможность удерживать груз в пространстве с помощью магнитного поля.

Одно из возможных применений сверхпроводимости – это создание мощных магнитов. Однако надежды первых экспериментаторов получить большие магнитные поля без затрат энергии были разрушены. Выяснилось, что сверхпроводник переходит в нормальное состояние, если магнитное поле достигает значения выше некоторого критического уровня H_c. Так что, для получения нулевого сопротивления нужно удерживать ниже определенного порогового значения не только температуру, но и напряженность магнитного поля. Эксперименты с различными сверхпроводниками показали, что зависимость критического магнитного поля от температуры хорошо описывается формулой:

$$H_c = H_0\left(1-\left(\frac{T}{T_c}\right)^2\right), \qquad (2.64)$$

где H_c – критическое значение напряженности магнитного поля, при котором сверхпроводимость разрушается; H_0 – критическое значение напряженности магнитного поля при температуре абсолютного нуля; T_c – критическая температура, при которой сверхпроводимость разрушается. Соответствующая кривая построена на рис. 2.14 и имеет параболическую форму. На рисунке видно, что выше критической кривой проводник находится в нормальном состоянии, ниже – в сверхпроводящем. Сверхпроводимость разрушается также и электрическим током, когда последний превосходит известный предел (I_c – критический ток). Однако этот эффект является простым следствием рассмотренного выше эффекта. Действительно, ток, протекающий по сверхпроводнику, возбуждает магнитное поле. Когда поле достигает критического значения, начинается разрушение сверхпроводимости. К примеру, для сверхпроводников 1-го рода предельная плотность тока обусловлена достижением на поверхности образца критической напряженности магнитного поля. В случае длиной прямолинейной проволоки круглого сечения радиуса r предельный ток определяется формулой:

$$I_c = 2\pi r H_c(T). \qquad (2.65)$$

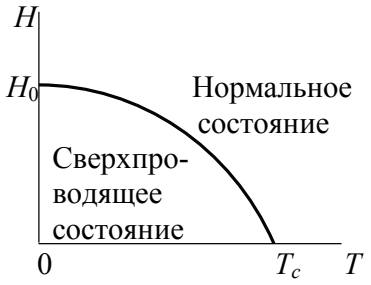


Рис. 2.14 – Зависимость критического магнитного поля от температуры сверхпроводников 1-го рода

Поскольку в сверхпроводящих элементах ток проходит в тонком поверхностном слое, средняя плотность тока, отнесенная ко всему поперечному сечению, уменьшается с увеличением диаметра провода.

Форма образца играет важную роль при распаде сверхпроводящего состояния. Например, в цилиндрических образцах, если магнитное поле однородно и направлено вдоль оси цилиндра, разрушение сверхпроводящего состояния при достижении H_c происходит одновременно во всем объеме образца. Образец целиком перейдет в нормальное состояние. Это связано с тем, что магнитное поле на поверхности имеет одинаковое значение во всех точках.

Сложнее себя ведут образцы другой формы, так как в этих случаях критическое поле H_c в различных точках поверхности образца достигается при различных значениях магнитного поля. Если образец имеет форму шара, то разрушение сверхпроводимости в магнитном поле происходит путем «распада» шара на чередующиеся тонкие слои из сверхпроводящих и нормальных фаз. В случае тел более сложной формы распад происходит на малые области (*домены*) из сверхпроводящих и нормальных фаз. Такое состояние распада сверхпроводника на сверхпроводящие и нормальные домены, связанное с его формой и обусловленное магнитным полем, называется *промежуточным состоянием*. Его существование было доказано прямыми опытами.

Одним из существенных факторов, определяющим поведение сверхпроводника, является *поверхностная энергия*, связанная с наличием границ между нормальной и сверхпроводящей фазой. Эта энергия аналогична энергии поверхностного натяжения на границе раздела двух жидкостей. Однако в случае сверхпроводников поверхностная энергия может быть как положительной, так и отрицательной. Сверхпроводники, для которых поверхностная энергия положительна, называют *сверхпроводниками* 1-*го рода*, а сверхпроводники с отрицательной поверхностной энергией – *сверхпроводниками* 2-*го рода*.

Допустим, что сверхпроводящий образец имеет форму длинного цилиндра. В таком случае никакого промежуточного состояния, связанного с формой не возникает. Однако при возрастании магнитного поля сверхпроводники 1-го и 2-го рода будут вести себя по-разному. Если внешнее приложенное магнитное поле достигает критического значения H_c, то у сверхпроводников проводящая фаза становится неустойчивой и весь образец переходит в более устойчивую фазу (нормальную). Дробление сверхпроводника на малые и сверхпроводящие домены не наблюдается, так как на образование границ требуется дополнительная энергия. В соответствии с этим на рис. 2.14 показаны только две области: сверхпроводящая и нормальная.

Совершенно другой вид будет иметь подобный график в случае сверхпроводников 2-го рода. Так как у них поверхностная энергия отрицательна, то имеется возможность понижения полной свободной энергии путем дробления образца на сверхпроводящие и нормальные домены. Такое дробление не связано с формой образца, а является внутренним свойством сверхпроводников 2-го рода. Кроме того, в отличие от промежуточного состояния, где доменная структура сравнительно грубая и видна невооруженным глазом, домены в сверхпроводниках 2-го рода значительно меньше (порядка 10^{-7}м и менее). Описанное состояние сверхпроводника 2-го рода, называется *смешанным состоянием*. На рис. 2.15 различают три области: *сверхпроводящую*, *область смешанного состояния* и *нормальную область*. Границы между этими областями определяются нижней H_{c1} и верхней H_{c2} критическими напряженностями магнитного поля.

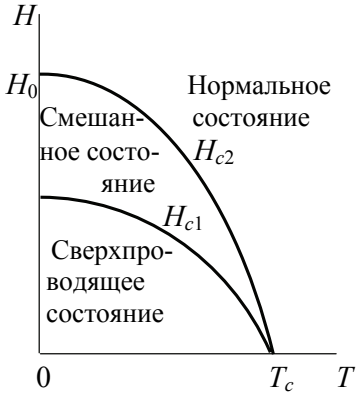


Рис. 2.15 – Зависимость критического магнитного поля от температуры для сверхпроводников 2-го рода

Сверхпроводники отличаются также и по характеру перехода металла из обычного состояния в сверхпроводящее состояние (в зависимости от магнитного поля). У сверхпроводников 1-го рода этот переход происходит скачком (рис. 2.16, *а*), а у сверхпроводников 2-го рода – постепенно (рис.

2.16, *б*). Критическая напряженность магнитного поля зависит от температуры. При $T = T_c$ она равняется нулю и возрастает до максимума при $T = 0$ К. Критическая напряженность для сверхпроводников 1-го рода достигает 10^5 А/м, а 2-го рода – 10^7 А/м. Для сверхпроводников 2-го рода формула (2.65) не выполняется и связь между I_c и H_c носит более сложный характер.

Ранее утверждалось, что критическое значение напряженности магнитного поля разрушало проводимость. Это было сделано на основе представлений о диамагнитном характере материала. Критическое поле было определено только для случая, когда магнитное поле выталкивается.

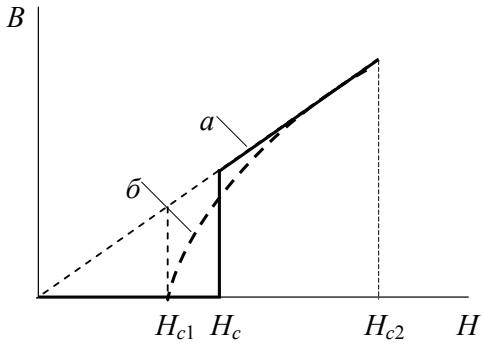


Рис. 2.16 – Зависимость индукции магнитного поля внутри сверхпроводника от напряженности внешнего магнитного поля для сверхпроводников 1-го рода (*а*) и 2-го рода (*б*)

Однако для некоторых материалов сверхпроводимость может существовать и при бóльших напряженностях магнитных полей. Вплоть до значений H_{c1} сверхпроводник является диамагнетиком (рис. 2.17), где намагниченность ($-J$) построена как функция внешнего поля. При напряженностях, превышающих H_{c1}, магнитное поле начинает проникать в материал (за «диамагнитную» глубину проникновения), а при H_{c2}, когда материал становится нормальным, происходит полное проникновение вглубь. Материалы, описываемые такой кривой намагниченности, относятся к сверхпроводникам 2-го рода. Те же сверхпроводники, которые выталкивают магнитное поле до тех пор, пока не перейдут в нормальное состояние, относятся к сверхпроводникам 1-го рода.

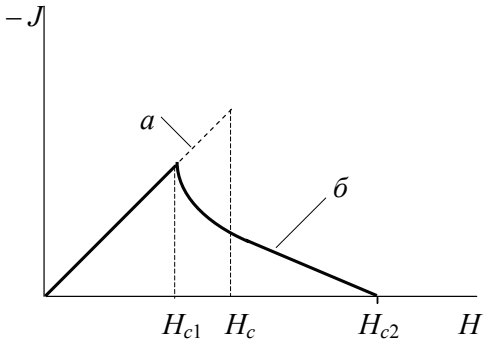


Рис. 2.17 – Кривые намагниченности для для сверхпроводников 1-го рода (*а*) и 2-го рода (*б*). Площади под обеими кривыми одинаковы

Двумерный анализ сверхпроводников 2-го рода показывает, что интенсивность магнитного поля изменяется периодически, имея отчетливо выраженные максимумы. В области максимумов магнитного поля концентрация сверхпроводящих электронов равна нулю. Можно утверждать, что в сверхпроводниках 2-го рода имеются периодически чередующиеся нормальные области, окруженные вихревыми токами. Число таких вихрей велико, они отстоят друг от друга приблизительно на 1мкм. Эти области намагниченности в сверхпроводниках можно увидеть под микроскопом, если посыпать поверхность сверхпроводника порошком из железа.

Эффекты Джозефсона. В 1962 году Б. Джозефсоном (Нобелевская премия 1973 г.) была предсказана возможность проявления двух необычных эффектов, которые вскоре были экспериментально обнаружены. Различают *стационарный* и *нестационарный эффекты Джозефсона*.

Стационарный эффект состоит в том, что, если создать туннельный контакт, образованный двумя сверхпроводниками, разделенными тонким слоем (порядка 10^{-9} м) диэлектрика, то через контакт будет протекать постоянный ток без источника тока. Электроны проводимости проходят через диэлектрик благодаря туннельному эффекту, т.е. диэлектрик становится сверхпроводником. Причина этого состоит в том, что туннелируют куперовские пары, а не отдельные электроны. Если ток через контакт Джозефсона не превышает определенного значения, называемого *критическим током контакта*, то на контакте отсутствует падение напряжения. Такое явление характерно только для стационарного эффекта Джозефсона. (Напомним, что в обычных металлах, при отсутствии в них электрического поля, макроскопического тока нет).

Если же через контакт пропускать ток, больший критического, то на контакте возникнет падение напряжения U и контакт начнет излучать электромагнитные волны. В этом случае реализуется *нестационарный эффект Джозефсона*. Излучение электромагнитных волн возможно только переменным током. Следовательно, через контакт Джозефсона при постоянном напряжении протекает переменный ток. Частота излучения ν связана с падением напряжения U соотношением $\nu = \dfrac{2eU}{h}$, где e – заряд электрона. Причина излучения состоит в следующем. Объединенные в пары электроны, создающие сверхпроводящий ток, при переходе через контакт приобретают дополнительную по отношению к основному состоянию энергию $2eU$. Единственная возможность для пары электронов вернуться в основное состояние сверхпроводника – это излучить фотон с энергией $h\nu = 2eU$, откуда $\nu = \dfrac{2eU}{h}$.

Нестационарный эффект Джозефсона является экспериментальным доказательством существования электронных пар в сверхпроводниках. Об этом говорит удвоенный заряд электрона в выражении для частоты ν-излучения.

Если соединить параллельно два джозефсоновских перехода, то максимальный сверхпроводящий ток, который может протекать через них, окажется периодической функцией магнитного потока, т.е.

$$I_{\max} = 2I_j \left| \cos \frac{\pi \Phi}{\Phi_0} \right|, \qquad (2.66)$$

где I_j – постоянная величина тока, зависящая от параметров перехода; Ф – захваченный магнитный поток, а Φ_0 – так называемый квант потока, равный $\frac{h}{2e} = 2 \cdot 10^{-15}$ Вб.

Сверхпроводниковые материалы. При низких температурах сверхпроводимостью обладают 26 металлов, но для большинства из них критическая температура очень мала: T_c < 4,2 К. Например, для ванадия T_c < 5,13 К, для индия T_c < 3,405 К, для ниобия T_c < 9,22 К, для свинца T_c < 7,2 К. Самой высокой критической температурой среди химических элементов обладает ниобий (T_c = 9,22 К), а наиболее низкой – иридий (T_c = 0,140 К). Сверхпроводимость – это свойство не отдельных атомов, а коллективный эффект, связанный со структурой всего образца. Об этом свидетельствует целый ряд фактов. Например, серое олово – полупроводник, а белое олово – металл, переходящий в сверхпроводящее состояние при температуре 3,72 К. Две кристаллические модификации лантана (α-La и β-La) имеют разные критические температуры перехода в сверхпроводящее состояние (для α-La T_c = 4,8 К, для β-La T_c = 5,95 К).

Все металлы являются сверхпроводниками 1-го рода, а металлические сплавы – сверхпроводниками 2-го рода. Следует заметить, что сверхпроводимостью не обладают металлы, являющиеся наилучшими проводниками в нормальных условиях. К ним относятся золото, медь, серебро. Малое сопротивление этих материалов указывает на слабое взаимодействие электронов с решеткой. Такое слабое взаимодействие не создает вблизи абсолютного нуля достаточного межэлектронного притяжения, способного преодолеть кулоновское отталкивание, поэтому и не происходит их перехода в сверхпроводящее состояние.

Металлические сплавы имеют более высокую температуру перехода в сверхпроводящее состояние. Например, V_3Ga – 14,8 К, Nb_3Ga – 20,3 К, Nb_3Sn – 21...24,3 К, Nb_3Ge – 22,3 К. Температура перехода таких сплавов ниже температуры жидкого водорода (22 К). Это ограничивает использование сверхпроводников.

Деление веществ на сверхпроводники 1-го и 2-го рода не является абсолютным. Любой сверхпроводник 1-го рода можно превратить в сверхпроводник второго рода, если создать в нем достаточную концентрацию дефектов кристаллической решетки. Например, у чистого олова T_c = 3,7 К, но если вызвать в нем резко неоднородную механическую деформацию, то критическая температура возрастет до 9 К, а критическая напряженность магнитного поля увеличится в 70 раз.

Сверхпроводимость никогда не наблюдается в системах, в которых существует ферро- и антиферромагнетизм. Образованию сверхпроводящего состояния в полупроводниках и диэлектриках препятствует малая концентрация свободных электронов. Однако в материалах с большой диэлектрической проницаемостью силы кулоновского отталкивания между электронами в значительной степени ослаблены. Поэтому некоторые из них также проявляют свойства сверхпроводников при низких температурах.

Примером может служить титанит стронция ($SrTiO_3$). Такие полупроводники, как кремний, германий, теллур, сурьма и др. проявляют сверхпроводящие свойства при высоких давлениях. Ряд полупроводников, таких как GeTe, SnTe, CuS, Au_2Bi и др., удается перевести в сверхпроводящее состояние добавкой большой концентрации легирующих примесей.

В 1986 году К. Мюллер и Дж. Беднорец в Цюрихе (Швейцария) открыли керамику барий-лантан-оксид меди ($La_{1,35}Ba_{0,15}CuO_4$) с критической температурой T_c = 35 К. Через год, в 1987 году, в Японии и США была обнаружена сверхпроводимость керамики иттрий-барий-оксид меди ($YBa_2Cu_3O_7$) с критической температурой между 90 и 100 К. Сверхпроводимостью обладают и керамические материалы состава $La_{1,8}Sr_{0,2}CuO_4$. Температура перехода в сверхпроводящее состояние таких материалов достигает 100…110 К, тогда как температура жидкого азота – 77 К. Это значительно облегчает использование сверхпроводников.

Попытки создания сверхпроводящей керамики при комнатных температурах не нашли должного подтверждения. Оказывается, трудность состоит в том, что сверхпроводящие материалы очень нестабильны: сверхпроводимость при комнатной температуре может существовать в них около минуты, а затем пропасть и уже не появиться вновь. Большинство исследователей считают, что решающую роль в высокотемпературной сверхпроводимости играет взаимодействие между ионами меди и кислорода. Большое значение имеет и структура кристаллитов, так как сверхпроводимостью обладают лишь некоторые конфигурации, другие же оказываются изоляторами или нормальными металлами. Однако, до настоящего времени реальный механизм появления сверхпроводимости при комнатных температурах не известен.

2.2. КЛАССИФИКАЦИЯ ПРОВОДНИКОВ

К проводникам относят металлы и их сплавы. По характеру применения в радиоэлектронной технике металлические материалы разделяют на металлы *высокой проводимости* и сплавы *высокого сопротивления*.

2.2.1. Материалы с высокой проводимостью

К материалам этого типа предъявляют следующие требования:

удельное сопротивление в нормальных условиях меньше 0,1 мкОм·м; высокие механические свойства (предел прочности при растяжении); способность легко обрабатываться, что необходимо для изготовления проводов малых и сложных сечений; способность давать контакты с малым переходным сопротивлением при пайке, сварке и других методах соединения проводов; коррозионная стойкость.

Требование максимальной удельной проводимости материала является основным. Перечисленным условиям отвечают только медь, алюминий и их сплавы.

Медь. Преимуществами *меди* являются: малое удельное сопротивление (только серебро имеет меньшее удельное сопротивление); достаточно высокая механическая прочность; удовлетворительная коррозийная стойкость. Кроме того, медь хорошо обрабатывается, легко прокатывается в ленты, фольгу, протягивается в проволоку; медь хорошо паяется и сваривается.

Вакуумная медь марки МВ выплавляется в вакуумных индукционных печах и является наиболее чистым проводниковым материалом (наличие примесей не более 0,01%).

Большое влияние на свойства меди оказывает её отжиг. Характерно, что при отжиге более значительно изменяются механические свойства меди и слабее меняется ее удельное сопротивление. Как проводниковый материал используют *твердую медь* (МТ) и *мягкую медь* (ММ). Применение твердой и мягкой меди различно. Твердую медь используют там, где требуется обеспечить высокую механическую прочность, твердость и сопротивляемость истиранию: для изготовления коллекторных пластин электрических машин, для шин распределительных устройств, для контактных проводов и т.п. Мягкую медь в виде проволоки круглого сечения применяют главным образом в виде токопроводящих жил обмоточных проводов, кабелей (где важна гибкость и пластичность, а прочность не имеет существенного значения).

В отдельных случаях помимо чистой меди в качестве проводникового материала применяют ее *сплавы*: *бронзу* (сплавы меди с оловом, алюминием) и *латунь* (сплав меди с цинком) различного состава.

Наиболее высокими механическими свойствами (прочностью и твердостью) при хорошей электропроводности обладают кадмиевая и бериллиевая бронза. Бериллиевая бронза имеет предел прочности при растяжении до 1350 МПа. Бронзу применяют при изготовлении токопроводящих пружинных контактов и пружин точных приборов.

Латунь допускает большое относительное удлинение. Поэтому из латуни изготавливают токопроводящие части выключателей, реле, экранов. Она легко штампуется.

Алюминий. Алюминий имеет удельное сопротивление большее, чем медь, поэтому для того же тока площадь поперечного сечения алюминиевого проводника должна быть на 30% больше. Но алюминий примерно в 3,5 раза легче меди, поэтому алюминиевый проводник при одинаковом сопротивлении и длине в 2 раза легче, чем медный. Максимальные значения предела прочности алюминиевого провода при растяжении почти в 2 раза ниже, чем соответствующие значения медного. Преимущество алюминия как проводникового материала еще и в том, что он дешевле меди.

Для электро- и радиотехнических целей используют алюминий марки АЕ. Содержание примесей в нем не превышает 0,5%. Для изготовления алюминиевой фольги, электродов, корпусов электролитических конденсаторов и других изделий используют алюминий марки А97,

содержащий не более 0,03% примесей. Алюминий наивысшей чистоты А999, используемый в полупроводниковом производстве, содержит не более 0,001% примесей. Удельное сопротивление особо чистого алюминия может достигать значения $\rho \approx 0{,}0263$ мкОм·м.

Алюминий может быть прокатан в фольгу толщиной несколько микрометров. Эта фольга широко применяется в качестве обкладок конденсаторов, контактных пластин и соединений в ИМС.

Алюминиевые сплавы широко используются в приборостроении и радиотехнике. Наиболее широко применяемой группой алюминиевых сплавов являются *дюралюмины*. Это сплавы системы алюминий-медь-магний. Из сплавов алюминия для изготовления проводов используют *альдрей* (алюминий с небольшими добавками магния Mg, кремния Si и железа Fe). Он обладает значительно лучшими механическими свойствами, чем чистый алюминий.

Необходимо учитывать, что контакт алюминия и меди имеет возможность гальванической коррозии. Вследствие этого алюминий в месте контакта разрушается. Поэтому контакты меди и алюминия надо тщательно защищать от влаги.

2.2.2. Проводники с высоким удельным сопротивлением

Сплавы высокого сопротивления. К ним относят металлические сплавы с удельным сопротивлением больше чем 0,3 мкОм·м.

Сплавы высокого сопротивления применяют при изготовлении электроизмерительных приборов, образцовых резисторов и нагревательных устройств. При использовании сплавов в электроизмерительной технике от них требуется не только высокое удельное сопротивление, но и возможно меньшее значение температурного коэффициента α_ρ, а также малая термо-ЭДС относительно меди.

Проводниковые материалы в электронагревательных приборах должны работать на воздухе при температуре порядка 1000º С. Среди большого количества материалов для указанных целей наиболее распространенными в практике являются сплавы на медной основе – манганин и константан, а также хромоникелевые и железоникелевые сплавы.

Манганин – основной сплав на медной основе для электроизмерительных приборов и образцовых резисторов. Манганин отличается желтоватым оттенком, хорошо вытягивается в тонкую проволоку диаметром 0,02 мм. Из манганина изготавливают также ленту толщиной 0,01…1 мм и шириной 10-300 мм.

Константан – сплав 60% Ni с 40% Cu. Содержание никеля в сплаве примерно соответствует максимуму ρ и минимуму α_ρ для сплавов Cu-Ni. Поэтому сплав имеет большое сопротивление (~ 0,5 мкОм·м) и очень малый температурный коэффициент (меньше 10^{-6} $К^{-1}$), отсюда и его название. Константан хорошо поддается обработке; его можно протягивать в

проволоку и прокатывать в ленту тех же размеров, что и из манганина. Значение α_ρ константана близко к нулю и обычно имеет отрицательный знак.

Хромоникелевые сплавы (*нихромы*).

К хромоникелевым сплавам относятся сплавы Х15Н60 (55-61% Ni, 10-15% Cr, 1,5% Mn, остальное – Fe) и Х20Н80 (75-78% Ni, 20-23% Cr, 1,5% Mn, остальное – Fe). Из этих сплавов изготавливают проволоку диаметром от 0,02 мм и ленту сечением от 0,1·1,0 мм и более.

Высокую жаростойкость нихрома можно объяснить значительной стойкостью этого сплава к прогрессирующему окислению на воздухе при высоких температурах. Это связано с тем, что температурные коэффициенты линейного расширения сплавов и их окислов примерно одинаковы, поэтому окислы не растрескиваются и не отделяются от проволоки при ее нагревании и расширении.

Сплавы для термопар. Хотя многие неметаллические материалы (в первую очередь полупроводники) имеют большие потенциальные возможности (к примеру, значительно бóльшие коэффициенты термо-ЭДС по сравнению с металлами) для успешного применения в термоэлектрической термометрии, технология их изготовления недостаточно совершенна. Поэтому подавляющее большинство термопар изготавливают из металлических компонентов. Наиболее часто применяют следующие сплавы:

1) *копель* (44% Ni и 56%Cu), ρ = 0,465 мкОм·м;
2) *алюмель* (95% Ni, Al, Si и Mn), ρ = 0,305 мкОм·м;
3) *хромель* (90%Ni и 10%Cr), ρ = 0,66 мкОм·м;
4) *платинородий* (90%Pt и 10%Rh), ρ = 0,19 мкОм·м.

2.2.3. Другие проводящие материалы

Согласно классификации, к другим проводящим материалам относят: *тугоплавкие*, *легкоплавкие*, *благородные*, *припои* и *неметаллические проводники*.

Тугоплавкие металлы. К тугоплавким проводникам относят металлы с температурой плавления больше 1700 ºС. Как правило, они являются химически устойчивыми при низких температурах, но при повышенных температурах активно взаимодействуют с атмосферой. Поэтому эксплуатация их при высоких температурах возможна только в атмосфере инертных газов или в вакууме. Механическая обработка тугоплавких металлов затруднительна вследствие их повышенной твердости и хрупкости.

Вольфрам – светло-серый материал, обладающий наиболее высокой температурой плавления и очень большой плотностью. Для него характерна слабая механическая связь кристаллитов, поэтому вольфрамовые изделия с мелкозернистой структурой очень хрупкие. В результате механической обработки ковкой и волочением вольфрам приобретает волокнистую структуру, что обеспечивает механическую прочность и гибкость тонких вольфрамовых нитей, диаметр которых может быть менее 10 мкм.

Вольфрам является одним из важнейших материалов электровакуумной техники, т.к. он имеет достаточно высокую плотность

тока эмиссии. Из него изготавливают нити накала осветительных ламп, катоды прямого и косвенного накала мощных генераторных ламп, высоковольтных выпрямителей, рентгеновских трубок. В микроэлектронике вольфрам применяют для изготовления испарителей для нанесения в вакууме тонких пленок различных материалов. Нагретый вольфрам хорошо работает в вакууме и водороде, но сильно окисляется в воздухе.

Молибден – тугоплавкий материал, близкий к вольфраму по свойствам, хотя почти в 2 раза легче последнего. Его используют в электровакуумной технике при менее высоких температурах, чем вольфрам. В паре с вольфрамом молибден используют для изготовления термопар, для измерения температур до 2000 ^{0}С в инертных средах, вакууме.

Тантал широко применяется в электровакуумной технике при менее высоких температурах, чем вольфрам, характеризуется высокой пластичностью даже при комнатной температуре. Например, прокатыванием получают танталовую фольгу для электролитических конденсаторов толщиной в несколько микрон. К преимуществам тантала относят не только возможность его использования в качестве электровакуумного конструкционного материала, выдерживающего температуры до 1200 ^{0}С, но и его способность поглощать остаточные газы в электровакуумном приборе. Тантал широко применяют в производстве пленочных резисторов, электролитических конденсаторов, а также в качестве различных нагревателей и испарителей, используемых при нанесении тонких пленок.

Титан применяют в электровакуумной технике благодаря хорошим механическим свойствам и высокой газопоглощающей способности.

Легкоплавкие металлы. К ним относят металлы с температурой плавления не более 400 ^{O}С. Это ртуть, галлий, индий, кадмий, цинк, свинец, олово.

Применение и свойства таких металлов очень разнообразны. Следует помнить, что ртуть и свинец, весьма ядовитые металлы. Ртуть легко испаряется при температурах выше + 18 ^{O}С, а её пары очень вредны. Олово при температурах ниже -40 $^\circ$С переходит в серое порошковидное вещество (оловянная чума).

Благородные металлы. К ним относят серебро, золото, платину и палладий. Эти металлы наиболее химически стойкие. В настоящее время такие металлы вырабатывают очень высокой чистоты: золото – 99,999%; серебро – 99,999%; платина – 99,9998%; палладий – 99,94%. Ценность этих металлов велика по сравнению со всеми другими металлами.

Золото – блестящий металл желтого цвета, обладающий высокой пластичностью. В электронной технике золото используется как контактный материал. Существенным преимуществом золота как контактного материала является его стойкость против образования сернистых и оксидных пленок в атмосфере, как при комнатных температурах, так и при нагревании.

Серебро – белый, блестящий металл, стойкий к окислению при нормальной температуре. От других благородных металлов отличается наименьшим удельным сопротивлением. Серебро применяется для

получения контактов в электронной аппаратуре. Высокие значения удельных теплоемкости, теплопроводности и электрической проводимости серебра обеспечивают по сравнению с другими металлами наименьший нагрев контактов и быстрый отвод теплоты от контактных точек.

Платина – белый металл, практически не соединяющийся с кислородом и очень стойкий к химическим реагентам. Платина прекрасно поддается механической обработке, вытягивается в очень тонкие нити и ленты. В отличие от серебра платина не образует сернистых пленок при взаимодействии с атмосферой, что обеспечивает платиновым контактам стабильное сопротивление. Она практически не растворяет водород, пропуская его через себя в нагретом состоянии.

Вследствие малой твердости платина редко используется для контактов в чистом виде, но служит основой для некоторых контактных сплавов.

Палладий – по ряду свойств близок к платине и часто служит ее заменителем, так как дешевле ее в 4-5 раз. Использование палладия в технике обусловлено его способностью интенсивно поглощать большие объемы водорода. Твердый палладий поглощает более чем 850-кратный объем водорода по отношению к собственному объему. Это связано с тем, что палладий – единственный элемент в периодической системе, для которого все заполненные уровни (4*s*, 4*p*, 4*d*) по количеству электронов в точности отвечают уровням соответствующих элементов нулевой группы с замкнутыми оболочками, а на других возможных уровнях (5*s*, 5*p*) электроны отсутствуют.

Неметаллические проводники. К ним относят материалы на основе углерода и оксидов металлов. В качестве первых следует указать *природный графит*, *сажу*, *пиролитический углерод* и *бороуглеродные пленки*. Их используют в качестве проводящих материалов для изготовления непроволочных линейных резисторов.

Угольные порошки используют в микрофонах, графитовые и угольные изделия применяют в качестве контактных электродов электрических машин (угольные щетки), дуговых ламп, анодов гальванических элементов. Кроме того, углерод инертное вещество, поэтому его в виде графита используют для производства тиглей, нагревателей, экранов.

Композитные проводниковые материалы. Они представляют смесь проводникового наполнителя с диэлектрической связкой. Изменением состава и характера распределения компонентов можно в довольно широких границах изменять электрические свойства таких веществ. Как проводниковую фазу используют металлы, графит, сажу, некоторые окислы и карбиды. Диэлектрической связкой служат органические и неорганические диэлектрики. Из них наибольшее внимание заслуживают *контактолы* и *керметы*.

Контактолы – это токопроводящие лаки, клеи, эмали, которые имеют вид пастоподобных композиций. Контактолы используют для изготовления ИМС по толстопленочной технологии, контактов, гибких волноводов.

Керметы – это металлодиэлектрические композиции с неорганической связкой. Керметные пленки отличаются хорошей однородностью свойств, повышенной термостойкостью. Их используют для производства резисторов, толстопленочных ИМС и резисторных микросборок.

Проводниковые материалы на основе оксидов. Некоторые оксиды металлов имеют относительно низкое сопротивление вследствие нарушения стехиометрического состава. Прежде всего, это оксид олова SnO_2 и индия In_2O_3. Тонкие слои таких пленок имеют большую прозрачность в видимой части спектра. Так, пленка SnO_2, толщиной до 2 мкм, поглощает не более 3% видимой части спектра. Поэтому такие материалы используются как прозрачные обкладки конденсаторов, преобразователей изображения и пр.

Припои и флюсы. *Припои* – это специальные сплавы, которые используют при пайке. Пайку осуществляют, если необходимо получить механически прочный герметичный шов, или электрический контакт с малым переходным сопротивлением. При пайке место соединения и припой разогревают. Припой имеет температуру плавления значительно ниже, чем соединяемый металл, то есть, припой плавится, а основной металл остается твердым. На границе соприкосновения расплавленного припоя и твердого металла происходят различные физико-химические процессы. Припой смачивает металл, растекается по его поверхности и заполняет щели между соединенными металлами. При этом компоненты припоя диффундируют в металл, а металл растворяется в припое. Вследствие этого возникает промежуточная прослойка, которая после затвердевания, соединяет детали в единое целое.

Припои подразделяют на 2 группы: твердые и мягкие. К *твердым* относятся припои с температурой плавления выше 300 ^{O}C, а остальные – к *мягким*. Твердые припои имеют большую механическую прочность, поэтому они чаще используют для достижения механической прочности соединения, а мягкие – для получения хорошего электрического контакта.

Мягкие припои изготовляют на основе сплавов “олово-свинец” (марка ПОС), а твердые – на основе медно-цинковых сплавов (марка ПМЦ), или серебряных (марка ПСр) с разными примесями. Проводимость припоев марки ПОС составляет 9…15% от проводимости меди.

При пайке необходимо защитить поверхности металлов от окисления и загрязнения. Для этого используют флюсы, которые являются вспомогательными материалами. Они должны: растворять и выводить оксиды и загрязнения с поверхности спаиваемых металлов; защищать поверхность металлов во время пайки; уменьшать поверхностное натяжение расплавленного припоя и смазывать поверхности соединяемых деталей.

Все флюсы подразделяются на: *бескислотные*, *активированные*, *активные*, *антикоррозийные*. Бескислотными флюсами является канифоль и флюсы, изготовленные на ее основе с добавлением спирта или глицерина. Активированными называются флюсы на основе канифоли с добавкой активаторов (салициловой кислоты, солянокислого диэтиламина и др.).

2.3. ПРИМЕНЕНИЕ ПРОВОДНИКОВ

Металлы и сплавы широко применяются в производстве, в частности, в радио- и электронной технике. Знание физических процессов и законов, характерных для данного типа проводников, позволяет создавать приборы широкого назначения. Рассмотрим более подробно некоторые электронные устройства и покажем, как знание физических свойств и особенностей проводниковых материалов, а также умение контролировать их свойства, помогает при изготовлении этих устройств.

2.3.1. Применение сверхпроводников

Сверхсильные магниты. Одной из важнейших *областей использования сверхпроводников* является получение сверхмощных магнитных полей. Сверхпроводниковые соленоиды дают возможность получать однородные магнитные поля напряженностью более 10^7 А/м, тогда как обычные соленоиды создают магнитные поля с максимальной напряженностью до 10^5 А/м. К примеру, сверхпроводниковый соленоид размерами не более 0,12·0,2 м может создать поле с индукцией 7 Тл (~ $5{,}6 \cdot 10^6$ А/м). Обычный магнит, создающий такое же поле, потреблял бы несколько мегаватт электрической мощности и десятки кубометров охлаждающей воды в минуту.

Какого же типа материалы нужны для получения сильных магнитных полей? Очевидно, это должны быть сверхпроводники 2-го рода, так как они сохраняют сверхпроводимость до существенных значений напряженности магнитного поля. Задача состоит в том, чтобы удержать сильное магнитное поле внутри сверхпроводника. Ситуация напоминает ту, с которой мы столкнемся в 5-й главе при изучении жестких магнитных материалов, когда цель будет состоять в предотвращении движения стенок доменов. Для этого необходимо большое количество структурных дефектов, т.е. сверхпроводник необходимо сделать настолько «неидеальным», насколько это возможно. По аналогии такие сверхпроводники получили название жестких сверхпроводников.

Магнитные поля высокой напряженности необходимы для удержания плазмы в термоядерных реакторах, в синхрофазотронах, индукционных накопителях энергии, в электрических машинах, трансформаторах.

Кроме того, такие поля используются и в космонавтике для защиты людей и приборов от повышенной космической радиации. Сильные магнитные поля искривляют траекторию заряженных частиц и «уводят» радиацию от космического корабля.

Элементы памяти. Сверхпроводники применяются при создании вычислительных машин, поэтому их можно использовать в качестве запоминающих устройств, хранящих большие и легко считываемые объемы информации.

Нарушение сверхпроводимости материала внешним магнитным полем используется в конструкции прибора, который называют *криотроном*. На

рис. 2.18 схематически изображено устройство пленочного криотрона. В условиях $T < T_c$ пленка из олова остается сверхпроводящей до тех пор, пока магнитное поле, создаваемое током, пропущенным через свинцовый сверхпроводник, не превысит критического для олова значения. На криотронных элементах можно выполнить ячейки вычислительных машин. Два состояния с нулевым (сверхпроводимость) и конечным сопротивлением (обычная проводимость) естественно отождествить с позициями 0 и 1 в двоичной системе исчисления. Достоинством ячеек памяти на пленочных криотронах является высокое быстродействие и компактность, малые потери. В таких устройствах время переключения достигает $2 \cdot 10^{-9}$ с.

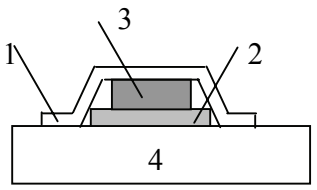


Рис. 2.18 – Схема пленочного криотрона:
1 – управляющая пленка из свинца; 2 – вентильная пленка из олова; 3 – изоляционный слой; 4 – подложка

Гораздо более эффективными являются сверхпроводящие элементы памяти, использующие эффекты Джозефсона. Джозефсоновский переход имеет два устойчивых состояния: одно с нулевым (стационарный эффект), а другое с конечным напряжением (нестационарный эффект). Увеличивая или уменьшая магнитное поле, пронизывающее переход, можно совершать переключение из одного состояния в другое. Достоинство устройств памяти на элементах Джозефсона связано с тем, что в них используется не фазовый переход из нормального состояния в сверхпроводящее, а лишь меняется характер туннелирования, что происходит значительно быстрее. Измеренное время составило 10^{-12} с.

Магнитометры. Другое важное применение джозефсоновские переходы находят в магнитометрах. В основе их работы лежит упоминавшийся ранее эффект, заключающийся в том, что максимальный сверхпроводящий ток в цепи из параллельно соединенных переходов Джозефсона зависит от магнитного потока, пронизывающего эту цепь. Согласно выражению (2.60), ток I_{max} проходит полный цикл, когда магнитный поток Ф изменяется от 0 до Φ_0. Если считать, что площадь контура равна 10^{-2} м^2, то максимальная измеримая индукция поля составит 10^{-11} Тл. Наилучший результат, полученный в лабораторных условиях, равен приблизительно 10^{-7} Тл. Технически осуществимые приборы могут иметь чувствительность приблизительно равную 10^{-6} Тл, что делает их самыми чувствительными из существующих в настоящее время магнитометров.

Магнитометры такого типа можно использовать для обнаружения подводных лодок. Вода хорошо поглощает микроволновое излучение, свет и звук, которые традиционно используют для обнаружения целей. Из-за этого обнаружить подводную лодку в погруженном состоянии практически невозможно. Однако магнитометры могут засечь малые изменения магнитного поля Земли при появлении поблизости подводной лодки.

Резонаторы. Перспективным является использование сверхпроводников для передачи электроэнергии, как на постоянном, так и переменном токе. В радиотехнике их используют как объемные резонаторы, которые имеют высокую добротность. Можно добиться стократного уменьшения потерь.

Принцип механического выталкивания сверхпроводников из магнитного поля положен в основу создания сверхскоростного железнодорожного транспорта на воздушной подушке. Этот же принцип используются для создания сверхпроводящих подвесов, которые применяются в гироскопах, двигателях и т.д.

Сверхпроводящие провода. В настоящее время промышленностью выпускается широкий ассортимент сверхпроводящих проволок и лент для различных целей. Изготовление таких проводников связано с большими технологическими трудностями. Они обусловлены плохими механическими свойствами многих сверхпроводников (особой хрупкостью), их низкой теплопроводностью и сложной структурой проводов. Поэтому вместо простых проволок и лент приходится создавать композицию из двух (обычно сверхпроводник с медью) или даже нескольких материалов. Для получения многожильных сверхпроводящих проводов из хрупких интерметаллидов особенно перспективен бронзовый метод (или метод твердофазной диффузии). По этому методу прессованием и волочением создается композиция из тонких нитей ниобия в матрице из оловянной бронзы. При нагреве олово из бронзы диффундирует в ниобий, образуя на его поверхности тонкую сверхпроводящую пленку станнида ниобия Nb_3Sn. Такой жгут может изгибаться, но пленки остаются целыми.

2.3.2. Провода

Большинство металлов обладают хорошей пластичностью, ковкостью и малым удельным сопротивлением. Пользуясь этими физическими свойствами металлов, их можно вытягивать в длинные провода разного сечения для подвода тока к различным приборам и установкам.

Все провода подразделяют на *монтажные провода*, *обмоточные провода* и *микропровода*.

Монтажные провода. *Монтажные провода* используются для соединения элементов схем, поэтому они имеют изоляционную оболочку. Она бывает полихлорвиниловой, полиэтиленовой, фторопластовой, волокнистой и резиновой. Наиболее термостойкими являются провода с фторопластовой изоляцией (до 250 ºС). Провод может быть одножильным и многожильным, причем, бóльшей гибкостью обладают многожильные провода.

В некоторых случаях монтажные провода объединяют в ленты – ленточные провода. Такие провода используют для межблочных соединений, в особенности, когда необходимо передвигать один или несколько блоков.

Обмоточные провода. *Обмоточные провода* используют для обмотки трансформаторов, катушек индуктивности, дросселей и др. При

использовании на низких частотах, такие провода имеют одну жилу, а на высоких – много жил (до 30 жил диаметром от 20 до 100 мкм), так как в многожильных проводах уменьшается поверхностное сопротивление. Обмоточные провода, как правило, имеют изоляцию из эмалей.

Микропровода. *Микропровода* имеют диаметр от 10 до 50 мкм. Они используются для обмоток микрокатушек индуктивности, для соединения "чипов" с рамкой выводов в ИМС и пр. Их изготовляют из меди, золота и алюминия.

СВЧ-проводники. При достаточно больших частотах ток проводит только поверхностный слой проводника, поэтому его сопротивление значительно выше, чем для постоянного тока. Так, при частоте 50 Гц для меди $\delta \cong 10^{-2}$ м, при 1МГц – 80 мкм, при 10 МГц – 30 мкм, 100 МГц – 7 мкм, 1000 МГц – 3 мкм. Вот почему поверхность СВЧ проводников играет такую значительную роль. Для уменьшения сопротивления СВЧ проводников их поверхность покрывают серебром, поскольку серебро имеет минимальное удельное сопротивление из всех металлов, а сам проводник делают пустотелым. Кроме того, покрытие серебром способствует защите проводника от коррозии.

2.3.3. Резисторы

Материалы для изготовления резисторов. Требования к материалам, которые используются для изготовления резисторов, следующие:

- Достаточно высокое удельное сопротивление. Это необходимо для уменьшения габаритов изделия.
- Малое значение температурного коэффициента удельного сопротивления. Это обеспечивает термостабильность резистора.
- Малое значение термо-ЭДС относительно меди. Это необходимо для уменьшения паразитных сигналов в резисторах за счет термоэффектов.

Таким требованиям отвечают тугоплавкие металлы (вольфрам, молибден, тантал, рений, хром), а также сплавы типа манганин, константан, нихромы.

Известно, что тонкие пленки из металла обладают намного бóльшими значениями удельного сопротивления ρ_d и малыми значениями температурного коэффициента α_ρ по сравнению с теми же параметрами объемного материала. Эти свойства тонких пленок нашли свое применение для изготовления тонкопленочных резисторов.

Поскольку характер зарождения и роста пленок зависит от множества случайных факторов, на практике трудно получить точное совпадение значений удельного сопротивления ρ_d для пленок одинаковой толщины. Поэтому для сравнительной оценки проводящих свойств тонких пленок пользуются параметром сопротивления квадрата $R_\square$ (или сопротивлением на безразмерный квадрат, или удельное поверхностное сопротивление), численно равного сопротивлению участка пленки, длина которого равна его

ширине при прохождении тока через две его противоположные грани параллельно поверхности подложки:

$$R_{\square} = \frac{\rho_d}{d} \ . \qquad (2.67)$$

Подбором толщины пленки d можно менять $R_{\square}$ независимо от удельного сопротивления. Вместе с тем, для определения $R_{\square}$ не требуется измерение толщины пленки.

Ввиду того, что $R_{\square}$ не зависит от величины квадрата, сопротивление тонкопленочного резистора легко рассчитать по формуле:

$$R_{\square} = R_{\square} \frac{l_0}{b_0}, \qquad (2.68)$$

где l_0 – длина резистора в направлении прохождения тока; b_0 – ширина пленки.

Для изготовления тонкопленочных резисторов обычно требуются пленки с поверхностным сопротивлением 500…1000 Ом/квадрат. Для нанесения этих пленок используют методы вакуумной технологии. Осаждение резистивной пленки происходит на ситалловом, стеклянном или другом диэлектрическом основании (подложке). В качестве исходного материала применяют металлический хром, нихром марки Х20Н80, обладающий из всех нихромов самыми низкими значениями температурного коэффициента α_ρ, танталовую фольгу.

Широко применяются для изготовления тонкопленочных резисторов резистивные сплавы, содержащие кремний. Это сплавы марок РС и МЛТ. Эти сплавы наносятся на диэлектрическую подложку методом термического вакуумного испарения.

Основные конструкции постоянных резисторов. На керамический цилиндр с внутренним отверстием, напыляют тонкую пленку резистивного

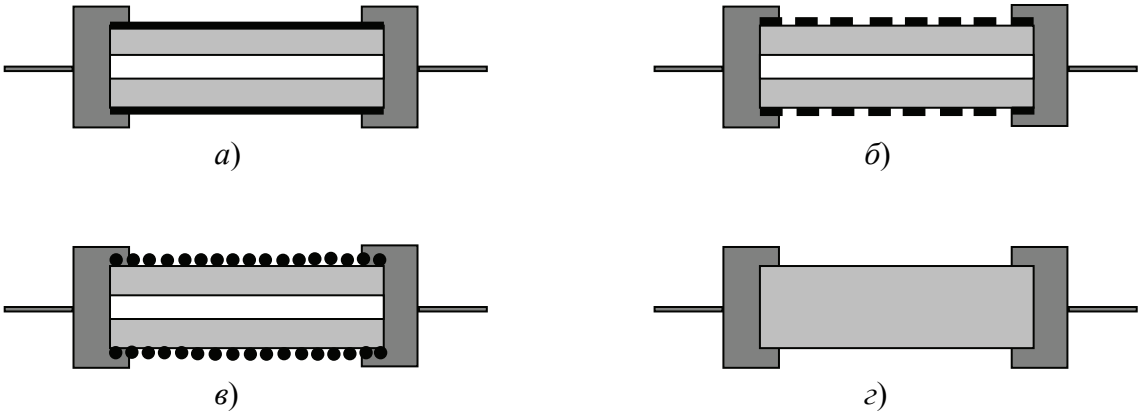


Рис. 2.19 – Основные конструкции постоянных резисторов: *а* – безындукционная; *б* – спиральная; *в* – проволочная; *г* – объемная

элемента (например, нихром, сплавы типа РС, МЛТ), а на торцах запаивают

контактные крышки. Это так называемые безындукционные резисторы (рис. 2.19, *а*).

Наиболее распространенная конструкция резисторов – спиральная. Так же как и для безындукционных резисторов, на керамический цилиндр с внутренним отверстием напыляют тонкую пленку резистивного элемента (например, нихром, сплавы типа РС, МЛТ). Затем на поверхности цилиндра делают спиральную канавку, а на торцах запаивают контактные крышки (рис. 2.19, *б*). Спиральная канавка, нарезанная на поверхности резистивного элемента, увеличивает сопротивление резистора. Это связано с тем, что сопротивление прямо пропорционально длине проводника и обратно пропорционально сечению проводника. Варьируя шаг спиральной нарезки можно получать различный диапазон сопротивлений для одной и той же заготовки. Но такая спиральная нарезка приводит к появлению индукционных явлений (аналогично индукционной катушке), вследствие чего проявляется инерционность таких сопротивлений.

Проволочные резисторы выполняют намоткой высокоомного провода на керамический цилиндр. Концы провода закрепляют чашками, или запаивают на контактах (рис. 2.19, *в*).

Композиционные объемные резисторы изготовляют из композиционных материалов. Контакты выводят из торцов резистора (рис. 2.19, *г*). Композиционные резистивные материалы представляют собой механические смеси из мелкодисперсных порошков металлов и их соединений с органической и неорганической связкой. В качестве проводящей фазы используются как проводники (порошки серебра, палладия), так и полупроводники (оксиды серебра и палладия, карбиды кремния и вольфрама). В качестве веществ, связующих мелкодисперсные порошки металлов, применяют диэлектрические материалы – термопластичные и термореактивные (полимеры, порошкообразное стекло, неорганические эмали).

2.3.4. Измерители температуры, радиаторы охлаждения

Термопары. Термоэлемент, который построен из двух различных металлических проводников *A* и *B* с замкнутой цепью, называют термопарой (рис. 2.20). Термо-ЭДС термопары зависит от температуры рабочего (измерительного) и свободного (опорного) спая и от состава материала проводников, образующих термопару. Вольтметр в этой цепи будет показывать разность потенциалов (термо-ЭДС), которая равна $U \approx \alpha_T (T_2 - T_1)$. Значение относительной удельной термо-ЭДС α_T приблизительно равняется нескольким мкВ/К и зависит от природы контактирующих материалов. Чувствительность термопар повышается, если их соединять последовательно; такие соединения называются термобатареями.

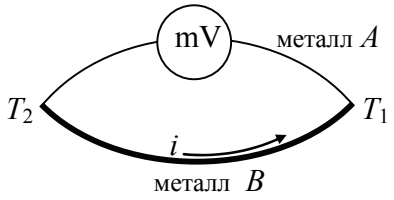


Рис. 2.20 – Схема устройства термопары ($T_2 > T_1$)

Термопары применяются как для измерения ничтожно малых разностей температур, так и для измерения очень высоких и очень низких температур (например, внутри дóменных печей или жидких газов). Если температуру холодного контакта поддерживать при 0 ºС, то вольтметр будет показывать напряжение пропорциональное температуре горячего контакта. Точность определения температуры с помощью термопар составляет, как правило, несколько кельвин, а у некоторых термопар достигает ≈ 0,01 К.

Достоинством термопар является высокая чувствительность и линейность, малая инерционность, возможность измерения температуры в широком интервале температур, независимость значения термо-ЭДС от длины проводников.

Вследствие того, что значение удельной термо-ЭДС α_T зависит от состава материала и незначительно от температуры, термопары градуируют, используя точки плавления металлов: свинца, олова, серебра и других.

Небольшие изменения состава сплава могут привести к значительным изменения термо-ЭДС. Однако это не ограничивает точности измерений, если только термопара не используется без предварительной градуировки.

Наиболее распространенными термопарами являются:

Хромель-копель (тип ХК). Позволяет измерять температуры до 600 ºС и имеет при этой температуре термо-ЭДС приблизительно 60 мВ.

Хромель-алюмель (типа ХА). Применяется к температурам порядка 1000 ºС и имеет при этой температуре термо-ЭДС приблизительно 40 мВ.

Медь-константан. Используется при низких температурах до 350 ºС. При предельной температуре термо-ЭДС достигает 15 мВ.

Платинородий-платина (тип ПП или ППР). Термопару применяют до температур 1600 ºС. Термо-ЭДС у этой термопары невелик (приблизительно 14 мВ при 1600 ºС), но она позволяет обеспечить наиболее точные и стабильные измерения температуры.

Явление термо-ЭДС имеет и отрицательные стороны. Если контактируют различные металлы, то в цепях возможно возникновение паразитной термо-ЭДС. Для устранения этого явления в цепях электроизмерительных устройств надо подбирать контактирующие металлы с малыми значениями термо-ЭДС.

Термометры сопротивления. На зависимости электрического сопротивления металлов от температуры основано действие *термометров сопротивления*. Эти устройства выполнены на основе платины или меди и позволяют по градуированной зависимости сопротивления от температуры измерять температуру с высокой точностью (до 0,003 К).

Применение же в качестве рабочего вещества термометра сопротивления полупроводников, приготовленных по специальной технологии, *термисторов* – позволяет отмечать изменение температуры в миллионные доли градуса и использовать термисторы для измерения температур в случае малых габаритов объектов измерения.

Контактные термометры. Принцип работы таких термометров основан на особенностях физических свойств ртути. Во-первых, ртуть при

комнатных температурах является жидким металлом, поэтому ее используют для наполнения капилляра термометра. Во-вторых, ртуть, как большинство металлов, при нагревании расширяется. В термометр, где рабочей средой служить ртуть, вставляется подвижная тонкая проволока. Расположение конца проволочки можно регулировать и устанавливать на требуемой температуре. Когда при повышении температуры ртуть в капилляре термометра расширяется и достигает конца проволочки, цепь замыкается и протекающий ток отключает нагрев. Как только температура падает, цепь размыкается, ток исчезает, и нагревательное устройство вновь включается. Таким образом, такие термометры используются в качестве своеобразных температурных «реле».

Радиаторы охлаждения. На участке комнатных температур удельная электропроводность падает обратно пропорционально температуре ($\sigma \sim \frac{1}{T}$), а согласно (2.9), удельная теплопроводность металлов λ не зависит от температуры ($\lambda = \sigma\beta T$ или $\lambda \sim \beta \frac{1}{T} T \sim \beta$). Это следствие из закона Видемана-Франца выполняется для большинства металлов. Такое свойство металлов важно при использовании металлов как радиаторов для охлаждения мощных полупроводниковых приборов.

Для таких радиаторов необходимо использовать металлы с большой удельной теплопроводностью. Чаще всего, это сплавы на основе алюминия (силумин), которые имеют хорошие тепловые, механические и антикоррозионные свойства. Медь для целей охлаждения используется реже вследствие её плохой коррозионной стойкости, а серебро – вследствие высокой стоимости.

ВОПРОСЫ.

1. В чем состоит физическая сущность классической теории Друде-Лоренца?
2. Каков физический смысл уровня Ферми? Как влияет температура на концентрацию свободных электронов в металле?
3. Какими свойствами обладает электронный газ в состоянии вырождения?
4. Почему в формулу для удельной проводимости металлов входит концентрация всех свободных электронов, хотя реально в электропроводности участвует лишь небольшая часть электронов, имеющих энергию, близкую к энергии Ферми?
5. Почему удельное сопротивление металлов растет с повышением температуры?
6. Что называют температурным коэффициентом удельного сопротивления? Является ли он константой для данного материала?
7. Как влияют примеси на удельное сопротивление металлов? Сформулировать правило Маттиссена.

8. Почему металлические сплавы обладают более высоким удельным сопротивлением, чем чистые компоненты, образующие сплав?

9. Почему ферромагнитные металлы обладают нелинейной зависимостью удельного сопротивления от температуры?

10. Почему при термической закалке удельное сопротивление металлов возрастает, а при термическом отжиге – уменьшается? Почему металлоидные примеси сильнее влияют на удельное сопротивление металлов, чем примеси металлических элементов?

11. Объясните поведение проводников на высоких частотах.

12. Что такое теплопроводность и теплоемкость металлов? Как объяснить низкую теплоемкость электронного газа в металлах с точки зрения квантовых представлений?

13. Как зависит удельное сопротивление тонких металлических пленок от их толщины? Объяснить, почему наблюдается такая зависимость.

14. Объяснить, почему тонкие пленки имеют отрицательный температурный коэффициент удельного сопротивления.

15. Что такое эффект Шоттки. Объяснить причину его появления.

16. Что такое внутренняя контактная разность потенциалов? При каких условиях возможно появление термо-ЭДС в замкнутой цепи?

17. Почему разность потенциалов, возникающую при контакте двух различных металлов, нельзя измерить с помощью вольтметра?

18. Что такое термоэлектрические явления? Назвать три основных термоэлектрические явления, объяснить их физическую сущность.

19. Какие металлы, и в каких условиях могут переходить в состояние сверхпроводимости? Что является причиной образования куперовских пар.

20. Как влияет магнитное поле на критическую температуру перехода в состояние сверхпроводимости? Чем отличаются друг от друга сверхпроводники 1-го и 2-го рода?

21. Что такое стационарный и нестационарный эффекты Джозефсона?

22. В каких материалах обнаружено явление высокотемпературной сверхпроводимости?

23. Почему не обладают сверхпроводимостью металлы с большой удельной проводимостью и ферромагнетики?

24. Как можно классифицировать проводниковые материалы?

25. Какие металлические сплавы высокого сопротивления нашли применение в электронной технике и почему?

26. Указать основные направления электронной техники, в которой применяются сверхпроводящие материалы.

27. Указать, какие материалы используются для создания термопар? Почему выбраны именно эти материалы?

ЗАДАЧИ.

1. Определить вероятность заполнения электронами энергетического уровня в металле, расположенного на $10kT$ выше уровня Ферми? (Ответ: $f(E) = 4{,}5 \cdot 10^{-5}$.)

2. Определить, как и во сколько раз изменится вероятность заполнения электронами в металле энергетического уровня, расположенного на 0,1 эВ выше уровня Ферми, если температуру металла повысить от 300 до 1000 К? (Ответ: Вероятность заполнения электронами увеличится в 11,4 раза.)

3. Вычислить концентрацию свободных электронов в кристалле меди, если известно, что плотность меди 8, $92\cdot10^3$ кг/м^3. При расчете полагать, что на каждый атом кристаллической решетки приходится один электрон. (Ответ: $n = 8{,}45\cdot10^{28}$ м$^{-3}$.)

4. Положение уровня Ферми для алюминия при $T \approx 0$ К соответствует энергия 11,7 эВ. Рассчитать число свободных электронов, приходящихся на один атом. Эффективную массу электрона принять равной массе свободного электрона. (Ответ: $n = 1{,}81\cdot10^{29}$ м$^{-3}$, $N_A = 6{,}02\cdot10^{28}$ м$^{-3}$, $n/N_A = 3$.)

5*. Определить энергию Ферми для лития, натрия, калия и цезия. Объяснить, почему она уменьшается с увеличением атомного номера элемента? (Ответ: E_F(Li) = 4,7 эВ; E_F(Na) = 3,16 эВ; E_F(K) = 2,05 эВ; E_F(Cs) =1,54 эВ.)

6. Как изменится интервал δE между соседними уровнями энергии свободных электронов в металле, если объем кристалла уменьшится в 10 раз? (Ответ: увеличится в 10 раз)

7. Вычислить длину свободного пробега электронов в меди при T = 300 К, если ее удельное сопротивление при этой температуре равно 0,017 мкОм/м. (Ответ: $\langle l \rangle = 3{,}89\cdot10^{-8}$ м.)

8. В медном проводнике под действием электрического поля проходит электрический ток плотностью 10^6 А/м^2. Определить скорость дрейфа и отношение ее к средней скорости (тепловой) движения электрона при температуре 300 К. (Ответ: $\upsilon_{др} = 7{,}4\cdot10^{-5}$ м/с; $\upsilon_{др}/\langle\upsilon\rangle = 6{,}1\cdot10^{-11}$).

9. Определить, во сколько раз изменится удельная теплопроводность меди λ_T при изменении температуры от 20 ºС до 300 ºС. Считать, что температурный коэффициент меди измерен при 20 ºС $\alpha_\rho = 4{,}3\cdot10^{-3}$ К$^{-1}$. (Ответ: в 1,12 раз.)

10. Катушка из медной проволоки имеет сопротивление 10,8 Ом. Масса медной проволоки 0,3 кг. Определить длину и диаметр намотанной на катушку проволоки. (Ответ: l = 146,3 м; d = 0,54 мм.)

11. Вычислить, во сколько раз сопротивление $R_\sim$ медного провода круглого сечения диаметром d = 1 мм на частоте f = 10 МГц больше сопротивления R_0 этого провода постоянному электрическому току. (Ответ: коэффициент увеличения удельного сопротивления K_R = 12).

12. Медный и алюминиевый провода равной длины имеют одинаковое сопротивление. Определить отношение диаметров этих проводов. Вычислить, во сколько раз масса алюминиевого провода меньше массы медного провода. (Ответ: $d_{Al}/d_{Cu} = 1{,}28$; $m_{Al}/m_{Cu} = 0{,}5$).

Глава III. ПОЛУПРОВОДНИКИ

3.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ПОЛУПРОВОДНИКАХ

К полупроводникам относится большое количество веществ, удельное сопротивление которых при нормальной температуре находится между значениями удельной электропроводности проводников и диэлектриков. Удельная проводимость полупроводников в отличие от металлов колеблется в широком диапазоне от 10^{-9} до 10^{5} $Ом^{-1}\cdot м^{-1}$, но такой количественный подход к классификации веществ условен и не отражает глубокого качественного различия между металлами и полупроводниками. Для уточнения этих отличий необходим глубокий анализ физических процессов в полупроводниках.

Особенностью полупроводников является их способность изменять свои свойства под влиянием различных внешних воздействий (изменение температуры, воздействие освещения, магнитного поля, внешнего давления и т.д.). В отличие от металлов полупроводники в широком интервале температур имеют отрицательный температурный коэффициент удельного сопротивления. Факт уменьшения удельного сопротивления при нагревании твердых тел впервые был установлен Фарадеем еще в 1833 г. при исследовании свойств сернистого серебра.

Свойства полупроводников сильно зависят от содержания примесей, даже в малых количествах присутствующих в кристалле. При введении примеси в полупроводник увеличивается значение проводимости, меняется тип проводимости, а также характер ее температурной зависимости. В этом также состоит качественное отличие полупроводников от металлов, в которых примеси всегда снижают проводимость.

Высокая «чувствительность» полупроводника к остаточным примесям, требование высокой степени чистоты и структурного совершенства исходных кристаллов явились одной из главных причин того, что длительное время потенциальные возможности полупроводников не использовались.

В настоящее время теория полупроводников и технология их получения достигли такого уровня, при котором стало возможным управлять свойствами материалов и приборов на их основе, заранее предсказывать их поведение в конкретных условиях эксплуатации.

Полупроводниками являются как элементы, так и химические соединения. Используемые в практике полупроводниковые материалы могут быть подразделены на: *простые полупроводники, полупроводниковые химические соединения, многофазные полупроводниковые материалы.*

К *простым полупроводникам* относятся 12 элементов: B (бор), C (алмаз), Si (кремний), Ge (германий), α-Sn (α-олово), P (фосфор), As (мышьяк), Sb (сурьма), S (сера), Se (селен), Te (теллур), I (йод). В современной технике наибольшее применение находят германий и кремний.

Полупроводниковыми химическими соединениями являются соединения элементов различных групп таблицы Д.И. Менделеева, соответствующие общей формуле A^mB^n .

К ним относятся химические соединения типа $A^{III}B^V$: арсенид и фосфид галлия, индия; антимониды индия, галлия; нитриды галлия. Из соединений $A^{II}B^{VI}$ – сульфиды цинка, кадмия, ртути; селениды и теллуриды цинка, кадмия, ртути. Из соединений $A^{IV}B^{IV}$ – прежде всего карбид кремния, а также сульфиды, селениды и теллуриды свинца.

К *твердым растворам* на основе соединения типа $A^{III}B^V$ относятся тройные твердые растворы $Ga_xAl_{1-x}As$, $Ga_xIn_{1-x}As$; $GaAs_xP_x$; InP_xAs_{1-x}.

К *многофазным полупроводниковым материалам* можно отнести материалы с полупроводящей или проводящей фазой из карбида кремния, графита и т.п., связанные керамической или другой связкой.

В зависимости от степени чистоты простые полупроводники и полупроводниковые химические соединения подразделяются на *собственные* и *примесные* полупроводники.

Собственные полупроводники – это такие полупроводники, в которых можно пренебречь влиянием примеси на их электрофизические свойства при данной температуре. В этих полупроводниках при температуре абсолютного нуля валентная зона полностью заполнена электронами, а зона проводимости свободна от электронов, то есть они ведут себя при абсолютном нуле как диэлектрики.

Примесные полупроводники – это такие полупроводники, электрофизические свойства которых в основном определяются примесями. Примеси создают локальные уровни в запрещенной зоне. При малой концентрации примесей расстояния между примесными атомами велики, поэтому их электронные оболочки не взаимодействуют между собою. Вследствие этого примесные уровни являются дискретными и не образуют зоны, как это имеет место для собственных атомов кристаллической решетки.

3.1.1. Структура и зонные диаграммы собственного полупроводника

В отсутствие внешнего электрического поля свободные электроны в полупроводниках, так же как и в металлах, совершают только хаотическое движение. Траектории движения электронов определяются их взаимодействием с тепловыми колебаниями атомов и с дефектами кристаллической решетки. При включении электрического поля $\vec{E}$ электроны дрейфуют в направлении, противоположном направлению поля.

Закон Ома в дифференциальной форме для полупроводников имеет вид:

$$j = \sigma \cdot E, \qquad (3.1)$$

где σ – проводимость полупроводника.

Рассмотрим собственный полупроводник, например, четырехвалентный германий. Этот элемент имеет на внешней электронной оболочке четыре электрона и при кристаллизации образует алмазоподобную модификацию гранецентрированной кубической решетки, в которой каждый атом,

расположенный в узле кристаллической решетки, окружен четырьмя ближайшими соседями и связан с ними ковалентными связями. При этом каждый внешний электрон принадлежит двум атомам, в результате чего внешние оболочки атомов достраиваются до восьми электронов. Для наглядности удобно рассмотреть плоскую сетку, которая фактически является проекцией кристаллической решетки на плоскость (100) (рис. 3.1). Как видно из рисунка 3.1, четыре валентных электрона каждого атома

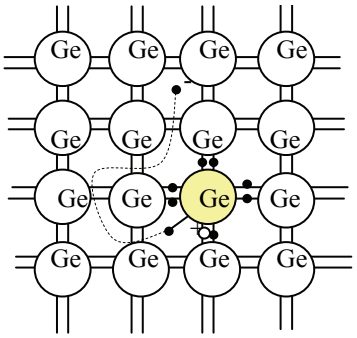
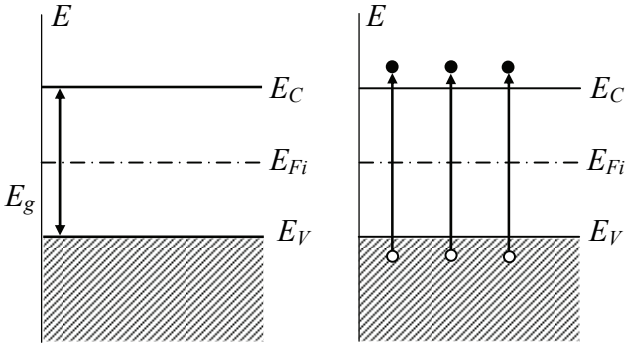


Рис. 3.1 – Плоская структура кристаллической решетки собственного полупроводника (германия)



а) $T = 0$ *б*) $T > 0$

● – электроны; ○ – дырки

Рис. 3.2 – Зонная диаграмма собственного полупроводника при температурах: *а*) $T = 0$; *б*) $T > 0$

участвуют в ковалентных связях, а энергетически находятся в валентной зоне, зона проводимости – свободна (рис. 3.2, *а*). При $T = 0$ К такой полупроводник не проводит электрический ток, так как нет свободных носителей заряда.

Выясним, как температура влияет на проводимость собственного полупроводника.

Рост температуры приводит к тому, что амплитуды колебаний атомов полупроводника увеличиваются, связь валентных электронов с атомом становится слабее. Если сообщить электрону *энергию активации* $E^{ак}$, равную энергии разрыва ковалентной связи, то связь разрушается, и электроны становятся свободными. В атоме, покинутом электроном, проявляется избыток положительного заряда (незаполненная валентная связь) или, как принято говорить, «положительная дырка» (рис. 3.1). Эта дырка ведет себя как положительный заряд, равный по величине заряду электрона. На освобожденное электроном место (дырку) может переместиться соседний электрон, а это равносильно тому, что переместилась дырка: она появится в том месте, откуда ушел электрон. Таким образом, дырку можно считать свободным носителем положительного заряда. Пока поле отсутствует, электроны и дырки передвигаются по кристаллу хаотически. Количество возникших свободных электронов равно количеству возникших свободных дырок, т.е. равновесная концентрация электронов n_i равна равновесной концентрации дырок p_i: $n_i = p_i$. Индекс i происходит от английского слова *intrinsic* – собственный.

На зонной диаграмме вышеописанный процесс можно отразить следующим образом. При некоторой температуре тепловые колебания атомов могут сообщить заметному количеству электронам валентной зоны энергию, достаточную для преодоления запрещенной зоны и попадания в зону проводимости. В валентной же зоне появятся свободные вакансии (дырки) (рис. 3.2, *б* – стрелками отмечен только процесс тепловой генерации носителей). Электронно-дырочные пары рождаются в одном акте с некоторой скоростью, зависящей от температуры. Это так называемая *тепловая генерация* носителей заряда. При этом всегда существует противоположный процесс исчезновения электронно-дырочных пар равной интенсивности, называемый *рекомбинацией*. Равенство этих процессов дает постоянную во времени концентрацию электронов и дырок.

Если приложить электрическое поле, то под его влиянием свободные электроны передвигаются в зоне проводимости против поля, а дырки в валентной зоне – по направлению поля. Таким образом, собственная проводимость осуществляется и электронами в зоне проводимости и дырками в валентной зоне:

$$\sigma = en_i\mu_n + ep_i\mu_p = en_i(\mu_n + \mu_p), \qquad (3.2)$$

где μ_n и μ_p – подвижности электронов и дырок соответственно.

Для нахождения концентрации собственных носителей (электронов и дырок) необходимо выяснить, как в собственных полупроводниках распределяются электроны в зонах.

Распределение электронов в полупроводниках, так же как и в металлах, описывается функцией распределения Ферми-Дирака:

$$f(E) = \frac{1}{e^{\frac{E-E_F}{kT}} + 1}, \qquad (3.3)$$

где E_F – энергия уровня Ферми (химический потенциал), k – постоянная Больцмана, T – абсолютная температура. Но если в металлах, для которых характерна высокая плотность свободных электронов, электронный газ становится невырожденным при очень высоких температурах (~ 10^5 К), то в случае с полупроводниками электронный газ, как правило, не вырожден уже при комнатных температурах. Это объясняется тем, что в полупроводниках число состояний в зоне проводимости порядка 10^{28} м$^{-3}$, а число свободных электронов обычно находится в пределах $10^{18}\ldots10^{24}$ м$^{-3}$, поэтому доля занятых состояний пренебрежимо мала и согласно формуле $f(E) = \frac{n(E)dE}{N(E)dE}$ вероятность заполнения электронами свободных уровней намного меньше единицы ($f << 1$). В этом случае в выражении (3.3) можно пренебречь единицей и получить следующую аппроксимацию:

$$f(E) = \exp\left(-\frac{E - E_F}{kT}\right), \qquad (3.4)$$

т.е. квантовая функция распределения Ферми-Дирака (3.3) переходит в классическую функцию распределения Максвелла-Больцмана (3.4), характерную для идеального газа. Такое состояние электронного газа называется невырожденным, а полупроводники – невырожденными полупроводниками.

Для того чтобы найти концентрацию свободных электронов в полупроводнике, необходимо умножить плотность состояний $N(E)$ в интервале от E до $E + dE$ в соответствующей зоне на вероятность их заполнения (3.4), а затем просуммировать (проинтегрировать) по всем состояниям в зоне. Расчеты показывают, что в собственном полупроводнике концентрации электронов в зоне проводимости и дырок в валентной зоне могут быть выражены таким образом:

$$n_i = N_C \cdot e^{-\frac{(E_C - E_F)}{kT}}, \tag{3.5}$$

$$p_i = N_V \cdot e^{-\frac{(E_F - E_V)}{kT}}, \tag{3.6}$$

где N_C и N_V – эффективные плотности состояний в зоне проводимости и в валентной зоне соответственно, E_C – энергия дна зоны проводимости; E_V – энергия потолка валентной зоны.

Значения эффективных плотностей состояний

$$N_C = \frac{2(2\pi m_n^* kT)^{3/2}}{h^3}, \tag{3.7}$$

$$N_V = \frac{2(2\pi m_p^* kT)^{3/2}}{h^3}, \tag{3.8}$$

где m_n^*, m_p^* – эффективные массы электронов и дырок.

Итак, невырожденный полупроводник графически можно представить в виде уровня с энергией E_C и концентрацией N_C и уровня с энергией E_V и концентрацией N_V.

Так как в собственном полупроводнике концентрации электронов n_i и дырок p_i равны, их можно найти по формуле:

$$n_i = p_i = \sqrt{n_i \cdot p_i} = \sqrt{N_C N_V} \cdot \exp\left(-\frac{E_g}{2kT}\right), \tag{3.9}$$

где $E_g = E_C - E_V$ – ширина запрещенной зоны.

Для решения многих задач необходимо знание положения уровня Ферми. Уровень Ферми находят из условия электронейтральности, которое в области собственной проводимости имеет вид:

$$n_i = p_i. \tag{3.10}$$

Для определенности будем отсчитывать уровень Ферми от дна зоны проводимости, т.е. $E_C = 0$. Тогда $E_g = E_C - E_V = -E_V$. Приравняв уравнения (3.5) и (3.6), получим: $N_C \cdot e^{\frac{E_F}{kT}} = N_V \cdot e^{-\frac{(E_F + E_g)}{kT}} \Rightarrow \frac{N_C}{N_V} = e^{\frac{-E_F - E_F - E_g}{kT}} \Rightarrow$

$$kT\ln\frac{N_C}{N_V} = -2E_F - E_g;\ E_F = -\frac{E_g}{2} - \frac{kT}{2}\ln\frac{N_C}{N_V} = -\frac{E_g}{2} + \frac{kT}{2}\ln\frac{N_V}{N_C}.$$

Так как $\frac{N_V}{N_C} = \left(\frac{m_p^*}{m_n^*}\right)^{3/2}$ (см. (3.7) и (3.8)), то

$$E_F = -\frac{E_g}{2} + \frac{kT}{2}\ln\frac{N_V}{N_C} = -\frac{E_g}{2} + \frac{3kT}{4}\ln\frac{m_p^*}{m_n^*} \quad . \qquad (3.11)$$

Из этого выражения видно, что в собственном полупроводнике при $T = 0$ К уровень Ферми находится точно в середине запрещенной зоны.

С повышением температуры положение уровня Ферми в собственном полупроводнике приближается к зоне носителей с меньшей эффективной массой (см. формулу (3.11)). Так как в большинстве полупроводников эффективная масса электронов меньше эффективной массы дырок, то с ростом температуры уровень Ферми приближается к зоне проводимости.

3.1.1.1. Основные параметры собственных полупроводников

К основным параметрам собственных полупроводников относятся:

- *ширина запрещенной зоны*;
- *эффективная масса носителей заряда*;
- *подвижность носителей заряда*;
- *концентрация носителей заряда*;
- *удельная проводимость* или *сопротивление*.

Ширина запрещенной зоны. *Шириной запрещенной зоны* называется величина энергетической щели, которая разделяет зону проводимости и валентную зону. Это энергия, которую должен получить валентный электрон, чтобы перейти из валентной зоны в зону проводимости. Принято считать, что ширина запрещенной зоны полупроводников находится в диапазоне $0 < E_g \leq 3$ эВ. Более подробно этот вопрос был рассмотрен в гл. I.

Полупроводники с шириной запрещенной зоны больше 2 эВ называют *широкозонными*, а с $E_g < 0{,}3$ эВ – *узкозонными*.

Эффективная масса носителей заряда. Эффективная масса может быть определена как коэффициент пропорциональности в законе, связывающем внешнюю силу, действующую на электрон в кристалле, с его ускорением (см. Гл. I.)

Хотя эффективные массы электронов и дырок не связаны с истинной массой электрона, их удобно выражать в долях истинной массы, т.е. вместо величин m_n^* и m_p^* (индексы *n* и *p* от английских слов *negative* и *positive* относятся к электронам и дыркам соответственно) зачастую используются отношения m_n^*/m_0 и m_p^*/m_0, где m_0 – масса электрона, равная $9{,}11\cdot10^{-31}$кг.

Подвижность носителей заряда. Подвижностью носителей заряда называется их дрейфовая скорость в электрическом поле единичной напряженности:

$$\mu = \frac{\upsilon_{др}}{E}, \tag{3.12}$$

где $\upsilon_{др}$ – дрейфовая скорость; E – напряженность электрического поля.

Подвижность носителей заряда можно найти по формуле:

$$\mu = \frac{e\tau}{m^*} = \frac{e}{m^*}\frac{\langle l \rangle}{\langle \upsilon \rangle}, \tag{3.13}$$

где $\tau = \dfrac{\langle l \rangle}{\langle \upsilon \rangle}$ – время релаксация электронов или дырок; $\langle l \rangle$ – средняя длина свободного пробега носителей заряда; $\langle \upsilon \rangle$ – средняя скорость теплового движения.

Подвижности электронов μ_n и дырок μ_p в собственном полупроводнике могут быть различными. Это связано с тем, что механизмы движения в электрическом поле свободных электронов и дырок различны. Так как в большинстве полупроводников эффективная масса электронов меньше эффективной массы дырок, то подвижность электронов выше, чем дырок. Например, в германии μ_n = 0,39 м^2/В·с, μ_p = 0,19 м^2/В·с, в кремнии – μ_n = 0,15 м^2/В·с, μ_p = 0,05 м^2/В·с.

Рассмотрим зависимость подвижности носителей заряда от температуры. В собственных полупроводниках подвижность носителей заряда в первую очередь зависит от тепловых колебаний кристаллической решетки (фононов). Это обусловлено рассеянием носителей заряда на тепловых колебаниях кристаллической решетки. С ростом температуры амплитуда тепловых колебаний узлов кристаллической решетки возрастает, поэтому длина свободного пробега электронов уменьшается, то есть уменьшается подвижность носителей заряда. Так как электронный газ в собственных полупроводниках является невырожденным, то в соответствии с классической статистикой Максвелла-Больцмана для средней тепловой скорости $\langle \upsilon \rangle$ носителей заряда имеем:

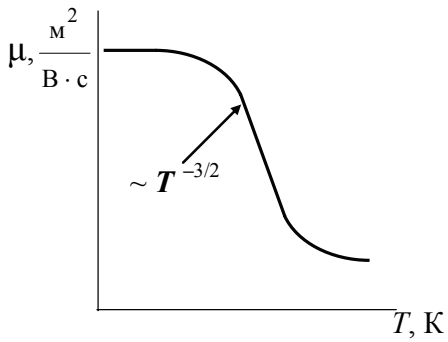


Рис 3.3 – Зависимость подвижности носителей заряда в собственном полупроводнике от температуры

$$\langle \upsilon \rangle = \sqrt{\frac{3kT}{m^*}}. \tag{3.14}$$

Так как $\langle l_T \rangle \sim \dfrac{1}{T}$ (см. ф-лу 2.11) и $\langle \upsilon \rangle \sim \sqrt{T}$ (см. ф-лу 3.14), то из (3.13) получим, что $\mu \sim \dfrac{\langle l \rangle}{\langle \upsilon \rangle} \sim \dfrac{1/T}{\sqrt{T}} = T^{-\frac{3}{2}}$. Таким образом, уменьшение подвижности подчиняется закону:

$$\mu = A \cdot T^{-\frac{3}{2}}, \qquad (3.15)$$

где A – некоторый коэффициент, не зависящий от температуры; T – абсолютная температура.

График температурной зависимости подвижности носителей заряда в собственных полупроводниках приведен на рис. 3.3.

Концентрация собственных носителей заряда. Концентрацию носителей в собственном полупроводнике можно найти по формуле (3.9)

$$n_i = p_i = \sqrt{n_i \cdot p_i} = \sqrt{N_C N_V} \cdot \exp\left(-\frac{E_g}{2kT}\right).$$

Эту зависимость удобно представить в координатах $\ln n_i = f(1/T)$. Прологарифмируем выражение (3.9) и получим:

$$\ln n_i = \ln \sqrt{N_C N_V} - \frac{E_g}{2k} \cdot \frac{1}{T}. \qquad (3.16)$$

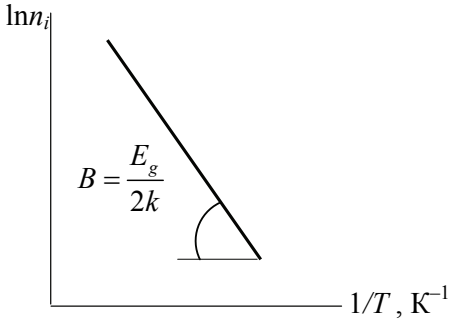


Рис 3.4 – Зависимость концентрации носителей заряда в собственном полупроводнике от температуры

Так как первое слагаемое практически от температуры зависит слабо, то на графике будем откладывать прямую линию с наклоном, равным $B = \frac{E_g}{2k}$ (рис. 3.4). Из этого графика можно найти концентрацию носителей заряда в собственных полупроводниках для данной температуры.

Удельная проводимость. В общем случае удельная проводимость собственного полупроводника определяется двумя типами носителей заряда: электронами и дырками, и ее можно найти по формуле (3.2).

Подстановка в (3.2) выражений (3.15) и (3.9) дает:

$$\sigma_i = e(\mu_n + \mu_p) n_i = e(A_n + A_p) T^{-3/2} \sqrt{N_C N_V} \exp\left(-\frac{E_g}{2kT}\right). \qquad (3.17)$$

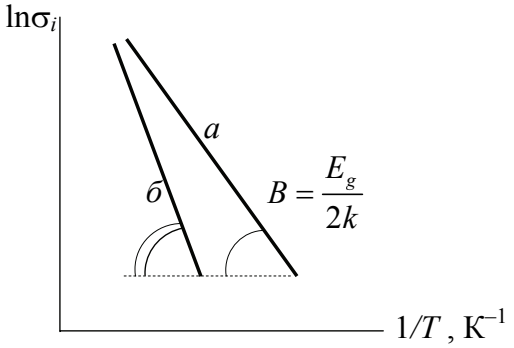


Рис 3.5 – Зависимость проводимости собственных полупроводников от температуры: *а* – для полупроводника с шириной запрещенной зоны E_{g1}; *б* – для полупроводника с шириной запрещенной зоны E_{g2}; $E_{g2} > E_{g1}$

Эту зависимость также более удобно представлять в координатах $\ln\sigma_i = f(1/T)$, где она будет прямой линией с наклоном $B = \frac{E_g}{2k}$ (рис. 3.5). Если экспериментально измерить зависимость $\ln\sigma_i = f(T)$ и построить ее в координатах $\ln\sigma_i$ (ось ординат) и $1/T$ (ось абсцисс), то, измеряя наклон прямой, можно рассчитать ширину запрещенной зоны E_g

данного полупроводника, пренебрегая более слабой зависимость μ_n и μ_p от температуры. С точки зрения простоты измерений этот метод является наиболее удобным. Из графика видно, что, чем меньше наклон линии, тем больше ширина запрещенной зоны (рис. 3.5).

Итак, удельная проводимость в собственных полупроводниках увеличивается с ростом температуры. Хотя, согласно формуле (3.15), подвижности электронов и дырок уменьшаются, рост удельной проводимости происходит за счет резкого увеличения концентрации собственных носителей заряда n_i, которая во много раз перекрывает уменьшение подвижности носителей заряда (см. формулу (3.17)).

3.1.2. Структура и зонные диаграммы примесных полупроводников

Примесные атомы в зависимости от их типа могут по-разному влиять на электрические свойства данного полупроводника. Процесс контролируемого введения в полупроводник необходимых примесей называется легированием. Из гл. I известно, что примеси могут быть в виде примесей замещения и примесей внедрения.

Примеси замещения. Роль примеси замещения рассмотрим на примере кристалла германия. Его кристаллическая решетка состоит из четырехвалентных атомов, связанных ковалентной связью, при которой валентные электроны основных атомов объединяются в электронные пары. Схематически это показано на рис. 3.1.

Если в такой решетке один из основных атомов заменить примесным атомом с бóльшей валентностью, например, пятивалентным атомом сурьмы (рис. 3.6), то четыре валентных электрона атома сурьмы будут участвовать в образовании ковалентной связи так же, как и основные атомы германия. Лишний пятый электрон может быть легко отщеплен от атома тепловыми колебаниями решетки, т. е. стать свободным. Образование свободного электрона не сопровождается нарушением ковалентной связи, как в случае с собственным полупроводником, следовательно, дырки не возникает. Проявившийся в данном случае избыточный положительный заряд обусловлен положительным ионом, который перемещаться по решетке не может.

Лишний пятый электрон не может находиться в валентной зоне кристалла из-за отсутствия в ней свободных энергетических состояний, поэтому этот электрон находится либо на примесном уровне в запрещенной зоне, либо в зоне проводимости. Если электрон находится на примесном уровне, то, чтобы перевести его в зону проводимости, примесный атом необходимо ионизировать, затратив на это энергию активации (ионизации) $E_D^{\text{акт}}$, намного меньшую E_g. Примеси, поставляющие электроны в зону проводимости, называются *донорами*, а энергетические уровни – *донорными*.

На зонной диаграмме при $T = 0$ донорный уровень заполнен электронами и проводимость полупроводника равна нулю (рис. 3.7, *а*), а при $T > 0$

электроны с донорного уровня переходят в зону проводимости (рис. 3.7, *б*), приводя к появлению свободных электронов.

