

Министерство образования и науки Российской Федерации

---

САНКТ-ПЕТЕРБУРГСКИЙ  
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО

---

*Н.М. Гнучев*

# ФИЗИЧЕСКИЕ ОСНОВЫ ЭЛЕКТРОНИКИ

Учебное пособие

Санкт-Петербург

2021

# ОГЛАВЛЕНИЕ

Введение.....	6
1. Электронная структура кристаллов.....	10
1.1. Энергетические уровни атомов.....	10
1.2. Энергетические зоны кристаллов.....	11
2. Электроны и дырки в полупроводниках.....	17
2.1. Собственная проводимость.....	17
2.2. Примесная проводимость. Полупроводник <i>n</i> -типа.....	19
2.3. Примесная проводимость. Полупроводник <i>p</i> -типа.....	21
2.4. Распределение электронов по энергиям в пределах энергетических зон.....	23
2.5. Эффективная масса электрона. Квантовомеханическое понятие электрона и дырки в кристалле.....	27
2.6. Равновесная концентрация свободных носителей в собственных и примесных полупроводниках.....	30
2.7. Уровень Ферми на энергетической диаграмме собственного полупроводника.....	33
2.8. Взаимосвязь концентраций подвижных носителей и уровень Ферми в примесных полупроводниках.....	33
3. Движение носителей заряда в полупроводниках.....	37
3.1. Хаотическое движение.....	37
3.2. Дрейф подвижных носителей заряда. Дрейфовые токи.....	37
3.3. Диффузия подвижных носителей заряда. Диффузионные токи	38
3.4. Уравнение непрерывности.....	39
3.5. Взаимосвязь диффузии и дрейфа заряженных частиц.....	42
3.6. Движение носителей в сильных электрических полях.....	44
3.7. Эффект Холла.....	45
3.8. Термоэлектрические явления. Эффекты Зеебека и Пельтье.....	48
4. Генерация и рекомбинация носителей заряда.....	55
4.1. Тепловая генерация.....	55
4.2. Ударная генерация.....	57
4.3. Полевая генерация .....	59
4.4. Световая генерация.....	60

4.5. Рекомбинация подвижных носителей заряда. Механизмы рекомбинации.....	61
4.6. Диффузионная длина неосновных носителей в примесных полупроводниках.....	63
4.7. Время жизни неосновных носителей в полупроводнике. Уравнение генерации–рекомбинации.....	65
5. Поверхностные явления в полупроводниках.....	68
5.1. Особенности электронного строения приповерхностной области полупроводника.....	68
5.2. Эффект поля в полупроводниках.....	70
5.2.1. Режим обеднения.....	70
5.2.2. Режим инверсии.....	72
5.2.3. Режим обогащения.....	73
6. Контактные явления в полупроводниках.....	75
6.1. Идеальный электронно-дырочный переход ( $p-n$ -переход).....	75
6.1.1. $P-n$ -переход в равновесном состоянии.....	77
6.1.2. Прямое включение $p-n$ -перехода.....	84
6.1.3. Обратное включение $p-n$ -перехода.....	87
6.1.4. Вольтамперная характеристика идеализированного $p-n$ -перехода.....	90
6.2. Контакт между полупроводниками одинакового типа проводимости с разной концентрацией примеси.....	96
6.3. Реальный электронно-дырочный переход.....	99
6.3.1. Обратное смещение. Токи генерации и утечки.....	99
6.3.2. Прямое смещение. Ток рекомбинации и влияние сопротивления базы.....	101
6.4. $P-n$ -переход на основе вырожденных полупроводников.....	103
6.5. Пробой электронно-дырочного перехода.....	110
6.5.1. Лавинный пробой.....	110
6.5.2. Туннельный пробой $p-n$ -перехода.....	112
6.5.3. Тепловой пробой $p-n$ -перехода.....	113

6.6. Емкость $p-n$ -перехода.....	115
6.6.1. Барьерная емкость.....	116
6.6.2. Диффузионная емкость.....	118
6.7. Переходные процессы в $p-n$ -переходе.....	119
6.7.1. Установление прямого напряжения при подаче на переход скачка прямого тока.....	119
6.7.2. Установление постоянного обратного тока перехода при ступенчатом переключении напряжения с прямого на обратное.....	122
6.8. Переходы металл – полупроводник.....	126
6.8.1. Выпрямляющий контакт (барьер Шоттки).....	126
6.8.2. Омический контакт металл- полупроводник.....	131
6.9. Гетеропереходы.....	132
7. Фотоэлектрические явления и светоизлучение.....	139
7.1. Взаимодействие оптического излучения с полупроводниками. Фотопроводимость.....	139
7.2. Фоторезистивный эффект. Фоторезистор.....	140
7.3. Фотоэлектрический эффект в $p-n$ -переходе.....	143
7.4. Излучение света полупроводниками.....	147
7.5. Инжекционные светодиоды с $p-n$ -переходами.....	148
7.6. Полупроводниковые инжекционные лазеры.....	151
Библиографический список.....	155

## ВВЕДЕНИЕ

Электроника — очень широкая область знаний. Трудно назвать точку отсчета, с которой начинается ее развитие. Если началом развития считать изобретение первых электронных приборов (начало 20-го века), то электроника уже «отпраздновала» свое столетие. Ни одна отрасль науки и техники не развивалась столь стремительно, как электроника. Опираясь на успехи атомной физики и физики твердого тела, инженеры-физики создали огромное, трудно поддающееся перечислению количество вакуумных, а затем и полупроводниковых приборов всевозможного назначения.

Настоящее время — это эпоха интегральной электроники. Ее начало можно установить абсолютно точно. Это — 1947 год, изобретение транзистора. Авторы изобретения Уильям Шокли, Уолтер Браттейн, Джон Бардин, сотрудники фирмы Bell Telephone Laboratories, в 1956 году получили Нобелевскую премию «За исследования полупроводников и открытие транзисторного эффекта».

Идея интегральной схемы была оглашена еще в 1952 году, до появления групповых методов изготовления полупроводниковых приборов. На ежегодной конференции по электронным компонентам, проходившей в Вашингтоне, сотрудник Британского королевского радиолокационного управления в Малверне Джеффри Даммер представил доклад о надежности элементов радиолокационной аппаратуры. В докладе он сделал пророческое утверждение: «С появлением транзистора и работ в области полупроводниковой техники можно себе представить электронное оборудование в виде твердого блока, не содержащего соединительных проводов. Блок может состоять из слоев изолирующих, проводящих, выпрямляющих и усиливающих материалов, в которых определенные участки вырезаны таким образом,

чтобы они могли непосредственно выполнять электрические функции».

Днем рождения интегральной схемы считается 12 сентября 1958 года. В этот день сотрудник фирмы Texas Instruments Джек Килби продемонстрировал руководству три странных прибора — склеенные пчелиным воском на стеклянной подложке устройства из двух кусочков кремния размером 11,1 x 1,6 мм. Это были объемные макеты — прототипы интегральной схемы генератора, доказывающие возможность изготовления всех элементов схемы на основе одного полупроводникового материала. В 2000 году Дж. Килби получил Нобелевскую премию «за вклад в изобретение интегральных схем».

Все микросхемы вплоть до 60-х годов 20 века делались с биполярными транзисторами, и если бы так оставалось и впредь — не видать бы нам никаких персональных компьютеров и мобильных телефонов. Но в 1960 году Джон Аталла и Дэвон Канг из Bell Laboratories изготовили полевой транзистор с изолированным затвором. Получился всем сегодня известный «бутерброд»: металлический (Al) затвор, подзатворный оксид ( $\text{SiO}_2$ ) и канал-полупроводник (Si). И хотя первые два элемента уже давно делаются из других материалов, мы все еще называем это МОП-транзисторами (структура металл-оксид-полупроводник).

В настоящее время благодаря стремительному развитию микроэлектронной технологии достигнута очень высокая плотность расположения элементов интегральной схемы: на площади кристалла кремния  $\sim 300 \text{ мм}^2$  помещается более трех миллиардов транзисторов! Легко подсчитать среднюю площадь, занимаемую одним транзистором:  $\sim 0,1 \text{ мкм}^2$ . Соответственно, его линейные размеры составляют несколько сотен нанометров. Следует отметить, что, несмотря на столь малые размеры активных элементов ИС, физические принципы, лежащие в основе их работы, такие же, как и у дискретных

полупроводниковых приборов: диодов, биполярных и полевых транзисторов, тиристоров и. т. д.

Цель данного учебного пособия — дать студентам первичные представления о физических процессах, определяющих принципы действия различных приборов дискретной и интегральной полупроводниковой электроники, а также об их свойствах, характеристиках и параметрах. Для облегчения восприятия материала использование математического аппарата ограничено лишь теми вопросами, рассмотрение которых без использования математических выкладок невозможно.

Пособие состоит из введения и 7 разделов. Первые четыре раздела посвящены изложению основ физики полупроводников, необходимых для понимания физических принципов работы электронных приборов. В первом разделе вводится понятие энергетических зон кристаллов, описано их различие для металлов, диэлектриков и полупроводников.

Во втором разделе рассматриваются особенности электропроводности собственных и примесных полупроводников; вводятся основные термины и понятия, описывающие электронные свойства полупроводников и их электронную структуру. Проведен краткий анализ причин видоизменения энергетических диаграмм при переходе от собственных полупроводников к *p*- и *n*-полупроводникам с различной концентрацией примесей.

Третий раздел содержит сведения о различном характере движения заряженных частиц в кристалле полупроводника. Рассмотрены диффузия и дрейф подвижных носителей заряда, приведены примеры взаимосвязи этих явлений. Описаны также особенности движения электронов и дырок в полупроводнике при воздействии магнитных полей и основы термоэлектрических явлений.

Четвертый раздел посвящен процессам генерации и рекомбинации подвижных носителей заряда. Проведена оценка вероятности межзонной излучательной и безызлучательной рекомбинации, а также рекомбинации с участием чужеродных атомов – «ловушек». Введены понятия «диффузионная длина» и «время жизни» неосновных носителей в примесных полупроводниках.

В пятом разделе обсуждаются особенности электронного строения приповерхностной области полупроводника, а также механизмы влияния поверхностного заряда и внешнего электрического поля на тип ее проводимости и концентрацию носителей.

В шестом разделе всесторонне рассматриваются физические принципы и основные свойства различных полупроводниковых переходов. Материал седьмого раздела — взаимодействие электромагнитного (светового) излучения с полупроводниками и  $p-n$ -переходами и излучение света  $p-n$ -переходом.

Каждый раздел сопровождается вопросами для самопроверки.

Данное пособие предназначено в качестве основного материала к учебным дисциплинам «Электронные приборы» и «Физические основы электроники» направлений подготовки бакалавров «Техническая физика» и «Электроника и наноэлектроника». Кроме того, пособие может использоваться как дополнительное для студентов других направлений подготовки, изучающих дисциплину «Электроника» и как материал для подготовки к лабораторным работам.

В рамках магистерских программ указанных направлений данное пособие соответствует содержанию аннотации дисциплины вариативной части «Практикум в лаборатории». Материал пособия может быть использован непосредственно в практической работе в научной лаборатории.

# 1. ЭЛЕКТРОННАЯ СТРУКТУРА КРИСТАЛЛОВ

## 1.1. ЭНЕРГЕТИЧЕСКИЕ УРОВНИ АТОМОВ

Атом любого вещества является весьма сложной физической системой. В рамках классического подхода его структура (в «нулевом» приближении — ядро с вращающимися вокруг него электронами) абсолютно неустойчива. Электроны, двигаясь по орбитам с центростремительным ускорением, должны терять энергию на электромагнитное излучение и неизбежно «упасть» на ядро.

Теория Бора позволила разрешить это противоречие. Согласно ее постулатам, электроны движутся по стационарным орбитам вокруг ядра, не излучая электромагнитные волны и, следовательно, не теряют энергию. Кроме того, переход электрона на более «высокую» или более «низкую» орбиту возможен только при поглощении или испускании кванта с энергией  $h\nu = E_m - E_n$ , где  $E_m$  и  $E_n$  — значения энергий электрона в атоме, соответствующие двум различным состояниям.

Расположение электронов по энергетическим состояниям и электронным оболочкам в атомах должно соответствовать основному правилу атомной физики — принципу запрета Паули. Согласно этому принципу, в атоме в данном состоянии может находиться только один электрон, состояние другого должно отличаться хотя бы одним квантовым числом. Под уровнем энергии будем понимать объединение электронов в двух состояниях с противоположными ориентациями спинов. Тогда можно считать, что на одном уровне энергии может быть не более двух электронов, и на электронных оболочках атомов могут находиться электроны различных энергетических уровней. На основе всех этих закономерностей построена периодическая система элементов.

Учитывая все, сказанное выше, проследим, как изменяется количество электронов на первых двух атомных «орбитах». На *K*-оболочке, ближайшей к ядру атома, может находиться не более двух электронов (один энергетический уровень,  $1s$ -состояние). На *L*-оболочке может находиться не более восьми электронов, из которых два — в  $2s$ -состоянии и шесть — в состоянии  $2p$ ; в этом состоянии объединены 3 уровня энергии. Так, например, в атоме простейшего металла (лития) 3 электрона (рис. 1.1), из которых два на *K*-оболочке ( $1s$ ), и третий на *L*-оболочке в состоянии  $2s$ . В атоме следующего металла, бериллия, 4 электрона: по два в состояниях  $1s$  и  $2s$ . Затем, начиная с атома бора вплоть до атома неона, происходит заполнение  $2p$ -состояний (от одного до шести электронов). Предшествующий неону в таблице Менделеева атом простейшего галогена (фтора) имеет на *L*-оболочке семь электронов: два в  $2s$ -состоянии и пять в состоянии  $2p$ .

## 1.2. ЭНЕРГЕТИЧЕСКИЕ ЗОНЫ КРИСТАЛЛОВ

При объединении отдельных атомов вещества в кристалл все электроны образуют систему, в которой действуют законы квантовой механики. Согласно принципу Паули, электроны, занимающие один и тот же энергетический уровень в отдельных атомах, не могут находиться в кристалле в абсолютно одинаковых энергетических состояниях. Например, в кристалле лития, составленном из  $N$  атомов,  $1s$ -электроны не могут занимать один и тот же уровень энергии. Поэтому при образовании кристалла происходит расщепление энергетических уровней, и образуется энергетическая зона ( $1s$ -зона лития), состоящая из  $N$  очень близко (с интервалом  $10^{-20} - 10^{-21}$  эВ) расположенных уровней. Таким же образом в кристалле образуются и остальные энергетические зоны. Ясно, что при такой «тесноте» уровней распределение электронов по энергиям в пределах зоны можно считать не дискретным, а квазинепрерывным.

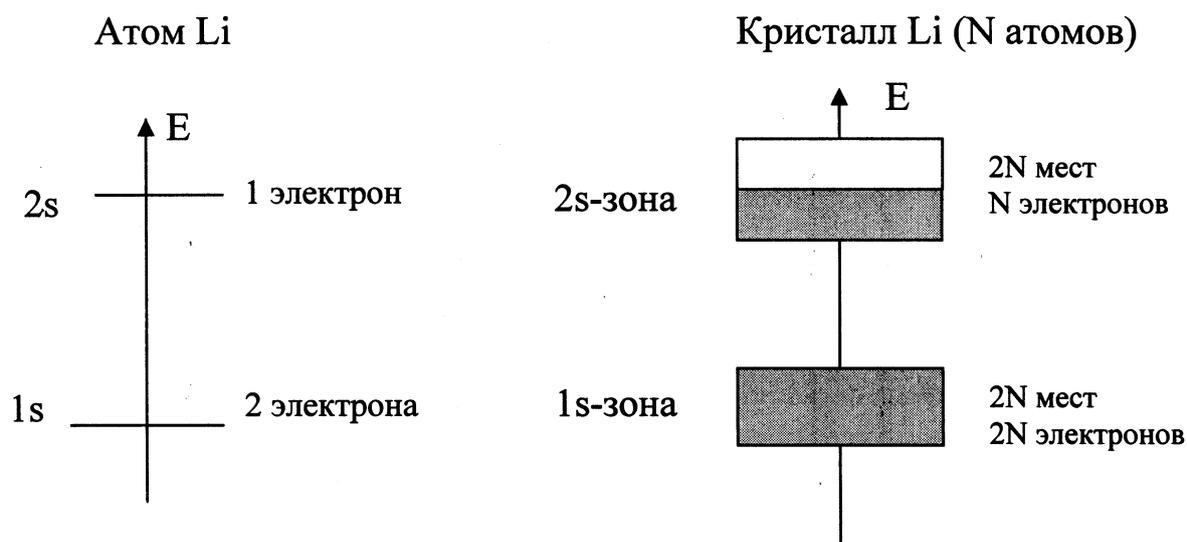


Рис. 1.1. Упрощенный принцип образования энергетических зон в кристалле лития, составленном из  $N$  атомов:  — части энергетических зон, занятые электронами;  — состояния, не занятые электронами в зоне

Принцип формирования энергетических зон в кристалле лития показан на рис. 1.1. В зоне  $1s$  все  $2N$  вакантных мест заполнены электронами, а зона  $2s$  оказывается заполненной лишь наполовину: на  $2N$  мест приходится только  $N$  электронов. Наличие частично заполненной энергетической зоны — характерный признак всех металлов, обеспечивающий их электропроводность.

Электроны частично заполненной зоны — это «свободные» электроны в металле. В отсутствие внешнего электрического поля их импульсы распределены в пространстве кристалла изотропно, что соответствует их хаотическому тепловому движению. При появлении электрического поля возникает асимметрия в пространственном распределении импульсов электронов, и как следствие, дрейф всей совокупности электронов в направлении, противоположном вектору электрического поля.

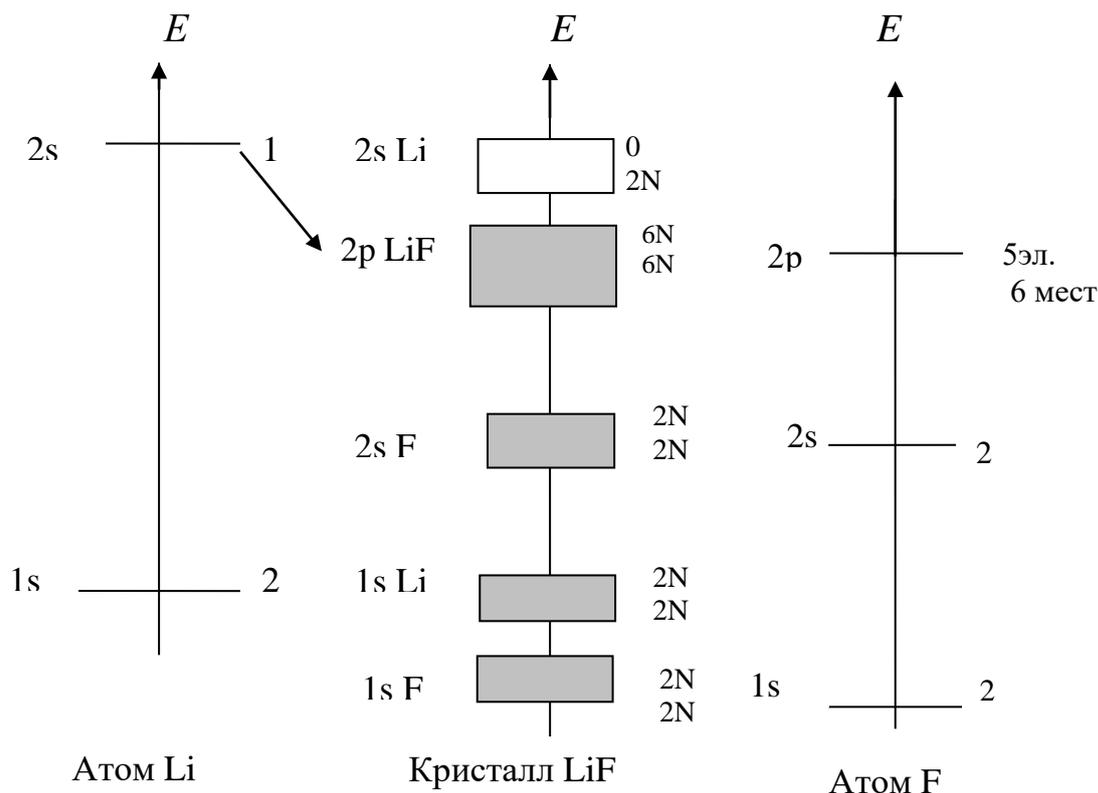


Рис. 1.2. Упрощенный принцип образования энергетических диэлектрика LiF, составленном из  $N$  атомов Li и  $N$  атомов F.