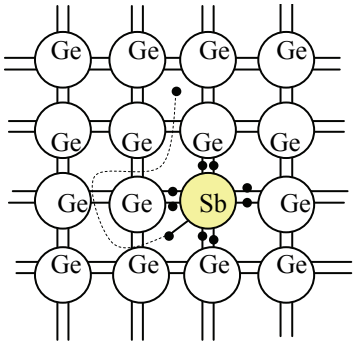
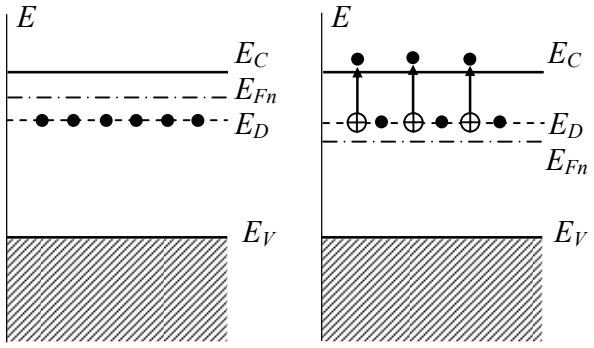


Рис. 3.6 – Плоская кристаллическая решетка примесного (электронного) полупроводника



а) $T = 0$ *б*) $T > 0$

● – электроны; ⊕ – ионизированные доноры

Рис. 3.7 – Зонная диаграмма примесного (электронного) полупроводника при: *а*) $T = 0$; *б*) $T > 0$

Таким образом, при некоторой небольшой температуре еще задолго до возникновения собственной проводимости, проводимость полупроводника электронного типа (*n*-типа) будет определяться донорными примесями.

Проводимость полупроводников *n*-типа можно найти по формуле:

$$\sigma_n = e\mu_n n_n, \tag{3.18}$$

где n_n – концентрация основных носителей заряда (электронов), определяющих тип проводимости.

Так как при комнатной температуре мелкие доноры полностью ионизированы, то в формуле (3.18) концентрацию электронов n_n можно приближенно считать равной концентрации доноров N_D ($n_n \approx N_D$). К примеру, энергия ионизации атомов сурьмы в германии $E_D^{\text{акт}}$ = 0,0127 эВ, и при комнатных температурах эти примеси будут полностью ионизированы.

Концентрация основных носителей (электронов) равна:

$$n_n = N_C \cdot e^{-\frac{(E_C - E_{F_n})}{kT}}. \tag{3.19}$$

Найдем значение энергии Ферми в запрещенной зоне электронного полупроводника. Уровень Ферми для полупроводника *n*-типа будем отсчитывать от дна зоны проводимости (E_C = 0), а n_n заменим N_D. Тогда $N_D = N_C \cdot e^{\frac{E_{F_n}}{kT}} \Rightarrow kT \ln \frac{N_D}{N_C} = E_{F_n}$, и энергия Ферми для электронного полупроводника (полупроводника *n*-типа) примет вид:

$$E_{F_n} = -kT \ln \frac{N_C}{N_D}. \tag{3.20}$$

Рассмотрим тот же кристалл, но в качестве замещающего атома возьмем атом с меньшей валентностью, например трехвалентный атом индия (рис. 3.8). В этом случае для образования устойчивой ковалентной связи атома индия с окружающими атомами основного полупроводника не будет

хватать одного электрона. Этот электрон атом индия способен захватить у соседнего атома основного вещества, где в результате этого образуется дырка. Образовавшийся в этом случае избыточный отрицательный заряд примесного атома связан с захваченным электроном, который в отличие от дырки перемещаться по решетке не может.

На зонной диаграмме (рис. 3.9 *а*, *б*) появлению дырки соответствует переход электрона из валентной зоны на примесный уровень. Для этого

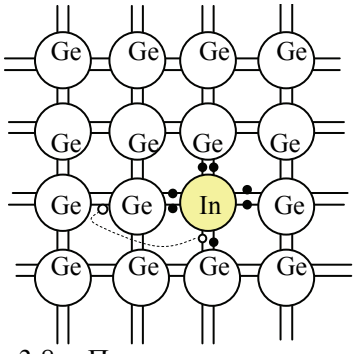
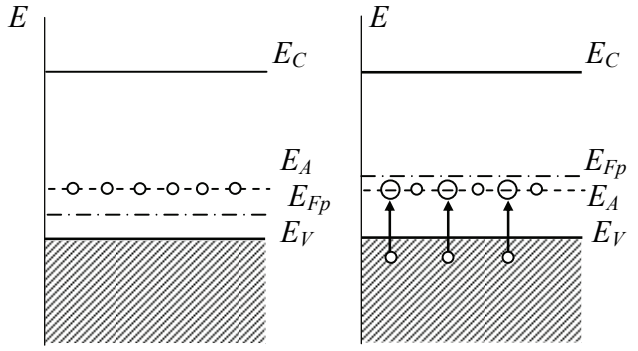


Рис. 3.8 – Плоская кристаллическая решетка примесного (дырочного) полупроводника



а) $T = 0$ *б*) $T > 0$

∘ – дырки; ⊖ – ионизированные акцепторы

Рис. 3.9 – Зонная диаграмма примесного (дырочного) полупроводника при: *а*) $T = 0$; *б*) $T > 0$

перехода необходимо затратить энергию активации (ионизации) $E_A^{\text{акт}}$, которая обычно намного меньше ширины запрещенной зоны (для индия E_A = 0,0117эВ). Примесный уровень E_A атома In располагается в запрещенной зоне вблизи потолка валентной зоны и при $T > 0$ электроны из валентной зоны могут быть захвачены примесным уровнем. Примеси, захватывающие электроны из валентной зоны, называют *акцепторами*, а их энергетические уровни – *акцепторным*. Возникшую проводимость называют *дырочной* или *p*-типа.

Проводимость полупроводников *p*-типа равна:

$$\sigma = e\mu_p p_p, \tag{3.21}$$

где p_p – концентрация основных носителей (дырок).

При комнатной температуре акцепторы почти всегда полностью ионизированы, поэтому в формуле (3.21) концентрацию дырок p_p можно заменить концентрацией полностью ионизированных доноров N_A ($p_p \approx N_A$).

Концентрацию основных носителей можно найти и по формуле:

$$p_p = N_V \cdot e^{-\frac{(E_{F_p} - E_V)}{kT}}. \tag{3.22}$$

Найдем значение энергии Ферми в запрещенной зоне дырочного полупроводника. Положение уровня Ферми в дырочном полупроводнике будем отсчитывать от потолка валентной зоны (E_V = 0), а p_p заменим N_A.

Тогда $N_A = N_V \cdot e^{-\frac{E_{F_p}}{kT}} \Rightarrow kT \ln \frac{N_A}{N_V} = -E_{F_p}$, откуда можно получить энергию Ферми в дырочном полупроводнике (полупроводнике *p*-типа)

$$E_{F_p} = kT \ln \frac{N_V}{N_A}. \tag{3.23}$$

Если уровень Ферми отсчитывать от дна зоны проводимости (E_C = 0), то формула для энергии Ферми в дырочном полупроводнике примет вид

$$E_{F_p} = -E_g + kT \ln \frac{N_V}{N_A}. \tag{2.24}$$

Помимо основных носителей, определяющих тип проводимости, в полупроводниках *n*-типа и *p*-типа существуют неосновные носители, концентрация которых равна p_n и n_p соответственно.

Концентрация неосновных носителей (дырок) в полупроводниках *n*-типа

$$p_n = N_V \cdot e^{-\frac{(E_{F_n} - E_V)}{kT}}. \tag{3.25}$$

Сравнив (3.19) и (3.25) получим, что $n_n >> p_n$, это неравенство следует из того, что $\left|E_C - E_{F_n}\right| << \left|E_{F_n} - E_V\right|$.

Концентрация неосновных носителей (электронов) в полупроводниках *p*-типа:

$$n_p = N_C \cdot e^{-\frac{(E_C - E_{F_p})}{kT}}. \tag{3.26}$$

Сравнив (3.22) и (3.26), получим, что $p_p >> n_p$, это неравенство следует из того, что $\left|E_{F_p} - E_V\right| << \left|E_C - E_{F_p}\right|$.

Итак, носители заряда, концентрация которых в данном полупроводнике больше, называются *основными*, концентрация которых значительно меньше – *неосновными*.

Найдем связь между концентрациями носителей тока в полупроводниках *n*-типа и *p*-типа и концентрацией в собственном полупроводнике. Перемножив n_n (3.19) и p_n (3.25), а также p_p (3.22) и n_p (3.26), получим соответствующие выражения:

$$\begin{aligned} n_n \cdot p_n &= N_C \cdot e^{-\frac{(E_C - E_{F_n})}{kT}} N_V \cdot e^{-\frac{(E_{F_n} - E_V)}{kT}} = N_C N_V \cdot e^{\frac{-E_C + E_{F_n} - E_{F_n} + E_V}{kT}} = \\ &= N_C N_V \cdot e^{-\frac{(E_C - E_V)}{kT}} = N_C N_V \cdot e^{-\frac{E_g}{kT}}; \end{aligned} \tag{3.27}$$

$$\begin{aligned} p_p \cdot n_p &= N_V \cdot e^{-\frac{(E_{F_p} - E_V)}{kT}} N_C \cdot e^{-\frac{(E_C - E_{F_p})}{kT}} = N_C N_V \cdot e^{\frac{-E_{F_p} + E_V - E_C + E_{F_p}}{kT}} = \\ &= N_C N_V \cdot e^{-\frac{(E_C - E_V)}{kT}} = N_C N_V \cdot e^{-\frac{E_g}{kT}}; \end{aligned} \tag{3.28}$$

Проанализировав выражения (3.27) и (3.28), придем к соотношениям, справедливым для состояния теплового равновесия:

$$n_n \cdot p_n = n_i^2; \qquad p_p \cdot n_p = n_i^2 , \tag{3.29}$$

где n_i – концентрация электронов в собственном полупроводнике, n_n – концентрация основных носителей (электронов), p_n – концентрация неосновных носителей (дырок), p_p – концентрация основных носителей (дырок), n_p – концентрация неосновных носителей (электронов). Интересно отметить, что свойства примесных полупроводников описываются через свойства собственных (n_i). Это происходит потому, что n_i является комбинационным параметром полупроводника, в который входят ширина запрещенной зоны и эффективные массы носителей.

Итак, атомы примесей замещения образуют в запрещенной зоне локальные уровни, проявляющиеся при меньшей валентности (по сравнению с валентностью основных атомов) в виде акцепторных уровней, а при большей валентности – в виде донорных уровней.

Следует обратить внимание на локальность примесных уровней. Она имеет место только до тех пор, пока примеси расположены на больших расстояниях друг от друга, т.е. пока их мало в кристалле. При увеличении их концентрации расстояние между примесными атомами уменьшается и при некоторой предельной концентрации $N_{\text{пр}}$ станет возможным перекрытие орбит валентных электронов этих атомов, произойдет их обобществление и примесные уровни сольются в примесную зону.

Оценим предельные значения концентрации примесей $N_{\text{пр}}$. Примем для простоты, что примесные атомы распределяются в решетке кристалла строго на равных расстояниях друг от друга, тогда расстояние между ними

$$a_{\text{п}} = \frac{1}{\sqrt[3]{N}}, \tag{3.30}$$

где N – концентрация атомов примеси в кристалле. При случайном хаотическом распределении N атомов примеси в кристалле среднее расстояние между ними составляет 0,7 от предыдущего

$$a_{\text{п}} = \frac{0,7}{\sqrt[3]{N}}. \tag{3.31}$$

Радиус электронной орбиты водородоподобного атома в кристалле из-за ослабления связи электрона со своим ядром за счет поляризации увеличится:

$$a = a_0 \frac{\varepsilon}{m^* / m_0}, \tag{3.32}$$

где ε – диэлектрическая проницаемость вещества; a_0 – радиус первой боровской орбиты для атома водорода ($a_0 = 0,53$ Å). Перекрытие орбит электронов соседних примесных атомов произойдет при условии $a_{\text{п}} = a$. Из этого условия следует, что образование примесной зоны начнется при предельной концентрации $N_{\text{пр}}$, определяемой по формуле:

$$N_{\text{пр}} \approx 1{,}32 \cdot 10^{30} \left(\frac{m^*}{m_0} \cdot \frac{1}{\varepsilon} \right)^3 ,(\text{м}^{-3}). \qquad (3.33)$$

К примеру, для германия ($\frac{m^*}{m_0} = 0{,}1$; ε = 16) предельная концентрация примеси порядка $N_{\text{пр}} \approx 3{,}2 \cdot 10^{24}$ м$^{-3}$.

Сложнее ведут себя те примеси замещения, валентность которых отличается от валентности атомов основного кристалла более чем на ±1.

Донорный или акцепторный характер примеси в этом случае по-прежнему определяется соотношением валентностей. Атомы с меньшей валентностью по сравнению с атомами кристалла будут вести себя как акцепторы, а с большей – как доноры. Однако детально теория таких примесей не разработана. Экспериментально установлено, что и донорные и акцепторные примеси такого типа, как правило, образуют уровни, лежащие далеко от дна зоны проводимости и соответственно от потолка валентной зоны. Эти уровни называют *глубокими*. Кроме того, такие примеси в ряде случаев образуют не один примесный уровень, а несколько, соответствующих различным зарядовым состояниям. Например, в германии атомы меди создают три уровня, соответствующих присоединению одного, двух или трех электронов, т. е. соответствуют трем типам ионов: Cu^{-1}, Cu^{-2}, Cu^{-3}.

Среди примесей замещения могут быть и такие, которые проявляют свойства *амфотерности*, т.е. они способны образовывать и донорные, и акцепторные уровни одновременно. Так ведет себя, например, примесь золота в кремнии и германии.

В случае полупроводниковых соединений положение еще более усложняется. Если имеется примесь замещения, то она будет замещать в соединении тот атом, валентность которого либо равна, либо на единицу больше или меньше валентности атома примеси. Рассмотрим соединение типа $A^{III}B^{V}$, например, арсенид галлия ($Ga^{III}As^{V}$). Если арсенид галлия легирован шестивалентным теллуром (Te^{VI}), атомы теллура будут замещать атомы мышьяка, поскольку валентность теллура на единицу больше атомов мышьяка. Это будет донорная примесь замещения. Если арсенид галлия легировать двухвалентным цинком (Zn^{II}), то последний будет замещать атомы галлия, поскольку его валентность на единицу меньше валентности галлия, поэтому возникнет дырочный полупроводник.

Примеси внедрения. Примеси внедрения также изменяют тип проводимости полупроводников. В этом случае электроны примесных атомов уже не участвуют в образовании валентных химических связей, а принадлежат только самому примесному атому. В такой ситуации донорный или акцепторный характер внедренного атома не зависит прямым образом от их валентности, а определяется значением электроотрицательности.

Если электроотрицательность примесных атомов $X_{\text{пр}}$ больше, чем атомов основного кристалла $X_{\text{осн}}$, то примесь проявляет акцепторные свойства. При обратном соотношении электроотрицательностей ($X_{\text{пр}} < X_{\text{осн}}$) атомы примеси

будут донорами. Атомы пяти- и семивалентных веществ обладают, как правило, более высокими значениями электроотрицательности, чем атомы первых трех групп периодической системы.

Атомы первой группы, внедряясь в междоузлие кристаллической решетки, становятся донорами, так как за счет слабой связи валентных электронов такие атомы легко ионизируется, вследствие чего возникает проводимость *n*-типа.

Электроотрицательные атомы (например, кислород) имеют небольшие размеры, и легко внедряясь в междоузлия решетки, захватывают электроны у атомов основного полупроводника, вследствие чего возникает проводимость *p*-типа. Если примесь может входить в решетку и в виде раствора замещения и в виде раствора внедрения, то может проявиться ее амфотерность. Например, тот же кислород в кремнии в замещенном состоянии является донором, а во внедренном состоянии – акцептором.

При производстве примесных полупроводников обычно используют примеси замещения. Это связано с тем, что технология примесей замещения лучше отработана и дает воспроизводимые результаты.

3.1.2.1. Основные параметры примесных полупроводников

Примесные полупроводники характеризуются следующими основными параметрами:

- *концентрация доноров или акцепторов;*
- *энергия ионизации примеси*;
- *концентрация свободных носителей заряда*;
- *подвижность носителей заряда*;
- *удельная проводимость*.

Концентрация доноров или акцепторов. Это количество примесных атомов в единице объема вещества. Обозначается: N_D – концентрация доноров, N_A – концентрация акцепторов.

Энергия ионизации примеси. Это энергия, которая необходима для перехода электрона с донорного примесного уровня в зону проводимости или электрона из валентной зоны на акцепторный примесный уровень. Для электронного полупроводника она отсчитывается от донорного уровня до дна зоны проводимости, а для дырочного – от потолка валентной зоны до акцепторного уровня.

Оценим значение энергии ионизации, например, для донорной примеси. Так как избыточный электрон донорного примесного атома движется по круговой орбите в кулоновском поле однократно заряженного положительного иона, донорный атом можно рассматривать как водородоподобный атом, помещенный в среду с диэлектрической проницаемостью ε.

Потенциальная энергия водородоподобного атома равна:

$$E_{\text{П}} = -\frac{m_0 Z^2 e^4}{8\varepsilon_0^2 h^2 n^2}, \qquad (3.34)$$

где m_0 – масса электрона; Z – заряд ядра; n = 1,2,3… – главное квантовое число. Для оценки энергии ионизации можно использовать формулу, аналогичную (3.34). Считая, что Z = 1 (для атома водорода) и n = 1, получим:

$$\Delta E_D = -\frac{m_n^* e^4}{8\varepsilon_0^2 \varepsilon^2 h^2}, \qquad (3.35)$$

где m_n^* – эффективная масса электрона; ε – относительная диэлектрическая проницаемость. Из сравнения этих двух формул можно прийти к следующему результату:

$$\Delta E_D = E_0 \frac{1}{\varepsilon^2} \frac{m_n^*}{m_0}, \qquad (3.36)$$

где E_0 – энергия ионизации атома водорода (E_0 = 13,52 эВ).

Для германия ε = 16, для кремния ε = 11,7; энергия ионизации примесей в них соответственно в 256 и 144 раз меньше энергии ионизации атома водорода. Поскольку $\frac{m_n^*}{m_0}$ несколько меньше единицы (для кремния ~ 0,2; для германия ~ 0,1), то можно заключить, что энергия ионизации доноров в германии менее 0,05 эВ, в кремнии – менее 0,10 эВ.

Аналогичная модель применима и для акцепторных примесей при подстановке эффективной массы дырок. Реальные значения энергии ионизации донорных и акцепторных примесей в германии и кремнии приводятся в табл. 3.1.

Таблица 3.1 – Энергия ионизации донорных и акцепторных примесей в полупроводниках, эВ

Полупроводник	Примеси						
	Доноры			Акцепторы			
	P	As	Sb	B	Al	Ga	In
Si	0,0453	0,0535	0,0425	0,0443	0,0684	0,0723	0,155
Ge	0,0127	0,014	0,0102	0,0106	0,0109	0,0111	0,0117

Из табл. 3.1 следует, что энергии ионизации различных примесей, в одном и том же кристалле близки, что позволяет сделать вывод о состоятельности водородоподобной модели.

По значению энергии ионизации примеси подразделяются на мелкие примеси и глубокие примеси. Если энергия ионизации примеси близка к половине ширины запрещенной зоны, то такую примесь называют *глубокой.* Если примесный уровень ближе к краю разрешенной зоны (меньше 0,1…0,05 эВ), то такую примесь называют *мелкой*. Судя по энергиям ионизации донорных и акцепторных примесей в кремнии и германии, их можно отнести к мелким примесям.

Температурная зависимость концентрации свободных носителей заряда. Рассмотрим зависимость концентрации свободных носителей заряда

от температуры в электронном полупроводнике. Для дырочного полупроводника ситуация будет аналогичной.

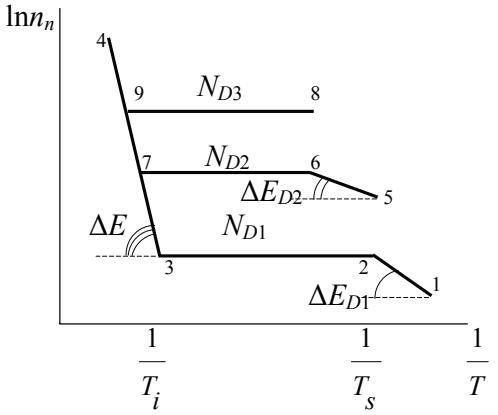


Рис. 3.11 – Зависимость концентрации свободных электронов от температуры в полупроводнике *n*-типа ($N_{D3} > N_{D2} > N_{D1}$)

Проанализируем кривую 1-2-3-4 (рис. 3.11), соответствующую относительно малой концентрации N_{D1}. В области низких температур увеличение концентрации электро-нов при нагревании полупроводника обусловлено возрастанием степени ионизации доноров (участок кривой 1-2). Каждый ионизированный донор можно рассматривать как центр, захвативший дырку. Так как концентрация доноров равна N_{D1}, то для концентрации ионизированных доноров N_{D1}^{+} можно записать

$$N_{D1}^{+} = N_{D1} \exp\left(-\frac{E_{D1} - E_F}{kT}\right), \tag{3.37}$$

где E_{D1} – положение донорного уровня на энергетической шкале.

При низкой температуре концентрация ионизированных доноров N_{D1}^{+} равна концентрации основных носителей (концентрации электронов):

$$N_{D1}^{+} = n_n = N_C \exp\left(-\frac{E_C - E_F}{kT}\right). \tag{3.38}$$

Перемножив выражения (3.37) и (3.38) получим, что

$$N_{D1}^{+} \cdot n_n = n_n^2 = N_C N_{D1} \exp\left(-\frac{E_{D1} - E_F}{kT}\right) \cdot \exp\left(-\frac{E_C - E_F}{kT}\right),$$

тогда

$$n_n = \sqrt{N_C N_{D1}} \exp\left(-\frac{E_C - E_{D1}}{2kT}\right) = \sqrt{N_C N_{D1}} \exp\left(-\frac{\Delta E_{D1}}{2kT}\right). \tag{3.39}$$

Из выражения (3.39) следует, что наклон прямой на участке 1-2 (рис. 3.11) характеризует *энергию ионизации примесей*. Этот участок называют участком «вымораживания примеси». В процессе дальнейшего нагревания при некоторой температуре, соответствующей точке 2, все доноры будут ионизированы. При этом вероятность ионизации собственных атомов полупроводника еще ничтожно мала. Поэтому в достаточно широком температурном диапазоне (участок 2-3) концентрация носителей заряда остается постоянной и практически равной концентрации доноров $n \approx N_D$. Этот участок графика принято называть *областью истощения примесей*.

При более высоких температурах (участок 3-4) доминирующую роль начинают играть переходы электронов через запрещенную зону. Этот участок графика называют *областью собственной проводимости*,

концентрацию которой можно рассчитать по формуле (3.9). Крутизна этой прямой определяется запрещенной зоной полупроводника: чем шире запрещенная зона, тем больше крутизна участка.

Для большинства примесных полупроводников температура ионизации собственных полупроводников T_i существенно превышает комнатную температуру. Для германия *n*-типа с концентрацией доноров $N_D = 10^{22}$ м$^{-3}$ температура T_i приблизительно равна 450 К. Значение T_i зависит от концентрации примеси и ширины запрещенной зоны полупроводника.

С увеличением концентрации примеси участок (5-6-7) смещаются вверх (см. формулу (3.37)). Понятно, что с увеличением концентрации примесных атомов уменьшается расстояние между ними. Это приводит к более сильному взаимодействию электронных оболочек примесных атомов и образованию примесной зоны. Поэтому можно сказать, что чем больше концентрация примесей в полупроводнике, тем меньше значение энергии ионизации этих примесей: если $N_{D3} > N_{D2} > N_{D1}$, то $\Delta E_{D1} > \Delta E_{D2} > \Delta E_{D3}$. Кроме того, чем больше концентрация примесей, тем выше температура их истощения T_S.

При достаточно большой концентрации доноров (N_{D3}) их энергия ионизации обращается в ноль, так как образовавшаяся примесная зона перекрывается зоной проводимости. Такой полупроводник является вырожденным. Температурная зависимость концентрации носителей заряда в этом случае характеризуется ломаной линией с двумя прямолинейными отрезками 8-9 и 9-4. Концентрация электронов в вырожденном полупроводнике *n*-типа постоянна во всем диапазоне примесной проводимости. Вырожденный полупроводник способен проводить электрический ток даже при очень низких температурах. Перечисленные свойства роднят вырожденные полупроводники с металлами. Поэтому их иногда называют полуметаллами.

Подвижность носителей заряда. В отличие от собственных полупроводников, в примесных полупроводниках кроме рассеяния носителей на тепловых колебаниях решетки имеет место еще один механизм рассеяния – на ионизированных атомах примеси. Роль дефектов и примесей возрастает по мере понижения температуры, когда число фононов быстро уменьшается. При этом заряженные частицы – электроны и дырки – более эффективно взаимодействуют благодаря кулоновскому взаимодействию с ионами примеси.

Процесс такого взаимодействия можно описать, используя соотношение Резерфорда, полученное им для объяснения экспериментов по рассеянию α-частиц. Из этого соотношения следует, что сечение рассеяния при прочих равных условиях, обратно пропорционально четвертой степени тепловой скорости частицы $S \sim \frac{1}{\langle \upsilon \rangle^4}$, где $\langle \upsilon \rangle = \sqrt{\frac{3kT}{m^*}}$. Известно также, что длина свободного пробега обратно пропорциональна сечению рассеяния, т. е.

$\langle l \rangle \sim \frac{1}{S} \sim \langle \upsilon \rangle^4$. Существенное влияние на рассеяние оказывает и концентрация заряженных примесей. Чем больше концентрация ионов $N_и$ (N_D или N_A), тем меньше расстояние между ними и тем ближе должны проходить электрон или дырка относительно заряженного центра. Поэтому длина свободного пробега $\langle l \rangle$ обратно пропорциональна концентрации примесей: $\langle l \rangle \sim \frac{1}{N_и}$. Оценить подвижность носителей $\mu_и$ можно по формуле (3.13) ($\mu \sim \frac{\langle l \rangle}{\langle \upsilon \rangle}$). Тогда, с учетом изложенных соображений, получим:

$$\mu_и \sim \frac{\langle l \rangle}{\langle \upsilon \rangle} \sim \frac{\langle \upsilon \rangle^4}{N_и \langle \upsilon \rangle} \sim \frac{\langle \upsilon \rangle^3}{N_и} \sim \frac{T^{\frac{3}{2}}}{N_и}, \qquad (3.40)$$

где $\mu_и$ – подвижность носителей заряда, зависящая от концентрации ионизированной примеси. Таким образом, в случае преобладания рассеяния носителей на ионизированных примесях подвижность возрастает с ростом температуры. Это связано с тем, что возрастание температуры приводит к росту тепловой скорости электронов, вследствие чего они в меньшей степени взаимодействуют с ионизированными атомами примеси.

Этот механизм доминирует при низких температурах (обычно, на участке “вымораживания примеси” и частично на участке “истощения примеси”).

Если в рассеянии носителей заряда участвуют оба механизма, то результирующая подвижность, по аналогии с (2.19), может быть найдена с помощью соотношения:

$$\frac{1}{\mu} = \frac{1}{\mu_т} + \frac{1}{\mu_и}. \qquad (3.41)$$

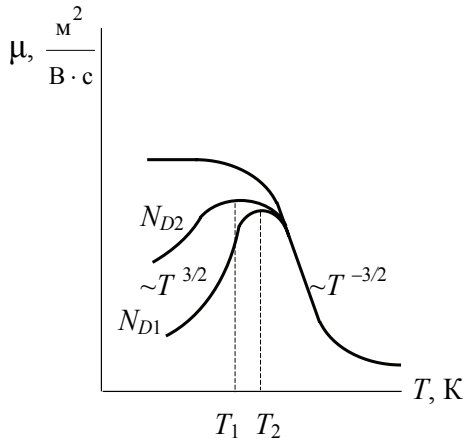


Рис 3.12 – Зависимость подвижности носителей заряда от температуры в донорном полупроводнике ($N_{D1} > N_{D2}$)

Ее температурная зависимость выражается кривой с отчетливо выраженным максимумом, как показано на рис. 3.12 для различных концентраций примесных атомов. С увеличением концентрации примеси максимум смещается в сторону более высоких температур.

Рассеяние носителей на нейтральных атомах примеси проявляется при очень низких температурах, когда атомы примеси фактически не ионизированы. Подвижность в этом случае не зависит от температуры и определяется только концентрацией примесных атомов. Еще меньшее влияние на величину подвижности имеют сдвиги, вызываемые пластической деформацией внутри кристалла.

При больших температурах, как и в случае собственных полупроводников, доминирует механизм рассеяния электронов на тепловых колебаниях кристаллической решетки, поэтому в этом случае подвижность носителей заряда с ростом температуры снижается согласно (3.15).

Температурная зависимость удельной проводимости примесных полупроводников. Рассмотрев влияние температуры на концентрацию и подвижность носителей заряда, можно представить и общий вид зависимости удельной проводимости от температуры.

Рассмотрим, к примеру, температурную зависимость проводимости полупроводника *n*-типа. График зависимости $ln\sigma$ от $1/T$ будет иметь вид ломаной линии (рис. 3.13). В области низких температур в полупроводнике практически существует только примесная проводимость (участок 1-2, рис. 3.13), увеличение которой с ростом температуры определятся тепловой ионизацией примесей (участок "вымораживания примеси"). На этом участке удельная проводимость полупроводника обусловлена только примесными электронами (участок 1-2), т. е. $\sigma_n = e\mu_n n_n$ (см. формулу (3.17)).

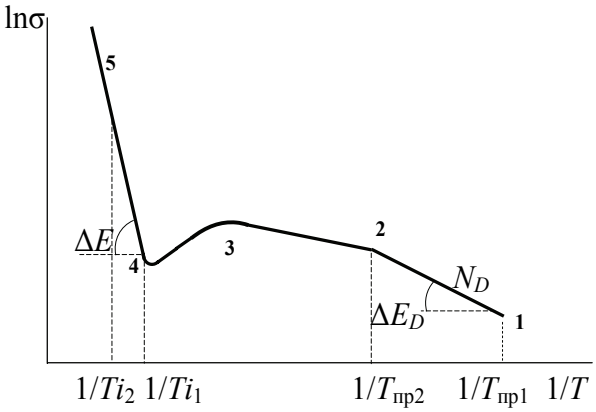


Рис. 3.13 – Зависимость удельной проводимости электронов от температуры в полупроводнике *n*-типа

Удельную проводимость на участке примесной проводимости можно представить в виде

$$\sigma_n = \sigma_{n0} \cdot \exp\left(-\frac{E_D}{2kT}\right), \qquad (3.42)$$

где σ_{n0} – удельная проводимость электронного полупроводника при температуре полной ионизации доноров; E_D – энергия донорного уровня в полупроводнике.

Участок 2-3-4 на рис. 3.13 (участок примесного насыщения или "истощения примеси") соответствует температурам, при которых все атомы примеси ионизированы, а собственная проводимость еще очень мала. Участок 2-3-4 играет очень важную роль в работе полупроводниковых приборов, так как является *областью рабочих температур*. Работа большинства полупроводниковых приборов нарушается при наступлении собственной проводимости, поэтому стараются использовать такие широкозонные полупроводники, чтобы участок 2-3-4 был как можно больше. В кремнии, например, участку примесного насыщения соответствует диапазон температур от 100 К до 500 К. Таким образом, рабочий температурный диапазон примесных полупроводников ограничен снизу температурой полной ионизации примесей, сверху – температурой, при которой в примесном полупроводнике начинает преобладать собственная

проводимость. В рабочем диапазоне можно считать все примесные атомы полностью ионизированными и пренебречь собственной концентрацией n_i, считая концентрации основных носителей равными концентрациям примесных атомов ($n_n \approx N_D,\ p_p \approx N_A$).

Итак, при низких температурах (в области рабочих температур) подвижность увеличивается с ростом температуры из-за преобладания механизма рассеяния носителей, например, электронов, на ионизированной примеси $\mu \sim \frac{T^{\frac{3}{2}}}{N_D}$. Из-за этого удельная проводимость полупроводника незначительно увеличивается (участок 2-3, рис. 3.13) и ее можно найти по формуле:

$$\sigma_n = e\mu_n n_n = eBT^{\frac{3}{2}}, \qquad (3.43)$$

где B – некоторый коэффициент, не зависящий от температуры.

При дальнейшем росте температуры (в области рабочих температур) подвижность носителей уменьшается из-за их рассеяния на тепловых колебаниях решетки $\mu \sim T^{-\frac{3}{2}}$. Поэтому удельная проводимость полупроводника незначительно уменьшается (участок 3-4, рис. 3.13) и ее можно найти по формуле:

$$\sigma_n = e\mu_n n_n = eAN_D T^{-\frac{3}{2}}, \qquad (3.44)$$

где A – некоторый коэффициент, не зависящий от температуры.

Для полупроводников с высокой концентрацией примесей, которые имеют небольшие значения энергии ионизации $\Delta E_{\text{пр}}$ (например, примеси элементов пятой и третьей групп в германии и кремнии, для которых $\Delta E_{\text{пр}}$ составляет сотые доли электрон-вольта), *температурная зависимость подвижности несущественна*, поэтому участок 2-3-4 будет представлять собой почти горизонтальную линию (полку).

Дальнейшее повышение температуры вызывает резкое увеличение проводимости, что связано с тепловой генерацией собственных электронов и дырок (участок 4-5, рис. 3.13). Чем меньше концентрация примесей, тем при более низких температурах начинает доминировать собственная проводимость. Так, чтобы наблюдать в германии собственную проводимость при комнатной температуре, необходимо, чтобы концентрация примеси в нем не превышала 10^{19} м$^{-3}$. У вырожденных полупроводников концентрация основных носителей не зависит от температуры (рис. 3.11, участок 8-9), а температурная зависимость удельной проводимости подобна температурному изменению удельной проводимости металлов.

Снимая экспериментальную зависимость $\ln\sigma = f(T)$ в широком интервале температур по наклону прямой на участке “вымораживания примеси” (рис. 3.13, участок 1-2), можно определить энергию ионизации примеси. По положению полки в области “истощения примеси” (рис. 3.13, участок 2-3-4), можно рассчитать концентрацию примеси, а по наклону

прямой в области собственной проводимости (рис. 3.13, участок 4-5) – ширину запрещенной зоны.

У реальных полупроводников температурная зависимость проводимости может отклоняться от рассмотренных выше зависимостей вследствие ряда причин. Одна из основных причин этого отклонения – наличие не одного, а нескольких видов примесных дефектов, у которых энергии ионизации могут быть различными.

Взаимная компенсация доноров и акцепторов. Случай примеси только одного типа, когда влиянием других примесей можно пренебречь, встречается сравнительно редко. Дело в том, что современная техника очистки полупроводниковых материалов, несмотря на очень высокий уровень разработки, даже для такого хорошо освоенного полупроводника, как германий, позволяет снизить концентрацию остаточных примесей до значений ~ 10^{19}…10^{18} м$^{-3}$, но не устранить их вовсе. Поэтому в реальных полупроводниках мы обычно имеем, кроме целенаправленно введенных доноров, некоторую концентрацию компенсирующих их акцепторов (или наоборот). Наличие же даже малой концентрации компенсирующей примеси может иногда существенно изменить свойства полупроводников.

Рассмотрим простой, но практически важный пример, когда в полупроводнике имеются мелкие доноры (с уровнями вблизи зоны проводимости) и мелкие акцепторы (с уровнями вблизи валентной зоны). Кроме того, предположим, что рассматриваемые интервалы температур и концентраций таковы, что уровень Ферми лежит между энергетическими уровнями доноров и акцепторов и удален от тех и других на несколько kT. Такая ситуация реализуется в германии и кремнии, содержащих доноры и акцепторы, концентрация которых меньше 10^{26} м$^{-3}$, в диапазоне температур от комнатной до температуры жидкого азота (~ 77 К). В этом случае все доноры практически лишены электронов (так как уровень Ферми лежит ниже уровня доноров) и заряжены положительно, а все акцепторы практически полностью заполнены электронами (так как уровень Ферми расположен выше их уровней) и заряжены отрицательно. Поэтому условие электронейтральности имеет вид

$$n - p = N_D - N_A. \qquad (3.45)$$

Если $N_D > N_A$, то $n > p$ и мы будем иметь полупроводник n-типа. Если температура не слишком велика, то концентрацией неосновных носителей можно пренебречь. Тогда

$$n_n = N_D - N_A \qquad (3.46)$$

Концентрация электронов в зоне проводимости оказывается такой, как если бы в полупроводнике имелись одни доноры, но с меньшей концентрацией, так как часть доноров компенсировалось акцепторами.

Если концентрация акцепторов больше концентрации доноров, то мы будем иметь полупроводник p-типа, и концентрация дырок будет

$$p_p = N_A - N_D. \qquad (3.47)$$

Если же $N_D = N_A$, то равенство (3.45) дает $n = p$, а это означает, что полностью компенсированный полупроводник по физическим свойствам похож на собственный. Так как для невырожденного полупроводника $np = n_i^2$, то

$$n = p = n_i . \qquad (3.48)$$

В этом случае носители, как и в случае собственного полупроводника, возникают только за счет возбуждений зона-зона. Тем не менее, между полностью компенсированными и собственными полупроводниками существуют отличия, например, в значениях подвижности носителей, времени жизни и др.

Полученные результаты показывают, что в суждении о степени чистоты полупроводниковых материалов по измерению концентрации носителей необходимо соблюдать осторожность, так как эта концентрация может быть мала не вследствие чистоты материала, а в результате взаимной компенсации доноров и акцепторов. Поэтому приходится прибегать к дополнительным исследованиям, например, к измерениям подвижности носителей.

3.1.3. Полупроводники в сильных электрических полях

3.1.3.1. Наклон энергетических зон в электрическом поле

Движение электрона во внешнем электрическом поле можно показать на картине энергетических зон. Двигаясь в электрическом поле, электрон меняет и свою координату x (горизонтальная ось), и энергию E (вертикальная ось), переходя с одного уровня (состояния) на другой (рис. 3.14, *а*). При этом его кинетическая энергия увеличивается на величину

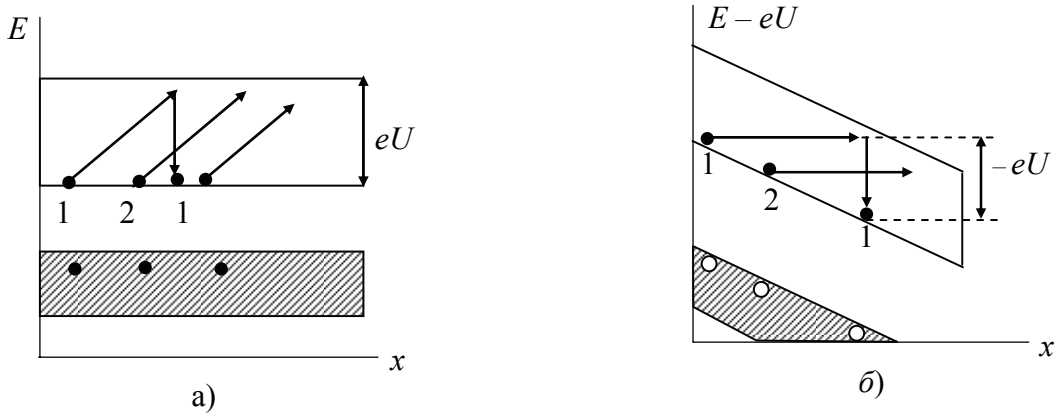


Рис. 3.14 – Движение электрона во внешнем электрическом поле: *а*) переход электрона с одного квазиуровня на более высокий квазиуровень; *б*) наклон энергетических зон с учетом внешнего поля

eU (где U – пройденная электроном разность потенциалов), а потенциальная энергия уменьшается на ту же величину, так что полная энергия не меняется. Накопленную энергию электрон может потерять при рассеянии, вернувшись в состояние с меньшей энергией (электрон 1).

Часто удобнее откладывать по вертикальной оси полную энергию с учетом внешнего электрического поля, т.е. ($E - eU$). Тогда движение электрона следует изображать горизонтальной линией, а энергетические уровни – наклонными (рис. 3.14, *б*). Тангенс угла наклона энергетических уровней при этом оказывается пропорциональным напряженности электрического поля, а энергетические уровни, соответствующие определенному значению полной энергии электрона, остаются горизонтальными.

3.1.3.2. Закон Ома в сильных электрических полях

При отсутствии внешнего электрического поля, равновесные носители заряда, имеющиеся в полупроводнике при данной температуре, движутся хаотически в разных направлениях. Если приложить внешнее электрическое поле с напряженностью E, то носители заряда начинают двигаться направленно, т. е. в полупроводнике возникает электрический ток. Плотность тока j удовлетворяет закону Ома, т. е. пропорциональна полю E ($j = ne\mu E = \sigma E$). Закон Ома требует, чтобы концентрация носителей n, подвижность μ и удельная проводимость σ не зависели от напряженности электрического поля E. Такая ситуация реализуется в области слабых полей.

В сильных же электрических полях нарушается пропорциональность между плотностью тока j и напряженностью внешнего электрического поля E. Напряженность поля, которую можно условно принять за границу между областью слабых и сильных полей, называют критической $E_{кр}$. Эта граница не является резкой и зависит от природы полупроводника, концентрации примесей и температуры окружающей среды. Для выяснения причин роста удельной проводимости σ рассмотрим влияние электрического поля отдельно на подвижность μ и на концентрацию носителей заряда n в полупроводнике.

Влияние напряженности поля на подвижность носителей заряда.

Независимость подвижности носителей заряда μ от напряженности поля определяется пренебрежимо малым изменением тепловых скоростей носителей заряда в полупроводнике от напряженности поля. Скорость носителей заряда – величина векторная и в электрическом поле может изменяться как ее абсолютное значение, так и направление: $\langle \vec{\upsilon}_0 \rangle = \langle \vec{\upsilon} \rangle + \vec{\upsilon}_{др}$. В области слабых полей, когда справедлив закон Ома, слабое электрическое поле изменяет скорость лишь по направлению. При достаточно больших напряженностях электрического поля приращение абсолютного значения скорости, получаемое на длине свободного пробега, станет сравнимо со значением тепловой скорости, т.е. $\upsilon_{др} \approx \langle \upsilon \rangle$. Таким образом, критерием слабого поля является выполнение неравенства: $\upsilon_{др} << \langle \upsilon \rangle$.

Оценим критическую напряженность поля, при которой возможны существенные отклонения от закона Ома. Свободный электрон, пройдя в полупроводнике под действием электрического поля расстояние $\langle l \rangle$,

увеличивает свою энергию на величину $e \cdot \langle l \rangle \cdot E$. Условие слабого поля можно записать в виде сопоставления этой дополнительной энергии с энергией теплового движения частиц:

$$e \cdot \langle l \rangle \cdot E << \frac{3}{2} kT \,. \tag{3.49}$$

Приравнивая обе составляющие энергии, найдем критическую напряженность поля $E_{кр}$:

$$E_{кр} = \frac{3}{2} \frac{kT}{e \cdot \langle l \rangle}. \tag{3.50}$$

Принимая, что длина свободного пробега приблизительно равна нескольким десяткам постоянной решетки $\langle l \rangle \approx 10^{-8}$ м при $T = 300$ К, найдем значение критической напряженности электрического поля: $E_{кр} \approx 10^6$ В/м.

В соответствии с выражением (3.13), резкое возрастание скорости в сильных полях должно привести к уменьшению времени свободного пробега и, следовательно, к уменьшению подвижности носителей заряда μ. Такая ситуация реализуется в атомных решетках с чисто валентными связями, как, например, в решетках алмаза, кремния, германия. В них средняя длина свободного пробега $\langle l \rangle$ носителей заряда не зависит от их скорости, а увеличение скорости сокращает время свободного пробега, а значит, уменьшается и подвижность. Опыт действительно показал, что при комнатных температурах уменьшение подвижности становится заметным для электронов в германии, начиная с $E_{кр} = 9 \cdot 10^4$ В/м, для дырок – с $E_{кр} = 1{,}4 \cdot 10^5$ В/м; в кремнии для электронов – с $E_{кр} = 2{,}5 \cdot 10^5$ В/м, для дырок – с $E_{кр} = 7{,}5 \cdot 10^5$ В/м.

Иные условия возникают в ионных решетках, в которых длина свободного пробега $\langle l \rangle$ носителей растет с увеличением их средней тепловой скорости. Как правило, это приводит к увеличению значения подвижности носителей заряда. Например, в ионных кристаллах Sb_2S_3 подвижность μ экспоненциально растет с ростом E в сильных полях порядка 10^8 В/м.

Рост подвижности в сильных полях может быть связан также с механизмом рассеяния носителей тока на ионизированных атомах примесей. Вероятность рассеяния и угол отклонения носителей убывают с увеличением скорости прохождения заряда вблизи иона. Поэтому средняя длина свободного пробега $\langle l \rangle$ в кристаллах с таким характером рассеяния, как и в ионных кристаллах, растет с увеличением скорости.

Итак, влияние сильного поля на изменение подвижности носителей заряда зависит от механизма рассеяния. В случае рассеяния на тепловых колебаниях решетки подвижность уменьшается, в случае рассеяния на ионизированных примесях – будет возрастать.

На практике не всегда удается наблюдать изменение проводимости в сильных электрических полях вследствие изменения подвижности носителей

заряда. Это объясняется тем, что в большинстве случаев возрастание напряженности поля приводит к значительному увеличению концентрации носителей.

Влияние напряженности поля на концентрацию носителей заряда.

Ударная ионизация. При возрастании напряженности поля E до некоторой величины $E_{кр}$ процесс возрастания концентрации носителей заряда будет частично компенсироваться обратным процессом рекомбинации. При напряженности электрических полей больших некоторого критического значения $E > E_{кр}$, электрон на длине свободного пробега приобретает энергию, равную ширине запрещенной зоны и при соударении с основными атомами кристаллической решетки происходит дополнительное выбивание электронов из валентной зоны в зону проводимости, т.е. образуются электронно-дырочные пары.

Появившийся в результате ионизации дополнительный электрон, ионизирует атом, и в результате в зоне проводимости оказываются два электрона, затем каждый из них набирает энергию от поля и снова ионизирует, в зоне проводимости оказываются уже четыре электрона и т.д. Этот процесс называют *ударной ионизацией полупроводника*. Ударная ионизация носит лавинообразный характер: при $E > E_{кр}$ рекомбинация уже не компенсирует генерацию электронно-дырочных пар, концентрация носителей заряда и плотность тока возрастают многократно, выделяется большое количество тепла и может произойти *пробой*, т. е. необратимое разрушение полупроводника. При меньших напряженностях лавина не может нарастать до бесконечности, так как при ударной ионизации образуются электронно-дырочные пары, и возрастает вероятность рекомбинации, препятствующей наступлению пробоя.

Уравновешивание процессов генерации и рекомбинации приводит к установлению некоторой стационарной повышенной концентрации носителей заряда, возрастающей с увеличением напряженности электрического поля.

Способность носителей заряда образовывать электронно-дырочные пары, характеризуется *коэффициентом ионизации*, который определяется как число электронно-дырочных пар, образуемых носителем заряда на единице длины. Коэффициент ионизации зависит от напряженности электрического поля и различен для разных материалов.

Значение $E_{кр}$ для различных полупроводниковых материалов достигает значений порядка $10^6 \ldots 10^8$ В/м. Такие значения напряженности поля могут возникнуть в тонких слоях полупроводника даже при небольших напряжениях. При производстве полупроводниковых приборов это накладывает ограничение на минимальную толщину полупроводниковых слоев и вынуждает также увеличивать удельное сопротивление полупроводниковых материалов.

Туннельный эффект. Сильному электрическому полю в полупроводнике соответствует большой наклон энергетических зон (рис. 3.15).

В этих условиях электроны могут проходить сквозь узкий потенциальный барьер (толщиной Δx) без изменения своей энергии – туннелировать благодаря своим квантовым свойствам. Напряженности электрического поля, при которых проявляется туннельный эффект, различны для разных материалов, так как толщина потенциального барьера Δx зависит от ширины запрещенной зоны. В типичных случаях эти напряженности поля находятся на уровне 10^8 В/м.

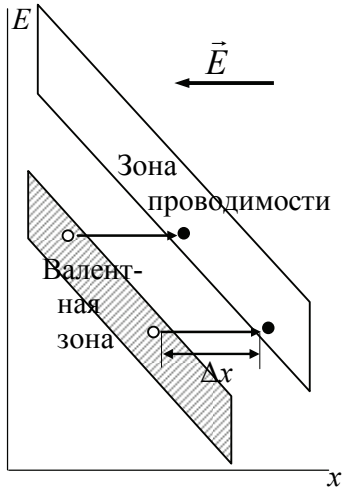


Рис 3.15 – Зонная диаграмма туннельного эффекта электрона под действием сильного поля $\vec{E}$

Необходимо отметить, что по аналогичному механизму могут туннелировать и электроны с примесных уровней, расположенных в запрещенной зоне, причем для этого требуются значительно меньшие напряженности поля.

Из всего выше сказанного ясно, что зависимость проводимости полупроводников вследствие влияния напряженности электрического поля, как на подвижность, так и на концентрацию носителей заряда в области сильных полей, может быть сложной. В зависимости от доминирования тех или иных механизмов изменения подвижности и концентрации носителей возможно возникновение в однородных полупроводниках вольтамперных характеристик *N*-типа или *S*-типа.

3.1.4. Физические явления в полупроводниках при внешних воздействиях

В данном разделе рассмотрим воздействие внешних сил (электрического и магнитного полей, градиента температуры, механического напряжения и т.д.) на полупроводники. При таких воздействиях носители заряда в полупроводнике находятся в неравновесных условиях: возникает направленное перемещение носителей – явление переноса.

Комбинируя различные внешние силы, действующие на полупроводниковые кристаллы, можно получать различные эффекты. Теоретически возможно существование сотен эффектов. Однако большинство эффектов, хотя и были рассмотрены теорией, не используются на практике. Наиболее широкое практическое распространение получили эффекты, рассмотренные ниже.

3.1.4.1. Эффект Холла

Как известно, внешнее электрическое поле, приложенное к однородному полупроводнику, вызывает электрический ток, обусловленный переносом носителей заряда в разрешенных зонах. К кристаллу могут быть приложены одновременно несколько внешних сил, например электрическое и магнитное поля, векторы напряженности которых перпендикулярны друг другу. Например, если пластинку полупроводника *n*-типа поместить во внешнее магнитное поле $\vec{B}$ и пропустить вдоль нее ток I (рис. 3.16), то, вследствие смещения движущихся зарядов к одной из граней пластинки, между боковыми гранями возникает поперечная разность потенциалов, называемая *ЭДС Холла*. Направление смещения свободных зарядов определяется направлением силы Лоренца $\vec{F}_{\text{Л}}$:

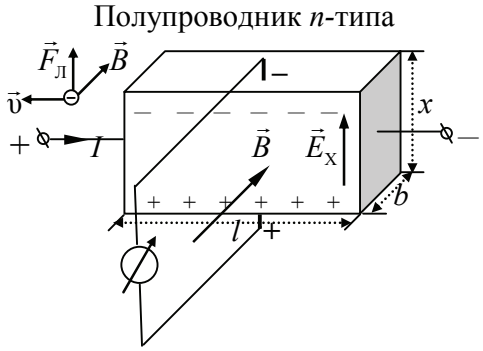


Рис. 3.16 – Определение типа проводимости полупроводника с помощью эффекта Холла

$$\vec{F}_{\text{Л}} = e\left[\vec{\upsilon}_{\text{др}} \cdot \vec{B}\right], \tag{3.51}$$

где e – заряд носителя, $\vec{\upsilon}_{\text{др}}$ – его дрейфовая скорость, $\vec{B}$ – индукция магнитного поля. Смещение носителей заряда в поперечном направлении (по отношению к $\vec{\upsilon}_{\text{др}}$ и $\vec{B}$) прекратится, когда сила Лоренца уравновесится силой действия возникшего поперечного электрического поля сместившихся зарядов (силой Кулона $\vec{F}_{\text{К}}$). Если $\vec{B} \perp \vec{\upsilon}_{\text{др}}$, то сила Лоренца равна $F_{\text{Л}} = e \cdot \upsilon_{\text{др}} \cdot B$, а сила действия на заряд поперечного электрического поля $F_{\text{К}} = eE_{\text{X}}$ (E_{X} – холловская напряженность). Приравняв эти силы, получим, что $e \cdot \upsilon_{\text{др}} \cdot B = eE_{\text{X}}$.

Для того, чтобы определить E_{X}, необходимо найти дрейфовую скорость из выражения $\upsilon_{\text{др}} = \dfrac{j}{e \cdot n}$, где j – плотность тока. Плотность тока: $j = \dfrac{I}{S} = \dfrac{I}{x \cdot b}$ (площадь поперечного сечения пластины равна $S = x \cdot b$, где x, b – поперечные размеры образца). Следовательно, $E_{\text{X}} = \dfrac{I \cdot B}{e \cdot n \cdot x \cdot b}$. Предполагая электрическое поле однородным, для холловской разности потенциалов U_{X} получаем:

$$U_{\text{X}} = E_{\text{X}} \cdot x = \frac{I \cdot B}{e \cdot n \cdot b} = \frac{1}{e \cdot n} \cdot \frac{I \cdot B}{b} = R_{\text{X}} \cdot \frac{I \cdot B}{b}, \tag{3.52}$$

где R_{X} – постоянная Холла

$$R_{\text{X}} = \frac{1}{e \cdot n}. \tag{3.53}$$

Знак постоянной Холла зависит от знака носителей тока.

Рассмотренный вывод для ЭДС Холла не является строгим. Он предполагает, что все электроны проводимости имеют одну и ту же скорость. Такое приближение справедливо лишь для металлов и сильно вырожденных полупроводников. В общем же случае следует учитывать распределение носителей тока по скоростям. Более строгий расчет приводит к следующему выражению для постоянной Холла: $R_X = \frac{A}{e \cdot n}$, где A – постоянная, зависящая от механизма рассеяния носителей. Для полупроводников, обладающих решеткой типа алмаза (германий, кремний, серое олово, InSb, AlAs и др.), имеющих носители тока одного знака и при условии, что основную роль играет рассеяние носителей на тепловых колебаниях решетки

$$A = \frac{3\pi}{8} = 1{,}17 . \qquad (3.54)$$

Для полупроводников, имеющих два типа носителей тока (электроны и дырки), расчеты, подобные изложенным выше, приводят к следующему выражению для постоянной Холла:

$$R_X = \frac{A}{e} \cdot \frac{p\mu_p^2 - n\mu_n^2}{\left(p\mu_p + n\mu_n\right)^2} . \qquad (3.55)$$

Для собственных полупроводников, в которых $n = p = n_i$, формула (3.55) приобретает следующий вид:

$$R_X = \frac{A}{e \cdot n_i} \cdot \frac{\mu_p - \mu_n}{\left(\mu_p + \mu_n\right)} . \qquad (3.56)$$

3.1.4.2. Термоэлектрические явления

Если на кристалл подействовать какой-либо другой внешней силой, например, создать на гранях разность температур ΔT, т.е. неравномерный нагрев однородного полупроводника, то в направлении от нагретой области в более холодную возникает диффузионный поток свободных зарядов. Это происходит из-за того, что неравномерный нагрев приводит к градиенту средней энергии носителей заряда в различных частях кристалла. Кроме того, перепад температур может влиять на изменение концентрации носителей заряда, чего не наблюдается в металлах. В разомкнутой цепи в стационарном состоянии плотность тока во всех точках равна нулю. Это означает, что образовавшееся электрическое поле E вызывает в каждой точке полупроводника ток, компенсирующий поток носителей заряда, т.е. образующееся внутреннее электрическое поле, препятствует дальнейшему разделению зарядов и способствует установлению равновесия.

На гранях кристалла возникает разность потенциалов, известная под названием термо-ЭДС $\mathcal{E}_T$

$$\mathcal{E}_T = \alpha_T \Delta T . \qquad (3.57)$$

Коэффициент пропорциональности α_T называют *коэффициентом дифференциальной термо-ЭДС*, т. е. ЭДС, отнесенной к единичной разности температур. Коэффициент дифференциальной термо-ЭДС α_T измеряется в В/К.

Эффект возникновения термо-ЭДС в полупроводниках при приложении к нему разности температур ΔT называется *термоэлектрическим эффектом*. (Следует отметить, что в полупроводниках, также как и в металлах, наблюдаются эффекты Зеебека, Пельтье и Томсона, подробно рассмотренные в параграфе 2.1.6).

В полупроводниках в общем случае в создании термо-ЭДС принимают участие носители заряда двух типов – электроны и дырки. Обусловленные ими составляющие термо-ЭДС противоположны по знаку. При отсутствии вырождения электронного газа значение коэффициента дифференциальной термо-ЭДС в собственных полупроводниках определяется выражением:

$$\alpha_T = \frac{k}{\sigma}\left\{ n\mu_n \left[2 + \ln \frac{2(2\pi m_n^* kT)^{3/2}}{nh^3} \right] + p\mu_p \left[2 + \ln \frac{2(2\pi m_p^* kT)^{3/2}}{ph^3} \right] \right\}, \qquad (3.58)$$

где первое слагаемое характеризует вклад, вносимый электронами, а второе – дырками. Для примесных полупроводников формула упрощается, так как одним из слагаемых можно пренебречь.

У полупроводников термоэлектрические свойства выражены значительно сильнее, чем у металлов. Это связано с тем, что в отличие от металлов, в которых электроны находятся в вырожденном состоянии и их концентрации и энергия Ферми слабо зависят от температуры, в полупроводниках те же параметры сильно меняются с изменением температуры. Поэтому, в полупроводниках коэффициент дифференциальной термо-ЭДС α_T в десятки раз больше, чем в металлах.

3.1.4.3. Тензочувствительность

Если подействовать на кристалл внешней механической силой, то можно наблюдать изменение сопротивления кристалла.

Как известно, при тепловом расширении, изменение межузельного расстояния решетки вызывает изменение ширины запрещенной зоны. Это имеет место и в случае, когда растяжение или сжатие происходят под действием приложенных механических напряжений. В собственных полупроводниках такие изменения невелики и могут привести лишь к малым изменениям сопротивления. Ситуация резко меняется, когда механическое напряжение приложено к примесному полупроводнику с ионизированными при данной температуре примесями. В этом случае даже очень малое изменение положения энергетических зон будет приводить к большому относительному изменению энергии активации примесей, поэтому сопротивление полупроводникового материала, испытывающего деформацию, будет существенно изменяться.

В общем случае сопротивление легированного полупроводникового кристалла равно

$$R = K\frac{\rho l}{S}, \tag{3.59}$$

где K – постоянная, зависящая от геометрических размеров элемента; l – длина образца; S – площадь поперечного сечения образца; ρ – удельное сопротивление полупроводника.

Для того, чтобы получить относительное изменение сопротивления полупроводникового образца при механическом напряжении, прологарифмируем, а затем продифференцируем выражение (3.59). Таким образом, получим, что

$$\frac{dR}{R} = \frac{dl}{l} - \frac{dS}{S} + \frac{d\rho}{\rho}.$$

Последнюю формулу можно переписать в виде

$$\frac{dR/R}{dl/l} = 1 - \frac{dS/S}{dl/l} + \frac{d\rho/\rho}{dl/l}. \tag{3.60}$$

Последний член $\frac{d\rho/\rho}{dl/l}$ называют *тензочувствительностью* и обозначают буквой G. Тензочувствительностью называется относительное изменение удельного сопротивления на единицу относительной деформации. Эта величина безразмерна и определяет изменение удельного сопротивления при деформации.

3.1.5. Оптические свойства полупроводников

3.1.5.1. Фотопроводимость

Поглощение света приводит к появлению в полупроводнике дополнительных, кроме имеющихся при данной температуре, *неравновесных носителей заряда*. Следовательно, проводимость полупроводника возрастет:

$$\sigma = \sigma_{темн} + \sigma_{ф}, \tag{3.61}$$

где $\sigma_{темн}$ – проводимость полупроводника в темноте; $\sigma_{ф}$ – фотопроводимость, появляющаяся при освещении полупроводника. *Фотопроводимость* – это явление увеличения проводимости кристаллов под действием света за счет внутреннего фотоэффекта. Внутренний фотоэффект (фоторезистивный эффект) – это процесс фотогенерации носителей тока в полупроводниках.

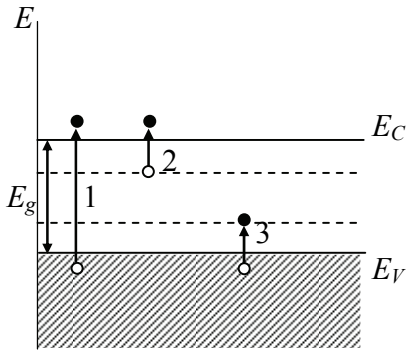


Рис. 3.17 – Зонная диаграмма полупроводника с переходами зона-зона (1) и уровень-зона (2, 3)

При внутреннем фотоэффекте первичным процессом является поглощение фотона с энергией, достаточной для перехода электрона из валентной зоны и с донорного уровня в зону проводимости, а также из валентной зоны на акцепторный уровень. Возможные типы переходов, приводящие к увеличению проводимости, приведены на рис. 3.17. Переход 1 приводит

к образованию пары электрон-дырка (собственное поглощение), а в результате других переходов образуются свободные носители заряда только одного знака (для перехода 2 – только электроны, для перехода 3 – только дырки). Переходы 2 и 3 вызывают появление примесной фотопроводимости.

Если оптический переход электронов происходит из валентной зоны в зону проводимости, то это называется собственной фотопроводимостью. При этом, энергия фотона $h\nu$ должна удовлетворять условию $h\nu \geq E_g$. Следовательно, существует некоторая граничная энергия фотона $h\nu_г$, при которой фотопроводность еще наблюдается. Она определяется соотношением

$$E = h\nu_г = h\frac{c}{\lambda_г}, \qquad (3.62)$$

где $h = 4{,}14\cdot10^{-15}$ эВ·с – постоянная Планка; $c = 3\cdot10^8$ м/с – скорость света; $\nu_г$ и $\lambda_г$ – частота и длина волны падающего света, называемая *красной границей фотопроводимости*. Если в выражение (3.61) подставить значение ширины запрещенной зоны, например кремния (E_g = 1,12 эВ), то получим граничную длину волны $\lambda_г = \frac{hc}{E_g} = \frac{4{,}14\cdot10^{-15}3\cdot10^8}{1{,}12} = 1{,}1$ мкм. Таким образом, кремний при межзонных переходах будет поглощать свет только с длиной волны, меньшей 1,1 мкм, и фотопроводимость возможна только при этих условиях.

Так как ширина запрещенной зоны для различных полупроводников изменяется от 0,1 до 3 эВ, то и пороговая длина волны поглощаемого света будет лежать в различных частях спектра: инфракрасной, видимой или ультрафиолетовой.

Изучение внутреннего фотоэффекта показало, что концентрация фотоносителей при заданной интенсивности светового потока зависит от его спектрального состава, так как у каждого полупроводника существует своя спектральная характеристика фоточувствительности. При увеличении интенсивности облучения количество свободных носителей заряда в полупроводнике возрастает. Одновременно увеличивается интенсивность обратного процесса – рекомбинации, и между процессами генерации и рекомбинации носителей заряда устанавливается динамическое равновесие.

Фотопроводимость у полупроводниковых материалов обычно наблюдается в относительно узком диапазоне длин волн, так как спектральная характеристика фотопроводимости, кроме резкого длинноволнового спада, имеет еще и более плавный коротковолновый спад (например, для GaAs при $\lambda < 0{,}9$ мкм). Это объясняется увеличением коэффициента поглощения света α и, как следствие, уменьшением *эффективной длины поглощения* $L_{эф} = \frac{1}{\alpha}$. Поглощение происходит во все более тонком поверхностном слое полупроводника, где и генерируется основное количество носителей тока. При этом возрастает роль быстрой

поверхностной рекомбинации, приводящей к уменьшению концентрации фотоносителей.

При длинах волн, бóльших граничной длины волны $\lambda_г$, энергии фотонов оказывается недостаточно для генерации электронно-дырочных пар, и тогда полупроводники можно считать прозрачными в этой области спектра (эффект оптического окна).

Процесс воздействия света на полупроводник имеет обратимый характер. После прекращения облучения проводимость полупроводника принимает прежнее значение $\sigma_{темн}$. Время, необходимое для рекомбинации неравновесных носителей заряда, для различных полупроводников составляет от наносекунд до нескольких часов. Этот процесс называют *релаксацией фотопроводимости*. Для его характеристики вводят понятие – *эффективное время жизни* неравновесных носителей заряда τ. *Эффективное время жизни* неравновесных носителей заряда – это время, за которое концентрация носителей уменьшается в *e* раз в данной точке полупроводника после прекращения внешнего воздействия. Типичные значения τ и $L_{эф}$ для кремния составляет соответственно 0,2 мкс и 0,1 мм. При получении полупроводников, предназначенных для высокочастотных приборов, значения τ и L необходимо уменьшить. Для этого полупроводник легируют примесями, создающими в запрещенной зоне глубокие энергетические уровни. Так, примесь золота в германии создает центры рекомбинации, позволяющие довести значение τ до 1 нс.

3.1.5.2. Виды поглощения света

При внутреннем фотоэффекте первичным процессом является процесс поглощения фотонов полупроводниковыми материалами. Рассмотрим подробнее виды поглощения света.

Обозначим через I интенсивность света, т.е. количество энергии, проходящей в единицу времени через нормальное к световому потоку единичное сечение полупроводника. Часть светового потока отражается от границы раздела. Доля отраженной энергии характеризуется *коэффициентом отражения* $R = I_R/I_0$. При нормальном падении световых лучей для слабо поглощающих сред коэффициент отражения может быть рассчитан по формуле:

$$R = \frac{(n-1)^2}{(n+1)^2}, \qquad (3.63)$$

где n – показатель преломления полупроводника.

Для большинства полупроводников n = 3…4. Этим значениям соответствует коэффициент отражения R = 0,25…0,36.

Свет, проходящий через полупроводник, ослабляется вследствие поглощения. Выделим на глубине x от поверхности полупроводника бесконечно тонкий слой dx. Количество световой энергии dI, поглощенное

слоем *dx*, пропорционально интенсивности света *I*, падающего на этот слой, и его толщине *dx*

$$dI = -\alpha I dx\,, \qquad (3.64)$$

где α – *коэффициент поглощения*; знак минус показывает убыль энергии при поглощении. Коэффициент поглощения α характеризует относительное изменение интенсивности излучения на единице длины. Интегрируя (3.64) от 0 до I_0 и учитывая коэффициент отражения *R*, получим зависимость интенсивности света от глубины его проникновения в полупроводник:

$$I(x) = I_0(1-R)\exp(-\alpha x)\,, \qquad (3.65)$$

где I_0 – интенсивность света, падающего на полупроводник. Формула (3.65) – это закон Бугера-Ламберта. Из него следует, что величина, обратная коэффициенту поглощения α, численно равна толщине слоя, на которой интенсивность проходящего света уменьшается в *e* раз. Коэффициент поглощения имеет размерность, обратную длине, т. е. $м^{-1}$.

Зависимость коэффициента поглощения α от длины волны λ или энергии фотонов *h*ν называют *спектром поглощения* вещества.

В полупроводниках различают несколько механизмов оптического поглощения. Каждому из них соответствует определенная область спектра.

Собственное поглощение. При освещении полупроводника светом в кристалле происходит взаимодействие квантов света с обеими подсистемами, составляющими кристалл, – атомной и электронной.

Взаимодействие света с электронной подсистемой весьма разнообразно. Прежде всего – это взаимодействие фотонов с электронами валентной зоны, т.е. с собственными электронами атомов, составляющих кристаллическую решетку. Фотоны определенной энергии способны отдать свою энергию этим электронам, оторвать их от атомов и перевести из валентной зоны на энергетические уровни в зоне проводимости. В этом случае фотоны сильно поглощаются в кристалле. Такой вид поглощения называют *собственным* или *фундаментальным*. Собственное поглощение возможно в том случае, если энергия фотонов превышает ширину запрещенной зоны. В зависимости от ширины запрещенной зоны оно проявляется в видимой или ближней инфракрасной области спектра. При больших энергиях фотонов весь свет поглощается в поверхностном слое полупроводника толщиной от одного до нескольких микрометров.

Из квантовой теории следует, что при оптических переходах электронов из одной энергетической зоны в другую существуют определенные правила $\vec{k}$-отбора (где $\vec{k}$ – волновой вектор). Разрешенными оптическими переходами являются лишь те, для которых волновой вектор (квазиимпульс) электрона остается неизменным. Иными словами, электрон и оставляемая им дырка должны иметь одинаковые, но противоположные по направлению квазиимпульсы. Такие переходы получили название *прямых переходов*.

Возможны переходы и не разрешенные правилами отбора. Однако вероятность их существенно меньше. При таких переходах закон сохранения

импульса выполняется благодаря тому, что в каждом акте поглощения принимают участие не две, а три «частицы»: фотон, электрон и фонон, т.е. квант тепловых колебаний. Последний и компенсирует разность значений импульса электрона в начальном и конечном состояниях. Такие переходы с участием фононов получили название *непрямых*. В этом случае избыточный импульс передается кристаллической решетке. Поскольку для осуществления непрямых переходов необходимо взаимодействие не двух, а трех «частиц» с согласующимися параметрами, то их вероятность меньше, чем вероятность прямых переходов. Соответственно меньше и показатель поглощения. На рис. 3.18 показаны прямые и непрямые оптические переходы.

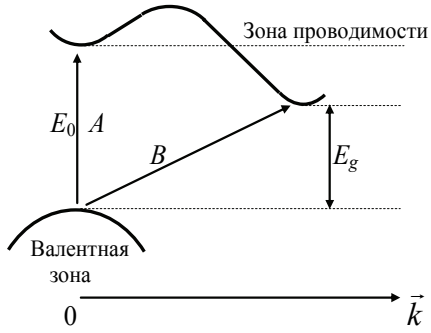


Рис. 3.18 – Прямые (*A*) и непрямые (*B*) оптические переходы

Так как с ростом температуры ширина запрещенной зоны у полупроводников уменьшается, то наблюдается смещение края собственного поглощения в сторону более низких энергий.

Экситонное поглощение. В некоторых полупроводниках при поглощении фотонов образуются особые возбужденные состояния, называемые экситонами. *Экситон* – это система связанных кулоновским взаимодействием электрона и оставленной им дырки. Он напоминает атом водорода, в котором роль ядра играет положительная дырка. Энергетические уровни возбужденного электрона, входящего в состав экситона и находящегося в центральном электростатическом поле дырки, лежат несколько ниже края зоны проводимости. В полупроводниках из-за большой диэлектрической проницаемости кулоновское притяжение мало, поэтому энергия связи в экситоне составляет лишь около 4 мэВ, а экситонные орбиты охватывают несколько элементарных ячеек кристалла (радиус орбиты порядка 15 нм).

Экситон может блуждать по кристаллу, передаваясь от одного атома (приходящего в нормальное состояние) к другому. Так как экситон представляет собой нейтральное образование, слабые электрические поля, приложенные к кристаллу не способны разрушить связь между электроном и дыркой в экситоне. Экситон при столкновениях с примесными центрами может либо «разорваться» и образовать два носителя заряда, либо рекомбинировать. Первое требует сообщения экситону энергии, необходимой для перевода электрона с экситонного уровня в зону проводимости; второй сопровождается либо излучением кванта энергии, либо чаще всего отдачей энергии экситона решетке полупроводника в виде фонона.

Поглощение на свободных носителях заряда. Фотоны могут взаимодействовать и со свободными носителями заряда в разрешенных зонах, т.е. с электронами в зоне проводимости и с дырками в валентной зоне. При этом энергия фотонов расходуется на перевод носителей заряда на более высокие энергетические уровни. Такой вид поглощения называют *поглощением на свободных носителях заряда*. Этот механизм поглощения обусловлен переходами электронов и дырок с одного уровня на другой внутри разрешенных энергетических зон (соответственно, зоны проводимости и валентной зоны). Электроны при столкновении с узлами решетки отдают накопленную кинетическую энергию. В результате энергия световой волны превращается в тепловую энергию решетки.

Интенсивность поглощения растет с увеличением длины волны падающего света, теоретическое рассмотрение показывает, что $\alpha \sim \lambda^2$. При фиксированной длине волны λ показатель поглощения тем больше, чем выше концентрация свободных носителей заряда n в полупроводнике, т.е. чем меньше удельное сопротивление материала ρ ($\alpha \sim n \sim 1/\rho$).

Примесное поглощение света. Этот важный вид поглощения света происходит при взаимодействии фотонов с примесными центрами (*поглощение примесными центрами*). В этом случае, во-первых, свет может отдавать энергию электронам этих атомов, отрывать их и переводить в зону проводимости (или переводить электрон из валентной зоны на примесный уровень), а во-вторых, свет может не ионизировать, а лишь возбуждать примесные атомы. Так как энергия ионизации примесей обычно намного меньше ширины запрещенной зоны, примесное поглощение смещено от края собственного поглощения в длинноволновую область спектра и экспериментально может наблюдаться лишь при низких температурах, когда бóльшая часть атомов примеси не ионизирована. Взаимодействие фотонов с примесными атомами носит резонансный характер, т.е. вероятность поглощения существенно падает, если энергия фотонов значительно отличается от энергии ионизации примеси.

Поглощение света решеткой. Взаимодействие света с атомами кристаллической решетки приводит к превращению энергии квантов света в энергию фононов. Такой тип взаимодействия получил название *решеточного* или *фононного поглощения*. При поглощении фотона в кристалле могут рождаться один (*однофононное поглощение*) или несколько фононов (*многофононное поглощение*). Однофононное поглощение присуще скорее бинарным и более сложным полупроводникам. В гомеополярных полупроводниках типа германия и кремния происходит многофононное поглощение света. Решеточное поглощение связано с увеличением колебательной энергии атомов. Оно проявляется в далекой инфракрасной области спектра и накладывается на примесное поглощение и поглощение свободными носителями заряда.

Из всех рассмотренных механизмов оптического поглощения света лишь собственное и примесное поглощение сопровождается генерацией

добавочных носителей заряда. Эти механизмы поглощения называются *фотоактивными*.

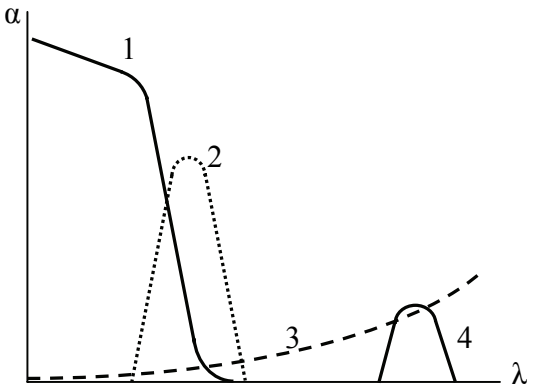


Рис. 3.19 – Зависимость коэффициента поглощения от длины волны излучения: 1 – собственное поглощение; 2 – экситонное поглощение; 3 – поглощение света свободными носителями заряда; 4 – примесное поглощение

Полный спектр поглощения полупроводника схематически показан на рис. 3.19. (Сплошными линиями выделены фотоактивные механизмы поглощения света).

Спектральная область между собственным и решеточным поглощениями у большинства полупроводников с малой концентрацией примеси структурных дефектов характеризуется высокой прозрачностью. Это позволяет использовать их в качестве оптических окон и светофильтров с резким краем поглощения в коротковолновой области.

3.1.5.3. Рекомбинация неравновесных носителей заряда

Процесс образования избыточных неравновесных носителей заряда называется процессом *генерации электронно-дырочных пар*. При генерации носителей заряда в полупроводниках с ярко выраженным типом проводимости (*p*- или *n*-тип) резко увеличивается концентрация неосновных носителей заряда.

Обратный процесс «сваливания» электронов из зоны проводимости в валентную зону называют процессом *рекомбинации носителей*. Очень важным параметром полупроводниковых материалов является *скорость рекомбинации R*, от значения которой зависят частотные характеристики полупроводниковых приборов.

Зависимость концентрации неравновесных носителей от времени после выключения возбуждения вследствие рекомбинации можно описать уравнением

$$n(t) = n_c e^{-\frac{t}{\tau}}, \qquad (3.66)$$

где n_c – стационарная концентрация электронов; τ – время жизни неравновесных носителей.

Равенство времен жизни электронов и дырок не обязательно. Поэтому вводят в рассмотрение разные времена жизни для электронов τ_n и дырок τ_p

$$\tau_n = \frac{\Delta n}{R_n}; \tau_p = \frac{\Delta p}{R_p}, \qquad (3.67)$$

где Δn и Δp – концентрации неравновесных носителей заряда, R_n и R_p – скорости рекомбинации электронов и дырок.

В общем случае τ_n и τ_p зависят от вида и механизма рекомбинации, состава полупроводника, температуры и т.д.

При генерации носители заряда получают энергию от внешнего воздействия, например, освещения. При обратном акте рекомбинации эта энергия выделяется. При этом не обязательно, чтобы поглощенная в виде светового кванта энергия выделялась также в виде светового кванта.

Классификацию механизмов рекомбинации принято осуществлять по способу отвода энергии, выделяющейся при акте захвата носителей заряда при рекомбинации.

Наиболее вероятными механизмами рекомбинации являются:

1) *излучательная рекомбинация*, при которой энергия выделяется в виде фотонов;

2) *многофононная рекомбинация*, связанная с непосредственной передачей выделяющейся энергии колебаниям атомов кристаллической решетки;

3) *ударная рекомбинация* (*эффект Оже*), при которой энергия передается ближайшему свободному электрону (или дырке), который затем отдает свою избыточную энергию либо колебаниям решетки, либо другим носителям.

Все эти три вида рекомбинации могут осуществляться как при рекомбинации зона-зона (*межзонная рекомбинация*), так и при рекомбинации через локальные центры (*рекомбинационные ловушки*).

Излучательная рекомбинация. *Межзонная излучательная рекомбинация* наиболее вероятна в прямозонных полупроводниках и сопровождается излучением фотонов. Время жизни τ носителей при межзонной излучательной рекомбинации зависит от степени легирования полупроводника только при малых уровнях инжекции. В достаточно сильно легированном полупроводнике времена жизни определяются концентрацией неосновных носителей заряда, т.е. в кристалле *n*-типа – дырками, а в кристалле *p*-типа – электронами. С увеличением степени легирования значение излучательного времени жизни τ уменьшается, так как растет концентрация основных носителей заряда, а концентрация неосновных носителей падает в соответствие с формулой $n_i^2 = n_n \cdot p_n = n_p \cdot p_p$.

Излучательная рекомбинация может происходить и на *локальных центрах*. Рекомбинационное излучение при переходе носителей заряда из соответствующих зон на энергетические уровни локальных центров экспериментально наблюдалась в германии и кремнии, а затем и в полупроводниках типа $A^{III}B^{V}$.

Излучательная рекомбинация носителей заряда происходит самопроизвольно, без внешнего электромагнитного воздействия. Такое *спонтанное* излучение имеет статистический характер и оно некогерентно. Это обусловлено тем, что акты спонтанного излучения происходят независимо друг от друга в разные моменты времени.

Переход электрона на более низкий энергетический уровень с излучением кванта света может произойти и под влиянием фотонов. Такую

рекомбинацию называют *вынужденной*, *индуцированной* или *стимулированной.* Таким образом, квант света с определенной частотой может не только поглощаться полупроводником, но и вызывать добавочное *индуцированное излучение*. Индуцированное излучение происходит в том же направлении, что и вызвавшее его излучение, в одной и той же фазе и с одинаковой поляризацией, т. е. индуцированное излучение является когерентным.

Безызлучательная рекомбинация. Безызлучательный механизм межзонной рекомбинации предполагает передачу энергии, освобождающейся в процессе рекомбинации, тепловым колебаниям решетки. Значение этой энергии, равное ширине запрещенной зоны, при этом должно размениваться на энергию группы фононов. Так, в GaAs E_g = 1,435 эВ при комнатной температуре, а $h\nu$ оптического фонона составляет приблизительно 0,03 эВ, т. е. в процессе безызлучательной рекомбинации должна участвовать группа из приблизительно 50 оптических фононов на каждую рекомбинирующую пару носителей, что маловероятно.

В настоящее время установлено, что значительная часть безызлучательных рекомбинационных переходов связана с локальными уровнями. В этом механизме рекомбинации существенную роль играет ионизация ловушки с энергией $E_и$. Если $E_и$ не превышает максимально возможной энергии фонона $(h\nu)_{max}$, то захват носителя сопровождается испусканием одного фонона. Если же $E_и$ больше $(h\nu)_{max}$, то рекомбинация сопровождается испусканием нескольких фононов. Обычно эффективными центрами являются глубокие уровни, энергия которых составляет несколько десятых долей электрон-вольта, тогда как величина $(h\nu)_{max}$ примерно в десять раз меньше.

В непрямозонных полупроводниках обычно преобладает безызлучательная рекомбинация.

***Ударная рекомбинация* (*Оже рекомбинация*)**. Процесс ударной рекомбинации предполагает участие трех носителей заряда: рекомбинирующую пару электрон-дырка и еще один электрон или дырку, которым передается энергия, выделяющаяся при рекомбинации.

Времена жизни носителей при ударной рекомбинации в легированных полупроводниках обратно пропорциональны квадрату равновесной концентрации носителей заряда, тогда как при излучательной рекомбинации они обратно пропорциональны первой степени концентрации. Это означает, что время жизни носителей при ударной рекомбинации больше времени жизни носителей при излучательной рекомбинации. В полупроводниках, по мере уменьшения ширины запрещенной зоны и увеличения температуры, вероятность ударной рекомбинации возрастает, а время жизни носителей уменьшается.

При наличии нескольких механизмов рекомбинации общее время жизни носителей заряда τ определяется по формуле:

$$\frac{1}{\tau} = \sum_{i}^{N} \frac{1}{\tau_i}, \qquad (3.68)$$

где τ_i – время жизни при i-м механизме рекомбинации; N – число каналов рекомбинации.

Отсюда следует правило, что при одновременном действии нескольких механизмов рекомбинации результирующее время жизни определяется наименьшим из времен жизни, соответствующем наиболее вероятному способу рекомбинации.

3.1.5.4. Люминесценция полупроводников

Все твердые тела в той или иной степени излучают свет. Следует различать тепловое излучение, которое является равновесным свойством тел и излучение электромагнитных волн нетеплового происхождения – люминесценцию. *Люминесценцией* называют электромагнитное нетепловое излучение, обладающее длительностью, значительно превышающей период световых колебаний.

Люминесценция представляет собой процесс выделения избыточной энергии, полученной кристаллом при переходе в неравновесное состояние. Более наглядно этот процесс был рассмотрен при изучении рекомбинации носителей заряда (п. 3.1.5.3). Излучательная рекомбинация – это один из видов люминесценции – рекомбинационное излучение полупроводника.

Вещества, способные люминесцировать, называются *люминофорами*. Для возникновения люминесценции атомы люминофора должны быть выведены из состояния термодинамического равновесия, т.е. возбуждены любыми энергетическими воздействиями. В зависимости от вида возбуждения люминофора различают *фотолюминесценцию* (возбуждение светом), *катодолюминесценцию* (возбуждение электронным пучком) и *электролюминесценцию* (свечение под действием электрического поля).

При люминесценции акты поглощения энергии полупроводником и излучения квантов света разделены во времени промежуточными процессами, что приводит к относительно длительному существованию свечения полупроводника после прекращения возбуждения.

Интенсивность люминесценции зависит от относительного вклада излучательной и безызлучательной рекомбинации. Для достижения высокой эффективности люминесценции необходимо уметь управлять механизмами рекомбинации, что представляет собой весьма сложную физическую и технологическую задачу. Генерация видимого люминесцентного излучения возможна лишь в прямозонных полупроводниках с шириной запрещенной зоны более 1,7 эВ.

3.2. КЛАССИФИКАЦИЯ ПОЛУПРОВОДНИКОВ

3.2.1. Простые полупроводники

Кремний. *Кремний* – основной материал современного полупроводникового производства. Он широко применяется для

изготовления дискретных полупроводниковых приборов, и является пока единственным полупроводниковым материалом для производства интегральных микросхем (ИМС). Это связано с тем, что чистый кремний имеет высокое удельное сопротивление, а наличие стабильного окисла на его поверхности позволяет широко использовать планарную технологию. Широкое применение кремния связано также и с тем, что технология массового производства интегральных микросхем на основе кремния значительно проще, а сам кремний распространен в природе гораздо шире других полупроводниковых материалов.

Кремний имеет кристаллическую решетку типа алмаза. Концентрация атомов кремния равна $5 \cdot 10^{28}$ в 1 $м^3$. На внешней оболочке атома кремния находятся 4 электрона. Основной параметр полупроводниковых материалов – ширина запрещенной зоны достаточно велика (1,12 эВ), вследствие чего приборы на его основе он имеют хорошую температурную стабильность. Кремний непрозрачен для видимого света, но для инфракрасных лучей с длиной волны более 1,2 мкм чистый кремний становится прозрачным. Наличие примесей ведет к увеличению коэффициента поглощения в инфракрасной области спектра.

Внешне кремний представляет собой темно-серый материал с металлическим блеском, твердый и хрупкий.

Кроме самого кремния в полупроводниковом производстве находит применение ряд его соединений: *двуокись кремния* SiO_2, *моноокись кремния* SiO, *нитрид кремния* Si_3N_4.

Получение и очистка кремния. Кремний – второй после кислорода элемент по распространенности в земной коре: его содержание достигает 29,5% по массе. Однако для полупроводниковой электроники кремний нашел широкое применение только во второй половине двадцатого столетия, после разработки эффективных способов его очистки.

В чистом виде кремний в природе не встречается, а входит в виде оксидов в большинство пород и минералов, и выделить его довольно трудно. Основными способами получения поликристаллического кремния являются восстановление тетрахлорида кремния парами водорода и термическое разложение моносилана SiH_4. Промышленность выпускает поликристаллический кремний семи марок КП-1 … КП-7, отличающихся содержанием примесей и размерами слитков.

Полученный поликристаллический кремний содержит 1…5% примесей, что слишком много для непосредственного использования в полупроводниковом производстве. Для получения кремния с удельной проводимостью, близкой к собственной, необходимо снизить содержание примесей в нем до $10^{19}…10^{18}$ атомов на кубический метр ($\sim 10^{-8}$ %). Очищенный *поликристаллический кремний* служит сырьем для выращивания монокристаллов и эпитаксиальных слоев кремния, а также применяется для изготовления некоторых типов полупроводниковых приборов.

Очистка кремния происходит методом *зонной плавки*. Идея метода

заключается в следующем. Большинство содержащихся в кремнии примесей (за исключением бора) гораздо лучше растворяются в жидкой фазе кремния, чем в твердой, поэтому при медленном остывании расплавленного кремния с одного конца примеси будут отталкиваться фронтом кристаллизации и оставаться в расплаве, а содержание их в твердой фазе будет меньше. Такое явление называется *сегрегацией*. Для осуществления сегрегации отдельную зону слитка кремния подвергают плавлению, постепенно передвигая ее вдоль слитка. В результате примеси вместе с расплавленной зоной будут сгоняться в конец слитка. Если повторить эту операцию несколько раз, можно добиться необходимой степени очистки кремния. Если кристалл расположить вертикально, а расплавленную фазу сделать не слишком большой, то она будет удерживаться силами поверхностного натяжения. Такой метод называется *бестигельной зонной плавкой*. Расплавленную зону можно создать с помощью печи сопротивления или с помощью катушки индуктивности, питаемой СВЧ током.

Для получения объемных *монокристаллов кремния* используют три основных метода: уже рассмотренный *метод бестигельной зонной плавки*, *метод вытягивания из расплава* (*метод Чохральского*) и *метод гарниссажной плавки.*

Метод вытягивания из расплава, или метод Чохральского, состоит в том, что очищенный и обезгаженный поликристаллический кремний расплавляют в тигле из особо чистого кварца в высоковакуумной печи. Затем с помощью штока вводят в расплав монокристаллическую затравку, ориентированную в нужном кристаллографическом направлении относительно поверхности расплава. Затем оплавленную затравку медленно поднимают, за ней тянется столбик расплава, удерживаемый силами поверхностного натяжения. По мере подъема столбик расплава остывает и кристаллизуется, причем с той же ориентацией, что и затравка.

Перспективным вариантом бестигельной зонной плавки является гарниссажный метод. При этом методе расплавленную зону создают за счет электронно-лучевого нагрева в глубоком вакууме. К преимуществам такой плавки относят большую однородность получаемых слитков и малое содержание кислорода (менее $5 \cdot 10^{11}$ м$^{-3}$), к недостаткам – повышенную плотность дислокаций $(3 \ldots 5) \cdot 10^{8}$ м$^{-2}$.

Использование монокристаллических слитков в полупроводниковом производстве сопряжено с большими потерями материала. Ведь действие большинства полупроводниковых приборов основано на процессах, происходящих в очень тонких граничных или поверхностных слоях полупроводника. Остальной объем монокристалла является подложкой и выполняет механические функции. Большая часть материала теряется при механической обработке слитков (резке, шлифовке и т.д.).

В настоящее время в полупроводниковом производстве наиболее широкое применение кремний находит в виде тонких монокристаллических слоев, получаемых осаждением на объемных монокристаллах, называемых

подложками. Полученные монокристаллические слои, сохраняющие кристаллографическую ориентацию подложки, называют *эпитаксиальными*. В качестве подложек используют монокристаллы самого кремния, а также сапфира, корунда и др.

Эпитаксиальные слои очень удобны при массовом производстве полупроводниковых приборов по наиболее распространенной *планарной технологии.* При использовании эпитаксии резко снижаются требования к чистоте и совершенству кристаллов, используемых в качестве подложек, а параметрами эпитаксиальных слоев легко управлять в процессе их выращивания.

В полупроводниковом производстве чистый монокристаллический кремний применяют редко. Действие большинства полупроводниковых приборов основано на использовании примесной электропроводности. Для придания кремнию примесной электропроводности в него вводят строго контролируемое количество необходимых примесей (легирование). В зависимости от характера влияния примеси на тип электропроводности все примесные элементы можно разделить на: нейтральные, акцепторные, донорные и создающие глубокие энергетические уровни. К нейтральным примесям относятся инертные газы, азот и элементы IV группы: кремний, свинец, олово. Основными акцепторными примесями в кремнии являются элементы III группы: галлий, индий, алюминий. Для создания донорных уровней в кремнии используют элементы V группы: мышьяк, сурьму, висмут, фосфор. Легирование кремния происходит одновременно с процессами получения монокристаллов и эпитаксиальных пленок.

Промышленность выпускает большое количество разнообразных марок монокристаллического кремния, отличающихся по величине и типу электропроводности, виду легирующих примесей, размерам слитков, величине удельного сопротивления и другим параметрам. В табл. 3.2 приведены основные параметры кремния.

Германий. Наличие германия в земной коре около $7 \cdot 10^{-4}$ %. Добывать его тяжело, так как минералов с большим содержанием германия немного. Основным сырьем для его получения являются продукты сгорания бурого угля, в одной тонне которого содержится до 0,1 кг германия.

Чистый германий имеет металлический блеск, твердость (хотя и меньшую, чем у кремния) и кристаллизуется в виде кубической решетки типа алмаза. Ширина запрещенной зоны невелика (при комнатной температуре E_g = 0,665 эВ), поэтому температурная стабильность приборов на базе германия мала. С этим связано и малое значение удельного сопротивления, что не эффективно для изготовления ИМС. Однако подвижность носителей заряда у германия высока, и это позволяет использовать его при изготовлении некоторых высокочастотных приборов.

Двуокись германия – неустойчивое соединение, растворяемое даже в воде, поэтому использовать его в качестве защитной пленки нельзя,

а это делает невозможным использование германия для производства интегральных микросхем по планарной технологии.

Для получения монокристаллического германия используют те же способы, что и для получения кремния. Однако, то, что германий даже в расплавленном состоянии не взаимодействует с углеродом и кварцем, позволяет в технологических процессах использовать тигли из графита и кварца.

Электрические свойства германия сильно зависят от количества содержащихся в нем примесей. Легирующие примеси в германии (как и в кремнии) делят на нейтральные, акцепторные, донорные и создающие глубокие энергетические уровни. Промышленность выпускает большое количество различных марок германия с электронным и дырочным типом электропроводности, легированных различными примесями с величиной удельного сопротивления 0,0001…45 Ом·см. В табл. 3.2 приведены основные параметры германия.

В связи с преимущественным развитием микроэлектроники германий все больше вытесняется кремнием и другими полупроводниковыми материалами. Но дискретные полупроводниковые приборы на основе германия, такие как высокочастотные транзисторы, импульсные диоды, фотодиоды, датчики Холла и др. по-прежнему находят широкое применение.

Для видимого света германий не прозрачен, для инфракрасных лучей относительно прозрачен при длине волны более 1,8 мкм. Рабочий интервал температур у германиевых приборов -60 ... +70 ^{0}C, поэтому более рационально использовать приборы на основе германия при низких температурах.

Селен. Из элементов VI группы периодической системы элементов Д. И. Менделеева полупроводниковыми свойствами обладают *сера, селен* и *теллур.* Из них в качестве простого полупроводника находит применение только селен, а теллур и серу используют лишь в составе их соединений – сульфидов и теллуридов, которые будут рассмотрены далее.

Селен (Se) существует в нескольких аллотропических формах, а также в аморфном виде. Свойства различных модификаций селена весьма сильно отличаются друг от друга. В табл. 3.2 приведены основные параметры наиболее устойчивой и широко применяемой гексагональной модификации селена.

Промышленность выпускает селен марок СЧ-1 и СЧ-2, содержащих соответственно 99,998% и 99,992% чистого селена с удельной проводимостью около 10^{-10} Ом$^{-1}$·м$^{-1}$. Для синтеза полупроводниковых соединений применяют особо чистый селен ОСЧ-А с содержанием примесей 10^{-5}…10^{-6} %, получаемый ректификацией селена СЧ-1. Монокристаллы селена получают осаждением из паровой фазы или выращиванием из расплава. Селен используют также в виде пленок толщиной 50…100 мкм, наносимых на подложки методом испарения в вакууме.

Аллотропическая форма существования селена зависит от способа получения и режима охлаждения расплавленного селена. Из аморфных модификаций селена наибольший интерес представляет стеклообразный селен черного цвета с плотностью 4300 кг/м^3. Стеклообразный селен нашел применение в фотографическом процессе ксерографии, где используется его свойство фотопроводимости.

Таблица 3.2 – Основные свойства кремния, германия и селена гексагональной модификации

Свойства	Кремний	Германий	Селен
Период решетки, нм	0,542	0,566	0,436
Плотность при 20 °С, кг/м^3	2320	5360	4800
Температурный коэффициент линейного расширения (0…100°С), К$^{-1}$	$4{,}2\cdot10^{-6}$	$6\cdot10^{-3}$	$25\cdot10^{-6}$
Удельная теплопроводность, Вт/(м·К)	80	55	3
Удельная теплоемкость (0…100°С), Дж/(кг·К)	710	333	330
Температура плавления, °С	1414	936	220
Теплота плавления, Дж/кг	$1{,}64\cdot10^{6}$	$4{,}13\cdot10^{5}$	$6{,}4\cdot10^{4}$
Коэффициент поверхностного натяжения при температуре плавления, Н/м	0,72	0,6	0,1
Собственное удельное сопротивление при температуре 20°С, Ом·м	$2\cdot10^{3}$	0,68	-
Собственная концентрация носителей заряда, м$^{-3}$	10^{16}	$2{,}5\cdot10^{19}$	-
Ширина запрещенной зоны, эВ при 0 К при 300 К	 1,165 1,12	 0,746 0,665	 2,5 2
Подвижность электронов, м^2/(В·с)	0,135	0,39	-
Подвижность дырок, м^2/(В·с)	0,048	0,19	$0{,}2\cdot10^{-4}$
Работа выхода электронов, эВ	4,3	4,8	2,85
Первый ионизационный потенциал, В	8,1	8,14	9,75
Диэлектрическая проницаемость	12,5	16	6,3

Для производства полупроводниковых приборов используют кристаллический гексагональный селен. Получают его нагреванием селена любой другой аллотропической формы до температуры 180…220 °С, близкой к температуре плавления. Кристаллохимические особенности гексагонального (серого) селена приводят к значительной анизотропии его механических, электрических и тепловых свойств. Например, проводимость и подвижность носителей тока в селене в зависимости от направления относительно кристаллографических осей может изменяться в 5 раз. По величине удельного сопротивления селен близок к диэлектрикам, но всегда обладает дырочной проводимостью, скорее всего, за счет неконтролируемых

примесей. В зависимости от способа измерений получают различные значения ширины запрещенной зоны E_g = 1,2…2 эВ. Селен обладает коэффициентом термо-ЭДС порядка 600 мкВ/К.

3.2.2. Полупроводниковые соединения. Общие свойства

Свойства простых полупроводников далеко не всегда отвечают требованиям современной полупроводниковой техники. Сложные полупроводники предоставляют широкие возможности для создания материалов с самыми различными свойствами.

Из сложных полупроводниковых материалов в настоящее время наиболее широкое применение нашли двойные кристаллические соединения типа A^mB^n, где $m+n=8$, например $A^{IV}B^{IV}$, $A^{III}B^{V}$, $A^{II}B^{VI}$. В этих соединениях кристаллическая решетка и среднее количество электронов на один атом такие же, как у кремния и германия, что предопределяет общность многих свойств. Кроме двойных кристаллических соединений широко используют соединения элементов IV группы свинца Pb и олова Sn с серой S, селеном Se, теллуром Te (элементы VI группы) – так называемые *халькогениды* (к ним же относятся и соединения типа $A^{II}B^{VI}$).

Для получения полупроводниковых материалов с промежуточными значениями параметров все чаще используют твердые растворы различных сложных полупроводников друг в друге: типа $A^{III}_x A^{III}_{1-x} B^{V}$ (например, $Al_xGa_{1-x}As$) и $A^{III}B^{V}_x B^{V}_{1-x}$ (например, GaP_xAs_{1-x}). В последнее время используются и четверные соединения $A^{III}B^{V}$ ($A^{III}_x A^{III}_{1-x} B^{V}_y B^{V}_{1-y}$), что позволяет управлять как шириной запрещенной зоны, так и постоянной кристаллической решетки.

Опыты по получению новых материалов этого типа проводятся даже на борту космических кораблей. Отсутствие силы тяжести и глубокий вакуум позволяют выращивать совершенные по структуре особо чистые кристаллы, получать растворы и сплавы несмешиваемых в земных условиях элементов. Так, на орбитальной станции «Салют» проводились эксперименты по получению монокристаллов твердых растворов Sn-Pb-Te из газовой фазы, монокристаллических пленок ZnS-CdS-ZnSe-CdSe и др.

Полупроводниковые свойства проявляются не только в кристаллических, но и в аморфных соединениях, в частности в *неорганических стеклах.* Стеклообразные полупроводники обладают только электронной электропроводностью. Введение примесей незначительно изменяет свойства стеклообразных полупроводников, примесная проводимость в них не проявляется. Полупроводниковыми свойствами обладают как *кислородсодержащие стекла,* получаемые сплавлением оксидов металлов (например, ванадий-фосфатные стекла типа V_2O_5–P_2O_5–ZnO), так и *бескислородные халькогенидные стекла,* получаемые сплавлением халькогенов с элементами III, IV и V групп периодической системы. Стеклообразные полупроводники пока еще не нашли широкого промышленного применения.

3.2.3. Полупроводниковые соединения типа $A^{IV}B^{IV}$

Единственным двойным соединением элементов IV группы периодической системы в твердой фазе является соединение кремния с углеродом – карбид кремния SiC. В состав карбида кремния стехиометрического состава входит 70% кремния и 30% углерода. Атомы углерода и кремния связаны ковалентной и частично ионной связью. Обладая очень большой энергией связи атомов, карбид кремния является исключительно прочным соединением, по твердости немного уступающим алмазу. Жаростойкость карбида кремния также очень высока.

Карбид кремния обладает *полиморфизмом*, при температуре выше 2000 ºС кубическая модификация β – SiC переходит в гексагональную α–SiC. При температурах 2600…2700 ºС карбид кремния возгоняется. Карбид кремния обладает очень высокой химической стойкостью. При комнатной температуре он вступает в реакцию только с расплавленной ортофосфорной кислотой и смесью азотной и плавиковой кислот, а также с расплавленными щелочами, которые используют в качестве травителей. Окисление карбида кремния происходит только при температуре более 800 ºС.

Карбид кремния имеет довольно сложную структуру энергетических зон. Для различных политипов α-модификации ширина запрещенной зоны E_g=2,72…3,34 эВ, а для β-SiC – E_g=2,39 эВ. Такое большое значение E_g позволяет создавать на основе карбида кремния полупроводниковые приборы, которые сохраняют работоспособность при температурах до 400…500 ºС. Собственная электропроводность карбида кремния из-за большой ширины запрещенной зоны наблюдается лишь при температурах более 1400 ºС. Подвижность носителей заряда в карбиде кремния низкая (не превышает 0,1 $м^2$/В·с для электронов и 0,02$м^2$/В·с для дырок).

Кристаллы чистого карбида кремния стехиометрического состава бесцветны. Наличие примесей окрашивает их в самые разнообразные цвета. Примеси элементов V группы (прежде всего азот), а также Li, O, являясь для карбида кремния донорами, придают ему электропроводность *n*-типа и окрашивают его в светло-зеленый цвет.

Акцепторными примесями для карбида кремния являются элементы III группы периодической системы (B, Al, Sc, Ga) и II группы (Be, Mg, Ca). Кристаллы карбида кремния с электропроводностью *p*-типа имеют голубую и черную окраску.

Роль примесей могут играть и сами атомы кремния и углерода. Избыток кремния относительно стехиометрического состава придает карбиду кремния электронную электропроводность, а избыток углерода – дырочную.

Диффузию примесей в карбиде кремния производят в закрытом объеме при температурах 1800…2200 °С. Донорные примеси имеют очень низкие коэффициенты диффузии в карбиде кремния, поэтому ввести их в карбид кремния в необходимом количестве методом диффузии не удается. Вследствие этого возможен только один путь создания электронно-дырочных

переходов: в карбид кремния, обладающий электронной проводимостью, производят диффузию необходимых акцепторных примесей.

Для создания омических контактов с карбидом кремния электронного типа проводимости используют сплав кремния с фосфором, для контактов с карбидом кремния дырочного типа – сплав кремния с алюминием или бором.

Чаще всего карбид кремния получают в виде поликристаллов (друз). Большинство кристаллов в друзах имеют незначительные размеры, но встречаются и кристаллы с площадью до 1,5…2 см^2. Карбид кремния имеет большую твердость.

3.2.4. Полупроводниковые соединения типа $A^{III}B^{V}$

Сложные полупроводники типа $A^{III}B^{V}$ представляют собой химические соединения металлов III группы периодической системы (алюминия, галлия, индия, бора) с элементами V группы (фосфором, мышьяком, сурьмой и азотом).

Обладая широким спектром различных электрофизических свойств, большой химической стойкостью и механической прочностью, соединения $A^{III}B^{V}$ позволяют выбрать среди них материалы с необходимыми свойствами. Соединения $A^{III}B^{V}$ используют для производства самых различных полупроводниковых приборов, особенно оптоэлектронных.

Наиболее широкое применение в полупроводниковом производстве нашли соединения мышьяка – арсениды галлия GaAs и индия InAs, соединения сурьмы – антимониды индия InSb и галлия GaSb и соединения фосфора – фосфид галлия GaP. Перспективны соединения алюминия с сурьмой и мышьяком AlSb и AlAs, а также с азотом A1N.

Основные параметры соединений. Основные параметры наиболее распространенных соединений $A^{III}B^{V}$ приведены в таблице 3.3.

Таблица 3.3 – Основные параметры соединений $A^{III}B^{V}$

Соединение	Постоянная решетки, нм	Плотность, кг/м^3	Ширина запрещенной зоны, эВ при T = 300 К	Температура плавления, °С	Подвижность, м^2/(В·с)		Низкочастотная диэлектрическая проницаемость
					электронов	дырок	
AlP	0,5463	2370	2,42	2000	0,008	0,003	9,8
GaP	0,5451	4070	2,25	1467	0,019	0,012	11,1
InP	0,5869	4780	1,28	1055	0,46	0,015	12,4
AlAs	0,5661	3600	2,16	1700	0,028	-	10,1
GaAs	0,5653	5320	1,435	1237	0,89	0,045	13,1
InAs	0,6058	5670	0,36	942	3,30	0,046	14,6
AlSb	0,6136	4280	1,58	1060	0,02	0,055	14,4
GaSb	0,6096	5650	0,72	712	0,40	0,14	15,7
InSb	0,6479	5780	0,18	525	7,80	0,075	17,7

Все перечисленные соединения $A^{III}B^{V}$ являются кристаллическими веществами со структурой типа *цинковой обманки* (*сфалерит*). Соединения $A^{III}B^{V}$ являются ближайшими электронными аналогами кремния и германия. Они образуются в результате взаимодействия элементов Ш-*б* подгруппы Периодической системы (бора B, алюминия Al, галлия Ga, индия In) с элементами V-*б* подгруппы (азотом N, фосфором P, мышьяком As и сурьмой Sb). Висмут Bi и таллий Tl не образуют соединений рассматриваемого ряда. Соединения $A^{III}B^{V}$ принято классифицировать по металлоидному элементу Соответственно, различают нитриды, фосфиды, арсениды и антимониды. Прочность связи атомов в кристаллической решетке внутри каждой группы (нитриды, фосфиды, арсениды и антимониды) уменьшается при увеличении суммарного атомного номера, что приводит к уменьшению ширины запрещенной зоны, температуры плавления и химической стойкости.

За исключением нитридов, для которых характерна структура гексагонального типа (*вюрцит*), все соединения $A^{III}B^{V}$ кристаллизуются в решетке *цинковой обманки* (*сфалерит*). В решетке того и другого типов атом элемента Ш группы находится в тетраэдрическом окружении четырех элементов V группы и наоборот. Структура сфалерита, в отличие от структуры алмаза, не имеет центра симметрии. Эта особенность приводит к различию в свойствах поверхностей (111) и $(\bar{1}\bar{1}\bar{1})$, целиком сложенных из разноименных атомов. Различия в поведении граней (111) и $(\bar{1}\bar{1}\bar{1})$ проявляется при травлении, окислении и при выращивании кристалла.

Для соединений $A^{III}B^{V}$ характерен особый вид химической связи, называемой донорно-акцепторной. Из четырех ковалентных связей, которыми каждый атом встраивается в решетку, три образуются обобществлением валентных электронов атомов A^{III} и B^{V}, а четвертая связь осуществляется неподеленной парой валентных атомов B^{V} (рис. 3.20). Образование этой связи соответствует энергетически выгодному переходу электронов от атома B^{V} в энергетическое состояние, общее для донора (атомов B^{V}) и акцептора (атомов A^{III}). В каждой ковалентной связи максимум электронной плотности смещен в сторону атома с более высокой электроотрицательностью, т. е. электронные облака стянуты к узлам решетки, где находятся атомы B^{V}. Благодаря такой поляризации связей атомы A^{III} приобретают некоторый эффективный положительный заряд, а атомы B^{V} – отрицательный. Величина этого эффективного заряда определяет степень ионности соединения, которая изменяется при переходе от одного химического соединения к другому. Таким образом, можно сказать, что отличие структуры сфалерита от кристаллической решетки кремния и германия состоит в том, что в узлах кристаллической решетки соединений

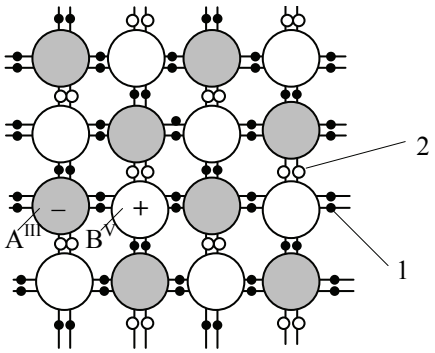


Рис. 3.20 – Плоская модель кристаллической решетки соединений $A^{III}B^{V}$: 1 – ионные связи; 2 – электронные связи

$A^{III}B^{V}$ находятся не нейтральные атомы, а поочередно расположенные положительно заряженные ионы B^{V} и отрицательно заряженные ионы A^{III}. Такая кристаллическая структура образуется за счет того, что один из пяти валентных электронов атома B^{V} переходит к атому A^{III}. В результате три межатомные связи носят электронный характер, а одна – ионный.

Соединения $A^{III}B^{V}$ химически весьма устойчивы, при комнатной температуре не взаимодействуют с водой и с кислородом воздуха.

Используя соединения с широкой запрещенной зоной (GaAs, GaP), можно изготовлять полупроводниковые приборы с рабочей температурой более 300 °С.

Характерной особенностью соединений $A^{III}B^{V}$ является высокая подвижность в них электронов. Наибольшего значения этот параметр достигает в антимониде индия, у которого μ_n = 7,8 м^2/В·с. Уникальным соотношением подвижности носителей обладает антимонид алюминия AlSb: подвижность электронов (0,02м^2/В·с) в нем меньше подвижности дырок (0,055м^2/В·с), хотя обычно в полупроводниках подвижность электронов, как правило, больше подвижности дырок. Высокая подвижность носителей заряда в соединениях $A^{III}B^{V}$ позволяет создавать высокочастотные полупроводниковые приборы: СВЧ-генераторы, детекторы микроволнового диапазона и др.

В соединениях $A^{III}B^{V}$ ярко выражены все оптические явления: поглощение электромагнитного излучения, фотопроводимость, спонтанное и стимулированное излучение. В зависимости от ширины запрещенной зоны материала эти явления наблюдаются в широком интервале длин волн, включая видимую и инфракрасную области спектра. На основе соединений $A^{III}B^{V}$ созданы самые разнообразные оптоэлектронные приборы: фоторезисторы, светодиоды, лазеры, приемники инфракрасного излучения, фотоумножители, оптические фильтры и др. (см. раздел 3.3).

В настоящее время все бóльшую роль играют твердые растворы соединений $A^{III}B^{V}$ друг в друге, так как это позволяет получать полупроводниковые материалы с варьируемыми значениями ширины запрещенной зоны E_g.

3.2.5. Полупроводниковые соединения типа $A^{II}B^{VI}$ (халькогениды)

Основные свойства соединений типа $A^{II}B^{VI}$. Полупроводниковые материалы этого типа представляют собой химические соединения халькогенов (серы S, селена Se, теллура Te) с металлами второй группы – цинком Zn, кадмием Cd, ртутью Hg.

Кристаллизуются соединения $A^{II}B^{VI}$ в структурах двух типов: кубической (типа сфалерита), образующейся при низких температурах, и гексагональной.

Каждый атом в этих решетках связан с четырьмя соседними, причем две связи носят ковалентный характер, а две – ионный. Так же, как и в соединениях типа $A^{III}B^{V}$, прочность связи атомов в кристаллической решетке

внутри каждой группы (сульфидов, селенидов и теллуридов) уменьшается при увеличении суммарного атомного номера. Соответственно уменьшаются ширина запрещенной зоны (от 3,7 эВ для ZnS до 0,02 эВ для HgTe), температура плавления (от 1830 до 670 °С) и химическая стойкость.

Основные параметры наиболее используемых соединений типа $A^{II}B^{VI}$ приведены для температур $T = 300$К в таблице 3.4.

Таблица 3.4 – Параметры полупроводниковых соединений $A^{II}B^{VI}$

Соединение	Тип структуры	Период решетки, нм	Плотность, кг/м^3	Температура плавления, °С	Ширина запрещенной зоны, эВ	Тип проводимости	Подвижность, м^2/(В·с)	
							электронов	дырок
ZnS	С*	0,541	4090	1020**	3,67	*n*	–	–
ZnS	В*	0,382(*a*) 0,626(*c*)	4100	1780	3,74	*n*	0,14	0,0005
CdS	В*	0,413(*a*) 0,675(*c*)	4820	1750	2,53	*n*	$\perp c$ 0,034 $\parallel c$ 0,035	0,011 0,0015
HgS	С*	0,584	7730	1480	1,78	*n*	0,07	–
ZnSe	С*	0,566	5420	1520	2,73	*n*	0,026	0,0015
CdSe	В*	0,430(*a*) 0,701(*c*)	5810	1264	1,85	*n*	$\parallel c$ 0,072	0,0075
HgSe	С*	0,608	8250	790	0,12	*n*	2,0	–
ZnTe	С*	0,610	6340	1239	2,23	*p*	0,053	0,003
CdTe	С*	0,648	5860	1041	1,51	*n, p*	0,12	0,006
HgTe	С*	0,646	8090	670	0,08	*n, p*	2,5	0,02

Примечание: * С – сфалерит; В – вюрцит; ** – температура фазового перехода из β в α-модификацию.

Соединений типа $A^{II}B^{VI}$ обладают высокой чувствительностью к излучению в различных частях спектра (от инфракрасного до рентгеновского), ярко проявляя фоторезистивные и люминесцентные свойства, чем и обусловлена их область применения. Они используются в основном в качестве люминофоров для всех видов люминесценции и материалов для фоторезисторов.

Технология получения соединений типа $A^{II}B^{VI}$ довольно проста. Поликристаллы чаще всего получают осаждением из водных растворов солей с помощью H_2S. Таким же образом получают сульфиды и теллуриды. Для получения теллурида кадмия и селенидов используют непосредственное сплавление исходных компонентов.

Соединения типа $A^{II}B^{VI}$ используют как в монокристаллическом, так и в поликристаллическом виде (пленки и прессованные таблетки). Монокристаллы получают выращиванием из расплава и направленной кристаллизацией. Эпитаксиальные пленки получают осаждением из газовой фазы на ориентированные подложки. Поликристаллические пленки соединений типа $A^{II}B^{VI}$ получают чаще всего методом катодного распыления в вакууме с последующей рекристаллизацией отжигом. Такие пленки

обладают высоким *пьезомодулем* и находят применение, наряду с *пьезоэлектриками*, в качестве преобразователей электромагнитных колебаний в аккустические в диапазонах ВЧ и СВЧ.

Сильное влияние на тип проводимости соединений $A^{II}B^{VI}$ оказывает отклонение их состава от стехиометрического. Избыток металла A^{II} приводит к появлению электронной электропроводности, избыток халькогена B^{VI} – к появлению дырочной электропроводности.

Халькогениды цинка. К ним относятся: *сульфид цинка* ZnS, *селенид цинка* ZnSe и *теллурид цинка* ZnTe. Из соединений $A^{II}B^{VI}$ наибольшее распространение получил сульфид цинка ZnS.

Удельное сопротивление ρ кристаллов сульфида цинка составляет 10^{12} Ом·м, а ширина запрещенной зоны около 3,7 эВ. Таким образом, по ширине запрещенной зоны и удельному сопротивлению сульфид цинка близок к диэлектрикам.

Нелегированные кристаллы сульфида цинка обладают электронной проводимостью. Для получения дырочной проводимости используют легирование элементами I группы.

Сульфид цинка в виде кристаллического порошка (реже в виде пленок) широко используют для изготовления люминофоров различных приборов: осциллографических и телевизионных электронно-лучевых трубок, люминесцентных осветительных ламп и др. Цвет свечения, длительность послесвечения и другие параметры, люминофоров можно изменять в широких пределах, изменяя тип активатора. В качестве активаторов используют Ag, Cu, Mn в количестве 0,001…3%, а также избыток цинка. Введение никеля в сульфид цинка позволяет получить очень малое время послесвечения.

Для повышения устойчивости люминофоров на основе ZnS используют твердые растворы типа $Zn_xCd_{1-x}S$ и ZnS_xSe_{1-x}. Частичное замещение сульфида цинка сульфидом кадмия увеличивает светоотдачу люминофора; введение селенида цинка делает люминофор устойчивым к влиянию различных примесей и к бомбардировке электронами.

Монокристаллы ZnS используют в качестве оптического материала, прозрачного в инфракрасной области спектра, для создания лазеров с длиной волны $\lambda = 0{,}32$ мкм, а также для пьезоусилителей акустических колебаний.

Халькогениды кадмия. Основным материалом для фоторезисторов и фотоэлементов является сульфид кадмия CdS. Он всегда обладает электронной проводимостью ввиду недостатка серы относительно стехиометрического состава. Удельное сопротивление под влиянием различных факторов (примесей, температуры, излучения) изменяется от 0,1 до 10^7 Ом·м. Сульфид кадмия является самым чувствительным фоторезистором в видимой части спектра: при освещении его сопротивление уменьшается в $10^4…10^6$ раз. Сульфид кадмия используют в качестве люминофора, для изготовления дозиметров ультрафиолетовых, рентгеновских и гамма-лучей, счетчиков частиц и т. д.

Халькогениды элементов других групп. Типичными полупроводниковыми соединениями типа $A^{IV}B^{VI}$ являются халькогениды свинца: сульфид свинца PbS, селенид свинца PbSe и теллурид свинца PbTe. Халькогениды свинца кристаллизуются в кубической решетке NaCl; характер химической связи атомов близок к ковалентному.

Халькогениды свинца не растворяются в воде, медленно окисляются на воздухе. Для их травления используют нагретую соляную и азотную кислоту, горячие растворы щелочей.

Халькогениды свинца являются узкозонными полупроводниками с E_g порядка 0,22…0,3 эВ.

Проводимость халькогенидов свинца очень чувствительна к нарушениям стехиометрического состава. Незначительный избыток свинца (0,003% для PbSe) приводит к появлению электронной проводимости, такой же избыток халькогена – к появлению дырочной проводимости. Акцепторными примесями в халькогенидах свинца являются металлы I группы (Ag, K, Na, Cu), замещающие атомы свинца. Роль доноров играют элементы III группы, замещающие атомы свинца, и галогены VI группы, замещающие атомы халькогена.

Халькогениды свинца обладают хорошими фоторезистивными свойствами. При низких температурах в халькогенидах свинца становится (эффективной) существенной излучательная рекомбинация носителей заряда, что позволяет использовать эти соединения для создания лазеров инфракрасного диапазона.

Теллурид свинца, обладающий высоким коэффициентом термо-ЭДС и малой теплопроводностью, является эффективным материалом для создания полупроводниковых термоэлементов, работающих при температурах 300…700 °С. Используют также твердые растворы типа PbS–PbSe–PbTe.

Халькогениды элементов V группы. Из полупроводниковых соединений типа $A^{V}_2B^{VI}_3$ нашли применение сульфид сурьмы Sb_2S_3, селенид и теллурид висмута Bi_2Se_3 и Bi_2Te_3.

Соединения типа $A^{V}_2B^{VI}_3$ кристаллизуются в виде ромбических и ромбоэдрических структур, обладающих низкой симметрией. Поэтому они обнаруживают сильную анизотропию свойств даже в пленочных образцах. Эти соединения синтезируют сплавлением исходных компонентов при пониженном давлении в запаянных кварцевых ампулах. Монокристаллы получают методом направленной кристаллизации. Проводимость соединений типа $A^{V}_2B^{VI}_3$ очень чувствительна к отклонениям от стехиометрического состава.

В заключение отметим, что сложные полупроводники вообще и соединения халькогенов в частности являются чрезвычайно перспективными материалами, исследование свойств которых продолжается и в настоящее время.

3.3. ПРИМЕНЕНИЕ ПОЛУПРОВОДНИКОВ

Возможность управления величиной и типом проводимости полупроводников в широком температурном диапазоне лежит в основе создания практически всех полупроводниковых устройств. Рассмотрим, как при конструировании наиболее широко применяемых электронных устройств используются физические свойства полупроводников, и как меняется работа этих устройств при целенаправленном изменении свойств полупроводниковых материалов.

3.3.1. Приборы на одном типе проводимости

На основе собственных полупроводников, легируемых одним типом примесей, можно создать целый ряд электронных устройств. Простейшие устройства, представляющие собой пластинку собственного полупроводника, легированного примесями одного типа до необходимой концентрации носителей, способны измерять значение магнитных полей, механических напряжений, температуру и выполнять многие другие функции.

3.3.1.1. Датчики Холла

Эффект Холла получил широкое практическое применение. Появление холловской ЭДС при помещении полупроводниковой пластины с током в поперечное магнитное поле используется в большом числе измерительных устройств и схем автоматики, в магнитометрах, предназначенных для измерения постоянных и переменных магнитных полей, токометрах для измерения токов, проходящих в электрических кабелях, ваттметрах, генераторах электрических колебаний и т.д. Преимуществами этих приборов является их безынерционность, отсутствие подвижных частей и малые габариты.

Холловское напряжение можно найти по формуле (3.52). Если учесть, что сила тока равна $I = \frac{U}{R}$, а сопротивление полупроводника $R = \rho\frac{l}{S} = \frac{\sigma l}{xb}$, то вместо (3.52) получим соотношение

$$\frac{U_X}{U \cdot B} = R_X \sigma \frac{x}{l}. \qquad (3.69)$$

Отношение $\frac{U_X}{U \cdot B}$ представляет собой относительную чувствительность датчика Холла. Из (3.69) видно, что чувствительность зависит от значения $R_X\sigma$ ($R_X\sigma = \frac{A}{en} \cdot en\mu = A \cdot \mu$), т.е. от подвижности основных носителей заряда в полупроводнике. Таким образом, чем больше подвижность μ носителей заряда в полупроводнике, тем больше чувствительность датчиков Холла. К недостаткам датчиков Холла можно отнести некоторое уменьшение их чувствительности с ростом температуры.

Высокой подвижностью обладают соединения индия InAs ($A^{III}B^{V}$) (μ = 3,3 м^2/(В·с)) и InSb (μ = 7,8 м^2/(В·с)); соединения ртути с селеном и теллуром ($A^{II}B^{VI}$) HgSe (μ = 2,0 м^2/(В·с)) и HgTe (μ = 2,5 м^2/(В·с)), твердые растворы $InAs_xP_{x-1}$.

Кроме того, одним из требований, предъявляемых к датчикам Холла, является их термостабильность. Тут успешно применяются кремний и германий, хотя эти полупроводники и обладают небольшой подвижностью носителей заряда.

Эффект Холла является также мощным экспериментальным средством для изучения свойств носителей тока в полупроводниках:

1) Измерив постоянную Холла, можно найти концентрацию основных носителей:

$$n = \frac{A}{R_X \cdot e}. \tag{3.70}$$

2) По полярности ЭДС Холла можно установить знак носителей.

3) Зная A и удельную проводимость примесного полупроводника σ ($\sigma = en\mu$), можно определить подвижность носителей по формуле:

$$\mu = \frac{R_X \cdot \sigma}{A}. \tag{3.71}$$

4) Можно вычислить энергию активации акцепторов или доноров. Как видно из формул (3.54-3.56) постоянная Холла R_X зависит от концентрации носителей заряда, а, следовательно, и от температуры. Измеряя ЭДС Холла в некотором диапазоне температур, получим экспериментальные данные для построения графика зависимости концентрации носителей от температуры. По графику можно вычислить энергию активации акцепторов или доноров.

3.3.1.2. Термоэлектрические устройства

Простейшее устройство, использующее эффект термо-ЭДС и предназначенное для определения типа проводимости, получило название термозонда. К примеру, если нагреть один конец (например, правый на рис. 3.21) полупроводника *n*-типа, то концентрация свободных электронов на этом конце возрастёт и электроны начнут диффундировать в сторону холодного конца. В результате термоэлектрического эффекта на левом холодном конце появится избыток отрицательных зарядов, на правом горячем конце – избыток положительных зарядов, т. е. на концах полупроводника появится термо-ЭДС. В полупроводнике *p*-типа в силу того, что перемещаться будут положительно заряженные дырки, знак термо-ЭДС изменится на противоположный. Следовательно, определить тип электропроводности можно по знаку термо-ЭДС (рис 3.21 *а*, *б*).

Численное значение коэффициента дифференциальной термо-ЭДС полупроводников в зависимости от температуры и концентрации примесных атомов колеблется в пределах $10^2 \ldots 10^3$ мкВ/К, т.е. на 2…3 порядка выше, чем у металлов. Такое различие в свойствах материалов приводит к тому, что в замкнутой цепи, составленной из полупроводника и металла, термо-ЭДС

слабо зависит от природы металлического проводника. Поэтому, приводя значения α_T для полупроводников, обычно не указывают, относительно какого металла она определена.

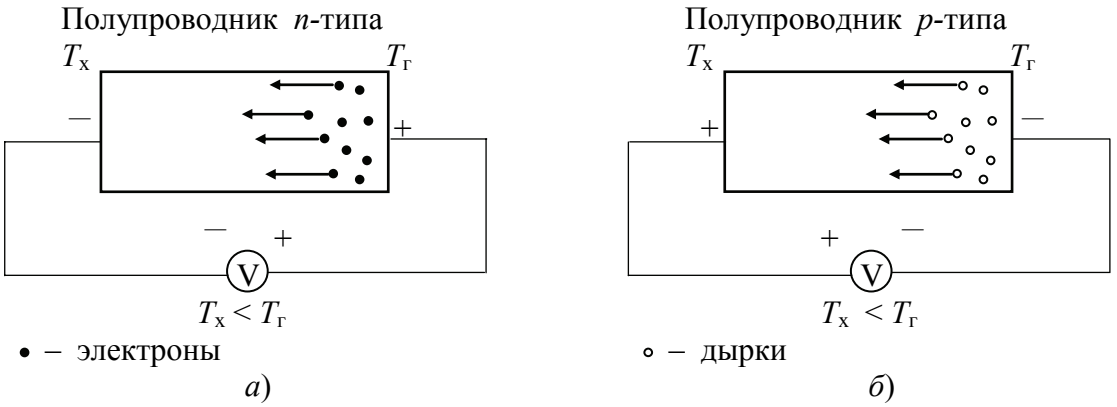


Рис. 3.21 – Методы определения типа электропроводности для полупроводников: *а* – *n*-типа; *б* – *p*-типа ($T_х$ – температура холодного конца полупроводника, $T_г$ – температура горячего конца полупроводника)

Большая термо-ЭДС полупроводников позволяет использовать их в качестве эффективных преобразователей тепловой энергии в электрическую. На основе термоэлектрического эффекта созданы термоэлектрические генераторы и солнечные батареи. Полупроводниковые материалы для этих целей должны обладать высоким коэффициентом термо-ЭДС α_T и низким коэффициентом теплопроводности χ. Сульфид сурьмы обладает высоким значением коэффициента термо-ЭДС – до 1000 мкВ/К. Такими же свойствами обладают фосфид галлия GaP, соединения висмута и селена BiSe, висмута и теллура BiTe и др.

3.3.1.3 Полупроводниковые тензометры.

Полупроводниковый датчик деформаций представляет собой кусочек полупроводника с двумя омическими контактами, который изменяет свое сопротивление при механической нагрузке; форму датчика подбирают такой, чтобы его было удобно приклеивать к испытываемому объекту.

Полупроводниковый материал для качественных тензометров должен обладать высокой тензочувствительностью, которая зависит от типа проводимости применяемого материала, величины его удельного сопротивления, кристаллографического направления, в котором приложена механическая сила.

Для кремния *p*-типа тензочувствительность лежит в пределах от 100 до 200, для германия *n*-типа – в пределах -147…-157, для германия *p*-типа – в пределах 100…105. Кроме кремния и германия, высокой тензочувствительностью обладают GaSb (для *n*-типа G = -226), InSb (для *p*-типа G = 212), PbTe и некоторые другие полупроводники. Тензочувствительность увеличивается с повышением удельного сопротивления, т. е. с уменьшением содержания примесей. Однако в этом случае тензодатчики становятся менее стабильными при изменении

температуры окружающей среды. Для изготовления полупроводниковых тензометров оказался пригодным и карбид кремния. Тензометры из карбида кремния способны стабильно работать при повышенных температурах.

Часто удобнее измерять величину, называемую *пьезосопротивлением* (или *тензосопротивлением*). Если X – механическое напряжение, $\frac{d\rho}{\rho}$ – относительное изменение удельного сопротивления образца при действии напряжения X, то пьезосопротивление можно найти по формуле:

$$\Pi = \frac{d\rho}{\rho \cdot X}. \qquad (3.72)$$

Обычно X выражают в Н/м^2, Π также имеет размерность м^2/Н. Величина пьезосопротивления Π различна в разных направлениях в кристалле. Однако всегда имеется преимущественное направление, в котором пьезосопротивление максимально.

3.3.2. Биполярные приборы

3.3.2.1. Электронно-дырочные переходы

Если пластинку собственного полупроводника легировать в одной части донорными примесями, а в другой части – акцепторными, получится новое устройство – электронно-дырочный переход (*p*-*n*-переход), т.е. электронно-дырочный переход в кристалле полупроводника создается изменением типа проводимости кристалла вдоль одного направления. Изучим подробнее физические свойства *p*-*n*-перехода, так как это устройство наиболее широко распространено в электронной технике, как в самостоятельной форме, так и в определенных сочетаниях. Для простоты будем считать распределение примесей постоянным в направлении x и с резкой границей между *p*- и *n*-областями, хотя есть *p*-*n*-переходы и с плавной границей – все зависит от целей создания данного устройства.

Для создания мелких примесей используются вещества, у которых энергии ионизации акцепторов и доноров малы, поэтому акцепторы и доноры практически полностью ионизированы. В *p*-области отрицательные заряды ионизированных неподвижных акцепторов скомпенсированы в каждой точке положительными зарядами подвижных дырок, т. е. $p_p = n_p + N_A$. Аналогично для электронов в *n*-области: $n_n = p_n + N_D$. Это и есть *условие электронейтральности*.

В легированных полупроводниках обычно концентрации основных носителей намного больше концентрации неосновных: $n_n >> p_n$ в *n*-области, $p_p >> n_p$ – в *p*-области; понятно, что и $p_p >> p_n$, а $n_n >> n_p$. Вследствие этого на границе *p*-*n*-перехода возникают диффузионные потоки дырок из *p*- в *n*-область и электронов из *n*-области в *p*-область.

Из-за ухода электронов в *n*-области вблизи границы *p*-*n*-перехода остается нескомпенсированный положительный заряд ионизированных доноров, а в *p*-области – нескомпенсированный отрицательный заряд ионизированных акцепторов. Таким образом, на границе *p*-*n*-перехода

образуется двойной электрический слой, препятствующий дальнейшей диффузии электрических зарядов – область объемного заряда (ООЗ). Из-за двойного электрического слоя и контактного поля на границах *p*-*n*-перехода устанавливается контактная разность потенциалов, которая сдвигает энергетические уровни в двух областях *p*-*n*-перехода на величину, равную разности уровней Ферми в *p*- и *n*-областях при отсутствии контакта. Этот сдвиг происходит из-за выравнивания уровней Ферми как термодинамической функции по обе стороны *p*-*n*-перехода (рис. 3.22).

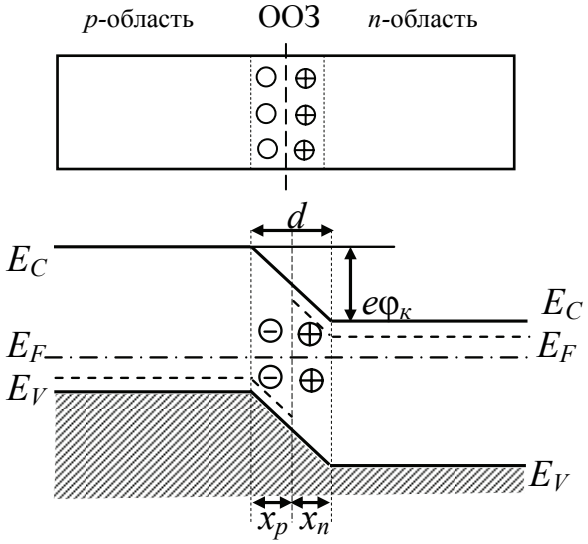


Рис. 3.22 – *p*-*n*-переход, находящийся в термодинамическом равновесии

Исходя из этого, определим контактную разность потенциалов по формуле:

$$\varphi_к = \frac{E_{F_n} - E_{F_p}}{e}, \qquad (3.73)$$

где энергии E_{F_n} и E_{F_p}, определяющие сдвиг зон на рис. 3.22, отсчитываются от дна зоны проводимости.

Зная уровни легирования в *p*- и *n*-областях, найдем значение контактной разности потенциалов следующим образом:

$$n_n p_p = N_C N_V \exp\left(-\frac{E_{F_P}\text{-}E_V}{kT}\right)\cdot\exp\left(-\frac{E_C\text{-}E_{F_n}}{kT}\right) = N_C N_V \exp\left(-\frac{E_g}{kT}\right)\exp\left(-\frac{E_{F_p}\text{-}E_{F_n}}{kT}\right),$$

откуда следует, что

$$n_n p_p = n_i^2 \cdot \exp\left(-\frac{E_{F_P} - E_{F_n}}{kT}\right) = n_i^2 \cdot \exp\left(\frac{e\varphi_к}{kT}\right).$$

Произведя преобразования и потенцируя вышеприведенное выражение, получим контактную разность потенциалов

$$\varphi_к = \frac{kT}{e}\cdot\ln\left(\frac{p_p \cdot n_n}{n_i^2}\right). \qquad (3.74)$$

Из (3.74) следует, что основное влияние на электрические параметры *p*-*n*-перехода оказывает степень легирования полупроводников донорами и акцепторами, так как от нее зависит положение уровней Ферми в запрещенной зоне. При малых и умеренных концентрациях примесей мы получаем *p*-*n*-переход, лежащий в основе работы выпрямительного диода. Варьируя концентрации доноров и акцепторов, можно менять значение контактной разности потенциалов.

Из подстановки (3.2) в (3.74) следует, что

$$\varphi_к = \frac{kT}{e}\cdot\ln\left(\frac{p_p\cdot n_n}{N_C\cdot N_V\cdot e^{-\frac{E_g}{kT}}}\right)=\frac{kT}{e}\left(\frac{E_g}{kT}\right)+\frac{kT}{e}\cdot\ln\left(\frac{p_p\cdot n_n}{N_C\cdot N_V}\right)=\frac{E_g}{e}-\frac{kT}{e}\ln\left(\frac{N_C\cdot N_V}{N_D\cdot N_A}\right), \qquad (3.75)$$

т.е. контактная разность потенциалов зависит в первую очередь от ширины запрещенной зоны: $\varphi_к \sim E_g$. Таким образом, для получения больших значений контактной разности потенциалов *p-n*-перехода необходимы полупроводники с большой шириной запрещенной зоны.

Анализируя формулы (3.74-3.75), можно определить минимальную и максимальную степени легирования *p-* и *n*-областей *p-n*-перехода.

1) Если принять, что концентрации доноров в *n*-области и акцепторов в *p*-области равны минимальным значениям $N_D = n_i$ и $N_A = p_i$, то из (3.74) следует

$$\varphi_к = \frac{kT}{e}\ln\frac{n_i^2}{n_i^2} = 0, \qquad (3.76)$$

т.е. контактная разность потенциалов будет равна нулю, и при таком низком уровне легирования *p-n*-переход не образуется.

2) Если же $N_D = N_C$ и $N_A = N_V$ (граница вырождения), то

$$\varphi_к = \frac{E_g}{e}-\frac{kT}{e}\ln\left(\frac{N_C\cdot N_V}{N_C\cdot N_V}\right)=\frac{E_g}{e}, \qquad (3.77)$$

т.е. уровень Ферми в *p*-области будет находиться на уровне потолка валентной зоны E_V, а уровень Ферми в *n*-области – на уровне дна зоны проводимости E_C, что соответствует началу вырождения в этих полупроводниках.

Таким образом, малые и умеренные концентрации примесей для получения *p-n*-перехода находятся в границах: $n_i < N_D < N_C$ и $p_i < N_A < N_V$.

Ширина p-n-перехода. Барьерная емкость. Электронно-дырочный переход можно представить состоящим из электронейтральных *p-* и *n*-областей (называемых базами) с переходной *областью объемного заряда* (ООЗ) между ними. Шириной *p-n*-перехода обычно называют ширину ООЗ $d = x_n + x_p$ (рис. 3.22). Изменяя значение ширины *p-n*-перехода, можно также изменять и его электрические свойства. Определим, от каких физических свойств полупроводникового материала зависит этот параметр.

Распределение потенциала в переходной области определяется из решения уравнения Пуассона, связывающего потенциал поля с объемной плотностью зарядов, создающих это поле. Решив его, можно найти, что контактная разность потенциалов равна:

$$\varphi_к = \frac{eN_AN_Dd^2}{2\varepsilon\varepsilon_0(N_A+N_D)}, \qquad (3.78)$$

где e – заряд электрона; ε – относительная диэлектрическая проницаемость; ε_0 – электрическая константа; N_A – концентрация акцепторов; N_D –

концентрация доноров; d – ширина ООЗ. Из (3.78) можно найти ширину ООЗ:

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0(N_A + N_D)\varphi_к}{eN_AN_D}} \ . \qquad (3.79)$$

Если концентрация носителей в одной области значительно больше, чем в другой, например $N_A >> N_D$ (так называемый несимметричный p-n-переход), выражение (3.79) примет вид:

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_к}{eN_D}} \ . \qquad (3.80)$$

Из анализа формул (3.79) и (3.80) видно, что ширина области объемного заряда зависит от степени легирования: чем сильнее легирован полупроводник, тем тоньше ООЗ.

Из решения уравнения Пуассона также следует, что $\frac{N_A}{N_D} = \frac{x_n}{x_p}$. В несимметричном p-n-переходе $N_A >> N_D$, $x_n >> x_p$, тогда ширина ООЗ $d \approx x_n$, т.е. область объемного заряда в основном распространяется в полупроводник с меньшей концентрацией примеси.

При приложении к p- и n-областям внешнего напряжения все оно будет падать на ООЗ, так как эта область обеднена носителями заряда и ее сопротивление велико по сравнению с сопротивлением остальных областей. Следовательно, внешнее напряжение будет суммироваться с контактной разностью потенциалов, и изменять высоту потенциального барьера. Если к p-области присоединить плюс, а к n-области – минус от внешнего источника ЭДС, то знак внешней ЭДС будет противоположен знаку $\varphi_к$ и потенциальный барьер уменьшится (рис. 3.23, *а*). Такая полярность внешнего напряжения называется *прямой*. Тогда из (3.79) для симметричного прямо смещенного p-n-перехода и (3.80) для несимметричного прямо смещенного p-n-перехода можно соответственно записать выражения для ширины ООЗ:

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0(N_A + N_D)(\varphi_к - U)}{eN_AN_D}}; \qquad (3.81)$$

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_к - U)}{eN_D}} \ . \qquad (3.82)$$

С увеличением внешнего прямого напряжения ширина p-n-перехода уменьшается (рис. 3.23, *а*).

При *обратной* полярности напряжения высота потенциального барьера растет, а ширина p-n-перехода увеличивается (рис. 3.24, *а*). Как уже отмечалось, в несимметричном p-n-переходе это увеличение происходит в основном в сторону области с меньшей концентрацией носителей.

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0(N_A + N_D)(\varphi_к + U)}{eN_AN_D}}; \qquad (3.83)$$

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_к + U)}{eN_D}}. \qquad (3.84)$$

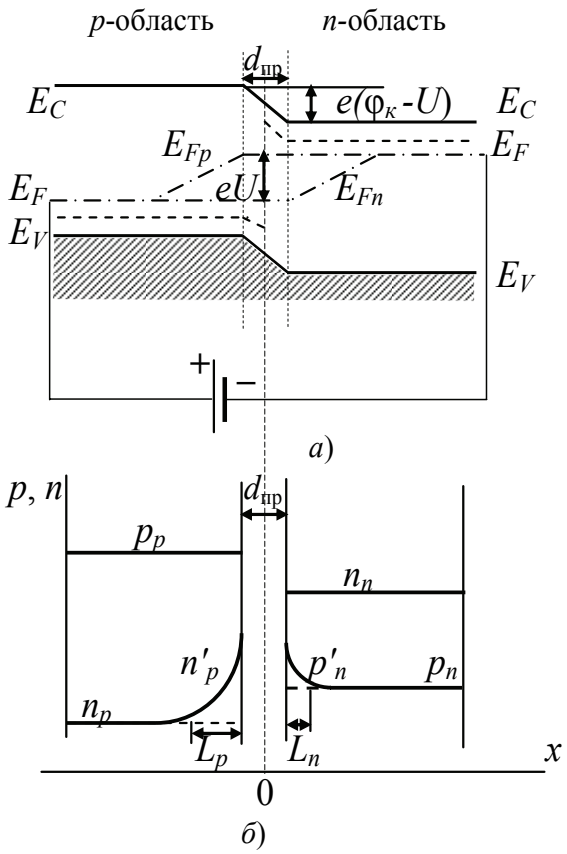


Рис. 3.23 – Зонная диаграмма (*а*) и распределение носителей (*б*) в прямо смещенном *p-n*-переходе

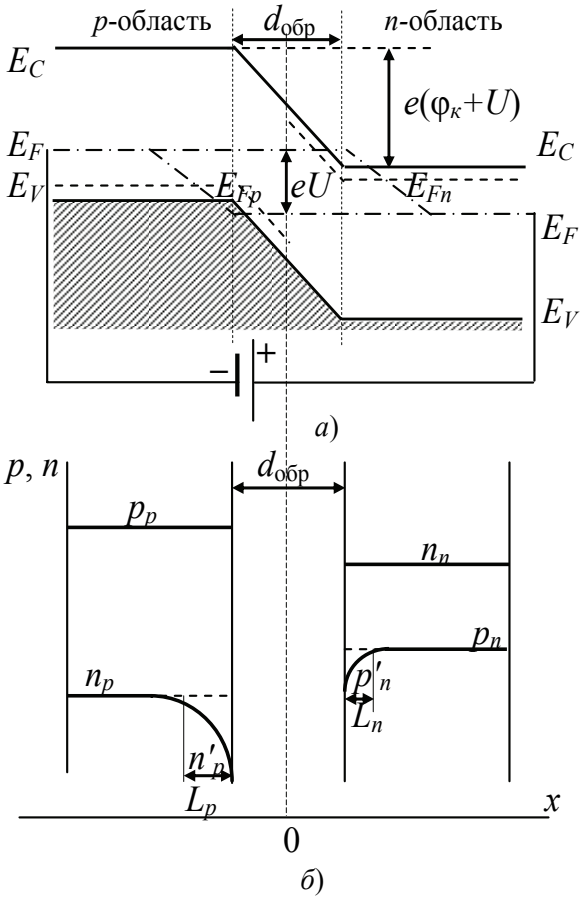


Рис. 3.24 – Зонная диаграмма (*а*) и распределение носителей (*б*) в обратно смещенном *p-n*-переходе

С увеличением ширины ООЗ увеличиваются и нескомпенсированные объемные заряды *p*- и *n*-областей. Такое изменение размеров и величины объемного заряда эквивалентно «поведению» емкости, т. е. *p-n*-переход ведет себя как конденсатор. Емкость *p-n*-перехода называется барьерной, и ее величина может быть определена по формуле емкости для плоского конденсатора $C_б = S(\varepsilon\varepsilon_0/d)$. Подставив в эту формулу значения ширины ООЗ, получим формулы для расчета барьерной емкости симметричного *p-n*-перехода (с концентрациями примесей в *p*- и *n*-областях одного порядка) и несимметричного *p-n*-перехода:

$$C_б = S \cdot \sqrt{\frac{\varepsilon_0\varepsilon}{2} \cdot \frac{eN_DN_A}{(N_A + N_D)\varphi_к}}; \qquad (3.85)$$

$$C_б = S \cdot \sqrt{\frac{\varepsilon_0\varepsilon}{2} \cdot \frac{qN_D}{\varphi_к}}. \qquad (3.86)$$

Для прямо смещенного *p*-*n*-перехода барьерная емкость будет расти, так как толщина ООЗ уменьшается. Барьерную емкость для симметричного и несимметричного прямо смещенных *p*-*n*-переходов найдем по формулам:

$$C_{б} = S \cdot \sqrt{\frac{\varepsilon_0 \varepsilon}{2} \cdot \frac{eN_D N_A}{(N_A + N_D)(\varphi_к - U)}}; \qquad (3.87)$$

$$C_{б} = S \cdot \sqrt{\frac{\varepsilon_0 \varepsilon}{2} \cdot \frac{qN_D}{(\varphi_к - U)}}. \qquad (3.88)$$

Для обратно смещенного *p*-*n*-перехода барьерная емкость будет уменьшаться, так как толщина ООЗ увеличивается. Барьерную емкость для симметричного и несимметричного обратно смещенных *p*-*n*-переходов найдем по формулам:

$$C_{б} = S \cdot \sqrt{\frac{\varepsilon_0 \varepsilon}{2} \cdot \frac{eN_D N_A}{(N_A + N_D)(\varphi_к + U)}}; \qquad (3.89)$$

$$C_{б} = S \cdot \sqrt{\frac{\varepsilon_0 \varepsilon}{2} \cdot \frac{qN_D}{(\varphi_к + U)}}. \qquad (3.90)$$

Как видно из этих выражений барьерная емкость *p*-*n*-перехода растет с увеличением концентрации примесей. Формулы (3.87-3.90) часто используется для определения концентрации примеси N_D или N_A по экспериментальной зависимости $1/C^2 = f(U)$, которая представляет собой прямую линию. Экстраполируя эту зависимость на ось напряжений до $I = 0$, можно экспериментально определить контактную разность потенциалов $\varphi_к$, так как при этом $1/C^2 = 0$.

Инжекция и экстракция неосновных носителей заряда. При включении *p*-*n*-перехода в прямом направлении потенциальный барьер уменьшается, следовательно, в *p*-область войдет добавочное количество электронов, а в *n*-область – дырок. Такое введение носителей заряда в область, где они являются неосновными, называется *инжекцией*. Т.е., можно сказать, что инжекция – это увеличение концентрации неосновных носителей заряда в области полупроводника, прилегающего к *p*-*n*-переходу, при подаче на него внешнего напряжения прямой полярности.

До появления инжектированных носителей (например, дырок в *n*-области) *n*-область была нейтральной, т. е. положительные и отрицательные заряды в каждом из достаточно малых объемов *n*-области в сумме были равны нулю. Дырки, инжектированные из *p*- в *n*-область, представляют собой положительный объемный заряд. Этот заряд создает электрическое поле, притягивающее к дыркам электроны, отрицательный заряд которых должен скомпенсировать положительный объемный заряд дырок. Такое перемещение электронов создает их недостаток в тех областях, откуда они ушли. Для восстановления электрической нейтральности

полупроводника из контакта к *n*-области должно войти дополнительное количество электронов. При инжекции неосновных носителей заряда в однородный полупроводник его электронейтральность практически сохраняется (за исключением граничных областей) и объемные заряды не возникают. Это условие принято называть условием *квазинейтральности*.

Так как заряды электронов и дырок равны, то количество вошедших электронов равно количеству инжектированных дырок. Таким образом, одновременно с появлением некоторого количества дырок – неосновных носителей заряда в *n*-области – появляется равное количество электронов – основных носителей заряда в этой области. И те, и другие являются *неравновесными*, так как создают дополнительную концентрацию, отличающуюся от концентрации в термодинамическом равновесии. Так как концентрация неосновных носителей мала, то дополнительное введение определенной концентрации неравновесных носителей существенно увеличит концентрацию именно неосновных носителей.

Инжектированные в *p*-область электроны и в *n*-область дырки будут диффундировать вглубь этих областей. Так как инжектированные носители неравновесны, то при диффузии концентрации инжектированных носителей будут экспоненциально уменьшаться в противоположные стороны от *p*-*n*-перехода за счет рекомбинации с основными носителями данной области (рис. 3.23, *б*). Диффузионная длина носителей L_n (L_p) – это расстояние, на котором концентрация носителей в стационарных условиях уменьшается в *e* раз.

При обратном смещении *p*-*n*-перехода концентрация неосновных носителей на границе *p*-*n*-перехода и полупроводника будет меньше их равновесной концентрации в глубине (*экстракция*) (рис. 24, *б*) и при напряжении $|U| \geq 4kT$ практически равна нулю.

Вольтамперная характеристика и способы управления ею.

Вольтамперная характеристика связывает ток (плотность тока) через *p*-*n*-переход и напряжение на нем. Найдем вольтамперную характеристику *p*-*n*-перехода из следующих соображений.

При прямом смещении высота потенциального барьера уменьшается на *eU* (рис. 23, *а*), где U – приложенное напряжение. В этих неравновесных условиях распределение электронов и дырок описывается с помощью двух квазиуровней: E_{F_n} и E_{F_p} отдельно для электронов и дырок. В области расщепления уровня Ферми на два квазиуровня наблюдается неравновесное состояние неосновных носителей. В областях постоянства уровней Ферми для соответствующих носителей существует равновесное распределение носителей по энергиям:

$$n = n_{n0} e^{-\frac{E-E_C}{kT}} \quad \text{и} \quad p = p_{p0} e^{-\frac{E_V-E}{kT}}, \qquad (3.91)$$

где n_{n0} и p_{p0} – равновесные концентрации основных носителей у краев соответствующих зон (рис. 23, *б*; рис. 24, *б*). В частности, отсюда можно найти равновесные концентрации неосновных носителей в отсутствие смещения:

$$n_{p0} = n_{n0} \cdot e^{-\frac{e\varphi_к}{kT}} \quad \text{и} \quad p_{n0} = p_{p0} \cdot e^{-\frac{e\varphi_к}{kT}}, \tag{3.92}$$

где $e\varphi_к$ – высота потенциального барьера в несмещенном *p*-*n*-переходе.

При наличии внешнего прямого смещения высота потенциального барьера становится равной

$$\varphi = \varphi_к - U, \tag{3.93}$$

и уровню E_C в начале *p*-области уже соответствует неравновесная концентрация электронов

$$n_p{}' = n_{n0} \cdot e^{-\frac{e(\varphi_к - U)}{kT}} = n_{p0} \cdot e^{\frac{eU}{kT}}, \tag{3.94}$$

уровню E_V в начале *n*-области – неравновесная концентрация дырок

$$p_n{}' = p_{p0} \cdot e^{-\frac{e(\varphi_к - U)}{kT}} = p_{n0} \cdot e^{\frac{eU}{kT}}. \tag{3.95}$$

Избыточные концентрации электронов и дырок в этих точках:

$$\Delta n_0 = n_p{}' - n_{p0} = n_{p0}\left(e^{\frac{eU}{kT}} - 1\right); \tag{3.96}$$

$$\Delta p_0 = p_n{}' - p_{n0} = p_{n0}\left(e^{\frac{eU}{kT}} - 1\right). \tag{3.97}$$

Избыточные неосновные носители диффундируют в толщу *n*- и *p*-областей, где вследствие рекомбинации устанавливаются их пространственные распределения вида:

$$\Delta n = \Delta n_0 \cdot e^{-\frac{x}{L_n}}, \tag{3.98}$$

$$\Delta p = \Delta p_0 \cdot e^{-\frac{x}{L_p}}, \tag{3.99}$$

где L_n и L_p – диффузионные длины электронов и дырок, соответственно. Координаты x отсчитываются для электронов и дырок как расстояние от области объемного заряда в толщу. При этом возникают диффузионные потоки неосновных неравновесных электронов и дырок в соответствии с законом Фика:

$$\Phi_n = -D_n \frac{d(\Delta n)}{dx} \text{ и } \Phi_p = -D_p \frac{d(\Delta p)}{dx}, \tag{3.100}$$

где D_n и D_p – диффузионные длины электронов и дырок, соответственно.

При $x = 0$

$$\frac{d(\Delta n)}{dx} = \frac{\Delta n}{L_n}, \quad \frac{d(\Delta p)}{dx} = \frac{\Delta p}{L_n}. \tag{3.101}$$

Разнонаправленным диффузионным потокам электронов и дырок соответствуют однонаправленные составляющие электрического тока:

$$j_n = eD_n \frac{\Delta n}{L_n}, \tag{3.102}$$

$$j_p = eD_p \frac{\Delta p}{L_p}. \tag{3.103}$$

Подставив значения Δn из (3.98) в (3.102) и Δp из (3.99) в (3.103) и сложив токи, получим:

$$j = e\left(D_n \frac{\Delta n}{L_n} + D_p \frac{\Delta p}{L_p} \right)\left(e^{\frac{eU}{kT}} - 1 \right), \tag{3.104}$$

где

$$j_s = e\left(D_n \frac{\Delta n}{L_n} + D_p \frac{\Delta p}{L_p} \right) \tag{3.105}$$

– плотность тока насыщения.

Тогда

$$j = j_s\left(e^{\frac{eU}{kT}} - 1 \right) \tag{3.106}$$

– компактная форма вольтамперной характеристики *p*-*n*-перехода (ВАХ). Учитывая, что $n_p = \frac{n_i^2}{p_p}$, $p_n = \frac{n_i^2}{n_n}$ (см. формулу (3.29)), и для мелких примесей $p_p = N_A$ и $n_n = N_D$, выражение ВАХ можно записать в виде:

$$j = e\left(\frac{D_n}{N_A \cdot L_n} + \frac{D_p}{N_D \cdot L_p} \right) n_i^2 \left(e^{\frac{eU}{kT}} - 1 \right), \tag{3.107}$$

или, с учетом того, что

$$D_n = \frac{kT}{e}\mu_n, \; D_p = \frac{kT}{e}\mu_p, \tag{3.108}$$

$$L_n = \sqrt{D_n \tau_n}, \; L_p = \sqrt{D_p \tau_p}, \tag{3.109}$$

получим ВАХ в виде:

$$j = \sqrt{ekT}\left(\frac{1}{N_A}\sqrt{\frac{\mu_n}{\tau_n}} + \frac{1}{N_D}\sqrt{\frac{\mu_p}{\tau_p}} \right) n_i^2 \left(e^{\frac{eU}{kT}} - 1 \right). \tag{3.110}$$

Обычно N_A, N_D, μ_n, μ_p, n_i известны точно, τ_n и τ_p – оценочно. Таким образом, форму ВАХ можно целенаправленно создавать для каждого конкретного прибора.

Проанализируем зависимость плотности тока насыщения от температуры в *p*-*n*-переходе. Для простоты, рассмотрим несимметричный *p*-*n*-переход ($p_P >> n_n$ и $p_n >> n_P$). Плотность тока насыщения j_s определяется

параметрами неосновных носителей в слаболегированной области *p*-*n*-перехода и равна $j_s = e(\frac{D_p}{L_p n_n})n_i^2$. Диффузионная длина и коэффициент диффузии от температуры зависят слабо, а так как все примеси при комнатной температуре ионизированы, то n_n не зависит от температуры. От температуры сильно зависит только концентрация собственных носителей $n_i \sim \exp\left(-\frac{E_g}{2kT}\right)$, т. е. чем больше температура, тем больше концентрация собственных носителей. Это значит, что и плотность тока насыщения ($j_s \sim n_i^2 \sim \exp\left(-\frac{E_g}{kT}\right)$) увеличивается с ростом температуры. *P*-*n*-переходы выполненные из полупроводников с бóльшей шириной запрещенной зоны имеют гораздо меньшие токи насыщения j_s.

При прямом смещении и напряжении $U > 3kT$ (~ 0,08 В) с погрешностью менее 5% единицей в (3.106) можно пренебречь и считать зависимость чисто экспоненциальной до $U < \varphi_к$. Тогда (3.106) будет иметь вид $j = j_s \cdot e^{\frac{eU}{kT}}$. При обратном смещении и напряжении $U > 3kT$ обратный ток мал и практически не зависит от напряжения $j = -j_s$. Это происходит потому, что ток при обратном смещении переносится неосновными носителями, т.е. дырками из *n*-области в *p*-область и электронами из *p*- в *n*-область. Уже при небольших обратных напряжениях практически все термически генерированные на расстоянии диффузионной длины от *p*-*n*-перехода неосновные носители экстрагируются полем *p*-*n*-перехода, и дальнейшее увеличение напряжения не приводит к росту тока. Вид вольтамперной характеристики *p*-*n*-перехода показан на рис. 3.25 (излом в точке 0 является следствием разного масштаба при прямом и обратном смещениях).

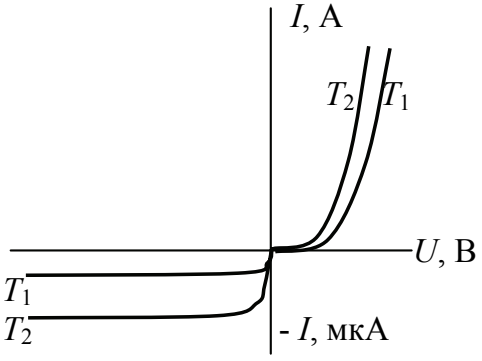


Рис. 3.25 – ВАХ *p*-*n*-перехода при различных температурах: $T_1 < T_2$

В связи с тем, что в прямом направлении ток через *p*-*n*-переход во много раз больше, чем в обратном, контакты *p*- и *n*-полупроводников называют *выпрямляющими*, т. е. пропускающими ток в одном направлении.

Диоды на основе *p*-*n*-переходов способны выполнять в схемах различные функции: выпрямления переменного тока, управляемого реактивного сопротивления, детектирования сигналов, модуляции, генерирования СВЧ колебаний и др. Функционирование диодов основано на ряде физических явлений, протекающих в полупроводниках: нелинейности

ВАХ *p-n*-перехода, пробое *p-n*-перехода, нелинейности реактивного сопротивления *p-n*-перехода, фоточувствительности и т.д.

Рассмотрим выпрямительные или силовые диоды – приборы, предназначенные для выпрямления тока. Физические параметры этих диодов (контактная разность потенциалов, ширина обедненного слоя, барьерная емкость и др.) зависят от свойств полупроводниковых материалов (ширины запрещенной зоны, степени легирования примесями, подвижности носителей и др.).

По типу применяемого материала выпрямительные диоды подразделяют на германиевые и кремниевые. Наибольшее распространение получили кремниевые диоды, как в дискретном, так и в интегральном исполнении. В кремниевых выпрямительных диодах обратные токи на несколько порядков меньше ($j_s \sim \exp\left(-\frac{E_g}{kT}\right)$), а допустимые обратные напряжения U существенно выше, чем в германиевых (у кремниевых диодов

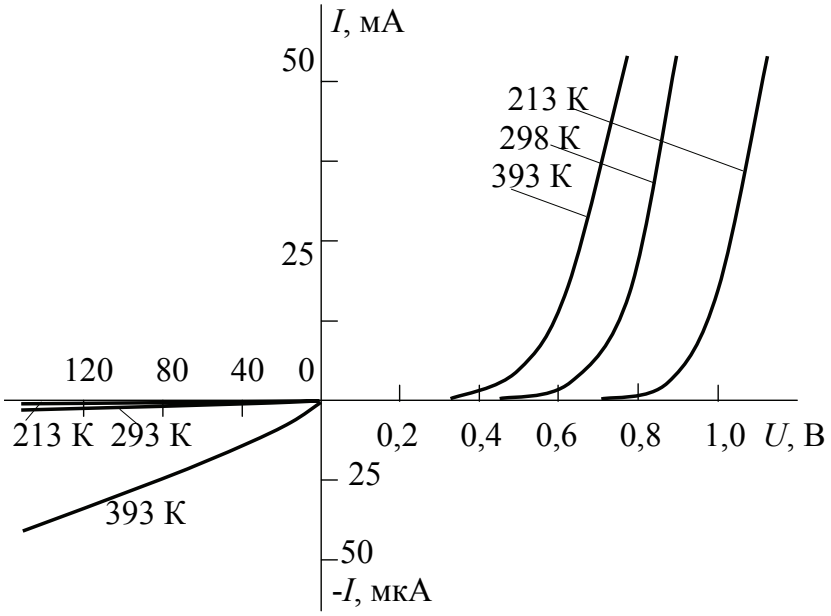


Рис. 3.26 – ВАХ кремниевого диода Д223 Б при различных температурах окружающей среды

до 2000 В, у германиевых – до 400 В). Следует отметить, что на обратной ветви ВАХ кремниевого диода отсутствует участок явного насыщения обратного тока. Это обусловлено более интенсивной генерацией носителей в области объемного заряда, которая расширяется с увеличением обратного напряжения (рис. 3.26).

Рабочие диапазоны температур кремниевых диодов 213 К … 398 К, германиевых – 213 К … 358 К. Невысокий верхний предел рабочей температуры германия обусловлен тем, что при $T > 358$ К тепловая генерация собственных носителей заряда становится настолько высокой, что

происходит резкое возрастание обратного тока и эффект выпрямления практически исчезает (рис. 3.27). Несмотря на перечисленные преимущества кремниевых диодов, германиевые диоды целесообразно применять в выпрямительных устройствах низкого напряжения. У германиевого диода прямая ветвь ВАХ значительно круче, чем у кремниевых, в результате при одинаковом токе нагрузки значение остаточного сопротивления германиевых диодов в прямом смещении будет в полтора-два раза меньше. Следовательно, и потери мощности, рассеиваемой в германиевом диоде, будут меньше, чем в кремниевом диоде. Таким образом, выбор полупроводникового материала, применяемого в диодах, зависит от поставленных технических целей: в одних случаях необходимы высокие допустимые обратные напряжения, в

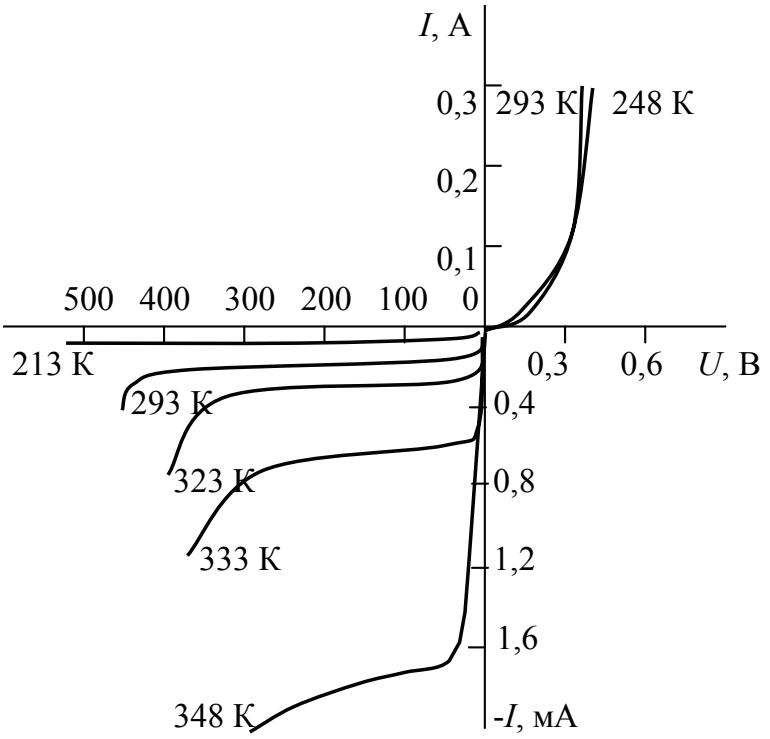


Рис. 3.27 – ВАХ германиевого диода при различных температурах окружающей среды

других – малые потери мощности и т.д.

Практический интерес к такому полупроводниковому материалу как карбид кремния, связан с разработкой мощных диодов, варикапов, способных стабильно работать при высоких температурах.

Селен также используют для изготовления выпрямителей. Усовершенствование технологии производства этого материала будет, вероятно, способствовать появлению новых приборов микроэлектроники.

Электронно-дырочные переходы получают чаще всего в процессе сплавления, диффузии и эпитаксии. Эпитаксиальные *p*-*n*-переходы наиболее совершенны.

Ниже покажем, как изменением параметров полупроводника *p*-*n*-перехода, можно получить диоды, работающие на принципиально других физических основах и обладающие особыми свойствами.

3.3.2.2. Туннельные диоды

До сих пор мы рассматривали легированные полупроводники с малым и средним содержанием примесных атомов: в типичном случае их относительная концентрация составляет менее одной миллионной доли. Мы связали тип примесной проводимости и концентрацию примеси с положением уровня Ферми, показав, что во всех случаях он располагается в

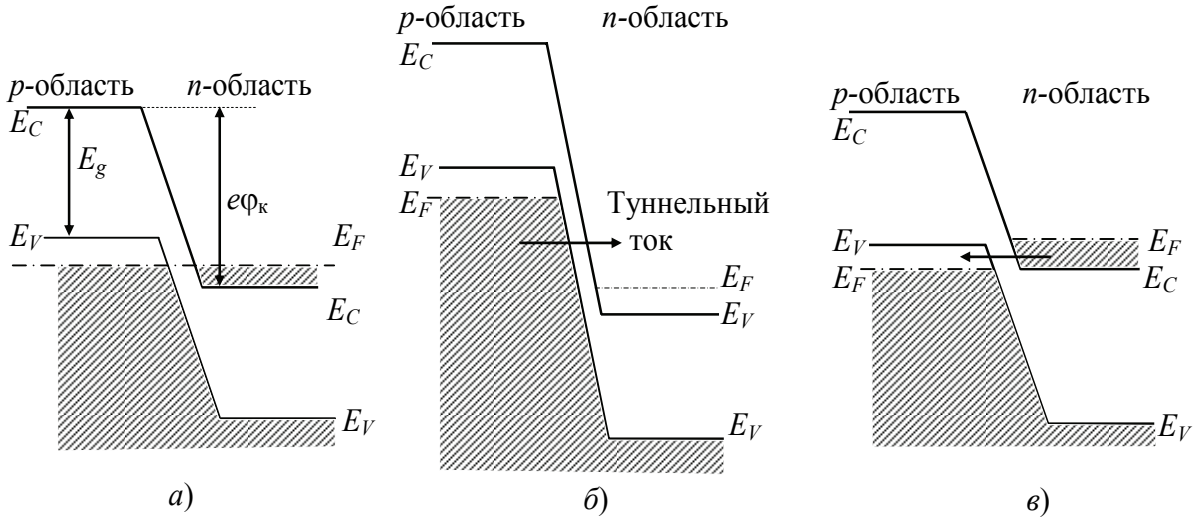


Рисунок 3.28 – Энергетическая диаграмма туннельного диода: *а*) – тепловое равновесие; *б*) – обратное смещение; *в*) – прямое смещение

пределах запрещенной зоны. Однако при высокой концентрации легирующей примеси (типичное значение составляет 10^{24} м$^{-3}$, или 0,01%) уровень Ферми будет располагаться в зоне проводимости (см. формулу (3.77)), если полупроводник *n*-типа или в валентной зоне, если полупроводник *p*-типа. В этих случаях говорят, что полупроводник становится «вырожденным». Туннельный диод является единственным рассматриваемым нами прибором, в котором как *p*-, так и *n*-стороны перехода выполнены из вырожденных полупроводников. На рис. 3.28, *а* представлена энергетическая зонная диаграмма несмещенного туннельного диода. Значение разности между уровнем Ферми E_{F_p} и потолком валентной зоны E_V, и уровнем Ферми E_{F_n} и дном зоны проводимости E_C для простоты выбраны одинаковыми с обеих сторон. Интересно отметить, что высота потенциального барьера превышает ширину запрещенной зоны. Поэтому число электронов, проходящих над потенциальным барьером в условиях теплового равновесия, должно быть незначительным. Из-за этого значение надбарьерного прямого тока I оказывается малым, что обусловливает протяженный вид характеристики выпрямления (рис. 3.29, *а*).

Однако в диодах такого типа существует и другой механизм, ведущий к появлению потока электронов. Так как туннельные диоды изготовлены из сильно легированных материалов, то большая концентрация примесей приводит, в свою очередь, к малой толщине переходной области (порядка 10^{-8} м). Поэтому электроны могут не только перемещаться поверх

потенциального барьера, но и туннелировать сквозь этот узкий барьер. Количественный анализ показывает, что доминирует именно туннельный ток. Следовательно, мы можем рассматривать тепловое равновесие как такое состояние, когда равные по величине токи туннелирования протекают в противоположных направлениях.

Если приложить обратное напряжение (рис. 3.28, *б*), то число электронов, туннелирующих через *p*-*n*-переход из *p*-области в *n*-область, возрастает, так как электронам в *p*-области соответствует большое число свободных состояний в *n*-области. При этом обратный туннельный ток резко возрастает с напряжением и обратная ветвь ВАХ идет круто вниз, т.е. туннельный диод, в отличие от выпрямительного, обладает высокой проводимостью при обратном включении.

При прямом смещении электроны начнут туннелировать из *n*-области в *p*-область (рис. 3.28, *в*). Если увеличивать приложенное напряжение, то ООЗ сужается, и начинает расти число доступных состояний в *p*-области, а это в свою очередь приведет к росту тока. Максимум тока получается в том случае, когда электроны из *n*-области достигают всех вакантных состояний в *p*-области, т.е. если уровень Ферми в *n*-области совпадает с потолком валентной зоны в *p*-области. Если еще увеличить напряжение смещения, то это приведет к (дальнейшему) росту числа электронов, стремящихся туннелировать в запрещенную зону. Однако за счет туннельного эффекта

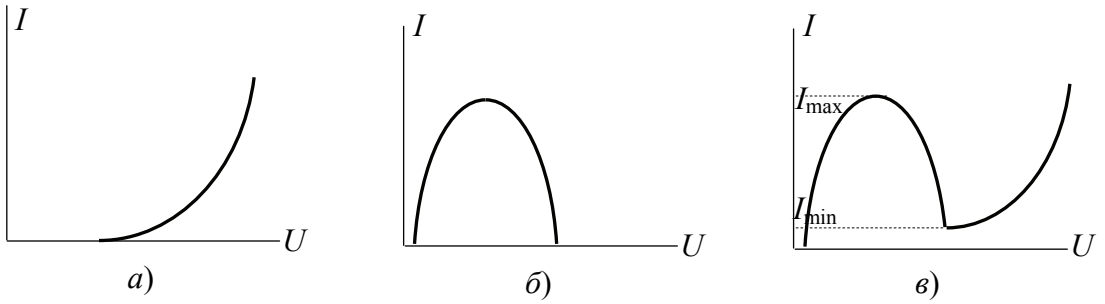


Рис. 3.29 – Ток в туннельном диоде: *а*) – прямой ток; *б*) – туннельный ток; *в*) – полный ток

эти электроны переместиться не смогут, так как там нет свободных энергетических уровней. В результате ток будет уменьшаться и достигнет минимума, когда потолок валентной зоны *p*-области совместится по энергии с дном зоны проводимости *n*-области. Вольтамперная характеристика будет иметь вид, показанный на рис. 3.29, *б*. Однако здесь отражен не весь ток, а лишь та его часть, которая связана с туннельным эффектом. Чтобы найти весь ток, нужно просто алгебраически просуммировать токи, представленные на рис. 3.29, *а* и *б* – туннельный и надбарьерный. Складывая эти токи, мы получаем вольтамперную характеристику (3.29, *в*), которую можно наблюдать в опытах.