Справа от зон (снизу вверх) указаны: число мест для электронов и число электронов в зоне. Зона  $2p$  LiF содержит  $5N$  электронов с уровней  $2p$  фтора и  $N$  электронов с уровней  $2s$  лития

Рассмотрим теперь, каким образом образуются энергетические зоны в кристалле простейшего диэлектрика, фторида лития LiF. Будем считать, что кристалл составлен из  $N$  атомов лития и  $N$  атомов фтора (рис. 1.2). В соответствии с рассмотренными выше принципами  $1s$ -уровень лития,  $1s$ - и  $2s$ - уровни фтора расщепляются в энергетические зоны. В каждой из этих зон будет по  $2N$  мест, полностью заполненных электронами.

При образовании зон  $2p$  фтора и  $2s$  лития в кристалле диэлектрика LiF происходит перераспределение электронов.  $2s$ -электроны лития переходят на нижележащие свободные уровни  $2p$ -зоны фтора, в которой изначально имеется  $N$  свободных мест. В результате  $2p$ -зона

заполняется полностью, в  $2s$ -зоне Li все  $2N$  состояний остаются незанятыми, и между свободными и занятыми электронными энергетическими состояниями возникает интервал запрещенных значений энергии — «запрещенная зона».

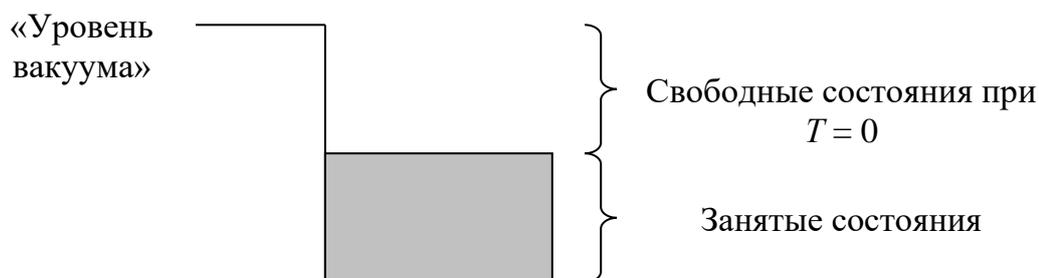


Рис. 1.3. Простейшая одномерная энергетическая диаграмма металла. «Уровень вакуума» — энергия, которой обладает электрон, вышедший из металла и находящийся в состоянии покоя с нулевой кинетической энергией вблизи его поверхности (формально — в бесконечно удаленной точке пространства)

В качестве одномерной энергетической диаграммы используют самые верхние зоны энергий электронов, содержащие занятые и свободные энергетические уровни. В нашем примере для металла (лития) — это частично заполненная  $2s$ -зона, для диэлектрика LiF — полностью заполненная зона  $2p$  фтора и зона свободных состояний  $2s$  Li.

Таким образом, простейшая одномерная энергетическая диаграмма металла (рис. 1.3) — это квазинепрерывный ряд занятых и свободных состояний.

Простейшая энергетическая диаграмма диэлектрика или идеального полупроводника (рис. 1.4) состоит из двух зон: валентной (полностью заполненной) и зоны проводимости, в которой при  $T = 0$  нет электронов. В валентной зоне находятся энергетические

состояния электронов, обеспечивающих химическую связь между атомами в диэлектрике.

Энергетические состояния в зоне проводимости соответствуют «свободным» электронам кристалла, обеспечивающим его электропроводность. Эти электроны появляются в кристалле диэлектрика при повышении температуры до значений, соответствующих энергиям, необходимым для разрыва валентных связей (см. раздел 2).

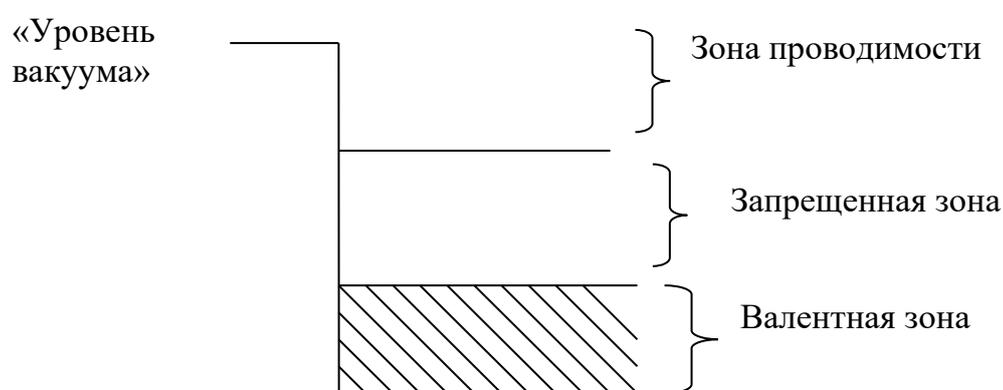


Рис. 1.4. Одномерная энергетическая диаграмма идеального кристалла диэлектрика или беспримесного (собственного) полупроводника

Валентная зона и зона проводимости на энергетической диаграмме (рис. 1.4) отделены друг от друга интервалом запрещенных для электронов значений энергии — «запрещенной зоной»  $\Delta E_z$ . Физический смысл этой величины очевиден: это — наименьшая энергия, которую необходимо сообщить связанным электронам для перевода их в «свободное» состояние в кристалле, когда они могут обеспечить электропроводность.

### *Вопросы к разделу 1*

1. Почему планетарная модель атома не может существовать в рамках классической физики, и каким образом теория Бора позволила сохранить эту модель?
2. Каковы основные закономерности построения периодической системы элементов?
3. Каковы физические причины расщепления уровней энергии атомов в энергетические зоны при объединении отдельных атомов вещества в кристалл?
4. Поясните, каким образом формируются энергетические зоны лития. Какова особенность энергетической структуры всех металлов?
5. Поясните, каким образом формируются энергетические зоны простейшего диэлектрика, фторида лития. Какова особенность энергетической структуры всех диэлектриков?
6. По какому принципу формируются простейшие энергетические диаграммы? Каковы энергетические диаграммы металла и диэлектрика?
7. Каков физический смысл величины  $\Delta E_z$  — ширины запрещенной зоны на энергетической диаграмме диэлектрика и полупроводника?
8. Используя простейшие энергетические диаграммы, объясните различие в характере температурной зависимости электропроводности металлов и диэлектриков.

## 2. ЭЛЕКТРОНЫ И ДЫРКИ В ПОЛУПРОВОДНИКАХ

### 2.1. СОБСТВЕННАЯ ПРОВОДИМОСТЬ

Рассмотрим кристалл четырехвалентного полупроводника, например, кремния. Кристаллическая решетка таких веществ построена на основе ковалентной (гомеополярной) связи; при этом происходит обобществление валентных электронов соседних атомов.

При сообщении кристаллической решетке некоторой энергии, например за счет нагрева кристалла, отдельные электроны могут покинуть валентные связи и превратиться в свободные носители заряда. Величина этой энергии должна быть равной или большей, чем энергия связи валентного электрона. Так как изначально электрон был общим для двух атомов (Si), то нельзя сказать, что ионизирован один из этих атомов. Уход электрона приведет к частичной ионизации двух соседних атомов. Поэтому появляющийся при этом положительный единичный заряд, равный заряду электрона, относят не к тому или иному атому, а к дефектной связи, оставленной электроном. Такой положительный заряд принято называть дыркой (рис. 2.1).

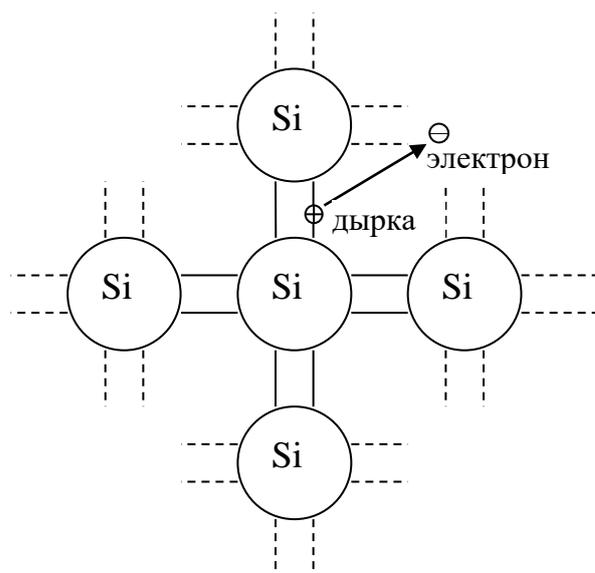


Рис. 2.1. Образование пары «электрон-дырка» в собственном полупроводнике

С уходом электрона в одной из валентных связей появляется вакансия, которую может замещать один из валентных электронов соседних связей. При таком переходе электрона из заполненной связи в дефектную происходит перемещение дырки в обратном направлении. Этот процесс взаимного перемещения электронов и дырок носит случайный, беспорядочный характер. Однако если поместить кристалл полупроводника в электрическое поле, то более вероятными будут такие переходы электронов из связи в связь, при которых дырка (положительный заряд) передвигалась бы в направлении вектора напряженности электрического поля.

Таким образом, разрыв валентной связи за счет передачи энергии кристаллу извне приводит к появлению пары свободных носителей заряда — электрона и дырки. Электропроводность, возникающая в кристалле полупроводника за счет нарушения валентных связей, называется собственной электропроводностью. При этом концентрация подвижных носителей (электронов и дырок) одинакова.

Энергетическая диаграмма идеального собственного полупроводника по своей структуре точно такая же, как на рис. 1.4. Имеется лишь одно отличие — в ширине запрещенной зоны. Полупроводником считается (весьма условно!) вещество, имеющее ширину запрещенной зоны  $\Delta E_z < 3-3,5$  эВ. Характерные значения величины  $\Delta E_z$  для наиболее распространенных полупроводников приведены в табл. 2.1.

Таблица 2.1

**Ширина запрещенной зоны некоторых широко используемых полупроводниковых материалов**

Полупроводник	Ширина запрещенной зоны, эВ
Германий (Ge)	0,66
Кремний (Si)	1,12
Арсенид галлия (GaAs)	1,42
Нитрид галлия (GaN)	3,39

## 2.2. ПРИМЕСНАЯ ПРОВОДИМОСТЬ. ПОЛУПРОВОДНИК N-ТИПА

Если вместо одного из атомов четырехвалентного полупроводника (Si, Ge) ввести в кристалл атом вещества 5-й группы периодической системы элементов (Sb, P), то этот примесный атом образует только 4 ковалентные связи с четырьмя соседними атомами. Пятый электрон оказывается связанным со своим атомом значительно слабее. Для того чтобы оторвать его от атома и превратить в свободный носитель заряда, требуется значительно меньше энергии, чем для высвобождения электрона из валентной связи атома основного вещества. Такие примесные атомы называются донорами.

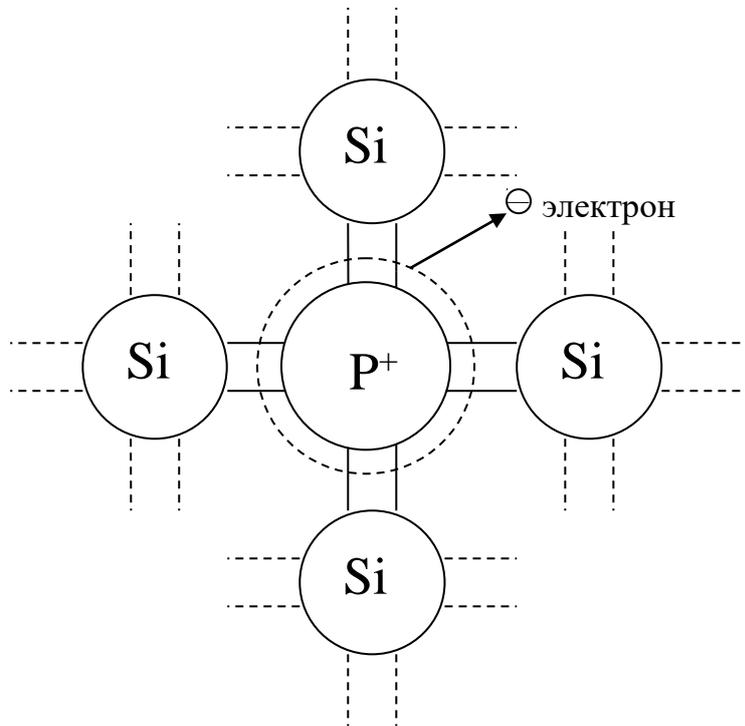


Рис. 2.2. Возникновение «свободного» электрона в полупроводнике *n*- типа за счет ионизации атома донорной примеси при повышении температуры.

Штриховой круг около атома фосфора обозначает локализацию 5-го электрона при  $T = 0 \text{ K}$

В рамках зонной модели это значит, что для того, чтобы перевести такой электрон в зону проводимости, необходимо значительно меньшее количество энергии, чем ширина запрещенной зоны  $\Delta E_z$ . Иными

словами, донорный уровень, на котором будут находиться такие электроны в невозбужденном состоянии, должен располагаться в запрещенной зоне вблизи дна зоны проводимости (рис. 2.3). При температурах, близких к абсолютному нулю, все избыточные электроны находятся на этом уровне (т. е. «привязаны» к своим примесным атомам), но уже при незначительном повышении температуры они получают достаточное количество энергии для перехода в зону проводимости. При этом атомы примеси фосфора превращаются в положительные ионы  $P^+$ .

Увеличивая концентрацию примесных атомов, мы тем самым увеличиваем количество свободных электронов в полупроводнике. Поскольку каждый атом примеси может дать один свободный электрон, в идеальном случае концентрация электронов проводимости равна концентрации примеси.

Количество дырок при этом не возрастает; оно определяется концентрацией собственных электронно-дырочных пар при данной температуре. В таком кристалле полупроводника ток в основном переносится электронами. Электроны в этом случае будут основными носителями, дырки — неосновными.

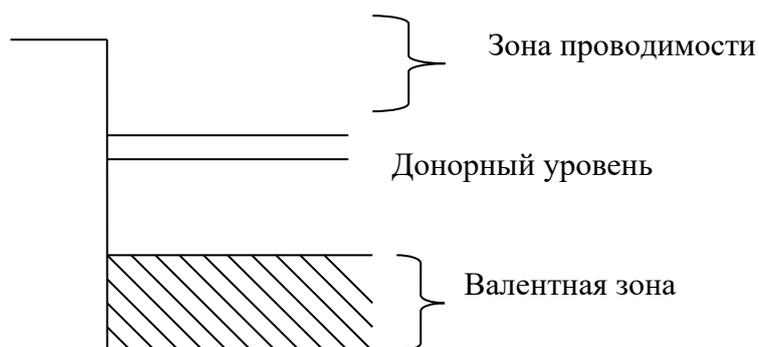


Рис. 2.3. Энергетическая диаграмма донорного полупроводника при невысоких концентрациях примеси ( $< 10^{18}$  атомов/см<sup>3</sup>)

Донорные примеси обеспечивают электронную проводимость полупроводника. Такие примесные полупроводники называются донорными, электронными или полупроводниками *n*-типа.

### 2.3. ПРИМЕСНАЯ ПРОВОДИМОСТЬ. ПОЛУПРОВОДНИК P-ТИПА

Введем теперь в кристалл вместо одного из атомов кремния атом 3-й группы периодической системы, например, атом бора. Его внешняя оболочка содержит три валентных электрона, и трехвалентный атом бора образует только три заполненные валентные связи. Четвертая связь остается не заполненной, но никакого заряда она не несет. Таким образом, при  $T = 0$  К атом бора и соседние атомы Si являются электрически нейтральными.

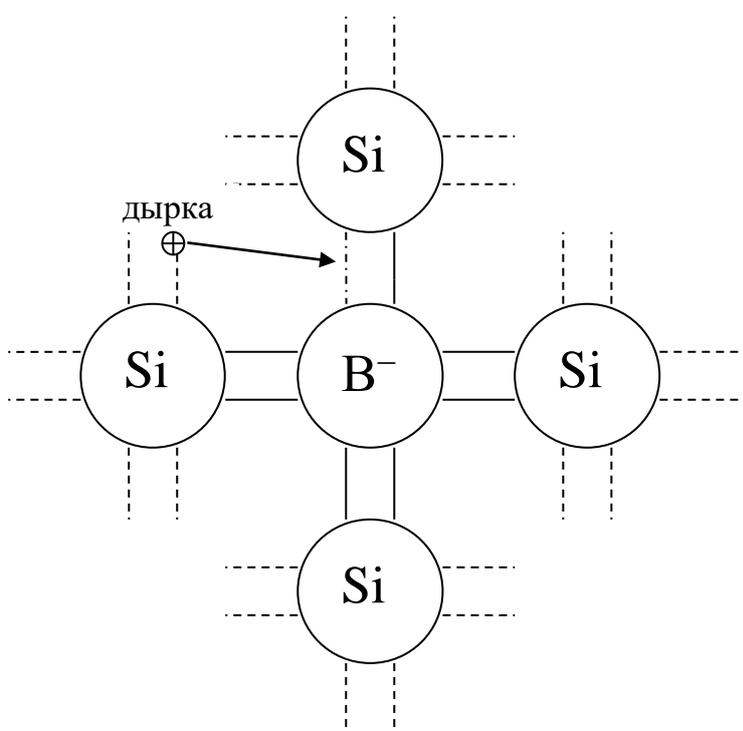


Рис. 2.4. Возникновение дырки в полупроводнике *p*-типа за счет ионизации трехвалентного атома примеси при повышении температуры.

Стрелкой показан переход валентного электрона атома кремния к атому примеси, приводящий к его ионизации

При небольшом тепловом возбуждении электрон с одной из соседних валентных связей (рис. 2.4) может перейти в эту не заполненную связь. При этом во внешней оболочке атома бора появляется лишний электрон; этот атом превращается в отрицательный ион, а в

связи, откуда ушел электрон, появляется дырка, несущая (условно) эффективный положительный заряд  $+e$ . Такие примесные атомы называются акцепторами.

Вводя в германий или кремний примеси акцепторов, атомов III группы, мы получаем возможность повышать концентрацию только дырок. Такой полупроводник будет обладать, в основном, дырочной проводимостью. Дырки будут основными носителями, электроны — неосновными.

Так как переход электронов из валентных связей к атому примеси не требует больших энергий, сравнимых с шириной запрещенной зоны, то введение атомов трехвалентной примеси в решетку кремния или германия приводит к появлению незанятого при  $T = 0$  К акцепторного уровня  $E_A$  в запрещенной зоне вблизи потолка валентной зоны (рис. 2.5).

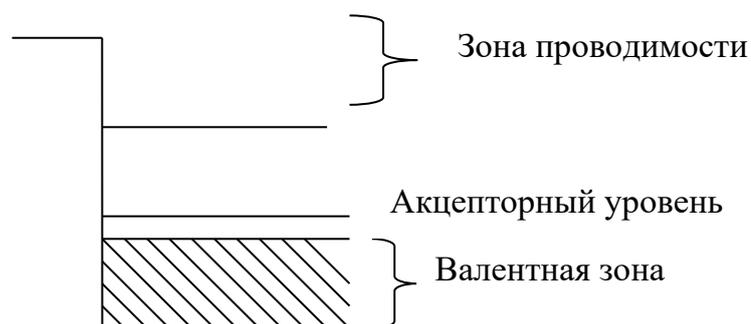


Рис. 2.5. Энергетическая диаграмма акцепторного полупроводника при невысоких концентрациях примеси ( $\leq 10^{17}$  атомов/см<sup>3</sup>)

При температурах, близких к абсолютному нулю, акцепторный уровень остается свободным. При небольшом повышении температуры один из валентных электронов покидает валентную зону и занимает этот уровень, оставляя после себя в валентной зоне свободный уровень — «дырку». Иначе говоря, акцепторный уровень принимает на себя все электроны, заполнившие недостающую, четвертую связь атомов акцепторов.

Таким образом, акцепторные примеси, принимая валентные электроны на акцепторные уровни, приводят к появлению дырочной проводимости. Полупроводники, в которых основными носителями заряда являются дырки, называются акцепторными, дырочными, или полупроводниками *p*-типа.

Концентрация примесей в полупроводниках обычно невысока:  $10^{-3} - 10^{-5} \%$ . Это означает, что при равномерном распределении примеси по объему полупроводника на  $10^5 - 10^7$  атомов вещества приходится один атом примеси. Тогда расстояние между соседними примесными атомами составляет  $10^5 - 10^7$  постоянных решетки, и их можно считать изолированными и не взаимодействующими друг с другом. Поэтому при таких концентрациях энергетические состояния всех атомов примеси характеризуются единым энергетическим уровнем.

В ряде случаев концентрацию примесей в полупроводниках доводят до  $10^{-1}\% - 1\%$  ( $10^{19} - 10^{20}$  атомов/см<sup>3</sup>). При этом примесные атомы образуют систему взаимодействующих частиц, и дискретный энергетический уровень примеси расщепляется в примесную зону. При мелкозалегающих уровнях примеси эта зона перекрывается с соседней зоной (валентной зоной *p*-полупроводника или зоной проводимости полупроводника *n*-типа). При этом для подвижных носителей образуется непрерывная энергетическая диаграмма, характерная для металлов. Такие примесные полупроводники называются вырожденными; их электрические свойства приближаются к свойствам металлов (подробнее см. раздел 6.4).

## **2.4. РАСПРЕДЕЛЕНИЕ ЭЛЕКТРОНОВ ПО ЭНЕРГИЯМ В ПРЕДЕЛАХ ЭНЕРГЕТИЧЕСКИХ ЗОН**

Для того, чтобы выяснить, каков характер распределения электронов по энергиям в пределах данной энергетической зоны, необходимо знать, во-первых, как распределены электроны по энергетическим уровням («по состояниям») в зоне, т. е. функцию плотности

заполнения состояний электронами  $\frac{dN}{dZ}$  и, во-вторых, по какому закону эти уровни размещены внутри зоны, или, иначе говоря, какова функция плотности состояний  $\frac{dZ}{dE}$ . Перемножив эти два выражения, мы и получим функцию распределения электронов по энергиям:

$$\frac{dN}{dZ} \cdot \frac{dZ}{dE} = \frac{dN}{dE} = f(E). \quad (2.1)$$

Для начала обратимся к энергетической диаграмме металла. Она представляет собой так называемую «открытую зону», в которой отсутствует верхняя граница энергетических состояний электронов. Электроны в твердых телах представляют собой квантовомеханическую систему, поэтому распределение таких электронов по состояниям подчиняется квантовой статистике Ферми-Дирака.

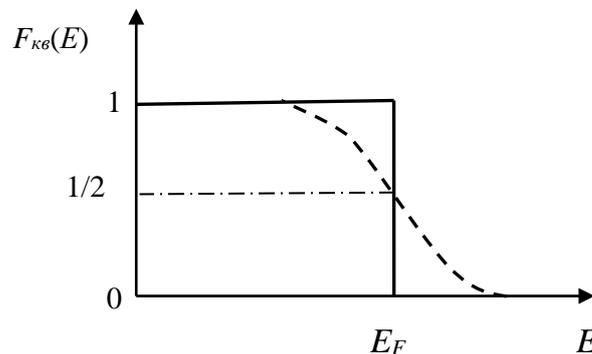


Рис. 2.6. Функция Ферми при  $T = 0$  (сплошная кривая) и при  $T > 0$  (штриховая линия)

Вероятность того, что уровень с данной энергией  $E$  при данной температуре  $T$  будет занят электроном, описывается функцией Ферми:

$$\frac{dN}{dZ} = F_{\text{кв}}(E) = \frac{1}{e^{\frac{E-E_F}{kT}} + 1}. \quad (2.2)$$

Рассмотрим конкретный вид функции Ферми при  $T = 0$  К (рис. 2.6). Для случая  $E > E_F$   $F_{\text{кв}}(E) = 0$ ; при  $E < E_F$   $F_{\text{кв}}(E) = 1$ ,

наконец, при  $E = E_F$  функция Ферми не определена. При  $T > 0$  и  $E = E_F$  функция определена и равна  $F_{\text{кв}}(E) = \frac{1}{2}$ .

Таким образом, при температуре абсолютного нуля электроны металла занимают все энергетические уровни в интервале  $0 \leq E \leq E_F$ . Уровни, расположенные выше  $E_F$ , при  $T = 0$  свободны. Уровень, соответствующий энергии электрона, равной  $E_F$ , называется уровнем Ферми. Физический смысл уровня Ферми в металле ясен: это — наибольшее значение энергии, которую может иметь свободный электрон в металле при температуре абсолютного нуля.

При повышении температуры резкая граница распределения Ферми при  $E = E_F$  «размывается»: вероятность нахождения электрона на уровнях, примыкающих к  $E_F$  со стороны больших энергий, становится отличной от нуля, и, соответственно, вероятность нахождения электрона на уровнях, примыкающих к  $E_F$  со стороны меньших энергий, становится меньше единицы. Поэтому, часть электронов с энергиями, меньшими или равными  $E_F$ , получив энергию порядка  $kT$ , переходят на уровни энергии, лежащие выше уровня Ферми.

Теперь обратимся к функции плотности состояний  $\frac{dZ}{dE}$ . Требуется определить закон изменения числа энергетических состояний  $Z$  внутри малого интервала  $\Delta E$  в зависимости от расположения этого интервала на оси энергии. Эта задача решена в физике твердого тела:

$$S_{\text{кв}}(E) = \frac{dZ}{dE} = A \sqrt{|E - E_{\text{гр}}|}, \quad (2.3)$$

где  $A = \frac{4\pi}{h^3} (2m^*)^{\frac{3}{2}}$ ,  $E_{\text{гр}}$  — величина, соответствующая началу отсчета энергии,  $h$  — постоянная Планка,  $m^*$  — эффективная масса электрона (см. раздел 2.5).

Для металла  $E_{\text{гр}} = 0$ , и распределение электронов проводимости по энергиям описывается выражением:

$$\frac{dN}{dE} = \Phi_{\text{кв}}(E) = \frac{4\pi}{h^3} (2m^*)^{\frac{3}{2}} \frac{\sqrt{E}}{e^{\frac{E-E_F}{kT}} + 1}. \quad (2.4)$$

График функции распределения электронов по энергиям в металле имеет вид, представленный на рис. 2.7.

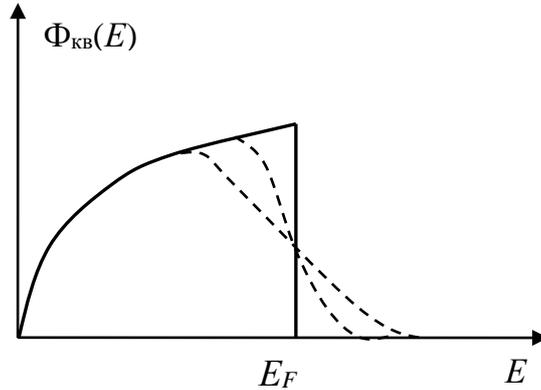


Рис. 2.7. Распределение электронов проводимости металла по энергиям при  $T = 0$  (сплошная кривая) и при  $T > 0$  (штриховые линии).

С ростом температуры происходит размытие границы при  $E = E_F$ : «хвост» распределения увеличивается

В собственном полупроводнике рассматриваются состояния электронов в валентной зоне и в зоне проводимости. Валентная зона — «закрытая», т. е. имеет верхнюю и нижнюю границы; при  $T = 0$  все уровни в ней заняты электронами. Зона проводимости является открытой, и при  $T = 0$  полностью свободна. В закрытой и полностью заполненной зоне вид функции плотности состояний  $\frac{dZ}{dE} = A\sqrt{|E - E_{\text{гп}}|}$  и, соответственно, распределение электронов по энергиям известны только на краях зоны вблизи от граничных значений энергии.

Положение уровня Ферми на энергетической диаграмме собственного полупроводника определяется из условия равенства концентраций электронов и дырок (см. ниже). Расчет показывает, что уровень Ферми и, соответственно, граница функции Ферми  $E_F$  в этом

случае расположены посередине запрещенной зоны. Расположение уровня Ферми на энергетической диаграмме и изменения функций распределения  $F_{\text{кв}}(E)$ ,  $S_{\text{кв}}(E)$  и  $\Phi_{\text{кв}}(E)$  с ростом температуры показаны на рис. 2.8.

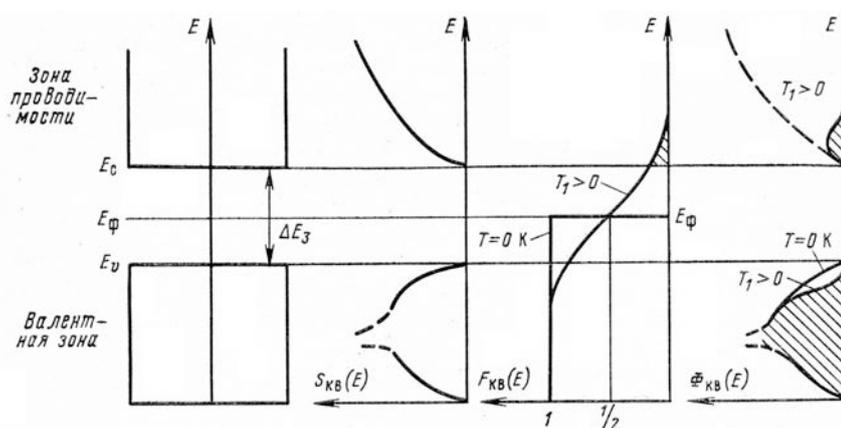


Рис. 2.8. Положение уровня Ферми и изменение функций распределения с температурой в собственном (беспримесном) полупроводнике

Заштрихованные площади под графиками соответствуют занятым электронным состояниям. При  $T > 0$  часть электронов, занимающих верхние энергетические уровни валентной зоны, переходят на уровни вблизи дна зоны проводимости (см. правую часть рисунка). Освобожденные в результате таких переходов состояния соответствуют появлению уровней дырок у потолка валентной зоны.

## 2.5. ЭФФЕКТИВНАЯ МАССА ЭЛЕКТРОНА. КВАНТОМЕХАНИЧЕСКОЕ ПОНЯТИЕ ЭЛЕКТРОНА И ДЫРКИ В КРИСТАЛЛЕ

Введение эффективной массы  $m^*$  вместо массы свободного электрона связано с тем, что электроны в кристалле находятся в периодическом поле кристаллической решетки. Электроны перемещаются в кристалле под действием двух сил: внешней силы  $\mathbf{F}$ , например, при наличии внешнего электрического поля, и сил, действующих со стороны решетки. Такое движение можно описать как движение свободного электрона под действием только внешней силы  $\mathbf{F}$ , и тогда

взаимосвязь между приложенной внешней силой и ускорением записывается в форме второго закона Ньютона:  $\mathbf{F} = m^* \cdot \mathbf{a}$ , где  $m^*$  — эффективная масса электрона. Этот подход применяется в физике твердого тела также и для описания движения дырок в кристалле полупроводника.

В рамках квантовой механики импульс частицы  $p$  пропорционален ее волновому вектору:  $p = \hbar k$ . Зависимость энергии от волнового вектора свободной частицы в кристалле, например, электрона или дырки в полупроводнике, можно описать с помощью формулы

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}. \quad (2.5)$$

Производная энергии по волновому вектору:

$$\frac{dE}{dk} = \frac{\hbar^2 k}{m}, \text{ при этом } \frac{1}{\hbar} \frac{dE}{dk} = \frac{\hbar k}{m} = V.$$

Данное выражение показывает, что скорость частицы  $V$  в кристалле пропорциональна первой производной энергии по волновому вектору частицы, или, что то же самое, по ее импульсу  $\hbar k$ . Тогда ускорение частицы

$$a = \frac{dV}{dt} = \frac{\hbar}{m} \frac{dk}{dt} = \frac{1}{\hbar} \frac{d}{dt} \left( \frac{dE}{dk} \right), \text{ откуда } a = \frac{1}{\hbar} \frac{dk}{dt} \frac{d^2 E}{dk^2}. \quad (2.6)$$

Для того, чтобы для свободной частицы в кристалле выполнялся закон Ньютона  $dp/dt = F = ma$ , из выражения (2.6) следует, что масса свободной частицы в кристалле должна определяться выражением:

$$m = \frac{\hbar^2}{d^2 E / dk^2} = m^*. \quad (2.7)$$

Так вводится понятие эффективной массы свободной частицы в кристалле. Это — сугубо квантовая величина, которую ни в коем случае нельзя отождествлять с массой свободной частицы в вакууме.

В соответствии с элементарными представлениями физики твердого тела зависимость энергии от модуля волнового вектора имеет вид, показанный на рис. 2.9. Верхняя кривая соответствует электронам зоны

проводимости, нижняя кривая – дыркам в валентной зоне. Кривые разделены энергетическим промежутком — запрещенной зоной.

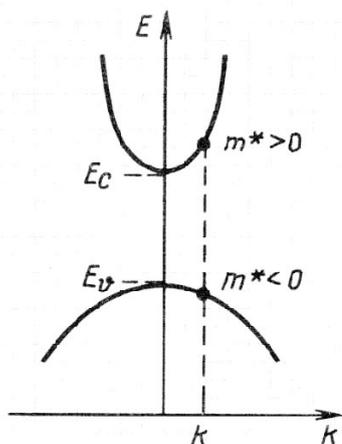


Рис.2.9. Диаграмма энергетических зон полупроводника или диэлектрика в форме зависимости энергии от волнового вектора.

$E_C$  — наименьшее возможное значение энергии электрона проводимости,  $E_V$  — та же величина для дырок

Если кривая  $E(k)$  имеет выпуклость вниз, то, согласно определению (2.7) эффективная масса положительна. Такой знак  $m^*$  соответствует электронам проводимости. Противоположный случай ( $m^* < 0$ ) соответствует дыркам в валентной зоне. При отрицательной эффективной массе частица будет ускоряться в направлении, противоположном ускорению электрона, т. е. она будет вести себя как некоторая «воображаемая» частица с положительной массой и положительным зарядом. Именно эта воображаемая частица в физике твердого тела называется дыркой.

Очевидно, что такое определение отличается от того, которое вводится в модели ковалентных связей. В этой модели движение дырки равносильно переходам электронов из связи в связь в противоположном направлении. Поэтому, строго говоря, в этом случае нет необходимости говорить о существовании двух типов носителей заряда. Тем не менее, это — доказанный экспериментальный факт: опыты по измерению эффекта Холла в полупроводниках (см. раздел 3.6) подтвердили наличие подвижных носителей обоих знаков.

Таким образом, квантовомеханическое понятие эффективной массы позволяет рассматривать дырки и электроны как классические заряженные частицы с эффективными массами  $m_n^*$  и  $m_p^*$

## 2.6. РАВНОВЕСНАЯ КОНЦЕНТРАЦИЯ СВОБОДНЫХ НОСИТЕЛЕЙ В СОБСТВЕННЫХ И ПРИМЕСНЫХ ПОЛУПРОВОДНИКАХ

Очевидно, что равновесную концентрацию электронов в зоне проводимости полупроводника можно определить как интеграл от функции  $\Phi_{\text{кв}}(E)$ . В этом случае  $E_{\text{гр}} = E_C$ , т. е. значение  $E_{\text{гр}}$  в формуле (2.3) соответствует энергии электрона у дна зоны проводимости.

$$N = n_0 = \int_{E_C}^{\infty} \frac{A\sqrt{E - E_C}}{e^{\frac{E - E_F}{kT}} + 1} dE. \quad (2.8)$$

Этот интеграл в общем случае не имеет аналитического решения. Однако для собственных полупроводников, а также для примесных с относительно невысокой концентрацией доноров и акцепторов ( $< 10^{18} \text{ см}^{-3}$ ), т. е. для невырожденных полупроводников, можно получить аналитическое выражение для  $n_0$ .

Для таких полупроводников величина  $E - E_F$  составляет как минимум десятые доли эВ, в то время как при комнатной температуре  $kT = 0,026$  эВ (1 электронвольт (эВ) равен  $1,6 \cdot 10^{-19}$  Дж). Поэтому единицей в знаменателе можно пренебречь, и тогда интеграл преобразуется к виду:

$$n_0 = \frac{4\pi}{h^3} (2m^*)^{\frac{3}{2}} \int_{E_C}^{\infty} \sqrt{E - E_C} \cdot e^{-\frac{E - E_F}{kT}} dE. \quad (2.9)$$

Это упрощение означает, что электроны в зоне проводимости собственных и примесных невырожденных полупроводников с очень хорошим приближением подчиняются классической статистике Максвелла-Больцмана.

Введя новую переменную  $x = \frac{E - E_C}{kT}$ , получаем выражение, в котором имеется табличный интеграл, равный  $\frac{\sqrt{\pi}}{2}$ :

$$n_0 = \frac{4\pi}{h^3} (2m^*kT)^{\frac{3}{2}} \cdot e^{-\frac{E_C-E_F}{kT}} \int_0^{\infty} \sqrt{x} \cdot e^{-x} dx.$$

Окончательно, имеем:

$$n_0 = \frac{2}{h^3} (2\pi m^*kT)^{\frac{3}{2}} \cdot e^{-\frac{E_C-E_F}{kT}} = N_C e^{-\frac{E_C-E_F}{kT}}. \quad (2.10)$$

Коэффициент при экспоненте  $N_C = \frac{2}{h^3} (2\pi m_n^*kT)^{\frac{3}{2}}$  — эффективная плотность состояний электронов в зоне проводимости. При условии  $E_F \rightarrow E_C$  величина  $N_C$  — это максимально возможная концентрация электронов в зоне проводимости невырожденного полупроводника. Следовательно, экспонента  $e^{-\frac{E_C-E_F}{kT}}$  имеет смысл вероятности нахождения свободных электронов в зоне проводимости.