В действительности эта характеристика получается несколько иной из-за неизбежного присутствия некоторых энергетических уровней в запрещенной

зоне («хвоста плотности состояний»). По этой причине возникает дополнительное туннелирование, особенно заметное в окрестности минимума тока. Но даже с учетом этого эффекта на практике ток в максимуме I_{max} отличается от тока в минимуме I_{min} в несколько раз. На рисунке 3.30 приведена вольтамперная характеристика туннельного диода. Приведены также ВАХ туннельных диодов на основе различных полупроводниковых материалов при температуре 300 К (рис. 3.31). Величина отношения тока проводимости I к максимальному току I_{max} (току в пике ВАХ) равна 8 для Ge, 4 для Si и 12 для GaAs. Отношение I/I_{max} является одним из важных параметров туннельного диода. В общем случае эта величина для данного полупроводника возрастает с уровнем легирования как n-, так и p-областей. Предельное значение этого соотношения зависит,

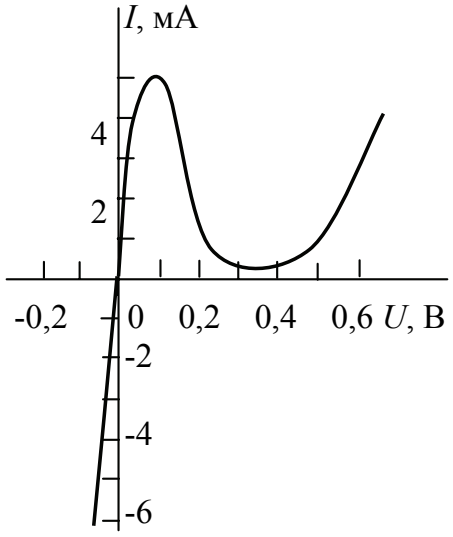


Рис. 3.30 – ВАХ туннельного диода

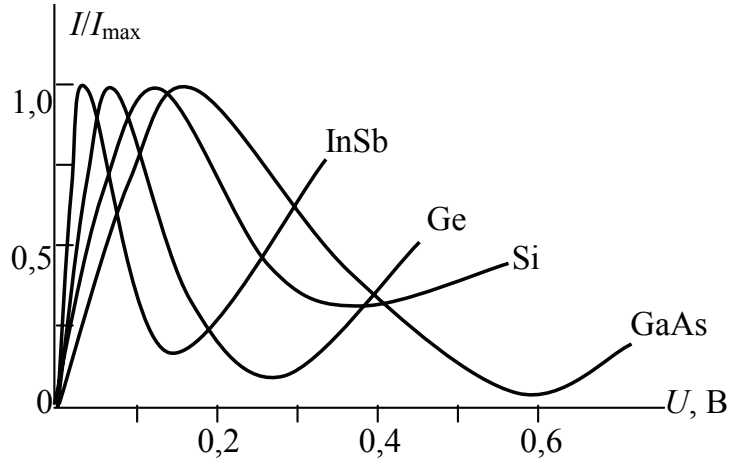


Рис. 3.31 – ВАХ туннельных диодов на основе различных полупроводниковых материалов при 300 К

во-первых, от пикового тока, который определяется шириной запрещенной зоны и эффективной массой туннелирующих носителей, и, во-вторых, от минимального тока, который определяется распределением и концентрацией локализованных состояний в запрещенной зоне.

Итак, туннельные диоды имеют ВАХ N–типа с участком *отрицательной проводимости*, на котором с ростом прямого смещения прямой ток уменьшается. На участке с отрицательной проводимостью дифференциальная проводимость равна

$$G = -\frac{dI}{dU}. \qquad (3.111)$$

Особое достоинство туннельных диодов состоит в том, что у них очень тонкие области перехода, вследствие этого туннельные диоды способны работать на сверхвысоких частотах (вплоть до частот порядка 100 ГГц). В туннельных диодах, вследствие малой инерционности, отрицательная проводимость практически не зависит от частоты вплоть до миллиметрового диапазона. Наличие отрицательной проводимости и резкая нелинейность ВАХ туннельных диодов позволяет использовать их в усилителях,

генераторах и смесителях с усилением в диапазоне волн вплоть до миллиметровых. На туннельных диодах строятся схемы триггеров, мультивибраторов, переключателей с очень малыми временами переключения. Так как туннельные диоды изготавливают на основе сильнолегированных полупроводниковых материалов, их параметры, в первую очередь удельное сопротивление, слабо зависят от температуры (рис. 3.11, участок 9-8-4). Поэтому туннельные диоды могут работать в широком диапазоне температур – от близких к абсолютному нулю до нескольких сотен градусов Цельсия.

Если подобрать соответствующие концентрации легирующих примесей, то можно получить новый тип диода, который называются *обращенным*. В обращенном диоде значение сопротивления при обратном смещении низкое, а при прямом смещении, наоборот, высокое. Уровень легирования должен быть снижен по сравнению с легированием туннельного диода ровно настолько, чтобы в отсутствие смещения потолок валентной зоны E_V в *p*-области и дно зоны проводимости E_C в *n*-области расположились на одном уровне. Из-за этого при прямом смещении ток оказывается очень малым, так как эффект туннелирования практически не наблюдается. Однако при обратном смещении в таком приборе может возникать значительный туннельный ток. Обращенные диоды являются разновидностью туннельных диодов и используются в радиотехнических устройствах как детекторы при малых уровнях сигналов.

3.3.2.4. Гетеропереходы

Все приборы, речь о которых шла раньше, выполнялись из одного и того же полупроводникового материала. Некоторые свойства этого материала

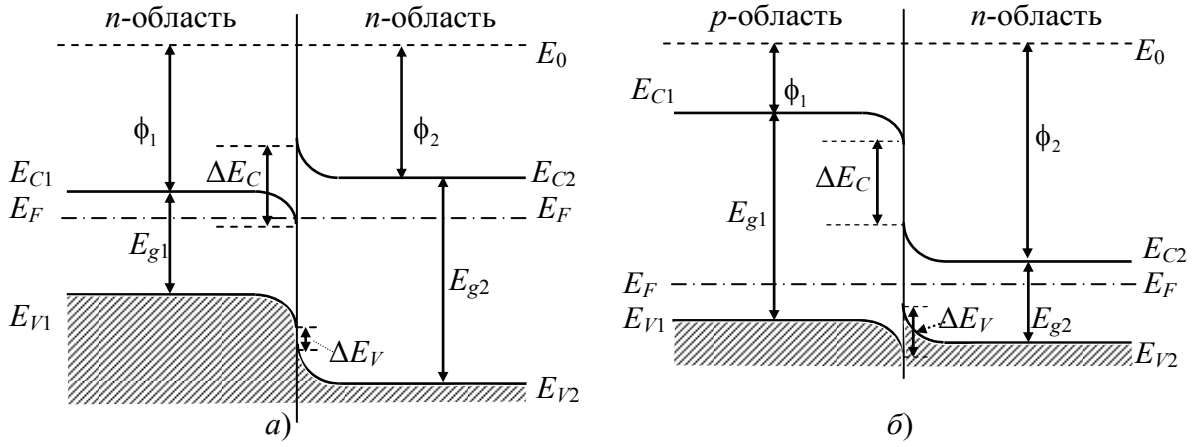


Рис. 3.32 – Зонная структура кристалла с гетеропереходом: *а* – переход типа *n-n* (изотипный переход); *б* – переход типа *p-n* (анизотипный переход)

модифицировались, например, за счет легирования: одна часть образца превращалась в полупроводник *n*-типа, а другая – в полупроводник *p*-типа. Ширина запрещенной зоны при этом оставалась неизменной. Такие

структуры называются *гомоструктурами* в отличие от *гетероструктур*, которые образованы полупроводниками с разными значениями ширины запрещенной зоны. Итак, гетеропереходом называется переход, образованный двумя различными полупроводниками. Если такие полупроводники имеют одинаковый тип проводимости, то они образуют *изотипный* гетеропереход (рис. 3.32, *а*). Если тип их проводимости различен, то получается *анизотипный* гетеропереход (рис. 3.32, *б*). Главную особенность гетероперехода, в отличие от обычного *p*-*n*-перехода (его еще называют *гомопереходом*), нетрудно понять из энергетических диаграмм, показанных на рис. 32 *а*, *б*.

Работа выхода χ определяется для полупроводника, как и для металла, величиной энергии, требуемой для перевода электрона, находящегося на уровне Ферми, через поверхность, ограничивающую объем, на уровень энергии вакуума за пределами материала. Поскольку работа выхода зависит от положения уровня Ферми, она меняется при легировании. Удобнее пользоваться такой физической величиной, как *электронное сродство* ϕ. Электронное сродство ϕ – это энергия, необходимая для перевода электрона со дна зоны проводимости на «уровень вакуума». Эта величина представляет собой свойство материала, не зависящее от легирования.

В гомопереходе работы выхода электрона для обеих его частей одинаковы, а в гетеропереходе – различны. Поэтому, откладывая от уровня электрона в вакууме (E_0 на рис. 3.32) вниз значения электронного сродства ϕ_1 и ϕ_2, получаем значения E_{C1} и E_{C2} и видим, что на границе раздела ($x = 0$) в гетеропереходе образуется разрыв в зоне проводимости. «Достраивая» схему до валентных зон, нетрудно убедиться в образовании разрыва и в валентной зоне. В равновесии, когда уровни Ферми выравниваются, эти разрывы сохраняются, а энергетическая диаграмма приобретает вид, показанный на рис. 3.32, *а*. Характер разрывов потолка валентной зоны E_V и дна зоны проводимости E_C различен. Первый часто называют разрывом вида «крюк», а второй – вида «стенка».

В случае анизотипного гетероперехода, например *p*-*n*-гетероперехода (рис. 3.32, *б*), разрыв «стенка» возникает в зоне проводимости, а «крюк» – в валентной зоне.

Одним из граничных условий является непрерывность электрической индукции на границе раздела, т.е. $\varepsilon_1 E_1 = \varepsilon_2 E_2$. Это условие означает, что электрическое поле, обусловленное наклоном зон на границе раздела, будет претерпевать разрыв, если имеется разница в величинах диэлектрических проницаемостей.

Таким образом, в гетеропереходах в отличие от *p*-*n*-переходов потенциальные барьеры для электронов и дырок могут значительно отличаться, что существенным образом сказывается на их инжекционных свойствах. Приложенное смещение в пропускном направлении к *p*-*n* (*n*-*p*)-гетеропереходу будет обеспечивать одностороннюю эффективную инжекцию из материала с бóльшей шириной запрещенной зоны. Так, для

p-*n*-гетероперехода (*p*-область широкозонная) наличие разрыва ΔE_C будет препятствовать инжекции электронов в широкозонный полупроводник. Наличие разрыва в валентной зоне ΔE_V уменьшает барьер для инжектируемых дырок.

Токи инжектированных электронов и дырок в первом приближении отличаются друг от друга на величину $\exp\left(\frac{\Delta E_C + \Delta E_V}{kT}\right)$. Если учесть, что kT при 300 К порядка 0,026 эВ, а (ΔE_C + ΔE_V) может составлять величины в несколько десятых электронвольта, то нетрудно убедиться, что в таких гетеропереходах режим односторонней инжекции легко достижим.

Ширину ООЗ в каждом полупроводнике можно найти, решив уравнение Пуассона для резкого перехода с каждой стороны границы раздела. Затем можно определить барьерную емкость, которая равна

$$C = \sqrt{\frac{eN_{D1}N_{A2}\varepsilon_1\varepsilon_2}{2(\varepsilon_1 N_{D1} + \varepsilon_2 N_{A2})(\varphi_к - U)}}, \qquad (3.112)$$

где N_{D1} – концентрация донорной примеси; N_{A2} – концентрация акцепторной примеси; U – приложенное прямое напряжение ($U = U_1 + U_2$); $\varphi_к$ – полный контактный потенциал ($\varphi_к = \varphi_{к1} + \varphi_{к2}$); $\varphi_{к1}$ и $\varphi_{к2}$ – электростатические потенциалы равновесного состояния первого и второго полупроводника.

Отношение напряжений на каждом полупроводнике составляет

$$\frac{\varphi_{к1} - U_1}{\varphi_{к2} - U_2} = \frac{N_{A2}\varepsilon_2}{N_{D1}\varepsilon_1}, \qquad (3.113)$$

Очевидно, что полученные выражения преобразуются к выражению для обычного гомогенного *p*-*n*-перехода, если полупроводники по обе стороны гетероперехода одинаковы.

Вольтамперная характеристика гетероперехода имеет вид:

$$I = I_s\left(1 - \frac{U}{\varphi_к}\right)\left[\exp\left(\frac{eU}{kT}\right) - 1\right], \qquad (3.114)$$

где I_s – ток насыщения, равный

$$I_s = \frac{eA^*T\varphi_к}{k}\exp\left(-\frac{e\varphi_к}{kT}\right), \qquad (3.115)$$

где A^* – постоянная Ричардсона.

Обратный ток гетероперехода, так же как и гомоперехода, зависит от температуры, но не имеет насыщения, а при больших напряжениях линейно возрастает с напряжением. В прямом направлении зависимость тока от напряжения допускает аппроксимацию экспоненциальной функцией, т.е. $I \sim \exp\left(\frac{eU}{nkT}\right)$, где n – некий коэффициент порядка единиц.

Следует отметить, что на внутренней границе раздела в гетеропереходе могут возникать поверхностные состояния, искажающие идеальные диаграммы, показанные на рис. 3.32. В этом случае переход носителей заряда

через гетерограницу может идти через поверхностные уровни, причем из зон на поверхностные уровни носители могут попасть при помощи туннельного эффекта. Рекомбинация носителей заряда на поверхностных уровнях ухудшает запирающие свойства гетероперехода, затрудняет осуществление большого уровня инжекции и т.д.

Поверхностные состояния могут иметь различную природу. Одна из них та же, что приводила к образованию таммовских уровней (см. гл. I), но чаще поверхностные уровни появляются за счет так называемых дислокаций несоответствия, которые возникают из-за различий в периодах кристаллической решетки и термических коэффициентов линейного расширения полупроводников, составляющих гетеропереход.

Поэтому изготовление качественных гетеропереходов является сложной технологической задачей. Выбор идеальной пары весьма непрост. Необходимо выполнение многих условий «совместимости» по: механическим, кристаллохимическим и термическим свойствам; кристаллической и электронной структуре контактирующих веществ. Важнейшим параметром совместимости является близость постоянных решеток сращиваемых кристаллов. Достаточно им отличаться на 1…2% и свойства перехода уже не зависят от объемных параметров, а полностью определяются дефектами, дислокациями и созданными ими заряженными центрами на границе раздела. Если рассогласование решетки составляет несколько процентов, то плотность состояний на границе раздела составляет порядка 10^{18} $м^{-2}$, и свойства гетероперехода могут полностью определяться этими состояниями. Это подтверждается измерениями на переходах Si-Ge, где были обнаружены барьеры, величину которых нельзя объяснить различиями в величинах электронного сродства ϕ.

Поскольку процесс получения гетеропереходов происходит при сравнительно высоких температурах, требуется совпадение коэффициентов теплового расширения контактирующих веществ в широком интервале температур. Если имеются большие различия в величинах коэффициентов расширения, то это может привести к заметному разупорядочению вследствие напряжений, возникающих на границе раздела при охлаждении.

Многие полупроводники, перспективные с точки зрения получения гетеропереходов, обладают структурой алмаза. Например, пара GaAs и Ge могут образовывать потенциально интересные гетеропереходы, так как при 300 К постоянные решетки у этих веществ почти идентичны (с точностью 0,07%), и коэффициенты линейного расширения при 300 К очень близки ($5{,}8\cdot10^{-6}$ $°C^{-1}$ и $5{,}7\cdot10^{-6}$ $°C^{-1}$ соответственно). Другими примерами гетеропереходов на основе:

полупроводник $A^{III}B^{V}$ – простой полупроводник является гетеропереход GaP-Si;

полупроводник $A^{II}B^{VI}$ – простой полупроводник являются гетеропереходы CdS-Si, CdS-Ge, ZnSe-Ge и др.

Возможно получение гетеропереходов между гексагональными и кубическими модификациями различных веществ, например, ZnTe-GaSb.

Тем не менее, некоторые методы получения гетеропереходов дают результаты, определяемые скорее технологическим процессом, чем фундаментальными свойствами полупроводников. К примеру, в гетеропаре GaAs-Ge неизбежны эффекты перекрестного легирования, поскольку Ga и As являются легирующими примесями для Ge, а Ge представляет собой легирующую примесь для GaAs. А вот для пары $Ge_{0,9}Si_{0,1}$-Ge таких проблем не возникает.

Иногда возникают проблемы совершенно другого свойства. Один взгляд на табл. 3.2 свойств полупроводниковых соединений $A^{III}B^{V}$ может определить «идеальную пару». Это арсенид галлия GaAs и арсенид алюминия AlAs. Арсениды алюминия и галлия имеют одинаковую постоянную решетки (разница в третьем знаке: 5,653 Å для GaAs и 5,661 Å для AlAs, коэффициент рассогласования 0,11%) при значительной разнице в ширине запрещенной зоны: 2,16 эВ у AlAs и 1,43 у GaAs. Коэффициенты линейного расширения при 300 К также очень близки ($5{,}8\cdot10^{-6}$ °C^{-1} и $5{,}3\cdot10^{-6}$ °C^{-1}). Однако, такой «идеальный» переход имеет один существенный недостаток: арсенид алюминия нестабилен на воздухе и во влажной среде он корродирует, подвергаясь гидролизу. Эту проблему долгое время никто не мог преодолеть. Решение оказалось необычайно простым. Стабилизирующим элементом оказался сам галлий, а твердые растворы тройного соединения $Al_xGa_{1-x}As$ вплоть до величины $x \approx 0{,}8$ коррозийно устойчивы. Таким образом, твердые растворы на основе тройных соединений оказались очень перспективными, в том числе и потому, что вариация состава x позволяет не только значительно уменьшить различия в структурных и термических характеристиках компонентов гетероперехода, но и подобрать необходимую ширину запрещенной зоны. Для создания гетеропереходов успешно используются тройные соединения $GaAs_xP_{1-x}$ и $ZnSe_xTe_{1-x}$.

Наиболее перспективными для создания гетеропереходов являются многокомпонентные твердые полупроводниковые растворы на основе четверных соединений, таких как $Ga_xIn_{1-x}As_ySb_{1-y}$.

Успешное применение гетеропереходов в различных приборах в первую очередь обязано эпитаксиальной технологии выращивания пленок, согласованных с решеткой подложки. В получаемых изотипных и анизотипных гетеропереходах практически отсутствуют поверхностные состояния на границе раздела.

Отмеченные выше особенности инжекции в гетеропереходе (эффект «суперинжекции») делает гетеропереход уникальным по эффективности инжектором, особенно перспективным для лазеров и электролюминесцентных приборов.

Наличие широкозонного окна для фотонов, энергия которых больше ширины запрещенной зоны одного и меньше ширины запрещенной зоны другого материала, позволяет использовать гетеропереход в качестве

эффективного фотоэлемента и обеспечивает вывод рекомбинационного излучения без поглощения в полупроводниковых источниках света.

3.3.2.4. Биполярные транзисторы

Биполярный транзистор – это система двух взаимодействующих *p-n*-переходов, имеющих общую *n*- или *p*-область. Крайние области называются эмиттером и коллектором, а средняя область – базой. При этом *p-n*-переход эмиттер-база включается в прямом направлении, а *p-n*-переход база-коллектор – в обратном. Такой режим включения является основным и называется активным (нормальным) режимом. На рис. 3.33 приведено распределение концентраций носителей тока в *p-n-p*-транзисторе в активном режиме. Из рис. 3.33 видно, что уровень легирования базы значительно меньше уровня легирования эмиттера ($n_n << p_{p\text{э}}$), а уровень легирования коллектора еще меньше. Концентрация инжектированных дырок в базе характеризуется большим градиентом, вызванным не столько рекомбинацией, сколько экстракцией дырок коллектором.

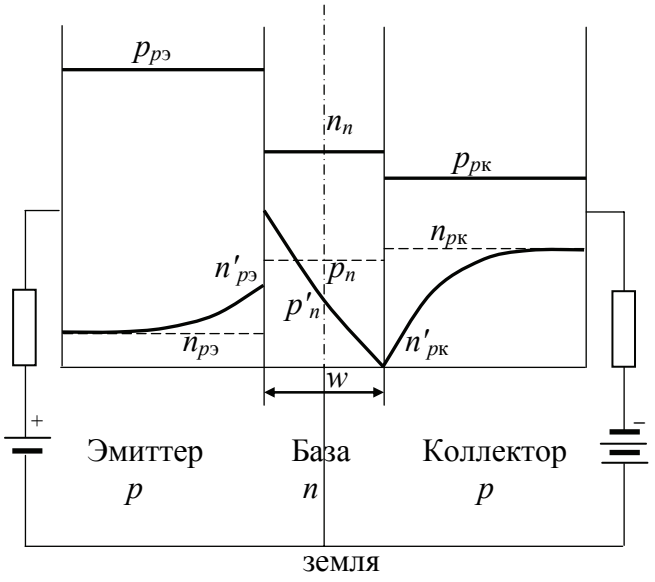


Рис. 3.33 – Распределение концентраций носителей тока в *p-n-p*-транзисторе в активном режиме:
$p_{p\text{э}}$, n_n, $p_{p\text{к}}$ – основные носители;
$n_{p\text{э}}$, $n_{p\text{к}}$ – неосновные носители;
$n'_{p\text{э}}$ – электроны, инжектированные из базы;
p'_n – дырки, инжектированные из эмиттера;
$n'_{p\text{к}}$ – экстрагированные электроны в коллекторе;

Видно также, что концентрация электронов, инжектированных из базы в эмиттер, меньше концентрации дырок, инжектированных из эмиттера в базу. Это является следствием несимметричного легирования перехода эмиттер-база. Коллектор выполняется слаболегированным, чтобы объемный заряд *p-n*-перехода база-коллектор был бы более протяженным, и в нем создавалось поле высокой напряженности.

Найдем основные соотношения в *p-n-p*-транзисторе на малых частотах. Так как током входной цепи является $I_\text{э}$, а током выходной цепи $I_\text{к}$, то величину

$$\alpha = \left.\frac{dI_\text{к}}{dI_\text{э}}\right|_{U_\text{к}=\text{const}} \qquad (3.116)$$

можно назвать *коэффициентом усиления по току* или *коэффициентом передачи тока*. Эта величина является внешним параметром транзистора и определяется произведением внутренних параметров – эффективностью эмиттера γ, эффективностью переноса β* и эффективностью коллектора α*.

Эффективность эмиттера определяется отношением изменения дырочной составляющей тока к изменению полного тока эмиттера:

$$\gamma = \left(\frac{dI_{p\text{э}}}{dI_{\text{э}}} \right)_{U_{\text{к}}=\text{const}} . \qquad (3.117)$$

Коэффициент переноса, характеризующий долю переданных от эмиттера к коллектору дырок через область базы:

$$\beta^* = \left(\frac{dI_{p\text{к}}}{dI_{p\text{э}}} \right)_{U_{\text{к}}=\text{const}} . \qquad (3.118)$$

Эффективность коллектора характеризует рост тока через коллекторный переход за счет воздействия подошедших дырок:

$$\alpha^* = \left(\frac{dI_{\text{к}}}{dI_{p\text{к}}} \right)_{U_{\text{к}}=\text{const}} . \qquad (3.119)$$

Очевидно, что произведение (3.117), (3.118) и (3.119) дает (3.116), т.е.

$$\alpha = \gamma \beta^* \alpha^* . \qquad (3.120)$$

Не вдаваясь в подробности вывода, отметим, что для малых сигналов и низких частот эффективность переноса

$$\beta^* = 1 - \frac{w^2}{2L_p^2}, \qquad (3.121)$$

где w – ширина базы; L_p – диффузионная длина дырок в базе;
и эффективность инжекции

$$\gamma = \frac{1}{1 + \frac{w\sigma_{\text{б}}}{L_n \sigma_{\text{э}}}}, \qquad (3.122)$$

где $\sigma_{\text{б}}$ – удельная проводимость базы; $\sigma_{\text{э}}$ – удельная проводимость эмиттера.

В случае, когда $\sigma_{\text{э}} >> \sigma_{\text{б}}$, а $w << L_p$,

$$\gamma = 1 - \frac{\sigma_{\text{б}}}{\sigma_{\text{э}}} \cdot \frac{w}{L_n}. \qquad (3.123)$$

Для плоскостных транзисторов, за исключением специальных случаев, эффективность коллектора $\alpha^* = 1$. Поэтому, пренебрегая поверхностной рекомбинацией, находим

$$\alpha = 1 - \frac{\sigma_{\text{б}}}{\sigma_{\text{э}}} \cdot \frac{w}{L_n} - \frac{w^2}{2L_p^2}. \qquad (3.124)$$

Из этого выражения следует, что, для того чтобы коэффициент усиления тока α стремился к единице, нужно, чтобы ширина базы w была бы возможно меньшей по сравнению с диффузионными длинами L_p и L_n и удельная проводимость эмиттера $\sigma_{\text{э}}$ была намного больше удельной проводимости базы $\sigma_{\text{б}}$. У хороших транзисторов α достигает значений 0,995…0,998.

Существует три схемы включения транзисторов: схема с общей базой, схема с общим эмиттером и схема с общим коллектором.

Так как $I_к = \alpha I_э$, $I_б = I_э - I_к = I_э(1-\alpha)$, легко подсчитать коэффициент усиления тока для схем с общим эмиттером и общим коллектором.

Коэффициент усиления тока можно определить как отношение тока выходной цепи к току входной:

для схемы с общим эмиттером

$$\beta = \frac{I_к}{I_б} = \frac{\alpha}{1-\alpha}; \qquad (3.125)$$

для схемы с общим коллектором

$$\alpha_к = \frac{I_э}{I_б} = \frac{1}{1-\alpha}. \qquad (3.126)$$

Например, если α = 0,99, то $\alpha_к$ = 99, а если α = 0,998, то $\alpha_к$ = 499. Таким образом, чем ближе коэффициент усиления по току α к единице, тем выше усилительные свойства транзистора как для схемы с общим эмиттером, так и для схемы с общим коллектором.

Чтобы иметь высокий коэффициент усиления, уровень легирования эмиттера должна быть во много раз больше, чем уровень легирования базы, т. е. $\frac{N_э}{N_б} >> 1$. Однако если уровень легирования эмиттера превысит некоторое граничное значение, то коэффициент усиления не увеличится, а уменьшится. Это обусловлено значительными негативными изменениями в кристаллической структуре эмиттера и появлением в запрещенной зоне различных центров захвата носителей, что ухудшает эффективность эмиттера.

Основными материалами для производства транзисторов служат кремний и германий. Из других полупроводников наиболее часто применяется арсенид галлия.

3.3.3. Униполярные приборы

3.3.3.1. Контакты металл-полупроводник. Диоды Шоттки

В предыдущем параграфе был рассмотрен широко применяемый в электронике переход, возникающий при контакте двух различных полупроводниковых материалов с различной шириной запрещенных зон (гетеропереход). А что получится, если рассмотреть контакт двух различных материалов, например, контакт металл-полупроводник?

Оказывается, что свойства перехода между полупроводниками *p*- и *n*-типа отличаются от свойств перехода между полупроводником и металлом, так как работа выхода полупроводника не равна работе выхода металла.

Рассмотрим случаи, когда полупроводник *n*-типа приведен в контакт с металлом, у которого работа выхода больше, чем у полупроводника ($\chi_м > \chi_п$) и полупроводник *p*-типа, приведенный в контакт с металлом, у которого работа выхода меньше, чем у полупроводника ($\chi_м < \chi_п$) (рис. 3.34, *а*). Различают две работы выхода: внешнюю, отсчитываемую от дна зоны проводимости ϕ, и термодинамическую, отсчитываемую от уровня Ферми $\chi_п$.

В последнее время вместо термина «внешняя работа выхода» все более широко употребляют термин электронное сродство, или сродство к электрону. Работа выхода полупроводника определяется как разность

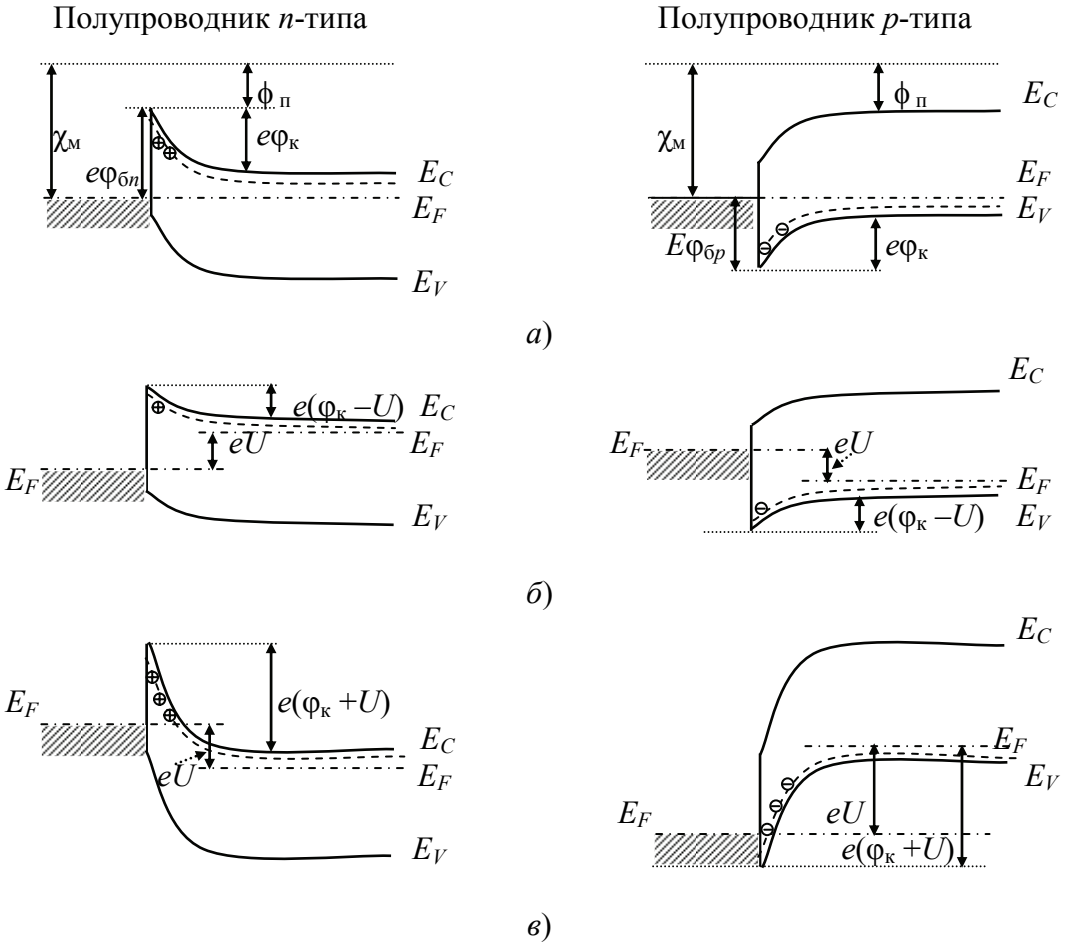


Рис. 3.34 – Зонные энергетические диаграммы контактов металла с полупроводниками *n*-типа при $\chi_м > \chi_п$ и *p*-типа при $\chi_м < \chi_п$: *а* – при термодинамическом равновесии; *б* – при прямом смещении; *в* – при обратном смещении

энергий электрона на уровне Ферми и на вакуумном уровне. Если вещества соприкасаются друг с другом, то можно считать, что уровни Ферми везде одинаковы (при термодинамическом равновесии уровни Ферми в контакте металл-полупроводник выравниваются, как и в случае контакта полупроводника *p*-типа с полупроводником *n*-типа (рис. 3.34, *а*)). Картина распределения уровней в металле останется прежней, так как можно считать, что никакого «изгиба энергетических зон» в металле не происходит. (В металле нескомпенсированные заряды могут существовать только на поверхности). Поэтому все падение потенциала будет на участке полупроводника. Это вполне обоснованно, поскольку число носителей в полупроводнике существенно меньше чем в металле. Подобное распределение существует и в несимметричных *p*-*n*-переходах: p^+-*n* или *p*-n^+.

Зонные энергетические диаграммы контактов металла с полупроводниками *n*-типа при $\chi_м > \chi_п$ и *p*-типа при $\chi_м < \chi_п$ для разных видов смещений приведены на рис. 3.34, *а*, *б*, *в*.

При идеальном контакте между металлом и полупроводником *n*-типа, высота барьера $e\varphi_{бn}$ равна разности между работой выхода металла и электронным сродством полупроводника:

$$e\varphi_{бn} = \chi_м - \phi_п . \qquad (3.127)$$

При идеальном контакте между металлом и полупроводником *p*-типа высота барьера $e\varphi_{бp}$ определяется аналогичным выражением

$$e\varphi_{бp} = E_g - (\chi_м - \phi_п) . \qquad (3.128)$$

При равновесии число электронов, переходящих через барьер из металла в полупроводник, равно числу носителей, переходящих со стороны полупроводника. Можно говорить, что ток течет в обоих направлениях. Приложим к переходу некоторое напряжение; в зависимости от его полярности потенциальная энергия электронов на полупроводниковой стороне либо возрастает, либо уменьшается. При прямом смещении высота барьера уменьшается (рис. 3.34, *б*). В этом случае все электроны с энергиями выше $\chi_м - \phi_п - eU$ могут попасть в металл. По аналогии с *p*-*n*-переходом отсюда следует, что число носителей, способных перейти из полупроводника в металл, растет пропорционально множителю $\exp\left(\frac{eU}{kT}\right)$; протекающий ток пропорционален такому же множителю. Поскольку ток из металла в полупроводник остается прежним, то общий ток равен

$$I = I_0\left(e^{\frac{eU}{kT}} - 1\right). \qquad (3.129)$$

Это означает, что рассматриваемый прибор представляет собой выпрямитель.

На вопросе о потенциальном барьере остановимся несколько подробнее. Одним из электродов в переходе служит металл, вблизи поверхности которого сосредоточены заряженные частицы. Эмиссии электронов из металла препятствует потенциальный барьер, образующийся за счет электрических сил изображения. Понижение этого барьера по мере увеличения приложенного внешнего напряжения называется *эффектом Шоттки* (см. параграф 2.1.5). Для описания этого эффекта в контакте металл-полупроводник можно воспользоваться формулами (2.50, 2.51), только величину ε_0 следует заменить на $\varepsilon_0\varepsilon$, где ε – относительная диэлектрическая проницаемость полупроводника. Например, для кремния $\varepsilon = 12$, так что эффект выражен слабее. Высота потенциального барьера, который должен преодолеет электрон, понижается на величину U_{max}

$$U_{max} = -e\left(\frac{eE}{4\pi\varepsilon_0\varepsilon}\right)^{\frac{1}{2}} . \qquad (3.130)$$

Эффективная работа выхода уменьшается от χ до $\chi_{эфф}$

$$\chi_{эфф} = \chi - e\left(\frac{eE}{4\pi\varepsilon_0\varepsilon}\right)^{\frac{1}{2}}. \qquad (3.131)$$

Хотя эффект Шоттки невелик, именно он дал свое имя рассматриваемым переходам. Обычно их называют *диодами Шоттки* или *диодами с барьером Шоттки*.

ВАХ контакта металл-полупроводник в соответствии с диодной теорией выпрямления определяется уравнением (3.118). На рис. рис. 3.35 приведены ВАХ для тонкого и толстого слоев объемного заряда в контакте металл-полупроводник. При одинаковой концентрации свободных носителей заряда тонкий слой объемного заряда имеет меньшее сопротивление, поэтому прямая ветвь кривой 1 идет круче ветви кривой 2, на обратной кривой отсутствует насыщение и обратный ток возрастает пропорционально $U^{1/2}$.

На рис. 3.36 приведены прямые ветви ВАХ выпрямительного диода с барьером Шоттки (1) и диода с *p*-*n*-переходом (2). Из рисунка видно, что падение напряжения на диодах Шоттки существенно меньше падения напряжения на диодах с *p*-*n*-переходом.

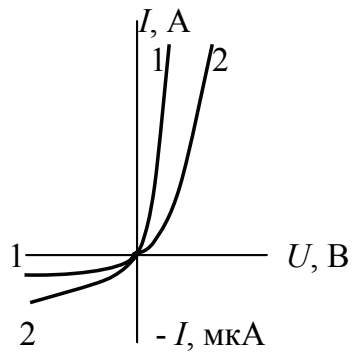


Рис. 3.35 – ВАХ для толстого и тонкого объемного слоя в контакте металл-полупроводник

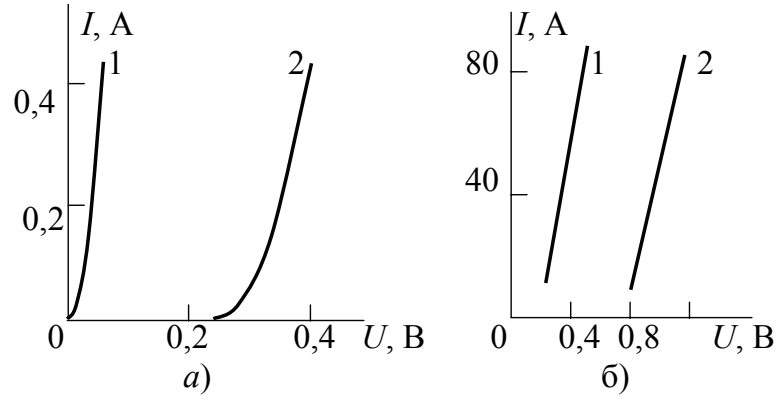


Рис. 3.36 – ВАХ выпрямительного диода с барьером Шоттки (1) и с *p*-*n*-переходом (2) для малых (а) и больших (б) токов, измеренных в импульсном режиме

Таким образом, отличительными особенностями диода Шоттки являются высокое быстродействие, малое падение напряжения при прямом смещение и высокий КПД выпрямления.

Известно, что палладий может адсорбировать большие объемы водорода. Если воспользоваться этим свойством палладия и напылить его на кремниевую подложку, то можно получить датчики водорода на основе диодов Шоттки [1-3], работающих как МП-конденсатор. Водород способен уменьшать плотность поверхностных состояний на границе раздела палладий-кремний, всвязи с чем емкость диода в области рабочей точки на вольтфарадной характеристике при постоянном напряжении уменьшается. Обнаружена связь между концентрацией водорода в среде и емкостью

палладиевого диода Шоттки (при увеличении концентрации емкость уменьшается).

Некоторые экспериментальные значения высоты барьера Шоттки представлены в табл. 3.5.

Таблица 3.5 – Экспериментальные значения высоты барьера Шоттки при 300 К (в вольтах)

Полу-провод-ник	Тип	E_g, эВ	Ag	Al	Au	Cr	Cu	In	Mg	Ni	Pd	Pt	Ti
Алмаз	*p*	5,47			1,71					0,38			
Ge	*n*	0,66	0,54	0,48	0,59		0,52	0,64					
Ge	*p*		0,50		0,30			0,55	0,40		0,81	0,90	0,50
Si	*n*	1,12	0,78	0,72	0,80	0,61	0,58			0,55			0,61
Si	*p*		0,54	0,58	0,34	0,50	0,46						
SiC	*n*	3,00		2,00	1,95							1,00	
AlAs	*n*	2,16			1,2								
AlSb	*p*	1,63			0,55								
BN	*p*	7,5			3,10								
BP	*p*	6,00			0,87								
GaSb	*n*	0,67			0,60								
GaAs	*n*	1,42	0,88	0,80	0,90		0,82					0,84	
GaAs	*p*		0,63		0,42								
GaP	*n*	2,24	1,20	1,07	1,30	1,06	1,20		1,04			1,45	1,12
GaP	*p*				0,72								
InSb	*n*	0,16	0,18*		0,17*								
InAs	*p*	0,33			0,47*								
InP	*n*	1,29	0,54		0,52								
InP	*p*				0,76								
CdS	*n*	2,43	0,56	Омич.	0,78		0,50			0,45	0,62	1,10	0,84
CdSe	*p*	1,7	0,43		0,49		0,33					0,37	
CdTe	*n*		0,81	0,76	0,71							0,76	
ZnO	*n*	3,2		0,68	0,65		0,45	0,30			0,68	0,75	
ZnS	*n*	3,6	1,65	0,80	2,00		1,75	1,50	0,82		1,87	1,84	
ZnSe	*n*		1,21	0,76	1,36		1,10	0,91				1,40	
PbO	*n*		0,95					0,93		0,96			

* При 77 К

Рассмотрим теперь случай, когда значение работы выхода металла ниже, чем у полупроводника *n*-типа ($\chi_м < \chi_п$) и выше, чем у полупроводника *p*-типа ($\chi_м > \chi_п$). Зонные диаграммы этих переходов даны на рис. 3.37 *а*, *б*. Например, для контакта металл-полупроводник *n*-типа электроны для достижения равновесия должны переместиться из металла в полупроводник, создавая в полупроводнике некоторую область накопления заряда. С какой бы стороны не рассматривать такой переход, в нем не удается обнаружить никакого

потенциального барьера. Вследствие этого ток через переход не зависит

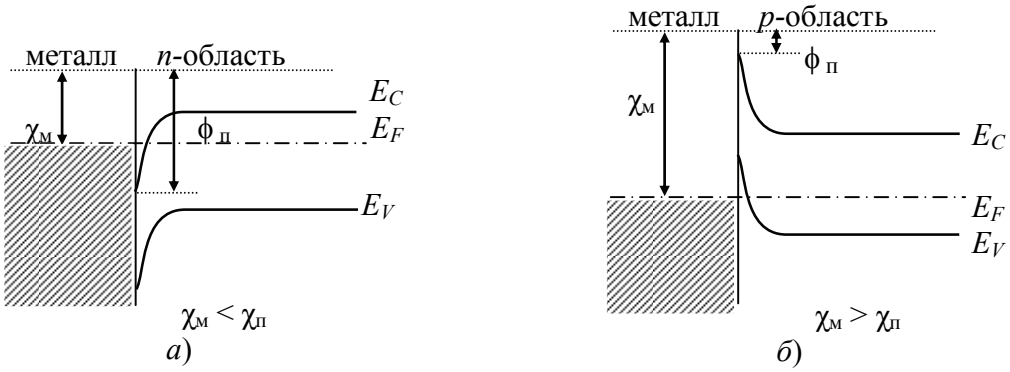


Рис. 3.37 – Образование антизапирающего слоя в области контакта металла с полупроводником *n*-типа при $\chi_м < \chi_п$ (*а*) и металла с полупроводником *p*-типа при $\chi_м > \chi_п$ (*б*)

существенно от приложенного напряжения и поэтому такой переход не обладает выпрямляющими свойствами, т.е. является хорошим омическим контактом. Аналогичная ситуация будет наблюдаться и для контакта металл-полупроводник *p*-типа. Омическим контактом называют контакт металл-полупроводник, сопротивление которого пренебрежимо мало по сравнению с объемным сопротивлением полупроводника.

В контактах металл-полупроводник с более высоким уровнем легирования преобладает туннельная компонента тока (рис. 3.38). Удельное сопротивление такого контакта в области туннелирования экспоненциально зависит от $\frac{\varphi_{бп}}{\sqrt{N_D}}$. При концентрации доноров $N_D > 10^{25}$ м$^{-3}$ удельное сопротивление омического контакта определяется в основном туннельными процессами и быстро уменьшается по мере повышения степени легирования.

Таким образом, для получения малых сопротивлений нужна либо высокая степень легирования, либо малая высота барьера, либо то и другое вместе. Именно из этих соображений исходят при изготовлении омических контактов.

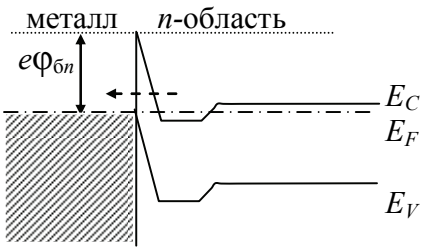


Рис. 3.38 – Зонная энергетическая диаграмма омического контакта металла с полупроводником *n*-типа высокой степени легирования при $\chi_м < \chi_п$

Хороший омический контакт не должен приводить к существенному изменению характеристик прибора, а падение напряжения на таком контакте при пропускании через него требуемого тока должно быть достаточно мало по сравнению с падением напряжения на активной области прибора. Наиболее надежными омическими контактами являются вплавленные контакты. Для вплавления применяются различные сплавы на основе олова, индия, свинца, алюминия и золота. ВАХ

омического контакта при прямом и обратном смещении является практически линейной и показывает низкое сопротивление.

3.3.3.2. Полевые транзисторы. Контакты металл-диэлектрик-полупроводник. МДП-транзисторы. Приборы с зарядовой связью

Одной из разновидностей транзисторов является полевой транзистор. Его структура показана на рис. 3.39. В кремниевой пластине *p*-типа формируют канал проводимости *n*-типа с двумя омическими контактами, обозначенными на рис. 3.39 *И* и *С*, что означает исток и сток. По каналу *n*-типа протекает ток основных носителей заряда – электронов от истока к стоку. (Полевые транзисторы могут иметь каналы проводимости как *n*-, так и *p*-типа). К верхней *p*-области структуры изготовляют омический контакт *З*, называемый затвором. Аналогичный контакт изготовляют к нижней *p*-области. Таким образом, в рассматриваемой структуре есть два *p*-*n*-перехода.

В полевых транзисторах с управляющим *p*-*n*-переходом подложка обычно технологически соединяется с верхним затвором. В этом случае изменение величины проводимости канала осуществляется под действием напряжения, приложенного к *p*-*n*-переходам верхнего низкоомного затвора и

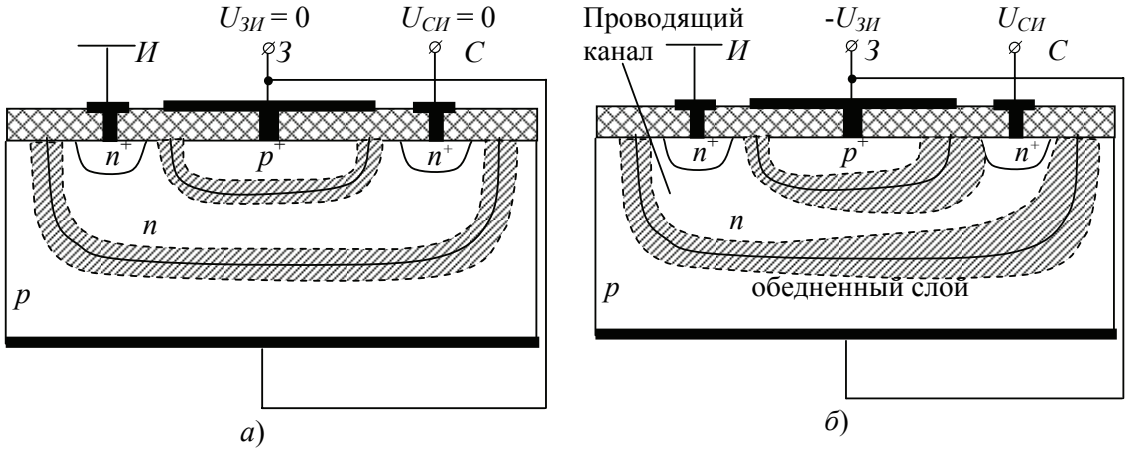


Рис. 3.39 – Полевой транзистор с управляющим *p*-*n*-переходом: *а*) – режим равновесия; б) – рабочий режим

высокоомной подложки – нижнего затвора. Рис. 3.39 поясняет принцип действия полевого транзистора, полярность напряжений которого соответствует транзистору с каналом *n*-типа. Исток обычно заземлен и относительно него измеряют напряжение на стоке и затворе, на затвор подается знак «–» источника тока. Толщины обедненных слоев *p*-*n*-переходов, ограничивающих проводящий канал в отсутствие смещений, минимальны (рис. 3.39, *а*).

При приложении напряжения к стоку ($U_{СИ} > 0$) и при напряжении на затворе, равном нулю ($U_{ЗИ} = 0$), по каналу течет ток, создаваемый основными носителями заряда (в данном случае электронами). Ток стока растет пропорционально напряжению. Увеличение напряжения на стоке приводит к

увеличению разности потенциалов между каналом и затвором, что, в свою очередь, увеличивает толщины обедненных слоев *p*-*n*-перехода и уменьшает сечение канала *n*-типа. Толщина обедненных слоев минимальна у истока и максимальна у стока (рис. 3.39, *б*).

При некотором напряжении на стоке обедненные слои вблизи него смыкаются, и наступает состояние перекрытия канала. Соответствующие напряжения на стоке называются напряжением насыщения ($U_{СИ\text{нас}}$). Дальнейшее увеличение не приводит к росту тока стока, а лишь увеличивает напряженность поля в обедненном слое, при этом точка смыкания обедненных слоев смещается в сторону истока.

Полевой транзистор – это прибор, в котором ток через канал управляется электрическим полем, возникающим при приложении электрического напряжения между затвором и истоком. Увеличение обратного напряжения на затворе приводит к расширению объемного заряда, сужению проводящего канала и уменьшению тока стока. Обратные токи *p*-*n*-перехода невелики и для управления достаточны малые мощности, потребляемые от источника сигнала в цепи затвора. Поэтому полевой транзистор способен усиливать электрические сигналы, как по мощности, так и по току и напряжению.

Полевые транзисторы имеют существенно более высокое входное сопротивление, чем биполярные транзисторы, что облегчает их сопряжение с СВЧ-устройствами. Так как полевые транзисторы являются униполярными приборами, они слабо чувствительны к эффектам накопления неосновных носителей, и поэтому имеют более высокие граничные частоты и скорости переключения.

Высокочастотные характеристики полевых канальных транзисторов зависят, в основном, от двух факторов: времени пролета τ и характерного времени зарядки затвора *RC*. Временем пролета называется временной интервал, который требуется носителям для преодоления расстояния от истока до стока.

Граничная частота полевых канальных транзисторов определяется геометрией приборов и параметрами материала:

$$f_{\text{гр}} = \frac{1}{2\pi\tau} = \frac{\upsilon_s}{2\pi l}, \qquad (3.132)$$

где l – длина канала; υ_s – скорость насыщения (максимальная дрейфовая скорость носителей в сильных полях). В кремнии Si и арсениде галлия GaAs электроны имеют более высокую подвижность, чем дырки. Поэтому в СВЧ-схемах используются только *n*-канальные полевые транзисторы. Кроме того, так как в GaAs подвижность электронов при малых полях примерно в 5 раз выше, чем в кремнии, граничная частота в GaAs-транзисторах больше, чем в аналогичных кремниевых приборах.

Если нанести металлический электрод не на чистый полупроводник, а на окисел этого полупроводника, к примеру, на оксид кремния SiO_2, то получим так называемую структуру металл-диэлектрик-полупроводник или МДП-структуру (МОП-структуру). SiO_2 на поверхности кремния обычно выращивается с помощью термического окисления, но могут быть

использованы и другие способы нанесения – плазменное или лазерное оксидирование. Вместо оксида кремния можно использовать и другие диэлектрики, например нитрид кремния.

МДП-структуры играют как самостоятельную, так вспомогательную роль, если используются в качестве затворов в МДП-транзисторах, которые являются важным видом полевых транзисторов.

В транзисторах со МДП-структурой принцип работы основан на модуляции сопротивления проводящего канала под воздействием поля. МДП-транзисторы являются униполярными приборами, работа которых основана на использовании только основных носителей заряда. Процессы инжекции в МДП-транзисторах не используются. На рис. 3.40 схематически показана конструкция МДП-транзистора двух типов. В полупроводниковой подложке *n*-типа сформированы две высоколегированные p^+-области: исток и сток. Металлический электрод – затвор, отделен от подложки тонким слоем диэлектрика, в рассматриваемом случае – оксидом кремния SiO_2. В диэлектрическом слое делаются «окна» для изготовления контактов истока и стока. На поверхность слоя SiO_2 наносят слой алюминия, служащего затвором.

В МДП-структуре, изображенной на рис. 3.40, *а*, предварительно до оксидирования создают слаболегированную *p*-область. Потом при

Рис. 3.40 – Структура МДП-транзисторов с каналом *p*-типа: *а* – со встроенным каналом; *б* – с индуцированным каналом

вплавлении (или диффузии) алюминия в окна под истоком и стоком эти области легируются дополнительно алюминием и становятся p^+-областями. Таким образом, в МДП-транзисторе этого вида проводящий канал *p*-типа оказывается встроенным технологически.

На рис. 3.40, *б* показана иная МДП-структура, в которой заметный ток от истока к стоку начинает проходить только при определенном *пороговом напряжении* $U_{пор}$ между затвором и истоком. При этом электрическое поле проникает под затвором в полупроводник и у поверхности образуется обедненный основными носителями заряда (электронами) слой. А при напряжении U, превышающем пороговое значение, обедненный слой вообще станет инверсионным, т.е. приобретет проводимость *p*-типа. Иными словами, под затвором проводящий канал *p*-типа в этой структуре *индуцируется*, т.е. его не получают технологически.

Управляющей цепью в МДП-транзисторах является цепь затвора, управляемой – цепь истока-стока. Управляющая цепь практически не потребляет тока, так как в нее входит участок с диэлектриком, поэтому в МДП-транзисторах получается значительное усиление мощности – намного больше, чем в биполярных транзисторах.

В 1970 году были предложены и сегодня широко используются *приборы с зарядовой связью* (ПЗС), которые представляют собой матрицу близко расположенных и взаимодействующих между собой МДП-структур (рис. 3.41). ПЗС применяют в запоминающих устройствах, устройствах аналоговой обработки информации, в качестве формирователей видеосигналов (ФВС) и др. ФВС на основе ПЗС позволяет преобразовывать оптическое изображение в последовательность электрических видеосигналов, заменяя передающие вакуумные телевизионные трубки.

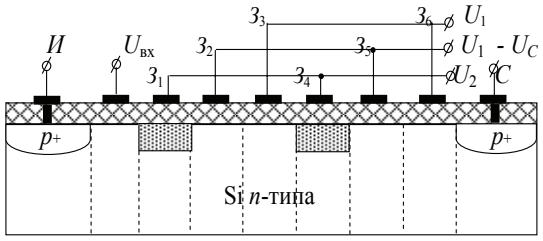


Рис. 3.33 – Схема прибора с зарядовой связью

Принцип работы ПЗС состоит в том, что в каждой отдельной МДП-структуре можно создать локальный приповерхностный заряд неосновных носителей и перемещать его вдоль поверхности от одной МДП-структуры к другой, меняя соответствующим образом последо-вательность тактовых импульсов, подаваемых на затворы. ПЗС обычно строят на основе кремния *n*-типа и на затворы подают отрицательное напряжение, по модулю меньше порогового. В полупроводнике *n*-типа образуется обедненная основными носителями заряда область в виде потенциальной ямы. В потенциальной яме скапливаются неосновные носители заряда (дырки), образующие зарядовый пакет. Если на затворах $З_3$ и $З_5$ действует напряжение U_1, а на среднем затворе $З_4$ – более отрицательное напряжение U_2, то на границах затворов $З_2$ и $З_5$ образуются электрические поля, препятствующие перемещению дырок из потенциальной ямы. Пакет дырок под затвором может сохраняться в течение определенного времени. Ввод зарядового пакета под затвор называют режимом записи информации, а напряжение U_2, обеспечивающее такой ввод, – напряжением записи.

Возможность накопления зарядов в ПЗС определила три главных области их применения: в качестве запоминающих устройств ЭВМ и персональных компьютеров; устройств обработки аналоговой информации; устройств, в которых электрические сигналы преобразуются в оптическое изображение. В последнем случае кроме кремния используются полупроводники типа $A^{III}B^{V}$ и их твердые растворы.

Главным объектом разработки и научных исследований микроэлектроники являются интегральные микросхемы: ИС (интегральные микросхемы), БИС (большие интегральные микросхемы), СБИС (сверхбольшие интегральные микросхемы) и УБИС (улучшенные большие

интегральные микросхемы). Они представляет собой изделия, состоящие из нескольких компонентов, изготовленных в одном кристалле и связанных друг с другом так, что все изделие в целом предназначено выполнять определенные функции преобразования электрических сигналов. Есть микросхемы для преобразования оптических сигналов видимого и инфракрасного участков спектра.

Всю интегральную схему характеризуют *плотностью упаковки* – числом элементов в 1 см^3 полупроводника и *степенью интеграции* – числом элементов, входящих в состав одной микросхемы. Дальнейшее развитие микроэлектроники связано с увеличением этих параметров. В настоящее время наиболее актуальны исследования и разработки элементов для СБИС и УБИС, проводимые в трех основных направлениях. Первое развивается на основе кремниевых МДП-транзисторов и позволяет получать максимальную степень интеграции (число элементов $10^6…10^9$) при достаточно высоком быстродействии (быстродействие 0,5…1 нс). Во втором направлении используются гетероструктурные кремниевые биполярные транзисторы Si-Ge и обеспечивается повышенное быстродействие порядка 0,1…0,5 нс, но при меньшей степени интеграции. Третье направление позволяет достигать сверхвысокого быстродействия порядка 50…100 пс при числе элементов $10^3…10^6$, оно развивается на основе арсенид-галлиевых транзисторов и гетероструктурных транзисторов с высокой подвижностью электронов.

3.3.4. Светодиоды

3.3.4.1. Физические принципы работы светодиода

Светодиоды принадлежат к классу люминесцентных приборов. Мы будем рассматривать в основном электролюминесценцию, и, прежде всего инжекционную электролюминесценцию, которая представляет собой оптическое излучение, возникающее при инжекции неосновных носителей в область полупроводникового *p-n*-перехода, где излучательные переходы наиболее эффективны.

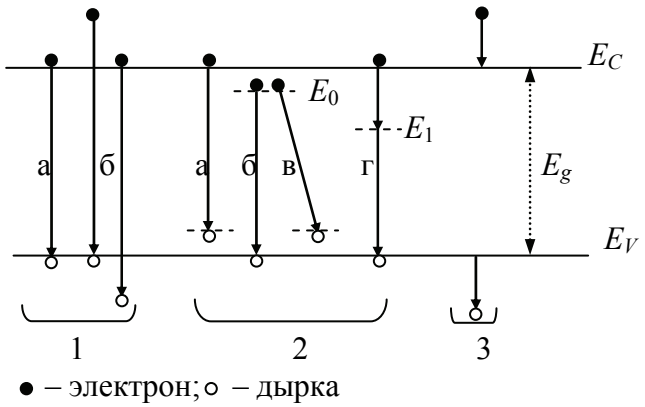


Рис. 3.42 – Основные переходы электронов в полупроводнике

На рис. 3.42 схематически представлены основные переходы в полупроводнике. Эти переходы можно характеризовать следующим образом.

Во-первых, это межзонные переходы, вызывающие: а) собственное излучение с энергией, очень близкой к ширине запрещенной зоны, которое может сопровождаться возбуждением фотонов или экситонов; б) излучение с более высокой энергией с участием

энергичных или «горячих» носителей, которое иногда может быть связано с лавинным пробоем.

Во-вторых, это переходы с участием химических примесей или физических дефектов: а) между зоной проводимости и акцепторным уровнем; б) между донорным уровнем и валентной зоной; в) между донорными и акцепторными уровнями (межпримесное излучение); г) через глубокие уровни.

В-третьих, это внутризонные переходы, которые вызывают излучение, называемое иногда тормозным, и которые протекают с участием «горячих» носителей.

Эффективными с точки зрения люминесценции являются такие материалы, в которых излучательные переходы преобладают над безызлучательными, т.е. прямозонные полупроводники.

Рассмотрим работу светодиода, испускающим при *прямом смещении* спонтанное рекомбинационное излучение в оптическом диапазоне (ультрафиолетовой, видимой и инфракрасной области спектра).

Излучение возникает (переходы 1-2 на рис. 3.43) вследствие рекомбинации инжектированных из эмиттера носителей в одной из частей *p-n*-перехода, называемой базой (на рис. 3.43 база – это *p*-область, прилегающая к границе раздела *p-n*-перехода). Рекомбинация происходит при переходе электронов с верхних уровней на нижние. Длина волны излучения при переходе электрона из зоны проводимости в валентную зону

$$\lambda = \frac{c}{\nu} = \frac{ch}{E_g}, \qquad (3.133)$$

где c – скорость света; E_g – ширина запрещенной зоны полупроводника.

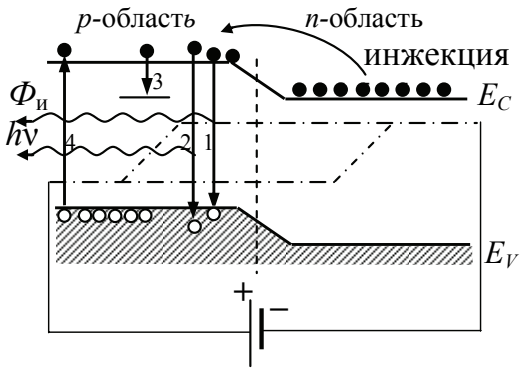


Рис. 3.43 – Электронные переходы на упрощенной зонной диаграмме светодиода при прямом смещении

Приведенная формула определяет лишь длину волны в максимуме спектра излучения λ_{max}. В реальных случаях рекомбинация часто происходит между электронами и дырками, расположенными на двух группах энергетических уровней.

3.3.3.2. Характеристики светодиодов

Основной физической величиной, характеризующей качество светодиодов, является *внешний квантовый выход* $\eta_{внш}$ – отношение числа выходящих из светодиода фотонов $\Phi_и$, к полному числу носителей заряда, протекающих через *p-n*-переход N_{p-n}:

$$\eta_{внш} = \frac{\Phi_и}{N_{p-n}} . \qquad (3.134)$$

Внутренняя квантовая эффективность генерации в базе $\eta_{внт}$ – отношение числа фотонов $\Phi_г$ к числу инжектированных в базу носителей $N_и = N_{p\text{-}n}\cdot\gamma$ (γ – *коэффициент инжекции*):

$$\eta_{внт} = \frac{\Phi_г}{N_и}. \qquad (3.135)$$

Число фотонов, излученных светодиодом, определяется также *коэффициентом вывода света из базы* $K_б$ (оптической эффективностью): $\Phi_и = \Phi_г \cdot K_б$.

Тогда внешний квантовый выход

$$\eta_{внш} = \eta_{внт}\cdot\gamma \cdot K_б. \qquad (3.137)$$

Оказывается, что $0 < \eta_{внш} < 1$, так как каждый из сомножителей в (3.137) больше нуля, но не может быть больше единицы.

Очевидно, чем больше внешний квантовый выход $\eta_{внш}$, тем выше качество светодиода. В связи с этим проведем анализ величин, входящих в (3.137).

1) Основной фактор, который уменьшает *внутреннюю квантовую эффективность* $\eta_{внт}$ – это безызлучательная рекомбинация на дефектах структуры (переход 3 на рис. 3.35)

Число фотонов, генерированных в *p*-области $\Phi_г = N_и \cdot r_и \cdot \tau$, где τ – эффективное время жизни инжектированных в базу электронов

$$\tau = \frac{1}{r_и + r_б}, \qquad (3.138)$$

определяется скоростью излучательной рекомбинации $r_и$ и безызлучательной рекомбинации $r_б$ в базе. Согласно определению $\eta_{внт}$ с использованием приведенных выше формул получим при $r_и > r_б$:

$$\eta_{внт} = \frac{N_и \cdot r_и \cdot \tau}{N_и} = \frac{r_и}{r_и + r_б} = \frac{1}{1 + \frac{r_б}{r_и}} \approx 1 - \frac{r_б}{r_и}. \qquad (3.139)$$

Видно, что для увеличения внутренней квантовой эффективности необходимо повышение доли излучательных переходов, что достигается внесением излучательных центров в базу (*p*-область) и созданием в ней достаточной концентрации основных носителей p_p, необходимых для рекомбинации с электронами.

2) Обычно эффективное излучение возникает при рекомбинации электронов, инжектированных из *n*-области в *p*-базу. Тогда в (3.137) *коэффициент инжекции электронов* γ определяется как отношение электронного тока инжекции к полному току через диод, который при условии $I_n >> I_p$ имеет вид:

$$\gamma = \frac{I_n}{I_n + I_p} = \frac{1}{1 + \frac{I_p}{I_n}} \approx 1 - \frac{I_p}{I_n}. \qquad (3.140)$$

Для повышения коэффициента инжекции необходимо увеличение концентрации основных носителей n_n в эмиттере. Поэтому в качестве светодиода чаще используется n^+-p-переход, в котором сильно легирована n^+-область.

3) Увеличение *коэффициента вывода света из базы* $K_б$ связано с уменьшением трех видов потерь света:

- *самопоглощения* света, генерированного в базе полупроводниковым материалом светодиода (переход 4 на рис. 3.34);
- поглощения боковой поверхностью базы и задней стенкой светодиода;
- полного внутреннего отражения света от поверхности светодиода или от границы раздела базы с защитной оболочкой светодиода.

Другой важной характеристикой, которая должна учитываться при конструировании светодиодов для оптических систем связи, является *диапазон рабочих частот*.

Предельная рабочая частота светодиода

$$f_{\max} = \frac{1}{2\pi\tau} \qquad (3.141)$$

определяется временем жизни τ инжектированных в p-область электронов.

3.3.3.3. Материалы для светодиодов

Мы рассмотрим вначале полупроводниковые материалы для светодиодов видимого участка спектра, которые широко применяются в информационных каналах, обеспечивающих связь электронной аппаратуры с ее пользователями, а затем материалы для инфракрасных светодиодов, которые эффективно используются в оптронах и в волоконно-оптических линиях связи (ВОЛС).

Материалы для светодиодов видимого диапазона. Так как глаз чувствителен к свету с энергией $h\nu \geq 1{,}8$ эВ (~ 0,7 мкм), то полупроводники, которые могут быть использованы для создания светодиодов видимого диапазона, должны иметь ширину запрещенной зоны больше этого значения. Прямозонные полупроводники особенно важны для электролюминесцентных приборов, так как излучательная рекомбинация в них является процессом первого порядка (протекает без участия фононов) и ожидаемая квантовая эффективность должна быть более высокой, чем для полупроводников с непрямой зоной, в которых при переходах возбуждаются фононы. Среди полупроводников наибольший интерес для вышеозначенных целей представляют полупроводники смешанного состава, например, $GaAs_{1-x}P_x$, относящиеся к соединениям $A^{III}B^{V}$.

Ширина запрещенной зоны в этих полупроводниках зависит от содержания компонент x. Поэтому, меняя значения x, можно получить необходимую ширину запрещенной зоны. В диапазоне $0 < x < 0{,}45$ полупроводник является прямозонным, а ширина запрещенной зоны увеличивается от $E_g = 1{,}424$ эВ при $x = 0$ до $E_g = 1{,}977$ эВ при $x = 0{,}45$. При $x > 0{,}45$ полупроводник становится непрямозонным. В таких полупроводниках вероятность межзонных переходов очень мала. Поэтому

для усиления излучательных процессов в непрямозонных полупроводниках, таких, например, как фосфид галлия GaP, специально создаются (вводятся) рекомбинационные центры.

Эффективные центры излучательной рекомбинации в $GaAs_{1-x}P_x$ могут быть созданы путем внедрения специальных примесей, например азота. Азот, вводимый в полупроводник, замещает атомы фосфора в узлах решетки. Азот и фосфор имеют одинаковую внешнюю электронную структуру (оба относятся к V группе элементов периодической системы), а структуры их внутренних оболочек сильно отличаются. Это приводит к возникновению вблизи зоны проводимости электронного уровня захвата. Полученный таким образом рекомбинационный центр называется *изоэлектронным центром*. В нормальном состоянии изоэлектронные центры нейтральны. В материале *p*-типа инжектированный электрон вначале захватывается на центр. Затем заряженный отрицательно центр захватывает дырку из валентной зоны, формируя связанный экситон. Последующая аннигиляция этой электронно-дырочной пары приводит к рождению фотона с энергией, примерно равной разности между шириной запрещенной зоны и энергией центра связи ($h\nu = E_g - \Delta E$). Так как захваченный электрон локализован на центре, его импульс рассеивается. Таким образом, обеспечивается закон сохранения квазиимпульса, вследствие чего вероятность непрямого перехода существенно возрастает. В непрямозонных материалах, таких, как GaP, описанный механизм излучательной рекомбинации является преобладающим.

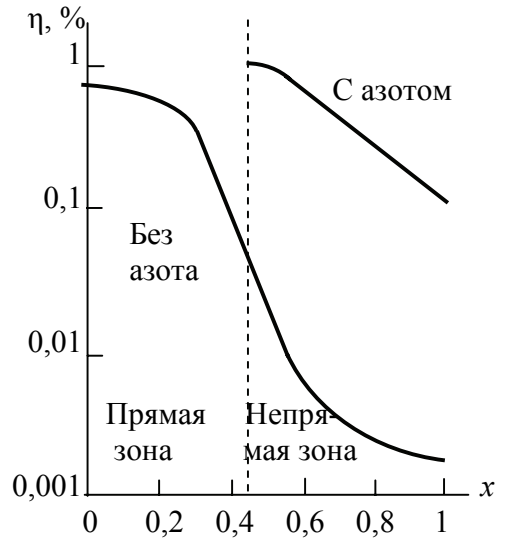


Рис. 3.44 – Зависимость квантовой эффективности η полупроводника $GaAs_{1-x}P_x$ от состава соединения x при наличии и в отсутствие изоэлектронной примеси азота при температуре $T = 300$ К

На рис. 3.44 приведена зависимость квантовой эффективности от состава $GaAs_{1-x}P_x$ при наличии и в отсутствие изоэлектронной примеси азота. Резкое уменьшение эффективности излучения $GaAs_{1-x}P_x$ без примеси азота в диапазоне $0,4 < x < 0,5$ обусловлено изменением механизма рекомбинации: при $x = 0,45$ и меньше преобладающими становятся прямые переходы. Наличие азота обеспечивает более высокие значения эффективности при $x > 0,45$, тем не менее, эффективность все же постепенно уменьшается с ростом x вследствие увеличения разности глубин прямой и непрямой запрещенных зон.

На рис. 3.45 представлены относительные спектры излучения некоторых типичных светодиодов при комнатной температуре. Красные светодиоды могут быть изготовлены из прямозонного соединения $GaAs_{0,6}P_{0,4}$, а также на основании GaP с примесью ZnO, которая, обладая таким же числом общих

валентных электронов, как и пара соседних атомов галлий-фосфор, замещает

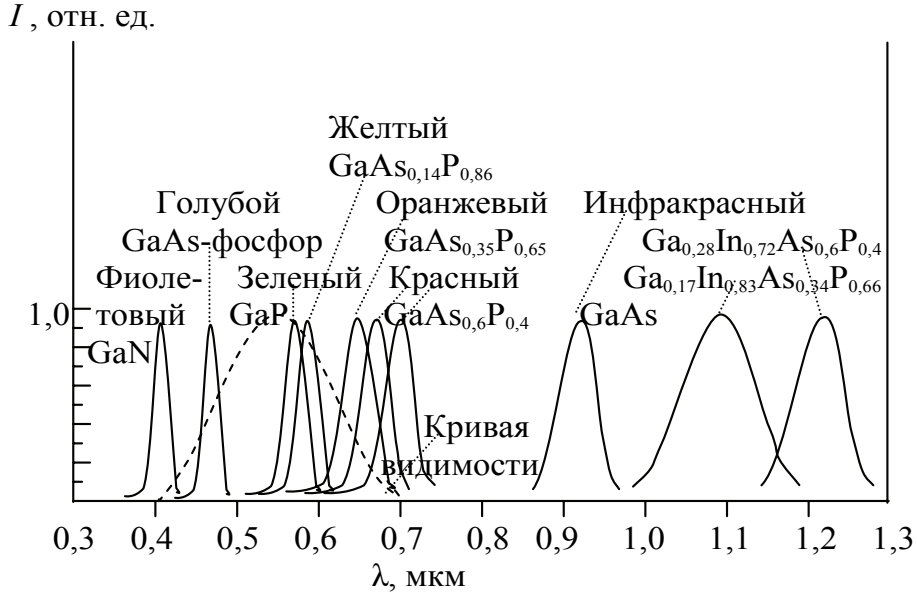


Рис. 3.45 – Относительные спектры излучения различных светодиодов видимого и инфракрасного диапазонов

их в решетке полупроводника. Уровень рекомбинации, обусловленный этой примесью, расположен на 0,3 эВ ниже дна зоны проводимости GaP, в результате чего испускаются кванты с энергией 1,95 эВ. Светодиоды, излучающие оранжевый, желтый и зеленый свет, создаются на основе непрямозонных соединений $GaAs_{1-x}P_x$ или GaP, легированных изоэлектрической примесью азота. Светодиоды, излучающие голубой свет, могут быть изготовлены из соединений сульфида цинка ZnS, карбида кремния SiC, а также в виде преобразователей инфракрасного излучения в видимое излучение. В фиолетовой части спектра расположено излучение светодиода на основе GaN. Твердый раствор $Al_xGa_{1-x}As$ используется в лазерных диодах, солнечных батареях, $GaAs_xP_{1-x}$ служит для изготовления светодиодов красного, желтого и оранжевого цвета. На твердых растворах $Ga_{1-x}Al_xAs$-GaAs, $Al_xGa_{1-x}Sb$-GaSb строят эффективные гетеропереходы. Инжекционные гетеропары InP-$Ga_xI_{1-x}As_{1-y}P_y$ перспективны для использования в волоконно-оптических линиях связи (ВОЛС), транзисторах и ИМС.

Фосфид галлия широко используется в производстве светодиодов, причем внешний квантовый выход достигает $\eta = 12\%$ при излучении с длиной волны $\lambda = 0{,}69$ мкм. Для управления цветом излучения производят контролируемое легирование фосфида галлия примесями.

В производстве светодиодов ультрафиолетового, фиолетового и голубого цветов используются такие широкозонные полупроводниковые материалы как сульфид кадмия CdS, селенид кадмия CdSe, оксид цинка ZnO, сульфид цинка ZnS. Эти полупроводники являются прямозонными и имеют высокий квантовый выход излучательной рекомбинации, однако они существуют только в виде полупроводников *n*-типа, так как легирование не

может изменить тип проводимости из-за эффекта самонасыщения. Поэтому создать *p*-*n*-переход в таких полупроводниках невозможно. А так как для инжекции в эти полупроводники необходимы дырки, то возникает необходимость формировать либо гетеропереход, либо барьер Шоттки. Формирование гетероперехода в данном случае технологически проблематично. Поэтому технологичнее сформировать контакт металл-полупроводник, причем необходимо выбрать металлы с дырочной проводимостью (например, ванадий, хром, марганец, см. гл. II) для инжекции положительных носителей.

На рис. 3.46, *а* показана основная структура преобразователя, в

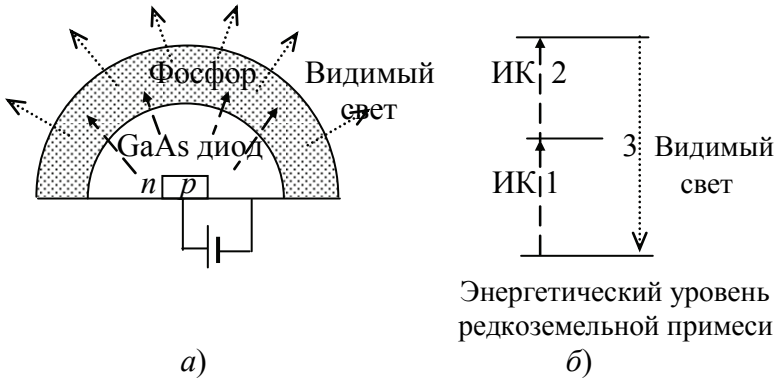


Рис. 3.46 – Преобразователь, в котором излучение GaAs-диода возбуждает испускание видимого света фосфором (*а*) и энергетические уровни преобразования (*б*)

котором инфракрасное излучение светодиода на основе арсенида галлия GaAs поглощается слоем фосфора с добавками редкоземельных элементов, например иттербия (Yb^{3+}) или эрбия (Er^{3+}). Характеристики преобразования зависят от процесса последовательного поглощения двух фотонов в инфракрасной области спектра, с последующим испусканием одного фотона в видимой области спектра (рис. 3.46, *б*).

Важнейшим и широко используемым свойством карбида кремния является его способность к *люминесценции* в видимой части спектра. Из-за большой запрещенной зоны карбид кремния эффективно используют для изготовления светодиодов видимого спектра излучения, от красного до фиолетового. Достоинством таких светодиодов является высокая стабильность параметров.

Излучательная рекомбинация носителей заряда, обусловленная примесями, наблюдается в виде *фотолюминесценции* или *электролюминесценции* при инжекции носителей через *p*-*n*-переход. Для этого используют полированные с двух сторон кристаллы гексагонального карбида кремния различных политипов. Монокристаллы, легированные азотом, приобретают электронный тип электропроводности. Вводя методом диффузии различные акцепторные примеси, создают инжекционные светодиоды, максимум излучения которых лежит в различных частях видимой области спектра. Для повышения эффективности электролюминесценции используют метод последовательной диффузии

различных примесей, например, вначале проводят диффузию алюминия, а затем – бора.

Материалы для светодиодов инфракрасного диапазона. Первые образцы инфракрасных светодиодов создавались на основе *n*-GaAs, в котором *p*-область образовывалась диффузией цинка Zn. Большое число дефектов в такой структуре возникало с одной стороны, вследствие сильного легирования *n*-области (n^+-*p*-переход) для достижения достаточного *коэффициент инжекции* γ, и с другой стороны – сильного легирования *p*-области (для увеличения скорости излучательной рекомбинации). Поэтому внешний квантовый выход $\eta_{внш}$ не превышал 1…2%.

Значительно бóльшим внешним квантовым выходом $\eta_{внш}$ (до 20…28%) обладают наиболее распространенные в настоящее время светодиоды, создаваемые эпитаксиальным наращиванием арсенида галлия GaAs *p*-типа, легированного кремнием (Si), на *n*-тип GaAs. Это обусловлено тем, что в области, сильно легированной кремнием, через примесные уровни излучаются кванты света с энергией 1,31…1,34 эВ. Эта энергия меньше ширины запрещенной зоны E_g = 1,43 эВ нелегированного арсенида галлия, что уменьшает самопоглощение при выводе через *n*-область.

Были также изучены светодиоды из арсенида галлия [4], изготовленные методом жидкостной эпитаксии на подложке *n*-типа, предварительно легированной теллуром в качестве донора до концентрации 10^{22} м$^{-3}$. Эпитаксиальная *p*-область была легирована германием в качестве акцептора до концентрации 10^{24} м$^{-3}$. Особенностью таких светодиодов было то, что при малых токах почти все излучение генерировалось в *n*-области в полосе с бóльшей энергией (в максимуме 1,42 эВ), а при больших токах – в полосе с мéньшей энергией (в максимуме 1,33 эВ) в p^+-области. Таким образом, спектр излучения таких светодиодов можно эффективно перестраивать с помощью изменения силы тока.

Для изготовления инфракрасных светодиодов используются многие другие полупроводники, имеющие ширину запрещенной зоны менее 1,5 эВ. К ним относятся твердые растворы, в состав которых входят три или четыре элемента III и IV групп периодической системы.

Получение высокоэффективных источников излучения на λ = 1,3 мкм и λ = 1,55 мкм для ВОЛС (волоконно-оптических линий связи) стало возможным благодаря полученным недавно данным по внутреннему квантовому выходу четверных твердых растворов. Внутренний квантовый выход в них не ниже, чем в арсениде галлия, а метод газофазной эпитаксии позволяет получать изопериодические ненапряженные структуры InGaAsP. Эффективность люминесценции таких структур намного больше тройных растворов GaAsP благодаря уменьшению скорости безызлучательной рекомбинации.

Увеличению внутреннего квантового выхода также способствует гетерогенная структура из твердых растворов, в широкозонных микроучастках которой осуществляется поглощение с последующей

люминесценцией в узкозонных микроучастках. При этом число поглощенных фотонов значительно уменьшается по сравнению с гомогенными полупроводниками. Специально разработанные способы эпитаксии позволяют компенсировать механические напряжения в структурах, что дает возможность увеличить в несколько раз внешний квантовый выход рекомбинационного излучения по сравнению с тройными соединениями GaAsP. Это связано с резким уменьшением дефектов в кристалле, к которым так чувствителен квантовый выход.

В еще бóльшей степени указанные проблемы решаются применением гетеропереходов (см. п. 3.3.2.3), например, GaAs-$Ga_{1-x}Al_xAs$, которые характеризуются различной шириной запрещенных зон *p*- и *n*-областей. Значение коэффициента инжекции γ, близкое к 1, достигается использованием в качестве области излучательной рекомбинации полупроводника GaAs с шириной запрещенной зоны E_{gp} мéньшей, чем ширина запрещенной зоны E_{gn} инжектирующей области $Ga_{1-x}Al_xAs$. Соотношение $E_{gp} < E_{gn}$ обеспечивает отсутствие самопоглощения в *n*-области при выводе излучения через нее.

3.3.5. Полупроводниковые лазеры

Полупроводниковые лазеры представляют собой совершенно особый класс приборов, поэтому остановимся на них несколько подробнее. Есть целый ряд важных причин такого внимания к ним:

- технология и свойства полупроводниковых материалов разработаны лучше, чем любого другого семейства материалов;
- лазерный эффект возникает за счет инжекции носителей заряда, поэтому полупроводниковыми лазерами легко управлять электрическим способом;
- имеют высокий КПД;
- работают при низких напряжениях;
- миниатюрны;
- устойчивы к механическим воздействиям;
- имеют большой срок службы;
- технология их изготовления удобна для массового производства, т.е. потенциально они имеют низкую стоимость;
- они могут работать на тех длинах волн, на которых потери и дисперсия в оптическом волокне близки к минимальным.

Основная идея работы полупроводникового лазера состоит в использовании *стимулированной излучательной рекомбинации* в полупроводниках с прямыми переходами, которая и приводит в результате к лазерной генерации. В прямозонных полупроводниках вероятность перехода со дна зоны проводимости на потолок валентной зоны намного больше, чем в полупроводниках с непрямыми переходами. В кристаллах с непрямыми оптическими переходами при рекомбинации наряду с фотонами образуются и фононы. В этом случае излучательная рекомбинация происходит менее

интенсивно вследствие конкуренции между рекомбинациями двух разных типов.

Для лазеров нужен материал, во-первых, с множеством электронов в зоне проводимости, стремящихся ее покинуть, и, во-вторых, с множеством вакансий у потолка валентной зоны, готовых эти электроны принять. Очевидно, что однородный полупроводник для этих целей не подходит, потому что оба эти условия в нем нельзя выполнить одновременно. Известно, что большое количество электронов можно получить в вырожденном полупроводнике *n*-типа (см. параграф 3.3.2). Аналогично в вырожденном полупроводнике *p*-типа существует большое количество дырок. Поэтому можно ожидать, что если соединить их вместе, то в области *p*-*n*-перехода оба условия окажутся выполненными одновременно. В середине *p*-*n*-перехода существует область перекрытия, где велики концентрации, как электронов, так и дырок. Эта область называется активной, так как именно здесь происходит излучательная рекомбинация. Чтобы поддерживать этот процесс, необходим достаточно высокий уровень инжекции.

Как известно, в отсутствии смещения на *p*-*n*-переходе энергии Ферми *p*- и *n*-областей E_{F_n} и E_{F_p} находятся на одном уровне (рис. 3.47, *а*). При малых прямых смещениях ток через *p*-*n*-переход мал, а значит и уровень инжекции невелик, инверсии населенностей не будет и индуцированное излучение не возникает.