Для дырок в валентной зоне справедливы те же упрощения, и их концентрация в невырожденных полупроводниках описывается выражением, аналогичным (2.10):

$$p_0 = N_V e^{-\frac{E_F-E_V}{kT}}; \quad (2.11)$$

где  $N_V = \frac{2}{h^3} (2\pi m_p^*kT)^{\frac{3}{2}}$  — эффективная плотность состояний дырок в валентной зоне,  $E_V$  — уровень, соответствующий началу отсчета энергии для дырок. Физический смысл  $N_V$  и экспоненциального множителя тот же, что и для электронов зоны проводимости:  $N_V$  — максимально возможная концентрация дырок;  $e^{-\frac{E_F-E_V}{kT}}$  — вероятность нахождения дырок в валентной зоне.

Введем следующие обозначения:  $n_n, p_p$  — концентрации основных носителей,  $p_n, n_p$  — концентрации неосновных носителей в  $n$ - и  $p$ -полупроводнике соответственно. Из предпосылок вывода формул (2.10) и (2.11) следует, что формула (2.10) описывает также концентрации электронов  $n_n$  и  $n_p$ , а формула (2.11) — концентрации дырок  $p_p$  и  $p_n$  в примесных невырожденных полупроводниках.

Собственная концентрация подвижных носителей  $n_i$  является важным параметром полупроводника, определяющим не только уровень его проводимости, но и, как будет показано ниже, взаимосвязь концентраций основных и неосновных носителей в примесных полупроводниках.

В собственных и компенсированных полупроводниках число электронов проводимости равно числу дырок:

$$n_0 = p_0 = n_i, \text{ или } n_i^2 = n_0 \cdot p_0. \quad (2.12)$$

Подставив сюда выражения (2.10) и (2.11) и, учитывая, что  $E_C - E_V = \Delta E_z$ , получим:

$$n_i = \sqrt{n_0 p_0} = \sqrt{N_C N_V} \cdot e^{-\frac{\Delta E_z}{2kT}}. \quad (2.13)$$

Отсюда видно, что, как и следовало ожидать, концентрация носителей в собственном полупроводнике сильно зависит от двух параметров: ширины запрещенной зоны  $\Delta E_z$  и температуры, причем температурная зависимость определяется тем, что величина  $T$  стоит в знаменателе экспоненты.

Собственная концентрация подвижных носителей в полупроводниках при температуре 300 К является табличной величиной; ее значения  $n_i$  для некоторых полупроводников приведены в табл. 2.2.

Таблица 2.2

**Взаимосвязь ширины запрещенной зоны и собственной концентрации подвижных носителей в полупроводниках**

Полупроводник	$\Delta E_z$ , эВ	$n_i$ , см <sup>-3</sup>
Ge	0,66	$2,4 \cdot 10^{13}$
Si	1,12	$1,4 \cdot 10^{10}$
GaAs	1,42	$1,8 \cdot 10^6$

Из данных табл. 2.2 видно, что увеличение ширины запрещенной зоны чуть больше, чем в 2 раза (Ge и GaAs) приводит к возрастанию собственной концентрации подвижных носителей на 7 порядков.

## 2.7. УРОВЕНЬ ФЕРМИ НА ЭНЕРГЕТИЧЕСКОЙ ДИАГРАММЕ СОБСТВЕННОГО ПОЛУПРОВОДНИКА

Концентрации дырок и электронов в собственном или компенсированном полупроводнике одинаковы:  $n_o = p_o$ :

$$N_C e^{\frac{E_{F_i} - E_C}{kT}} = N_V e^{\frac{E_V - E_{F_i}}{kT}}, \quad (2.14)$$

откуда следует, что

$$E_{F_i} = \frac{E_C + E_V}{2} + \frac{kT}{2} \ln \frac{N_V}{N_C}, \quad (2.15)$$

или,

$$E_{F_i} = \frac{E_C + E_V}{2} + \frac{3}{4} kT \ln \frac{m_p^*}{m_n^*}. \quad (2.16)$$

Это справедливо практически для любых температур, так как эффективные массы электронов проводимости и дырок в нулевом приближении примерно одинаковы:  $m_p^* \sim m_n^*$ . Таким образом,

$$E_{F_i} \cong \frac{E_V + E_C}{2}. \quad (2.17)$$

Можно считать с хорошим приближением, что уровень Ферми в собственных полупроводниках при любой температуре находится по середине запрещенной зоны.

## 2.8. ВЗАИМОСВЯЗЬ КОНЦЕНТРАЦИЙ ПОДВИЖНЫХ НОСИТЕЛЕЙ И УРОВЕНЬ ФЕРМИ В ПРИМЕСНЫХ ПОЛУПРОВОДНИКАХ

Для примесного невырожденного полупроводника, как следует из соотношений (2.10), (2.11) и (2.13), произведение равновесных концентраций основных и неосновных носителей оказывается всегда равным квадрату собственной концентрации носителей в полупроводнике без примесей:

$$n_n p_n = p_p n_p = n_i^2. \quad (2.18)$$

Это соотношение не зависит от концентрации примесей в полупроводнике.

Рассмотрим причину этого в рамках зонной модели. Возрастающие концентрации основных носителей, например, дырок при

увеличении количества акцепторов, приводит к тому, что увеличивается количество свободных энергетических уровней электронов у потолка валентной зоны. Следовательно, для появления в зоне проводимости неосновных носителей (электронов) необходимо сообщить связанным валентным электронам энергию, превышающую ширину запрещенной зоны в тем большей степени, чем выше уровень легирования  $p$ -полупроводника. Поэтому концентрация неосновных носителей при увеличении концентрации примесей (и, соответственно, основных носителей) будет пропорционально уменьшаться.

Для нахождения положения уровня Ферми в примесных полупроводниках необходимо решить уравнения, соответствующие условию равенства концентраций основных носителей ( $n_n, p_p$ ) и равновесных концентраций ( $n_0, p_0$ ). Формулы получаются очень громоздкими, но для наиболее важных случаев возможны упрощения. Так, для области низких температур, когда еще не все примесные атомы превратились в ионы (см. разделы 2.2 и 2.3), получаются следующие формулы:

Для полупроводника  $n$ -типа:

$$E_{F_n} = \frac{E_C + E_D}{2} - \frac{kT}{2} \ln 2 \frac{N_C}{N_D}, \quad (2.19)$$

для полупроводника  $p$ -типа:

$$E_{F_p} = \frac{E_V + E_A}{2} + \frac{kT}{2} \ln 2 \frac{N_V}{N_A}. \quad (2.20)$$

Из этих соотношений следует, что при температуре абсолютного нуля уровень Ферми в полупроводнике  $n$ -типа располагается посередине между дном зоны проводимости и донорным уровнем, а в  $p$ -полупроводнике посередине между потолком валентной зоны и акцепторным уровнем:

$$E_{F_n} = \frac{E_C + E_D}{2}, \quad E_{F_p} = \frac{E_V + E_A}{2} \quad (T = 0 \text{ К}). \quad (2.21)$$

Изменение положения уровня Ферми при увеличении температуры от абсолютного нуля до относительно невысоких значений ( $\leq 100 \text{ К}$ ) описывается формулами (2.19) и (2.20). Проследим эти из-

менения с учетом того, что логарифмы в этих формулах всегда положительные.

С ростом температуры уровень Ферми в *n*-полупроводнике смещается вниз. Известно, что при комнатной температуре, когда все атомы донорной примеси ионизованы, он располагается в верхней части запрещенной зоны ниже донорного уровня. Полная ионизация доноров означает, что при комнатной температуре донорный уровень не занят электронами.

В *p*-полупроводнике при тех же условиях уровень Ферми поднимается вверх, и при комнатной температуре находится в нижней части запрещенной зоны выше акцепторного уровня. При этом все электроны, перешедшие от атомов основного вещества в связи к примесным атомам и превратившие их в отрицательные ионы, находятся на акцепторном уровне.

Общий характер изменения положения уровня Ферми показан на рис. 2.10.

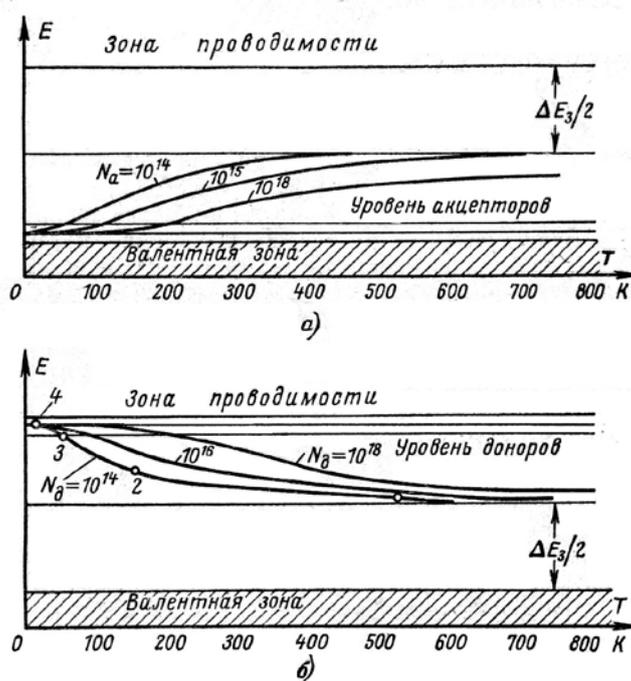


Рис. 2.10. Изменение положения уровня Ферми в акцепторном (а) и донорном (б) полупроводниках с ростом температуры для различных концентраций примесей

При относительно высоких температурах (500-600 К) уровень Ферми в полупроводнике  $n$ -типа  $E_{Fn}$  снижается, а уровень Ферми в полупроводнике  $p$ -типа  $E_{Fp}$  поднимается к середине запрещенной зоны (см. рис. 2.10). Это связано с тем, что при таких условиях из-за резкого увеличения концентрации собственных электронно-дырочных пар (т. е. из-за возрастания скорости тепловой генерации) в примесных полупроводниках выравниваются концентрации электронов и дырок, вследствие чего начинает преобладать собственная проводимость.

### ***Вопросы к разделу 2***

1. Поясните механизм собственной проводимости полупроводника. Какова зависимость проводимости полупроводника от температуры?
2. Каким образом собственный полупроводник можно преобразовать в полупроводник  $n$ -типа? Каковы требования к сорту атомов примеси?
3. Каким образом собственный полупроводник можно преобразовать в полупроводник  $p$ -типа? Каковы требования к сорту атомов примеси?
4. Объясните, каким образом обеспечивается дырочная проводимость полупроводников.
5. Каковы энергетические диаграммы полупроводников  $n$ - и  $p$ -типа? По какой причине не наблюдается расщепление примесных уровней в энергетические зоны при невысоких концентрациях примеси?
6. При каком условии образуются энергетические зоны примесей?
7. Каков физический смысл уровня Ферми на энергетической диаграмме металла?
8. От каких факторов зависит равновесная концентрация подвижных носителей в полупроводниках?
9. Каково положение уровня Ферми на энергетических диаграммах собственного полупроводника и примесных полупроводников?
10. Как и почему изменяется положение уровня Ферми примесных полупроводников при повышении температуры?

## 3. ДВИЖЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

### 3.1. ХАОТИЧЕСКОЕ ДВИЖЕНИЕ

В отсутствие электрического поля подвижные носители в кристалле находятся в хаотическом движении. Оно характеризуется тем, что направления перемещений частиц равновероятны, то есть направленное движение отсутствует. Распределение их по энергиям определяется законами квантовой статистики Ферми-Дирака (см. соотношение (2.4)). В частном случае полупроводника с невысокой концентрацией примесей оно преобразуется в классическое энергетическое распределение Максвелла-Больцмана (2.6).

### 3.2. ДРЕЙФ ПОДВИЖНЫХ НОСИТЕЛЕЙ ЗАРЯДА. ДРЕЙФОВЫЕ ТОКИ

Направленное движение носителей заряда под действием электрического поля называется дрейфом. Электроны, получая ускорение в электрическом поле, приобретают на длине свободного пробега дополнительную кинетическую энергию. В результате, совокупность хаотически движущихся электронов будет перемещаться в направлении, противоположном вектору напряженности электрического поля  $\vec{\varepsilon}$ . Такое движение электронов в полупроводнике приводит к появлению дрейфового тока:

$$J_{n\varepsilon} = env_n = en \frac{v_n}{\varepsilon} \varepsilon = en\mu_n \varepsilon = \sigma_n \varepsilon, \quad (3.1)$$

где  $\sigma_n$  — удельная проводимость,  $\mu_n$  — подвижность электронов, т. е. величина, равная средней скорости их движения в электрическом поле с напряженностью, равной единице: ( $\mu = \frac{v_n}{\varepsilon}$ ).

Аналогичным образом появится дырочная составляющая дрейфового тока:

$$J_{p\varepsilon} = \sigma_p \varepsilon = e p \mu_p \varepsilon . \quad (3.2)$$

Знаки в правых частях (3.1) и (3.2) одинаковы, так как противоположные направления движения разноименно заряженных частиц соответствуют току одного направления.

Полная плотность дрейфового тока при наличии свободных электронов и дырок равна:

$$J_\varepsilon = J_{n\varepsilon} + J_{p\varepsilon} = e(n\mu_n + p\mu_p)\varepsilon . \quad (3.3)$$

### 3.3. ДИФФУЗИЯ ПОДВИЖНЫХ НОСИТЕЛЕЙ ЗАРЯДА. ДИФФУЗИОННЫЕ ТОКИ

*Диффузия* — это направленное движение частиц (нейтральных или заряженных) из-за их неравномерного распределения по объему, т. е. из-за наличия градиента концентрации. Движущей силой диффузии является хаотическое тепловое движение частиц.

Поток частиц через плоскость раздела двух элементарных объемов с концентрациями  $N_1$  и  $N_2$ :

$$\begin{aligned} \Phi &= -D \frac{dN}{dx}; \\ \frac{dN}{dx} &= \frac{N_1 - N_2}{2\ell_{cp}}, \end{aligned} \quad (3.4)$$

где  $D = \frac{1}{3} v_{cp} \ell_{cp}$  — коэффициент диффузии,  $\ell_{cp}$  — средняя длина свободного пробега частиц.

Тогда плотности диффузионного тока электронов и дырок:

$$J_{n\text{Диф}} = eD_n \frac{dn}{dx}; \quad J_{p\text{Диф}} = -eD_p \frac{dp}{dx}. \quad (3.5)$$

Для полупроводника, содержащего свободные электроны и дырки, плотность суммарного диффузионного тока:

$$J_{\text{Диф}} = J_{n\text{Диф}} + J_{p\text{Диф}} = e \left( D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right). \quad (3.6)$$

Физической причиной этих токов является неравномерность распределения концентраций электронов и дырок по объему полупроводника. Эти токи существуют в отсутствие внешнего электрического поля. Диффузионные токи электронов и дырок могут быть как противоположно направленными (при одинаковом направлении градиентов их концентраций), так и сонаправленными (если направления градиентов противоположны).

Таким образом, движение электронов и дырок в полупроводниках обусловлено двумя процессами: диффузией под действием градиента их концентрации и дрейфом под действием электрического поля. Полный ток состоит из четырех составляющих, и его плотность можно представить следующим образом:

$$J_{n\text{Диф}} + J_{p\text{Диф}} + J_{n\varepsilon} + J_{p\varepsilon}. \quad (3.7)$$

### 3.4. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

В общем случае поведение неравновесных подвижных носителей заряда описывается уравнением непрерывности. Рассмотрим упрощенно его содержание. Пусть в полупроводнике  $p$ -типа в направлении оси  $x$  имеется градиент концентрации электронов  $dn/dx > 0$  и действует постоянное электрическое поле напряженностью  $E_x$ .

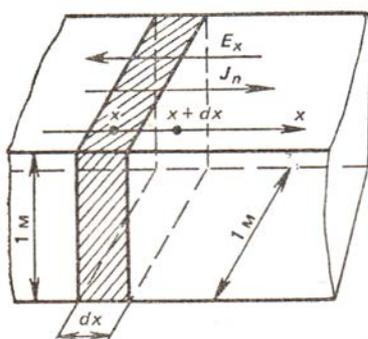


Рис. 3.1. К выводу уравнения непрерывности.  
 $E_x$  — напряженность электрического поля,  
 $J_n$  — суммарная плотность тока электронов

Выделим в полупроводнике слой толщиной  $dx$ , расположенный перпендикулярно оси  $x$ , с площадью сечения  $1 \text{ м}^2$  (рис. 3.1). Таким

образом, объем этого слоя равен  $dx$ , и задача сводится к одномерному случаю.

Концентрации носителей заряда зависят от двух факторов: координаты  $x$  и времени  $t$ . Обозначим концентрацию электронов в слое в момент времени  $t$  через  $n(x, t)$ , а в момент времени  $t + dt$  — через  $n(x, t + dt)$ . Изменение числа электронов в слое за время  $dt$  составит:

$$[n(x, t + dt) - n(x, t)]dx = \frac{\partial n}{\partial t} dt dx. \quad (3.8)$$

Оно вызывается протекающими в слое процессами генерации и рекомбинации, а также диффузией и дрейфом носителей заряда. В результате генерации за время  $dt$  в слое объемом  $dx$  возникает  $g dx dt$  электронов ( $g$  — скорость генерации).

Вследствие рекомбинации в единицу времени в единице объема полупроводника исчезнет  $r = -\frac{n - n_0}{\tau_n}$  электронов. Убыль электронов за время  $dt$  в объеме  $dx$  составит  $-\frac{n - n_0}{\tau_n} dx dt$ .

Наличие градиента концентрации и внешнего электрического поля в направлении оси  $x$  приведет к тому, что поток электронов  $\Phi_n(x)$ , втекающий в слой  $dx$ , будет больше потока  $\Phi_n(x + dx)$  на выходе из слоя. Изменение число электронов за время  $dt$ , вызванное различием этих потоков, можно описать соотношением

$$[\Phi_n(x) - \Phi_n(x + dx)]dt = -\frac{\partial \Phi_n}{\partial x} dx dt. \quad (3.9)$$

Полное изменение числа электронов в слое за время  $dt$  составит

$$\frac{\partial n}{\partial t} dt dx = \left( -\frac{\partial J_n}{\partial x} + g - \frac{n - n_0}{\tau_n} \right) dt dx. \quad (3.10)$$

Сокращая обе части уравнения (3.10) на  $dt dx$ , получим уравнение непрерывности для избыточных электронов полупроводнике  $n$ -типа:

$$\frac{\partial n}{\partial t} = -\frac{\partial \Phi_n}{\partial x} + g - \frac{n - n_0}{\tau_n}. \quad (3.11)$$

Аналогичное уравнение непрерывности для избыточных дырок в полупроводнике  $p$ -типа:

$$\frac{\partial p}{\partial t} = -\frac{\partial \Phi_p}{\partial x} + g - \frac{p - p_0}{\tau_p}. \quad (3.12)$$

Слагаемые  $-\frac{\partial \Phi_n}{\partial x}$  и  $-\frac{\partial \Phi_p}{\partial x}$  в правых частях уравнений (3.11) и (3.12) отражают влияние диффузии и дрейфа носителей заряда. Рассмотрим более подробно эти процессы и получим уравнения непрерывности, в которых они будут отражены явным образом.