Для создания состояния с инверсией населенности на *p*-*n*-переход необходимо подать такое прямое смещение *U*, чтобы разность уровней Ферми превышала ширину запрещенной зоны, т. е. $eU = E_{F_n} - E_{F_p} > E_g$ (рис. 3.47, *б*). При некотором пороговом напряжении $U_{\text{П}}$ выполняется

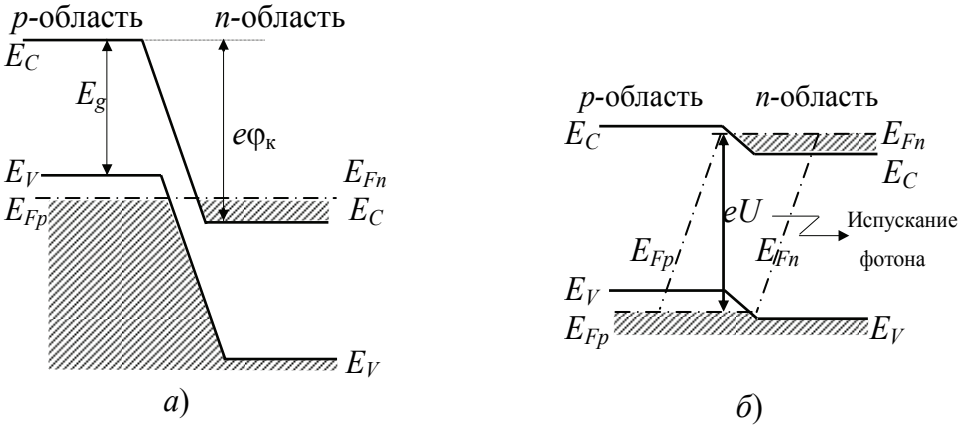


Рис. 3.47 – Вырожденный *p*-*n*-переход: *а* – при тепловом равновесии; *б* – при прямом смещении

условие генерации, и диод начинает излучать когерентный свет, близкий к монохроматическому.

Рассмотрим широко распространенную структуру лазера на *p*-*n*-переходе (рис. 3.48). Две боковые грани структуры скалываются и полируются перпендикулярно плоскости перехода. Две другие грани делаются шероховатыми для того, чтобы исключить излучение в

направлениях, не совпадающих с главным. Такая структура называется резонатором Фабри-Перо. Действие лазера основано на том, что при прямом

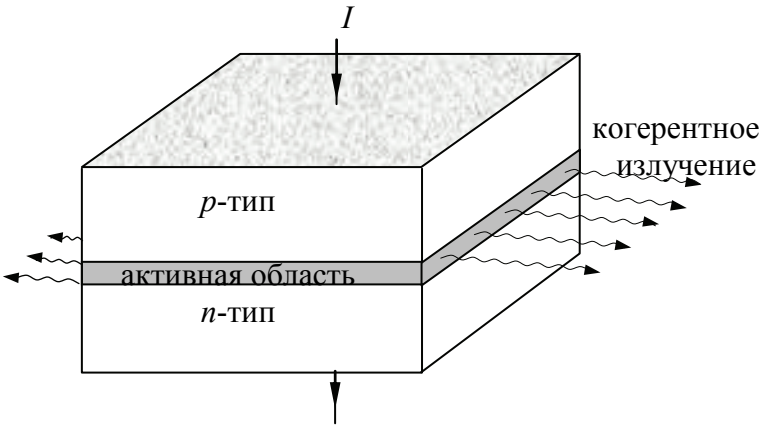


Рис. 3.48 – Основная структура полупроводникового лазера на *p*-*n*-переходе в виде резонатора Фабри-Перо

смещении электроны инжектируются в *p*-область базы, где происходит их излучательная рекомбинация с имеющимися там дырками. Для создания состояния с инверсионной населенностью в базе необходимо, чтобы инжекция электронов в *p*-область базы превышала инжекцию дырок в *n*-область, где рекомбинация не так эффективна (ток дырок из *p*- в *n*-область в данном случае целиком относится к потерям). Поэтому концентрация донорной примеси N_D в *n*-области должна значительно превышать концентрацию акцепторной примеси N_A в *p*-области, т. е. $n_n > p_p$. Вместе с тем, для увеличения вероятности процесса излучательной рекомбинации необходима большая концентрация дырок в валентной зоне базы, что достигается увеличением концентрации легирующей акцепторной примеси в базе. В лазере благодаря использованию оптического резонатора все световые волны, направленные под углом к оси активной области, сразу или после нескольких отражений отклоняются от оси резонатора и не усиливаются. Усиливаются только те волны, которые многократно проходят активную область вблизи границы *p*-области. Излучателем является узкая часть *p*-области, прилегающая к *p*-*n*-переходу. Например, толщина излучающей части *p*-области лазерного диода на арсениде галлия порядка 2 мкм.

Набор полупроводниковых материалов, на базе которых создаются лазеры, постоянно растет.

В связи с тем, что GaAs стал первым лазерным полупроводниковым материалом, наиболее интенсивное развитие получили родственные ему твердые растворы элементов III и V групп. Диоды, изготовленные на базе соединений $A^{III}B^{V}$, таких как PbS, PbTe, PbSe и родственных им твердых растворов, также обладают «лазерными свойствами». К наиболее важным из сложных соединений, обладающих подобными свойствами, относятся твердые растворы $Al_xGa_{1-x}As_ySb_{1-y}$ и $Gaxln_{1-x}As_ySb_{1-y}$.

Хотя теоретически КПД лазерных диодов может приближаться к 100%, практически он намного меньше. Главная причина этого – необходимость в сильном легировании *p*- и *n*- областей перехода. Одновременно с введением примесей в полупроводнике образуется большое количество дефектов в активной области, что ведет к значительным потерям излучения. Растет также и поглощение на свободных носителях. Вследствие этих причин для возникновения индуцированного излучения требуется большая плотность тока через *p*-*n*-переход, которая приводит к деградации лазерных диодов, а также к необходимости создания громоздких систем охлаждения. Поэтому суммарный квантовый выход при комнатной температуре составляет не больше 2…3%; пороговые плотности токов порядка 10^8 А/м^2. При работе в режиме максимальной мощности излучения могут возникать трещины в активной области лазера и разрушаться зеркала резонатора, которыми чаще всего являются грани сколов.

Указанные недостатки устранены в *p*-*i*-*n*-диоде с гетеропереходами. В таком лазере средним активным слоем *i* служит материал с мѐньшей шириной запрещенной зоны, являющийся собственным полупроводником, а *p*- и *n*-эмиттерами – материалы с бóльшей шириной запрещенной зоны, обладающие, соответственно, дырочной и электронной проводимостью. Материалами для *p*-, *i*- и *n*-областей являются тройные соединения (например, p^+-$Al_xGa_{1-x}As$, *i*-$Al_yGa_{1-y}As$, *n*-$Al_zGa_{1-z}As$). Различная ширина запрещенной зоны получается за счет изменения соотношения компонентов *x*, *y*, *z*.

Как отмечалось ранее, гетеропереходы характеризуются большим коэффициентом инжекции (отношением инжекционного тока из широкозонного полупроводника в узкозонный к обратному току). При включении диода в прямом направлении инверсия населенности достигается высоким уровнем инжекции электронов и дырок в узкозонную часть, где они излучательно рекомбинируют. Так как инверсионная населенность достигается за счет неравновесных носителей тока, нет необходимости в высоком уровне легирования активной области. В результате этого резко снижаются потери излучения в активной области, а, следовательно, и пороговая плотность тока, которая в настоящее время для гетеролазера, работающего при комнатной температуре, порядка $9 \cdot 10^6$ А/м^2. Внешняя квантовая эффективность у лучших образцов достигает 70%. Ширина активной области в гетеролазере может быть много больше ширины активной области лазера на гомо-*p*-*n*-переходе, поэтому в гетеролазере могут быть получены гораздо бóльшие интенсивности излучения. Снижение пороговых токов обеспечивает снижение рабочего тока, что приводит к уменьшению деградационных эффектов.

3.3.6. Фотодетекторы

Зависимость проводимости полупроводников от освещения используют для создания различных фоточувствительных приборов для различных длин волн (от инфракрасных до ультрафиолетовых). Использование твердых

растворов нескольких полупроводниковых материалов и введение нужных примесей позволяет изготавливать приборы с необходимыми спектром поглощения, инерционностью и другими характеристиками.

Фотодетекторы представляют собой полупроводниковые приборы, которые могут детектировать оптические сигналы посредством протекающих в них электронных процессов. Расширение спектрального диапазона излучения когерентных и некогерентных источников света в дальнюю инфракрасную область, с одной стороны, и в ультрафиолетовую – с другой, привело к увеличению потребности в быстродействующих чувствительных фотодетекторах. В общем случае в фотодетекторе протекает три основных процесса:

1) генерация носителей под действием внешнего излучения;

2) перенос носителей и их умножение за счет того или иного механизма усиления тока, характерного для данного полупроводника;

3) взаимодействие тока с внешней цепью, обеспечивающее получение выходного сигнала.

Фотодетекторы играют незаменимую роль в волоконно-оптических системах связи ближнего инфракрасного (ИК) диапазона (0,8...1,6 мкм). Они преобразуют изменение интенсивности принятого излучения в электрические сигналы, которые соответствующим образом усиливаются и подвергаются последующей обработке. Фотодетекторы, предназначенные для таких систем, должны обладать высокой *чувствительностью* в рабочем диапазоне волн или частот, высоким *быстродействием* и низким *уровнем шумов*. Кроме того, они должны иметь малые размеры, низкие управляющие напряжения или токи, а также высокую надежность в работе.

Фотодетекторы по физическим принципам работы подразделяются на три основных типа:

- в фоторезисторах происходит только увеличение проводимости за счет генерации светом дополнительных фотоэлектронов и фотодырок – эффект фотопроводимости (эффект инерционен, не используется на высоких частотах модуляции);

- в фотодиодах и фотоэлементах генерированные светом носители разделяются внутренним электрическим полем *p*-*n*-перехода – фотогальванический эффект (фотоэлементы обладают хорошими пороговыми свойствами (малошумящие); фотодиоды – высокой скоростью ответа).

- в фототранзисторах, лавинных и инжекционных фотодиодах происходит умножение разделенных фотоносителей за счет того или иного механизма усиления фототока, характерного для данного устройства.

Рассмотрим более подробно работу фотодиода как элемента, широко применяющегося в оптоволоконной связи, к которому приложено смещение в обратном направлении. Выбор знака смещения связан с тем, что для повышения чувствительности необходимо получить большую величину относительного изменения тока. Действительно, если удвоение тока в один микроампер можно легко заметить, то изменения на один микроампер при

силе тока в один миллиампер зафиксировать очень трудно. Если фотоны с подходящей длиной волны (фотоактивный для данного полупроводника свет) попадают в область с проводимостью *p*-типа, они генерируют электроны, которые являются для этой области неосновными носителями, дрейфующими через *p*-*n*-переход под действием смещения. Так как для более эффективной работы *p*-*n*-переход должен располагаться в непосредственной близости к освещаемой поверхности, используется чувствительный фотодетектор с тонким слоем полупроводника *p*-типа, нанесенном на подложку из полупроводника *n*-типа.

Фотоактивным для данного полупроводника является свет, энергия фотонов которого больше или равна ширине запрещенной зоны. Однако энергия фотонов поглощаемого света не может более чем вдвое превышать ширину запрещенной зоны, так как при этом свет поглощается вблизи поверхности, где генерированные носители сразу рекомбинируют через поверхностные состояния. Поэтому полупроводник с шириной запрещенной зоны E_g будет поглощать свет с энергией максимума излучения $h\nu$ в соответствии с неравенством:

$$E_g \le h\nu \le 2E_g . \qquad (3.142)$$

Этот критерий можно разбить на две части:

- $h\nu \ge E_g$ – жесткий критерий;
- $h\nu \le 2E_g$ – мягкий критерий.

Поглощаемое излучение удобно характеризовать *квантовой интенсивностью* Ф – количеством фотонов, падающих на 1 м^2 поверхности полупроводника за одну секунду.

При освещении *p*-*n*-перехода, из-за наличия контактной разности потенциалов $\varphi_к$, генерированные светом избыточные носители тока разделяются электрическим полем *p*-*n*-перехода. Электроны переносятся электрическим полем в *n*-область, а дырки – в *p*-область. Так как из обеих областей через *p*-*n*-переход уходят только неосновные носители, то можно считать, что генерируемые светом носители заряда увеличивают обратный ток *p*-*n*-перехода, так как именно он образуется за счет неосновных носителей. Вольтамперную характеристику *p*-*n*-перехода при освещении можно записать в виде алгебраической суммы обычного тока I (см. ф-лу (3.106)) и фототока $I_ф$, создаваемого носителями, возбужденными светом:

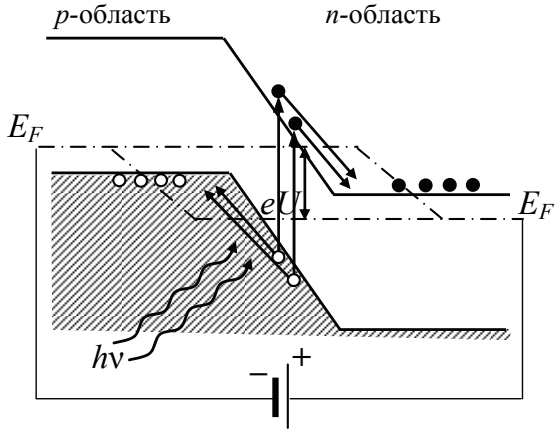
$$I = I_s\left(\mathrm{e}^{\frac{eU}{kT}} - 1\right) - I_ф . \qquad (3.143)$$

В фотодиодном режиме (обратное смещение, $U < 0$) потенциальный барьер возрастает (рис. 3.49), в формуле (3.143) величина $\mathrm{e}^{\frac{eU}{kT}} << 1$ и ею можно пренебречь, тогда ток через переход будет состоять из обратного тока, который протекает в отсутствие излучения (темновой ток), и фототока:

$$I = I_s + I_ф . \qquad (3.144)$$

Вольтамперная характеристика фотодиода для световых потоков с разной интенсивностью показана на рис. 3.50.

Рассмотрим, какие физические свойства полупроводниковых материалов можно использовать для улучшения наиболее важных характеристик фотодиодов: *квантовой эффективности* и *скорости фотоответа*.



● – электроны; ○ – дырки

Рис. 3.49 – Обратносмещенный *p-n*-переход в фотодиодном режиме

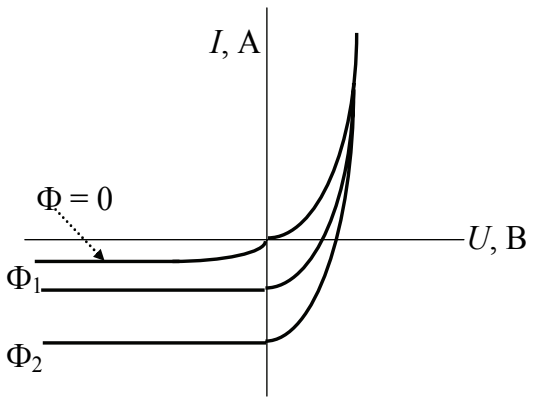


Рис. 3.50 – ВАХ фотодиода для световых потоков разной интенсивности ($Ф_2 > Ф_1$)

Квантовая эффективность представляет собой число фотогенерированных электронно-дырочных пар, отнесенное к числу падающих фотонов:

$$\eta = \frac{I_ф h\nu}{eP_{опт}} \ , \tag{3.145}$$

где $I_ф$ – фототок, обусловленный поглощением падающего оптического излучения с мощностью $P_{опт}$ и длиной волны λ (соответствующей энергии $h\nu$).

Сравнительным критерием качества является *чувствительность* фотодиода, которая определяется как отношение фототока к оптической мощности

$$R = \frac{I_ф}{P_{опт}} = \frac{\eta e}{h\nu} = \frac{\eta\lambda e}{hc} = \frac{\eta\lambda}{1,24} \ , \tag{3.146}$$

где λ дана в микрометрах.

Таким образом, для данной квантовой эффективности чувствительность растет с ростом λ. Для идеального фотодиода ($\eta = 1$) $R = \lambda/1,24$ (А/Вт).

Одним из основных факторов, определяющих квантовую эффективность, является коэффициент поглощения α. Вследствие того, что α сильно зависит от длины волны, для данного полупроводника область длин волн, в которой можно получить заметный фототок, ограничена. ***Длинноволновая граница*** определяется шириной запрещенной зоны полупроводника в соответствии с уравнением $\lambda_г = \frac{hc}{E_g} = \frac{1,24}{E_g(\text{эВ})}$ и

приблизительно равна 1,7 мкм для Ge и 1,1 мкм для Si. *Коротковолновая граница* фотоответа обусловлена тем, что в области коротких волн коэффициент α слишком велик ($\geq 10^7$ м$^{-1}$) и излучение поглощается в непосредственной близости к поверхности, где времена поверхностной рекомбинации слишком малы. Вследствие этого фотоносители рекомбинируют раньше, чем доходят до *p*-*n*-перехода.

В ультрафиолетовой и видимой областях спектра высокой квантовой эффективностью обладают диоды со структурой металл-полупроводник. В ближнем ИК-диапазоне (на длинах волн ~ 0,8…0,9 мкм) в кремниевых диодах с просветляющим покрытием может быть достигнута квантовая эффективность около 100%. В области длин волн 1,0…1,6 мкм высокой квантовой эффективностью характеризуются фотодиоды на Ge, а также на основе тройных (например, GaInAs) и четверных (например, GaInAsP) соединений элементов III и V групп. Для обеспечения высокоэффективной работы в длинноволновой области спектра фотодиоды охлаждают до низких температур (например, до 77 К).

Скорость фотоответа ограничивается такими факторами:

- временем диффузии $\tau_б$ неосновных носителей, возбужденных светом в объеме полупроводника, по направлению к области объемного заряда;

- временем пролета $t_{пр}$ через обедненную область:

- постоянной времени $\tau_{RC} = C_з \cdot R$, где R – сопротивление нагрузки, $C_з$ – емкость объемного заряда.

Носители, генерированные за пределами обедненной области, диффундируют к *p*-*n*-переходу, что существенно увеличивает время задержки фотоответа. Чтобы уменьшить время диффузии носителей, необходимо сформировать *p*-*n*-переход вблизи освещаемой поверхности. Поэтому толщину базы фотодиода d выполняют меньше диффузионной длины L_n, чтобы бóльшая часть генерируемых светом неосновных носителей доходила до *p*-*n*-перехода.

Скорость фотоответа определяет одну из основных характеристик работы фотодиода – *максимальную частоту* f_{max}. Эта частота равна

$$f_{max} = \frac{1}{2\pi\tau_б}, \qquad (3.147)$$

и определяется временем выхода носителей из базы $\tau_б$

$$\tau_б = \frac{d^2}{2D_n} = \frac{d^2 e}{2kT\mu_n}\ , \qquad (3.148)$$

где d – толщина базы фотодиода; $D_n = \frac{kT}{e}\mu_n$ – коэффициент диффузии электронов в данном полупроводнике. При этой частоте сигнал уменьшается относительно низкочастотного уровня до уровня 0,7. Так как коэффициент диффузии электронов D_n обычно больше коэффициента диффузии дырок D_p, то освещаемой частью фотодиода выбирается полупроводник *p*-типа. Таким способом достигается уменьшение времени выхода $\tau_б$ фотоэлектронов из

p-базы и увеличение максимальной рабочей частоты f_{max}. Время пролета $t_{пр}$ и постоянную времени τ_{RC} можно уменьшать различными способами, но эти времена в фотодиоде малы по сравнению со временем выхода носителей из базы $\tau_б$ и существенно не влияют на быстродействие диода.

Повысить быстродействие фотодиодов, сохраняя при этом высокую чувствительность, возможно при использовании более сложных структур.

Одним из преимуществ фотодиода на основе гетероперехода состоит в том, что его квантовая эффективность не зависит критическим образом от расстояния, на котором расположен переход от поверхности, так как материал с более широкой шириной запрещенной зоны можно использовать в качестве «окна», пропускающего оптическое излучение. Кроме того, гетеропереход дает уникальную возможность выбора такого сочетания материалов, которое обеспечивает оптимальные для заданной длины волны оптического сигнала значений квантовой эффективности и быстродействия.

Для того, чтобы получить гетеропереход с низкими токами утечки, необходимо обеспечить хорошее согласование постоянных решеток обоих полупроводников. Тройные полупроводниковые соединения $Al_xGa_{1-x}As$ (с прямой зоной при $x < 0{,}4$), выращенные в виде эпитаксиальных слоев на подложке GaAs, формируют гетеропереходы с совершенным согласованием решеток (с постоянной решетки 5,653 Å). Эти гетеропереходы играют важную роль в оптоэлектронных приборах, работающих в спектральном диапазоне 0,65…0,85 мкм.

Для больших длин волн (1…1,6 мкм) могут быть использованы тройные соединения, такие как $Ga_{0,47}In_{0,53}As$ (E_g = 0,73 эВ) и четверные соединения, такие как $Ga_xIn_{1-x}As_yP_{1-y}$ (например, $Ga_{0,27}In_{0,73}As_{0,63}P_{0,37}$ с E_g = 0,95 эВ), которые имеют совершенное согласование решеток с подложкой InP (постоянная решетки 5,8686 Å). Типичные значения квантовой эффективности фотодиодов на гетероструктурах составляют 55…60%.

Фотодиоды со структурой металл-полупроводник (фотодиоды Шоттки) также позволяют повысить быстродействие до 10^{-11} с. Свет, проходя через тонкую прозрачную пленку из металла, поглощается сразу в области объемного заряда, поэтому время пролета обедненной области перехода металл-полупроводник $t_{пр}$ ничтожно мало. А малое удельное сопротивление металла еще более уменьшает сопротивление базы *R*, следовательно, и τ_{RC}. Фотоносителями в этом случае являются дырки, которые практически мгновенно рекомбинируют с электронами в металле, внося свой вклад в фототок. Фотодиоды с барьером Шоттки особенно эффективны в видимой и ультрафиолетовой областях спектра. Для этих спектральных областей коэффициент поглощения α в наиболее распространенных полупроводниках очень велик (~ 10^5 см$^{-1}$ и более) и эффективная глубина поглощения мала ($1/\alpha \approx 0{,}1$ мкм и менее). Поэтому оказывается возможным подобрать металл и просветляющее покрытие так, чтобы бóльшая часть падающего излучения

поглощалась вблизи поверхности полупроводника (например, Au-*n*-Si-фотодиод с толщиной пленки золота 107 Å).

В *p*-*i*-*n*-структуре поглощение света происходит в *i*-области, которая имеет удельное сопротивление в 10^7 раз больше, чем сопротивление *p*- и *n*-областей. Это позволяет сделать электрическое поле *E* перехода достаточно большим, и все оно будет сосредоточено в *i*-области. Сильное электрическое поле быстро разделяет генерированные фотоносители и время $t_{пр} = d/\mu E$ становится очень малым. Большая толщина *i*-области обеспечивает малое значение $C_з$, что приводит к уменьшению τ_{RC}. Все это позволяет повысить частоту преобразования оптического сигнала до нескольких ГГц.

Одним из новых направлений в области конструирования фотодиодов есть создание *инжекционных* фотодиодов [5-8] – детекторов с внутренним усилением на основе *p*-*i*-*n*-структуры с длинной базой (толщина базы *d* превышает диффузионную длину носителей заряда *L*). Под *инжекционным усилением* имеют в виду такое изменение уровня инжекции носителей заряда, которое в десятки и сотни раз превышает действие первичного возбуждающего фактора (света). Чувствительность «длинных» диодов к оптическому излучению намного больше, чем у соответствующих безынжекционных аналогов. Спектральная чувствительность инжекционных фотодиодов распространяется в область примесного поглощения, т. е. на их основе могут быть созданы фотоприемники для ИК-области спектра.

Биполярные и униполярные (МДП) транзисторы, которые так же, как и диоды, могут выполнять функции фотодетекторов, называются фототранзисторами. Транзисторный эффект обеспечивает высокое усиление фоточувствительности, но по сравнению с фотодиодом, фототранзистор более сложен в изготовлении и уступает ему в быстродействии.

В настоящее время наиболее перспективными материалами для фотодетекторов, в том числе и фотодиодов, являются тройные и четверные твердые растворы на основе полупроводниковых соединений $A^{III}B^{V}$, оптические характеристики которых более подробно обсуждались в параграфе 3.3.4.

Тем не менее, фотодетекторы на простых полупроводниковых материалах не утратили своей роли.

Такой полупроводниковый материал как германий для видимого света не прозрачен, однако для инфракрасных лучей он относительно прозрачен при длине волны более 1,8 мкм. Оптические свойства германия используют при изготовлении фототранзисторов и фотодиодов. Рабочий интервал температур у германиевых приборов –60...+70ºС. Наиболее выгодно германиевые приборы использовать при низких температурах.

Селен также находит широкое применение при производстве фотоэлектрических приборов: фоторезисторов и фотоэлементов. Это связано с тем, что спектральные характеристики селеновых фотоэлектрических приборов почти полностью совпадают со спектральной характеристикой глаза. После прекращения освещения селена требуется значительное время,

чтобы электропроводность его приняла прежнее значение, т. е. *время релаксации фотопроводимости* селена велико. Селен почти прозрачен в инфракрасной области спектра. Это позволяет использовать его в качестве фильтров и защитных покрытий в приборах инфракрасного диапазона.

ВОПРОСЫ.

1. Что такое собственные полупроводники? Какими свойствами они обладают? Основные параметры собственных полупроводников.
2. Что такое примесные полупроводники? Какими свойствами они обладают? Какие примеси в полупроводниках являются донорами, а какие акцепторами?
3. Какие из дырок обладают большей энергией: в глубине валентной зоны или у потолка валентной зоны?
4. Почему при вхождении в кристаллическую решетку полупроводника энергия ионизации примесного атома, играющего роль донора, уменьшается в десятки раз?
4. Как концентрация и удельная проводимость примесных полупроводников зависит от температуры?
5. Каким соотношением связаны между собой концентрация электронов и дырок в невырожденном полупроводнике при термодинамическом равновесии?
6. Какие основные механизмы рассеяния ограничивают подвижность носителей заряда в полупроводниках?
7. Назовите основные механизмы поглощения света в полупроводниках. Какие из них являются фотоактивными?
8. В чем различия между прямыми и непрямыми оптическими переходами?
9. Как с помощью эффекта Холла определить тип проводимости полупроводника?
10. Как изменяется ширина запрещенной зоны кремния и германия при понижении температуры?
11. Почему для изготовления большинства полупроводниковых приборов требуются монокристаллические материалы и не могут быть использованы поликристаллические образцы?
12. Объясните, почему при одинаковом содержании легирующих примесей поликристаллический кремний обладает гораздо более высоким удельным сопротивлением, чем монокристаллический материал?
13. Какие преимущества кремния обусловливают его широкое применение при изготовлении планарных транзисторов и интегральных микросхем?
14. Что такое электронно-дырочный переход (*p*-*n*-переход)? Какие электронно-дырочные переходы называют симметричными, а какие несимметричными?
15. Построить энергетические диаграммы для симметричного резкого *p*-*n*-перехода для следующих случаев: а) внешнее напряжение отсутствует; б) прямое смещение перехода; в) обратное смещение перехода.

16. Как и почему изменяется высота потенциального барьера *p*-*n*-перехода с изменением температуры и с изменением концентрации примесей в прилегающих к переходу областях?
17. Нарисуйте энергетические диаграммы контактов металл-полупроводник для двух случаев: а) $\chi_м > \chi_п$; б) $\chi_м < \chi_п$. Какими физическими свойствами обладают такие контакты?
18. Что такое гетеропереход? Объяснить, почему в *p*-*n*-гетеропереходах инжекция основных носителей происходит всегда из широкозонного в узкозонный полупроводник?
19. Какие полупроводниковые материалы наиболее перспективны для создания гетеропереходов со свойствами идеального контакта?
20. Какие преимущества имеют многокомпонентные твердые растворы при создании высококачественных гетеропереходов? Приведите примеры таких гетеропереходов.
21. Физические принципы работы инжекционных лазеров и светодиодов. Какие материалы используются для изготовления инжекционных лазеров и светодиодов?
22. Объяснить физические принципы работы фотодетекторов. Какие материалы используются для изготовления фотодиодов?

ЗАДАЧИ.

1. Найти положение уровня Ферми в собственном полупроводнике (германий) при 300 К, если известно, что ширина его запрещенной зоны E_g=0,665 эВ, а эффективные массы для дырок и электронов соответственно равны $m_p^* = 0{,}388m_0$; $m_n^* = 0{,}55m_0$, где m_0 – масса свободного электрона. (Ответ: $E_F = 0{,}326$ эВ выше потолка валентной зоны)
2. Определить вероятность заполнения электронами энергетического уровня, расположенного на 10 kT выше уровня Ферми. Как изменится вероятность заполнения этого уровня электронами, если температуру увеличить в два раза? (Ответ: $f(E_{T1}) = 4{,}54\cdot10^{-5}$; $f(E_{T2}) = 6{,}74\cdot10^{-3}$)
3. Оценить тепловую и дрейфовую скорости электронов при T = 300 К в германии *n*-типа с концентрацией доноров $N_D = 10^{22}$ м$^{-3}$, если плотность тока $j = 10^4$ А/м^2, а эффективная масса электронов проводимости $m_n = 0{,}12m_0$. (Ответ: $\langle\upsilon\rangle = 3{,}37\cdot10^5$ м/с; $\upsilon_{др} = 6{,}25$ м/с).
4. При напряженности электрического поля E = 100 В/м плотность тока через полупроводник $j = 6\cdot10^4$ А/м^2. Определить концентрацию электронов проводимости в полупроводнике, если их подвижность μ_n = 0, 375 м^2/(В·с). Дырочной составляющей тока пренебречь. (Ответ: $n = 10^{22}$ м$^{-3}$).
5. Образец кремния содержит в качестве примеси фосфор с концентрацией атомов $N_D = 2\cdot10^{20}$ м$^{-3}$. Какую нужно создать концентрацию атомов галлия в этом полупроводнике, чтобы тип проводимости изменился на противоположный, а удельное сопротивление стало равным ρ = 0,5 Ом·м?

При расчетах полагать, что подвижность дырок μ_p = 0, 05 м^2/(В·с). (Ответ: $N_A = 4{,}5\cdot10^{20}$ м$^{-3}$).

6. Образец кремния *n*-типа при температуре T_1 = 300 К имеет удельное сопротивление ρ = 0,05 Ом·м, причем концентрация электронов в нем не изменяется при нагревании до температуры T_2 = 500 К. Определить, во сколько раз изменяется концентрация неосновных носителей заряда в этом температурном диапазоне. При температуре T_1 подвижность электронов μ_n = 0,14 м^2/(В·с). Значение собственной концентрации носителей заряда $n_i = 7\cdot10^{15}$ м$^{-3}$. Коэффициент температурного изменения ширины запрещенной зоны $\alpha = 2{,}84\cdot10^{-4}$ эВ/К. (Ответ: в $5{,}2\cdot10^8$ раз).

7.* Определить длину волны излучения, возникающего при прямой межзонной рекомбинации в твердых растворах $Ga_{0,8}In_{0,2}As$ при температуре T = 300 К, если известно, что ширина запрещенной зоны твердого раствора $Ga_{0,5}In_{0,5}As$ равна 0,74 эВ. (Ответ: λ = 1,11 мкм. Изменение запрещенной зоны в зависимости от состава может быть аппроксимировано плавной кривой параболического вида: $E_g(x) = E_{g1}\cdot x + E_{g2}\cdot(1-x) - bx\cdot(1-x)$, где E_{g1} и E_{g2} – ширина запрещенных зон бинарных соединений, образующих твердый раствор; x – состав твердого раствора; b – эмпирический коэффициент нелинейности).

8. Имеется несимметричный германиевый *p-n*-переход с концентрацией примесей $N_D = 10^3\ N_A$, причем на каждые 10^8 атомов германия приходится один атом акцепторной примеси. Определить контактную разность потенциалов при температуре T = 300 К. Концентрацию атомов германия и ионизованных атомов принять равными $4{,}4\cdot10^{28}$ м$^{-3}$ и $2{,}5\cdot10^{19}$ м$^{-3}$ соответственно. (Ответ: $\varphi_к$ = 0,326 В).

9. Обратный ток насыщения *p-n*-перехода при комнатной температуре равен $I_0 = 10^{-14}$ А. При повышении температуры до 125 ° С обратный ток насыщения увеличился в 10^5 раз. Определить напряжение U на переходе при комнатной температуре и температуре 125 °С, если прямой ток через него I = 1 мА. (Ответ: при комнатной температуре U = 0,66 В; при температуре 125 °С U = 0,5 В).

10. Для изготовления светодиода использовалось тройное соединение $Al_xGa_{1-x}Sb$, у которого x = 0,505. Ширина запрещенной зоны определяется в соответствии с соотношением вида: $E_g(x) = (0{,}77 + 1{,}37x)$. Время выхода электронов из базы τ = 10^{-9} с. Определить ширину запрещенной зоны тройного соединения, длину волны электролюминесценции и предельную частоту работы светодиода. (Ответ: $E_g(x) = 1{,}46$ эВ; λ = 850 нм; f_{max} = 159 Мгц).

ГЛАВА IV. ДИЭЛЕКТРИКИ

4.1. ФИЗИЧЕСКИЕ ПРОЦЕССЫ В ДИЭЛЕКТРИКАХ

4.1.1. Основные понятия

4.1.1.1. Поляризация диэлектриков

Согласно зонной теории твердого тела диэлектрики – это вещества, у которых *запрещенная зона* настолько велика, что в нормальных условиях проводимость в них практически отсутствует. *Диэлектриками* называют материалы, основным электрическим свойством которых является способность к *поляризации* и в которых возможно существование электрического поля, вызванного внешними или внутренними причинами.

Реальный (технический) диэлектрик тем больше приближается к идеальному, чем меньше его удельная проводимость и чем слабее у него выражены виды поляризации, связанные с рассеиванием электрической энергии и выделением тепла.

Поляризация – это процесс, состоящий в том, что при внесении диэлектрика во внешнее электрическое поле в любом малом объеме диэлектрика ΔV возникает в результате той или иной поляризации молекул отличный от нуля суммарный дипольный электрический момент.

Количественной мерой поляризации диэлектрика служит вектор $\vec{P}$, называемый *поляризованностью* (вектором поляризации) и равный отношению электрического дипольного момента малого объема диэлектрика к этому объему ΔV:

$$\vec{P} = \frac{1}{\Delta V}\sum_{i=1}^{n}\vec{p}_i\,, \qquad (4.1)$$

где $\vec{p}_i$ – электрический дипольный момент i-й молекулы. Объем ΔV должен быть столь малым, чтобы в его пределах электрическое поле можно было считать однородным. В то же время число молекул в объеме ΔV должно быть достаточно велико, чтобы к ним можно было применять статистические закономерности ($n >> 1$). Единица измерения поляризованности - Кл/м^2.

В пределах малого объема ΔV все молекулы неполярного диэлектрика приобретают в электрическом поле одинаковые индуцированные моменты $\vec{p}$. Поэтому поляризованность неполярного диэлектрика в электрическом поле напряженности $\vec{E}$ равна

$$\vec{P} = n_0\vec{p} = n_0 q_д l_0\,, \qquad (4.2)$$

где $q_д$ – заряд диполя; l_0 – расстояние между зарядами в диполе; n_0 – концентрация молекул ($n_0 = n/\Delta V$).

Рассмотрим элемент объема диэлектрика площадью dS и длиной dl в электрическом поле $\vec{E}$ (рис. 4.1). Молекулярные диполи с зарядами $q_д$ и длиной l_0 расположены цепочкой вдоль поля. Если в одной цепочке n диполей, то общий электрический момент цепочки $p = nq_д l_0 = q_д l$. Обозначив через N число выходящих на поверхность S цепочек, получим

$$P = \frac{Np}{V} = \frac{Nq_д l}{V} = \frac{Nq_д}{S} = \frac{q'}{S} = \sigma', \qquad (4.3)$$

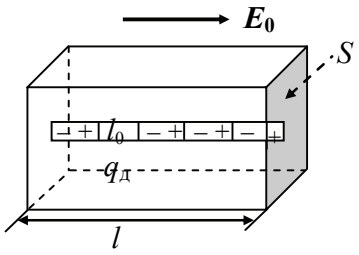


Рис. 4.1 – Диэлектрический параллелепипед с поляризованными цепочками атомов

где $q' = Nq_д$ – число выходящих на поверхность зарядов, а σ' – их поверхностная плотность. Таким образом, *поляризованность определяется поверхностной плотностью поляризационных зарядов.* Это объясняется тем, что в глубине диэлектрика положительные и отрицательные заряды взаимно компенсируются, а нескомпенсированные электрические заряды останутся только на поверхности диэлектрика. Эта формула поясняет также смысл единицы измерения поляризованности как отношение единицы заряда к единице поверхности (Кл/м^2).

Напряженность электрического поля $\vec{E}$ в диэлектрике, помещенном, например, в конденсатор, определяется разностью между полем, создаваемым зарядами на обкладках (без диэлектрика) $E_0 = \sigma_0/\varepsilon_0$ и полем, создаваемым поляризационными зарядами

$$E = E_0 - E' = (\sigma_0 - \sigma')/\varepsilon_0, \qquad (4.4)$$

где разность $\sigma_0 - \sigma'$ называют поверхностной плотностью свободных зарядов. Следовательно, при внесении в отключенный заряженный конденсатор диэлектрика поле внутри него уменьшается.

Опыт показывает, что в изотропных диэлектриках поляризованность прямо пропорциональна напряженности электростатического поля в данной точке

$$P = \sigma' = \kappa\varepsilon_0 E \ , \qquad (4.5)$$

где κ – безразмерный положительный коэффициент, называемый диэлектрической восприимчивостью. Подставив (4.5) в (4.4), получим выражение $E = E_0 - \kappa E$, откуда найдем значение поля внутри диэлектрика

$$E = \frac{E_0}{1+\kappa} = \frac{E_0}{\varepsilon} \ , \qquad (4.6)$$

где

$$\varepsilon = 1 + \kappa \ - \qquad (4.7)$$

относительная диэлектрическая проницаемость.

Электрическое смещение D в диэлектрике:

$$D = \varepsilon_0\varepsilon E = \varepsilon_0 E + \kappa\,\varepsilon_0 E = \varepsilon_0 E + P, \qquad (4.8)$$

где ε_0 – электрическая постоянная, равная $8{,}854 \cdot 10^{-12}$ Ф/м. Единица измерения электрического смещения D измеряется в Кл/м^2.

Формула (4.7) показывает, что относительная диэлектрическая проницаемость любого вещества больше единицы: диэлектрическая восприимчивость любого вещества конечна и положительна, поэтому $\varepsilon > 1$, и только для вакуума, где не происходит никакой поляризации, $\kappa = 0$, и, следовательно, $\varepsilon = 1$.

Диэлектрические проницаемости различных веществ значительно отличаются одна от другой. Величина ε вакуума равна единице, газов – близка к единице. Для большинства практически применяемых жидких и твердых диэлектриков ε – величина порядка нескольких единиц, реже – несколько десятков и иногда более 100. Наибольшую диэлектрическую проницаемость имеют некоторые сегнетокерамические материалы, диэлектрическая проницаемость ε которых в определенных условиях может достигать десятков тысяч.

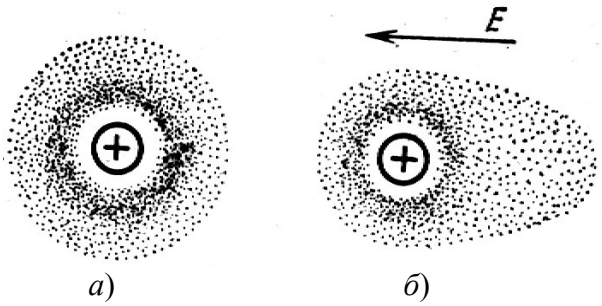
Для большинства диэлектриков, называемых *линейными диэлектриками*, поляризованность P прямо пропорциональна напряженности электрического поля E в данной точке диэлектрика (см. формулу (4.5))

У *нелинейных диэлектриков*, к которым, в частности, относятся сегнетоэлектрики, прямой пропорциональности между величинами P и E не наблюдается.

4.1.1.2. Виды поляризации диэлектриков

Известно много физических механизмов поляризации. Отметим из них лишь основные.

Электронная поляризация представляет собой упругое смещение и деформацию электронных оболочек атомов и ионов (рис. 4.2). Электронная поляризация устанавливается при наложении внешнего электрического поля за чрезвычайно короткое время (порядка 10^{-15} с). Она наблюдается у всех диэлектриков (независимо от возможного наличия в них и других видов поляризации) и практически не связана с потерями энергии до резонансных частот.



а) *б*)

Рис. 4.2 – Схематическое изображение электронной поляризации: *а* – в отсутствии внешнего поля; *б* – во внешнем поле

Поляризуемость частиц при электронной поляризации не зависит от температуры, а диэлектрическая проницаемость уменьшается с повышением температуры в связи с тепловым расширением диэлектрика и уменьшением числа поляризующихся частиц в единице объема.

Ионная поляризация наблюдается у диэлектриков ионного строения и возникает вследствие упругого смещения связанных ионов из положения

равновесия на расстояние, меньшее постоянной кристаллической решетки. С повышением температуры поляризованность возрастает, поскольку тепловое расширение, удаляя ионы друг от друга, ослабляет силы их взаимодействия. Время установления ионной поляризации около 10^{-15} с и она, так же как и электронная, практически не связана с потерями энергии и не зависит от частоты, вплоть до частот инфракрасного диапазона. Для λ =1 мкм (ИК) $t = 3{,}33 \cdot 10^{-15}$ с.

Ионная поляризация характерна для кристаллических диэлектриков ионной структуры с плотной упаковкой ионов. Величина поляризации с повышением температуры возрастает в результате ослабления упругих сил, действующих между ионами, так как расстояния между ними при тепловом расширении увеличиваются.

Дипольно-релаксационная поляризация заключается в повороте (ориентации) дипольных молекул в направлении внешнего электрического поля. Дипольные молекулы, находящиеся в хаотическом тепловом движении, ориентируются преимущественно в направлении внешнего электрического поля, создавая эффект поляризации диэлектрика. Она принадлежит к числу «медленных» видов поляризации; время установления ее значительно больше, чем время установления ионной или электронной поляризации.

С увеличением температуры молекулярные силы ослабляются, что должно усиливать дипольно-релаксационную поляризацию. В то же время возрастает энергия теплового движения молекул, что уменьшает ориентирующее действие поля. Поэтому величина дипольно-релаксационной поляризации с увеличением температуры сначала возрастает, пока ослабление молекулярных сил сказывается сильнее, чем возрастание хаотического теплового движения, затем, когда хаотическое движение становится интенсивнее, начинает падать. При снятии поля поляризация начинает исчезать из-за хаотического теплового движения молекул, и поляризованность P спадает со временем по экспоненциальному закону:

$$P(\tau) = P_0 e^{-\frac{t}{\tau_0}}, \qquad (4.9)$$

где P_0 – поляризованность в момент снятия напряжения; τ_0 – постоянная времени, называемая *временем релаксации*.

Обычно τ_0 для дипольно-релаксационной поляризации имеет порядок $10^{-6}\ldots10^{-10}$ с, следовательно, дипольно-релаксационная поляризация проявляется лишь на частотах ниже $10^{6}\ldots10^{10}$ Гц. Дипольно-релаксационная поляризация связана с потерями энергии, так как поворот диполей в направлении поля требует преодоления некоторого сопротивления и существенно зависит от температуры.

Ионно-релаксационная поляризация обусловлена смещением слабо связанных ионов под действием внешнего электрического поля на расстояние, превышающее постоянную кристаллической решетки. При этом виде поляризации возникают потери энергии; поляризация заметно усиливается с повышением температуры. Ионно-релаксационная

поляризация наблюдается в неорганических диэлектриках ионной структуры с неплотной упаковкой ионов. Поляризация этого типа носит замедленный характер и при высоких частотах несущественна.

Электронно-релаксационная поляризация возникает за счет возбужденных тепловой энергией избыточных «дефектных» электронов или «дырок». Она характерна для диэлектриков с высоким показателем преломления, большим внутренним полем и электронной проводимостью, например, рутила TiO_2, загрязненного примесями ниобия, кальция, бария. Электронно-релаксационной поляризации соответствует относительно высокое значение диэлектрической проницаемости, а также наличие максимума в температурной зависимости ε.

Миграционная поляризация – дополнительный механизм поляризации, проявляющийся в твердых телах неоднородной структуры; в диэлектриках с включениями проводящих частиц и т.д.; на границах раздела различных диэлектрических материалов; на границе диэлектрика и электрода и т. п. Миграционная поляризация сводится к переносу зарядов и накоплению их на поверхностях раздела материалов, имеющих различные параметры. Миграционная поляризация, как и дипольно-релаксационная, принадлежит к числу медленных (релаксационных) видов поляризации; проявляется при низких частотах и связана со значительными потерями.

Резонансная поляризация наблюдается в диэлектриках при световых частотах. Она зависит от физико-химических особенностей вещества, может относиться к собственной частоте электронов в атомах или ионах (при очень высоких частотах) или к характеристической частоте дефектных электронов (при более низких частотах).

Частотная зависимость составляющих комплексной диэлектрической проницаемости показана на рис. 4.3.

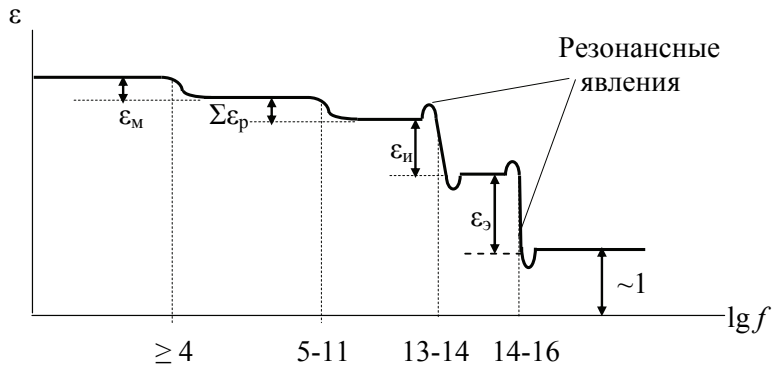


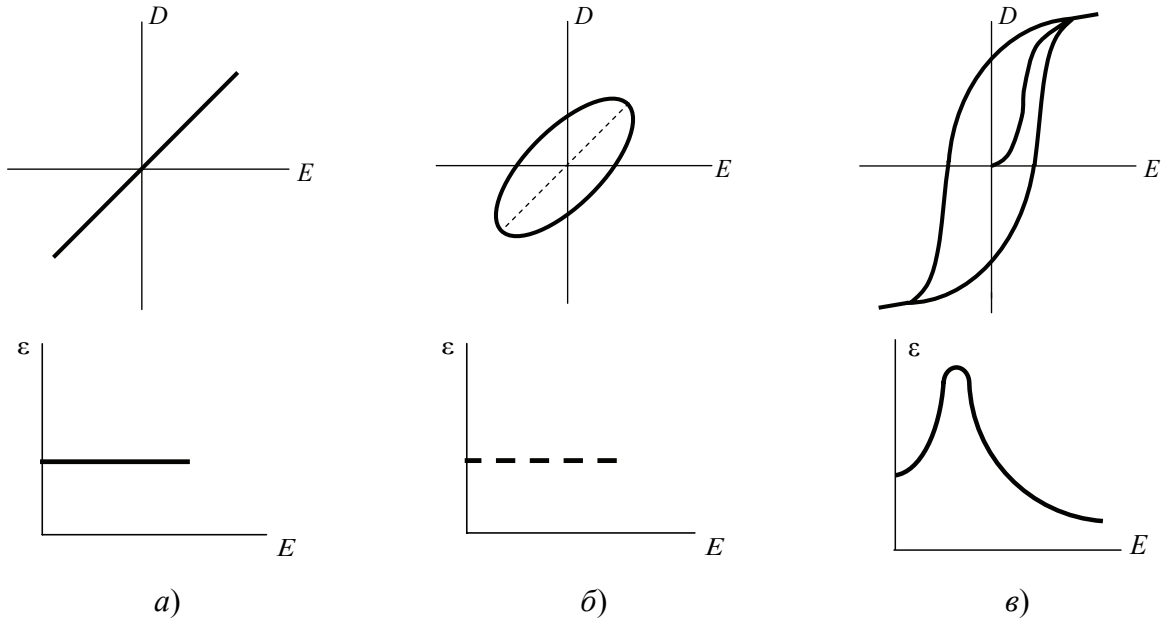
Рис. 4.3 – Частотная зависимость диэлектрической проницаемости: $\varepsilon_м$ – миграционная диэлектрическая проницаемость; $\Sigma\varepsilon_р$ – суммарная релаксационная поляризация; $\varepsilon_и$ – ионная поляризация; $\varepsilon_э$ – электронная поляризация

Самопроизвольная (*спонтанная*) *поляризация* существует только у группы твердых диэлектриков, обладающих такими же особенностями, как и сегнетова соль, а потому получивших название *сегнетоэлектрики*.

Особенности такой поляризации будут рассмотрены ниже.

4.1.1.3. Классификация диэлектриков по механизмам поляризации

В зависимости от влияния напряженности электрического поля переменной полярности на величину диэлектрической проницаемости, все диэлектрики подразделяются на *линейные* и *нелинейные*. Для линейных



а) *б*) *в*)

Рис. 4.4 – Зависимость электрического смещения и диэлектрической проницаемости от напряженности переменной полярности для: *а* – линейных диэлектриков; *б* – релаксационных механизмов поляризации; *в* – сегнетоэлектриков

диэлектриков с быстрыми видами поляризации (*электронной* и *ионной*) зависимость вектора электрической индукции D от напряженности электрического поля E переменной полярности показана на рис. 4.4, *а*, а для материалов с замедленными (релаксационными) видами поляризации на рис. 4.4, *б*. Площадь эллипса пропорциональна количеству энергии, рассеиваемой диэлектриком за один цикл изменения напряженности вследствие наличия у него релаксационных механизмов поляризации. Для нелинейных диэлектриков (сегнетоэлектриков) кривая зависимости D от E приобретает вид петли гистерезиса (рис. 4.4, *в*).

На рис. 4.4 приведены также соответствующие зависимости диэлектрической проницаемости ε от напряженности электрического поля E, показывающие различия линейных и нелинейных диэлектриков. Если емкость конденсатора с линейным диэлектриком зависит только от его геометрических размеров, и не меняется в случае приложенной разности потенциалов, то при нелинейном диэлектрике она становится управляемой электрическим полем. Поэтому нелинейные диэлектрики иногда называют активными (управляемыми) диэлектриками.

Если в основу классификации положить различия в механизмах поляризации, то все диэлектрики можно подразделить на несколько групп.

Неполярные диэлектрики. К этой группе относят диэлектрики, не содержащие электрических *диполей*, способных к переориентации во внешнем электрическом поле. Неполярным диэлектрикам свойственна в

основном электронная поляризация. Они применяются как электроизоляционные материалы в технике высоких и сверхвысоких частот. К ним относятся *полистирол*, *полиэтилен*, *политетрафторэтилен* (фторопласт-4), *бензол*, *воздух*, *парафин* и др.

Полярные диэлектрики. В эту группу входят диэлектрики, содержащие электрические диполи, которые способны к ориентации во внешнем электрическом поле. В полярных диэлектриках кроме электронной наблюдают и дипольно-релаксационную поляризацию. К ним можно отнести целлюлозу, *поливинилхлорид*, *политрифторхлорэтилен* (фторопласт-3), *полиметилметакрилат* (органическое стекло) и др.

Диэлектрики с ионной структурой. В эту группу входят твердые неорганические диэлектрики с электронной, ионной, ионно-релаксационной и электронно-релаксационной поляризацией. Здесь выделяют две подгруппы материалов в зависимости от величины потерь, расходуемых на поляризацию:

– Диэлектрики с электронными и ионными видами поляризации, при которых потери электрической энергии очень малы. К ним относят слюду, кварц, корунд (Al_2O_3), рутил (TiO_2) и др.

– Диэлектрики с ионно-релаксационной и электронно-релаксационной поляризацией, при которой имеются существенные потери электрической энергии. Например, неорганические стекла, керамика, микалекс и др.

Сегнетоэлектрики. В эту группу входят материалы, обладающие, прежде всего *спонтанной поляризацией*. К ним можно отнести сегнетову соль, титанат бария ($BaTiO_3$), титанат стронция ($SrTiO_3$) и др.

4.1.1.4. Поляризация сегнетоэлектриков

Сегнетоэлектриками называются вещества, обладающие спонтанной поляризацией, величина и направление которой может быть изменены с помощью внешнего электрического поля.

Сначала приведем важнейшие диэлектрические свойства сегнетоэлектриков.

1. Диэлектрическая проницаемость ε, которая у «классических» диэлектриков в большинстве случаев меньше 10 и редко превышает 100, у сегнетоэлектриков может достигать значений порядка $10^4 \ldots 10^5$. Например, для сегнетовой соли ε достигает 10000, для титаната бария ε порядка 6000…7000.

2. Диэлектрическая проницаемость ε не является постоянной величиной, а зависит от напряженности электрического поля E.

3. Функция $\varepsilon = \varepsilon(E)$ зависит от предыстории материала, вследствие чего наблюдается *диэлектрический гистерезис*.

4. Сегнетоэлектрические свойства сильно зависят от температуры. Для каждого сегнетоэлектрика существует определенная температура,

называемая точкой Кюри $T_К$, выше которой сегнетоэлектрик становится обычным диэлектриком.

При температуре, превышающей критическую температуру $T_К$, т.е. точку Кюри, диэлектрическая восприимчивость κ (κ = ε – 1) уменьшается с возрастанием температуры в соответствии с законом Кюри-Вейсса

$$\kappa = \frac{C}{T - T_К}, \qquad (4.10)$$

где C – константа. Температура Кюри для используемых в технике сегнетоэлектриков лежит в интервале от 300 до 400^0 С. Диэлектрическая восприимчивость κ из-за анизотропии кристалла является тензором. Для простоты будем считать, что электрическое поле $\vec{E}$ параллельно одной из трех главных осей тензора κ. В этом случае тензорные свойства диэлектрической восприимчивости не проявляются и ее можно считать скаляром.

6. Сегнетоэлектричество является свойством только твердых кристаллических веществ. Оно обнаруживается лишь у веществ 11-ти кристаллических групп (из 32 существующих), называемых полярными, или пироэлектрическими группами. Все кристаллы полярных (пироэлектрических) групп имеют особую полярную ось.

Впервые сегнетоэлектричество было исследовано на двух веществах: сегнетовой соли (от названия этой соли и произошел термин сенето-электрик) и дигидрофосфате калия.

7. Сегнетоэлектрики имеют доменную структуру. Домены представляют собой области размерами порядка 10^{-6} … 10^{-3} м, обладающие спонтанной (самопроизвольной) поляризацией. В отсутствие внешнего электрического поля направление электрических моментов у доменов равновероятно, поэтому поляризованность образца кристалла равна нулю. В принципе, если кристалл имеет очень маленькие размеры, то он может состоять только из одного домена. Но крупные образцы всегда разбиваются на большое количество доменов, поскольку однодоменное состояние энергетически не выгодно.

Внешнее электрическое поле изменяет направление электрических моментов доменов так, что происходит переориентация дипольных моментов по полю и это создает эффект очень сильной поляризации. Этим объясняются свойственные сегнетоэлектрикам сверхвысокие значения диэлектрической проницаемости. Следствием доменного строения сегнетоэлектриков является нелинейная зависимость их электрической индукции $\vec{D}$ от напряженности электрического поля $\vec{E}$ (рис. 4.5) и явление диэлектрического гистерезиса. (Зависимость поляризованности $\vec{P}$ от напряженности $\vec{E}$ также нелинейна). При воздействии слабого электрического поля зависимость $\vec{D}$ от $\vec{E}$ носит приблизительно линейный

характер (участок ОА). На этом участке преобладают процессы обратимого смещения доменных границ. В области более сильных полей смещение границ носит необратимый характер (участок АВ). При этом разрастаются домены с преимущественной ориентацией, у которых вектор спонтанной поляризации образует наименьший угол с направлением поля. При некоторой напряженности поля, соответствующей точке С на графике, все домены оказываются ориентированными по полю (монодоменное состояние), что соответствует области насыщения. Кривую ОАВС называют *начальной кривой поляризации* сегнетоэлектрика. В каждом, даже малом кристалле материала, встречаются напряжения, дислокации, примеси и другие дефекты. И при всех полях за исключением лишь очень слабых, стенки доменов при своем движении наталкиваются на них. Между доменной стенкой и дислокацией (или границей зерна, или примесью) возникают взаимодействия. В результате, когда стенка наталкивается на препятствие, она задерживается на нем до тех пор, пока внешнее электрическое поле не достигнет определенной величины. Затем, когда поле несколько подрастет, стенка внезапно срывается. Таким образом, движение доменной стенки оказывается не плавным, как в идеальном кристалле, а скачкообразным. Эти скачки в микроскопическом масштабе проявляются на кривой поляризации и называются скачками Баркгаузена, а такое явление – эффектом Баркгаузена. Эти крайне малые скачки можно заметить лишь в специальных опытах на образцах очень малого поперечного сечения.

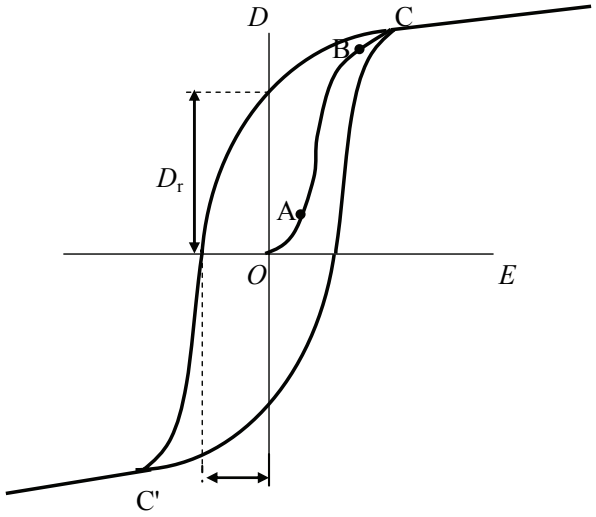


Рис. 4.5 – График поляризации сегнетоэлектрика во внешнем электрическом поле: ОА – граница упругого смещения стенок доменов; АВ – граница необратимого разрастания доменов; выше В – участок насыщения, монодоменное состояние

Если в поляризованном до насыщения образце уменьшить напряженность электрического поля до нуля, то электрическая индукция в ноль не обратится, а примет некоторое остаточное значение D_r. Отставание изменения индукции от изменения напряженности электрического поля называется *диэлектрическим гистерезисом*. При воздействии полем противоположной полярности индукция быстро уменьшается и при некоторой напряженности поля изменяет свое направление. Напряженность поля, при которой индукция проходит через ноль, называется *коэрцитивной силой*. Дальнейшее увеличение напряженности поля противоположной полярности вновь переводит образец в состояние насыщения (точка С').

Диэлектрический гистерезис обусловлен необратимым смещением доменных границ под действием электрического поля и свидетельствует о дополнительном механизме диэлектрических потерь, связанных с затратами энергии на переориентацию доменов. Площадь петли гистерезиса численно равна удельной плотности энергии, затраченной на один цикл переполяризации образца. Вследствие потерь на гистерезис сегнетоэлектрики характеризуются весьма большим тангенсом угла диэлектрических потерь (см. § 4.1.3.3), который в типичных случаях принимает значение порядка 0,1.

Нелинейный характер зависимости $D = f(E)$ с участком насыщения приводит к нелинейной зависимости диэлектрической проницаемости ε от напряженности электрического поля E (рис. 4.6). С ростом напряженности электрического поля относительная диэлектрическая проницаемость увеличивается и при определенных значениях поля достигает максимума, а затем начинает уменьшаться. Наличие участка насыщения на зависимости $D = f(E)$ при дальнейшем возрастании напряженности электрического поля соответствует постепенному уменьшению ε до единицы.

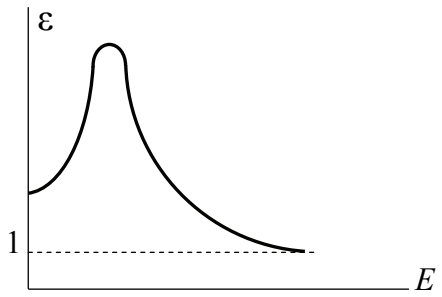


Рис. 4.6 – Зависимость диэлектрической проницаемости ε от напряженности электрического поля E для диэлектриков

Специфические свойства сегнетоэлектриков имеют место только в определенном интервале температур. На рис. 4.7 представлена температурная зависимость диэлектрической проницаемости сегнетовой соли вдоль полярной оси. (Полярная ось в сегнетоэлектриках – это ось, при наложении электрического поля вдоль которой наблюдается максимальное значение диэлектрической проницаемости.) Сегнетоэлектрики могут иметь несколько точек Кюри. В рассматриваемом нами случае наблюдается две точки Кюри.

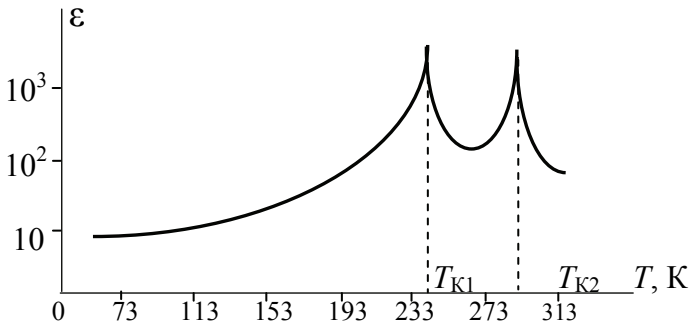


Рис. 4.7 – Зависимость диэлектрической проницаемости ε сегнетовой соли от температуры

В сильном электрическом поле сегнетоэлектрики переходят в монодоменное состояние, поэтому повышение температуры слабо сказывается на поляризации. Но после температуры Кюри происходит распад доменного состояния, и диэлектрик теряет спонтанную поляризацию.

4.1.1.5. Механизм спонтанной поляризации

Этот механизм рассмотрим на примере метатитаната бария ($BaTiO_3$). При температуре выше температуры Кюри (T_K = 393 К) $BaTiO_3$ имеет кристаллическую решетку, показанную на рис. 4.8, *а*.

Основу этой структуры составляют кислородные октаэдры, в центре которых находятся ионы титана, а в вершинах – ионы кислорода. Размер элементарной ячейки больше удвоенной суммы ионных радиусов титана и кислорода. Поэтому ион титана имеет некоторую свободу передвижения в границах кислородного октаэдра.

Если температура высока ($T > T_K$), то вследствие интенсивного теплового движения ион титана непрерывно перебрасывается от одного кислородного иона к другому, так что усредненное во времени его положение совпадает с центром элементарной ячейки, то есть такая ячейка

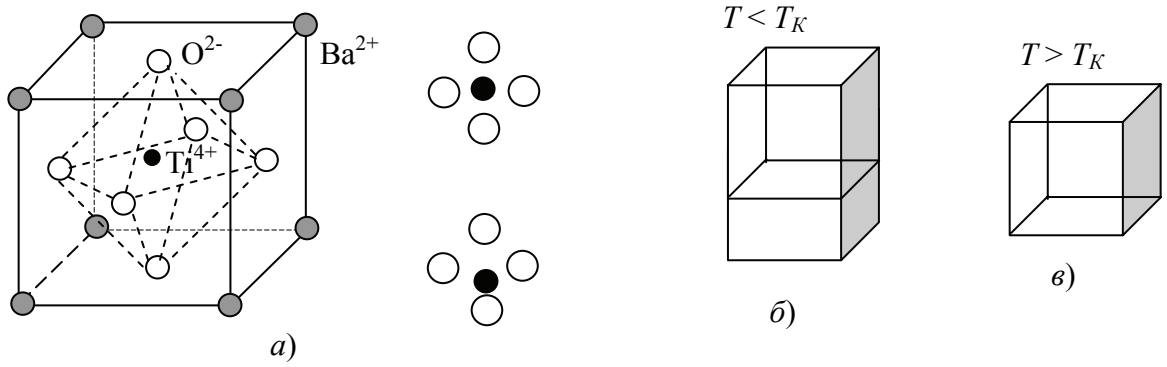


Рис. 4.8 – Пояснение механизма поляризации в метатитанате бария: *а* – элементарная кристаллическая ячейка метатитанита бария; *б* – деформация кристаллической ячейки при $T < T_K$; *в* – кристаллическая ячейка при температуре $T > T_K$; T_K – температура Кюри

не имеет электрического момента (рис. 4.8, *в*). При этом домены отсутствуют, и спонтанная поляризация не наблюдается.

При понижении температуры ниже температуры Кюри ($T < T_K$ = 393 К), энергии теплового движения недостаточно для переброса ионов титана из одного положения равновесия в другое, поэтому ион титана локализован вблизи одного из окружающих ионов кислорода. Вследствие этого кубическая симметрия в расположении заряженных частиц нарушается, и элементарная ячейка приобретает электрический момент (рис. 4.8, *б*). Одновременно с этим изменяется и ячейка – она вытягивается вдоль направления электрического момента, приобретая тетрагональную симметрию. Электрический момент ячейки действует на ионы титана в соседних ячейках и смещает последние в том же направлении. Вследствие этого возникает домен, то есть макроскопический участок, который имеет суммарный электрический момент. Одновременно в другом месте возникает аналогично другой домен, с иным направлением электрического момента. Поскольку в тетрагональной модификации $BaTiO_3$ смещение ионов может проходить в направлении любого ребра элементарной ячейки, то возможны шесть направлений спонтанной поляризации. Так как направления

электрических моментов доменов случайны, то суммарный электрический момент вещества в отсутствие внешнего электрического поля равен нулю.

4.1.1.6. Пироэлектричество

Пироэлектрическим эффектом называют изменение спонтанной поляризованности кристаллических диэлектриков при изменении температуры. Пироэлектрики, так же как и сегнетоэлектрики, имеют особую полярную ось. С увеличением температуры происходит анизотропное расширение кристалла, связанное с обратимым изменением поляризации. На противоположных гранях, лежащих на концах полярной оси, появляются заряды противоположного знака.

Уравнение пироэлектрического эффекта записывают в виде:

$$-dP_{\text{сп}} = p \cdot dT\,, \tag{4.11}$$

где $P_{\text{сп}}$ – спонтанная поляризованность диэлектрика, p – пироэлектрический коэффициент; dT – изменение температуры.

При неизменной температуре спонтанный электрический момент диэлектрика скомпенсирован свободными зарядами противоположного знака за счет процессов электропроводности или адсорбции заряженных частиц из окружающей атмосферы. Изменение спонтанной поляризованности сопровождается освобождением некоторого заряда на поверхности диэлектрика. Благодаря этому в замкнутой цепи возникает электрический ток:

$$i = -\frac{S \cdot dP_{\text{сп}}}{dt} = \frac{S \cdot p \cdot dT}{dt}\,, \tag{4.12}$$

где S – площадь поверхности пироэлектрика, $\frac{dT}{dt}$ – скорость изменения температуры.

Температурное изменение спонтанной поляризованности обусловлено двумя основными причинами. С одной стороны, повышение температуры нарушает упорядоченность в расположении элементарных дипольных моментов (первичный или истинный пироэффект). А с другой стороны, нагревание вызывает изменение линейных размеров диэлектрика и пьезоэлектрическую поляризацию, обусловленную деформацией (вторичный пироэффект). Пироэлектрический коэффициент p учитывает оба эти фактора.

Качество пироэлектрического материала принято характеризовать приведенным физическим параметром:

$$R = \frac{p}{\varepsilon c}\,, \tag{4.13}$$

где ε – диэлектрическая проницаемость, c – удельная теплоемкость. Чем больше значение R, тем большую разность потенциалов можно получить на образце при одной и той же поглощаемой мощности.

Пироэлектрическими свойствами обладают все сегнетоэлектрические материалы и некоторые *линейные диэлектрики* (например, турмалин, сульфат лития). Особенностью линейных диэлектриков является то, что в них, в

отличие от сегнетоэлектриков, направление спонтанной поляризованности не может изменяться с помощью внешнего электрического поля.

Сегнетоэлектрики проявляют пироэлектрические свойства только в *монодоменизированном* состоянии, для которого характерна одинаковая ориентация спонтанной поляризованности всех доменов. В *полидоменном* образце, для которого характерна различная ориентация спонтанной поляризованности всех доменов, суммарная поляризованность равна нулю, и поэтому пироэффект отсутствует. Монодоменизация сегнетоэлектрика осуществляется путем выдержки его в постоянном электрическом поле при температуре несколько ниже точки Кюри. Создание и закрепление монодоменизированного состояния в сегнетоэлектрических кристаллах является одной из важнейших проблем при их использовании в качестве пироэлектриков.

Из термодинамической теории сегнетоэлектричества вытекает следующая зависимость спонтанной поляризованности от температуры:

$$P_{\text{сп}} = A\sqrt{T_{\text{К}} - T}\,, \tag{4.14}$$

где А – некоторая константа для данного материала, $T_{\text{К}}$ – температура Кюри. Отсюда следует, что пироэлектрический коэффициент p возрастает по мере приближения к температуре фазового перехода второго рода ($T_{\text{К}}$):

$$p = \frac{dP_{\text{сп}}}{dT} = \frac{A}{2\sqrt{T_{\text{К}} - T}}\,. \tag{4.15}$$

4.1.1.7. Пьезоэлектрический эффект

Сущность *прямого пьезоэлектрического эффекта*, наблюдаемого только на кристаллах, не имеющих центра симметрии, состоит в том, что при действии механических напряжений наблюдается появление электрических зарядов на противоположных гранях диэлектрических кристаллов. Это явление было открыто братьями Кюри в 1880 г. Возникающий на каждой из поверхностей диэлектрика электрический заряд Q изменяется по линейному закону в зависимости от механического напряжения:

$$Q = d \cdot F\,, \tag{4.16}$$

где d – пьезомодуль, F – сила механического воздействия. Тогда поляризованность, связанную с пьезоэффектом можно найти по формуле:

$$P = \frac{Q}{S} = \frac{d \cdot F}{S} = d \cdot \sigma = q_S\,, \tag{4.17}$$

S – площадь пьезоэлектрика, на которую оказано механическое воздействие, q_S – поверхностный заряд, P – поляризованность, σ – механическое напряжение диэлектрика. Из (4.17) видно, что пьезомодуль d численно равен заряду, возникающему на единице поверхности пьезоэлектрика при приложении к нему единицы давления. Значение пьезомодуля используемых на практике пьезоэлектриков составляет около 10^{-19} Кл/Н.

Пьезоэффекты возникают только в веществах с гетерополярной химической связью, то есть пьезоэлектриками могут быть лишь *ионные* или

сильные полярные диэлектрики. Кроме того, необходимым условием наличия пьезоэффекта является *отсутствие центра симметрии* кристалла. Пьезоэлектриками могут быть только вещества с большим удельным сопротивлением, иначе пьезоэлектрическая поляризация быстро компенсируется свободными носителями заряда.

Известно более тысяч веществ, обладающих пьезоэлектрическими свойствами, в том числе – все сегнетоэлектрики. Среди них одно из важных мест занимает монокристаллический кварц. Это одна из модификаций двуокиси кремния. Пьезоэлектрические свойства существуют лишь у α-кварца, относящегося к тригональной системе; при температурах до 200 ºС пьезоэлектрические свойства α-кварца практически от нее не зависят. С дальнейшим повышением температуры пьезоэлектрический эффект медленно убывает вплоть до температуры 576 ºС, выше которой α-кварц претерпевает фазовое превращение и переходит в β-модификацию. Кристаллы β-кварца относятся к той кристаллографической группе гексагональной структуры, в которой есть центр симметрии, поэтому пьезоэлектрические свойства в этих кристаллах не наблюдаются.

α-Кварц имеет форму шестигранной призмы. В структуре кварца атомы кремния располагаются по винтовой линии, параллельной оси симметрии третьего порядка (ось *Z*). В кристаллах кварца принято различать три главные оси, образующие систему координат: *X* – электрическую ось, проходящую через вершины шестиугольного сечения (таких осей имеется три); *Y* – механическую ось, перпендикулярную стороне шестиугольника поперечного сечения (таких осей тоже три); *Z* – оптическую ось, проходящую через вершины кристалла и перпендикулярную к осям *X* и *Y*.

Возникновение пьезоэлектрического эффекта легко понять с помощью модельного представления. Химическая форма кварца имеет вид SiO_2. Его кристаллическая решетка состоит из положительных ионов кремния и отрицательных ионов кислорода. Каждый ион кремния несет по четыре, а каждый ион кислорода – два элементарных заряда. В первом приближении можно представить, что ионы кремния и кислорода расположены в шестигранных ячейках, одна из которых изображена на рис. 4.9, если смотреть на кристалл вдоль оптической оси *Z*, перпендикулярной плоскости рисунка. Ионы кремния изображены на рисунке белыми шариками 1, 2, 3, ионы кислорода – серыми. Те и другие ионы расположены по спирали, направление вращения которой определяется либо левым, либо правым кварцем (рис. 4.9 относится к левому кварцу). Ион

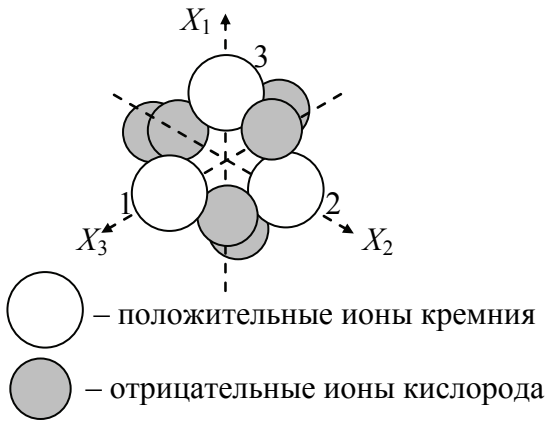


Рис. 4.9 – Шестигранная ячейка кварца

кремния 3 лежит несколько глубже иона 2, а ион 2 – глубже иона 1. В целом ячейка нейтральна и не имеет дипольного момента.

Для упрощения рассуждений заменим каждую пару соседних ионов кислорода одним отрицательным ионом с удвоенным зарядом. Мы придем к упрощенной модели решетки, изображенной на рис. 4.10, *а*. Если подвергнуть такую ячейку сжатию вдоль полярной (электрической) оси X_1 (рис. 4.10, *б*), то ион кремния 3 и ион кислорода 4 вклинятся между окружающими их боковыми ионами. В результате на плоскости *A* появится отрицательный, а на плоскости *B* – положительный заряды (*продольный пьезоэлектрический эффект*).

При сжатии в боковом направлении, т. е. перпендикулярно к оптической

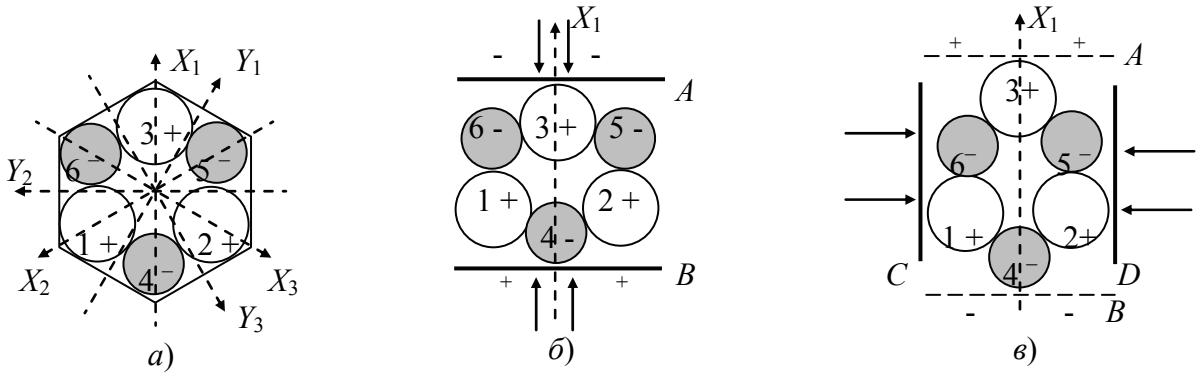


Рис. 4.10 – Упрощенная модель ячейки кварца (*а*); продольный пьезоэффект (*б*); поперечный пьезоэффект (*в*)

оси *Z* и электрической оси X_1 (рис. 4.10, *в*), ионы кремния 1 и 2 получают одинаковые, но противоположно направленные смещения внутрь ячейки. Так же ведут себя ионы кислорода 5 и 6. При этом сохраняется симметрия ячейки относительно плоскости, проходящей посредине между плоскостями *C* и *D*, и на этих плоскостях не возникает никаких зарядов. Однако ион кремния 3 и ион кислорода 4 смещаются наружу. Благодаря этому возникает дипольный момент, направленный в сторону полярной оси X_1. На плоскости *A* появится положительный, а на плоскости *B* – отрицательный заряды (*поперечный пьезоэлектрический эффект*). Таким образом, знаки зарядов в продольном и поперечном эффектах противоположны.

Из рассматриваемой модели видно также, что замена сжатия растяжением приводит к изменению знаков электрических зарядов при пьезоэлектрическом эффекте, и что поляризация пропорциональна деформации кристалла (когда деформации малы). Наконец, из модели видно, что сжатие или растяжение кристалла в направлении оптической оси *Z* никакими пьезоэлектрическими эффектами не сопровождаются. Все эти заключения подтверждены опытами.

Известен также *обратный* пьезоэлектрический эффект. При нем происходит изменение размеров диэлектрика $\Delta l/l$ в зависимости от приложенной напряженности электрического поля E по линейному закону:

$$\frac{\Delta l}{l} = \delta = d \cdot E, \qquad (4.18)$$

где l – длина пьезоэлектрика; Δl – относительное изменение длины пьезоэлектрика; δ – относительная деформация.

Доказано, что пьезомодули d прямого и обратного пьезоэффектов для одного и того же материала равны между собой.

При обратном пьезоэлектрическом эффекте деформация пьезоэлектрика зависит от направления электрического поля и меняет знак при изменении направления последнего. При приложении к пьезоэлектрику синусоидального электрического поля у него возникают синусоидальные деформации той же частоты.

Обратный пьезоэлектрический эффект не следует путать с явлением электрострикции, которое имеет место в любых диэлектриках при любых механизмах поляризации.

Электрострикцией называют изменение геометрических размеров диэлектрика $\Delta L/L$ вследствие смещения электрических зарядов его структуры, пропорциональное квадрату напряженности поля E и которое не зависит от его направления:

$$\Delta L/L = \chi \cdot E^2, \qquad (4.19)$$

где χ – модуль электрострикции.

Если к образцу в этом случае приложить синусоидальное напряжение, то деформация образца будет проходить с двойной частотой.

Явление электрострикции не имеет обратного эффекта, т. е. при механической деформации никакого заряда на противоположных гранях образца не образуется.

4.1.1.8. Электретное состояние в диэлектриках

Электретом называют тело из диэлектрика, длительно сохраняющее поляризацию и создающее в окружающем его пространстве электрическое поле, т. е. электреты являются формальными аналогами постоянных магнитов. Термин электрет был предложен английским физиком Хевисайдом в 1896 г., а образцы электретов были впервые изготовлены японским физиком Егучи в 1922 г.

В зависимости от способа формирования заряда различают *термоэлектреты*, *фотоэлектреты*, *радиоэлектреты*, *электроэлектреты*, *трибоэлектреты*.

Термоэлектретами принято называть электреты, получаемые охлаждением расплавленного или нагретого диэлектрика в сильном электрическом поле. После охлаждения диэлектрика в электрическом поле подвижность полярных молекул или свободных зарядов резко уменьшается, и диэлектрик может длительно сохранять сформированную остаточную поляризацию и заряды.

Фотоэлектреты получают из материалов, обладающих фотопроводимостью (сера, сульфид кадмия, сульфид цинка и др.), при совместном воздействии электрического поля и световой энергии. После

завершения светового воздействия носители разноименных зарядов оказываются «замороженными» у противоположных сторон диэлектрика, который становится электретом. Фотоэлектреты, при условии их хранения в темноте, способны удерживать заряды в течение нескольких месяцев.

Радиоэлектреты получают при воздействии на диэлектрик радиоактивного излучения. В результате ударной ионизации поверхностного слоя или внедрения в поверхностный слой диэлектрика ускоренных заряженных частиц поверхность диэлектрика оказывается заряженной.

Электроэлектреты получают воздействием на диэлектрик только электрического поля при комнатной температуре. В зазоре между электродом и диэлектриком свободные носители заряда (ионы и электроны), ускоренные полем, бомбардируют поверхность диэлектрика, создавая тем самым заряженный поверхностный слой.

Трибоэлектретами называют электреты, получаемые контактной электризацией. Образование поверхностного заряда объясняется различной работой выхода электронов. При плотном контакте двух диэлектриков электроны диэлектрика с меньшей работой выхода переходят в диэлектрик с большей работой выхода.

Рассмотрим состояние электрических зарядов на примере термоэлектрета. На рис. 4.11 видно, что на каждой из поверхностей электрета, находящегося под поляризующими электродами, образуются электрические заряды обоих знаков. Заряды, перешедшие из поляризующего электрода или из воздушного зазора на поверхность твердого диэлектрика и имеющие тот же знак заряда, что и на электроде, называют *гомозарядами*. Заряды со знаками противоположными полярности электродов, возникают в электрете за счет различных релаксационных механизмов поляризации и называются *гетерозарядами*.

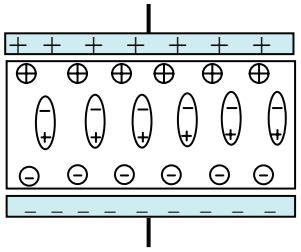


Рис. 4.11 – Состояние зарядов в электрете

Разность гетеро- и гомозарядов определяет результирующий заряд поверхности электрета. Преобладанием того или иного заряда объясняется инверсия его знака на поверхности электрета с течением времени. Сразу после окончания поляризации преобладает гетерозаряд, а спустя некоторое время, когда тепловое движение дезориентирует диполи, преобладающим оказывается гомозаряд. Опыты показывают, что гомозаряды сохраняются в течение более длительного времени по сравнению с гетерозарядами. Поверхностная плотность зарядов, наблюдающаяся у различных электретов, может составлять $10^{-6}…10^{-4}$ Кл/м^2. При напряженности поляризующего поля $E < 0{,}5$ МВ/м, как правило, обнаруживается только гетерозаряд, при $E > 1$ МВ/м – гомозаряд.

У органических полярных электретов преобладают гетерозаряды, у неорганических (керамических) материалов и органических неполярных диэлектриков – гомозаряды. У такого известного неполярного диэлектрика как фторопласт-4 (пленочный), существует очень прочно удерживаемый и

значительный по величине гомозаряд. Гомозаряд локализован только в поверхностных слоях электрета, тогда как гетерозаряд распределен по всему объему электрета, что и подтверждает физическую природу этих зарядов.

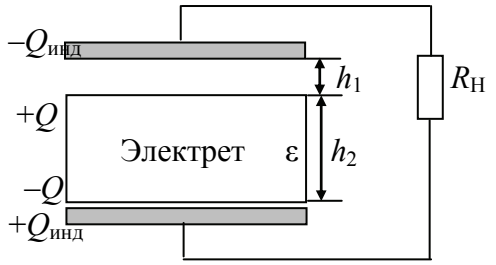


Рис. 4.12 – Электрет в электрическом поле

Если электрет поместить между металлическими обкладками, как показано на рис 4.12, то величина индуцированного заряда на них

$$Q_{инд} = \frac{Q}{\frac{\varepsilon h_1}{h_2} + 1}, \qquad (4.20)$$

где Q – заряд на поверхности электрета; h_1 – величина зазора между одной из поверхностей электрета и электродом (принимаем, что зазор между другой поверхностью и электродом бесконечно мал); h_2 – толщина электрета; ε – диэлектрическая проницаемость материала электрета. При изменении зазора h_1 величина индуцированного заряда также изменяется. При периодическом изменении зазора в цепи между электродами течет переменный ток, частота которого равна частоте изменения зазора. Так как внутреннее сопротивление системы электрет-электрод при этом велико (10^7…10^8 Ом), величина тока не зависит от внешней нагрузки, и получается своеобразный генератор переменного тока.

С течением времени величина заряда электрета изменяется, что связано с изменением остаточной поляризации. Для характеристики этого явления вводится понятие *время жизни* электрета. В настоящее время электретное состояние наблюдается более чем у 70 диэлектриков.

4.1.2. Физические свойства гетерогенных и слоистых диэлектриков

4.1.2.1. Зависимость свойств диэлектриков от температуры

Для оценки температурной зависимости ε диэлектрических материалов и емкости C конденсаторов пользуются величинами температурных коэффициентов: *температурного коэффициента диэлектрической проницаемости*, $К^{-1}$

$$\alpha_\varepsilon = \frac{1}{\varepsilon} \cdot \frac{d\varepsilon}{dT} \qquad (4.21)$$

и *температурного коэффициента емкости*, $К^{-1}$

$$\alpha_C = \frac{1}{C} \cdot \frac{dC}{dT}. \qquad (4.22)$$

Величину α_ε обычно находят методом графического дифференцирования кривой $\varepsilon = F(T)$, как показано на рис. 4.13**.**

При конструировании радиоаппаратуры, работающей в условиях меняющейся температуры, требуется обеспечить практическую

независимость емкости конденсатора от температуры, т.е. создать термокомпенсированный конденсатор. Для решения этой задачи существуют два пути.

Во-первых, возможно применить систему двух (или более) параллельно или последовательно соединенных друг с другом конденсаторов, α_C которых имеют различные знаки (один – положительный, а другой – отрицательный).

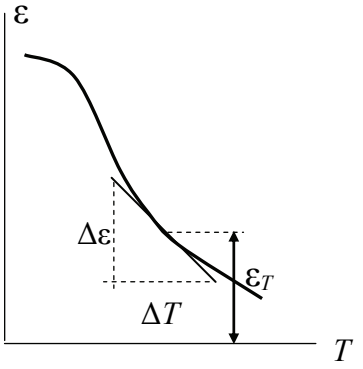


Рис. 4.13 – Определение α_ε методом графического дифференцирования

Для случая параллельного соединения конденсаторов, имеющих емкости C_1 и C_2, температурные коэффициенты, емкости которых соответственно равны α_{C_1} и α_{C_2}, результирующая емкость равна: $C = C_1 + C_2$. Температурный коэффициент результирующей емкости можно найти по формуле:

$$\alpha_C = \frac{C_1\alpha_{C_1} + C_2\alpha_{C_2}}{C_1 + C_2}. \qquad (4.23)$$

Условие температурной компенсации: $\alpha_C = 0$, откуда

$$C_1\alpha_{C_1} + C_2\alpha_{C_2} = 0. \qquad (4.24)$$

Для случая последовательного соединения двух конденсаторов

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}.$$

Температурный коэффициент результирующей емкости

$$\alpha_C = \alpha_{C_1} + \alpha_{C_2} - \frac{C_1\alpha_{C_1} + C_2\alpha_{C_2}}{C_1 + C_2} \qquad (4.25)$$

и условие температурной компенсации

$$C_1\alpha_{C_2} + C_2\alpha_{C_1} = 0. \qquad (4.26)$$

Во-вторых, задачу температурной компенсации емкости можно разрешить, применяя диэлектрик, представляющим собой смесь двух материалов, имеющих различные знаки α_ε. Это будет показано ниже.

4.1.2.2. Напряженность электрического поля в многослойных диэлектриках

Величина диэлектрической проницаемости важна для расчета напряженности поля в многослойных диэлектриках. Рассмотрим вначале изменение напряженности поля на примере двухслойного диэлектрика при переменном напряжении. Простейший случай плоского двухслойного конденсатора представлен на рис. 4.14, где h_1 и h_2 – толщины слоев; ε_1, ε_2 – диэлектрические проницаемости материалов слоев; E_1 и E_2 –

напряженности поля в слоях; U_1 и U_2 – напряжения на слоях. Напряжение на двухслойном конденсаторе определим по формуле:

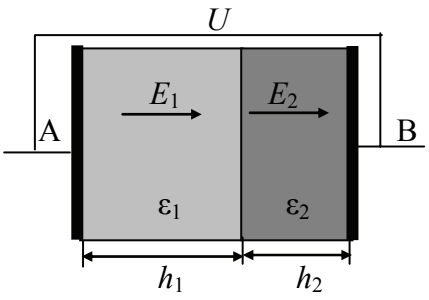


Рис. 4.14 – Схема двухслойного плоского конденсатора

$$U = U_1 + U_2 = E_1 \cdot h_1 + E_2 \cdot h_2 . \qquad (4.27)$$

где U – полное напряжение между электродами А и В конденсатора. Комбинируя условие последовательного соединения слоев и условие непрерывности вектора электрического смещения при переходе из одного слоя в другой, получим условие обратной пропорциональности напряженностей в слоях диэлектрическим проницаемостям слоев

$$\frac{E_1}{E_2} = \frac{\varepsilon_2}{\varepsilon_1}, \qquad (4.28)$$

и найдем напряженности поля в обоих слоях

$$E_1 = \varepsilon_2 \frac{U}{h_1\varepsilon_2 + h_2\varepsilon_1}, \quad E_2 = \varepsilon_1 \frac{U}{h_1\varepsilon_2 + h_2\varepsilon_1} . \qquad (4.29)$$

Напряжение на слоях U_1 и U_2 получим, умножив E_1 на h_1 и E_2 на h_2:

$$U_1 = \frac{\varepsilon_2 h_1 U}{h_1\varepsilon_2 + h_2\varepsilon_1}, \quad U_2 = \frac{\varepsilon_1 h_2 U}{h_1\varepsilon_2 + h_2\varepsilon_1}. \qquad (4.30)$$

Из формулы (4.30) понятно, что при $h_1 = h_2$ напряжение U обратно пропорционально диэлектрической проницаемости ε ($\frac{U_1}{U_2} = \frac{\varepsilon_2}{\varepsilon_1}$). Это нужно учитывать при расчете и конструировании слоистой структуры, так как бóльшее напряжение будет падать на слоях с мéньшей диэлектрической проницаемостью. В особенно опасном положении оказываются воздушные прослойки внутри диэлектрической структуры или между изолятором и электродами при их неплотном прилегании. Из-за малой величины ε и малой электрической прочности газов, в таких прослойках легко возникают частичные разряды. Поэтому в неоднородном поле для уменьшения электрической нагрузки электроизоляционных материалов следует в места с наибольшей напряженностью поля помещать материалы с наибольшей диэлектрической проницаемостью ε.

При постоянном напряжении соотношение между напряжениями на слоях можно выразить следующим образом:

$$\frac{U_1}{U_2} = \frac{R_1}{R_2} = \frac{\rho_1 h_1}{\rho_2 h_2}. \qquad (4.31)$$

Тогда напряжения на слоях равны

$$U_1 = \frac{\rho_1 h_1 U}{\rho_1 h_1 + \rho_2 h_2}; \qquad U_2 = \frac{\rho_2 h_2 U}{\rho_1 h_1 + \rho_2 h_2}, \tag{4.32}$$

а напряженности на тех же слоях можно вычислить по формулам:

$$E_1 = \frac{U_1}{h_1} = \frac{\rho_1 U}{\rho_1 h_1 + \rho_2 h_2}; \quad E_2 = \frac{U_2}{h_2} = \frac{\rho_2 U}{\rho_1 h_1 + \rho_2 h_2}, \tag{4.33}$$

где ρ_1 и ρ_2 – удельные сопротивления диэлектрических слоев 1 и 2.

4.1.2.3. Диэлектрическая проницаемость слоистых и гетерогенных диэлектриков

На практике часто приходится встречаться с задачей определения эффективной диэлектрической проницаемости ε композиционного диэлектрика, представляющего собой смесь двух (или большего числа) компонентов.

Легко рассчитать ε для модели плоского конденсатора, диэлектрик которого состоит из двух сплошных диэлектриков, имеющих различные диэлектрические проницаемости. Обозначая диэлектрические проницаемости компонентов ε_1 и ε_2 и их объемные доли υ_1 и υ_2 (очевидно, что $\upsilon_1 + \upsilon_2 = 1$), получаем:

а) для параллельного соединения

$$\varepsilon = \upsilon_1 \varepsilon_1 + \upsilon_2 \varepsilon_2; \tag{4.34}$$

б) для последовательного соединения (двухслойный диэлектрик)

$$\varepsilon = \frac{\varepsilon_1 \varepsilon_2}{\upsilon_1 \varepsilon_2 + \upsilon_2 \varepsilon_1}. \tag{4.35}$$

Формулы (4.34) и (4.35) могут использоваться в ряде практических случаев. Было установлено, что для конденсаторной бумаги достаточно хорошо подходит модель последовательно соединенных слоев клетчатки и воздуха.

Во многих практических случаях (пластические массы, керамика и т. п.) мы имеем дело с неупорядоченной, хаотической («статистической») смесью компонентов, называемой гетерогенными системами.

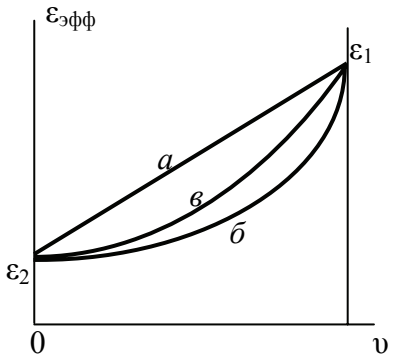


Рис. 4.15 – Зависимость диэлектрической проницаемости смеси двух компонентов от их объемного содержания в смеси: *а* – для параллельного соединения компонентов; *б* – для последовательного соединения; *в* – эффективная диэлектрическая проницаемость смеси

Как показал Винер, диэлектрические свойства гетерогенных систем любой структуры лежат между предельными значениями диэлектрической проницаемости для параллельного или последовательного расположения слоев:

$$\left(\upsilon_1 / \varepsilon_1 + \upsilon_2 / \varepsilon_2\right) \le \varepsilon \le \left(\upsilon_1 \varepsilon_1 + \upsilon_2 \varepsilon_2\right). \tag{4.36}$$

В этом случае истинное значение эффективной диэлектрической проницаемости смеси ε лежит между значениями, определяемыми формулами (4.34) и (4.35), как это схематически представлено на рис. 4.15.

Существует большое число приближенных формул для расчета эффективной диэлектрической проницаемости ε гетерогенных систем. Чаще всего пользуются формулой

$$\varepsilon^k = \varepsilon_1^k \upsilon_1 + \varepsilon_2^k \upsilon_2 , \qquad (4.37)$$

где показатель k может принимать любые значения от –1 до +1. Значение k = –1 соответствует слоям, перпендикулярным полю, а k = +1 – параллельным ему.

Для расчетов эффективной диэлектрической проницаемости смесей широко используется формула Лихтенеккера:

$$\lg \varepsilon = \upsilon_1 \lg \varepsilon_1 + \upsilon_2 \lg \varepsilon_2 , \qquad (4.38)$$

где, как и выше, ε_1 и ε_2 – диэлектрические проницаемости компонентов, а υ_1 и υ_2 – их объемные содержания в смеси.

Температурный коэффициент диэлектрической проницаемости смеси получается дифференцированием по температуре формулы (4.38)

$$\frac{1}{\varepsilon} \cdot \frac{d\varepsilon}{dT} = \frac{\upsilon_1}{\varepsilon_1} \cdot \frac{d\varepsilon_1}{dT} + \frac{\upsilon_2}{\varepsilon_2} \frac{d\varepsilon_2}{dT} , \qquad (4.39)$$

что может быть записано в виде

$$\alpha_\varepsilon = \alpha_{\varepsilon_1} + \alpha_{\varepsilon_2} . \qquad (4.40)$$

Из формулы (4.40) следует, что и гетерогенные диэлектрические системы пригодны для изготовления термокомпенсированных конденсаторов, т. е. они имеют $\alpha_\varepsilon = 0$ при условии

$$\upsilon_1 \alpha_{\varepsilon_1} + \upsilon_2 \alpha_{\varepsilon_2} = 0 , \qquad (4.41)$$

откуда определяются соответствующие температурной компенсации значения концентраций υ_1 и υ_2.

Для случая гетерогенного диэлектрика, состоящего из n компонентов, формулы (4.34), (4.35) и (4.38) записываются в виде:

а) для параллельного соединения компонентов

$$\varepsilon = \sum_{i=1}^{n} \upsilon_i \varepsilon_i \ ; \qquad (4.42)$$

б) для последовательного соединения компонентов

$$\frac{1}{\varepsilon} = \sum_{i=1}^{n} \frac{\upsilon_i}{\varepsilon_i} ; \qquad (4.43)$$

в) для неупорядоченной смеси компонентов (по формуле Лихтенеккера)

$$\lg \varepsilon = \sum_{i=1}^{n} \upsilon_i \lg \varepsilon_i . \qquad (4.44)$$

В настоящее время для расчетов эффективной проницаемости гетерогенной системы наиболее удачными считаются формула Максвелла-Гарнетта (известная также как формула Винера-Вагнера)

$$\frac{\varepsilon-\varepsilon_1}{\varepsilon+2\varepsilon_1}=\upsilon_2\frac{\varepsilon_2-\varepsilon_1}{\varepsilon_2+2\varepsilon_1} \qquad (4.45)$$

и формула Бруггемана

$$\left(1-\upsilon_2\right)\frac{\varepsilon_1-\varepsilon}{\varepsilon_1+2\varepsilon}+\upsilon_2\frac{\varepsilon_2-\varepsilon}{\varepsilon_2+2\varepsilon}=0\,. \qquad (4.46)$$

Первоначально формулы (4.45) и (4.46) были получены для разбавленных систем со сферическими вкраплениями. В работе [9] было проведено обобщение этих формул на случай концентрированных многокомпонентных гетерогенных систем, состоящих из неоднородных несферических частиц или частей.

4.1.3. Электрические свойства диэлектриков

4.1.3.1. Общие сведения о проводимости диэлектриков

В диэлектрике, находящемся под напряжением, происходят поляризационные процессы, связанные с перемещением электрических зарядов, следовательно, в диэлектрике возникают электрические токи. Ток, сопутствующий электронной и ионной поляризации, называют *током смещения* $I_{см}$. Он проходит в очень малые промежутки времени (10^{-13} – 10^{-15}с) и является чисто реактивным. Релаксационные виды поляризации вызывают прохождение в диэлектрике *тока абсорбции* $I_{аб}$. Ток абсорбции – это часть тока через диэлектрик, которая экспоненциально затухает с течением времени. Природа тока абсорбции сильно зависит от типа диэлектрика. Наличие в диэлектрике небольшого числа свободных зарядов обуславливает возникновение небольшого по величине *сквозного тока* $I_{ск}$. Таким образом, ток, проходящий через диэлектрик, представляет собой сумму токов смещения, абсорбции и сквозного:

$$I = I_{см} + I_{аб} + I_{ск}. \qquad (4.47)$$

В первый момент приложения постоянного напряжения ток значительно больше, чем некоторое время спустя, когда в диэлектрике протекает лишь сквозной ток. Это объясняется тем, что токи смещения и абсорбции вскоре прекращаются, так как они обусловлены быстропротекающими процессами поляризации.

В случае переменного напряжения все три вида токов имеют место в течение всего времени нахождения материала в электрическом поле.

Рассмотрим векторную диаграмму токов в диэлектрике, показанную на рис. 4.16, к которому приложено переменное напряжение U. Ток смещения $I_{см}$ не вызывает потерь энергии, носит чисто емкостной характер и опережает напряжение по фазе на 90^{o}. Ток абсорбции также опережает напряжение, но, учитывая вызываемые им потери, на угол

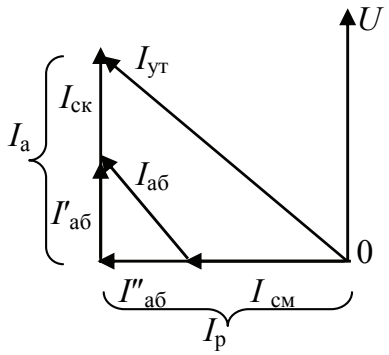


Рис. 4.16 – Векторная диаграмма токов в диэлектрике

меньший 90^{o}. Поэтому, вектор тока абсорбции можно разложить на активную $I'_{аб}$ и реактивную $I''_{аб}$ составляющие. Сквозной ток, изображенный вектором $I_{ск}$, имеет чисто активный характер и совпадает по фазе с напряжением. Полный ток через диэлектрик (обычно называемый током утечки) равен векторной сумме активной $I_{а}$ и реактивной $I_{р}$ составляющих тока:

$$I_{ут} = I_{а} + I_{р} = (I'_{аб} + I_{ск}) + (I''_{аб} + I_{см}). \quad (4.48)$$

Следовательно, проводимость диэлектрика величина комплексная. При переменном напряжении активная составляющая определяется как сквозным током, так и активной составляющей тока абсорбции, а при постоянном напряжении – только сквозным током.

Удельная проводимость диэлектриков в нормальных условиях составляет $10^{-10}…10^{-16}$ Ом$^{-1}$м$^{-1}$.

4.1.3.2. Проводимость твердых диэлектриков

В большинстве случаев активная проводимость диэлектриков ионная, реже – электронная. Сопротивление диэлектрика, заключенного между двумя электродами, при постоянном напряжении, т. е. сопротивление изоляции $R_{из}$, можно вычислить по формуле:

$$R_{из} = \frac{U}{I_{ут} + I_{аб}}, \quad (4.49)$$

где $I_{ут}$ – ток утечки; $I_{аб}$ – ток абсорбции.

В твердых изоляционных материалах различают *объемную проводимость* и *поверхностную проводимость.* Для сравнительной оценки объемной и поверхностной проводимости разных материалов используют удельное объемное сопротивление ρ_V и удельное поверхностное сопротивление ρ_S.

Удельное объемное сопротивление ρ_V равно сопротивлению куба с ребром в 1 м, мысленно вырезанного из исследуемого материала, если ток проходит через две противоположные грани этого куба; ρ_V выражается в Ом·м.

В случае плоского образца материала в однородном поле удельное объемное сопротивление рассчитывают по формуле:

$$\rho_V = R\frac{S}{h}, \quad (4.50)$$

где R – объемное сопротивление, Ом; S – площадь электрода, м^2; h – толщина образца, м.

Удельное поверхностное сопротивление ρ_S равно сопротивлению квадрата (любых размеров), мысленно выделенного на поверхности материала, если ток проходит через две противоположные стороны этого квадрата; измеряется в Омах. Удельное поверхностное сопротивление

$$\rho_S = R_S\frac{d}{l}, \quad (4.51)$$

где R_S – поверхностное сопротивление образца материала между параллельно поставленными электродами шириной *d,* отстоящими друг от друга на расстоянии *l* (рис. 4.17). Удельному объемному сопротивлению соответствует *удельная объемная проводимость* $\sigma = 1/\rho_V$ [Ом$^{-1}$м$^{-1}$], а удельному поверхностному сопротивлению – *удельная поверхностная проводимость* $\sigma = 1/\rho_S$ [Ом$^{-1}$].

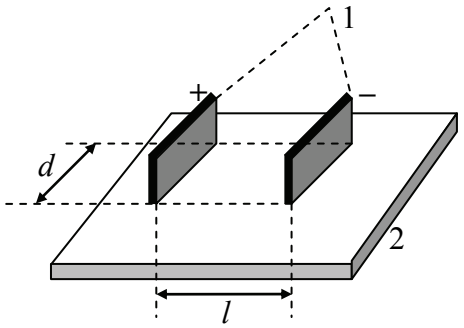


Рис. 4.17 – Схема для измерения ρ_S: 1 – электроды; 2 – диэлектрик

Полная проводимость твердого диэлектрика, соответствующая его сопротивлению $R_{из}$, складывается из объемной и поверхностной проводимостей. Соответственно, полное сопротивление твердого диэлектрика, измеряемое в Омах, складывается из параллельно включенных объемного и поверхностного сопротивлений, определяемых по формуле:

$$R_d = \frac{R_V R_S}{R_V + R_S}. \qquad (4.52)$$

Произведение сопротивления изоляции диэлектрика конденсатора на его емкость, называют *постоянной времени* конденсатора:

$$\tau_0 = R_{из} C. \qquad (4.53)$$

Физическая сущность объемной проводимости твердых диэлектриков. Проводимость твердых тел обуславливается как перемещением ионов самого диэлектрика, так и ионов случайных примесей, а у некоторых материалов вызвана наличием свободных электронов.

Вид проводимости устанавливают экспериментально, используя закон Фарадея для электролиза. Ионная проводимость сопровождается переносом вещества на электроды. При электронной проводимости это явление не наблюдается.

В твердых диэлектриках ионного строения проводимость обусловлена главным образом перемещением ионов, которые вырываются из решетки под влиянием флюктуаций теплового движения. При низких температурах дрейфуют слабо закрепленные ионы, в частности ионы примесей, а при высоких температурах – ионы кристаллической решетки.

В диэлектриках с атомными или молекулярными решетками проводимость сильно зависит от наличия примесей. В каждом конкретном случае о механизме электропроводности судят по энергии активации носителей зарядов.

Собственная проводимость твердых тел и изменение ее в зависимости от температуры определяются структурой вещества и его составом. В телах кристаллического строения с ионными решетками проводимость зависит от валентности ионов. Кристаллы с одновалентными ионами имеют бóльшую удельную проводимость, чем кристаллы с многовалентными ионами.

В анизотропных кристаллах удельная проводимость неодинакова по разным его осям. Например, в кварце удельная проводимость в направлении, параллельном главной оси, приблизительно в 1000 раз больше, чем в направлении, перпендикулярном этой оси.

Величина удельной проводимости аморфных тел связана, прежде всего, с их составом. Высокомолекулярные органические полимеры имеют удельную проводимость, которая зависит в значительной степени от ряда факторов: от химического состава и наличия примесей; от степени полимеризации. Органические неполярные аморфные диэлектрики, например, полистирол, отличаются малой удельной проводимостью.

Большую группу аморфных тел составляют неорганические стекла. Удельная проводимость стекол зависит главным образом от химического состава; это позволяет в ряде случаев получать достаточно высокое удельное сопротивление.

У твердых пористых диэлектриков при наличии в них влаги даже в очень малых количествах, значительно увеличивается удельная проводимость.

Физическая сущность поверхностной проводимости твердых диэлектриков. Поверхностная проводимость обусловлена наличием адсорбированной влаги, загрязнениями и другими дефектами поверхности диэлектрика.

Удельная поверхностная проводимость тем ниже, чем меньше полярность вещества и чем чище поверхность диэлектрика. Причем, наличие загрязнений на поверхности относительно мало влияет на удельную поверхностную проводимость *гидрофобных* диэлектриков и сильно влияет на проводимость *гидрофильных* диэлектриков.

Наибольшее увеличение удельной поверхностной проводимости наблюдается у полярных диэлектриков, частично растворимых в воде, у которых на поверхности образуется пленка электролита. Кроме того, на поверхности полярных диэлектриков могут адсорбироваться различные загрязнения, которые также приводят к росту поверхностной проводимости. Высокую поверхностную проводимость имеют и объемно-пористые материалы, так как процесс поглощения влаги в глубине материала стимулирует также и образование её пленки на поверхности.

4.1.3.3. Диэлектрические потери

Диэлектрическими потерями называют электрическую мощность, затраченную на нагревание диэлектрика, находящегося в электрическом поле.

Мощность, рассеиваемая в единице объема, называется удельными потерями p. Но чаще всего для характеристики способности диэлектрика рассеивать энергию в электрическом поле используют угол диэлектрических потерь δ или тангенс этого угла $tg\delta$.

Углом диэлектрических потерь δ называют угол, дополняющий до 90^0 угол сдвига фаз между током и напряжением в емкостной цепи.

В случае идеального диэлектрика δ = 0. Чем больше рассеиваемая в диэлектрике мощность, переходящая в тепло, тем больше угол диэлектрических потерь δ и его функция *tg*δ.

Тангенс угла диэлектрических потерь непосредственно входит в формулу для значения рассеиваемой в диэлектрике мощности, поэтому практически наиболее часто пользуются этой характеристикой.

Схемы замещения диэлектриков с потерями. При изучении поведения диэлектрика с потерями при переменном напряжении часто оказывается целесообразным заменить рассматриваемый диэлектрик емкостью без потерь и активным сопротивлением, соединенными между собой параллельно (параллельная схема, рис. 4.18, *а*) или последовательно (последовательная схема, рис. 4.18, *б*). Иногда такая замена является чисто формальной, но в ряде случаев она отражает реальные физические процессы, протекающие в диэлектрике.

Обе схемы эквивалентны, если при равенстве полных сопротивлений

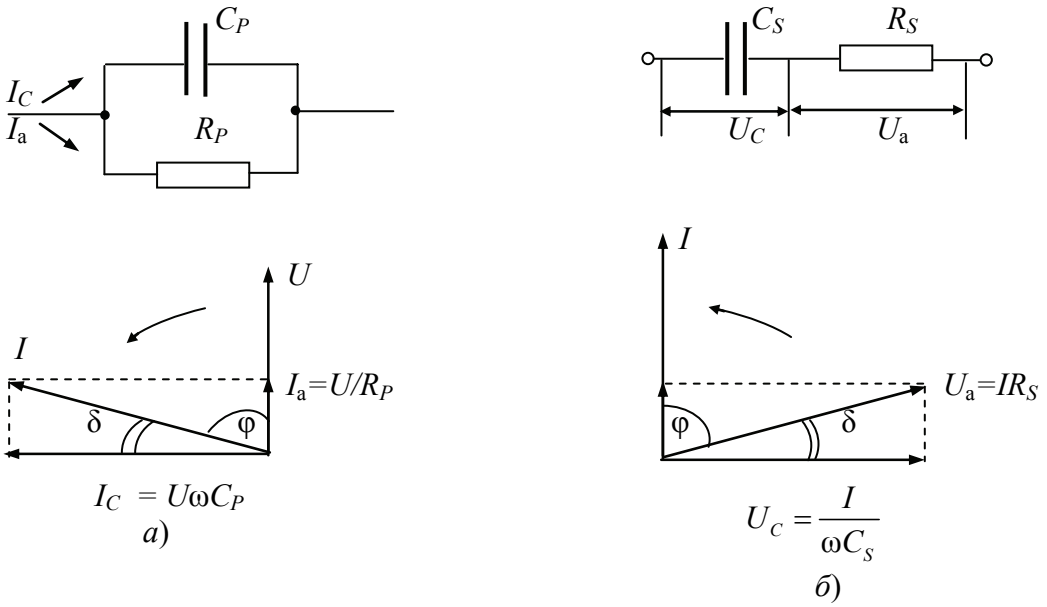


Рис. 4.18 – Параллельная (*а*) и последовательная (*б*) эквивалентные схемы диэлектрика с потерями и векторные диаграммы для них

$Z_1 = Z_2 = Z$ равны соответственно их активные и реактивные составляющие (равны *tg*δ). Рассмотрим схему, эквивалентную конденсатору с диэлектриком, обладающим потерями. Эта схема должна быть выбрана с таким расчетом, чтобы активная мощность, расходуемая в данной схеме, была равна мощности, рассеиваемой в диэлектрике конденсатора, а ток был бы сдвинут относительно напряжения на тот же угол, что и в рассматриваемом конденсаторе. В дальнейшем мы будем ориентироваться на параллельную схему (рис. 4.18, *а*).

Из рисунка 4.18, *а* видно, что $I_a = \frac{U}{R}$ и $I_C = U\omega C_P$, тогда *tg*δ можно найти по формуле:

$$tg\delta = \frac{I_a}{I_C} = \frac{1}{\omega C_P R}, \qquad (4.54)$$

где C_P – емкость в параллельной эквивалентной схеме.

Диэлектрические потери обусловлены *сквозным током* и активной составляющей *токов поляризации*, а также наличием посторонних проводящих или полупроводящих включений. Потери энергии в диэлектриках наблюдаются как при постоянном напряжении, так и при переменном. Мощность, рассеиваемая в диэлектрике

$$P_a = U \cdot I_a. \qquad (4.55)$$

Из формулы (4.54) можно найти значение активной составляющей тока

$$I_a = I_C \cdot tg\delta = U\omega C_P tg\delta. \qquad (4.56)$$

Подставив формулу (4.56) в формулу (4.55), получим значение активной мощности

$$P_a = U \cdot I_a = U^2 \omega C_P tg\delta, \qquad (4.57)$$

где P_a выражается в ваттах, U – в вольтах, ω – в (секундах)$^{-1}$, C – в фарадах.

Если для данного диэлектрика известно, что потери в нем определяются только потерями от сквозной проводимости в широком диапазоне частот, то угол потерь конденсатора с таким диэлектриком может быть вычислен для любой частоты. Потери в таком конденсаторе равны

$$P_a = \frac{U^2}{R}, \qquad (4.58)$$

где R – сопротивление при некоторой частоте.

В случае же неоднородного электрического поля удельные потери в разных точках диэлектрика различны, так как различна величина напряженности электрического поля в разных точках; кроме того, если диэлектрик неоднороден, при расчете удельных потерь необходимо учитывать и различие в параметрах диэлектрика в разных частях объема изоляции. В этом случае используется формула для удельных диэлектрических потерь. Выражение для удельных диэлектрических потерь, т.е. мощность, рассеиваемая в 1 м^3 диэлектрика (Вт/м^3) с учетом того, что емкость плоского конденсатора равна $C = \varepsilon_0 \varepsilon \frac{S}{h}$, имеет вид:

$$p = \frac{P_a}{V} = \frac{U^2 \omega C_P tg\delta}{S \cdot h} = \frac{U^2 \omega \varepsilon_0 \varepsilon S tg\delta}{S \cdot h^2} = E^2 \omega \varepsilon_0 \varepsilon tg\delta, \qquad (4.59)$$

где V – объем диэлектрика между плоскими электродами (м^3), $E = \frac{U}{h}$ – напряженность электрического поля (В/м). Величина E в формуле (4.59) выражена в В/м. Формула пригодна для любой картины неоднородного поля.

Так как емкость между противоположными гранями куба со стороной 1 м

$$c_1 = \varepsilon_0 \varepsilon, \qquad (4.60)$$

а если учесть, что реактивная и активная удельные проводимости равны $\sigma_c = \omega C_P$, $\sigma_{\text{а}} = \dfrac{1}{R}$, то реактивная составляющая удельной проводимости

$$\sigma_c = \omega c_1 = \omega \varepsilon_0 \varepsilon , \tag{4.61}$$

а активная составляющая при данной частоте (см. формулы (4.54) и (4.61))

$$\sigma_{\text{а}} = \sigma_c \cdot tg\delta = \omega \varepsilon_0 \varepsilon tg\delta . \tag{4.62}$$

Из формулы (4.62) можно получить тангенс потерь, выраженный через активную и реактивную проводимости:

$$tg\delta = \frac{\sigma_{\text{а}}}{\sigma_c} . \tag{4.63}$$

Выражение (4.62) используют для вычисления активной удельной проводимости диэлектрика при различных частотах по измеренным значениям диэлектрической проницаемости и тангенса угла потерь материала при этих значениях частоты.

Добротностью конденсатора называют величину

$$Q_C = \frac{1}{tg\delta} . \tag{4.64}$$

Если учесть, что величина

$$\sigma_{\text{а}} = \omega \varepsilon \varepsilon_0 tg\delta \tag{4.65}$$

представляет собой удельную объемную активную проводимость диэлектрика (при переменном напряжении), то формула (4.59) может быть переписана в виде

$$p = E^2 \sigma_a . \tag{4.66}$$

Обычно (при одном и том же значении E) потери при переменном напряжении больше, чем потери при постоянном напряжении.

Для параллельной схемы из векторной диаграммы (рис. 4.20, *а*) были найдены $tg\delta$ (ф-ла (4.54)) и $P_{\text{а}}$ (ф-ла (4.57)):

$$tg\delta = \frac{1}{\omega C_P R_P} , \quad P_{\text{а}} = U^2 \omega C_P tg\delta .$$

Для последовательной схемы (рис. 4.20, *б*)

$$tg\delta = \omega C_S R_S , \tag{4.67}$$

где C_S – емкость в последовательной эквивалентной схеме.

$$P_{\text{а}} = \frac{U^2 \omega C_S tg\delta}{1 + tg^2\delta} . \tag{4.68}$$

Связи между C_P и C_S и между R_P и R_S при переходе от параллельной модели диэлектрика к последовательной и обратно:

$$C_S = C_P(1 + tg^2\delta) \ ; \ C_P = \frac{C_S}{1 + tg^2\delta} \tag{4.69}$$

$$R_S = \frac{R_P}{1 + \dfrac{1}{tg^2\delta}} \ ; \ R_P = R_S\left(1 + \frac{1}{tg^2\delta}\right) \tag{4.70}$$

Сопротивление R_P в параллельной схеме, как следует из выражения (4.70), во много раз больше сопротивления R_S. Угол потерь от выбора схемы не должен зависеть.

Для доброкачественных диэлектриков можно пренебречь значением $tg\delta$ по сравнению с единицей в формуле (4.69) и считать $C_P \approx C_S$. Выражения для мощности, рассеиваемой в диэлектрике, в этом случае будут также одинаковы у обеих схем:

$$P_a = U^2 \omega C tg\delta. \qquad (4.71)$$

Следует отметить, что при переменном напряжении, в отличие от постоянного, емкость конденсатора с диэлектриком, у которого большие потери, становится условной величиной и зависит от выбора эквивалентной схемы. Отсюда и диэлектрическая проницаемость материала с большими потерями при переменном напряжении также условна.

4.1.3.4. Виды диэлектрических потерь

Диэлектрические потери могут быть вызваны следующими основными причинами: сквозной проводимостью, релаксационной поляризацией, ударной ионизацией, а также явлениями резонанса.

Потери на проводимость существенны в диэлектриках, имеющих заметную *объемную проводимость* или *поверхностную проводимость*. Если при этом потери от других механизмов несущественны, то частотные зависимости P_a и $tg\delta$, могут быть получены при использовании параллельной эквивалентной схемы замещения реального диэлектрика. Диэлектрические потери этого вида не зависят от частоты приложенного напряжения, а $tg\delta$ уменьшается с частотой по гиперболическому закону.

Потери при сквозной электропроводности возрастают с ростом температуры по экспоненциальному закону

$$P = Ae^{-b/T} \qquad (4.72)$$

где A, b—постоянные материала;

В зависимости от температуры $tg\delta$ изменяется по тому же закону, так как можно считать, что реактивная мощность от температуры практически не зависит.

Потери, обусловленные *релаксационными видами поляризации*, имеют место в диэлектриках с дипольной структурой и в диэлектриках с ионной структурой с неплотной упаковкой ионов. Тангенс угла диэлектрических потерь полярных диэлектриков с ионной структурой с неплотной упаковкой ионов зависит от температуры и частоты (рис. 4.19). Наличие температурного максимума $tg\delta$ в полярном диэлектрике можно объяснить на примере дипольно-релаксационной поляризции (рис. 4.19, *а*). При невысоких температурах межмолекулярные связи достаточно велики, молекулы не успевают следовать за изменением поля, дипольно-релаксационная поляризация практически отсутствует, $tg\delta$ мал. С ростом температуры растет количество полярных молекул, участвующих в поляризации, расходуется дополнительная энергия на их поляризацию, поэтому $tg\delta$ возрастает.

Достигнув максимального значения, $tg\delta$ начинает уменьшаться, так как дальнейшее повышение температуры настолько усиливает хаотическое тепловое движение полярных молекул, что затрудняет их поворот в электрическом поле. Достигнув наименьшего значения, $tg\delta$ вновь возрастает, что связано с увеличением сквозного тока.

На рис (4.19, *б*) показан график зависимости $tg\delta$ от частоты для полярного диэлектрика. Частотный максимум $tg\delta$ объясняется следующим образом. При возрастании частоты число поворотов диполей за единицу времени увеличивается, следовательно, увеличивается и рассеиваемая мощность и $tg\delta$. При достаточно высокой частоте $tg\delta$ начинает уменьшаться, так как полярные молекулы не успевают следовать за изменением электрического поля. Положение частотного максимума (критической частоты $f_{кр}$) определяется из условия

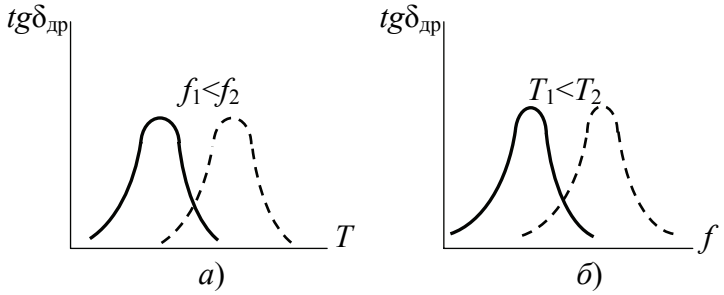


Рис. 4.19 – Зависимость релаксационных потерь в диэлектриках при дипольно-релаксационной поляризации: *а* – от температуры; *б* – от частоты

$$f_{кр}=1/2\pi\tau_0=\frac{kT}{8\pi\eta r^3}, \qquad (4.73)$$

где τ_0 – время релаксации, с; k – постоянная Больцмана, Дж/К; T – температура, К; η – динамическая вязкость вещества, Па·с; r – радиус молекулы вещества, условно принимаемой за шарообразную, м.

С повышением частоты максимум $tg\delta$ смещается в область более высоких температур, так как с повышением температуры межмолекулярные связи ослабляются и время релаксации снижается.

Потери, вызванные *ударной ионизацией*, свойственны пористым и слоистым диэлектрикам с газовыми включениями. При повышении напряжения сверх определенного значения $U_{ион}$, называемого порогом ионизации, в газовых включениях и других дефектах внутри диэлектрика возникает процесс ударной ионизации, который приводит к рассеиванию энергии электрического поля. При напряжении $U_к$, соответствующем окончанию процесса ионизации, газ уже ионизирован, поэтому энергия поля больше не расходуется на ионизацию, вследствие чего $tg\delta$ уменьшается.

Ионизация газовых включений особенно опасна для неорганических диэлектриков с закрытыми порами, так как может вызвать местный нагрев изделий и их разрушение. Для повышения качества такой электрической изоляции ее пропитывают, заполняя поры маслами, лаками, компаундами.

Резонансные потери наблюдаются в некоторых газах при строго определенной частоте и выражаются в интенсивном поглощении энергии электромагнитного поля. Резонансные потери возможны и в твердых телах, если частота вынужденных колебаний, вызываемая электрическим полем,

совпадает с частотой собственных колебаний частиц твердого вещества. Наличие максимума в частотной зависимости $tg\delta$ характерно и для резонансного механизма потерь, но при изменении температуры максимум $tg\delta$ не смещается.

4.1.3.5. Пробой твердых диэлектриков

Диэлектрик, находясь в электрическом поле, может потерять свойства изоляционного материала, если напряженность поля превысит некоторое критические значение. Явление необратимого образования проводящего канала в диэлектрике называют *пробоем*.

Существует два различных вида пробоя: *тепловой* и *электрический*. Оба явления могут происходить одновременно (наиболее часто встречающийся случай) и, более того, поддерживать друг друга. Однако можно создать такие условия, которые позволят экспериментально изучить один или другой механизм в чистом виде.

Тепловой пробой. Тепловой пробой возникает в том случае, когда количество тепловой энергии, выделяющейся в диэлектрике за счет диэлектрических потерь, превышает то количество энергии, которое может рассеиваться в данных условиях; при этом нарушается тепловое равновесие, а процесс приобретает лавинный характер.

Явление теплового пробоя приводит к разогреву материала в электрическом поле до температур, соответствующих плавлению, обугливанию и пр. Пробивное напряжение, обусловленное нагревом диэлектрика (так называемое *тепловое пробивное* напряжение) зависит от частоты приложенного напряжения, условий охлаждения, температуры окружающей среды и т.д. Кроме того, оно зависит и от *нагревостойкости* материала: органические диэлектрики (например, полистирол) имеют более низкие значения тепловых пробивных напряжений, чем неорганические (кварц, керамика) при прочих равных условиях.

Механизм теплового пробоя более вероятен при повышении температуры, когда преобладающими становятся потери от сквозной проводимости. Типичными признаками теплового пробоя являются экспоненциальное уменьшение пробивного напряжения с ростом температуры, а также снижение электрической прочности с увеличением времени выдержки диэлектрика в электрических полях. Значение рассеиваемой в диэлектрике мощности определяется выражением (4.71). При использовании экспоненциальной зависимости тангенса угла потерь от температуры

$$tg\delta = tg\delta_0 \cdot \exp\left[\alpha\left(T - T_0\right)\right] \qquad (4.74)$$

и после преобразования, с учетом формулы (4.74), выражения (4.71) для рассеиваемой мощности получим

$$P_a = \frac{U^2 \cdot f \cdot \varepsilon \cdot S \cdot tg\delta_0 \cdot \exp\left[\alpha\left(T - T_0\right)\right]}{1{,}8 \cdot 10^{10} \cdot h}, \qquad (4.75)$$

где U – приложенное напряжение, f – частота, ε – диэлектрическая проницаемость, $tg\delta_0$ – тангенс угла потерь диэлектрика при температуре окружающей среды, α – температурный коэффициент тангенса угла диэлектрических потерь, T – температура нагретого за счет диэлектрических потерь материала, T_0 – температура окружающей среды, S – площадь электрода, h – толщина диэлектрика.

Так как теплопроводность металла электродов, за редким исключением, на два-три порядка больше, чем теплопроводность диэлектрика, будем считать, что теплоотвод из нагревающегося объема в окружающую среду происходит через электроды (теплоотводом через торцевую поверхность диэлектрика пренебрегаем). Мощность, отводимую от диэлектрика, выразим с помощью формулы Ньютона:

$$P_T = 2\sigma S(T - T_0), \tag{4.76}$$

где σ – суммарный коэффициент теплопередачи от диэлектрика во внешнюю среду.

Напряжение теплового пробоя можно определить по двум условиям:

$$P_a = P_T, \tag{4.77}$$

$$\frac{\partial P_a}{\partial T} = \frac{\partial P_T}{\partial T}. \tag{4.78}$$

Условие (4.78) справедливо только для граничного режима, тогда как условие (4.77) выполняется для всех случаев устойчивой работы диэлектрика под напряжением.

Подставив выражения (4.75) и (4.76) в (4.77) и (4.78), получим

$$\frac{U^2 \cdot f \cdot \varepsilon \cdot S \cdot tg\delta_0 \cdot \exp\left[\alpha\left(T - T_0\right)\right]}{1{,}8 \cdot 10^{10} \cdot h} = 2\sigma S(T - T_0), \tag{4.79}$$

$$\frac{\alpha \cdot U^2 \cdot f \cdot \varepsilon \cdot S \cdot tg\delta_0 \cdot \exp\left[\alpha\left(T - T_0\right)\right]}{1{,}8 \cdot 10^{10} \cdot h} = 2\sigma S. \tag{4.80}$$

Разделив (4.79) на (4.80), получим

$$\frac{1}{\alpha} = T - T_0. \tag{4.81}$$

Подставляя (4.81) в (4.80) и решая его относительно U, найдем напряжение теплового пробоя

$$U_{\text{пр}} = K\sqrt{\frac{\sigma \cdot h}{f \cdot \varepsilon \cdot \alpha \cdot \text{tg}\delta_0}}, \tag{4.82}$$

где K – числовой коэффициент, равный $1{,}15 \cdot 10^5$, если все величины, имеющие размерности, выражены в системе СИ.

Разновидностью теплового пробоя можно считать *ионизационный пробой*. Он характерен для твердых пористых диэлектриков и обусловлен ионизацией газа в порах. За счет ионизационных потерь разогревается поверхность закрытых пор, возникают локальные перепады температуры в диэлектрике и связанные с ними термомеханические напряжения. Такие процессы особенно опасны в хрупких материалах, поскольку

термомеханические напряжения могут превзойти предел прочности материала и вызвать растрескивание диэлектрика.

Электрический пробой. С помощью вышеописанного теплового механизма пробоя невозможно объяснить все наблюдаемые при пробое явления. Например, этот механизм не дает ответа на вопрос, почему при низких температурах напряженность поля пробоя практически не зависит от температуры и начинает уменьшаться лишь при повышенных температурах. Это показывают измерения, проведенные на каменной соли. Кроме того, необходимые для теплового пробоя условия выполняются лишь при малой толщине слоя, а также при достаточно длительной нагрузке. Электрическое сопротивление изолятора уменьшается не только в случае увеличения температуры, но и из-за наличия электрического поля. Можно предположить, что происходит образование лавин носителей, когда источником ионизации служат ионные и электронные столкновения. Оценка напряженности поля, необходимого для создания новых ионов в кристалле при ионных столкновениях, приводит к напряженности пробоя порядка 10^{10} В/м, что примерно в 100 раз превышает измеренные значения. Если учесть, что в столкновениях участвуют и свободные электроны, средняя длина свободного пробега которых приблизительно в 100 раз больше, чем у ионов, то приходим к выводу, что пробой происходит в основном в результате электронных столкновений.

Откуда же в диэлектрике берутся свободные электроны? С точки зрения зонных представлений (см. гл. III) в зоне проводимости беспримесного изолятора содержатся электроны, концентрация которых определяется по формуле $n_i = \sqrt{N_C N_V} \cdot e^{-\frac{E_g}{2kT}}$. В таком случае их концентрация исчезающе мала. Но так как в технических диэлектриках всегда имеются загрязнения, дающие примесные центры, то происхождение свободных электронов объяснимо.

Различают несколько видов электрического пробоя: *электрический пробой однородных диэлектриков*, *электрический пробой неоднородных диэлектриков*, *электрохимический пробой*.

Электрический пробой однородных диэлектриков по своей природе является чисто электронным процессом, заключающимся в образовании в диэлектрике электронной лавины из небольшого числа электронов под действием сильного электрического поля. Развитие лавины происходит вследствие ударной ионизации, которая ускоряет образование проводящего канала. Ускоренные полем электроны при столкновении передают свою энергию узлам решетки и разогревают ее вплоть до плавления. В разрядном канале создается значительное давление, которое может привести к появлению трещин или разрушению диэлектрика.

Этот вид пробоя протекает за время не более 10^{-7} секунд. Электрическая прочность при электрическом пробое в некоторой степени зависит от

температуры и сопровождается в начальной стадии разрушением диэлектрика в очень узком канале.

Большое значение имеет неоднородность электрического поля. Пробивные напряжения в однородном и резко неоднородном полях сильно различаются. Таким образом, одни и те же образцы в полях разной степени неоднородности обнаруживают разные значения электрической прочности.

Любой диэлектрик может быть использован только при напряженностях поля, не превышающих некоторого предельного значения. Если напряженность поля превысит некоторое критическое значение, произойдет *пробой диэлектрика*, при этом сквозной ток резко возрастает, а сопротивление диэлектрика падает до очень низкой величины. Схематически зависимость тока через диэлектрик от напряжения может быть представлена графиком (рис. 4.20).

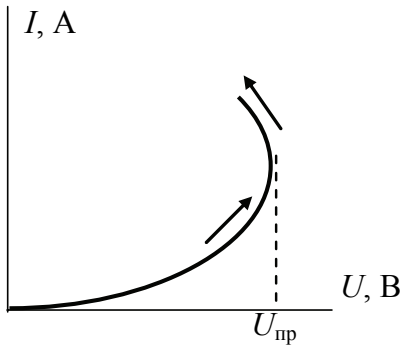


Рис. 4.20 – Зависимость тока от напряжения при пробое диэлектриков

В месте пробоя возникает искра или даже электрическая дуга, которая может вызвать оплавление, обгорание, растрескивание и тому подобные изменения, как диэлектрика, так и электродов. После снятия напряжения в пробитом твердом диэлектрике можно обнаружить след в виде пробитого (откуда и название явления пробой), проплавленного, прожженного и т. п. отверстия, обычно неправильной формы. При повторном приложении напряжения к ранее подвергавшейся пробою твердой изоляции, пробой происходит при более низком напряжении по месту прежнего пробоя.

Наибольшее значение напряжения, которое было приложено к диэлектрику в момент пробоя, называют *пробивным напряжением* $U_{пр}$, а соответствующую напряженность поля – *электрической прочностью* диэлектрика $E_{пр}$. Электрическая прочность определяется величиной пробивного напряжения, отнесенного к толщине диэлектрика h в месте пробоя

$$E_{пр} = U_{пр}/h. \qquad (4.83)$$

Электрическую прочность удобно выражать в МВ/м.

В большинстве случаев при возрастании h величина $E_{пр}$ уменьшается, т. е. величина $U_{пр}$ возрастает с увеличением толщины изоляции не линейно, а медленнее. Это связано с увеличивающейся вероятностью появления дефектного участка диэлектрика. Таким образом, тонкие пленки могут обладать существенно более высокой электрической прочностью, чем массивные образцы. Это свойство получило название *электрического упрочнения материалов*. Его использование позволяет повысить надежность пленочной изоляции микроэлектронных элементов и устройств, так как эксплуатационные напряженности поля в тонких пленках порядка 10^6 В/м.

В объемных образцах такие напряженности приводили бы к электрическому пробою. Однако при переходе к особо тонким слоям начинают сказываться неизбежные неоднородности толщины материала и $E_{пр}$ опять снижается.

Электрический пробой неоднородных диэлектриков. Пробой неоднородных диэлектриков чаще всего наблюдают в пористых твердых диэлектриках. Пористость определяется самой природой диэлектрика, например у волокнистых материалов, или появляется в процессе эксплуатации, например, в слоистой изоляции на основе слюды. Наличие газовых (воздушных) включений в таких диэлектриках приводит при повышенных напряжениях к развитию ударной ионизации газа. Очаги внутренней ионизации в порах действуют разрушающе на основной твердый компонент изоляции вследствие бомбардировки ионами и электронами. Электрический пробой неоднородных диэлектриков также отличается достаточно быстрым развитием. Пробивные напряжения невысоки и весьма мало отличаются друг от друга в однородном или неоднородном полях.

У многих технических диэлектриков электрическая прочность практически не зависит от температуры до некоторого ее значения. Выше этого значения она резко падает с увеличением температуры, что говорит о появлении электротеплового пробоя.

На значение электрической прочности диэлектрика непосредственно влияет вид напряжения и время его воздействия. Так, при кратковременных импульсах пробой происходит при бóльших напряжениях, чем в случае постоянного или длительно приложенного переменного напряжения.

Таким образом, пробивное напряжение при электрическом и тепловом пробоях зависит от ряда факторов: частоты электрического поля, условий охлаждения, температуры окружающей среды. Кроме того, напряжение электротеплового пробоя связано с *нагревостойкостью* материала. *Нагревостойкость* – способность диэлектрика выдерживать действие повышенной температуры в течение времени, сравнимого со временем эксплуатации, без недопустимого ухудшения его свойств. Для электроизоляционных материалов установлено семь классов нагревостойкости. Органические диэлектрики вследствие малой нагревостойкости при прочих равных условиях имеют более низкие значения пробивных напряжений, чем неорганические. Электротепловой пробой наиболее вероятен при повышенных температурах (как самого диэлектрика, так и окружающей среды), при высокой частоте и при длительном воздействии напряжения. При кратковременных импульсных напряжениях электротепловой пробой обычно не успевает развиться.

В условиях, когда принципиально возможен пробой любого вида, надо применять электроизоляционные материалы с малым тангенсом потерь $tg\delta$, высокой температурной стабильностью и повышенной теплопроводностью.

Электрохимический пробой обусловлен химическими процессами, происходящими в диэлектрике под воздействием электрического поля. Химические изменения приводят к деградации диэлектрика, т. е. к

постепенному необратимому уменьшению сопротивления, и, следовательно, к постоянному снижению электрической прочности. Деградация особенно заметно проявляется при постоянном напряжении и низких частотах, в условиях повышенных температуры и влажности воздуха. Электрохимический пробой для своего развития требует длительного времени.

Старение свойственно, в большей степени, органическим диэлектрикам. Оно обусловлено, прежде всего, развитием ионизационного процесса в воздушных включениях с выделением озона и оксидов азота. Это приводит к постепенному химическому разрушению изоляции. Старение может иметь место и в некоторых неорганических диэлектриках. Например, в титановой керамике, содержащей окислы металлов переменной валентности, электрохимический пробой встречается значительно чаще, чем в керамике, состоящей из окислов алюминия, кремния, магния, бария.

4.1.4. Оптические свойства диэлектриков

Большой интерес для практического использования в электронике и оптической связи представляют оптические явления, происходящие в диэлектриках. Поэтому рассмотрим более подробно некоторые оптические явления, без которых принцип работы многих приборов будет неясен.

4.1.4.1. Дисперсия света

Основным материалом для изготовления оптических волокон является чистый плавленый кварц (SiO_2), показатель преломления которого может быть уменьшен путем добавления примесей фтора (F) или окисла бора (B_2O_3), или увеличен за счет введения примесей окисла германия (GeO_2), фосфорного ангидрида (P_2O_5) и др.

Экспериментальные результаты по измерению показателя преломления n от длины волны λ в объемных образцах компонентного стекла на основе кварца с высокой точностью апроксимируются в диапазоне оптической прозрачности диэлектрика трехчленной формулой Селмейера:

$$n^2 - 1 = \sum_{j=1}^{3} \frac{A_j \lambda^2}{\lambda^2 - l_j^2}, \qquad (4.84)$$

где A_j, l_j – *коэффициенты Селмейера* (постоянные для данного диэлектрика величины), которые находятся подгонкой измеренных значений показателей преломления объемного образца стекла. Можно найти спектральные зависимости показателя преломления в объемном кварцевом стекле с определенной концентрацией различных легирующих примесей при использовании интерполяционных многочленов Селмейера как функции мольной концентрации примеси. За узловые точки интерполяции Селмейера берутся экспериментально найденные значения этих коэффициентов для известных концентраций выбранной примеси. Так, например, для длины волны λ = 0,8 мкм показатель преломления для стекол равен: n = 1,453 (100%

SiO_2); n = 1,458 (3,1% GeO_2 , 96,9% SiO_2); n = 1,466 (7,9% GeO_2, 92,1% SiO_2); n = 1,475 (13,5% GeO_2, 86,5% SiO_2). Видно, что при увеличении концентрации окисла германия (GeO_2) показатель преломления увеличивается. Если необходимо уменьшить показатель преломления, то необходимо добавить примеси фтора (F), тогда n = 1,449 (1% F, 99% SiO_2) или примесь окисла бора (B_2O_3), тогда n = 1,448 (13,3% B_2O_3, 86,7% SiO_2). Таким образом, влияя на свойства материала путем добавления различных примесей, мы можем управлять его оптическими характеристиками.

Среда является дисперсной, если показатель преломления зависит от частоты волны, т.е. $n = n(\omega)$. Показатель преломления n в данной точке среды равен значению отношения скорости света в вакууме к фазовой скорости монохроматической плоской электромагнитной волны

$$n = \frac{c}{\upsilon_ф} = \frac{\beta}{k} \rightarrow \upsilon_ф = \frac{c}{n}, \qquad (4.85)$$

где $k = \frac{\omega}{c} = \frac{2\pi}{\lambda}$ есть волновое число в вакууме, λ – длина волны в вакууме.

Диапазон частот, в котором показатель преломления растет с ростом частоты волны, называется участком с *нормальной дисперсией среды*. Например, нормальная дисперсия наблюдается в веществах, прозрачных для света, когда потери энергии световой волны незначительны. В линиях оптической связи, при использовании многокомпонентного стекла на базе плавленого кварца, работает диапазон длин волн ближнего и среднего инфракрасного диапазона. В этом случае показатель преломления медленно и монотонно возрастает с ростом частоты (т. е. уменьшается с ростом длины волны λ). Такое стекло имеет нормальную дисперсию показателя преломления.

Дисперсия среды называется *аномальной*, если с ростом частоты волны показатель преломления уменьшается. Аномальная дисперсия наблюдается в областях, которые отвечают узким полосам интенсивного поглощения света в данной среде. Например, для обычных стекол эти полосы находятся в инфракрасном и ультрафиолетовом участках спектра оптического излучения.

4.1.4.2. Полное внутреннее отражение

Если свет распространяется из среды с большим показателем преломления n_1 (оптически более плотной) в среду с меньшим показателем преломления n_2 (оптически менее плотную), например, из стекла в воду, то согласно закону Снеллиуса $\frac{\sin i_2}{\sin i_1} = \frac{n_1}{n_2} > 1$ преломленный луч удаляется от нормали и угол преломления i_2 больше, чем угол падения i_1. С увеличением угла падения увеличивается угол преломления до тех

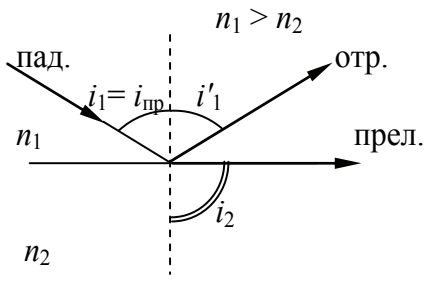


Рис. 4.21 – Полное внутреннее отражение луча при $i_1 = i_{пр}$

пор, пока при некотором угле падения ($i_1 = i_{пр}$) угол преломления не окажется равным $\pi/2$. Угол $i_{пр}$ называется *предельным углом*. При углах падения $i_1 > i_{пр}$ весь падающий свет полностью отражается от поверхности раздела сред.

По мере приближения угла падения к предельному углу, интенсивность преломленного луча уменьшается, а отраженного растет. Если $i_1 = i_{пр}$, то интенсивность преломленного луча обращается в нуль, а интенсивность отраженного луча равна интенсивности падающего (рис. 4.21). Таким образом, при углах падения от $i_{пр}$ до $\pi/2$ луч не преломляется, а полностью отражается в первую среду, причем интенсивность падающего и отраженного луча одинаковы. Это явление называется *полным внутренним отражением*.

Предельный угол $i_{пр}$ определим из формулы Снеллиуса при подстановке в нее $i_2 = \pi/2$. Тогда

$$\sin i_{пр} = \frac{n_2}{n_1} = n_{21}. \qquad (4.86)$$

Это уравнение удовлетворяет значениям угла $i_{пр}$ при $n_2 \le n_1$.

Явление полного внутреннего отражения используется в волоконной оптике (световодах), в рефрактометрах (позволяет найти показатель преломления тел) и т. д.

4.1.4.3. Поглощение света

Поглощением света называется явление потери энергии световой волной, проходящей через вещество, вследствие преобразования энергии в другие формы (в основном в тепловую). В результате поглощения интенсивность света при прохождении через вещество уменьшается.

Поглощение света в веществе описывается законом Бугера:

$$I = I_0 e^{-\alpha x}, \qquad (4.87)$$

где I_0 и I – интенсивности плоской монохроматической волны на входе и на выходе слоя поглощающего вещества толщиной x, α – коэффициент поглощения, зависящий от длины волны света, химической природы и состояния вещества и не зависящий от интенсивности света.

Коэффициент поглощения для диэлектриков невелик (примерно 10^{-3} – 10^{-5} см$^{-1}$), однако у них наблюдается селективное поглощение света в определенных интервалах длин волн, когда α резко возрастает, и наблюдаются сравнительно широкие полосы поглощения, это связано с явлением резонанса электронов в атомах диэлектрика.

4.1.4.4. Поляризация света

Поляризация света – это физическая характеристика оптического излучения, описывающая световую анизотропию электромагнитных волн, т. е. неэквивалентность различных направлений в плоскости, перпендикулярной световому лучу. Поперечность электромагнитной волны лишает ее осевой симметрии относительно направления ее распространения из-за наличия выделенных направлений (вектора $\vec{E}$ – напряженности

электрического поля, вектора $\vec{H}$ – напряженности магнитного поля) в плоскости, перпендикулярной направлению вектора скорости волны.

Все типы поляризации света можно рассмотреть на примере монохроматической электромагнитной волны, составляющие которой (вектора $\vec{E}$ и $\vec{H}$) меняются вдоль осей x и y по гармоническому закону. Рассмотрим поляризацию электрической составляющей электромагнитной волны (обычно все рассуждения проводятся для вектора $\vec{E}$, называемого световым, подразумевая также наличие и вектора $\vec{H}$). Если разность фаз компонентов $\vec{E}$ равна нулю, мы получим стабильные годографы $\vec{E}$, которые могут иметь вид, изображенный на рис. 4.22: *а*) свет, полностью поляризованный в вертикальной плоскости; *б*) свет, полностью поляризованный в горизонтальной плоскости; *в*) естественный (неполяризованный) свет; *г*) свет, частично поляризованный в вертикальной плоскости.

Свет, несущий равную энергию в волне с любой ориентацией вектора $\vec{E}$, называется естественным или неполяризованным светом. Такой свет излучают нагретые тела.

Свет, несущий различную энергию в волне с различным положением $\vec{E}$, называют частично поляризованным. Частично поляризованный свет можно представить как наложение неполяризованного света на полностью

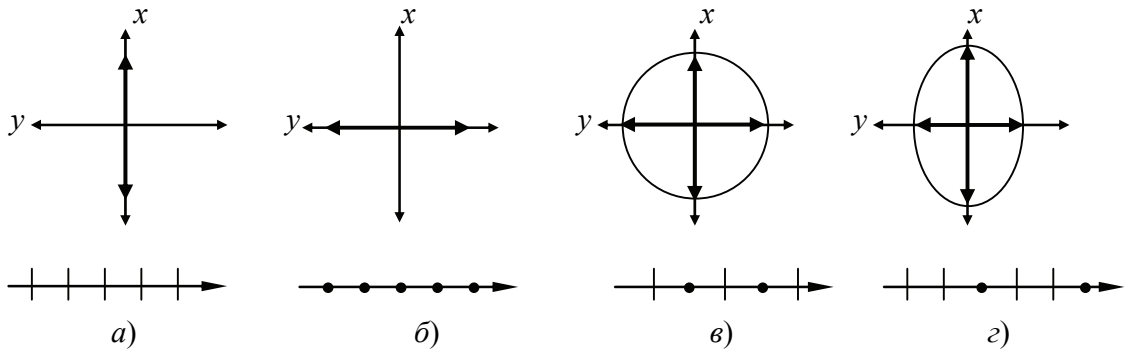


Рис. 4.22 – Годографы: *а*) свет, полностью поляризованный в вертикальной плоскости; *б*) свет, полностью поляризованный в горизонтальной плоскости; *в*) естественный (неполяризованный) свет; *г*) свет, частично поляризованный в вертикальной плоскости. Ниже годографов показаны условные обозначения соответствующих волн на лучах

поляризованный свет. Он характеризуется коэффициентом поляризации

$$P = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \qquad (4.88)$$

где $I_{\max}$ – интенсивность света в волне, несущей наибольшую энергию; $I_{\min}$ – в волне, несущей наименьшую энергию. Для случая *а*) и *б*) $P = 1$; для случая *в*) $P = 0$; для случая *г*) $0 < P < 0$.

Поляризацию света принято называть правой, если вектор $\vec{E}$ совершает вращение по часовой стрелке при наблюдении навстречу световому лучу, и левой – если вектор $\vec{E}$ совершает вращение против часовой стрелки.

Для поляризации естественного света используются специальные устройства, называемые поляризаторами. *Поляризаторы* – это оптические приборы, которые при освещении любым светом пропускают без ослабления только полностью поляризованный свет, плоскость поляризации которого совпадает с главной плоскостью поляризатора. Прошедший через поляризатор свет подчиняется закону Малюса

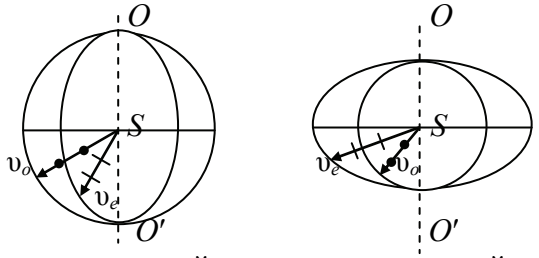
$$I = I_0 \cos^2 \varphi, \qquad (4.89)$$

где I_0 – интенсивность падающего линейно поляризованного света, I – интенсивность линейно поляризованного света после прохождения поляризатора, φ – угол между плоскостью поляризации падающей волны и главной плоскостью поляризатора. Естественный свет, прошедший через поляризатор, становится линейно поляризованным (плоскость его поляризации совпадает с главной плоскостью поляризатора), а его интенсивность уменьшается в два раза, так как среднее значение квадрата косинуса для углов в пределах 360 º равно 0,5.

4.1.4.5. Оптическая анизотропия

Все прозрачные кристаллы за исключением центральносимметричных обладают способностью двойного лучепреломления. Это явление впервые обнаруженно датским ученым Э. Бартолином в 1669 г. для исландского шпата (разновидность кальцита $CaCO_3$).

Допустим, что в точке S внутри одноосного кристалла находится точечный источник света. На рис. 4.23 показано распространение обыкновенного и необыкновенного лучей в кристалле (главная плоскость совпадает с плоскостью чертежа, OO' – направление оптической оси). Волновой поверхностью обыкновенного луча (он распространяется со скоростью υ_o = const) является сфера, необыкновенного луча ($\upsilon_e \neq$ const) – эллипсоид вращения. Наибольшее расхождение волновых поверхностей обыкновенного и необыкновенного лучей наблюдается в направлении, перпендикулярном оптической оси. Эллипсоид и сфера касаются друг друга в точках их пересечения с оптической осью OO'. Если $\upsilon_e < \upsilon_o$ ($n_e > n_o$), то эллипсоид необыкновенного луча вписан в сферу обыкновенного луча и одноосный кристалл называется положительным (рис. 4.23, *а*). Если $\upsilon_e > \upsilon_o$ ($n_e < n_o$), то эллипсоид описан вокруг сферы и одноосный кристалл называется отрицательным (рис. 4.23,



положительный для кварца
$n_{e\,\max} = 1{,}555$
$n_o = 1{,}546$
а)

отрицательный для кальцита
$n_{e\,\max} = 1{,}488$
$n_o = 1{,}662$
б)

Рис. 4.23 – Волновые поверхности в положительном (*а*) и отрицательном (*б*) кристаллах

б). Исландский шпат относится к отрицательным кристаллам, кварц – к положительным кристаллам.

Пусть на кристаллическую пластинку, вырезанную параллельно оптической оси, падает под прямым углом линейно поляризованный свет. Внутри пластинки он разбивается на обыкновенный и необыкновенный лучи, которые в кристалле пространственно не разделены, хотя движутся с разными скоростями. На выходе из кристалла эти лучи складываются.

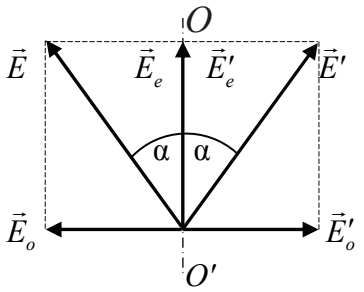


Рис. 4.24 – Изменение направления поляризации света после прохождения полуволновой пластинки

Между обыкновенным и необыкновенным лучами в пластинке возникает оптическая разность хода

$$\Delta_{\text{опт}} = (n_o - n_e)d\,, \qquad (4.90)$$

или разность фаз

$$\Delta\varphi = \frac{2\pi}{\lambda_0}(n_o - n_e)d\,, \qquad (4.91)$$

где d – толщина пластины, λ_0 – длина волны в вакууме.

Пластинка, для которой

$$(n_o - n_e)d = \pm(m + \frac{1}{2})\lambda_0;\;\; m = 0,1,2\ldots, \qquad (4.92)$$

называется полуволновой пластинкой. На выходе из такой пластинки колебания векторов $\vec{E}'_o$ и $\vec{E}'_e$ сдвинуты по фазе на π ($\Delta\varphi = \frac{2\pi}{\lambda_0}\cdot\frac{\lambda_0}{2} = \pi$). Свет, выходящий из пластинки, остается линейно поляризованным. Однако направление колебаний векторов падающего света $\vec{E}$ и проходящего света $\vec{E}'$ симметричны относительно главной плоскости пластины (рис. 4.24). Таким образом, пластинка в полволны поворачивает плоскость колебаний прошедшего через нее света на угол 2α (α – угол между плоскостью колебаний в падающем луче и осью пластинки).

4.1.4.6. Искусственная оптическая анизотропия

Особый интерес для технического применения представляет искусственная оптическая анизотропия в диэлектриках, так как этими явлениями можно управлять при помощи электрических и магнитных полей.

Еще в начале прошлого столетия Т. Зеебек (1813) и Д. Брюстер (1816) обнаружили явление *фотоупругости*, состоящее в том, что оптически изотропное твердое тело под влиянием механической деформации становится оптически анизотропным.

Например, при одностороннем сжатии или растяжении стеклянной пластинки она приобретает свойства одноосного кристалла, оптическая ось которого совпадает с направлением сжатия или растяжения. Разность показателей преломления обыкновенного и необыкновенного лучей в направлении оптической оси, пропорциональна нормальному напряжению σ:

$$n_o - n_e = k\sigma, \qquad (4.93)$$

где n_o – показатель преломления для обыкновенного луча; n_e – показатель преломления для необыкновенного луча; k – коэффициент, зависящий от свойств вещества.

Явление искусственной оптической анизотропии при деформациях используется для обнаружения внутренних напряжений, которые могут возникать в изделиях из стекла и других прозрачных изотропных материалов вследствие механической нагрузки и несоблюдения технологии их изготовления. Оптический метод изучения распределения внутренних напряжений в различных непрозрачных частях машин и сооружений на прозрачных моделях широко применяется в современной технике.

Появление двойного лучепреломления диэлектрика во внешнем электрическом поле называется линейным электрооптическим *эффектом Поккельса*. Устройства, применяющие эффект Поккельса, нашли свое применение в оптоволоконной технике в качестве высокочастотных модуляторов света (более подробно см. разд. 4.3).

Эффект Керра – название трех явлений, два из которых были открыты Дж. Керром в 1875 г. (электрооптический эффект) и в 1876 г. (магнитооптический эффект); после появления лазеров в сильных электромагнитных полях был замечен эффект, аналогичный электрооптическому эффекту Керра.

Электрооптический эффект Керра – это квадратичный электрооптический эффект, появляющийся в результате двойного лучепреломления в оптически изотропных веществах (газах, жидкостях, кристаллах с центром симметрии, стеклах) под действием внешнего однородного электрического поля. Оптически изотропная среда, помещенная в электрическое поле, становиться анизотропной и приобретает свойства одноосного кристалла, оптическая ось которого направлена вдоль электрического поля. Разность показателей преломления вещества для необыкновенного и обыкновенного лучей монохроматического света при его распространении в направлении, перпендикулярном вектору $\vec{E}$, пропорциональна E^2:

$$n_e - n_{\text{o}} = B\lambda_0 E_{\text{вн}}^2, \qquad (4.94)$$

где λ_0 – длина волны света в вакууме, B – *постоянная Керра*. Постоянная Керра зависит от природы вещества и быстро уменьшается с увеличением температуры. Кроме того, постоянная Керра обладает дисперсией (обычно увеличивается при уменьшении λ) и может быть положительной или отрицательной.

Классическая теория эффекта Керра для веществ с неполярными анизотропными молекулами (у анизотропной молекулы ее поляризуемость не одинакова в разных направлениях) была разработана П. Ланжевеном (1910) и развита для веществ с полярными молекулами М. Борном (1918). Неполярные анизотропные молекулы поляризуются во внешнем электрическом поле и стремятся ориентироваться так, чтобы направления их

максимальной поляризуемости совпадали с направлением $E_{\text{вн}}$. В связи с этим в электрическом поле относительная диэлектрическая проницаемость ε и показатель преломления *n* начинают зависеть от направления, среда становится оптически анизотропной (волновая поверхность необыкновенных лучей становится эллипсоидом вращения).

Если вещество состоит из полярных молекул, обладающих постоянными электрическими моментами, то во внешнем электрическом поле возникает преимущественная ориентация векторов электрических моментов молекул по направлению $E_{\text{вн}}$, что также приводит к оптической анизотропии среды. Ориентационный механизм установления оптической анизотропии среды применим к газам и в меньшей степени к жидкостям, где значительную роль начинают играть неучтенные в теории межмолекулярные взаимодействия.

В случае сферически симметричных молекул, а также в твердых телах, где ориентационные степени свободы молекул «заморожены», эффект Керра носит чисто поляризационный характер. Действие поляризационного механизма сводится к тому, что исходно оптически изотропная молекула, поляризованная внешним электрическим полем, обнаруживает различия в оптической поляризуемости в направлениях вдоль и поперек поля. Фактически это уже нелинейный эффект взаимодействия поля с веществом.

Строгое теоретическое объяснение эффекта Керра может быть проведено лишь в рамках квантовой механики.

Эффект Керра обладает чрезвычайно малой инерционностью: время релаксации ~ 10^{-11} – 10^{-12} с. Это нашло широкое применение при создании быстродействующих *оптических затворов* и *модуляторов света*, необходимых для лазерной техники и скоростной фотографии.

Эффектом Коттона-Мутона называется возникновение оптической анизотропии у некоторых изотропных веществ (жидкостей, стекол, коллоидов) при помещении их в сильное внешнее магнитное поле.

Если молекулы среды анизотропны и обладают постоянными магнитными моментами, то они могут преимущественно ориентироваться магнитным полем. В достаточно сильных однородных магнитных полях вещество приобретает оптические свойства одноосного кристалла, оптическая ось которого совпадает по направлению с вектором $\vec{H}$ напряженности поля. Разность показателей преломления вещества для необыкновенного и обыкновенного лучей монохроматического света при его распространении в направлении, перпендикулярном вектору $\vec{H}$, пропорциональна H^2:

$$n_e - n_{\text{o}} = C\lambda_0 H^2, \qquad (4.95)$$

где *C* – постоянная Коттона-Мутона; λ_0 – длина волны света в вакууме. Значение *C* зависит от природы вещества, длины волны λ_0 и температуры. Двойное лучепреломление здесь можно обнаружить и измерить так же, как это делается при изучении эффекта Керра, так как физические законы обоих явлений аналогичны.

Эффект Фарадея. М. Фарадей (1845) экспериментально установил, что оптическая среда приобретает под действием внешнего магнитного поля способность вращать плоскость поляризации света, распространяющегося вдоль направления поля.

Это явление называется эффектом Фарадея, или магнитным вращением плоскости поляризации света. Угол поворота φ плоскости поляризации пропорционален длине пути света в веществе и напряженности магнитного поля: $\varphi = VHl$. Коэффициент пропорциональности V называется постоянной Вердье. Она зависит от природы вещества и длины волны света λ_0.

4.2. КЛАССИФИКАЦИЯ ДИЭЛЕКТРИКОВ

4.2.1. Пассивные диэлектрики

4.2.1.1. Твердые органические диэлектрики

В электроизоляционной технике применяют как синтетические полимеры, так и природные. Синтетические материалы могут быть получены двумя принципиально различными способами: полимеризацией и поликонденсацией.

Полимеризация – это химическая реакция, при которой из низкомолекулярного соединения (мономера) получают высокомолекулярное соединение (полимер) без изменения элементарного химического состава вещества.

Поликонденсация – это химическая реакция между различными низкомолекулярными соединениями (мономерами), при которой образование высокомолекулярного соединения идет с выделением побочных веществ, например, воды, водорода и др. Диэлектрики, полученные в результате поликонденсации, как правило, обладают несколько пониженными электроизоляционными свойствами, по сравнению с диэлектриками, полученными в результате полимеризации.

1. *Полимеризационные синтетические полимеры.*

- *Полимерные углеводороды*. К ним относят полистирол, полиэтилен, полипропилен, поливинилхлорид и др.

Полистирол – твердый прозрачный материал. Он является *неполярным диэлектриком*, с чем и связаны его высокие электроизоляционные свойства. Полистирол – термопластичный диэлектрик, размягчающийся при температуре 110…120°С.

Недостатками полистирола являются хрупкость при пониженных температурах, склонность к образованию поверхностных трещин, малая стойкость к действию растворителей. *Ударопрочные полистиры*, которые представляют собой смесь полистирола с синтетическими каучуками, отличаются повышенными значениями ударной вязкости и диэлектрической проницаемости. Остальные параметры практически остаются на том же уровне, что и у обычных видов полистирола.

В промышленности полистирол применяют главным образом в качестве высокочастотной изоляции, благодаря малому значению $tg\delta$. Ударопрочный

полистирол применяют для изготовления каркасов катушек индуктивности, оснований и изоляторов для электроизмерительных приборов и т. д.

В оптической промышленности, кроме неорганических материалов, также находят применение прозрачные пластмассы. К ним, прежде всего, относятся полистирол (аналогичен флинту) и полиметилметакрилат (аналогичен крону).

Способом вытягивания размягченного полистирола получают гибкие полистирольные пленки.

Полиэтилен – неполярный, термопластичный диэлектрик белого или светло-серого цвета, получаемый в результате реакции полимеризации газа этилена.

Полиэтилен практически не гигроскопичен, гибок. Электроизоляционные свойства полиэтиленов находятся на таком же высоком уровне, что и у полистиролов, и отличаются высокой стабильностью. Благодаря высоким электроизоляционным свойствам полиэтилен нашел широкое применение при изготовлении деталей радиотехнической аппаратуры, электроизоляционных пленок, его используют в качестве изоляции в различных проводах и кабелях (радиочастотных, монтажных, наземных и подводных телефонных и др.).

Полипропилен – это линейный неполярный полимер, получаемый из газа – пропилена CH_2=CH–CH3.

Электроизоляционные свойства полипропилена аналогичны свойствам полиэтилена. Однако полипропилен более холодостоек и гибок, чем полиэтилен. Из полипропилена могут быть получены пленки, волокна, ткани и фасонные изделия методом литья под давлением. Полипропилен можно применять как комбинированный бумажно-пленочный диэлектрик в силовых конденсаторах, как пленочный диэлектрик в обмоточных проводах и т. д.

Поливинилхлорид – линейный термопластичный полимер, полученный в результате полимеризации газообразного мономера–винилхлорида H_2C=CH–C1.

Поливинилхлорид – *полярный диэлектрик*, что обусловлено асимметрией строения его молекул из-за наличия атомов хлора. Вследствие полярного строения, поливинилхлорид имеет пониженные электрические свойства по сравнению с неполярными полимерами**,** но отличается большей стабильностью при изменении частоты. Его удельное сопротивление почти не изменяется при повышении температуры вплоть до 90 °С. Наряду с поливинилхлоридом (винипластом) в электроизоляционной технике широко применяют пластифицированный поливинилхлорид (пластикат). Пластикат обладает большой эластичностью, более высокой холодостойкостью (до –50°С), чем непластифицированный поливинилхлорид. Его недостатком является значительная зависимость удельного объемного сопротивления от температуры.

Полиметилметакрилат (органическое стекло, плексиглас) – это прозрачный бесцветный материал, получаемый в результате полимеризации эфиров метакриловой кислоты.

Полиметилметакрилат – полярный термопластичный диэлектрик с малой гигроскопичностью и значительной химической стойкостью.

Органическое стекло хорошо поддается механической обработке, легко сваривается в специальных устройствах при температуре 140…150 °С с применением давления на свариваемые поверхности 0,5…1 МПа, склеивается полярными растворителями. Применяют органическое стекло для изготовления корпусов приборов, шкал, линз и т. д.

- *Фторорганические полимеры*. Из этой группы наиболее широкое применение нашли *политетрафторэтилен* и *политрифторхлорэтилен*. Рассмотрим *политетрафторэтилен*, который получают путем полимеризации тетрафторэтилена $F_2C=CF_2$.

Благодаря симметричному расположению атомов в молекулах политетрафторэтилен неполярен. Промышленность выпускает этот материал под названием *фторопласт*-4 (тефлон). Цифра 4 указывает на число атомов фтора в молекуле мономера. Фторопласт-4 обладает необычайно высокой для органического вещества *нагревостойкостью* (порядка +250 °С), что объясняется высокой энергией связи между углеродом и фтором и экранирующим влиянием атомов фтора на связи между атомами углерода.

Фторопласт-4 – белый или сероватый материал с более высокой плотностью, чем плотность обычных органических полимеров. По электроизоляционным свойствам фторопласт-4 принадлежит к лучшим из известных диэлектриков. По химической стойкости он превосходит благородные металлы, что позволило широко использовать его при изготовлении изоляции, работающей в агрессивных средах. Он совершенно негорюч, практически не гигроскопичен, не смачивается водой и другими жидкостями.

К недостаткам политетрафторэтилена можно отнести его малую твердость и стойкость в зоне электрической дуги.

Широкому внедрению фторопласта-4 препятствует его высокая стоимость и сложность технологии. Это обстоятельство заставило разработать ряд разновидностей фторопласта-4.

Политрифторхлорэтилен является полимером трифторхлорэтилена $F_2C=CFCl$, известен под названием фторопласт-3.

Замена в элементарном звене одного атома фтора атомом хлора вызывает появление дипольного момента, в силу чего фторопласт-3 полярен и обладает значительными диэлектрическими потерями по сравнению с фторопластом-4. Нагревостойкость его ниже, чем у фторопласта-4 и составляет 125 °С, нижний предел рабочей температуры составляет -195 °С.

Технология получения фторопласта-3 проще, он дешевле, чем фторопласт-4, применяется при производстве кабелей.

2. *Поликонденсационные синтетические полимеры*. Из этой группы высокомолекулярных соединений в качестве электроизоляционных материалов наиболее широкое применение получили полиэфирные, кремнийорганические, эпоксидные, фенолоформальдегидные, полиамидные, полиуретановые полимеры (смолы).

- *Полиэфирные смолы* – продукт поликонденсации различных многоатомных спиртов (гликоля, глицерина и др.) и органических кислот (фталевой, малеиновой и др.) или их ангидридов. К полиэфирным смолам относят полиэтилентерефталат, поликарбонаты, глифталевые смолы и др.

Полиэтилентерефталат (*лавсан)* – прозрачный высокополимерный диэлектрик кристаллического или аморфного строения Лавсан имеет значительную механическую прочность и высокую температуру размягчения (примерно 260 °С). Его применяют для изготовления синтетических волокон, пряжи, тканей, тонких электроизоляционных пленок.

Поликарбонаты – полиэфиры угольной кислоты, обладающие хорошими механическими и электрическими свойствами. Поликарбонаты размягчаются при 140 °С, имеют хорошую химическую стойкость и невысокую гигроскопичность.

Поликарбонаты применяют для изготовления слоистых пластиков, компаундов, а пленки из поликарбоната используют в конденсаторном производстве и в качестве изоляции в электрических машинах и аппаратах.

Глифталевые смолы термореактивны и имеют ярко выраженные *дипольно-релаксационные потери.* Отличительными качествами этих смол являются их высокая клеящая способность при хороших электрических характеристиках, стойкость к поверхностным разрядам и высокая нагревостойкость (до 130 °С).

Глифталевые смолы в электротехнике используют как основу для клеящих, пропиточных и покрывных лаков, пленки которых оказываются стойкими к нагретому минеральному маслу. Клеящие глифталевые лаки широко применяют для склеивания слюды в производстве твердой и гибкой слюдяной изоляции (миканитов, микалент).

- *Кремнийорганические полимеры* представляют собой соединения, основу структурной формулы которых составляет цепь чередующихся атомов кремния и кислорода.

Свободные связи кремния заполнены органическими радикалами (этильными, фенольными и др.). Молекулы кремнийорганических полимеров могут иметь линейную структуру (термопластичные смолы) и пространственную структуру (термореактивные смолы).

Особенностью кремнийорганических соединений, выгодно отличающей их от органических и неорганических полимеров, является их высокая нагревостойкость (до +250 °С) и холодостойкость (до -60 °С). Они обладают весьма малой гигроскопичностью, химически инертны, являются слабополярными диэлектриками с хорошими электрическими параметрами, которые мало меняются при нагреве.

3. ***Компаунды***. *Компаунды* – это электроизоляционные материалы, состоящие из смеси различных смол, битумов, масел и др. В момент применения компаунды представляют собой жидкости, которые постепенно отвердевают, превращаясь в монолитный твердый диэлектрик. Наибольшее распространение получили синтетические компаунды (эпоксидные,

кремнийорганические, полиэфирные и др.). В состав компаундов, кроме основы, входят обычно наполнители и пластификаторы.

Эпоксидные компаунды представляют собой продукты модификации эпоксидных смол отвердителями, пластификаторами и наполнителями. Эпоксидные компаунды отличаются хорошей адгезией к металлам, керамике, пластмассам и к волокнистой изоляции. Они обладают повышенной механической прочностью и малой объемной усадкой (0,5…2,5%), что важно в производстве полупроводниковых приборов и интегральных схем.

Ассортимент эпоксидных компаундов в настоящее время очень широк. Так, в производстве полупроводниковых приборов и интегральных схем для герметизации кристаллов с *p-n*-переходами широко применяют эпоксидно-полиэфирные, эпоксидно-тиоколевые, эпоксидно-каучуковые и эпоксидно-кремнийорганические компаунды.

4. *Слоистые пластики и фольгированные материалы*. *Слоистые пластики* являются одной из разновидностей пластмасс, где связующим веществом служит полимер, а наполнителем – листовые волокнистые материалы. Из слоистых пластиков наиболее широко применяют *стеклотекстолит*, *гетинакс* и *текстолит*.

Фольгированные материалы – это слоистые пластики или синтетические пленки, облицованные с одной или двух сторон металлической (обычно медной) фольгой. Фольгированные материалы используют в качестве оснований для печатных схем. Эти материалы должны отвечать следующим требованиям: иметь высокие удельное сопротивление и электрическую прочность, большой диапазон рабочих температур, малые значения температурного коэффициента линейного расширения *TKl*, обладать достаточной механической прочностью, легко поддаваться технологической обработке. Фольгированные материалы используются для изготовления печатных плат радиоэлектронной аппаратуры широкого применения в диапазоне частот до 10^6 Гц.