Поток электронов  $\Phi_n$  можно выразить через плотность тока:  $\Phi_n = -J_n/e$ , где  $e$  — заряд электрона. Плотность дрейфового тока электронов:  $J_n = en\mu_n E_x$ , где  $\mu_n$  — подвижность электронов,  $n$  — их концентрация.

Плотность диффузионного тока электронов пропорциональна градиенту концентрации:

$$J_{n \text{ Диф}} = eD_n \frac{\partial n}{\partial x},$$

где  $D_n$  — коэффициент диффузии.

Плотность тока, обусловленная диффузией и дрейфом электронов, есть сумма двух составляющих:

$$J_n = J_{n_E} + J_{n \text{ Диф}} = e \left( n\mu_n E_x + D_n \frac{\partial n}{\partial x} \right). \quad (3.13)$$

Соответственно, плотность тока, обусловленная диффузией и дрейфом дырок в полупроводнике  $p$ -типа:

$$J_p = e \left( p\mu_p E_x - D_p \frac{\partial p}{\partial x} \right). \quad (3.14)$$

Знак «минус» у диффузионной составляющей плотности тока дырок указывает на то, что направление диффузионного тока противоположно градиенту их концентрации. Заменяя в (3.11) и (3.12) потоки  $\Phi_n$  и  $\Phi_p$  плотностями токов (формулы (3.13) и (3.14)), получим уравнение непрерывности в виде:

$$\text{Для электронов: } \frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} + E_x \mu_n \frac{\partial n}{\partial x} + g - \frac{n - n_0}{\tau_n}; \quad (3.15)$$

$$\text{Для дырок: } \frac{\partial p}{\partial t} = D_p \frac{\partial^2 p}{\partial x^2} - E_x \mu_p \frac{\partial p}{\partial x} + g - \frac{p - p_0}{\tau_p}. \quad (3.16)$$

При учете координатной зависимости напряженности электрического поля  $E_x(x)$  уравнения непрерывности для одномерного случая приобретают окончательный вид:

$$\frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} + E_x \mu_n \frac{\partial n}{\partial x} + n \mu_n \frac{\partial E_x}{\partial x} + g - \frac{n - n_0}{\tau_n}; \quad (3.17)$$

$$\frac{\partial p}{\partial t} = D_p \frac{\partial^2 p}{\partial x^2} - E_x \mu_p \frac{\partial p}{\partial x} - p \mu_p \frac{\partial E_x}{\partial x} + g - \frac{p - p_0}{\tau_p}. \quad (3.18)$$

### 3.5. ВЗАИМОСВЯЗЬ ДИФФУЗИИ И ДРЕЙФА ЗАРЯЖЕННЫХ ЧАСТИЦ

Диффузионное и дрейфовое движение заряженных частиц, и в частности электронов и дырок в полупроводнике, взаимосвязано. Диффузионные процессы выравнивания концентрации сопровождаются перераспределением зарядов, следовательно, при этом возникает электрическое поле  $\vec{\varepsilon}$ . Это вызовет дрейфовое движение частиц в направлении, противоположном диффузионному движению. В равновесном состоянии плотности этих потоков равны. Например, для электронов:

$$e n \mu_n \varepsilon = e D_n \frac{dn}{dx}. \quad (3.19)$$

Электроны, находящиеся в электрическом поле, обладают потенциальной энергией, равной  $-e\varphi(x)$ , где  $\varphi(x)$  — потенциал поля в точке  $x$ . Для невырожденных полупроводников, в соответствии с законом Больцмановского распределения в условиях равновесия, концентрация электронов в области, охватывающей точку  $x$ , равна:

$$n(x) = C \exp\left(\frac{e\varphi(x)}{kT}\right), \quad (3.20)$$

где  $C$  — коэффициент пропорциональности. Найдем производную от концентрации по  $x$  с учетом того, что  $\mathcal{E} = -d\varphi(x)/dx$ :

$$\frac{dn}{dx} = C \exp\left(\frac{e\varphi(x)}{kT}\right) \frac{e}{kT} \mathcal{E}. \quad (3.21)$$

Подставив выражение (3.21) в (3.19), для коэффициента диффузии электронов и дырок получим:

$$D_n = \frac{kT}{e} \mu_n, \quad D_p = \frac{kT}{e} \mu_p. \quad (3.22)$$

Эти выражения называются соотношениями Эйнштейна. Они связывают коэффициенты диффузии носителей заряда с их подвижностью и математически отражают взаимосвязь диффузии и дрейфа заряженных частиц.

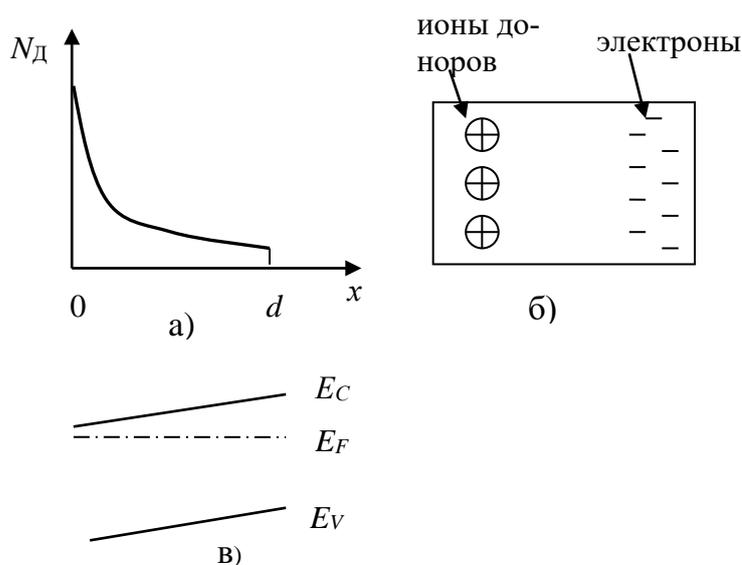


Рис. 3.2. Полупроводник  $n$ -типа с неравномерным по объему распределением примеси. *а)* изменение концентрации доноров по глубине; *б)* «встроенное» электрическое поле в полупроводнике, обусловленное разделением зарядов после установления равновесия; *в)* энергетическая диаграмма  $n$ -полупроводника для такого случая

Примером проявления такой взаимосвязи является полупроводник с неравномерным распределением примеси по объему. Пусть в полупроводнике  $n$ -типа создано неравномерное по глубине

распределение доноров, а значит и электронов проводимости (см. рис. 3.2, а). Тогда электроны начнут диффундировать в область с меньшей концентрацией донорной примеси, и в полупроводнике произойдет разделение зарядов. Область, куда пришли дополнительные электроны, приобретает избыточный отрицательный заряд, а на месте ушедших электронов остается некомпенсированный положительный заряд ионов доноров (рис. 3.2, б).

Возникает диффузионное электрическое поле и встречный дрейфовый ток электронов. Когда суммарный ток окажется равным нулю, установится равновесное состояние; при этом в полупроводнике будет существовать так называемое «встроенное» электрическое поле.

Энергетическая диаграмма такого полупроводника показана на рис. 3.2, в. Она содержит наклонные зоны, отражающие наличие электрического поля в кристалле и горизонтальный уровень Ферми, свойственный равновесным системам. Подвижные носители, попадающие извне в такой полупроводник (например, за счет диффузии), будут ускоряться или замедляться в зависимости от знака их заряда.

### 3.6. ДВИЖЕНИЕ НОСИТЕЛЕЙ В СИЛЬНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

В электрических полях с величиной напряженности  $\varepsilon \leq 10^3$  В/см носители на длине свободного пробега приобретают относительно малую энергию. Поэтому их распределение по энергетическим уровням соответствует распределению при данной температуре кристаллической решетки. Иными словами дрейфовая скорость частиц в несильных электрических полях много меньше средней тепловой скорости:

$$(v_{др} \ll v_{ср} = \sqrt{\frac{3kT}{m}}).$$

Следовательно, в слабых электрических полях подвижность не зависит от напряженности поля, т. е. полупроводник подчиняется закону Ома (см. начальный участок кривой на рис. 3.3).

При  $\varepsilon > 10^3 \text{ В/см}$  средняя энергия носителей уже соответствует температуре, превышающей температуру кристаллической решетки. Число столкновений с атомами решетки увеличивается; это приводит к замедлению нарастания дрейфовой скорости и к уменьшению подвижности при увеличении напряженности электрического поля. Поэтому при  $10^3 \text{ В/см} < \varepsilon < 10^4 \text{ В/см}$  полупроводник обладает свойством нелинейного сопротивления.

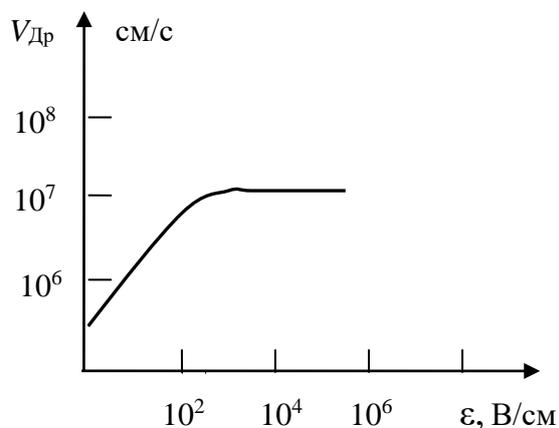


Рис. 3.3. Зависимость дрейфовой скорости подвижных носителей от напряженности электрического поля

При  $\varepsilon > 10^4 \text{ В/см}$  дрейфовая скорость  $v_{Др}$  оказывается не зависящей от напряженности электрического поля  $\varepsilon$ . В полупроводнике наблюдается насыщение тока, и соответствующая напряженность поля называется критической. Для большинства полупроводников  $\varepsilon_{кр} \sim 10^4 \text{ В/см}$ .

### 3.7. ЭФФЕКТ ХОЛЛА

В последнее время значительное развитие получила полупроводниковая магнитоэлектроника, основанная на гальваномагнитных явлениях. Эти явления представляют собой результат воздействия магнитного поля на электрические свойства полупроводников, по которым протекает электрический ток.

Важнейшее из гальваномагнитных явлений — эффект Холла. Он состоит в том, что при протекании тока в полупроводнике возникает поперечная разность потенциалов (напряжение, или ЭДС Холла),

если на этот полупроводник действует магнитное поле, вектор индукции которого перпендикулярен направлению тока.

Эффект был обнаружен в 1879 году американским физиком Э. Холлом в металле (пластинках золота) и долгое время не находил практического применения. Положение резко изменилось с середины 20 века, когда началось бурное развитие полупроводниковой электроники.

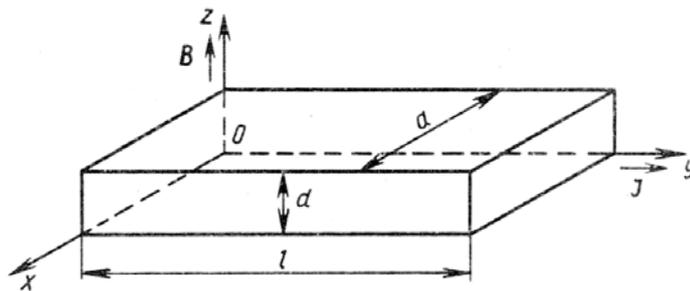


Рис. 3.4. Полупроводниковый образец, в котором наблюдается эффект Холла

Рис. 3.4 поясняет возникновение поперечной разности потенциалов. Через полупроводник протекает электрический ток плотностью  $J$ . Образец помещен в слабое постоянное магнитное поле, вектор индукции которого  $B$  перпендикулярен вектору  $J$ . Электрическое поле возникает в направлении оси  $x$ , т. е. вектор напряженности поля перпендикулярен плоскости, заданной векторами  $B$  и  $J$ .

Пусть образцом является идеальный собственный полупроводник с равными концентрациями электронов и дырок. При подключении напряжения ток будет обусловлен их встречным движением, следовательно, сила Лоренца, действующая со стороны магнитного поля на движущиеся заряды, будет направлена одинаково для обоих типов носителей — в сторону положительной оси  $x$  (рис. 3.4). В результате к одной грани пластинки ( $ld$ ) придут носители разных знаков в одинаковом количестве, накопления заряда одного знака не произойдет, и ЭДС Холла будет равна нулю. Для реальных беспримесных полупроводников возникающая ЭДС будет зависеть от соотношения

концентраций и подвижностей электронов и дырок. В случае примесных полупроводников эффект Холла проявляется очень отчетливо, и знак ЭДС Холла будет зависеть от типа их проводимости.

Возьмем для определенности образец полупроводника  $n$ -типа и рассмотрим процессы, происходящие в нем при указанных выше условиях (рис. 3.4). Сила Лоренца отклоняет движущиеся электроны к грани ( $ld$ ):

$$F_x = ev_y B_z, \quad (3.23)$$

где  $v_y$  — дрейфовая скорость электронов.

Так как образец вдоль оси  $x$  изолирован, и ток вдоль этой оси не может быть замкнутым, то для уравнивания силы (3.23) на противоположной грани пластины должны накапливаться положительные заряды. В результате возникает электрическое поле  $\mathbf{\epsilon}_x$ , перпендикулярное плоскости, в которой лежат векторы  $\mathbf{B}$  и  $\mathbf{J}$ . Сила, действующая на электроны со стороны поля  $\mathbf{\epsilon}_x$ , уравнивает силу Лоренца (3.23):

$$ev_y B_z = e\epsilon_x. \quad (3.24)$$

Плотность тока через образец в направлении оси  $y$ :  $J_y = ev_y n$ , следовательно, можно записать:

$$e\epsilon_x = \frac{J_y B_z}{n}, \quad (3.25)$$

где  $n$  — концентрация электронов.

Отсюда выводится постоянная Холла  $R_H$ :

$$R_H = \frac{\epsilon_x}{J_y B_z} = \frac{1}{ne}. \quad (3.26)$$

Видно, что знак постоянной Холла совпадает со знаком носителей заряда: для электронов она отрицательна, для дырок — положительна.

Ток электронов вдоль оси  $y$ :  $I = en\mu_n \epsilon_y ad$ , следовательно

$$\epsilon_y = \frac{I}{en\mu_n nad}. \quad (3.27)$$

Заменив в выражении (3.25)  $v_y$  на произведение  $\mu_n \varepsilon_y$ , можно записать:

$$\varepsilon_x = \varepsilon_y \mu_n B_z = \frac{I \mu_n B_z}{ade \mu_n n} = \frac{I B_z}{ad n e}. \quad (3.28)$$

Возникающее напряжение между боковыми гранями пластины называется ЭДС Холла:

$$U_H = -a \varepsilon_x = -\frac{I B_z}{d n e}. \quad (3.29)$$

Видно, что знак ЭДС Холла определяется типом проводимости полупроводника, а ее величина прямо пропорциональна величине магнитной индукции и зависит также от концентрации носителей заряда и геометрии образца.

Применения эффекта Холла в электронике чрезвычайно разнообразны. Датчики или преобразователи Холла используются для измерения магнитной индукции, концентрации и подвижности носителей заряда, определения типа проводимости примесных полупроводников. Они также применяются для разнообразных бесконтактных электрических измерений и выработки управляющих электрических сигналов в различных цепях.

### 3.8. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ. ЭФФЕКТЫ ЗЕЕБЕКА И ПЕЛЬТЬЕ

Под термоэлектрическими явлениями понимается прямое преобразование тепловой энергии в электрическую в проводниках (металлах, полупроводниках), а также обратные явления нагревания и охлаждения спаев двух проводников проходящим током.

**Эффект Зеебека** состоит в том, что в замкнутой цепи, состоящей из разнородных проводников, возникает электродвижущая сила (термоЭДС), если места контактов поддерживают при разных температурах. На основе этого эффекта в металлических проводниках работают всевозможные термопары — датчики температуры.

Установлено, что абсолютные значения всех термоэлектрических коэффициентов растут с уменьшением концентрации носителей;

поэтому в полупроводниках они в десятки и сотни раз больше, чем в металлах и сплавах. Именно поэтому термоэлектрические полупроводниковые материалы нашли в настоящее время широкое применение для создания различных приборов, принцип действия которых основан на термоэлектрических эффектах.

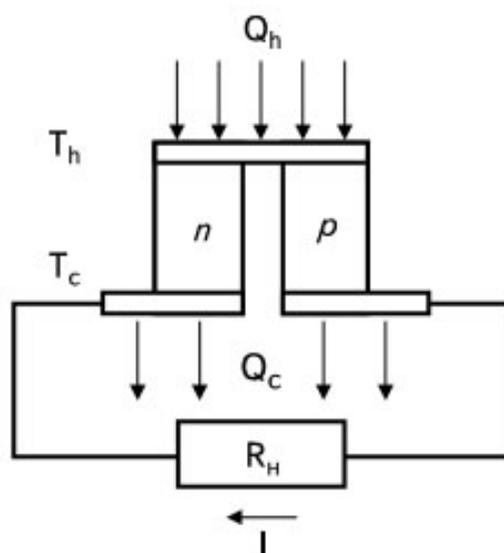


Рис. 3.5. Устройство простейшего полупроводникового термоэлемента.

1 — отрицательная ветвь ( $n$ -тип),

2 — положительная ветвь ( $p$ -тип),

3 — металлическая пластина («горячий спай»).

$Q_h$  — тепло, подводимое к горячему спаю,

$Q_c$  — тепло отводимое от холодного спаю при температуре  $T_c$

Рис. 3.5 поясняет принцип термоэлектрического преобразования тепловой энергии с применением полупроводников  $p$ - и  $n$ -типа. Горячий спай двух полупроводников поддерживается при температуре  $T_h$ . В разрыв холодного спаю, находящегося при температуре  $T_c$ , подключен резистор нагрузки  $R_n$ , через который протекает термоток в направлении, указанном на рисунке, и соответственно, возникает падение напряжения. Если цепь, соединяющую холодные концы  $n$ - и  $p$ -областей разорвать, то между ними установится максимально возможная разность потенциалов — термоЭДС. При замыкании холодных концов через гальванометр потечет наибольший возможный ток короткого замыкания.

Для полупроводников основной причиной, вызывающей усиление термотока в эффекте Зеебека, является диффузия носителей. В дырочных полупроводниках на холодном контакте скапливаются дырки, а на горячем остается некомпенсированный отрицательный заряд. В полупроводниках  $n$ -типа наблюдается обратная картина: на холодном контакте скапливаются электроны, а на горячем — положительные ионы доноров.

В термоэлементе, состоящем из дырочного и электронного полупроводников (рис. 3.5), термоЭДС складываются. Кроме того, если вдоль проводника существует градиент температур, то подвижные носители на горячем конце приобретают более высокие энергии и скорости, чем на холодном; в полупроводниках в дополнение к этому концентрация подвижных носителей растет с температурой. В результате возникает диффузия этих носителей от горячих концов к холодным. Накопление заряда продолжается до тех пор, пока возникшая разность потенциалов не вызовет поток носителей в обратном направлении. В конце концов, установится равновесие, и возникнет основная составляющая термоЭДС — диффузионная.

В полупроводниках со смешанной проводимостью к холодному контакту диффундируют и электроны и дырки, и их заряды взаимно компенсируются. Если концентрации и подвижности электронов и дырок равны, то термоЭДС равна нулю.