5. *Пленочные электроизоляционные материалы*. Эти материалы представляют собой тонкие пленки, изготовленные различными способами в зависимости от исходного полимера. Для повышения механической прочности пленки применяют тот или иной прием ориентации молекул полимера, в результате чего они вытягиваются в определенном направлении и закрепляются в таком состоянии.

Благодаря высоким электрическим свойствам, малой толщине, достаточной механической прочности и во многих случаях высокой влагостойкости эти пленки представляют собой весьма ценный и прогрессивный диэлектрический материал для производства конденсаторов, кабельных изделий и т. д.

Свойства пленок обусловлены свойствами исходных полимеров, поэтому пленки можно разделить на два больших класса: неполярные и полярные.

- *Неполярные пленки*. К важнейшим неполярным пленкам относят: полистирольную, пленку из сополимеров стирола, политетрафторэтиленовую, полиэтиленовую, полипропиленовую.

Полистирольную пленку (ПС) получают методом выдавливания размягченного (при 140...160°С) полистирола с одновременным растяжением в продольном и поперечном направлениях. Недостатком этой пленки является сравнительно невысокая температура размягчения и низкая механическая прочность при надрыве.

Политетрафторэтиленовая пленка (ПТФЭ) является наиболее нагревостойкой из неполярных синтетических пленок. Она имеет толщину 20…150 мкм и используется в качестве изоляции в проводах, кабелях, микромашинах. Широкое применение ПТФЭ ограничивает ее высокая стоимость.

Полиэтиленовая пленка (ПЭ) является наиболее дешевой, она подвержена воздействию нефтяного масла и при достаточно высокой температуре растворяется в нем. Это обстоятельство ограничивает применение пленки в конденсаторах, кабелях и монтажных проводах.

Полипропиленовая пленка (ПП) по свойствам во многом схожа с полиэтиленовой, что объясняется сходством исходных полимеров. Полипропиленовую пленку используют в высоковольтных конденсаторах в качестве комбинированной бумажно-пленочной изоляции.

- *Полярные пленки*. Из полярных синтетических электроизоляционных пленок наиболее известны полиэтилентерефталатная пленка, пленки из эфиров целлюлозы, поликарбонатная, полиамидная, поливинилхлоридная. Их используют в конденсаторном и кабельном производстве, а также для изоляции обмоток низковольтных электрических приборов.

6. *Волокнистые электроизоляционные материалы*. *Волокнистые материалы* состоят из отдельных тонких, обычно гибких волокон, отличающихся большой величиной отношения длины к толщине. Их можно разделить на природные и синтетические материалы.

К *природным волокнам* относят материалы растительного происхождения (хлопок, бумагу), животного происхождения (шелк, шерсть) и минерального происхождения (асбест).

Группу *синтетических волокон* составляют полистирольные, полиэтилентерефталатные, полиамидные, полиэфирные, полиэтиленовые и другие волокна, изготавливаемые путем вытягивания соответствующих полимеров из растворов и расплавов. В эту группу следует также отнести ацетатный и медно-аммиачный шелка, стеклянные волокна.

Электроизоляционные бумаги. Из древесины, путем ее химической переработки, получают техническую целлюлозу, или клетчатку, которая является сырьем для изготовления электроизоляционных бумаг. По назначению их делят на *конденсаторную*, *кабельную*, *пропиточную* и др.

Малая толщина и возможность пропитки конденсаторной бумаги позволяют получить достаточно высокое значение удельной емкости бумажного конденсатора. В настоящее время в конденсаторах, применяемых

в электронной технике, бумага постепенно вытесняется синтетическими пленками.

4.2.1.2. Твердые неорганические диэлектрики

1. *Стекла*. *Стекла* – неорганические аморфные вещества, представляющие собой сложные системы различных оксидов

Параметры стекол определяются их химическим составом и режимом тепловой обработки и поэтому колеблются в весьма широких пределах. Так, при нормальной температуре, удельное сопротивление стекол $\rho = 10^6 \ldots 10^{15}$ Ом·м (у некоторых стекол может составлять 10^3 Ом·м), $\varepsilon = 4 \ldots 25$; $tg\delta = 0{,}0002 \ldots 0{,}01$; $E_{пр} = 23 \ldots 500$ МВ/м.

Предел прочности стекла при растяжении невелик (100…300 МПа) и увеличивается с повышением содержания в нем оксидов кремния SiO_2. Щелочные оксиды снижают прочность стекла. Стекло противостоит сжатию гораздо лучше, чем растяжению, предел прочности при сжатии составляет 6000…21000 МПа.

Из тепловых свойств стекол наибольшее значение имеют температура размягчения и температурный коэффициент линейного расширения (*ТКl*). Необходимо, чтобы значения *ТКl* стекла и соединяемых с ним материалов были приблизительно одинаковыми, иначе при смене температуры может произойти растрескивание стекла, нарушение герметичности в месте ввода металлических ножек в радиолампах.

Типы стекол. В зависимости от назначения различают следующие основные виды электротехнических стекол: электровакуумные, изоляторные, стеклоэмали и др.

Электровакуумные стекла применяют для изготовления баллонов электронных и газоразрядных ламп, оболочек рентгеновских ламп, горловин кинескопов и т. п. Кроме того, их широко применяют в полупроводниковой промышленности для получения металлостеклянных спаев в корпусах маломощных приборов.

Электровакуумные стекла делят на группы по признаку спаиваемости с определенным металлом или сплавом. Так, стекла молибденовой группы имеют температурный коэффициент расширения (*ТКl*), близкий к *ТКl* молибдена, и при пайке с ним образуют вакуумно-плотные спаи. Стекла вольфрамовой группы имеют *ТКl*, близкий к *ТКl* вольфрама, а стекла платиновой группы имеют *ТКl*, близкий к *ТКl* платины. Названия стекол «молибденовые», «вольфрамовые», «платиновые» определяются не составом стекла, а только тем, что значения *ТКl* этих стекол близки к *ТКl* молибдена, вольфрама, платины.

Изоляторные стекла используют в качестве герметизированных вводов в некоторых типах конденсаторов, терморезисторов, в полупроводниковом производстве для изготовления изоляторов в металлических корпусах таблеточного типа и изоляторов в мощных приборах. Эти стекла используют также для изоляции штырьковых вводов в корпусах кремниевых и германиевых транзисторов.

Боросиликатные и алюмосиликатные стекла применяют для изготовления изоляционных подложек, которые служат основой для расположения активных и пассивных элементов интегральных микросхем.

Стеклоэмалями называют стекловидные покрытия, наносимые на поверхность изделий с целью защиты от коррозии, электрической изоляции, а также для получения красивого внешнего вида. При эмалировке порошок стекла измельчают, наносят на поверхность изделия и обжигают. В результате плавления эмаль растекается по поверхности изделия, и после охлаждения остается на ней в виде тонкого (0,1…1,0 мм) сплошного стекловидного покрытия.

Стеклоэмали применяют в качестве электроизоляционного материала для трубчатых резисторов, в которых на наружную поверхность керамической трубки нанесена проволочная обмотка. Наносимый поверх нее слой эмали создает изоляцию между отдельными витками, между обмоткой и окружающей средой, а также защищает ее от воздействия влаги, от окисления и т. п.

Стекловолокно получают из расплава стекла, чаще из бесщелочного алюмоборосиликатного. Расплавленная стекломасса под действием собственного веса медленно вытекает из платиновых «лодочек» через фильеры диаметром около 1 мм. Выходящая из фильеры нить наматывается с большой скоростью (около 30 м/с) на барабан и, не успевая охладиться, вытягивается в тонкое волокно диаметром в несколько микрон. При вытягивании происходит ориентация молекул, способствующая повышению гибкости и механической прочности стекла.

Из стеклянных нитей изготавливают световоды, изоляцию монтажных, обмоточных и микропроводов, стеклянные ткани, используемые в производстве нагревостойких стеклолакотканей и стеклотекстолитов. Короткое стекловолокно применяют как наполнитель в пресс-порошках, идущих на изготовление пластмасс, претерпевающих тысячи отражений, ослабление света обусловлено лишь поглощением в материале и отражением от торцов. Поэтому основной путь снижения потерь в световодах – чистота исходных материалов и стерильность производства.

Оптическое стекло получается в результате варки смеси неорганических веществ. Стекло не является кристаллом. Отличительной особенностью оптического стекла является его однородность. Стекла и остальные оптические материалы обладают дисперсией: показатель преломления зависит от длины волны.

С увеличением длины волны λ и смещением ее от фиолетовой области спектра к красной области спектра, показатель преломления n уменьшается. Функция $n(\lambda)$ является основной характеристикой сорта стекла.

График функции $n(\lambda)$ (дисперсионная кривая) близок к гиперболе. Показатели преломления достаточно просто измеряются с точностью до 6 десятичных разрядов только для определенных длин волн (380…780 нм). Для любой другой длины волны λ видимой и примыкающих к ней УФ- и ИК-

областях спектра показатель преломления можно найти по дисперсионной формуле. Например, можно использовать степенной ряд

$$n^2 = A_0 + A_1\lambda^2 + A_2\lambda^{-2} + A_3\lambda^{-4} + A_4\lambda^{-6} + A_5\lambda^{-8}. \qquad (4.96)$$

Значение коэффициентов A_0, …, A_s для каждого сорта стекла можно найти в специальных справочниках. Точность расчета показателя преломления n по этой формуле достигает несколько единиц в шестом десятичном знаке.

Необходимо также кратко упомянуть наиболее важные характеристики стекла (показатель преломления и дисперсия). Показатель преломления n_e для длины волны λ_e = 546,1 нм (желто-зеленый свет) принят в качестве основного показателя преломления. Для характеристики дисперсии используют число Аббе

$$\nu_e = \frac{n_e - 1}{n_{F'} - n_{C'}}, \qquad (4.97)$$

где n_F – показатель преломления стекла для длины волны λ_F = 486,1 нм; n_C – показатель преломления стекла для длины волны λ_C = 656,3 нм. Для разных типов стекол n_F и n_C имеют различные значения.

Разность показателей преломления $n_{F'} - n_{C'}$ (можно найти в справочниках) называется средней дисперсией. Разность показателей преломления между другими линиями спектра называют частичными дисперсиями.

Для краткого описания оптических свойств указывают основной показатель преломления n_e и число Аббе ν_e. Большая величина числа Аббе соответствует низкой дисперсии.

Стекла делятся на 2 вида: *кроны* и *флинты*. Кронами называют все стекла с $\nu_e > 55$. Флинтами называют все стекла с $\nu_e < 50$. Между ними лежат несколько переходных групп.

В оптических стеклах могут встречаться отклонения от однородности. Свили являются лентовидными, узловидными или нитевидными включениями в стекло, имеющими показатель преломления, отличный от показателя преломления остальной массы стекла. Так как при варке стекла газ не полностью выделяется из вязкой массы, то в стекле могут присутствовать мелкие пузыри.

Прозрачность стеклянной пластины описывается спектральным коэффициентом пропускания (без учета потерь на отражение на поверхностях образца) $\tau_i(\lambda)$, который можно найти по формуле:

$$\tau(\lambda) = \frac{(\Phi_{e\lambda})_\tau}{\Phi_{e\lambda}}, \qquad (4.98)$$

где $\Phi_{e\lambda}$ – поток излучения, падающий на поверхность; $(\Phi_{e\lambda})_\tau$ – спектральный поток излучения после прохождения пластины.

Стекла для особенно высококачественных систем должны обладать незначительным эффектом двойного лучепреломления, который возникает из-за остаточных напряжений в стекле. При использовании стекол

необходимо также учитывать их химические свойства (устойчивость к климатическим и атмосферным воздействиям, кислотостойкость), твердость и коэффициент теплового расширения.

В видимой и ближней инфракрасной области спектра наибольшее распространение получили световодные волокна, изготовленные из оптического стекла. Для световедущих жил используют стекла типа тяжелых *флинтов* ТФ, баритовых флинтов БФ и сверхтяжелых кронов СТК, а для изоляционной оболочки применяют стекла типа *крона* К или типа легкого крона ЛК. (В менее ответственных случаях, например для целей освещения, вместо стеклянных световодных волокон применяют более дешевые пластмассовые волокна из полиметилметакрилата, полистирола и других пластмасс).

2. *Ситаллы*. *Ситаллы* – это стеклокристаллические материалы, получаемые с помощью специальной термообработки стекла, приводящей к его частичной кристаллизации. При изготовлении ситаллов в стекломассу кроме обычных оксидов вводят тонкодисперсные примеси, служащие для образования центров кристаллизации. Таким образом, ситаллы состоят из очень мелких кристаллических зерен, скрепленных стекловидной прослойкой. Содержание кристаллической фазы в ситаллах достигает 95%, размеры оптимально развитых кристаллов составляют 0,01…1 мкм. Ситаллы непрозрачны или частично пропускают свет и имеют значительно более низкий коэффициент теплового расширения, чем стекло или плавленый кварц.

Возможность изменения химического состава исходного стекла позволяет получать ситаллы с редким сочетанием физико-химических свойств: высокой механической прочностью, твердостью, термической и химической стойкостью, а по электроизоляционным свойствам ситаллы, как правило, превосходят стекла того же химического состава. Интервал рабочих температур может составлять от –50 до +700 °С. Доступность сырья и простая технология обеспечивают невысокую стоимость изделий.

Свойства ситаллов различных марок позволяют применять их при производстве многих приборов электронной техники, работающих в широком диапазоне частот и температур, например в СВЧ- приборах в качестве подложек для аттенюаторов, опор для крепления разрядников, для изготовления стержней электровакуумных ламп и деталей электронной техники и т.д.

3. *Оксидные электроизоляционные пленки*. В качестве неорганического электроизоляционного материала в электролитических конденсаторах и элементах интегральных схем, а также для изоляции алюминиевых проводов и лент, нашли широкое применение пленки оксидов алюминия ($A1_2O_3$), тантала (Ta_2O_5), титана (TiO_2), ниобия (Nb_2O_3) и др.

Тонкая неорганическая пленка может быть образована двумя основными способами:

1) в результате химической или электрохимической реакции поверхностного слоя проводникового материала; 2) путем осаждения диэлектрика на поверхность изолируемого материала.

Наиболее широко применяют оксидную изоляцию в электролитических конденсаторах. В качестве диэлектрика в этих конденсаторах используют следующие оксиды: $A1_2O_3$ (ε ≈10, $E_{пр}$ = 400…800 МВ/м); Ta_2O_5 (ε ≈ 25, $E_{пр}$ = 400…500 МВ/м); Nb_2O_5 (ε ≈ 40, $E_{пр}$ ≈ 400 МВ/м); TiO_2 (ε ≈ l80, $E_{пр}$ ≈ 400 МВ/м). Оксидные конденсаторы при малых размерах и массе имеют высокую емкость, благодаря малой толщине (до 1 мкм) диэлектрика и значительной диэлектрической проницаемости. Оксиды кремния, тантала, ниобия применяют также для изготовления тонкопленочных конденсаторов, диэлектриком в которых является пленка из этих оксидов, наносимая методами вакуумной технологии на проводящую подложку.

Алюминиевые провода и ленты, покрытые пористой оксидной пленкой, полученной анодированием, применяют как заменители обмоточных медных проводов с обычными видами изоляции. Они имеют высокую нагревостойкость (рабочая температура достигает 600 °С), малую толщину оксидной изоляции, высокие теплопроводность и химическую стойкость; процесс изолирования дешев и прост. Недостатком такой оксидной изоляции является ее малая гибкость и заметная гигроскопичность.

В полупроводниковой технологии для защиты *p-n*-переходов от воздействия окружающей среды (прежде всего влаги и кислорода) кроме лаков, эмалей, компаундов, нитридных пленок применяют пленки на основе оксидов алюминия, титана и др. Пленки для защиты *p-n*-переходов полупроводниковых кристаллов кроме высоких электроизоляционных свойств должны обладать малой химической активностью; иметь хорошее сцепление с полированной поверхностью полупроводникового материала; не вызывать значительных механических напряжений при изменениях температуры; сохранять защитные свойства во всем температурном диапазоне работы полупроводниковых приборов.

4. ***Керамические электроизоляционные материалы***. *Керамика* – твердый плотный материал, получаемый спеканием неорганических солей с минералами и оксидами металлов.

Керамические материалы представляют собой многофазную систему и включают в себя кристаллическую, *стекловидную* и *газовую* фазы. Основной фазой является кристаллическая фаза. Керамические материалы нагревостойки, обладают высокими электрическими свойствами, механически прочны, стабильны и надежны в эксплуатации. Преимуществом керамики перед другими материалами является возможность получения заранее заданных электрофизических параметров материала в широком диапазоне изменением исходного состава массы и технологии производства изделий.

Радиотехнические керамические материалы, в зависимости от назначения, изготовляют следующих типов: А – высокочастотные для конденсаторов; Б – низкочастотные для конденсаторов; В – высокочастотные

для установочных изделий. Каждый тип керамического материала подразделяют на классы, общее число которых десять.

Высокочастотная конденсаторная керамика типа А отличается высоким содержанием кристаллической фазы и небольшим содержанием бесщелочной аморфной фазы. Кристаллическую фазу формируют в соответствии с желаемым температурным коэффициентом диэлектрической проницаемости, стремясь получить наибольшее значение диэлектрической проницаемости и низкое значение тангенса диэлектрических потерь. Значение ε для керамики типа А изменяется от 14 до 250; $tg\delta \le 6\cdot10^{-4}$ (при $t = 20$ °С, $f = 10^6$ Гц); $E_{пр} = 635$ МВ/м.

К керамикам типа А относятся стронциевая керамика, перовскитовая керамика, станнатная керамика и лантановая керамика.

Низкочастотная конденсаторная керамика типа Б отличается высокой диэлектрической проницаемостью ε (900…9800), относительно большими диэлектрическими потерями $tg\delta = 2\cdot10^{-3}\ldots5\cdot10^{-2}$, небольшим значением электрической прочности ($E_{пр} = 4\ldots15$ МВ/м). Эта керамика изготавливается на основе твердых растворов титаната и цирконата бария, ниобатов свинца, стронция и ряда других и используется для изготовления разделительных и блокировочных конденсаторов.

Высокочастотная установочная керамика типа В используется в качестве материала для изготовления различных установочных деталей радиоэлектронной аппаратуры, работающих в поле высокой частоты и несущих большую механическую нагрузку. Эта керамика имеет низкое значение тангенса диэлектрических потерь в области высоких частот, обладает слабой зависимостью тангенса диэлектрических потерь от температуры и частоты, имеет высокую механическую прочность. К ней относится форстеритовая керамика, глиноземная керамика и стеатит.

5. *Слюда и материалы на ее основе*. Слюда представляет собой природный минерал с кристаллической структурой, который легко расщепляется по плоскостям спайности на пластинки толщиной до 5 мкм. Известно более 30 разновидностей слюды, все они являются водными алюмосиликатами щелочноземельных металлов. Несмотря на большое количество разновидностей слюды, для электрической изоляции применяют только две: *мусковит* ($K_2O\cdot3Al_2O_3\cdot6SiO_2\cdot H_2O$) и *флогопит* ($K_2O\cdot6MgO\cdot Al_2O_3\cdot6SiO_2\cdot2H_2O$). Стоящая в формулах вода является кристаллизационной, входящей в состав кристаллической решетки.

При нагревании вода начинает выделяться, что вызывает вспучивание слюды, увеличение ее толщины за счет расслаивания по плоскостям спайности. Кроме того, слюда теряет прозрачность, резко снижаются ее электрические и механические свойства. Температура начала обезвоживания и резкого ухудшения свойств у мусковитов 500… 600 °С, у флогопитов 800… 900 °С.

Слюды являются широко распространенными минералами. Первоначальная обработка слюды состоит в ее очистке от посторонних

примесей. Затем слюду раскалывают на пластинки толщиной 0,1…0,6 мм и обрезают дефектные участки. В дальнейшем пластины расщепляют на листки толщиной 5…50 мкм. Такую слюду называют щипаной и применяют для клееной слюдяной изоляции, а из лучших сортов мусковита готовят *конденсаторную слюду*, представляющую собой листочки прямоугольной формы толщиной 20…55 мкм, длиной 7…60 мм, шириной 4…50 мм. Конденсаторную слюду применяют для изготовления слюдяных конденсаторов постоянной емкости.

Слюдяная бумага производится по технологии, близкой к технологии обычной бумаги. Слюду размельчают, готовят пульпу, на бумагоделательных машинах раскатывают листы слюдяной бумаги.

Слюду также применяют и в виде штампованных фасонных изделий для крепления и электрической изоляции внутренней арматуры в электронных лампах; в качестве тепловой защиты цоколей мощных электрических ламп накаливания; для изоляции центральных электродов и экранов свечей зажигания авиационных двигателей и т. д.

В настоящее время в технике получила распространение синтетическая слюда – *фторфлогопит*, отличающаяся от природной тем, что в ее кристаллической решетке ионы ОН замещены ионами фтора. Фторфлогопит химически устойчивее обычного флогопита, обладает более высокими электрическими параметрами, более нагревостоек, практически не вспучивается при нагревании. Он выдерживает смены температур от –70 °С до +1000 °С.

К синтетическим слюдяным материалам также относят *микалекс,* представляющий собой твердый материал, получаемый путем горячего прессования смеси порошкообразной природной слюды и легкоплавкого борно-свинцового и борно-боритового стекла. Микалекс применяют для изготовления изоляционных деталей колебательных контуров мощных радиопередатчиков, каркасов индуктивных катушек, плат переключателей, различных деталей вакуумных приборов и др.

4.2.2. Активные диэлектрики

Активными диэлектриками называют такие диэлектрики, свойствами которых можно управлять с помощью внешних факторов и использовать эти факторы для создания функциональных элементов электроники. Чаще всего активные диэлектрики классифицируют согласно физическим эффектам, используемым для управления свойствами материалов. То есть можно выделить сегнето-, пьезо-, пироэлектрики, электреты, и др. Но, хотя такая классификация является логической, очень нелегко отделить одну группу материалов от других. Это связано с тем, что большое количество активных диэлектриков имеют высокую чувствительность по отношению к нескольким видам энергетических воздействий. Наибольшую универсальность свойств имеют сегнетоэлектрики. Они могут иметь свойства пьезо-, пироэлектриков, электретов, нелинейно-оптических материалов.

4.2.2.1 Сегнетоэлектрики, пьезоэлектрики, электреты

Основные особенности сегнетоэлектриков и механизм спонтанной поляризации были рассмотрены ранее.

Классификация сегнетоэлектриков. По типу химической связи и физическим свойствам сегнетоэлектрики подразделяют на две группы:

1. *Ионные сегнетоэлектрики.* Эти материалы имеют кристаллическую решетку, основой которой является кислородный октаэдр. К ним относят метатитанат бария ($BaTiO_3$), ниобат калия ($KNbO_3$), ниобат лития ($LiNbO_3$), танталат лития ($LiTaO_3$), барий-натриевый ниобат ($BaNaNb_5O_{15}$). Эти сегнетоэлектрики имеют большую механическую прочность, не растворяются в воде, легко изготавливаются с использованием керамической технологии, имеют довольно высокую температуру Кюри, (например, $LiNbO_3$ – 1483 K, $BaTiO_3$ – 393 K), и большое значение относительной диэлектрической проницаемости.

2. *Дипольные сегнетоэлектрики.* Эти материалы имеют готовые полярные группы атомов, которые занимают различные положения равновесия. К ним относят сегнетову соль ($NaKC_4H_4O_6 \cdot 4H_2O$), триглинсульфат $(NH_2CH_2COOH)_3 \cdot H_2SO_4$, дигидрофосфат калия ($KH_2PO_4$), нитрит натрия ($NaNO_2$) и прочие. Эти диэлектрики имеют низкие механические свойства, легко растворяются в воде, имеют низкую температуру Кюри (часто ниже комнатной). Их очень легко выращивать из водных растворов в виде монокристаллов.

Подавляющее большинство сегнетоэлектриков первой группы имеют более высокую температуру Кюри и большее значение спонтанной поляризованности по сравнению с сегнетоэлектриками второй группы.

Пьезоэлектрики. Известно более тысячи веществ, обладающих пьезоэлектрическими свойствами, в том числе – все сегнетоэлектрики. Однако практическое применение в пьезотехнике находит ограниченный круг материалов. Среди них одно из важных мест занимает монокристаллический кварц. Это одна из модификаций двуокиси кремния.

Ввиду ограниченных запасов природного кварца основные потребности пьезотехники удовлетворяются искусственно выращиваемыми кристаллами. Из других веществ следует отметить кристаллы турмалина (сложные алюмоборосиликаты), используемые для резонаторов на частотах больших, чем кварцевые резонаторы.

Помимо кварца широкое применение в качестве пьезоэлектрического материала находит сегнетоэлектрическая керамика. В обычном состоянии сегнетокерамика является изотропной средой вследствие хаотического расположения кристаллических зерен и деления их на домены с различным направлением спонтанной поляризованности. Однако если подвергнуть сегнетокерамику воздействию сильного электрического поля, то поляризованность доменов получит преимущественную ориентацию в одном направлении. После снятия поля сохраняется устойчивая остаточная поляризованность. По своим свойствам поляризованный

сегнетокерамический образец близок к пьезоэлектрическому кристаллу, т. е. обладает высокой пьезоактивностью.

Поляризованную пьезокерамику, предназначенную для использования в пьезоэлектрических целях, называют пьезокерамикой. Пьезокерамика имеет перед монокристаллами то преимущество, что из нее можно изготовить активный элемент практически любого размера и формы. Основным материалом для изготовления пьезокерамических элементов являются твердые растворы $PbZrO_3$-$PbTiO_3$ (цирконат-титанат свинца). Значительную конкуренцию кварцу, превосходя его по добротности на больших частотах, представляют ниобат и танталат лития ($LiNbO_3$ и $LiTaO_3$). Но для того чтобы в этих керамиках были пьезоэлектрические эффекты, их приводят в монодоменное состояние отжигом в сильном электрическом поле.

Электреты. В качестве электретных материалов могут быть использованы как органические, так и неорганические диэлектрики.

Электреты из *органических материалов* можно условно разбить на две группы.

К первой группе относят электреты, полученные из природных смол (канифоль, шеллак, янтарь) и их смесей, а также электреты из сахара, асфальта, эбонита, полиметилметакрилата и др. Электреты этой группы имеют время жизни не более одного года и практического применения не имеют.

Ко второй группе относят электреты из тонких органических полимерных пленок. Для их получения используют слабополярные и нейтральные полимеры с высокими диэлектрическими свойствами, например, полиэтилентерефталат, поликарбонаты, полипропилен, политетрафторэтилен.

К неорганическим электретам в первую очередь относят электреты из керамики, которые по ряду своих свойств ($\rho = 10^{10}\ldots10^{12}$ Ом·м$^{-1}$, $\varepsilon \geq 20$) представляют интерес для промышленного использования. Электреты из керамики изготавливают из титаносодержащего керамического материала на основе TiO_2, $CaTiO_3$, SrO-Bi_2O_3-TiO_2, $BaTiO_3$, $MgTiO_3$, $ZnTiO_3$, $SrTiO_3$ и некоторых других. Наибольшим и наиболее устойчивым зарядом обладают электреты из $CaTiO_3$. Значение поверхностной плотности заряда керамических электретов, как правило, не превышает 10^{-4} Кл/м^2.

В группу неорганических электретов входят также электреты, изготавливаемые из ситаллов и различных стекол, например кварцевого, малощелочных боросиликатных и др.

В заключение этого параграфа можно сделать следующие обобщения. Некоторые пьезоэлектрические кристаллы могут поляризоваться без какой-либо механической нагрузки, а только в результате изменения температуры. Поляризация обнаруживается по скоплению зарядов на электродах. Это явление и называется пироэлектричеством. Механизм пироэффекта подобен механизму пьезоэлектрического эффекта. Оба эффекта основаны на деформации ионной кристаллической решетки, причем безразлично,

возникает ли эта деформация под действием внешней силы или в результате теплового расширения. Существенное различие заключается лишь в том, что пироэлектрические кристаллы должны иметь постоянную поляризацию, тогда как для пьезоэлектрического эффекта это условие не является необходимым. Отсюда ясно, что не все пьезоэлектрические кристаллы являются и пироэлектрическими; другими словами, пироэлектрические вещества представляют собой подгруппу пьезоэлектрических веществ. Если отнести сюда сегнетоэлектрические вещества, которые также характеризуются наличием постоянной поляризации, то мы получим схему классификации диэлектрических материалов, представленную на рис. 4.25.

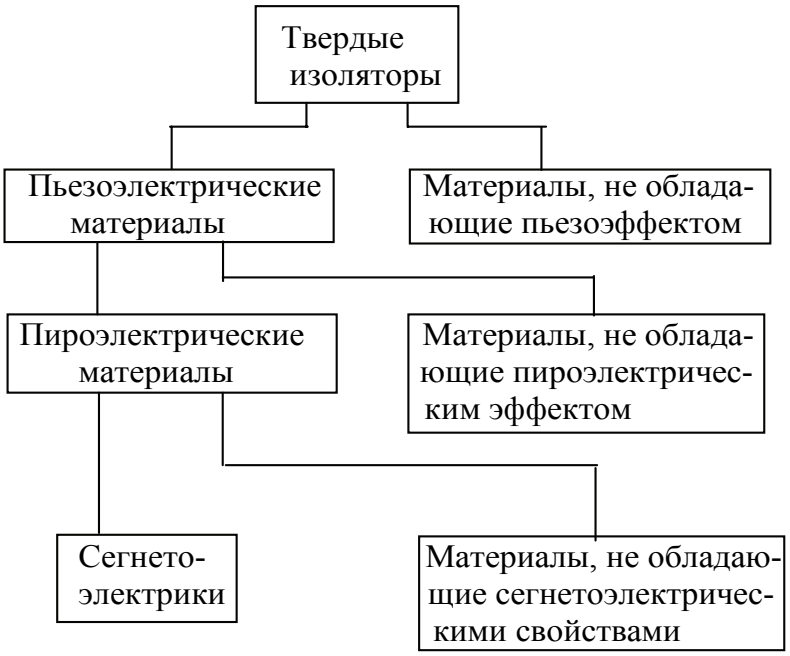


Рис. 4.25 – Схема классификации диэлектриков

4.3. ПРИМЕНЕНИЕ ДИЭЛЕКТРИКОВ

В данном разделе рассмотрим применение пассивных и активных диэлектриков в электронике и оптической связи. В последние годы во всем мире происходит мощное развитие информационных технологий. Поэтому вопрос качественной передачи информации стоит очень остро. Особое значение имеет скорость и надежность передачи информации.

Нас будет интересовать, как, изменяя свойства диэлектрических материалов, можно сконструировать различные приборы и улучшить их характеристики. В первую очередь рассмотрим ряд приборов, используемых в оптической связи: световоды, модуляторы и твердотельные лазеры, используемые как источники когерентного излучения.

Вначале дадим краткую характеристику применения активных диэлектриков, таких как пьезоэлектрики, пироэлектрики, сегнетоэлектрики и электреты.

Применение пьезоэлектриков. Оптические свойства пьезоэлектриков широко используются для создания оптических модуляторов и квантовых генераторов когерентного света (лазеров). Более подробно об этом написано ниже в параграфах 4.3.2 и 4.3.3.

Наглядным примером применения пьезоэлектриков является кварцевый резонатор. Плоскопараллельная кварцевая пластинка с электродами и держателем представляет собой пьезоэлектрический резонатор, т.е. является колебательным контуром с определенной резонансной частотой колебаний. Частота этих колебаний зависит от толщины пластины и направления кристаллографического среза. Преимуществами кварцевых резонаторов являются малый тангенс угла потерь $tg\delta$ и высокая механическая добротность (т.е. очень малые механические потери). В лучших кристаллах кварца механическая добротность может составлять $10^6 \ldots 10^7$. Это обеспечивает высокую частотную избирательность кварцевых резонаторов. Если в таком резонаторе возбудить колебания на резонансной частоте, то их затухание будет происходить в течение длительного времени. Кварцевый элемент, поставленный во входную цепь электрического генератора, навязывает ему собственную резонансную частоту.

Пьезоэлектрики используют для построения мощных ультразвуковых излучателей для дефектоскопии, механической обработки материалов, гидроакустики. Двойное превращение энергии используют для фильтров, линий задержки и пьезотрансформаторов. Последние позволяют получать большие напряжения и используются для питания электронно-лучевых трубок, счетчиков Гейгера и др.

Для изготовления малогабаритных микрофонов, телефонов, высокочастотных громкоговорителей, слуховых аппаратов, детонаторов, различных устройств для поджига газа, датчиков давления, деформации, вибрации, ускорения широко используется пьезокерамика.

Применение пироэлектриков. Пироэлектрики используют как преобразователи тепловой энергии в электрическую. Значительный пироэффект в сегнетоэлектриках используется для создания тепловых датчиков и приемников лучистой энергии, предназначенных, в частности, для регистрации инфракрасного и СВЧ-излучения. Принцип действия пироэлектрических фотоприемников очень прост: лучистая энергия, попадая на зачерненную (поглощающую) поверхность сегнетоэлектрического кристалла, нагревает его. В результате нагревания изменяется спонтанная поляризованность кристалла и возникает ток (см. формулу (4.21)), который регистрируется электронной схемой. Основным свойством таких тепловых фотоприемников является неселективность спектральной чувствительности. Существенное преимущество их состоит в том, что они не требуют охлаждения при детектировании излучения даже в далекой ИК-области спектра. Наряду с этим они обладают достаточно высоким быстродействием (способны работать в частотном диапазоне до 10 МГц), однако по чувствительности уступают полупроводниковым фотоприемникам.

Применение сегнетоэлектриков. В технике наиболее часто используют ионные сегнетоэлектрики, что связано с их механическими и электрическими свойствами. Основными направлениями использования сегнетоэлектриков являются следующие: низкочастотные конденсаторы с большой удельной емкостью, вариконды (нелинейные конденсаторы), ячейки памяти для ЭВМ.

Благодаря диэлектрическому гистерезису сегнетоэлектрики применяют в запоминающих устройствах ЭВМ. Для этих целей необходим материал с возможно более прямоугольной петлей гистерезиса (ППГ), что характерно для монокристаллов (например, триглицинсульфата). Однако более массовое применение для целей запоминания и хранения больших массивов информации нашли именно ферромагнитные материалы (см. гл. V).

Следует отметить, что кроме ионных сегнетоэлектриков очень широкое распространение получила сегнетокерамика. Этот ряд материалов, изготовленных на основе $BaTiO_3$ с разными примесями по керамической технологии, позволяет изменять в требуемых пределах ε, $tg\delta$, α_ε и другие характеристики.

Применение электретов. Электреты могут быть использованы для изготовления микрофонов. В электретном микрофоне электретная пленка из полмиера одновременно выполняет функции источника постоянного электрического напряжения и мембраны. Это является важным преимуществом, так как исчезает необходимость в источнике постоянного электрического напряжения. В настоящее время электретные микрофоны все более широко используются в звуковой акустике и успешно конкурируют с аналогами других типов.

Электреты очень удобны для использования в клавишных устройствах компьютеров, телефонов и т. д., где используются физические свойства механоэлектретов. Клавиша является электромеханическим преобразователем – нажатие на клавишу на выходе преобразуется в электрический сигнал. Конструкция клавиши включает в себя два электрода и электрет, расположенный между электродами, один из которых может перемещаться относительно электрета.

Кроме вышеперечисленного, электреты разных классов (по методу получения) могут быть использованы для измерения механических вибраций, дозиметров радиации, измерителей атмосферного давления и влажности, электрометров, в электростатических вольтметрах, в электрофотографии и многих других устройствах.

4.3.1. Световоды

Световодом называется оптический объект, при прохождении через который свет многократно отражается от боковых поверхностей, канализируется в узком пучке. В простейшем случае световод представляет собой цилиндрический стержень из стекла или пластмассы с плоскими торцами. Благодаря явлению *полного внутреннего отражения* луч, при распространении в таком световоде, многократно отражается от боковой поверхности – границы стержня с воздухом. В таких световодах потери

световой энергии на поверхности раздела жилы и оболочки малы даже для лучей, претерпевающих тысячи отражений. Ослабление света обусловлено лишь поглощением в материале и отражением от торцов, поэтому основной путь снижения потерь в световодах – чистота исходных материалов и стерильность производства.

Важным преимуществом световода является то, что он может быть изогнут. Однако одиночный световод не может передавать изображений. Это связано с тем, что если на входной поверхности имеется неравномерное распределение яркости, то выходная поверхность будет освещена более или менее равномерно.

Очень тонкие световоды из стекла или пластмассы обладают двумя преимуществами: они представляют собой гибкие волокна и их можно использовать для переноса изображений (каждое волокно переносит свой элемент изображения предмета). Волокна со светоизолирующими оболочками диаметром 2…15 мкм параллельно укладывают в пучки и спекают, получая *многожильные световоды* диаметром 0,3…3 мм. Общее число световедущих жил может доходить до 100 000. При соединении таких световодных волокон в жгут необходимо принять меры, чтобы на боковых поверхностях не нарушалось условие полного внутреннего отражения из-за соприкосновения волокон. Для достижения этой цели каждое отдельное волокно следует светоизолировать от других. Для обеспечения светоизоляции каждое световолокно делают составным (сердцевина+ оболочка). Во избежание просачивания света из одной световедущей жилы в соседние толщина светоизолирующей прослойки между ними должна составлять не менее 1 мкм. Световоды выпускают как прямыми, так и изогнутыми или закрученными. Жгуты из тонких волокон, соединенных только по концам, обладают большой гибкостью. Их изгиб не нарушает передачи по ним изображения. Световоды широко применяют для передачи различной информации в вычислительной технике, телевидении, фототелеграфии и т. д.

Чтобы понять, как свойства материала влияют на характеристики работы волоконно-оптических элементов, рассмотрим оптические характеристики одиночного световолокна.

На рис. 4.26 показан ход луча в прямом цилиндрическом волокне. Для обеспечения светоизоляции сердцевина из стекла с показателем преломления n_1 окружена стеклянной оболочкой с более низким показателем преломления n_2. Условие полного внутреннего отражения на границе сердцевина–оболочка имеет вид $n_1 \geq n_2$. Тогда половинный *апертурный угол* $u = \varepsilon_{0max}$ (максимальный угол падения на торцевую поверхность) находится

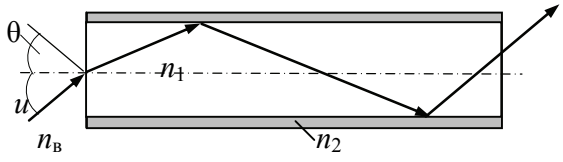


Рис. 4.26 – Числовая апертура прямого цилиндрического волокна

из уравнения:

$$\sin u = \sqrt{n_1^2 - n_2^2}\ . \qquad (4.99)$$

(Числовая *апертура* равна: $A = n\sin u$, где n – показатель преломления в пространстве предметов (чаще всего $n = 1$); u – половинный угол при вершине пучка лучей, исходящих из предметной точки.)

Числовая апертура прямого цилиндрического волокна (рис. 4.26) равна $A = \sin u$ ($n = 1$, так как торец волокна граничит с воздухом). Для прямого цилиндрического волокна входная апертура равна выходной.

В случае изгиба цилиндрического волокна (рис. 4.27) максимально возможная входная апертура рассчитывается по формуле:

$$\sin u = \sqrt{n_1^2 - n_2^2\left(1 + \frac{D}{2r}\right)^2}\ , \quad (4.100)$$

где D – диаметр сердцевины волокна, r – радиус изгиба. Апертура изогнутого волокна меньше, чем апертура прямого волокна. При $r = \infty$ формула (4.100) эквивалентна формуле (4.99).

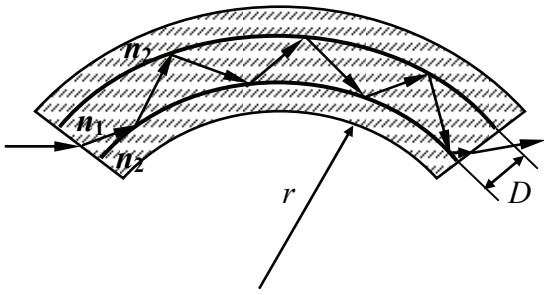


Рис. 4.27 – Схема прохождения лучей в световоде в случае изгиба

Примеры.

1. Чаще всего для волокон используют комбинацию стекол со следующими показателями преломления: $n_1 = 1{,}62$; $n_2 = 1{,}52$. Нужно рассчитать предельный угол падения на боковую поверхность ε_0, числовую апертуру волокна A, максимальный апертурный угол $2u$, при котором выполняется условие полного внутреннего отражения для всех лучей пучка.

Решение. $A = \sin u = \sqrt{1{,}62^2 - 1{,}52^2} = 0{,}56$; $2u$=75,73;

$$\varepsilon_0' = 90^\circ - \arcsin\frac{0{,}56}{1{,}62} = 90^\circ - 22{,}5^\circ = 67{,}5^\circ .$$

2. Очень часто используемые волокна с D = 70 мкм; $n_1 = 1{,}62$; $n_2 = 1{,}52$ должны изгибаться так сильно, как это допускает прочность (r = 12 мм). Какую апертуру тогда можно использовать?

Решение. $A = \sin u = \sqrt{1{,}62^2 - 1{,}52^2\left(1 + \frac{70\cdot 10^{-6}}{24\cdot 10^{-3}}\right)^2} = 0{,}548$. Сравнение с ранее полученным результатом для прямого волокна показывает, что даже при сильном изгибе апертура волокна уменьшается незначительно.

Для обеспечения малых потерь при полном внутреннем отражении оболочка волокна не должна быть очень тонкой. Толщины оболочки составляет примерно от 0,5 до 2 мкм. Полное внутренне отражение нельзя заменить нанесением зеркального слоя на волокно: так как число отражений

на 1 м длины составляет от 10^3 до 10^4, то почти весь световой поток будет поглощаться зеркальным слоем.

Элементы с волокнами из пластмассы дешевле, так как вместо полировки торцов достаточно гладко отрезать волокно. Поломка волокна практически исключена. Пластиковые волокна имеют несколько бóльшие диаметры, чем волокна из стекла, например 250 мкм. В результате старения коэффициент пропускания пластиковых волокон уменьшается. Могут встречаться минимумы коэффициента пропускания в районе 900 и 1000 нм, вследствие чего излучение GaAs-диода ослабляется.

Область пропускания обычных волокон из стекла соответствует диапазону от 400 до 2000 нм. За пределами этой области используют инфракрасные волокна (из специальных стекол) и ультрафиолетовые волокна (сердцевина из кварца, а оболочка из пластмассы).

Принцип действия градиентных волокон несколько иной, в них показатель преломления уменьшается от центра к краю поперечного сечения сердечника пропорционально квадрату радиуса. Градиентные волокна изготавливаются путем ионного обмена из специального стекла диаметром от 0,5 до 2 мм. Траектории лучей в градиентном волокне имеют синусоидальную форму. Свет проходит по такому волокну, периодически фокусируясь в точку на оптической оси, как через ряд миниатюрных линз («самофокусирующееся волокно»). Следовательно, возможна передача изображений одиночным волокном. Оптические волокна для передачи информации на большие расстояния должны иметь возможно меньшие потери. Их *спектральный коэффициент поглощения* или *удельный коэффициент потерь* рассчитывается по формуле:

$$\beta_0(\lambda) = 10 \cdot L^{-1} \lg(\tau_i(\lambda)) \text{ [дБ/км]}, \qquad (4.101)$$

где L – длина оптоволоконной линии; $\tau_i(\lambda)$ – чистый спектральный коэффициент пропускания. Например, если $\tau_i = 0{,}5$ ($\lambda = 800$ нм), то удельный коэффициент потерь $\beta_0 = 3$ дБ/км.

Полное затухание в оптоволокне определяется по формуле:

$$\beta = \beta_0(\lambda) \cdot L. \qquad (4.102)$$

Большое значение имеет величина пропускной способности по частоте волоконно-оптической линии. При передаче цифровой информации, на вход подается последовательность световых импульсов с малыми временными интервалами между ними. На другом конце линии приемник должен воспринимать импульсы раздельно. Многомодовые волокна со ступенчатым профилем для этой цели подходят плохо по следующей причине. Распространение излучения в волокне характеризуется модой, которая определяется типом распространения электрического поля. В лучевом представлении моды соответствуют различным зигзагообразным траекториям. Так как эти траектории имеют разную длину, то это приводит к уширению (расплыванию) длительности импульса. Для рассмотренного выше примера 1 при $L = 1$ км значение $\Delta t = 0{,}33$ мкс, а это очень большая величина для передачи информации.

В многомодовом градиентном оптическом волокне отдельные моды также имеют траектории различной геометрической длины, но их оптические пути и время прохождения импульса почти равны. Это происходит из-за того, что более длинные траектории в краевой области волокна компенсируются за счет уменьшения показателя преломления к периферии. При оптимальном профиле показателя преломления уширение реального импульса света никогда не станет равным нулю из-за дисперсии материала. В пределах малой ширины полосы $\Delta\lambda$ лазерной линии (или более широкого спектрального диапазона излучения полупроводникового световода) дисперсия $n(\lambda)$ приводит к различному времени прохождения различных спектральных составляющих излучения. Общая величина Δt лежит в интервале от нано- до пикосекунд.

В сердечниках волокон со ступенчатым профилем показателя преломления, чей диаметр соизмерим с длиной волны, может распространяться только одна мода. Эти волокна обладают особенно благоприятными свойствами и применяются для передачи информации.

Количество мод, которое может распространяться в волокне, определяет параметр волокна V

$$V = 2\pi \frac{r_k}{\lambda} \sqrt{n_1^2 - n_2^2}\,, \qquad (4.103)$$

где r_k – радиус сердцевины.

Очень короткий импульс спектральной ширины $\Delta\lambda = \lambda_2 - \lambda_1$, введенный в световод в одномодовом режиме, когда фронт волны перпендикулярен оси световода, выходит из него несколько размытым во времени. Это происходит из-за оптической дисперсии, когда сигнал с наибольшей длиной волны придет к выходу быстрее, чем сигнал с меньшей длиной волны (нормальная дисперсия). Такой эффект накладывает фундаментальное ограничение на верхнюю границу частоты следования импульсов. Частота следования импульсов не может быть больше граничной частоты $f_{\max}$, при которой в процессе распространения импульсов по световоду, длиноволновой край λ_2 одного импульса догонит коротковолновой край λ_1 предыдущего импульса, иначе говоря наименьший период следования импульсов $T_{\min}$ должен быть не больше разности времен распространения коротковолнового t_1 и длинноволнового t_2 импульсов:

$$T_{\min} = t_1 - t_2 = \frac{n(\lambda_1)\cdot l}{c} - \frac{n(\lambda_2)\cdot l}{c} = \frac{\left[n(\lambda_1) - n(\lambda_2)\right]\cdot l}{c} = \frac{\dfrac{dn}{d\lambda}\cdot \Delta\lambda \cdot l}{c}, \qquad (4.104)$$

где $n(\lambda_1)$ и $n(\lambda_2)$ – показатели преломления для коротковолнового и длинноволнового краев спектрального импульса; l – длина световода; c – скорость света в вакууме; λ_1 и λ_2 находят по спектральной полуширине $\Delta\lambda$ – ширине на половине интенсивности.

Максимальную частоту следования импульсов в световоде можно найти по формуле:

$$f_{\max} = \frac{c}{\frac{dn}{d\lambda} \cdot \Delta\lambda \cdot l}. \qquad (4.105)$$

Это частота абсолютного среза сигнала для очень коротких импульсов. В реальном случае требуется значительно снизить $f_{\max}$,

Все световоды имеют так называемые окна прозрачности, т.е. минимумы поглощения для определенных длин волн λ, распространяющихся в оптическом волокне.

Первым рабочим диапазоном волоконно-оптическиих линий связи (ВОЛС) на многомодовом волокне, был диапазон волн в окрестности 850 нм. Первые полупроводниковые излучатели света (лазеры и светодиоды) были разработаны для длины волны 850 нм. В *первом окне прозрачности* 760…860 нм величина коэффициента затухания современных волокон составляет 2…3 дБ/км и в некоторых случаях это окно продолжают использовать для передачи оптических сигналов на небольшие расстояния, например, в локальных вычислительных сетях. С совершенствованием волоконно-оптических технологий были освоены другие диапазоны длин волн, отвечающие *второму* (1260…1360 нм, затухание 0,5 дБ/км) и *третьему* (1530…1565 нм, затухание 0,2 дБ/км) окнам прозрачности. Во втором окне прозрачности работают многомодовые и некоторые одномодовые оптические волокна.

Первые два окна прозрачности расположены между пиками поглощения света в кварцевом стекле гидроксильными группами OH^-, а третье – расположено справа от пика этих потерь на резонансной длине волны 1383 нм. Это окно отвечает локальному уровню потерь, полученному спектральными кривыми потерь при рассеяние Рэлея ($\sim\lambda^{-4}$). Собственные потери в оптоволокне – потери через рэлеевское рассеяние на микронеоднородностях и поглощение на электронных или атомных резонансах в материале, из которого изготовлено волокно, составляют основную часть потерь в волокне.

Уменьшить рэлеевские потери в оптоволокне и получить меньшее затухание можно с помощью использования лазеров (светодиодов) с бóльшей длиной волны, так как короткие волны сильнее подвержены рэлеевскому рассеянию. Можно избавиться от рэлеевского рассеяния, перейдя к кристаллическим материалам. Известно, что кристаллы фтороида циркония (ZrF), селенида мышьяка (AsSe) и иодида калия (KI) обладают очень малым затуханием в области 2…5 мкм.

Использование лазеров для передачи сигналов по волокну позволяет сильно уменьшить размытие импульсов из-за дисперсии, так как спектральная ширина лазерного излучения значительно меньше, чем у источника некогорентного светодиода.

В последние годы изготовлены и освоены излучатели и фотоприемники, которые позволяют работать в длинноволновом диапазоне – четвертое окно прозрачности (1565…1625 нм). С подавлением пика потерь из-за поглощения

фотонов на гидроксилах OH^- на длине волны 1383 нм, появилась возможность использования пятого окна прозрачности (1460…1530 нм, коротковолновой диапазон), а также расширенного диапазона (1360…1460 нм). Этот тип потерь существенно зависит от технологии изготовления оптических волокон и имеет тенденцию к снижению. Уровень потерь для современных оптических волокон составляет значение порядка 0,4 дБ/км, а для некоторых типов волокон – до 0,31 дБ/км. Диапазон применения таких волокон от 1280 до 1675 нм.

К преимуществам волоконно-оптических линий относится их нечувствительность к электромагнитным полям, возможность проведения измерений во взрывоопасных помещениях и вблизи высоковольтных электрических установок (волокно является хорошим изолятором), незначительные габариты и сравнительно низкая цена.

4.3.2. Модуляторы

Возникновение двойного лучепреломления вещества во внешнем электрическом поле называется электрооптическим эффектом Поккельса. Наложение внешнего электрического поля, малого по сравнению с внутренним полем кристалла, приводит к перераспределению связанных зарядов и небольшой деформации ионной решетки, что сопровождается изменением показателя преломления. В этом эффекте изменение показателей преломления пропорционально первой степени внешнего поля $\vec{E}_0$, в отличие от эффекта Керра, где они пропорциональны квадрату поля.

Эффект Поккельса может наблюдаться только в кристаллах, не обладающих центром симметрии, называемых пъезоэлектриками. Дело в том, что этот эффект линеен относительно внешнего поля $\vec{E}_0$. Поэтому, при изменении направления поля $\vec{E}_0$ на противоположное, должен меняться на противоположный и знак изменения показателя преломления Δn. Но в кристаллах с центром симметрии это невозможно, так как оба взаимно противоположных направления внешнего поля физически эквивалентны. При использовании пьезоэлектриков, обладающих естественным двойным лучепреломлением, необходимо, чтобы кристалл до наложения внешнего электрического поля не давал двойного лучепреломления. Этого можно достичь, если оптически одноосный кристалл вырезать так, чтобы одна из граней параллелепипеда была перпендикулярна к оптической оси, а свет направить вдоль этой оси.

Симметрия кристалла накладывает определенные ограничения на электрооптические коэффициенты Поккельса, часть из них обращается в нуль, некоторые могут оказаться равными между собой. Материал считается обладающим значительным электрооптическим эффектом, если его коэффициенты r порядка $10^{-10}…10^{-12}$ м/В. Поэтому при обычных внешних полях порядка 10^6 В/м линейное изменение показателя преломления составляет ~ 10^{-5}. Это означает, что существенные изменения оптической длины под действием эффекта Поккельса могут быть получены только в тех

случаях, когда длина кристалла в направлении распространения света ориентировочно в 10^5 раз превышает длину волны света.

Из механизма явления ясно, что эффект Поккельса столь же безынерционен, что и эффект Керра (быстродействие $10^{-9}…10^{-11}$ с). Из-за линейности и низкого рабочего напряжения в технике оптической связи чаще всего используют эффект Поккельса.

Квадратичный и другие эффекты более высокого порядка намного меньше линейного эффекта Поккельса, однако в центросимметричных кристаллах эффект Поккельса обращается в нуль и основную роль начинает играть квадратичный эффект Керра.

К электрооптическим материалам предъявляется ряд требований, позволяющий выбирать наиболее перспективные из них. Эти требования таковы:

- возможность выращивания монокристаллов оптического качества и достаточно крупных размеров с использоваением промышленных методов синтеза;
- высокая химическая стойкость кристаллов по отношению к действию окружающей атмосферы, нерастворимость их воде и других растворителях;
- достаточно высокая твердость, обеспечивающая хорошую обрабатываемость (микротвердость должна быть выше 1 ГПа);
- минимальная анизотропия теплового расширения, обеспечивающая высокую стабильность пьезоэлектриков, физические характеристики которых, в том числе и электрооптические постоянные, должны обладать малой температурной зависимостью и быть линейными по характеру;
- хорошие диэлектрические характеристики, в том числе в сильных полях, в широком частотном и температурном диапазонах.

Учет всех требований и оптимизация их сочетаний заметно сужают круг потенциально интересных для применения веществ. Применение многих перспективных электрооптических веществ ограничивается трудностями их получения в виде монокристаллов.

При выборе электрооптических материалов критерием их качества служат, в основном, полуволновое напряжение $U_{\lambda/2}$ и коэффициент электрооптической эффективности $\frac{1}{\varepsilon U_{\lambda/2}^2}$, позволяющий оценить управляющие мощности; учитываются также данные об электрооптическом коэффициенте Поккельса r, показателе преломления n и области оптической прозрачности кристалла λ.

Практически используемые материалы с эффектом Поккельса можно разделить на:

1) одноосные монокристаллы тетрагональной системы: KDP (KH_2PO_4 – дигидрофосфат калия, $r = 11\cdot10^{-12}$ м/В, $\lambda = 0{,}25…1{,}2$ мкм); DKDP (KH_2PO_4 – дейтерозамещенный дигидрофосфат калия, $r = 11\cdot10^{-8}$ м/В); ADP (NiH_4PO_4 – дигидрофосфат аммония, $r = 11\cdot10^{-12}$ м/В) и одноосные оксидные монокристаллы тригональной системы: LNO ($LiNbO_3$ – метаниобат лития,

$r = 30{,}8 \cdot 10^{-12}$ м/В, $\lambda = 0{,}5 \ldots 0{,}65$ мкм); LTO ($LiTaO_3$ – тантанат лития, $r = 33 \cdot 10^{-12}$ м/В, $\lambda = 0{,}5 \ldots 0{,}65$ мкм);

2) изотропные полупроводниковые соединения кубической системы: ZnTe – теллурид цинка, $r = 1{,}4 \cdot 10^{-12}$ м/В, $\lambda = 0{,}5 \ldots 0{,}6$ мкм; ZnSe – селенид цинка, $r = 1{,}1 \cdot 10^{-12} \ldots 2 \cdot 10^{-12}$ м/В, $\lambda = 0{,}6$ мкм; GaAs – арсенид галлия, $r = 1{,}6 \cdot 10^{-12}$ м/В, $\lambda = 1{,}0$ мкм; CdTe – теллурид кадмия, $r = 6{,}8 \cdot 10^{-12}$ м/В, $\lambda = 0{,}6$ мкм.

3) изотропные оксидные монокристаллы кубической системы, например, оксиды висмута: $Bi_{12}XO_{20}$, где X = Ge, Si, Ti. ($Bi_{12}GeO_{20}$ – $r = 2{,}8 \cdot 10^{-12} \ldots 3{,}1 \cdot 10^{-12}$ м/В, $\lambda = 0{,}85$ мкм; $Bi_{12}SiO_{20}$ – $r = 4{,}35 \cdot 10^{-12}$ м/В, $\lambda = 0{,}87$ мкм).

Каждый из материалов для ячеек может работать лишь в определенном диапазоне волн, где он незначительно поглощает энергию проходящего излучения. Так, например, кристаллы KDP могут работать в диапазоне 0,25…13 мкм, в пределах которого коэффициент пропускания превышает 20 %.

Датчики на материалах тетрагональных (KDP, KDA) и тригональных (LNO, LTO) систем обладают высокой чувствительностью, но вследствие температурной зависимости коэффициента естественного двойного лучепреломления нуждаются в температурной компенсации, обычно достигаемой высокоточной обработкой кристалла.

Перспективным является использование и полупроводниковых соединений кубической системы. Благодаря большим показателям преломления полупроводники оказываются подходящим метериалом для изготовления модуляторов. Хотя, например, GaAs обладает меньшим электрооптическим коэффициентом, чем конкурирующий материал KDP, тем не менее при волне $\lambda = 10{,}6$ мкм первый обладает бóльшим фактором качества: $(n^3 r)_{AsGa} = 5{,}9 \cdot 10^{-11}$ м/В и $(n^3 r)_{KDP} = 3{,}3 \cdot 10^{-11}$ м/В. CdTe обладает еще большим фактором качества – $(n^3 r)_{AsGa} = 12 \cdot 10^{-11}$ м/В при сходных характеристиках пропускания. Так как к электрооптическим модуляторам прикладывают сильные электрические поля ($E \geq 10^5$ В/м), то для них требуются кристаллы с высоким удельным сопротивлением ($\geq 10^9$ Ом·м), поэтому при комнатной температуре можно использовать лишь те полупроводники, ширина запрещенной зоны которых превышает 1,3 эВ. Однако, при увеличении температуры удельное сопротивление полупроводников уменьшается, что приводит к значительному снижению чувствительности модуляторов.

Оксидные монокристаллы кубической системы не имеют естественного двойного лучепреломления и обладают высоким сопротивлением, следовательно, высокой чувствительностью, что делает их наиболее пригодным материалом для ячеек Поккельса.

Таким образом, в современной оптоэлектронной технике распространены оксидные материалы, что связано, в частности, с их высокими механической и электрической прочностями, стойкостью к действиям среды, низкими диэлектрическими потерями.

Значительное увеличение электрооптических коэффициентов Поккельса происходит в сегнетоэлектрических кристаллах при приближении к точке Кюри. Известно, что электрооптические коэффициенты r для некоторых кристаллов, например, KH_2PO_4 и KD_2PO_4, в точке Кюри увеличиваются в ~ 1500 раз по сравнению с комнатной температурой, что позволяет снизить управляющие напряжения. Однако трудности охлаждения кристаллов и поддержание с высокой точностью их температуры ограничивают применение устройств, работающих при температуре, близкой к температуре Кюри.

На практике эффект Поккельса часто маскируется вторичным электрооптическим эффектом, обусловленным деформациями пьезокристалла при воздействии светового излучения за счет обратного пьезооптического эффекта. Эти деформации из-за наличия фотоупругости приводят к изменению показателя преломления, которое складывается с первичным эффектом Поккельса. На очень высоких частотах деформации кристалла малы и имеет место только первичный эффект Поккельса.

Особенно резко увеличивается деформация на частотах, соответствующих собственным колебаниям кристалла. Когда частота внешнего электрического поля совпадает с одной из собственных частот, деформации увеличиваются в Q раз, где Q – добротность соответствующего колебания. При таком резонансе электрооптический коэффициент может возрасти в 1000 раз, что позволяет во столько же раз снизить управляющее напряжение. Однако это явление наблюдается лишь в узкой полосе частот и сильно зависит от температуры. Для улучшения частотной характеристики широкополосной модуляции света с помощью эффекта Поккельса приходится специально гасить собственные колебания электрооптического кристалла, однако и в этом случае переход от низких частот к более высоким сопровождается уменьшением коэффициента Поккельса за счет пьезоэффекта.

На частотах около 0,6 МГц имеет место собственный резонанс кристалла. Высокочастотное значение постоянных Поккельса обычно не меняется вплоть до частот 10^{12} Гц, соответствующих частоте резонанса кристаллической решетки.

В качестве примера рассмотрим физические принципы работы широко используемого в электрооптике модулятора на базе ячейки Поккельса (рис. 4.28). Для ячейки Поккельса можно взять как пьезоэлектрические кристаллы KDP, обладающие естественным двулучепреломлением, так оксидные монокристаллы кубической системы $Bi_{12}XO_{20}$, являющиеся до наложения электрического поля изотропными. Для определенности рассмотрим ячейку Поккельса на базе изотропного кубического кристалла.

Модулятор представляет собой электрооптический кристалл, находящийся между двумя поляризаторами, установленными так, что их главные оптические оси скрещены под прямым углом. Если на ячейку не подается напряжение, то свет через нее не проходит. Так как разность фаз пропорциональна напряжению (линейный электрооптический эффект), то

ячейку Поккельса можно рассматривать как электрически управляемую фазовую пластину.

Практическое применение электрооптического эффекта основано на том,

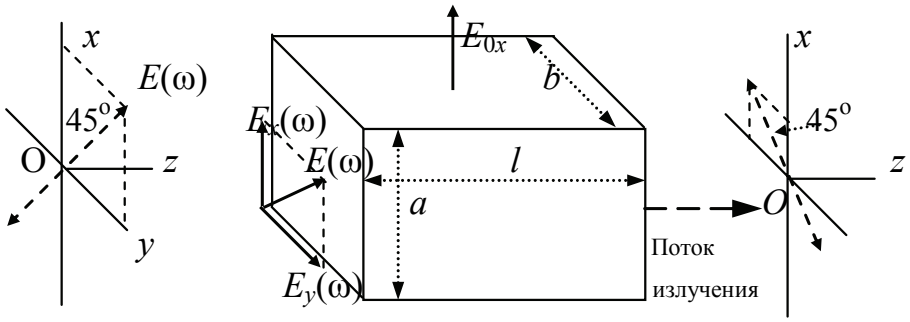


Рис. 4.28 – Электронный модулятор на базе ячейки Поккельса

что показатель преломления различен для света, поляризованного перпендикулярно и параллельно внешнему электрическому полю, и эта разность зависит от величины поля. На рис. 4.28 входящий свет, направленный вдоль оси z, поляризован под углом 45˚ по отношению к внешнему электрическому полю $\vec{E}_{0x}$. В кристалле происходит разложение падающей световой волны на волны двух типов: обыкновенной $E_x(\omega)$, поляризованной в вертикальной плоскости и необыкновенной $E_y(\omega)$, поляризованной в горизонтальной плоскости. При наличии внешнего поля $\vec{E}_{0x}$ возникает разность показателей преломления для продольной к внешнему электрическому полю $E_x(\omega)$ и поперечной $E_y(\omega)$ компонент волны

$$\Delta n = \frac{n^3 r \cdot \vec{E}_{0x}}{2}, \qquad (4.106)$$

где n – показатель преломления материала ячейки в отсутствие электрического поля; r – электрооптический коэффициент Поккельса; $n^3 r$ – фактор качества электрооптического материала; $\vec{E}_{0x}$ – напряженность внешнего электрического поля, направленного вдоль оси 0x. В результате при выходе световой электромагнитной волны из кристалла длиной l компоненты электромагнитного поля $E_x(\omega)$ и $E_y(\omega)$ будут обладать оптической разностью хода

$$\Delta_{\text{опт}} = \Delta n \cdot l = \frac{n^3 r \cdot \vec{E}_{0x} \cdot l}{2}. \qquad (4.107)$$

Соответствующая разность фаз между компонентами $E_x(\omega)$ и $E_y(\omega)$ будет равна

$$\Delta\varphi = \frac{2\pi}{\lambda}\Delta_{\text{опт}} = \frac{\pi \cdot n^3 r \cdot \vec{E}_{0x} \cdot l}{\lambda}. \qquad (4.108)$$

Если оптическая разность хода $\Delta_{\text{опт}} = \frac{\lambda}{2}$, то разность фаз $\Delta\varphi = \pi$, и в результате суперпозиции компонент $E_x(\omega)$ и $E_y(\omega)$ плоскость поляризации выходящего света будет повернута на 90° по отношению к плоскости поляризации входящего пучка, т. е. кристалл превратится в полуволновую пластинку. Таким образом, если кристалл с поперечными размерами $a \cdot b$ помещен между скрещенными поляризаторами, то выходной сигнал отсутствует при отсутствии напряжения и достигает максимума, когда приложено внешнее «полуволновое» напряжение:

$$U_{\lambda/2} = \vec{E}_{0x} a = \frac{\lambda a}{n^3 r \cdot l}. \qquad (4.109)$$

Ячейка Поккельса позволяет создавать световые импульсы при подаче полуволнового напряжения и реализовать импульсно-кодовую и широтно-импульсную модуляцию светового сигнала. Частота модуляции может превышать 10 ГГц. Данный тип модулятора характерен для дискретной оптики.

Эффект Поккельса широко применяется при создании различных устройств управления оптическим излучением, таких, как модуляторы света в волоконно-оптических линиях связи, переключателях оптических каналов и т.п.

4.3.3. Оптические квантовые генераторы

Во многих сегнетоэлектриках проявляются *нелинейные оптические эффекты*, в основе которых лежит нелинейная поляризация среды, т.е. зависимость показателя преломления от напряженности поля самой световой волны. Нелинейность оптических свойств сегнетоэлектрических кристаллов позволяет осуществить генерацию лазерного излучения.

Рубин является наиболее освоенным в лазерной технике монокристаллом, пригодным для возбуждения стимулированного когерентного излучения. Вообще для этих целей могут быть использованы многие кристаллы, отображающие законы строгой повторяемости и порядка в пространственном расположении частиц вещества с включением в их решетку активных атомов или ионов. Такие ионы могут быть получены из элементов с недостроенными внутренними электронными оболочками, например, элементы группы железа имеют недостроенную 3*d*-оболочку, лантаноиды – 4*f*-оболочку, актиноиды – 5*f*-оболочку.

Эти ионы и являются собственно активными, в то время как кристаллическая решетка основного материала выполняет роль матрицы. Вследствие экранирующего действия внешних электронов поле кристаллической решетки оказывает незначительное воздействие на уровни энергии активных ионов, в результате чего такие кристаллы обладают узкими и интенсивными линиями люминесценции в спектре излучения.

Существует много материалов, применяемых в качестве активных сред в лазерах. Сюда относятся различные диэлектрические кристаллы, стекла,

газы, полупроводники и даже плазма. Эта сторона дела подробно изучается в квантовой электронике. Ограничимся описанием устройства и принципа действия только твердотельного рубинового лазера.

Первый квантовый генератор света был создан в 1961 г. Мейманом на рубине. Рубин – это драгоценный камень красного или розового цвета, очень твердый, тугоплавкий, химически инертный, с высокими оптическими свойствами (показателем преломления). Природные рубины, как правило, невелики и могут быть различной формы и размеров. В радиоэлектронике в качестве активного тела в оптических квантовых генераторах стимулированного излучения (лазерах) применяют искусственные рубины, представляющие собой достаточно большие монокристаллы $A1_2O_3$ (корунд), в которых часть ионов алюминия Al^{3+} замещена ионами хрома Cr^{3+}. Количеством хрома определяется цвет рубина. Розовый рубин содержит около 0,05 мас.%, красный – 0,5 мас. % хрома. Для активных элементов оптических квантовых генераторов применяют розовый рубин. Искусственные кристаллы рубина обычно выращивают в печах по методу Вернейля. По этому методу тщательно размельченный порошок окиси алюминия с добавкой Cr_2O_3 медленно падает в пламя водородно-кислородной горелки. Отдельные частички порошка, проходя через пламя, расплавляются и затем кристаллизуются на затравочном кристалле, помещенном вне пламени. Полученную заготовку отжигают, а потом обрабатывают, придавая ей необходимую форму и размеры. Высококачественные кристаллы рубина могут быть получены и методом вытягивания из расплава.

Ионы хрома играют основную роль в работе квантового генератора. Корунд – это диэлектрик с широкой запрещенной зоной между валентной зоной и зоной проводимости. Энергетические уровни хрома в корунде лежат в этой запрещенной зоне. Они были изучены задолго до создания лазеров. Основным (невозбужденным) уровнем является уровень E_1 (рис. 4.29). Он имеет сложную структуру, которая, однако, не играет роли в работе квантового генератора. Выше лежат близкие возбужденные уровни $E_{2а}$ и $E_{2б}$. Это узкие уровни. При переходе с них на основной уровень E_1 излучается красный свет с длинами волн $\lambda = 694{,}3$ нм и $\lambda = 692{,}9$ нм. Этот свет и придает рубину характерную розовую или красную окраску (в зависимости от концентрации ионов хрома). Более интенсивна линия $\lambda = 694{,}3$ нм. По этой причине только она и усиливается при работе лазера. Вторая линия не играет роли. Выше уровней $E_{2а}$ и $E_{2б}$ расположены две сравнительно широкие полосы энергий E_3 и E_4.

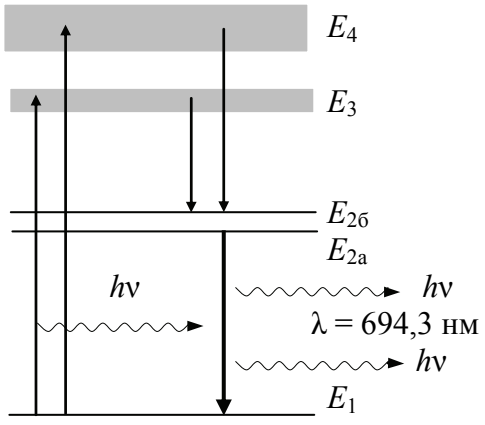


Рис. 4.29 – Зонная диаграмма энергетических уровней в рубиновом лазере

Переходы между уровнями этих полос сопровождаются излучением зеленого и голубого света соответственно.

Инверсная заселенность создается между уровнями E_1 и E_2. Для этого используется так называемая *оптическая накачка*, т. е. освещение кристалла рубина мощной вспышкой света. Рубину придают форму цилиндрического стержня с диаметром 0,001…0,02 м и длиной от 0,02 до 0,20 м и более. Концы стержня тщательно отполированы, их серебрят и тогда они могут служить зеркалами. Зеркала могут быть и внешними. Для освещения рубинового стержня применяют импульсные ксеноновые газоразрядные лампы.

Если бы энергетический спектр состоял только из двух линий E_1 и E_2, то с помощью оптической накачки создать инверсионную заселенность этих уровней было бы невозможно. Это связано с тем, что индуцированные переходы атомов с нижнего уровня на верхний и обратно идут с одинаковыми скоростями. Поэтому лампа самое большее могла бы лишь уравнять населенности обеих уровней. Положение меняется благодаря наличию третьего, широкого энергетического уровня, состоящего из полос E_3 и E_4. Лампа переводит атомы хрома из невозбужденного состояния в возбужденное, т.е. в энергетические полосы E_3 и E_4. Значительная ширина этих полос имеет большое практическое значение, так как благодаря большой ширине полос на их возбуждение уходит 10…15% лучистой энергии лампы. (Лампа излучает свет, близкий к белому). На уровнях E_3 и E_4 возбужденные атомы хрома живут очень короткое время ($\sim 10^{-8}$ с). За это время они переходят на один из уровней E_2. При таком переходе атомы хрома расходуют свою энергию на возбуждение колебаний кристаллической решетки. Возможность обратного возвращения атома из полос E_3 и E_4 на уровень E_1, хотя и существует, но вероятность такого процесса пренебрежимо мала сравнительно с вероятностью перехода на уровни E_2. Уровни E_2 метастабильны, время жизни атома на них $\sim 10^{-3}$ с, что по атомным масштабам очень велико. Если переводить атомы с уровня E_1 на урони E_3 и E_4 достаточно быстро, то на уровни E_2 можно перевести больше половины атомов хрома. Тогда на уровнях E_2 окажется больше атомов, чем на уровне E_1, т.е. возникнет инверсионная заселенность этих уровней.

Значительная доля лучистой энергии (более 50%), поглощенной рубиновым стержнем, тратится на его нагревание. При температурах, превышающих примерно 1000 К, рубиновый лазер перестает генерировать. Поэтому в некоторых конструкциях лазеров предусмотрено охлаждение водой или даже жидким азотом.

Кристалл рубина обладает оптической анизотропией и имеет почти кубическую симметрию, несколько искаженную вдоль одной из пространственных диагоналей, в результате чего истинная симметрия кристалла – ромбоэдрическая. Рубин – одноосный кристалл. Обыкновенный и необыкновенный показатели преломления светло-красного рубина, употребляемого в лазерах, равны соответственно n_o = 1,7653 и n_e = 1,7573 (для λ = 656 нм). Рубиновый лазер может давать поляризационный свет без

каких либо специальных поляризационных приспособлений. Для этого оптическая ось рубинового стержня не должна совпадать с его геометрической осью. Возникновение линейной поляризации нельзя объяснить различием коэффициентов отражения обыкновенного и необыкновенного лучей, так как разность $n_o - n_e$ слишком мала. Поляризация возникает потому, что в обыкновенной волне направление волновой нормали и луча совпадают между собой, а в необыкновенной – не совпадают. Чтобы зеркала лазера действовали эффективно, как резонатор, необходимо, чтобы волновые нормали были к ним перпендикулярны. Но тогда в рубине только обыкновенный луч будет распространяться параллельно геометрической оси стержня, а необыкновенный пойдет под углом к ней, попадая на боковую поверхность стержня. Поэтому резонатор будет в основном усиливать обыкновенные лучи. Из лазера будут преимущественно выходить обыкновенные лучи, в которых электрический вектор электромагнитной волны перпендикулярен к оптической и геометрической осям кристалла. Если эти оси совпадают, то поляризации не возникает. Если же, как показали наблюдения, угол между ними заключен в пределах от 60 до 90˚, то линейная поляризация полная.

Наиболее распространен импульсный режим работы рубинового лазера.

Кроме рубина в качестве материалов рабочих элементов лазера используют некоторые другие кристаллы, например, гранаты, а также различные стекла с присадками соответствующих активаторов.

Применение гранатов в современной квантовой электронике существенно расширилось, что объясняется удачным сочетанием их механических, теплофизических и оптических свойств. Этот класс материалов представляет собой более 30 различных по составу кристаллических матриц, лазерные свойства которым придают ионы группы железа (Cr^{3+}, Ni^{2+}) и шесть редкоземельных илнов (Nd^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, Yb^{3+}).

Гранаты кристаллизуются в кубическую структуру. Катионы располагаются в трех подрешетках с различной координацией по кислороду. Кристаллы гранатов имеют важное преимущество перед другими классами лазерных кристаллов, заключающееся в наличии трех различных кристаллографических положений для катионов в структуре граната. Это дает возможность в широких пределах изменять состав как самих матриц, так и легирующих добавок, позволяя тем самым создавать матрицы с необходимыми спектральными характеристиками.

Наибольшее значение как активный элемент получил иттрийалюминиевый гранат $Y_3Al_5O_{12}$, легированный неодимом Nd^{3+}, который является основным материалом промышленных твердотельных лазеров. Кристалл иттрийалюминиевого граната (ИАГ) оптически изотропен и имеет кубическую решетку структурного типа граната.

У лазеров на основе ИАГ низкая пороговая энергия возбуждения при комнатной температуре, высокая механическая прочность и хорошая теплопроводность, что обеспечивает надежность лазеров, работающих в

непрерывном и высокочастотном режимах. Длина волны излучения лазера на основе ИАГ, легированного неодимом – 1,06 мкм.

Следует отметить, что создание оптически совершенных кристаллов с однородным распределением примесей как по сечению, так и по длине кристалла представляет значительную технологическую трудность, и стоимость их довольно высока.

К преимуществам стекол как лазерных активных материалов относятся: технологичность и простота изготовления образцов больших размеров; дешевизна сырья; высокая оптическая однородность и изотропность свойств образцов большого размера; возможность введения активатора в необходимых концентрациях с равномерным распределением его по объему.

В то же время по сравнению с кристаллами стекла имеют и недостатки, к которым относятся: низкая теплопроводность и большой температурный коэффициент линейного расширения; ограниченная область прозрачности (0,33…2,5 мкм), которая при введении примесей может еще более сужаться; сравнительно слабая фотохимическая стойкость.

Наиболее часто применяется стекла с неодимом и эрбием. Длины волн излучения лазеров на основе неодимового и эрбиевого стекол – 1,06 мкм и 1,54 мкм соответственно.

Сравнение свойств кристаллов и стекол показывает, что эти материалы дополняют друг друга и, следовательно, одинаково важны для лазерной техники.

4.3.4. Конденсаторы

На использовании пассивных свойств диэлектрических материалов основано их применение в электротехнике. При использовании *диэлектрических материалов в качестве диэлектрика конденсатора* необходимо иметь материал с как можно бóльшей величиной диэлектрической проницаемости.

В зависимости от назначения конденсаторы условно разделяют на конденсаторы общего и специального назначения. Конденсаторы общего назначения используют в основном в приборах, где к ним не предъявляются специальные требования. К специальным конденсаторам относят: высоковольтные, импульсные, помехозащитные, пусковые и прочие.

Емкость постоянных конденсаторов фиксирована и в процессе эксплуатации никак не регулируется. Конденсаторы переменной емкости допускают изменение емкости в процессе функционирования аппаратуры. Емкость подстроечных конденсаторов возможно изменять при разовом или периодическом регулировании.

Важную роль в эксплуатации конденсаторов играет тип диэлектрического материала.

Полиэтилентерефталатные конденсаторы имеют очень малые токи утечки и, соответственно, малые потери. Поэтому их выгодно использовать, в цифровых и аналоговых преобразователях, таймерах, генераторах низких частот.

Полистирольные и фторопластовые конденсаторы тоже имеют малые токи утечки. Кроме того, их свойства очень мало изменяются с изменением частоты. Поэтому такие конденсаторы используют в контурах, где важную роль играет стабильность параметров.

Бумажные конденсаторы имеют большую реактивную мощность. Поэтому их широко используют для защиты от индустриальных помех, как искрогасящие.

Оксидные конденсаторы имеют большую удельную емкость. Поэтому их очень выгодно использовать в сглаживающих фильтрах.

Керамические конденсаторы имеют малую индуктивность. Это позволяет использовать их в качестве блокировочных и высокочастотных конденсаторов. В последнем случае их используют для термокомпенсации при фиксированной настройке контуров.

ВОПРОСЫ.

1. Что называют поляризацией диэлектрика? Какие виды поляризации можно считать мгновенными, а какие замедленными?
2. В чем различие между ионной и ионно-релаксационной поляризациями? Что характеризует время релаксации и от каких факторов оно зависит?
3. Что понимают под линейными и нелинейными, полярными и неполярными диэлектриками?
4. Что называют диэлектрическими потерями? Назовите механизмы диэлектрических потерь.
5. Какие диэлектрики называют активными? Какое отличие между активными и пассивными диэлектриками?
6. Как объяснить диэлектрический гистерезис и нелинейность зависимости заряда от напряжения у сегнетоэлектриков?
7. Что называют сегнетоэлектрической точкой Кюри?
8. Что такое прямой и обратный пьезоэффекты? В каких диэлектриках можно наблюдать эти явления?
9. Что такое пироэлектрический эффект?
10. Какова природа электретного состояния в диэлектриках?
11. Объяснить физические процессы, протекающие в слоистых диэлектриках. Что такое гетерогенные диэлектрики? Как находится эффективная диэлектрическая проницаемость гетерогенных диэлектриков?
12. Перечислить види пробоев в диэлектриках. Объяснить различия между ними.
13. Как объяснить оптическую анизотропию в кристаллах. Что такое положительные и отрицательные кристаллы?
14. Что такое искусственная анизотропия? Объяснить физическую сущность эффектов Поккельса и Керра.
15. Какие требования предъявляются к диэлектрику как лазерному материалу?
16. Каковы физические принципы работы световолокна? Назовите основные характеристики оптических волокон?

17. Описать физические принципы работы оптических модуляторов. Какие материалы применяются для оптических модуляторов на основе ячейки Поккельса?
18. Применение активных и пассивных диэлектриков в конденсаторах. Перечислить типы конденсаторов, указать их назначение.

ЗАДАЧИ.

1. Вычислить поляризованность монокристаллов каменной соли, считая, что смещение ионов под действием электрического поля от положения равновесия составляет 1% расстояния между ближайшими соседними ионами. Элементарная ячейка кристалла имеет форму куба, рассояние между соседними ионами $a = 0{,}28$ нм. (Ответ: поляризованность $P = 0{,}02$ Кл/м^2).
2. Между пластинами плоского конденсатора без воздушных промежутков зажат лист гетинакса толщиной $h = 1$ мм. На конденсатор подано напряжение $U = 200$ В. Найти поверхностную плотность заряда на пластинах конденсатора σ_1 и на диэлектрике $\sigma_д$. Диэлектрическая проницаемость гетинакса $\varepsilon = 6$. (Ответ: $\sigma_1 = 10^{-5}$ Кл/м^2; $\sigma_д = 8{,}85\cdot10^{-6}$ Кл/м^2).
3. Композиционный термокомпенсированный керамический материал изготовлен на основе двух диэлектриков с диэлектрическими проницаемостями $\varepsilon_1 = 40$ и $\varepsilon_2 = 80$. Предполагая хаотическое распределение компонентов, определить состав керамики, если температурные коэффициенты соответственно равны $\alpha_{\varepsilon1} = 2\cdot10^{-4}$ К$^{-1}$ и $\alpha_{\varepsilon2} = -1{,}5\cdot10^{-3}$ К$^{-1}$. Найти диэлектрическую проницаемость композитного диэлектрика. (Ответ: $\varepsilon_2 = 43{,}2$).
4. Тангенс угла диэлектрических потерь tgδ неполярного диэлектрика на частоте 50 Гц равен 10^{-3}. Вычислить активную мощность рассеяния $P_а$ в конденсаторе из этого диэлектрика на частоте $f = 1$ кГц при напряжении 1 кВ, если емкость конденсатора C равна 1000 пФ. (Ответ: В неполярном диэлектрике активная мощность рассеяния не зависит от частоты и равна $P_а = 3{,}14\cdot10^{-4}$ Вт).
5. Определить удельные диэлектрические потери в плоском конденсаторе, изготовленном из пленки полистирола толщиной 20 мкм, если на конденсатор подано напряжение 2 В частотой 2 МГц (для полистирола $\varepsilon = 2{,}5$; $\mathrm{tg}\delta = 2\cdot10^{-4}$). (Ответ: $P_а = 600$ Вт/м^3).
7. Известно, что при тепловом пробое диэлектрик толщиной 4 мм пробивается при напряжении 15 кВ на частоте 100 Гц. При каком напряжении промышленной частоты пробьется такой же диэлектрик толщиной 2 мм? (Ответ: $U_{пр} = 15$ кВ).
8. Спонтанная поляризованность монокристаллов титанита бария при комнатной температуре равна 0,25 Кл/м^2. Предполагая, что причиной возникновения спонтанной поляризации является только смещение иона титана из центра элементарной кубической ячейки, определить это смещение. Период решетки титаната бария $a = 0{,}4$ нм. (Ответ: $\Delta l = 0{,}025$ нм).

9. На пластину *X*-среза пьезоэлектрического кварца толщиной h = 1 мм вдоль оси *X* воздействует механическое напряжение $\sigma = 10^5$ Н/м^2. Определить разность потенциалов между противоположными плоскостями пластины, если в направлении оси *X* пьезомодуль продольного пьезоэффекта $d_{11}= 2{,}3\cdot10^{-12}$ Кл/Н. Диэлектрическую проницаемость кварца принять равной 4,6. (Ответ: U = 5,77 В).

10. Определить напряженности электрического поля E_1 и E_2 в каждом слое конденсатора, изображенного на рис. 4.14. В конденсаторе использована пленка политетрафторэтилена и пропитанная конденсаторная бумага одинаковой толщины ($h_1 = h_2 = 20$ мкм). Известны диэлектрические проницаемости ε и удельные объемные сопротивления ρ этих материалов: $\varepsilon_1 = 2$; $\varepsilon_2 = 4$; $\rho_1 = 10^{15}$ Ом·м; $\rho_2 = 10^8$ Ом·м. Расчет провести для постоянного напряжения 100 В и переменного напряжения амплитудой 100 В. (Ответ: На переменном напряжении E_1 = 3,33 МВ/м; E_2 = 1,67 МВ/м; На постоянном напряжении E_1 = 5 МВ/м; E_2 = 0,5 В/м U = 5,77 В).