Дополнительный вклад в величину термоЭДС вносит различие контактных разностей потенциалов на горячем и холодном спае. Концентрация основных носителей на горячем спае меньше, чем на холодном, а концентрации неосновных носителей на несколько порядков больше. Следовательно, в соответствии с закономерностью, определяемой формулами (6.6) и (6.7) для контактной разности потенциалов на  $p$ - $n$ -переходе, ее величина на горячем спае будет меньше, чем на холодном. Вследствие этого возникает добавочная разность потенциалов с полярностью, совпадающей с диффузионной термоЭДС.

Еще один физический механизм, дающий вклад в термоЭДС — увлечение подвижных носителей фононами — квазичастицами, которые являются квантами тепловой энергии. При наличии перепада температуры в  $p$ -ветви к холодному концу приходит дополнительное количество дырок, а в  $n$ -ветви — дополнительное количество электронов. В рамках обычных представлений можно сказать, что подвижные носители вытесняются из областей, где амплитуда колебаний атомов решетки относительно высока и движутся в область с меньшей температурой. Это явление дает заметный вклад в термоЭДС при низких температурах.

**Эффект Пельтье** — термоэлектрическое явление, обратное эффекту Зеебека: при пропускании электрического тока через контакт (спай) двух различных веществ (проводников или полупроводников) на контакте, помимо джоулева тепла, происходит выделение дополнительного тепла Пельтье  $Q_P$  при одном направлении тока и его поглощение при обратном направлении.

Величина выделяемого тепла  $dQ_P$  и его знак зависят от вида контактирующих веществ, и от протекшего через ветви термоэлемента заряда  $dq = Idt$ :

$$dQ_P = \pi_{12} Idt .$$

Здесь  $\pi_{12} = \pi_1 - \pi_2$  — коэффициент Пельтье для данного контакта, связанный с абсолютными коэффициентами Пельтье  $\pi_1$  и  $\pi_2$  контактирующих материалов. При этом считается, что ток идет от первого образца ко второму. При выделении тепла Пельтье имеем:  $Q_P > 0$ ,  $\pi_{12} > 0$ ,  $\pi_1 > \pi_2$ . При поглощении тепла Пельтье оно считается отрицательным и соответственно:  $Q_P < 0$ ,  $\pi_{12} < 0$ ,  $\pi_1 < \pi_2$ . Очевидно, что  $\pi_{12} = -\pi_{21}$ . Размерность коэффициента Пельтье  $[\pi]_{\text{СИ}} = \text{Дж/Кл} = \text{В}$ .

Вместо тепла Пельтье часто используют физическую величину, определяемую как тепловая энергия, ежесекундно выделяющаяся на контакте единичной площади. Эта величина, получившая название мощность тепловыделения, определяется формулой:  $w_P = \pi_{12} J$ , где  $J$  — плотность тока. Соответственно, размерность  $[w_P] = \text{Вт/м}^2$ .

Эффект Пельтье, как и все термоэлектрические явления, выражен особенно сильно в контактах, составленных из электронных ( $n$ -тип) и дырочных ( $p$ -тип) полупроводников.

Рассмотрим физический процесс, приводящий к выделению и поглощению тепла Пельтье в контактах, составленных из полупроводников  $p$ - и  $n$ -типа.

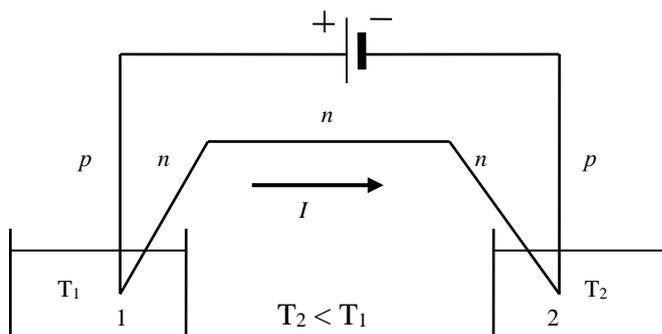


Рис. 3.6. Простейший термоэлемент для наблюдения эффекта Пельтье

Пусть ток в контакте идет от дырочного полупроводника к электронному (контакт 1, рис. 3.6). При этом электроны и дырки движутся навстречу друг другу в направлении спая и, встретившись, рекомбинируют. В результате рекомбинации освобождается энергия, которая выделяется в виде тепла, и температура контакта 1 повышается.

На контакте 2 на том же рисунке электроны и дырки движутся в разные стороны от спая. Убыль носителей тока в пограничной области восполняется за счет рождения новых электронно-дырочных пар. На образование таких пар требуется энергия, которая поставляется тепловыми колебаниями атомов решетки. Образующиеся электроны и дырки увлекаются в противоположные стороны от спая электрическим полем. Поэтому, пока через контакт 2 идет ток, непрерывно происходит тепловая генерация новых электронно-дырочных пар. В результате тепло в контакте 2 (рис. 3.6) поглощается, и, следовательно, этот спай будет охлаждаться. При смене полярности источника внешнего напряжения картина изменится на противоположную.

Рис. 3.7 поясняет использование полупроводников  $p$ - и  $n$ -типа в термоэлектрических холодильниках. Модуль Пельтье (рис. 3.7) представляет собой термоэлектрический элемент, состоящий из последовательно соединенных полупроводников  $p$ - и  $n$ -типа. Каждый из таких переходов имеет тепловой контакт с одним из двух радиаторов. В результате прохождения электрического тока указанной полярности образуется перепад температур между радиаторами модуля Пельтье: верхний радиатор работает как холодильник, нижний радиатор нагревается и служит для отвода тепла.

Объединение большого количества пар полупроводников  $p$ - и  $n$ -типа позволяет создавать эффективные охлаждающие элементы. Такие элементы называют полупроводниковыми холодильниками.

Типичный отдельный модуль холодильника обеспечивает значительный температурный перепад, который составляет несколько десятков градусов. При соответствующем принудительном охлаждении нагревающегося радиатора второй радиатор — холодильник — позволяет достичь отрицательных значений температур.

Для увеличения разности температур возможно каскадное включение термоэлектрических модулей Пельтье при обеспечении соответствующего охлаждения для отвода выделяющегося тепла. Это позволяет сравнительно простыми средствами получить значительный перепад температур и обеспечить эффективное охлаждение защищаемых элементов.

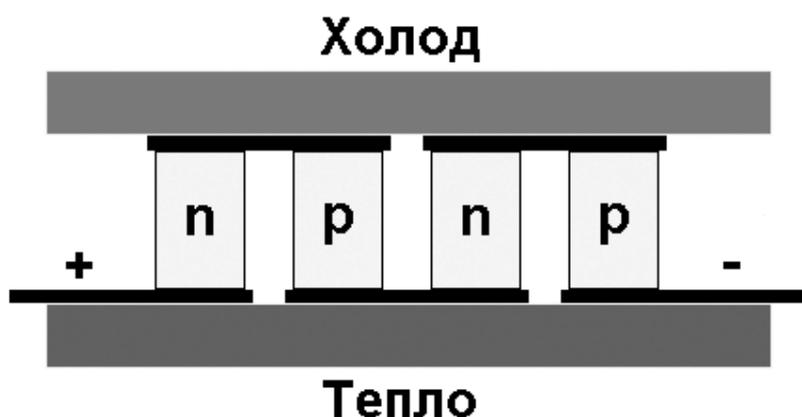


Рис. 3.7. Модуль Пельтье с тремя спаями

В качестве примера использования полупроводниковых термоэлектрических холодильников можно привести устройства охлаждения центральных процессоров мощных персональных компьютеров, так называемые «кулеры». В этих устройствах объединены вентиляторы, радиаторы для рассеивания выделяющегося тепла и модули Пельтье, у которых холодный радиатор (см. рис. 3.7) имеет очень хороший тепловой контакт с посадочным местом процессора. При достижении некоторой заранее установленной температуры процессора на модуль Пельтье подается напряжение, и холодильник обеспечивает поддержание безопасной температуры.

### ***Вопросы к разделу 3***

1. Каков механизм дрейфового движения подвижных носителей в полупроводниках?
2. Что является движущей силой диффузии, и какова физическая причина диффузионного движения подвижных носителей в полупроводниках?
3. Какова причина взаимосвязи диффузии и дрейфа подвижных носителей в полупроводниках?
4. Поясните механизм возникновения «встроенного» электрического поля в полупроводнике с неравномерным распределением примеси по объему.
5. Как изменяется дрейфовая скорость подвижных носителей при увеличении напряженности электрического поля?
6. В чем заключается физический механизм возникновения ЭДС Холла?
7. От каких физических параметров зависит знак и величина ЭДС Холла?
8. Какие физические механизмы обеспечивают появление ЭДС в термоэлементе на основе полупроводников  $p$ - и  $n$ -типа?
9. Поясните принцип действия полупроводникового термоэлектрического холодильника на основе эффекта Пельтье.

## 4. ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Процессы образования (генерации) и исчезновения (рекомбинации) подвижных носителей заряда в полупроводниках (электронов, дырок и электронно-дырочных пар) происходят непрерывно и в значительной степени определяют характеристики и параметры всевозможных полупроводниковых приборов.

Генерацией называют процесс образования подвижных носителей в полупроводниках при различных внешних воздействиях. Возникновение свободных носителей требует затраты энергии, необходимой для освобождения электронов из связанного состояния. Виды генерации определяются характером передачи энергии:

1. Тепловая генерация при достижении температуры выше некоторого критического значения. Принято считать, что при данной постоянной температуре в полупроводнике существует соответствующая равновесная концентрация подвижных носителей.

Перечисленные ниже виды генерации приводят к появлению неравновесных избыточных подвижных носителей заряда:

2. Ударная генерация или ударная ионизация: энергия передается в виде кинетической энергии движущихся частиц.
3. Полевая генерация (или ионизация) за счет энергии электрического поля.
4. Световая генерация за счет передачи связанным электронам энергии фотонов при облучении светом.

### 4.1. ТЕПЛОВАЯ ГЕНЕРАЦИЯ

Процесс тепловой генерации электронно-дырочных пар в собственном полупроводнике и тепловой генерации основных носителей в примесных полупроводниках обсуждался в разделе 2.

Рассмотрим более подробно, как происходит тепловая генерация подвижных носителей в полупроводнике  $n$ -типа при повышении

температуры от абсолютного нуля до температур, при которых уже происходит тепловая генерация собственных электронно-дырочных пар.

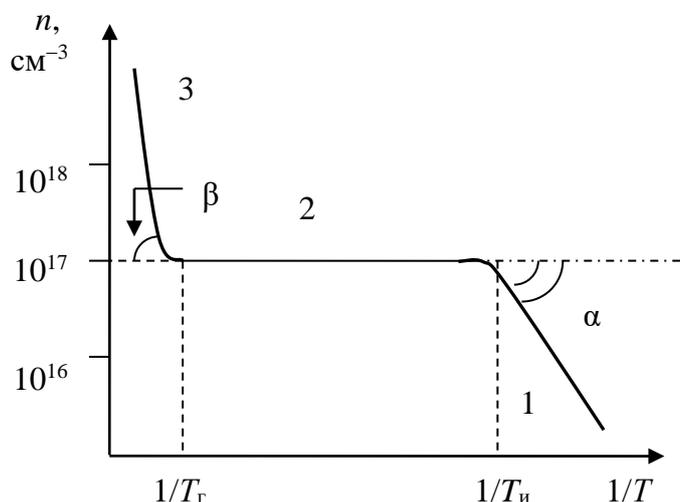


Рис. 4.1. Зависимость концентрации электронов полупроводника  $n$ -типа (в логарифмическом масштабе) от обратной температуры. Концентрация примеси равна  $10^{17} \text{ см}^{-3}$

На рис. 4.1 приведена температурная зависимость концентрации основных носителей, электронов проводимости. На кривой имеется три характерных участка, отражающих различные состояния донорного полупроводника. Часть 1 графика соответствует области низких температур, при которых происходит ионизация атомов донорной примеси (см. раздел 2.2.). Концентрация электронов растет с повышением температуры до значения, равного концентрации доноров  $10^{17} \text{ см}^{-3}$  и затем на участке 2 остается неизменной. Этот переход происходит при температуре  $T_{и}$  — температуре истощения примесей. При достижении этой температуры все атомы доноров оказываются ионизованными.

В физике полупроводников выводится формула для концентрации электронов в полупроводнике  $n$ -типа в данном интервале температур:

$$n_n = \sqrt{\frac{N_C N_D}{2}} \cdot e^{-\frac{E_C - E_D}{kT}}. \quad (4.1)$$

Таким образом, линейность участка 1 отражает экспоненциальную зависимость концентрации электронов проводимости, образовавшихся в результате ионизации доноров, от температуры. Из формулы (4.1) следует, что тангенс угла наклона  $\alpha$  этого прямолинейного участка к оси  $1/T$  пропорционален энергии ионизации доноров  $\Delta E_D = E_C - E_D$ .

Горизонтальная часть кривой (2) соответствует интервалу рабочих температур полупроводника, в котором генерация собственных электронно-дырочных пар практически отсутствует, и концентрация электронов остается постоянной. Температура  $T_r$ , выше которой начинается резкий рост концентрации электронов (участок 3), соответствует началу генерации собственных электронно-дырочных пар в полупроводнике. Тангенс  $\beta$ , угла наклона прямой на этом участке, пропорционален ширине запрещенной зоны (см. формулу (2.13)).

## 4.2. УДАРНАЯ ГЕНЕРАЦИЯ

Это явление наблюдается в полупроводнике при наличии сильных электрических полей ( $\sim 10^5 - 10^6$  В/см). В полупроводниковых приборах такие поля наблюдаются в областях электронно-дырочных переходов или других неоднородностей. Электроны на длине свободного пробега разгоняются в электрическом поле и приобретают кинетическую энергию, достаточную для разрыва валентных связей атомов в решетке полупроводника.

Если размеры кристалла значительно больше длины свободного пробега электрона, то произойдет лавинное размножение подвижных носителей: число вновь рожденных электронно-дырочных пар будет возрастать в геометрической прогрессии.

Для рассмотрения процесса ударной генерации в рамках зонной модели обратимся к рис. 4.2, а. При приложении к кристаллу полупроводника разности потенциалов, края энергетических зон

становятся наклонными, так же, как и в случае неравномерного распределения примеси по объему (см. раздел 3.4). Зонная диаграмма строится в координатах  $E - eU(x) = f(x)$ , где  $x$  — расстояние от поверхности вглубь кристалла.

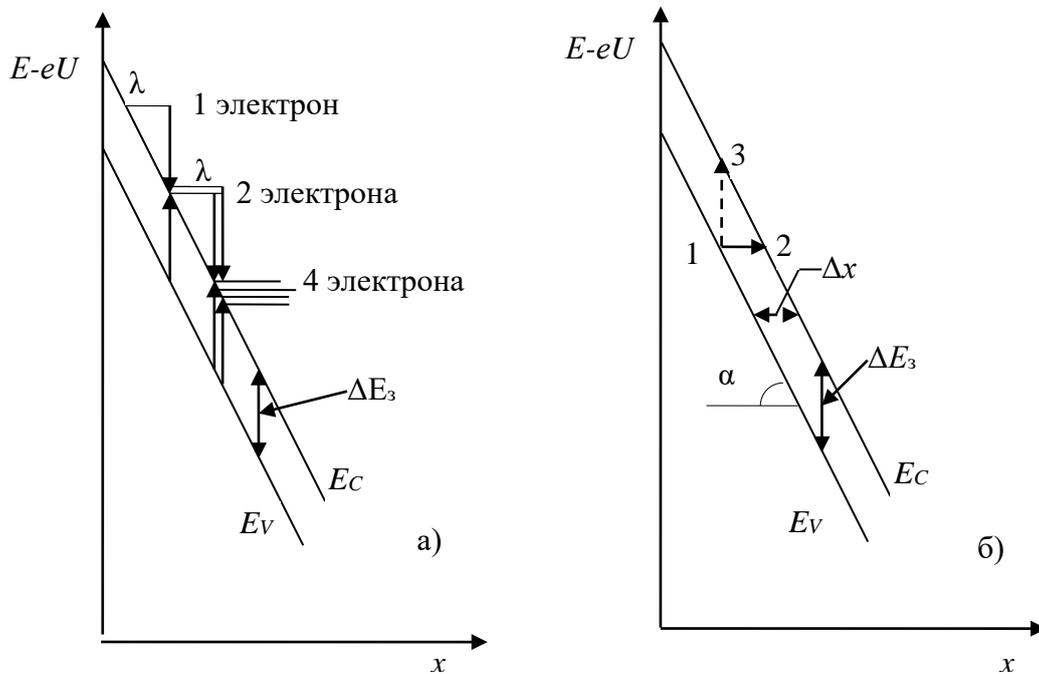


Рис. 4.2. Процесс ударной генерации (а) и полевой генерации (б) в полупроводнике в рамках зонной модели кристалла.

$\lambda$  — средняя длина свободного пробега электрона в кристалле полупроводника,

$\Delta x$  — ширина барьера для электронов потолка валентной зоны

Пусть внешняя разность потенциалов достаточна для того, чтобы было выполнено условие для начала процесса ударной генерации. Тогда энергия электрона относительно дна зоны проводимости на длине свободного пробега  $\lambda$  будет возрастать (см. верхнюю часть рис. 4.2, а), и в конце пути станет равной ширине запрещенной зоны.

Обладая такой энергией, электрон может разорвать валентную связь атома полупроводника, при этом образуется новая электронно-дырочная пара. В результате, на дне зоны проводимости окажутся

уже два электрона: один, потерявший энергию на разрыв валентной связи, и второй, появившийся в результате разрыва связи. Оба электрона, ускорившись на пути  $\lambda$ , так же потеряют энергию на разрыв двух связей, и на уровнях дна зоны проводимости окажутся теперь четыре электрона, и т. д.

### 4.3. ПОЛЕВАЯ ГЕНЕРАЦИЯ

Полевая генерация наблюдается в еще более сильных электрических полях ( $> 10^6$  В/см), и принципиально отличается от ударной генерации. В основе этого явления лежит квантовомеханический туннельный эффект.

Известно, что вероятность прохождения частицы через узкий потенциальный барьер без изменения своей энергии

$$P = e^{-\frac{4\pi}{h}\sqrt{2m^*\Delta E}\Delta x}.$$

Здесь  $\Delta E$  и  $\Delta x$  — высота и ширина потенциального барьера,  $m^*$  — эффективная масса электрона (см. раздел 2.5).

В нашем случае  $\Delta E = \Delta E_3$ , и, учитывая, что  $\operatorname{tg} \alpha = \frac{e\Delta U}{\Delta x} = e\varepsilon = \frac{\Delta E_3}{\Delta x}$  (см. треугольник 1-2-3 рис. 4.2, б), получим:

$$\Delta x = \frac{\Delta E_3}{e\varepsilon}.$$

Таким образом, вероятность полевой генерации определяется следующим выражением:

$$P = e^{-\frac{4\pi}{he\varepsilon}\sqrt{2m^*(\Delta E_3)^3}}. \quad (4.2)$$

В результате смещения энергетических зон (см. рис. 4.2, б) энергия электрона в точке 1 на потолке валентной зоны равна энергии электрона в точке 2 на дне зоны проводимости. Эти два состояния разделяет треугольный энергетический барьер с высотой  $\Delta E_3$  и шириной  $\Delta x$ . В сильных полях  $\Delta x$  мало и сравнимо с длиной волны

электрона в точке 1. Если уровень 2 свободен, то в таких условиях возможен туннельный перенос электрона из точки 1 в точку 2.