11. Определить максимально возможное значение показателя преломления материала светоизолирующей оболочки n_2 цилиндрического световода, при котором все световые лучи из конического пучка с углом u = 120º при вершине испытывают полное внутреннее отражение на границе сердцевина-оболочка (см. рис. 4.26). Показатель преломления сердцевины n_1 принять равным 1,6. Внешней средой является воздух ($n_в$ = 1). (Ответ: n_2 = 1,345).

12. Свет, излучаемый светодиодом, с длиной волны, полученной в задаче 10 (см. гл. III), распространяется по кварцевому одномодовому световоду. Окно прозрачности соответствует длине волны λ = 850 нм. Дисперсия $\frac{dn}{d\lambda}$ в оптоволокне зависит от длины волны в соответствии с соотношением $\frac{10^{-14}}{\lambda^3}$ м$^{-1}$. Найти максимальную длину световода, если предельная наибольшая частота следования импульсов в нем f_{max} = 200 МГц. Определить полное затухание в оптоволокне полученной длины, если удельный коэффициент потерь данного оптоволокна β_0 = 2,7 дБ/км. (Ответ: L = 3060 м; β = 8,26 дБ).

13. Свет, излучаемый светодиодом, с длиной волны, полученной в задаче 10 (см. гл. III), модулируется на ячейку Поккельса, фактор качества которой равен 10^{-9} м/В. Найти полуволновое напряжение $U_{\lambda/2}$ для данной ячейки Поккельса с поперечными размерами 0,15·0,15 мм и длиной 10 мм. Рассчитать разность показателей преломления двух составляющих проходящего света: поляризованного вдоль внешнего электрического поля и поперек него. (Ответ: $U_{\lambda/2}$ =15 В; $\Delta n = 3{,}8\cdot10^{-5}$).

Глава V. МАГНЕТИКИ

5.1. Физические процессы в магнетиках

5.1.1. Вещество в магнитном поле. Намагничивание

Подобно тому, как для количественного описания поляризации диэлектриков вводилась поляризованность, для количественного описания намагничивания магнетиков вводят векторную величину – *намагниченность*, определяемую магнитным моментом единицы объема магнетика:

$$\vec{J} = \vec{P}_m / V = \sum \vec{p}_a / V, \qquad (5.1)$$

где $P_m = \sum \vec{p}_a$ – магнитный момент магнетика, представляющий собой векторную сумму магнитных моментов отдельных атомов (молекул).

Рассматривая характеристики магнитного поля, мы вводили вектор магнитной индукции $\vec{B}$, характеризующий результирующее магнитное поле, создаваемое всеми макро- и микротоками, и вектор напряженности $\vec{H}$, характеризующий магнитное поле макротоков. Следовательно, магнитное поле в веществе складывается из двух полей: внешнего поля $\vec{B}_0$, создаваемого намагничивающим током проводимости I в вакууме, и поля $\vec{B}'$, создаваемого молекулярными токами I' намагниченного вещества

$$\vec{B} = \vec{B}_0 + \vec{B}', \qquad (5.2)$$

где $\vec{B}_0 = \mu_0 \vec{H}$.

Для описания поля, создаваемого молекулярными токами, рассмотрим магнетик в виде цилиндра с площадью сечения S и длиной l, внесенный в однородное внешнее поле с индукцией $\vec{B}_0$. Возникающее в магнетике магнитное поле молекулярных токов будет направлено противоположно внешнему полю для диамагнетиков и совпадать с ним по направлению для парамагнетиков. Плоскость сечения цилиндра и плоскости молекулярных токов перпендикулярны вектору внешнего поля $\vec{B}_0$. Если рассмотреть любое сечение цилиндра, перпендикулярное к его оси, то во внутренних участках сечения магнетика молекулярные токи соседних атомов направлены навстречу друг другу и взаимно компенсируются (см. рис. 5.1). Нескомпенсированными будут лишь молекулярные токи, выходящие на боковую поверхность цилиндра.

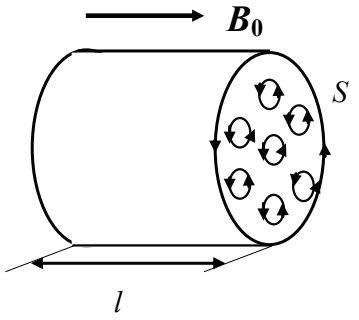


Рис. 5.1 – Намагничивание вещества во внешнем магнитном поле

Ток, текущий по боковой поверхности цилиндра, подобен току в соленоиде и создает внутри него поле, магнитную индукцию $\vec{B}'$ которого можно вычислить, учитывая, что $N = 1$ витку (принимаем, что для рассматриваемого цилиндра $\mu = 1$):

$$\vec{B}' = \mu_0 I'/l \,, \tag{5.3}$$

где I' – сила молекулярного тока; I'/l – линейная плотность тока.

Магнитный момент магнетика можно найти по формуле

$$P_m = \frac{I' \cdot l \cdot S}{l} = \frac{I' \cdot V}{l}, \tag{5.4}$$

где V – объем магнетика. Тогда намагниченность магнетика равна

$$\vec{J} = \vec{P}_m / V = I' / l \,. \tag{5.5}$$

Подставив (5.5) в (5.3) получим, что $\vec{B}' = \mu_0 \vec{J}$. Тогда из формулы (5.2) следует, что

$$\vec{B}_0 = \mu_0 \vec{H} + \mu_0 \vec{J} \tag{5.6}$$

или

$$\frac{\vec{B}}{\mu_0} = \vec{H} + \vec{J} \,. \tag{5.7}$$

Как показывает опыт, намагниченность в несильных полях прямо пропорциональна напряженности поля, вызывающего намагничивание:

$$\vec{J} = \chi \vec{H} \,, \tag{5.8}$$

где χ – безразмерная величина, называемая *магнитной восприимчивостью* вещества.

Для диамагнетиков $\chi < 0$ (поле молекулярных токов противоположно внешнему полю), а для парамагнетиков $\chi > 0$ (поле молекулярных токов совпадает с внешним полем).

Из формул (5.5) и (5.7) можно получить выражение:

$$\vec{B} = \mu_0 (1 + \chi) \vec{H} = \mu_0 \mu \vec{H} \,. \tag{5.9}$$

Безразмерная величина

$$\mu = (1 + \chi) \tag{5.10}$$

представляет собой *магнитную проницаемость* вещества.

Таким образом, для диамагнетиков $\chi < 0$, $\mu < 1$, для парамагнетиков и ферромагнетиков $\chi > 0$, $\mu > 1$.

Диа- и парамагнетики называются *слабомагнитными веществами*, *ферромагнетики* – сильномагнитными веществами. Ферромагнетики обладают высокой относительной магнитной проницаемостью (μ может

достигать значений 10^5) и спонтанной намагниченностью, т. е. способностью оставаться намагниченными даже в отсутствие внешнего магнитного поля

В этой главе основное внимание будет уделено именно ферромагнетикам, так как магнитные свойства диа- и парамагнетиков не находят широкого применения в технике.

5.1.2. Свойства ферромагнетиков

5.1.2.1 Природа ферромагнитного состояния

Первая количественная теория ферромагнетизма была разработана французским физиком Вейссом. Он предположил, что атомы ферромагнетика, как и парамагнетика, обладают магнитными моментами и взаимодействуют между собой силами, зависящими от угла между этими магнитными моментами. Эти силы стремятся установить магнитные моменты соседних атомов параллельно друг другу. В результате ориентации магнитных моментов атомов в определенном направлении и создается намагничивание ферромагнетика. В теории Вейсса силы взаимодействия между атомами сводятся к некоторому «эффективному» магнитному полю, которое и ориентирует атомы ферромагнетика. Эффективное поле является суммой обычного макроскопического поля в веществе $\vec{H}$ и некоторого гипотетического «молекулярного поля».

Теория Вейсса объяснила основные факты из области ферромагнетизма: спонтанную намагниченность, возникновение доменной структуры, существование температуры Кюри. Однако понятие молекулярного поля не давало объяснения о природе этого поля.

Физическая природа молекулярного поля была установлена Я. И. Френкелем в 1927 г. и почти одновременно с ним – В. Гейзенбергом на основе квантовой механики.

Таблица 5.1 – Заполнение электронных оболочек в атоме железа

Номер оболочки n	Обозначение состояния		
	s	p	d
1	↑↓		
2	↑↓	↑↓ ↑↓ ↑↓	
3	↑↓	↑↓ ↑↓ ↑↓	↑↓ ↑ ↑ ↑ ↑
4	↑↓		

Объяснение причин спонтанной намагниченности и природы сил, ее вызывающих, можно кратко изложить следующим образом.

Как правило, отличным от нуля магнитным моментом обладают лишь те атомы и ионы, которые имеют внутренние, не полностью заполненные электронами, оболочки. Определим, в чем же заключается особенности атомов с недостроен-

ными внутренними оболочками, а именно, атомов кобальта, никеля и железа. Рассмотрим более детально атом железа. В атоме железа (№ 26) его электроны распределены по четырем оболочкам. Первая и вторая оболочки заполнены полностью и содержат соответственно 2 и 8 электронов. Третья и четвертая оболочки не достроены: в третьей оболочке находится 14 электронов (вместо ожидаемых 18), а в четвертой – 2 электрона. 14 электронов третьей оболочки распределены по подоболочкам следующим образом: в первой *s* подоболочке – 2 электрона, во второй *p* подоболочке – 6 электронов, в третьей *d* подоболочке – 6 электронов. Спины электронов, принадлежащих к каждой подоболочке, могут быть ориентированы в двух противопо-ложных направлениях. В застроенных первых двух оболочках атома железа магнитные спиновые моменты электронов взаимно компенсируют друг друга. В третьей оболочке *s* и *p* подоболочки также характерны тем, что спиновые моменты электронов на этих подоболочках компенсируют друг друга. Что же касается третьей *d* подоболочки, то из шести находящихся на ней электронов пять имеют спины, ориентированные в одном направлении, и лишь один электрон имеет спин, ориентированный противоположно (табл. 5.1). Итак, в атоме железа спины четырех электронов в третьей оболочке (3*d*) остаются некомпенсированными. Это не противоречит принципу Паули, разрешающему двум электронам находиться в одном и том же состоянии в атоме, если только они имеют противоположные спины, так как состояния этих четырех электронов различны. Детальные расчеты, показывающие, почему электроны располагаются именно таким образом, очень сложны и являются задачей квантовой механики.

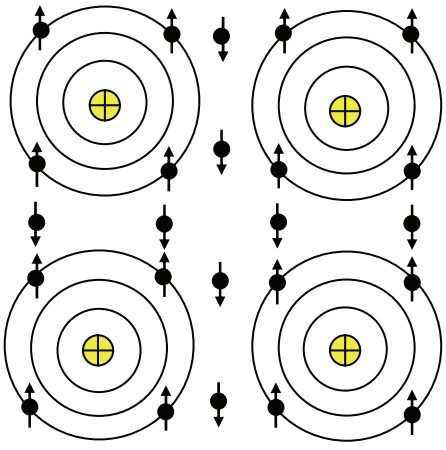


Рис. 5.2 – Ориентирующие воздействие спинов электронов проводимости на спины электронов недостроенных оболочек атомов

Когда атом железа окружен другими атомами железа, четыре неспаренных электрона на его внешней оболочке с $n = 3$ вынуждают ближайшие электроны проводимости приобрести противоположные спины. Спины электронов проводимости, в свою очередь, вынуждают стать противоположными спины электронов на внешней оболочке с $n = 3$ у соседнего атома (рис. 5.2). Следовательно, магнитные моменты соседних атомов выстраиваются в одном направлении не из-за магнитного воздействия их друг на друга, а благодаря своеобразному применению принципа Паули: электроны в одной и

той же пространственной области должны иметь противоположные спины. Такое двойное взаимодействие спинов электронов в недостроенных оболочках атомов через электроны проводимости называется *обменным*.

Обменное взаимодействие – это квантовомеханический эффект, не имеющий аналога в классической физике. Результирующая энергия взаимодействия, наряду с кулоновским членом, содержит добавочный член $E_{об}$, зависящий от взаимной ориентации спинов. Эта добавочная энергия и есть *энергия обменного взаимодействия* атомов сортов i и j, имеющих спиновые моменты $M_i^{сп}$ и $M_j^{сп}$. Ее находят по формуле:

$$E_{об} = -2A\left(M_i^{сп} M_j^{сп}\right), \qquad (5.11)$$

где A – *обменный интеграл*, имеющий размерность энергии. Численное значение и знак обменного интеграла зависят от степени перекрытия электронных оболочек атомов i и j, т. е. от расстояния между атомами в кристаллической решетке. Скалярное произведение спиновых моментов при параллельной ориентации спинов равно $\left(M_i^{сп} M_j^{сп}\right) = 1$, а при антипараллельной ориентации – $\left(M_i^{сп} M_j^{сп}\right) = -1$. Энергетически более выгодным является состояние с наиболее низкой энергией. Если обменный интеграл положителен ($A > 0$), то энергия обменного взаимодействия $E_{об}$ будет минимальной при условии, когда ей соответствует параллельная ориентация спиновых моментов. Действительно, из формулы (5.11) мы получаем $E_{об\,min} = -2A \cdot 1 = -2A$. Если же обменный интеграл отрицателен ($A < 0$), то энергия обменного взаимодействия будет минимальной при условии, когда ей соответствует антипараллельная ориентация спиновых моментов (из формулы (5.11) мы получаем $E_{об\,min} = -2(-A) \cdot -1 = -2A$).

Наличие внутренних незаполненных оболочек в атоме является необходимым, но еще недостаточным условием для возникновения ферромагнетизма. Например, внутренние незаполненные слои имеют атомы марганца, хрома, ванадия, все лантаниды, хотя они и не являются ферромагнетиками. Установлено, что значение обменного интеграла A зависит от отношения межатомного расстояния a к диаметру d недостроенной внутренней оболочки атома веществ. Ферромагнетизм наблюдается только для элементов, у которых обменный интеграл положителен ($A > 0$), что выполняется для соотношения $a/d > 1{,}5$ (рис. 5.3, область II). Из химически чистых элементов ферромагнитными свойствами обладают только кристаллы девяти: α-железо (α-Fe), кобальт (Co), никель (Ni), гадолиний (Gd) и при температурах ниже 0° С пять редкоземельных элементов (эрбий, диспозий, тулий, гольмий и тербий).

Однако в ряде случаев для металлов, не являющихся ферромагнетиками, при изменении постоянной решетки за счет легирования другими элементами можно добиться того, что они становится ферромагнетиками.

Такой эффект наблюдается у марганца (a/d = 1,47), при его легировании азотом в малых концентрациях, когда отношение a/d оказывается около 1,5.

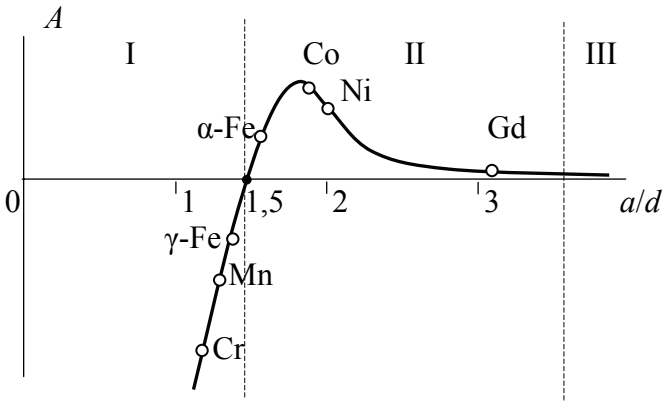


Рис. 5.3 – Зависимость обменного интеграла А от отношения межатомного расстояния к диаметру недостроенной внутренней оболочки

Марганец приобретает ферромагнитные свойства и в сплавах с такими неферромагнитными элементами, как висмут, олово (Mn-Cu-Al, соединения MnSb, MnBi, CrO_2, $MnOFe_2O_3$) и др., атомы которых «раздвигают» решетку марганца. В результате этого расстояние между атомами увеличивается, что приводит к необходимости взаимодействия через электроны проводимости. Число ферромагнитных материалов очень велико, так как к ферромагнитным материалам относятся сплавы самих ферромагнитных элементов и их сплавы с неферромагнитными элементами. Более того, известны ферромагнитные сплавы из неферромагнитных элементов.

При уменьшении расстояния между атомами интеграл обмена, пройдя максимальное значение, начинает убывать и становится отрицательным (рис. 5.3, область I). При отрицательном значении интеграла обмена спины электронов в атомах самопроизвольно устанавливаются антипараллельно друг другу, что приводит к особому явлению, называемому *антиферромагнетизмом*. Границей области ферромагнетизма и антиферромагнетизма является отношение a/d = 1,5. Кристалл марганца, несмотря на наличие пяти некомпенсированных спинов, является антиферромагнетиком, как и кристалл хрома. Антиферромагнитными свойствами обладает большое число соединений, например, соединения марганца, железа, хрома с другими элементами (MnO, MnS, FeO) и т.д.

При значительных расстояниях между атомами a, когда оно в 3…4 раза превышает диаметр незаполненной электронной оболочки d (рис. 5.3, область III), обменное взаимодействие стремится к нулю. В этом случае обменные силы уже не могут противодействовать тепловому движению и не в состоянии упорядочить расположение спинов. Соответственно, такие вещества должны проявлять парамагнитные свойства.

Ферромагнетизм и антиферромагнетизм представляют собой квантовое явление, являющееся свойством не отдельных атомов, а всего кристалла.

Для каждого ферромагнетика существует определенная температура $T_К$, называемая *точкой Кюри*, при которой он теряет свои сильные магнитные свойства. При нагревании ферромагнетика выше точки Кюри он превращается в парамагнетик. Чем выше интеграл обмена ферромагнетика (рис. 5.3), тем выше точка Кюри. Для кобальта, железа, никеля и гадолиния точка Кюри равна соответственно 1400 К, 1043 К, 631 К и 289 К. Переход вещества из ферромагнитного состояния в парамагнитное, происходящий в точке Кюри, не сопровождается поглощением или выделением теплоты, т.е. в точке Кюри происходит фазовый переход II рода. При остывании образца ниже точки Кюри ферромагнитные свойства восстанавливаются.

Выше уже указывалось, что в зависимости от структуры кристалла обменные силы могут вызывать параллельную или антипараллельную ориентацию электронных спинов соседних атомов. В простейшем случае (в отсутствие намагничивающего поля) электронные спины атомов образуют две пространственные подрешетки с противоположно направленными спинами, вставленные одна в другую. Если величина намагниченности обеих подрешеток одинакова, то их магнитные моменты компенсируют друг друга, так что кристалл в целом не обладает магнитным моментом. Такое явление, названное *антиферромагнетизмом*, было теоретически предсказано Неелем (в 1932 г.) и независимо от него Л. Д. Ландау (в 1933 г.). При низких температурах антиферромагнетики имеют ничтожно малую магнитную восприимчивость, которая увеличивается с ростом температуры. При некоторой температуре, называемой *антиферромагнитной точкой Кюри* или *точкой Нееля*, магнитное упорядочение спинов нарушается, и антиферромагнетик превращается в парамагнетик. Для хрома и марганца точка Нееля равна соответственно 475 К и 100 К.

Если величина намагниченности обеих спиновых подрешеток не одинакова, то возникает некомпенсированный антиферромагнетизм. Такие материалы называются *ферримагнетиками*. По отношению к внешнему магнитному полю ферримагнетик подобен ферромагнетику, и его относительная магнитная проницаемость намного больше 1 и может достигать нескольких тысяч. Поэтому ферримагнитные материалы – *ферриты*, состоящие из оксидов металлов, иногда называют неметаллическими ферромагнетиками. Ферритами является большая группа материалов, являющихся соединениями оксида железа Fe_2O_5 с оксидами других металлов.

5.1.2.2 Энергия магнитной анизотропии

В монокристаллах ферромагнетиков существуют направления *легкого* и *трудного намагничивания*. Число таких направлений определяется симметрией кристаллической решетки. В отсутствие поля магнитные моменты самопроизвольно ориентируются вдоль одной из осей легкого намагничивания. Например, элементарная ячейка железа представляет собой

объемно-центрированный куб. Направление легкого намагничивания совпадает с ребром куба [100], следовательно, в кристалле можно выделить шесть эквивалентных направлений легкого намагничивания. Направление пространственной диагонали куба [111] соответствует направлению трудного намагничивания для железа. У никеля, имеющего структуру гранецентрированного куба, диагональ [111], наоборот, является направлением легкого намагничивания; симметрия решетки никеля определяет восемь таких направлений. В то же время кобальт, обладающий гексагональной структурой, имеет лишь два направления легкого намагничивания, совпадающих с осью призмы, т.е. магнитные моменты доменов в отсутствии внешнего поля могут быть ориентированы лишь в двух противоположных направлениях [0001] .

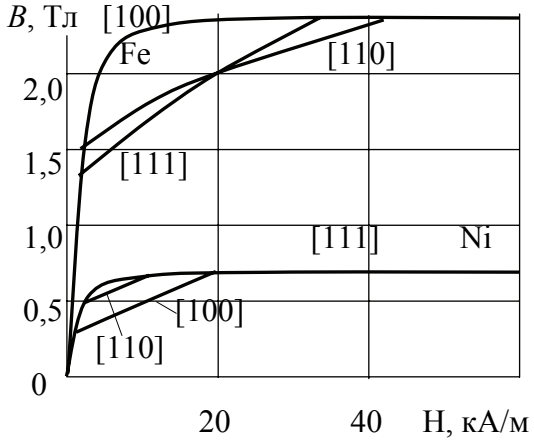


Рис. 5.4 – Зависимость энергии намагничивания от ориентации кристалла.

Кривые легкого, среднего и трудного намагничивания для железа и никеля приведены на рис. 5.4. Например, для железа дополнительная работа по намагничиванию образца до насыщения в направлении $[111]$ по сравнению с работой, требуемой для намагничивания образца в легком направлении, $[100]$ составляет примерно $1{,}4 \cdot 10^4$ Дж/м^3. Энергию, затрачиваемую внешним магнитным полем на поворот вектора намагниченности ферромагнитного кристалла из направления легкого намагничивания в направлении трудного намагничивания называют *энергией естественной магнитной кристаллографической анизотропии*. Эту энергию можно вычислить по площади фигуры, ограниченной кривыми намагничивания в направлении соответствующих кристаллографических осей.

Магнитная анизотропия влияет на доменную структуру, процессы намагничивания и другие свойства ферромагнетиков.

5.1.2.3 Доменная структура и намагничивание ферромагнетика

Экспериментально показано, что особые свойства ферромагнетиков обусловлены их доменным строением. Домены представляют собой микроскопические области, спонтанно намагниченные до насыщения даже в отсутствии внешнего магнитного поля. Спонтанная намагниченность доменов обусловлена параллельной ориентацией магнитных моментов атомов. Каждый домен обладает магнитным моментом, но ориентированы

эти моменты хаотично, поэтому суммарный магнитный момент ферромагнетика в отсутствие внешнего магнитного поля равен нулю.

Дробление ферромагнетиков на домены происходит потому, что этот процесс «энергетически выгоден». Если бы весь ферромагнетик был самопроизвольно намагничен в одном направлении, то получился бы минимум энергии обменного взаимодействия, но такому ферромагнетику соответствовала бы значительная энергия возбужденного им магнитного поля. При дроблении ферромагнетика на домены, ориентированные хаотично, магнитное поле, возбужденное ферромагнетиком, значительно ослабляется. Уменьшается и соответствующая ему энергия. Что же касается энергии обменного взаимодействия электронов, то благодаря короткодействующему характеру обменных сил эта энергия остается неизменной для всех электронов, за исключением электронов на границах доменов, где она возрастает из-за различной ориентации спинов электронов соседних доменов. Обменная энергия атомов, расположенных на границах доменов, может рассматриваться как поверхностная энергия, поскольку она пропорциональна полной площади поверхностей, вдоль которых граничат домены. По мере дробления доменов, поверхностная, а с ней и полная энергия обменного взаимодействия возрастает, а энергия магнитного поля убывает. Дробление доменов прекращается, когда сумма магнитной и обменной энергий достигает минимума. Условием этого минимума и определяется размер доменов.

Доменная структура ферромагнетиков доказана экспериментально. Прямым доказательством является получение так называемых порошковых фигур. Тщательно отполированная поверхность ферромагнетика покрывается тонким слоем жидкости, в которой взвешены мельчайшие частицы ферромагнитного порошка (например, Fe_2O_3). Частицы оседают преимущественно в местах максимальной неоднородности магнитного поля, т.е. на границах между доменами. Благодаря этому осевший порошок очерчивает границы доменов. Получившиеся «порошковые фигуры» хорошо видны в микроскоп при небольшом увеличении.

Другой, косвенный, метод основан на наблюдении эффекта Баркгаузена. Для его наблюдения ферромагнитный образец в виде длинного стержня или проволоки помещается в катушку, концы которой присоединены к осциллографу или громкоговорителю. При плавном увеличении тока через катушку в громкоговорителе слышен своеобразный шорох, а на осциллограмме появляются беспорядочные всплески. В этом и состоит эффект Баркгаузена, а причины его появления пояснены в разделе, посвященном диэлектрическому эффекту Баркгаузена. В ферромагнетиках происходит аналогичное скачкообразное движение стенок доменов через несовершенства кристалла.

Из осциллограмм эффекта Баркгаузена, а также с помощью порошковых фигур, оказалось возможным определить объем и линейные размеры

доменов. Длина доменов колеблется в диапазоне от единиц мкм до 2…3 мм, объем – от 10^{-18} до 10^{-9} м^3. Переходный слой, разделяющий домены, намагниченные в различных направлениях, называют «стенкой Блоха». В пределах такого слоя происходит постепенное изменение ориентации спинов, толщина «стенок Блоха» может достигать несколько сот межатомных расстояний (например, в железе она составляет около 100 нм).

Домены не следует смешивать с мельчайшими кристаллитами (зернами), из которых состоят поликристаллические ферромагнетики. В каждом из этих кристаллитов может содержаться много доменов.

Ферромагнетики, помимо способности сильно намагничиваться, обладают и другими свойствами, существенно отличающими их от диа- и парамагнетиков. Если для слабомагнитных веществ зависимость магнитной индукции $\vec{B}$ (намагниченности $\vec{J}$) от напряженности магнитного поля $\vec{H}$ линейна, то для ферромагнетиков эта зависимость является довольно сложной. Для ферромагнетиков по мере возрастания напряженности поля намагниченность сначала растет быстро, затем медленнее, и, наконец, достигается так называемое магнитное насыщение.

Сложный характер зависимости B от H обусловлен тем, что действие внешнего магнитного поля на домены на разных стадиях различно. На рис. 5.5 приведена доменная структура с призматическими замыкающими

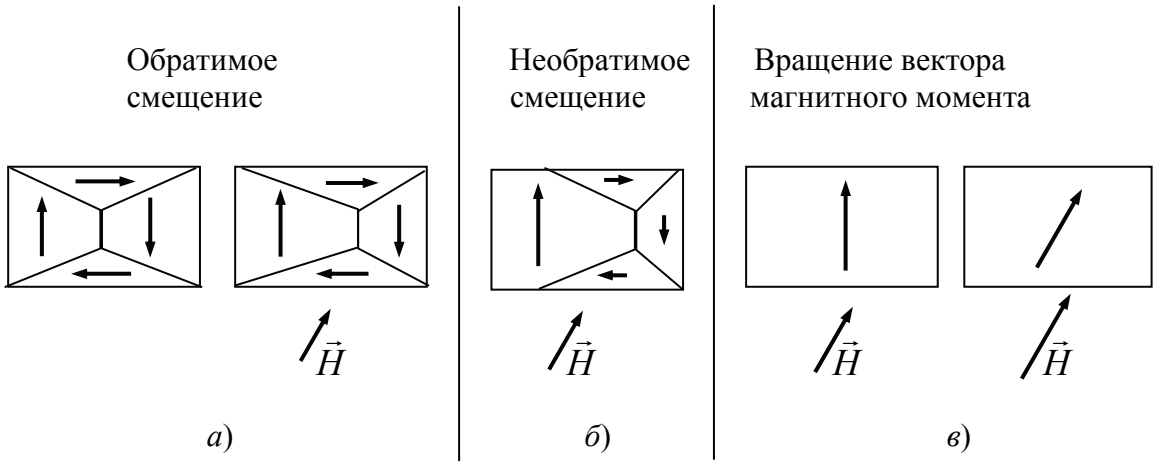


Рис. 5.5 – Смещение доменных границ с ростом напряженности магнитного поля $\vec{H}$: *а*) – процесс обратимого смещения доменных границ; *б*) – процесс необратимого смещения доменных границ; в) – процесс вращения вектора магнитного момента

доменами, характерными для ферромагнетиков с кубической кристаллической решеткой. В слабых полях домены, у которых магнитные моменты образуют меньший угол с внешним полем, растут за счет уменьшения доменов, у которых магнитные моменты образуют бóльший угол с внешним полем. Этот процесс обратимый, и называется *процессом обратимого смещения доменных границ* (рис. 5.5, *а*). Суммарная намагниченность образца становится отличной от нуля, что приводит к увеличению магнитной индукции B (область 1 на рис. 5.6). Увеличение

напряженности магнитного поля заставляют доменные стенки перемещаться на бóльшие расстояния. Наталкиваясь на дефекты кристалла, стенки некоторое время копят энергию, тратя ее на скачки Баркгаузена. Процессы, характерные для этой области полей, происходят необратимо, т. е. при снятии внешнего магнитного поля магнитное состояние уже не возвращается к исходному, поэтому его называют *процессом необратимого смещения доменных границ* (рис. 5.5, *б*). Перемагничивание приобретает необратимый характер – возникает остаточное намагничивание и гистерезис. На участке необратимого смещения доменных границ кривая намагничивания имеет наибольшую крутизну (область 2 на рис. 5.6). При дальнейшем росте напряженности магнитного поля границы доменов исчезают, и материал превращается в один домен, составляющий наименьший угол с направлением внешнего поля. Дальнейшее увеличение напряженности поля вызывает постепенный поворот вектора магнитного момента к направлению внешнего поля. Этот процесс называется *процессом вращения вектора магнитного момента* и может происходить как обратимо, так и необратимо (рис. 5.5, *в*). При таких полях ферромагнетик намагничивается до насыщения и магнитная индукция B растет линейно за счет роста B_0 (формула (5.2)), так как B' остается постоянной (область 3 на рис. 5.6).

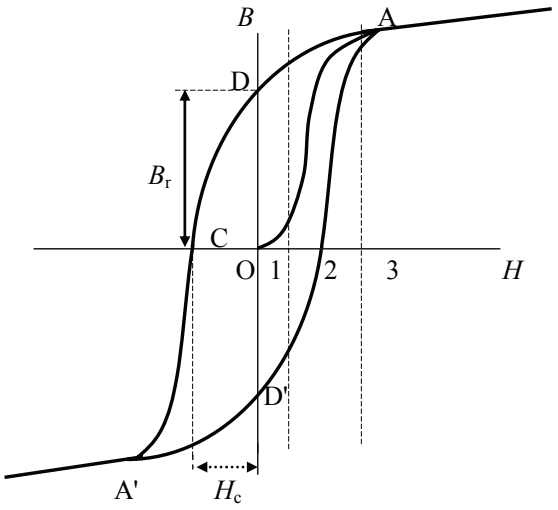


Рис. 5.6 – График кривой намагничивания ферромагнетика

Отставание изменения индукции от изменения напряженности магнитного поля называется *магнитным гистерезисом*. Рассмотрим подробнее это явление. Если ферромагнетик поместить во внешнее магнитное поле и изменять его напряженность, то индукция поля будет изменяться по сложной кривой OA (рис. 5.6). Эта кривая называется *начальной кривой намагничивания*. Если довести ферромагнетик до насыщения, а затем уменьшать напряженность, то индукция также будет уменьшаться, но уже по кривой AD. В результате, когда напряженность достигнет нулевого значения, индукция будет отлична от нуля и равна B_r. Это значение называется *остаточной индукцией*.

Чтобы размагнитить ферромагнетик, нужно создать магнитное поле противоположного направления. Напряженность H_c, при которой индукция становится равной нулю, называется *коэрцитивной силой*. При дальнейшем увеличении напряженности поля обратного направления ферромагнетик перемагничивается и при некотором значении H достигает насыщения (точка A'). Если затем уменьшать напряженность до нуля, то индукция будет

изменяться по кривой A'D'. При увеличении напряженности в прямом направлении, индукция изменяется по кривой D'A. Полученная замкнутая кривая называется *петлей гистерезиса*.

Если максимальные значения напряженности таковы, что намагничивание достигает насыщения, то петля называется *максимальной* (*предельной*) *петлей гистерезиса*. Если насыщение не достигнуто, то получаются частные петли гистерезиса, которые лежат внутри максимальной петли, а их вершины находятся на начальной кривой намагничивания OA. При дальнейшем возрастании напряженности внешнего магнитного поля площадь максимальной петли остается неизменной. Площадь петли гистерезиса численно равна удельной плотности энергии, затраченной на один цикл перемагничивания.

Разные ферромагнетики имеют разные гистерезисные петли. Ферромагнетики с малой (в пределах от нескольких тысячных до 100…200 А/м) коэрцитивной силой H_c (с узкой петлей гистерезиса) называются *мягкими*, с большой (от нескольких сотен до нескольких десятков тысяч ампер на метр) коэрцитивной силой (с широкой петлей гистерезиса) – *жесткими* (*твердыми*).

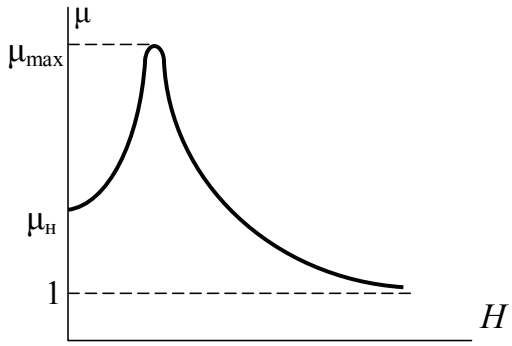


Рис. 5.7 – Зависимость магнитной проницаемости μ от напряженности магнитного поля H для ферромагнетиков

Существенная особенность ферромагнетиков – не только большие значения относительной магнитной проницаемости μ (например, для железа – 5000, для сплава супермаллоя – 800 000), но и сложная зависимость μ от H (рис. 5.7). Вначале относительная магнитная проницаемость μ растет с увеличением напряженности внешнего магнитного поля H и достигает максимума, затем начинает уменьшаться, стремясь в сильных полях к 1. Понять, почему в сильных полях μ стремится к 1, можно из анализа формулы $\mu = \dfrac{B}{\mu_0 H} = 1 + \dfrac{J}{H}$ (см. формулы (5.8) и (5.9)). При насыщении $J = J_{\text{нас}} = \text{const}$ и с ростом H отношение $\dfrac{J}{H} \to 0$, поэтому $\mu \to 1$.

Предельное значение магнитной проницаемости при напряженности магнитного поля, стремящейся к нулю, называют *начальной магнитной проницаемостью* $\mu_н$. Эта характеристика имеет важное значение для технического использовании магнитных материалов. Экспериментально ее определяют в слабых полях с напряженностью порядка 0,1 А/м.

Крутизна отдельных участков кривой намагничивания и ветвей петли гистерезиса характеризуется *дифференциальной магнитной проницаемостью*

$\mu_{диф} = \frac{1}{\mu_0}\frac{dB}{dH}$. Максимум дифференциальной магнитной проницаемости соответствует точке перегиба начальной кривой намагничивания.

Процесс намагничивания ферромагнетика сопровождается изменением его линейных размеров и объема. Это явление получило название *магнитострикции*. Величина и знак эффекта зависят от напряженности H намагничивающего поля, от природы ферромагнетика и ориентации его кристаллографических осей.

5.1.3. Магнитные потери

Перемагничивание ферромагнетиков в переменных полях сопровождается потерями энергии, вызывающими нагрев материала. В общем случае потери на перемагничивание складываются из потерь на гистерезис, на вихревые токи и на магнитное последействие. Вкладами последнего механизма в разогрев ферромагнетика обычно можно пренебречь. Удельные потери на гистерезис $w_Г$ (в Дж/м^3) за один цикл перемагничивания, отнесенные к единице объема вещества, определяются площадью статической петли гистерезиса, т.е. петли, полученной при медленном изменении магнитного потока:

$$w_Г = \oint_{ст} \vec{H} \times d\vec{B}. \qquad (5.12)$$

Мощность, обусловленная потерями на гистерезис:

$$P_Г = \oint_{ст} f \cdot \vec{H} \cdot d\vec{B}, \qquad (5.13)$$

где f – частота переменного магнитного поля.

Для практических целей наиболее важной характеристикой является активная мощность $P_Т$, выделяющаяся в ферромагнетике при его перемагничивании, т.е. энергия, расходуемая в единицу времени. Мощность, обусловленная потерями на вихревые токи, определяется эмпирической формулой следующего вида:

$$P_Т = w_Т f V = \xi f^2 B_{max}^2 V, \qquad (5.14)$$

где B_{max} – максимальная индукция, достигаемая в данном цикле, V – объем образца, ξ – коэффициент, пропорциональный удельной проводимости вещества и зависящий от формы и размера поперечного сечения намагничиваемого образца.

Так как величина $P_Т$ зависит от второй степени частоты, а величина $P_Г$ – от первой степени, при высоких частотах в первую очередь учитывают величину $P_Т$, т. е. потери на вихревые токи.

Потери на магнитное последействие обусловлены отставанием магнитной индукции от изменения напряженности магнитного поля. Это явление называют магнитной вязкостью.

В слабых полях и на высоких частотах динамическая петля гистерезиса вследствие отставания индукции от напряженности поля имеет форму

эллипса. Обозначив угол отставания через δ_μ, разложим магнитную индукцию на две составляющие:

$$B_{m1} = B_m \cos\delta_\mu \quad \text{и} \quad B_{m2} = B_m \sin\delta_\mu . \tag{5.15}$$

Первая составляющая совпадает по фазе с напряженностью магнитного поля, а вторая отстает от нее на угол $\pi/2$. Потери на перемагничивание обусловлены лишь составляющей B_{m2}. Физическую величину, определяемую соотношением

$$\mu' = \frac{B_{m1}}{\mu_0 H_m}, \tag{5.16}$$

называют упругой магнитной проницаемостью, а величину

$$\mu'' = \frac{B_{m2}}{\mu_0 H_m}, \tag{5.17}$$

называют вязкой магнитной проницаемостью.

Наиболее точно поведение ферромагнетиков в переменных полях описывает комплексная магнитная проницаемость:

$$\tilde{\mu} = \mu' - i\mu'' . \tag{5.18}$$

Угол δ_μ принято называть углом магнитных потерь. Из изложенного следует, что

$$tg\delta_\mu = \frac{\mu''}{\mu'} . \tag{5.19}$$

Тангенс угла магнитных потерь можно выразить через параметры эквивалентной схемы. Для этого индукционную катушку с сердечником из магнитного материала представляют в виде последовательной цепочки из индуктивности L и активного сопротивления r (см. рис. 5.8), отражающего все виды потерь на перемагничивание.

Собственной емкостью и сопротивлением обмотки обычно пренебрегают. Из векторной диаграммы (см. рис. 5.9) видно, что $U_a = I \cdot r$ и

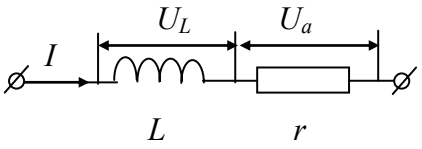


Рис. 5.8 – Эквивалентная схема потерь на перемагничивание (последовательное соединение)

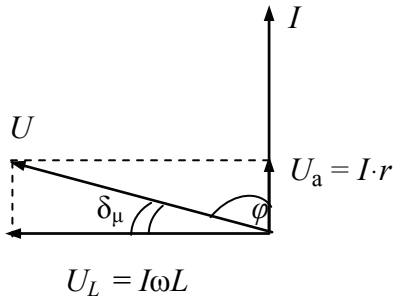


Рис. 5.9 – Векторная диаграмма для определения тангенса угла магнитных потерь

$U_L = I\omega L$, тогда $tg\delta$ можно найти по формуле:

$$tg\delta_{\mu} = \frac{U_{\text{а}}}{U_L} = \frac{r}{\omega L}. \quad (5.20)$$

Мощность, рассеиваемая в ферромагнетике

$$P_{\text{а}} = U_{\text{а}} \cdot I. \quad (5.21)$$

Если учесть, что $U_L = I\omega L$, то из формулы (5.20) можно найти значение активной составляющей напряжения

$$U_{\text{а}} = U_L \cdot tg\delta_{\mu} = I\omega L tg\delta_{\mu}. \quad (5.22)$$

Подставив формулу (5.22) в формулу (5.21), получим значение активной мощности

$$P_{\text{а}} = I^2 \omega L tg\delta_{\mu}. \quad (5.23)$$

Величину, обратную δ_{μ}, называют добротностью сердечника:

$$Q_{\mu} = \frac{1}{tg\delta_{\mu}}. \quad (5.24)$$

5.2. КЛАССИФИКАЦИЯ ФЕРРОМАГНЕТИКОВ

Применяемые в электронной технике магнитные материалы подразделяют на две основные группы: магнитомягкие и магнитотвердые. В отдельную группу выделяют материалы специального назначения.

К ***магнитомягким*** относят магнитные материалы с малой коэрцитивной силой. Они обладают способностью намагничиваться до насыщения в слабых магнитных полях, характеризуются узкой петлей гистерезиса и малыми потерями на перемагничивание. Магнитомягкие материалы используются в основном в качестве различных магнитопроводов: сердечников дросселей, трансформаторов, электромагнитов, магнитных систем измерительных приборов и т. п.

К ***магнитотвердым*** относят материалы с большой коэрцитивной силой H_c. Они перемагничиваются лишь в очень сильных магнитных полях и служат в основном для изготовления постоянных магнитов.

Условно магнитомягкими считают материалы, у которых $H_c < 800$ А/м, а магнитотвердыми – $H_c > 4000$ А/м. Необходимо, однако, отметить, что у лучших магнитомягких материалов коэрцитивная сила может составлять менее 1 А/м, а в лучших магнитотвердых материалах ее значение превышает 500 кА/м.

По масштабам применения в электронной технике среди ***материалов специального назначения*** следует выделить материалы с прямоугольной петлей гистерезиса (ППГ), ферриты для устройств сверхвысокочастотного диапазона и магнитострикционные материалы.

5.2.1. Магнитомягкие ферромагнетики

Помимо высокой магнитной проницаемости и малой коэрцитивной силой магнитомягкие материалы должны обладать большой индукцией насыщения, т. е. пропускать максимальный магнитный поток через заданную

площадь поперечного сечения магнитопровода. Выполнение этого требования позволяет уменьшить габаритные размеры и массу магнитной системы.

Магнитный материал, используемый в переменных полях, должен обладать возможно меньшими потерями на перемагничивание, которые складываются в основном из потерь на гистерезис и вихревые токи.

Для уменьшения потерь на вихревые токи в трансформаторах выбирают магнитомягкие материалы с повышенным удельным сопротивлением. Обычно магнитопроводы собирают из отдельных изолированных друг от друга тонких листов. Связано это с тем, что в массивных сплошных проводниках, помещенных в магнитное поле, возникают сильные индукционные токи. Эти токи оказываются замкнутыми в толще проводника и называются *вихревыми* (токи Фуко). В трансформаторах они приводят к разогреву и снижению КПД. Для уменьшения вихревых токов сердечник изготавливают из тонких пластин, покрытых изолятором так, чтобы изолирующие прослойки пересекали линии вихревых токов (рис. 5.10).

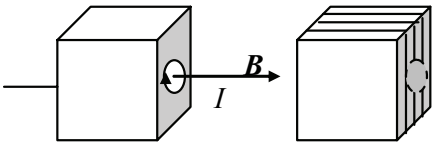


Рис. 5.10 – Использование пластинчатых образцов для уменьшения токов Фуко в магнитных материалах

Широкое применение получили ленточные сердечники, навиваемые из тонкой ленты с межвитковой изоляцией из диэлектрического лака.

К листовым и ленточным материалам предъявляется требование высокой пластичности, благодаря которой облегчается процесс изготовления изделий из них.

Важным требованием к магнитомягким материалам является обеспечение и стабильности их свойств, как во времени, так и по отношению к внешним воздействиям, таким как температура и механическое напряжение. Из всех магнитных характеристик наибольшим изменениям в процессе эксплуатации подвержена магнитная проницаемость и коэрцитивная сила.

Основным компонентом большинства магнитных материалов является ***железо***. Само по себе железо в элементарном виде представляет собой типичный магнитомягкий материал, магнитные свойства которого существенно зависят от содержания примесей. Среди элементарных ферромагнетиков железо обладает наибольшей индукцией насыщения (около 2,2 Тл).

Особо чистое железо, содержащее малое количество примесей (менее 0,05%) получают сложными технологическими способами. Примеси относительно слабо влияют на магнитные свойства железа, если их концентрации ниже предела растворимости. (Предел растворимости – это такая граничная концентрация примеси, при превышении которой примеси уже не растворяются). Низким пределом растворимости в железе обладает кислород, углерод, азот и сера, поэтому эти примеси и являются наиболее

вредными, так как даже незначительное количество этих примесей перестает растворяться в железе. Это приводит к тому, что при охлаждении железа после термообработки такие примеси из-за ограниченной растворимости выделяются в виде микровключений побочных фаз, которые затрудняют смещение доменных границ в слабом магнитном поле. Чистое железо редко применяют, так как магнитные свойства даже лучших разновидностей железа уступают свойствам магнитных материалов, полученных современными технологическими методами.

Кремнистая электротехническая сталь является основным магнитомягким материалом массового потребления. Добавление в состав этих сталей кремния приводит к повышению удельного сопротивления, что вызывает снижение потерь на вихревые токи. Кроме того, наличие кремния в составе сталей способствует выделению углерода в виде графита, а также удалению кислорода за счет образования окислов кремния, который в виде шлака удаляется из расплава. В результате легирование кремнием приводит к увеличению магнитной проницаемости μ, уменьшению коэрцитивной силы H_C и снижению потерь на гистерезис.

Сталь выпускается в виде рулонов, листов и лент. Листы тонкого проката предназначены в основном для использования в электромагнитных полях повышенной частоты (до 1 кГц). С уменьшением толщины листов уменьшаются потери на вихревые токи. Однако в очень тонких листах наблюдается резкое возрастание коэрцитивной силы, что приводит к резкому увеличению потерь на гистерезис.

В промышленности широко используются магнитные низкокоэрцитивные сплавы, так называемые *пермаллои*. Это железоникелевые сплавы, обладающие большой магнитной проницаемостью в области слабых полей и очень малой коэрцитивной силой. Пермаллои подразделяют на высоконикелевые (содержат 72…80% никеля) и низконикелевые – 40…50 % никеля.

Наибольшая магнитная проницаемость (более 10^5) получается у сплава, содержащего 78,5 % никеля. Очень легкое намагничивание этого сплава в слабых полях объясняется практическим отсутствием у него магнитной анизотропии и явления магнитострикции. Вследствие слабой анизотропии облегчается поворот магнитных моментов в направлении поля, а из-за отсутствия магнитострикции при намагничивании не возникает механических напряжений, затрудняющих смещение доменных границ под действием слабого поля.

Однако нужно заметить, что магнитная проницаемость пермаллоев сильно снижается с увеличением частоты. Это объясняется возникновением в материале заметных вихревых токов из-за небольшого удельного сопротивления. А так как удельное сопротивление высоконикелевых пермаллоев почти в три раза меньше, чем у низконикелевых, поэтому при повышенных частотах предпочтительнее использовать низконикелевые пермаллои. Кроме того, остаточная индукция высоконикелевых пермаллоев почти в два раза ниже, чем у электротехнической стали. Для придания

сплавам необходимых свойств в состав пермаллоев вводят ряд добавок (молибден, хром, кремний, марганец, медь для повышения удельного сопротивления и уменьшения чувствительности к механическим деформациям).

Ферриты. Как отмечалось ранее (см. раздел 5.1), ферриты представляют собой оксидные магнитные материалы, у которых спонтанная намагниченность доменов обусловлена нескомпенсированным антиферромагнетизмом.

Большое удельное сопротивление ρ, превышающее удельное сопротивление железа в $10^3 \ldots 10^{13}$ раз, следовательно, и относительно небольшие потери энергии в области повышенных и высоких частот наряду с хорошими магнитными свойствами обеспечивают ферритам широкое применение в радиотехнике.

Ферриты получают в виде керамики и монокристаллов. Благодаря невысокой стоимости и относительной простоте технологического цикла керамические материалы занимают ведущее место среди высокочастотных магнетиков.

В качестве магнитомягких материалов наиболее широко применяют никель-цинковые и марганец-цинковые ферриты. Они представляют собой твердые растворы замещения, образованные двумя простыми ферритами, один из которых ($NiFe_2O_4$ или $MnFe_2O_4$) является ферромагнетиком, а другой ($ZnFe_2O_4$) – немагнитен. В никель-цинковых ферритах возникает один очень интересный эффект: увеличение концентрации немагнитного компонента приводит к увеличению индукции насыщения и начальной магнитной проницаемости $\mu_н$, однако это же приводит к монотонному снижению температуры Кюри. Аналогичные закономерности наблюдаются и для марганец-цинковых ферритов. Значение начальной магнитной проницаемости $\mu_н$ и коэрцитивной силы H_c определяется не только составом материала, но и его структурой. Препятствиями, мешающими свободному перемещению доменных границ при воздействии на феррит слабого магнитного поля, являются макроскопические поры, включения побочных фаз, участки с дефектной кристаллической решеткой и др. Устранение этих структурных барьеров позволяет существенно повысить магнитную проницаемость материала. Большое влияние на значение начальной магнитной проницаемости ферритов оказывает размер кристаллических зерен. Марганец-цинковые ферриты с крупнозернистой структурой (со средним размером кристаллитов порядка 40 мкм) могут обладать начальной магнитной проницаемостью до 20000, что близко к начальной магнитной проницаемости лучших марок пермаллоя.

Рассмотрим некоторые магнитные свойства ферритов. Для ферритов, используемых в переменных полях, кроме начальной магнитной проницаемости одной из важнейших характеристик является тангенс угла потерь $tg\delta$. Благодаря низкой проводимости составляющая потерь на вихревые токи в ферритах практически мала, и ею можно пренебречь. В слабых полях оказываются незначительными и потери на гистерезис.

Поэтому изменение значения $tg\delta$ в ферритах на высоких частотах обусловлено другими механизмами (релаксационными и резонансными явлениями). Для оценки частотного диапазона, в котором может использоваться данный материал, вводят понятие критической частоты $f_{кр}$. Обычно под $f_{кр}$ понимают такую частоту, при которой $tg\delta$ достигает значения 0,1.

Инерционность смещения доменных границ, проявляющаяся на высоких частотах, приводит не только к росту магнитных потерь, но и к снижению магнитной проницаемости ферритов. Для сравнительной оценки качества магнитомягких ферритов используется относительный тангенс угла потерь $\frac{tg\delta}{\mu_н}$.

Из экспериментальных исследований температурной зависимости начальной магнитной проницаемости $\mu_н$ следует, что она повышается с ростом температуры до точки Кюри и затем резко падает. При этом, чем выше начальная проницаемость, тем ниже точка Кюри феррита.

Магнитные свойства ферритов зависят от механических напряжений, которые могут возникать при нанесении обмотки, крепления изделий и т.д. Чтобы не было ухудшения магнитных характеристик, ферриты следует оберегать от механических нагрузок.

По электрическим свойствам ферриты относятся к классу полупроводников или даже диэлектриков. Их проводимость обусловлена процессами электронного обмена между ионами переменной валентности. Концентрация электронов, участвующих в обмене практически не зависит от температуры. Но при повышении температуры экспоненциально увеличивается вероятность перескока между ионами переменной валентности, т. е. возрастает подвижность носителей заряда, что приводит к увеличению удельной проводимости, а значит к уменьшению удельного сопротивления. Изменение удельной проводимости и удельного сопротивления с достаточной точностью можно описать следующими формулами:

$$\sigma = \sigma_0 e^{-\frac{E_0}{kT}}, \qquad (5.25)$$

$$\rho = \rho_0 e^{\frac{E_0}{kT}}, \qquad (5.26)$$

где σ_0 и ρ_0 – постоянные величины для данного материала, E_0 – энергия активации проводимости.

Экспериментально установлено, что присутствие в ферритах ионов двухвалентного железа приводит к ослаблению анизотропии и магнитострикции; это увеличивает значение начальной магнитной проницаемости. Отсюда вытекает следующая закономерность: ферриты с высокой магнитной проницаемостью, как правило, обладают невысоким удельным сопротивлением.

Магнитомягкие ферриты с начальной магнитной проницаемостью 400…20000 в слабых полях во многих случаях выгодно заменяют листовые ферромагнитные материалы: пермаллой и электротехническую сталь. В средних и сильных магнитных полях замена листовых ферромагнетиков ферритами нецелесообразна, так как у ферритов меньше индукция насыщения.

Магнитомягкие ферриты широко применяются в качестве сердечников контурных катушек постоянной и переменной индуктивностей, фильтров в аппаратуре радио- и проводной связи, магнитных модуляторов и усилителей и т. д. Из них изготавливают стержневые магнитные антенны, индуктивные линии задержки и другие узлы электронной аппаратуры.

5.2.2. Магнитотвердые ферромагнетики

Магнитотвердые материалы отличаются от магнитомягких высокой коэрцитивной силой. Площадь петли гистерезиса у магнитотвердых материалов значительно больше, чем у магнитомягких.

По применению магнитотвердые материалы можно подразделить на материалы для постоянных магнитов и материалы для записи и длительного хранения звука, изображения и т.д.

Магнитные цепи с постоянными магнитами должны быть разомкнутыми, т. е. иметь рабочий воздушный зазор. При наличии зазора за счет свободных полюсов создается внутреннее размагничивающее поле H_d, которое уменьшает индукцию внутри магнита до значения B_d. Положение рабочей точки, характеризующей состояние магнитного материала, зависит от величины зазора.

При отсутствии внешнего магнитного поля макроскопические токи отсутствуют. В соответствии с законом полного тока

$$\oint Hdl = 0. \qquad (5.27)$$

Это справедливо для любого контура интегрирования, в частности, вдоль пути по всей оси магнита. Отсюда следует, что

$$H_d l_d - H_0 l_0 = 0, \qquad (5.28)$$

где l_d и l_0 – длины магнита и воздушного зазора соответственно, H_0 – напряженность магнитного поля в зазоре (см. рис. 5.11).

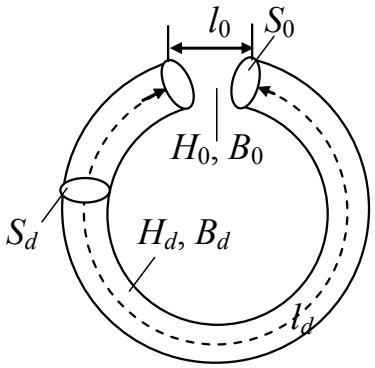


Рис. 5.11 – Тороидальный магнит с зазором l_0

Важнейшее требование к постоянному магниту состоит в том, чтобы получить *максимальную магнитную энергию* W_0 в рабочем зазоре:

$$W_0 = \frac{B_0 H_0}{2} V_0, \qquad (5.29)$$

где V_0 – объем зазора; $V_0 \approx l_0 S_0$.

Пренебрегая потоком рассеяния и учитывая, что магнитная индукция непрерывна, запишем:

$$B_d l_d = B_0 l_0 . \tag{5.30}$$

Из выражений (5.28) – (5.30) следует:

$$W_0 = \frac{B_d H_d}{2} V_d , \tag{5.31}$$

где $V_d \approx l_d S_d$ – объем магнита.

Формула (5.31) показывает, что удельная (т.е. отнесенная к единице объема магнита) магнитная энергия в воздушном зазоре определяется положением рабочей точки на кривой размагничивания:

$$W_d = \frac{W_0}{V_d} = \frac{B_d H_d}{2} . \tag{5.32}$$

Чем меньше длина магнита и относительно больше зазор, тем больше размагничивающее поле полюсов и меньше B_d.

На рис. 5.12 представлена кривая, выражающая зависимость удельной энергии W_d от значения индукции B_d.

При замкнутом магните $B_d = B_r$, причем энергия равна нулю, так как $H_d = 0$. Если зазор между полюсами очень велик, то энергия при этих условиях также стремится к нулю, так как $B_d = 0$, $H = H_c$.

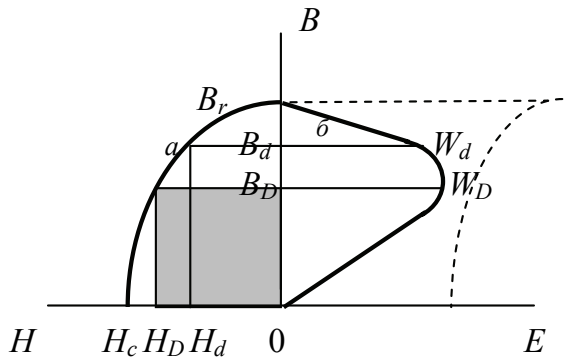


Рис. 5.12 – Кривые размагничивания (*а*) и магнитной энергии в воздушном зазоре (*б*)

Из рис. 5.9 видно, что при некоторых значениях B_D и H_D энергия достигает максимального значения.

Значение

$$W_D = \frac{B_D H_D}{2} = E_{d\max} \tag{5.33}$$

определяет наилучшее использование магнита и тем самым является наиболее важной характеристикой качества материалов, используемых для изготовления постоянных магнитов. Нередко для характеристики таких материалов пользуются произведением $B_D H_D$, опуская множитель 1/2.

Графически энергию W_d в определенном масштабе можно представить площадью прямоугольника со сторонами B_d и H_d (на рис. 5.8 выделен серым цветом прямоугольник, характеризующий $B_{d\max}$).

Форму кривой размагничивания принято характеризовать *коэффициентом выпуклости* η_B, под которым понимают отношение

$$\eta_B = \frac{(BH)_{\max}}{B_r H_c} . \tag{5.34}$$

С увеличением прямоугольности петли гистерезиса коэффициент выпуклости приближается к единице. Чем больше остаточная индукция B_r,

коэрцитивная сила H_c и коэффициент выпуклости η_B, тем больше максимальная энергия магнита.

Чтобы получить высокую коэрцитивную силу в магнитном материале, необходимо затруднить процесс перемагничивания. Это можно осуществить двумя путями: либо воспрепятствовать смещению доменных границ, либо вообще исключить эти границы, уменьшая размер кристаллических зерен. Большая коэрцитивная сила возникает в материале, состоящем из однодоменных частиц, у которых велика анизотропия формы.

Литые высококоэрцитивные сплавы Fe-Ni-Al и Fe-Ni-Co-Al, модифицированные различными добавками, являются активными элементами многих приборов. Высококоэрцитивное состояние этих сплавов обусловлено их дисперсным распадом на две фазы при охлаждении до определенной температуры. В системе Fe-Ni-Al одна из фаз близка по составу к чистому железу и является сильномагнитной; другая фаза состоит из никеля и алюминия и оказывается слабомагнитной. Материалы, имеющие такую структуру, намагничиваются в основном за счет процессов вращения магнитных моментов доменов. Однако без легирующих добавок эти сплавы не применяют. Добавки (например, медь, титан, ниобий) не только улучшают магнитные свойства, но и обеспечивают лучшую воспроизводимость свойств.

Недостатком сплавов подобного вида является трудность изготовления из них изделий точных размеров вследствие хрупкости и высокой твердости. Поэтому для производства постоянных магнитов чаще всего используют магнитные порошки, скрепленные тем или иным связующим веществом (металлокерамические магниты, металлопластические магниты и др.).

5.2.3. Ферромагнетики для хранения информации

Магнитные (магнитомягкие) материалы с прямоугольной петлей гистерезиса (ППГ) находят широкое применение в устройствах автоматики и вычислительной техники. Сердечники из материала с ППГ имеют два устойчивых магнитных состояния, соответствующие различным направлениям магнитной индукции. Именно благодаря этой особенности их можно использовать в качестве элементов для хранения и переработки двоичной информации. Запись и считывание информации осуществляются переключением сердечника из одного магнитного состояния в другое с помощью импульсов тока, создающих требуемую напряженность магнитного поля.

К материалам этого типа предъявляется ряд требований, а для их характеристики привлекают некоторые дополнительные параметры. Одним из таких параметров является *коэффициент прямоугольности* петли гистерезиса $K_{\text{ПУ}}$, представляющий собой отношение остаточной индукции B_r к максимальной индукции $B_{\max}$:

$$K_{\text{ПУ}} = \frac{B_r}{B_{\max}}. \qquad (5.35)$$

Для определенности B_{max} измеряют при $H_{max} = 5H_c$. Желательно, чтобы $K_{ПУ}$ был возможно ближе к единице. Для обеспечения быстрого перемагничивания сердечников они должны иметь небольшой коэффициент переключения S_q, численно равный количеству электричества на единицу толщины сердечника, необходимого для перемагничивания его из одного состояния остаточной индукции в другое.

Кроме того, материалы с ППГ должны обеспечивать малое время перемагничивания, большую температурную стабильность магнитных характеристик, следовательно, иметь высокую температуру Кюри.

Применяемые на практике кольцевые сердечники с внешним диаметром 0,5 мм позволяют достигать частоты циклов переключения порядка нескольких мегагерц при условии отсутствия ограничений, обусловленных монтажом. Если частота циклов должна быть порядка 10МГц и выше, то в качестве запоминающих элементов следует использовать тонкие пленки, обладающие направлением легчайшего намагничивания. В этом случае можно достичь того, чтобы все векторы намагниченности в пленке одновременно меняли свое направление на противоположное (когерентный поворот) – тогда перемагничивание происходит за чрезвычайно короткий промежуток времени. Поры, имеющиеся как в ферритах, так и в магнитных пленках, тормозят перемагничивание.

5.3. ПРИМЕНЕНИЕ ФЕРРОМАГНЕТИКОВ

5.3.1. Трансформаторы

Одним из широко применяемых на практике технических устройств, использующих вихревое электрическое поле, является *трансформатор*. Он состоит из замкнутого сердечника из магнитомягкого ферромагнетика, на котором расположены две обмотки – первичная (вход) и вторичная (выход). Концы первичной обмотки обычно подключаются к сети переменного тока, вторичной – к потребителю (рис. 5.13).

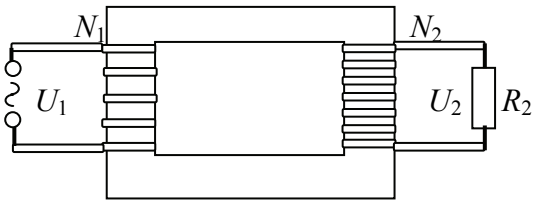


Рис. 5.13 – Вид трансформатора с первичной и вторичной обмотками

Переменный ток, протекающий в первичной обмотке, создает в сердечнике изменяющийся магнитный поток, который возбуждает во вторичной обмотке ЭДС самоиндукции. Согласно закону Ома, напряжение на первичной обмотке с числом витков N_1 равно

$$U_1 = r_1 I_1 - \varepsilon_1 = r_1 I_1 + N_1 d\Phi / dt\,, \qquad (5.36)$$

где $\varepsilon_1 = -N_1 d\Phi / dt$ – ЭДС самоиндукции, а I_1 , r_1 – ток в обмотке и ее сопротивление. Соответственно, напряжение на вторичной обмотке

$$U_2 = r_2 I_2 - \varepsilon_2 = r_2 I_2 + N_2 d\Phi / dt\,. \qquad (5.37)$$

При большом сопротивлении потребляющей цепи, подключенной ко второй обмотке, токи малы ($rI<<\varepsilon$), тогда $U_1 = -\varepsilon_1 = N_1 d\Phi / dt$ и $U_2 = -\varepsilon_2 = N_2 d\Phi / dt$, из этого следует, что

$$\frac{U_2}{U_1} = \frac{N_2}{N_1} = k, \qquad (5.38)$$

где k – коэффициент трансформации. В зависимости от соотношения числа витков в обмотках трансформатор может либо повышать напряжение, либо понижать.

Технические проблемы, подлежащие разрешению при разработке высококачественных трансформаторных материалов, сводятся к получению высокой магнитной проницаемости и малых энергетических потерь. В процессе разработки магнитомягких материалов выделились три основных типа таких материалов:

1) сплавы Fe-Si, используемые в силовой низкочастотной электротехнике;

2) сплавы Fe-Ni для высококачественной маломощной аппаратуры (например, для трансформаторов, используемых в электроакустике);

3) ферриты для использования в высокочастотных (частоты порядка мегагерц) устройствах.

Обратим особое внимание на требования к магнетикам, применяемым для изготовления трансформаторов и дросселей различного назначения. Первое требование, а именно, большая магнитная проницаемость, важно для расчета обмоток трансформатора. Высокая проницаемость позволяет достигать насыщения в слабых полях, поэтому первичная обмотка и ток в ней могут быть невелики. Кроме того, при высокой проницаемости петля гистерезиса обычно узка, поэтому малы прямые магнитные потери. Первое требование – это снижение общих потерь. Их можно снизить, уменьшив потери на вихревые токи за счет добавления в магнитный материал примесей, повышающих электросопротивление.

Добавление к чистому железу кремния улучшает его свойства двояким способом: увеличивает сопротивление, что снижает потери на вихревые токи, и уменьшает потери на гистерезис. Железо, содержащее 4% кремния, подвергают специальной термической обработке для того, чтобы кристаллиты поликристаллического листа были ориентированы примерно в одном направлении и лист в целом был близок к монокристаллическому образцу. Домены в таком листе получаются более крупными и имеют правильную форму. Такие отличия от обычного сплава Fe-Si резко снижают потери на гистерезис. Повышение содержания кремния снижает также и потери на вихревые токи. Оба эти эффекта можно было бы усилить повышением содержания кремния, но при содержании кремния, превышающем 6%, сплавы Fe-Si настолько тверды, что почти не поддаются обработке.

Весьма ценными свойствами для изготовления трансформаторов обладают широко применяемые сплавы Fe-Ni, называемые *пермаллоями*. Обычно широко применяются два сплава, содержащих 50% никеля и 78% никеля. Одна из причин отличного качества пермаллоя-78 заключается в том, что у него константа магнитострикции почти равна нулю, поэтому близка к нулю и магнитострикционная энергия. Сплав *супермаллой* – это в основном тот же сплав, только 5% железа в нем заменена молибденом. Этот материал имеет малые потери, как на гистерезис, так и на вихревые токи, так как добавка молибдена резко увеличивает его сопротивление.

Вследствие различия свойств низконикелевые и высоконикелевые пермаллои имеют несколько различные применения. Низконикелевые сплавы используют для изготовления сердечников малогабаритных силовых трансформаторов, дросселей, реле, сердечников импульсных трансформаторов и аппаратуры связи звуковых и высоких частот. Высоконикелевые сплавы используют для изготовления сердечников малогабаритных трансформаторов, реле и магнитных экранов, магнитных усилителей и бесконтактных реле.

Марганец-цинковый феррит, как и другие ферриты, хороший изолятор. Поэтому потери на вихревые токи ничтожно малы, даже если сердечник трансформатора не набран из листов. Ферриты особенно широко применяются на сверхвысоких частотах, когда металлы становятся практически бесполезными.

5.3.2. Элементы памяти

До пятидесятых годов магнитные материалы использовались главным образом в электрических машинах и трансформаторах. Совершенствование технологии открыло новые возможности их применения. Самым значительным из них явилось использование магнетизма для хранения информации. Винчестерный диск позволяет хранить информацию с плотностью более 3 миллионов бит/см^2. Физический принцип записи на винчестерный диск основан на намагничивании или размагничивании необходимых участков диска. Сохранению записанной информации способствует выпуклая, близкая к прямоугольной форме, кривая размагничивания. Двоичные элементы на магнитных сердечниках с ППГ характеризуются высокой надежностью, малыми габаритами, относительной стабильностью характеристик, низкой стоимостью. Они обладают практически неограниченным сроком службы, сохраняют записанную информацию при отключенных источниках питания. Магнитная память незаменима при хранении больших массивов информации, если не столь важно время обращения к памяти.

Магнитные ленты для видео- и звукозаписи также используются для хранения информации. Эти носители информации сделаны на основе магнитотвердых материалов. Обычно применяют тонкие металлические ленты из нержавеющих сплавов и ленты на пластмассовой основе с

порошковым рабочим слоем. Лента с записанной на ней информацией тем меньше подвержена саморазмагничиванию, чем выше коэрцитивная сила материала. Однако для облегчения записи желательно иметь небольшую коэрцитивную силу. Противоречивым требованиям наилучшим образом удовлетворяют значения H_c, лежащие в интервале 20-50 кА/м. В дополнение к этому материал для магнитной записи должен обладать возможно более высокой стабильностью магнитных параметров при изменениях температуры.

5.3.3. Ферриты для устройств СВЧ

Диапазон сверхвысоких частот (СВЧ) соответствует длинам волн от 1 м до 1 мм. В аппаратуре и приборах, где используются электромагнитные волны диапазона СВЧ, необходимо управлять этими колебаниями: переключать поток энергии с одного направления на другое, изменять фазу колебаний, поворачивать плоскость поляризации волны, частично или полностью поглощать мощность потока.

Электромагнитные волны могут распространяться в пространстве, заполненном диэлектриком, а от металлов они почти всегда отражаются. Поэтому металлические поверхности используют для направления волн, их концентрации и рассеяния. Электромагнитная энергия СВЧ чаще всего передается по волноводам, представляющими собой полые или частично заполненные твердыми материалами металлические трубы. В качестве твердых материалов для управления потоком энергии в волноводах используют ферриты СВЧ и некоторые немагнитные активные диэлектрики. Магнитными характеристиками первых можно управлять с помощью магнитного поля, электрическими свойствами вторых – за счет внешнего электрического поля.

ВОПРОСЫ

1. Как классифицируют вещества по магнитным свойствам?
2. Чем отличается обменное взаимодействие в ферро- и антиферромагнетиках?
3. Какие процессы происходят в ферромагнетике при его намагничивании внешним полем? Что называют основной кривой намагничивания магнитного материала? Что такое гистерезис?
4. Как изменяется статическая магнитная проницаемость ферромагнетиков от напряженности внешнего магнитного поля?
5. В чем заключается явление магнитострикции?
6. Каковы причины появления магнитных потерь при циклическом перемагничивании ферромагнетиков? Назовите способы уменьшения магнитных потерь.
7. Какие материалы называются ферритами?
8. Как классифицируют магнитные материалы по свойствам и техническому назначению?
9. Назовите важнейшие свойства магнитотвердых материалов.

10. Какие физические принципы положены в основу работы трансформаторов и дросселей? Какие магнитные материалы используются для этих целей?
11. Какие физические принципы положены в основу магнитной записи воспроизведения информации? Какие магнитные материалы используются для этих целей?

ЗАДАЧИ

1. В однородное магнитное поле индукцией B_0 перпендикулярно магнитному потоку помещена плоскопараллельная пластина из однородного изотропного ферромагнетика с магнитной проницаемостью μ. Определить магнитную индукцию B_1 и напряженность магнитного поля H_1 внутри ферромагнетика. (Ответ: $B_1 = B_0$; $H_1 = B_0/(\mu_0\,\mu)$.
2. Магнитная восприимчивость никеля при температурах 400 и 800 ºС равна соответственно $1{,}25\cdot10^{-3}$ и $1{,}14\cdot10^{-4}$. Определить температуру Кюри и магнитную восприимчивость при T = 600 К. (Ответ: $T_К$ = 359,9 ºС; $\chi_{(600\,^\circ C)} = 2{,}08\cdot10^{-4}$).
3. В сердечнике трансформатора суммарные удельные магнитные потери на гистерезис и вихревые токи при частотах 1 и 2 кГц составляют соответственно 2 и 6 Вт/кг (при неизменной максимальной индукции в сердечнике). Рассчитать магнитные потери на вихревые токи в сердечнике на частоте 2 кГц. (Ответ: $P_Т$ = 4 Вт/кг).
4. Найти удельные магнитные потери в ферритовом сердечнике марки 2000 НН, перемагничивающегося на частоте 0,1 МГц магнитным полем напряженностью H_m = 4 А/м, если в данных условиях $tg\delta$ = 0,2, магнитная проницаемость μ = 2500. (Ответ: $p_м = 1000$ Вт/м^3).
5. Найти индуктивность соленоида, имеющего N = 200 витков, намотанных на диэлектрическое основание, длиной l = 50 мм. Площадь поперечного сечения S = 50 мм^2. Как изменится индуктивность катушки, если в нее введен цилиндрический ферритовый сердечник, имеющий магнитную проницаемость μ = 400, определенную с учетом размагничивающего действия воздушного зазора? (Ответ: L_0 = 50,2 мкГн; L_1 = 20мГн).
6. Определить магнитную индукцию ферромагнитного сердечника, помещенного внутрь соленоида длиной l = 20 см с числом витков N = 800, если по обмотке проходит ток I = 0,2 А, а эффективная магнитная проницаемость сердечника μ = 200. (Ответ: B = 0,28 Тл).

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Акцепторные уровни 104
Акцепторы 104
Антиферромагнетизм 281

БКШ теория 68
Больцмана статистика 97
Барьер потенциальный 62
Бозе-Эйнштейна распределение 39

Валентная зона 35
Видемана-Франца закон 45
Вольт-амперная характеристика 158
Время жизни носителей 129
Вихревые токи 285
Волоконная связь 262

Германий 134
Гетеропереход 163
Гистерезис диэлектрический 203
– магнитный 280

Дебая температура 55
Дефекты решетки 27
– точечные 28
– линейные 30
– поверхностные 31
– сложные 32
Джозефсона эффект 74
Дислокации 31
Дисперсия 234
Диод 158
– силовой 158
– туннельный 159
– обращенный 162
Диффузионная длина 156
Длина свободного пробега 45, 54, 100
Доноры 103
Дырки 105
Диэлектрические материалы 196
Диэлектрическая проницаемость 203, 218
Домены 278

Емкость *p-n*-перехода 158
– барьерная 158

Зона валентная 35
– проводимости 35
– запрещенная 35

Излучательная рекомбинация 130
Излучение лазерное 185
Индексы Миллера 18
Инжекция носителей 155
Интегральная микросхема 175

Катодолюминесценция 131
Керра эффект 239
Кремний 132
Контакт металл-полупроводник 168
Контактная разность потенциалов 150
Концентрация носителей заряда 101,110
Коэрцитивная сила 280
Кубическая система 13, 18
Куперовские пары 69
Кюри температура 60, 205

Лазер полупроводниковый 184
– рубиновый 268
– на гетеропереходе 186
Лоренца сила 119
– число 50
Люминесценция 131

Магнитная восприимчивость 277
Магнитная проницаемость 277
Магнитострикция 281
Масса эффективная 36
МДП-структура 173
МДП-транзистор 173
Межзонные переходы 176
Мелкие уровни 109
МОП-транзистор 173

Нееля температура 281
Непрямой переход 126
Несимметричный *p-n*-переход 152

Область объемного заряда 151
Обратное напряжение 154
Обращенный диод 162
Одноэлектронное приближение 33
Ома закон 45
Отражение света 234
– полное внутреннее 234

Парамагнетизм 277
Паули принцип 33

Пельтье эффект 64
Переход *p-n* 149
Пермаллой 294
Пироэлектричество 212
Плотность состояний 98
– в валентной зоне 98
– в зоне проводимости 98
Поглощение света 124
– примесными центрами 127
– свободными носителями 127
– собственное 125
Подвижность 97, 103, 112
– зависимость от температуры 100
– от электрического поля 127
Показатель преломления 16, 233
Поккельса эффект 265
Полупроводники 95
– вырожденные 151
– примесные 95
– собственные 95, 96
– *n* типа 103
– *p*-типа 104
Поляризация света 236
Прибор с зарядовой связью 174
Пробой диэлектриков 227
– тепловой 227
– электрический 229
– электрохимический 232
Прямой переход 126
Пьезоэлектричество 208

Работа выхода 60
Рекомбинация носителей 128
– безызлучательная 130
– излучательная 130
– прямая межзонная 130

Сверхпроводники 68
– второго рода 72
– первого рода 71
– элементы памяти 84
– энергетическая щель 89
Световоды 259
Связь химическая 21
– Ван-дер-Ваальса 24
– ионная 22
– ковалентная 21
– металлическая 23
Сегнетоэлектрики 202
Смещение *p-n*-перехода 153
– обратное 153
– прямое 153
Симметричный *p-n*-переход 152
Собственный полупроводник 95
Сродство электронное 163
Супермаллой 294
Селен 135

Таммовские состояния 30
Тензочувствительность122
Теплоемкость 46
Теплопроводность 45
Термо-ЭДС 62,121
– коэффициент 121
Термопара 89
Транзистор 173
– полевой 173
Туннелирование 118
Туннельный диод 160

Удельное сопротивление металлов 54
– сплавов 57
Уровень Ферми 48, 98, 103, 104
Уровни энергетические
– глубокие 109
– мелкие 109

Ферми-Дирака распределение 39
Ферримагнетизм 281
Ферриты 281, 287
Ферромагнетики 277
Фононы 128
Фотолюминесценция 132
Фотопроводимость 122
– примесная 123
– собственная 123
Фотоэффект 122
Френкеля вакансии 28

Холла постоянная 120
– эффект 119

Шоттки вакансии 28
– эффект 62

Экситон 126
Электрическая проводимость 112
– в сильных электрических полях 118
Электреты 210
Электрострикция 210
Энергия ионизации примесей 108, 110
Эффективная масса 36

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Пасынков В.В. Материалы электронной техники/ В.В. Пасынков, В.С. Сорокин. – М.: Высшая школа, 1986. – 367 с.
2. Калинин Н.Н. Электрорадиоматериалы/ Калинин Н.Н., Скибинский Г.Л., Новиков П.П. – М.: Высшая школа, 1981. – 165 с.
3. Солимар Л. Лекции по электрическим свойствам материалов/ Л. Солимар, Д. Уолш. – М: Мир, 1991. – 501 с.
4. Детлаф А.А. Курс физики/ А.А. Детлаф, Б.М. Яворский. – М: Высшая школа, 1989. – 607 с.
5. Гусев В.А. Основы твердотельной электроники/ Гусев В.А. – Севастополь: СевНТУ, 2004. – 639 с.
6. Шредер Г. Техническая оптика / Г. Шредер, Х. Трайбер. – М.: Техносфера, 2006. – 423 с.
7. Вербицький В.Г. Розробка високоефективних мікро-, нанотехнологій оптоелектроніки і комунікаційних систем на їх основі/ [В.Г. Вербицький, І.М. Вікулін, П.П. Воробієнко та ін.]. – К.: Логос, 2009. – 262 с.

ЦИТИРУЕМАЯ ЛИТЕРАТУРА

1. Богородицкий Н.П. Электротехнические материалы/ Богородицкий Н.П., Пасынков В.В., Тареев Б.Н. – Л.: Энергоатомиздат, 1985. – 304 с.
2. Ашкрофт Н. Физика твердого тела. – Т.1/ Н. Ашкрофт, Н. Мермин. – М: Мир, 1979. – 399 с.
3. Трофимова Т.И. Курс физики/ Трофимова Т.И. – М: Высшая школа, 1990. – 478 с.
4. Ормонт Б.Ф. Введение в физическую химию и кристаллохимию полупровдников/ Ормонт Б.Ф. – М: Высшая школа, 1973. – 655 с.
5. Най Дж. Физические свойства кристаллов/ Най Дж. – М: Мир, 1967. – 385 с.
6. Фистуль В.И. Введение в физику полупроводников/ Фистуль В.И. – М: Высшая школа, 1984. – 352 с.
7. Зи С. Физика полупроводниковых приборов. – Т. 1/ Зи С. – М: Мир, 1984. – 455 с.
8. Зи С. Физика полупроводниковых приборов. – Т. 2/ Зи С. – М: Мир, 1984. – 459 с.
9. Зиненко В.Н. Основы физики твердого тела/ Зиненко В.Н., Сорокин Б.П., Турчин П.П. – М: Высшая школа, 1983. – 330 с.
10. Матвеев А.Н. Молекулярная физика/ А.Н. Матвеев – М: Высшая школа, 1981. – 395 с.
11. Справочник по электротехническим материалам: в 3-х т.
    Т.1; под ред. Ю.Б. Корицкого и др. – М.: Энергоатомиздат,1986. – 368 с.
    Т.2; под ред.Ю.Б. Корицкого и др. – М.: Энергоатомиздат,1987. – 464 с.
    Т.3; под ред.Ю.Б.Корицкого и др. – Л.: Энергоатомиздат,1988. – 728 с.
12. Милнс А. Гетеропереходы и переходы металл-полупроводник/ А. Милнс, Д. Фойхт. – М.: Мир, 1975. – 432 с.

13. Бонч-Бруевич В.П. Физика полупроводников/ В.П. Бонч-Бруевич, С.Г. Калашников – М: Наука, 1990. – 685 с.
14. Митрофанов О.В. Физические основы функционирования изделий микроэлектроники/ Митрофанов О.В., Симонов Б.М., Коледов Л.А. – М.: Высшая школа, 1997. – 167 с.
15. Сивухин Д.В. Общий курс физики: в 4-х т. – Т.3/ Сивухин Д.В. – М.: Наука, 1977. – 687 с.
16. Викулин И.М. Физика полупроводниковых приборов/ И.М. Викулин, В.И.Стафеев. – М.: Радио и связь, 1990. – 262 с.
17. Уэрт Ч. Физика твердого тела/ Ч. Уэрт, Р. Томсон. – М: Мир, 1969. – 558 с.
18. Полупроводниковые фотоприемники: ультрафиолетовый, видимый и ближний инфракрасный диапазоны спектра/ [И.Д. Анисимова, И.М. Викулин, Ф.А Заитов, Ш.Д. Курмашев]; под ред. Стафеева В.И. – М.: Радио и связь, 1984. – 216 с.
19. Мустель Е.П. Методы модуляции и сканирования света/ Е.П. Мустель, В.Н. Парыгин. – М: Наука, 1970. – 426 с.
20. Сонин А.С. Электрооптические кристаллы/ А.С. Сонин, А.С.Васильева. – М: Атом. Издат., 1971. – 347 с.
21. Антипов Б.Л. Материалы электронной техники: Задачи и вопросы/ Антипов Б.Л., Сорокин В.С., Терехов В.А. – М: Высшая школа, 1990. – 208 с.
22. Мосс Т. Полупроводниковая оптоэлектроника/ Мосс Т., Барел Г., Эллис. Б. – М: Мир, 1976. – 431 с.
23. Панков Ж. Оптические процессы в полупроводниках/ Ж. Панков – М: Мир, 1973. – 456 с.
24. Суорц К.Э. Необыкновенная физика обыкновенных явлений/ Суорц К.Э. – М: Наука, 1987. – 384 с.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

1. Викулин И.М. Чувствительность кремниевых МДП-структур к водороду/ И.М. Викулин, Н.И. Крамаренко, С.К. Криськив //Электронная техника.– 1989. – №1. – С. 56-62. – (Сер.2. «Полупроводниковые приборы»).
2. Крамаренко Н.И. Пассивация поверхностных состояний на границе раздела $Si\text{-}SiO_2$ при воздействии водорода на Pd-МДП-структуры/ Н.И. Крамаренко, С.К. Криськив, Я.О. Ройзин // Аморфные полупроводники и диэлектрики на основе кремния в электронике. – 1989. – С.62.
3. Крамаренко Н.И. Изменение электрических параметров МДП-структур с палладиевыми затворами при адсорбции водорода/ Н.И. Крамаренко, С.К. Криськив, Я.О. Ройзин //Поверхность. – 1990. – С.61-65.
4. Викулин И.М. Об управлении спектром светоизлучающих диодов/ И.М. Викулин, Н.А. Кальнев, Б.В. Коробицын //Электронная техника – 1988. – Вып.4 (195). – С.70-72. – (Сер.2. «Полупроводниковые приборы»).

5. Викулин И.М. Фотоприемники с инжекционным усилением/ И.М. Викулин, Ш.Д. Курмашев, В.И. Стафеев // Электроника. – 1989. – С.60.
6. Оптимизация геометрических характеристик *p-n*-структур для оптоэлектроники / И.М. Викулин, В.Э. Горбачев, Б.В. Коробицын [и др.] // ТКЭА. – 2004. – №3. – С.38-39.
7. Викулин И.М. Обработка и представление спектров оптоэлектронных приборов/ И.М. Викулин, Б.В. Коробицын, С.К. Криськив [и др.]// Наукоі праці ОНАЗ. –2005. – №1. – С.99-103.
8. Vikulin I.M. Injection-Based Photodetectors / I.M. Vikulin, Sh.D. Kurmashev, V.I. Stafeev// Phisics of semiconductor devis. – 2008. – Vol.42. – №1. – P.113-127.
9. Сушко М.Я. Метод компактных групп в теории диэлектрической проницаемости гетерогенных систем/ М.Я. Сушко, С.К. Криськив// Журнал технической физики. – 2009. – Т. 79. – Вып. 3. – С.97-101.