Вероятность туннельного перехода электронов из валентной зоны сквозь потенциальный барьер 1-3-2 в зону проводимости и обратно одинакова. Однако перенос электронов из валентной зоны (слева направо на рис. 4.2, б) преобладает, поскольку их там значительно больше. Следствием этого является резкое увеличение концентрации электронов при достижении указанной выше пороговой величины электрического поля в полупроводнике.

#### 4.4. СВЕТОВАЯ ГЕНЕРАЦИЯ

Различают биполярную и однополярную световую генерацию. В первом случае энергия  $h\nu$  поглощенного фотона расходуется на разрыв валентной связи собственных атомов полупроводника. Разрыв валентной связи может произойти лишь при условии, что энергия фотона превышает ширину запрещенной зоны:  $h\nu > \Delta E_3$ . При этом образуются и электроны, и дырки, т. е. возникают пары разноименно заряженных подвижных частиц.

Однополярная генерация происходит при фотоионизации примесных атомов при низких температурах. Энергия поглощаемого фотона при этом должна быть больше энергии ионизации доноров  $\Delta E_D$  для полупроводников  $n$ -типа или больше энергии ионизации акцепторов  $\Delta E_A$  для полупроводников  $p$ -типа. Образуются соответственно либо только электроны, либо дырки.

Примеси с мелкозалегающими уровнями при комнатной температуре почти полностью ионизованы. Поэтому в этих условиях однополярная генерация может при наличии в полупроводнике чужеродных атомов, в которых энергия связи электронов относительно велика, и соответствующие энергетические уровни лежат в запрещенной зоне далеко от ее границ. Иными словами, в запрещенной зоне должны быть глубокозалегающие примесные уровни.

## 4.5. РЕКОМБИНАЦИЯ ПОДВИЖНЫХ НОСИТЕЛЕЙ ЗАРЯДА. МЕХАНИЗМЫ РЕКОМБИНАЦИИ

Рекомбинацией называется процесс возвращения свободных носителей заряда в связанное состояние. Рекомбинация может происходить различными способами; при этом необходимо, чтобы одновременно выполнялись законы сохранения энергии и импульса. Тот или иной механизм рекомбинации реализуется в зависимости от того, каким образом расходуется энергия, выделившаяся при исчезновении электронно-дырочной пары.

Процесс рекомбинации может быть одноступенчатым и двухступенчатым. В первом случае происходит межзонная, или непосредственная рекомбинация: свободный электрон с энергетического уровня у дна зоны проводимости переходит на один из свободных верхних энергетических уровней валентной зоны; при этом выделяется энергия, равная ширине запрещенной зоны полупроводника. Это соответствует заполнению электроном зоны проводимости ранее незанятой валентной связи атома полупроводника.

Если энергия выделяется в виде кванта электромагнитного (светового) излучения:  $h\nu = \Delta E_z$ , то происходит излучательная рекомбинация. Ясно, что закон сохранения энергии выполняется: величина выделившейся энергии  $E_C - E_V$ , равна энергии кванта светового излучения. Для выполнения закона сохранения импульса при излучательной рекомбинации необходимо, чтобы векторная сумма импульсов электрона и дырки была равна импульсу фотона. Так как его величина  $\frac{h\nu}{c}$  ничтожно мала (близка к нулю), то для выполнения закона сохранения импульса необходимо, чтобы электрон и дырка оказались в одной точке кристалла с равными по модулю и противоположно направленными импульсами. Вероятность такого события, и, следовательно, вероятность межзонной излучательной рекомбинации невелика; в лучшем случае она достигает десятых долей процента.

Если выделившаяся при заполнении электроном валентной связи энергия передается решетке кристалла и идет на увеличение амплитуды тепловых колебаний атомов полупроводника, возникает безызлучательная межзонная рекомбинация. Как известно из физики твердого тела, кристаллическая решетка может воспринимать энергию лишь малыми порциями порядка нескольких  $kT$ , т. е. нескольких сотых долей эВ. Величина выделившейся при рекомбинации энергии  $\Delta E_z$  для различных полупроводников, по крайней мере, на порядок больше, и поэтому вероятность межзонной безызлучательной рекомбинации также мала.

Основной механизм рекомбинации подвижных носителей в полупроводниках с относительно широкой запрещенной зоной (например, в арсениде галлия) — рекомбинация через ловушки. Этот процесс протекает в два этапа. В примесных полупроводниках последовательность переходов зависит от типа проводимости. Например, в полупроводнике  $n$ -типа уровни ловушек, как правило, находятся ниже уровня Ферми в запрещенной зоне и заняты электронами. Поэтому первым шагом рекомбинации является переход электрона с уровня ловушки на энергетический уровень валентной зоны вблизи ее потолка (рис. 4.3, *а*, переход 1). Иначе говоря, происходит захват дырки (неосновного носителя) уровнем ловушки. На втором этапе электрон проводимости переходит на освободившийся уровень ловушки.

В полупроводнике  $p$ -типа уровни ловушек, расположенные выше уровня Ферми, в основном, свободны, и первым этапом рекомбинации является захват электрона (неосновного носителя) ловушкой (рис. 4.3, *б*). Вторая стадия рекомбинации — переход электрона с энергетического уровня ловушки на свободный уровень в валентной зоне.

Как видно, первым этапом рекомбинации в примесных полупроводниках является захват неосновного носителя атомом-ловушкой. Этот этап определяет преобладающую часть времени жизни неосновных носителей: их концентрация на много порядков

меньше концентрации основных носителей, и столкновения с атомами ловушек происходят относительно редко.

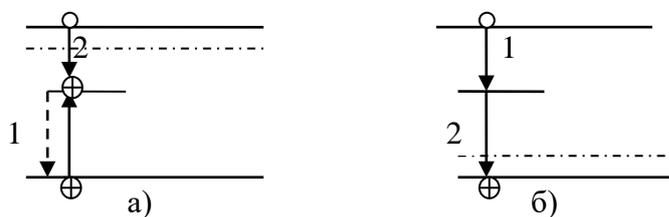


Рис. 4.3. Рекомбинация через ловушки в примесных полупроводниках: а) для полупроводника *n*-типа, б) для полупроводника *p*-типа

На втором этапе рекомбинации происходит захват основного носителя. В *n*-полупроводнике на уровень ловушки захватывается электрон (рис. 4.3, а, стрелка 2), в полупроводнике *p*-типа на уровень ловушки захватывается дырка (стрелка 2 на рис. 4.3, б). Это — относительно быстрый процесс, так как концентрация основных носителей велика.

Роль рекомбинационных ловушек могут играть специальные примесные атомы или ионы (Cu, Au, Cd, S, N<sub>2</sub>, O<sub>2</sub>), а также различные несовершенства кристалла: незаполненные узлы, трещины и т. п.

Поверхностная рекомбинация является разновидностью рекомбинации через ловушки. Из-за большого количества дефектов на поверхности полупроводника этот процесс идет значительно интенсивнее. Скорость поверхностной рекомбинации сильно зависит от качества обработки поверхности, т. е. от числа дефектов на единице площади. Например, для шлифованной поверхности кремния скорость рекомбинации в несколько тысяч раз больше, чем для поверхности после химического травления.

#### 4.6. ДИФФУЗИОННАЯ ДЛИНА НЕОСНОВНЫХ НОСИТЕЛЕЙ В ПРИМЕСНЫХ ПОЛУПРОВОДНИКАХ

Этот важный параметр можно определить как среднюю длину диффузионного перемещения неосновного носителя за время жизни, т. е. от момента возникновения за счет генерации до момента

исчезновения при рекомбинации. Для нахождения математической формулы, определяющей диффузионную длину, необходимо решить уравнение диффузии для стационарного случая, когда концентрация неосновных носителей в любой точке кристалла не изменяется со временем.

Пусть на поверхности полупроводника  $p$ -типа или на границе раздела с другим веществом (координата  $x = 0$ ) создается постоянная во времени избыточная концентрация неосновных носителей  $\Delta n_p(0)$ . Тогда в результате диффузии их концентрация в толще полупроводника при  $x > 0$  будет больше равновесной. При удалении от поверхности вглубь кристалла избыточные неосновные носители будут рекомбинировать, их концентрация будет уменьшаться, и в предельном случае при  $x \rightarrow \infty$  избыточные неосновные носители исчезнут ( $\Delta n_p(x) \rightarrow 0$ ).

Уравнение, описывающее диффузионное движение электронов в полупроводнике  $p$ -типа с учетом рекомбинации имеет вид:

$$D_n \frac{d^2 \Delta n_p}{dx^2} - \frac{\Delta n_p}{\tau_n} = 0. \quad (4.3)$$

Здесь  $D_n$  — коэффициент диффузии электронов в полупроводнике  $p$ -типа,  $x$  — координата вглубь кристалла,  $\tau_n$  — время жизни неосновных носителей в полупроводнике  $p$ -типа. Решая данное уравнение при указанных выше граничных условиях, получаем:

$$\Delta n_p(x) = \Delta n_p(0) \cdot e^{-\frac{x}{\sqrt{D_n \cdot \tau_n}}};$$

и

$$n_p(x) = n_{p0} + \Delta n_p(0) \cdot e^{-\frac{x}{\sqrt{D_n \cdot \tau_n}}}. \quad (4.4)$$

Знаменатель показателя экспоненты имеет размерность длины, и величина

$$L_n = \sqrt{D_n \cdot \tau_n} \quad (4.5)$$

называется диффузионной длиной неосновных носителей — электронов в полупроводнике  $p$ -типа. Аналогично рассматривается и величина  $L_p = \sqrt{D_p \cdot \tau_p}$  — диффузионная длина дырок в  $n$ -полупроводнике. Из (4.4) следует, что диффузионная длина — это расстояние от границы вглубь кристалла, на котором концентрация диффундирующих неосновных носителей уменьшается в  $e$  раз.

## **4.7. ВРЕМЯ ЖИЗНИ НЕОСНОВНЫХ НОСИТЕЛЕЙ В ПОЛУПРОВОДНИКЕ. УРАВНЕНИЕ ГЕНЕРАЦИИ–РЕКОМБИНАЦИИ**

Время жизни неосновных носителей — это средний промежуток времени от момента их появления в кристалле за счет какого-либо механизма генерации до исчезновения из-за рекомбинации. Иначе говоря, это — среднестатистическое время существования их в свободном состоянии. Генерация и рекомбинация подвижных носителей (основных и неосновных) происходят в полупроводнике непрерывно, поэтому скорость изменения их концентрации определяется разностью скоростей этих процессов. Наибольший интерес с точки зрения степени влияния на работу полупроводниковых приборов представляет генерация и рекомбинация неосновных носителей.

Рассмотрим полупроводник  $p$ -типа. Изменение концентрации неосновных носителей (электронов) при внешнем воздействии, достаточном для начала генерации электронно-дырочных пар, (например, при резком нагреве, при облучении светом, при наложении внешнего электрического поля) можно в грубом приближении описать следующим уравнением:

$$\frac{d\Delta n_p}{dt} = g - r. \quad (4.6)$$

Здесь  $g$  — скорость генерации,  $r$  — скорость рекомбинации, т. е. число неравновесных неосновных носителей, возникающих при

генерации и исчезающих при рекомбинации за 1 секунду в единице объема ( $\text{см}^{-3} \cdot \text{с}^{-1}$ ).

Скорость рекомбинации можно представить в виде:

$$r = \frac{\Delta n_p}{\tau_n}, \quad (4.7)$$

где  $\Delta n_p = \Delta n_p(t)$  — избыточная концентрация электронов в данный момент времени, а  $\tau_n$  — время жизни неосновных носителей в полупроводнике  $p$ -типа. После прекращения внешнего воздействия начинается процесс установления равновесия: генерация прекращается, и концентрация неравновесных носителей уменьшается за счет рекомбинации.

Если в некоторый момент времени внешнее воздействие на полупроводник прекратится, то скорость генерации электронно-дырочных пар станет равной нулю ( $g = 0$ ), и уравнение (4.6) примет вид:

$$\frac{d\Delta n_p}{dt} = -\frac{\Delta n_p}{\tau_n}. \quad (4.8)$$

Решив это уравнение с учетом начального условия (при  $t = 0$   $\Delta n_p = \Delta n_p(0)$ ), получим:

$$\Delta n_p(t) = \Delta n_p(0) \cdot e^{-\frac{t}{\tau_n}}. \quad (4.9)$$

Видно, что по прошествии промежутка времени  $\tau_n$  концентрация избыточных неосновных носителей уменьшается за счет рекомбинации в  $e$  раз. Таким образом, время жизни неосновных носителей является одним из важнейших параметров, определяющих скорость убывания их концентрации, или, как принято говорить, скорость их расщивания. Такие процессы в электронно-дырочных переходах определяют быстродействие полупроводниковых диодов различного назначения, биполярных транзисторов и тиристор.

#### *Вопросы к разделу 4*

1. Дайте определение явления генерации подвижных носителей и перечислите ее разновидности.
2. Как изменяется концентрация электронов в донорном полупроводнике при повышении температуры от 0 К в отсутствие других внешних воздействий?
3. Каков физический механизм ударной генерации подвижных носителей? Как отображается этот процесс на энергетической диаграмме?
4. По какой причине для объяснений физических механизмов ударной и полевой генерации используют «наклонные» энергетические зоны полупроводника?
5. Каков физический механизм полевой генерации подвижных носителей? Как отображается этот процесс на энергетической диаграмме?
6. Объясните, каким образом в полупроводниках возникает биполярная и однополярная световая генерация подвижных носителей.
7. Каковы необходимые и достаточные условия для рекомбинации электрона и дырки?
8. Поясните причину малой вероятности межзонной излучательной рекомбинации.
9. Что такое «атомы-ловушки» и как происходит рекомбинация с их участием в примесных полупроводниках?
10. Каков физический смысл термина «диффузионная длина неосновных носителей».
11. На основании каких предпосылок составляется уравнение генерации–рекомбинации?
12. Объясните физический смысл термина «время жизни неосновных носителей в примесном полупроводнике».

## 5. ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

### 5.1. ОСОБЕННОСТИ ЭЛЕКТРОННОГО СТРОЕНИЯ ПРИПОВЕРХНОСТНОЙ ОБЛАСТИ ПОЛУПРОВОДНИКА

Поверхность кристалла, т. е. граница между кристаллом и окружающей средой (обрыв решетки) является самым большим его «дефектом». На этой границе происходит обрыв связей между атомами в перпендикулярном направлении, и вследствие этого электронная структура приповерхностного слоя отличается от таковой для объема. На энергетической диаграмме вблизи поверхности кристалла в запрещенной зоне появляются дискретные поверхностные энергетические уровни. Электроны, занимающие эти уровни, не могут проникать вглубь кристалла и локализуются вблизи поверхности. Поверхностные уровни могут служить донорами, акцепторами или ловушками. Заполнение акцепторных уровней означает локализацию электронов, освобождение донорных уровней — локализацию дырок, и в результате поверхность заряжается отрицательно или положительно.

В соответствии с условием электронейтральности это должно сопровождаться образованием вблизи поверхности полупроводника объемного заряда, нейтрализующего поверхностный заряд. Это происходит, прежде всего, за счет оттеснения вглубь кристалла зарядов того же знака (электронов, дырок). В результате приповерхностный слой полупроводника оказывается обедненным носителями заряда одного знака с поверхностным зарядом и обогащенным носителями заряда противоположного знака, и электрическое поле поверхностного заряда тем самым экранируется.

Если в металлах, имеющих очень высокую концентрацию электронов ( $\sim 10^{22}$  см<sup>-3</sup>), нейтрализация поверхностного заряда происходит уже на глубине, равной нескольким параметрам кристаллической решетки, то в полупроводниках область объемного заряда простирается на глубину  $10^{-4}$  см и более. Обычно ее принимают равной

дебаевской длине (радиусу) экранирования  $L_D$ . Она определяется как расстояние, на протяжении которого напряженность электрического поля в веществе с подвижными носителями заряда уменьшается в  $e$  раз ( $e \approx 2,72$ ). Расчеты показывают, что для собственного полупроводника

$$L_D = \left( \frac{\varepsilon \varepsilon_0 k T}{2 e^2 n_i} \right)^{\frac{1}{2}}, \quad (5.1)$$

где  $\varepsilon$  — относительная диэлектрическая проницаемость полупроводника,  $\varepsilon_0$  — электрическая постоянная вакуума,  $e$  — заряд электрона. Для примесного полупроводника с концентрацией донорной или акцепторной примеси  $N$ :

$$L_D = \left( \frac{\varepsilon \varepsilon_0 k T}{2 e^2 N} \right)^{\frac{1}{2}}. \quad (5.2)$$

Поскольку  $N \gg n_i$ , дебаевский радиус в примесных полупроводниках намного меньше, чем в собственных. Так, при  $T = 300$  К в собственном полупроводнике (кремнии)  $L_D \approx 14$  мкм, а в  $n$ -кремнии с концентрацией доноров  $10^{16} \text{ см}^{-3}$   $L_D \approx 0,04$  мкм.

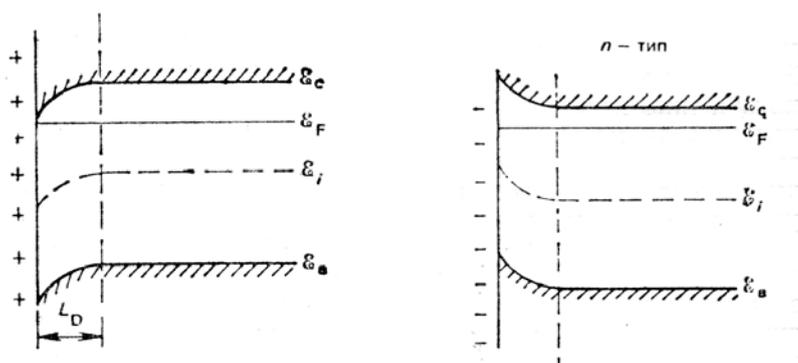


Рис. 5.1. Изгиб зон полупроводника  $n$ -типа при наличии положительного (а) и отрицательного (б) поверхностного заряда

Наличие на поверхности полупроводника электрического заряда обуславливает возникновение разности потенциалов между его

поверхностью и объемом. Следствием этого является искривление энергетических зон.

Отрицательный поверхностный заряд вызывает изгиб зон вверх (рис. 5.1, б), так как при перемещении электрона из объема к поверхности его энергия увеличивается. Соответственно, положительный заряд на поверхности вызывает изгиб энергетических зон вниз (рис. 5.1, а).

## 5.2. ЭФФЕКТ ПОЛЯ В ПОЛУПРОВОДНИКАХ

Аналогичная картина — обеднение или обогащение приповерхностного слоя полупроводника подвижными носителями заряда — наблюдается и под воздействием электрического поля, направленного перпендикулярно поверхности. Это явление называется эффектом поля. Состояние приповерхностного слоя определяется ориентацией и величиной вектора электрической напряженности.

Рассмотрим полупроводник *p*-типа, у которого отсутствует поверхностный заряд. Ясно (см. выше), что это предположение далеко от реальности. Тем не менее, примем его для того, чтобы явным образом проявилось влияние внешнего электрического поля. Вектор напряженности поля направлен перпендикулярно к поверхности, и поле проникает в полупроводник на глубину  $L_D$  (см. формулу (5.2)).

### 5.2.1. Режим обеднения

Если напряженность поля невелика и ее вектор направлен к поверхности полупроводника *p*-типа, (рис. 5.2, а), то основные носители (дырки) отесняются вглубь кристалла и их концентрация у поверхности заметно уменьшается (рис. 5.2, б). При этом неосновные носители (электроны) притягиваются из объема кристалла к поверхности, но из-за небольшой напряженности поля (рис. 5.2, в) их концентрация у поверхности остается пренебрежимо малой. В результате возникает слой толщиной  $L_{об}$ , обедненный основными носителями, в котором сосредоточены некомпенсированные отрицательные ионы акцепторов.

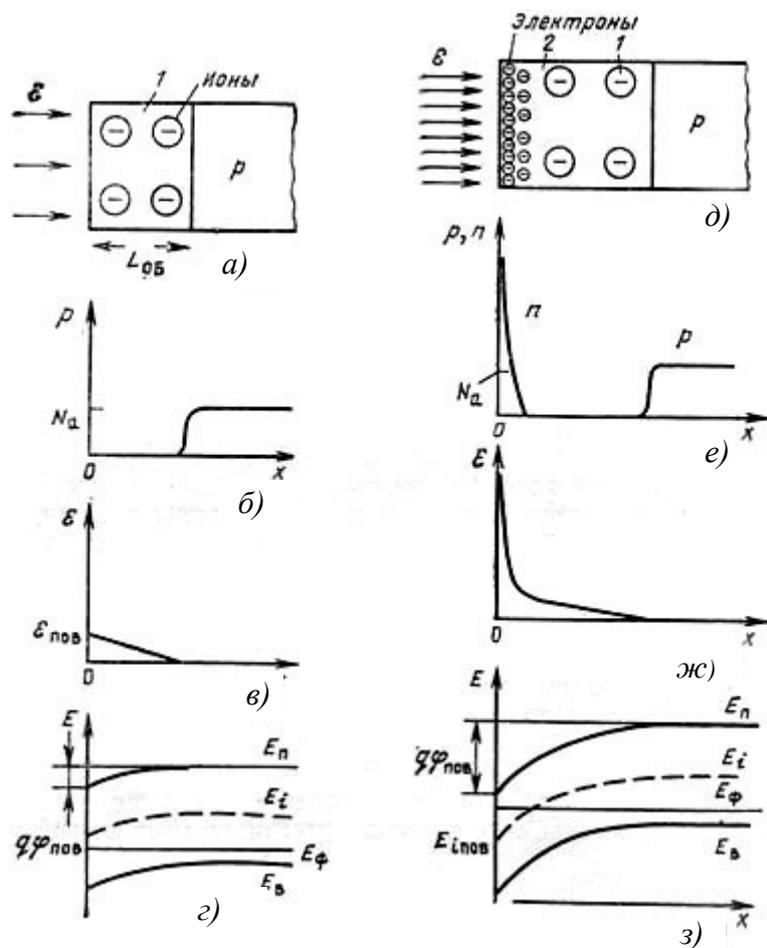


Рис. 5.2. Режим обеднения (а – в) и инверсии (д – ж) приповерхностной области полупроводника  $p$ -типа при воздействии внешнего электрического поля, перпендикулярного к поверхности

На рис. 5.2, в показана энергетическая диаграмма обедненного слоя. Энергетические зоны кристалла у поверхности в пределах слоя толщиной  $L_{об}$  изгибаются вниз, что соответствует увеличению энергии дырок и уменьшению энергии электронов. Уровень Ферми расположен ниже середины запрещенной зоны и остается постоянным, так как система находится в равновесии (ток через полупроводник отсутствует). Следовательно, тип проводимости полупроводника по всей толщине одинаков.

Видно, что расстояние между потолком валентной зоны и уровнем Ферми  $E_F - E_V$  увеличивается, а расстояние от уровня Ферми до

дна зоны проводимости  $E_C - E_F$  уменьшается на одинаковую величину  $e\varphi_{\text{пов}}$ . Соответственно, (см. формулы (2.10) и (2.11)) это приводит к уменьшению поверхностной концентрации дырок и к увеличению поверхностной концентрации электронов:

$$p_{\text{пов}} = p_{p_0} e^{-\frac{e\varphi_{\text{пов}}}{kT}}; \quad n_{\text{пов}} = n_{p_0} e^{\frac{e\varphi_{\text{пов}}}{kT}}.$$

Режим обеднения будет сохраняться, пока поверхностная концентрация электронов не превысит концентрацию дырок:  $n_{\text{пов}} > N_A$ . Равенству этих концентраций соответствует пороговое значение поверхностного потенциала:

$$\varphi_{\text{пор}} = \frac{kT}{e} \ln \frac{N_A}{n_{p_0}} = \frac{kT}{e} \ln \frac{N_A p_{p_0}}{n_i^2} = \frac{kT}{e} \ln \frac{N_A^2}{n_i^2}.$$

Окончательно,

$$\varphi_{\text{пор}} = 2 \frac{kT}{e} \ln \frac{N_A}{n_i}. \quad (5.3)$$

### 5.2.2. Режим инверсии

При дальнейшем увеличении напряженности поля поверхностный потенциал становится больше порогового, и в приповерхностном слое возникает режим инверсии типа проводимости (рис. 5.2,  $\delta - \varepsilon$ ). В этом режиме поверхностная концентрация неосновных носителей (в нашем случае электронов) значительно превышает концентрацию основных носителей (рис. 5.2,  $e$ ). Толщина инверсного слоя очень мала и определяется величиной радиуса дебаевского экранирования  $L_D$ . Например, как указывалось выше, в  $n$ -кремнии с концентрацией донорной примеси  $10^{16} \text{ см}^{-3}$   $L_D \approx 0,04 \text{ мкм}$ .

На рис. 5.2,  $\varepsilon$ , показана энергетическая диаграмма полупроводника  $p$ -типа в режиме инверсии. Видно, что изгиб энергетических зон вниз значительно больше, чем в режиме обеднения, и у поверхности уровень Ферми находится выше середины запрещенной зон. Такое положение соответствует энергетической диаграмме донорного полупроводника.

### 5.2.3. Режим обогащения

При изменении направления вектора напряженности электрического поля основные носители (дырки) будут притягиваться к поверхности; их концентрация станет выше, чем в объеме. Такое состояние приповерхностного слоя называется режимом обогащения. В обогащенном приповерхностном слое концентрация дырок значительно превышает концентрацию акцепторов.

Если ввести теперь в рассмотрение поверхностный заряд, то его знак и величина будут определять состояние приповерхностного слоя полупроводника в отсутствие внешнего электрического поля. Поверхностный заряд можно считать сосредоточенным на плоскости, поэтому вместе с индуцированным зарядом противоположного знака он создает однородное электрическое поле с напряженностью

$$|\vec{\epsilon}_{\text{пов}}| = \frac{Q_{\text{пов}}}{\epsilon_0 \epsilon}, \quad (5.4)$$

где  $Q_{\text{пов}}$  — поверхностный заряд, приходящийся на единицу площади поверхности,  $\epsilon_0$  — электрическая постоянная,  $\epsilon$  — диэлектрическая проницаемость полупроводника. Это означает, что заряд  $Q_{\text{пов}}$  изменяет концентрацию носителей в приповерхностной области полупроводника точно так же, как внешнее электрическое поле с напряженностью  $|\vec{\epsilon}| = \frac{Q_{\text{пов}}}{\epsilon_0}$ . Таким образом, только под воздействием поверхностного заряда в полупроводнике на глубине порядка дебаевского радиуса экранирования может образовываться либо обедненный, либо инверсный, либо обогащенный слой.

### Вопросы к разделу 5

1. Каков физический механизм экранирования поверхностного заряда полупроводника?

2. Какова толщина слоя, в пределах которого внешнее электрическое поле, перпендикулярное поверхности, изменяет электрические свойства веществ?
3. Дайте определение понятия «дебаевский радиус экранирования». От каких физических параметров, и каким образом зависит его величина?
4. Каковы физические причины появления изгиба энергетических зон у поверхности полупроводника?
5. Каким образом воздействует на подвижные носители внешнее электрическое поле, направленное перпендикулярно поверхности полупроводника?
6. Дайте определение понятия «эффект поля в полупроводнике».
7. Каковы условия, при которых в приповерхностной области полупроводника  $p$ -типа возникает режим обеднения?
8. Какому соотношению концентраций основных и неосновных носителей в приповерхностной области соответствует пороговое значение напряженности внешнего электрического поля?
9. Каковы условия, при которых в приповерхностной области полупроводника  $p$ -типа возникает режим инверсии типа проводимости?
10. Чем определяется глубина инверсного слоя в полупроводнике?
11. Изобразите энергетическую диаграмму полупроводника  $p$ -типа в режиме инверсии типа проводимости.
12. Каковы физические причины того, что под воздействием поверхностного заряда могут возникать режимы обеднения, инверсии и обогащения?

## 6. КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

При идеальном контакте двух проводящих тел на границе возникают переходные слои различной природы. Контактные явления в полупроводниках определяют принцип действия всех активных элементов полупроводниковой электроники: диодов, полевых и биполярных транзисторов, тиристоров. Переходные слои на границах раздела, которые существуют в полупроводниковых активных элементах, возникают на контактах различного типа:

- 1) контакты между полупроводниками одной природы с разными или одинаковыми типами примесных атомов ( $p-n$ - переходы, переходы  $n-n^+$ ,  $p-p^+$ );
- 2) контакты между полупроводниками различной природы (гетеропереходы),
- 3) контакты металл-полупроводник  $n$ - или  $p$ -типа;
- 4) контакты металл-диэлектрик-полупроводник.

В полупроводниковой электронике эти контакты часто называют электрическими переходами, так как при установлении равновесного состояния на границе раздела веществ возникает электрическое поле. Это поле определяет принцип действия конкретных полупроводниковых приборов, изготовленных на основе таких переходов.

### 6.1. ИДЕАЛЬНЫЙ ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД ( $P-N$ -ПЕРЕХОД)

$P-n$ -переход — это переходный слой между двумя областями полупроводника с разным типом электропроводности, в котором существует электрическое поле. Такой переход нельзя создать путем механического контакта двух кристаллов полупроводника, так как поверхности всегда загрязнены чужеродными атомами, окислами и т. п. Переходы создаются с помощью специальных технологических

приемов, например, методом диффузии примесей или методом сплавления.

Переходы различаются по двум основным параметрам:

- по характеру распределения концентрации примесей у границы раздела,
- по соотношению концентраций примесных атомов в смежных полупроводниковых областях.

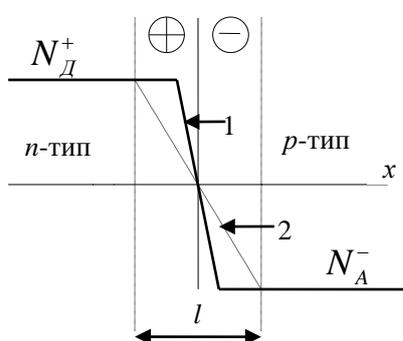


Рис. 6.1. Одно из возможных распределений концентраций примесей в резком (1) и плавном (2)  $p$ - $n$ -переходе.  $l$  – ширина области пространственного заряда ионов примесей (ширина перехода)

В первом случае характерными величинами являются: ширина области изменения концентрации примеси  $\Delta x$  у границы раздела и ширина перехода  $l$ . Если  $\Delta x \ll l$ , переход называется резким, а при условии  $\Delta x \sim l$  образуется плавный  $p$ - $n$ -переход. Резкому и плавному переходу могут соответствовать распределения концентрации примесей, показанные на рис. 6.1.

Если концентрации примесных атомов в  $p$ - и  $n$ -областях одинаковы, то переход называется симметричным (рис. 6.1). Несимметричный  $p$ - $n$ -переход образуется полупроводниками с различными концентрациями примесей и, следовательно, с различными концентрациями основных и неосновных носителей. Если концентрации основных носителей отличаются более чем на порядок, то переход называется односторонним.

### 6.1.1. *p-n-переход в равновесном состоянии*

В этом разделе рассматривается идеализированный *p-n*-переход, для которого действуют следующие ограничения:

- Концентрация примесей в смежных областях не превышает  $10^{17} \text{ см}^{-3}$ , т. е. переход изготовлен на основе невырожденных полупроводников.
- Токи малы и не вызывают заметного падения напряжения в толще полупроводника. Вследствие этого электрическое поле сосредоточено только в области *p-n*-перехода.
- Омические контакты к *p*- и *n*-областям идеальны, т. е. около них в полупроводнике существует всегда равновесная концентрация носителей заряда.
- В области перехода не происходят процессы генерации и рекомбинации носителей.
- Поверхностные явления несущественны.

Установление равновесного состояния на *p-n*-переходе происходит следующим образом. При идеальном контакте двух кристаллов одного и того же полупроводника с различным типом электропроводности из-за градиента концентрации основных носителей возникает их диффузия через плоскость раздела. Дырки из *p*-области, примыкающей к границе, начинают диффундировать в *n*-область; диффузия электронов протекает в обратном направлении. В результате нарушается условие электронейтральности в областях, примыкающих к границе. В *p*-области после ухода из нее дырок у границы раздела остаются некомпенсированные отрицательные ионы акцепторов, а в *n*-области — положительные ионы доноров. Образуется область объемного заряда, состоящая из двух разноименно заряженных слоев, в которой возникает электрическое поле  $\varepsilon_k$ , и между *n*- и *p*-областями устанавливается контактная разность потенциалов  $\varphi_k$ . Потенциал *n*-области становится положительным по отношению к потенциалу *p*-области.

Возникшее электрическое поле  $\epsilon_k$  создает потенциальный барьер на пути диффундирующих основных носителей, и скорость их перемещения в смежную область уменьшается. Вместе с тем под действием  $\epsilon_k$  возникает дрейфовое движение неосновных носителей заряда. Дырки движутся из  $n$ -области в  $p$ -область; электроны — в противоположном направлении.

Таким образом, в процессе установления равновесия на переходе через границу проходят встречные потоки одноименно заряженных частиц, т.е. текущие навстречу друг другу токи:

$$J_{\text{Диф}} = J_{n\text{Диф}} + J_{p\text{Диф}} \quad (\text{основные носители})$$

$$J_{\text{Др}} = J_{p\text{Др}} + J_{n\text{Др}} \quad (\text{неосновные носители})$$

Равновесие на переходе установится при условии, когда напряженность контактного электрического поля  $\epsilon_k$  достигнет такой величины, при которой диффузионный ток основных носителей полностью компенсируется встречным дрейфовым током неосновных носителей. Полный ток через переход при этом становится равным нулю:

$$J = J_{\text{Диф}} + J_{\text{Др}} = 0$$

Переход обеднен подвижными носителями зарядов, поэтому сопротивление его значительно выше сопротивления объемов полупроводника, лежащих за пределами слоя толщиной  $l$ .

Энергетическая диаграмма симметричного  $p$ - $n$ -перехода, находящегося в равновесии, приведена на рис. 6.2. За пределами области объемного заряда ионов примесей электрическое поле отсутствует, поэтому энергетические зоны тут горизонтальны. В пределах ширины перехода энергетические зоны имеют наклон, отражающий величину напряженности контактного электрического поля. Так как потенциал  $n$ -области положителен по отношению к потенциалу  $p$ -области, то на диаграмме энергетические зоны для  $n$ -области расположены ниже, чем для  $p$ -области.

Поскольку в равновесном состоянии уровень Ферми должен быть единым для всей системы, высота установившегося

потенциального барьера  $e\phi_k$ , препятствующего диффузии основных носителей, равна разности энергий Ферми  $p$ - и  $n$ -областей. Таким образом:

$$e\phi_k = E_{F_n} - E_{F_p}. \quad (6.1)$$

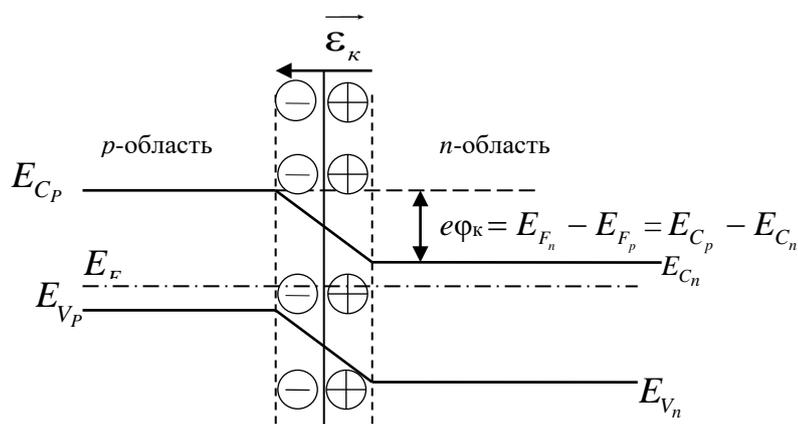


Рис. 6.2. Энергетическая диаграмма идеального симметричного электронно-дырочного перехода в состоянии равновесия

Будем считать, что объемные заряды в переходе обусловлены только ионизованными атомами доноров и акцепторов, и на основе этой идеальной модели рассмотрим основные физические параметры резкого  $p$ - $n$ -перехода в равновесном состоянии.

Для нахождения зависимости высоты потенциального барьера от концентрации основных и неосновных носителей воспользуемся соотношением (2.7) для концентрации электронов. В полупроводнике  $n$ -типа для основных носителей имеем:

$$n_n = N_C e^{-\frac{E_{C_n} - E_F}{kT}}. \quad (6.2)$$

Концентрация неосновных носителей (электронов) в  $p$ -полупроводнике равна:

$$n_p = N_C e^{-\frac{E_{C_p} - E_F}{kT}}. \quad (6.3)$$

Из рис. 6.2 видно, что потенциальный барьер для диффузии электронов в  $p$ -область равен разности уровней дна зоны проводимости в  $p$ - и  $n$ -областях:

$$e\varphi_k = E_{C_p} - E_{C_n}; \text{ тогда } E_{C_p} = e\varphi_k + E_{C_n}. \quad (6.4)$$

Подставляя (6.4) в выражение (6.3), получим:

$$n_p = N_C e^{-\frac{E_{C_n} + e\varphi_k - E_F}{kT}} = N_C e^{-\frac{E_{C_n} - E_F}{kT}} \cdot e^{-\frac{e\varphi_k}{kT}} = n_n e^{-\frac{e\varphi_k}{kT}}. \quad (6.5)$$

Окончательно имеем:

$$e\varphi_k = kT \ln \frac{n_n}{n_p}. \quad (6.6)$$

Аналогичным образом можно выразить  $e\varphi_k$  через концентрации дырок:

$$e\varphi_k = kT \ln \frac{p_p}{p_n}. \quad (6.7)$$

Оценим величину потенциального барьера  $e\varphi_k$  для кремниевого несимметричного  $p$ - $n$ -перехода при комнатной температуре ( $T = 300$  К). Пусть концентрация дырок в  $p$ -области равна  $10^{15} \text{ см}^{-3}$ , тогда концентрация неосновных носителей-электронов  $n_p$  в соответствии с формулой (2.18) и табл. (2.1) будет равна  $10^5 \text{ см}^{-3}$ . Концентрацию основных носителей  $n_n$  в полупроводнике  $n$ -типа выберем равной  $10^{17} \text{ см}^{-3}$ . Тогда в соответствии с формулой (6.6) получим:

$$e\varphi_k [\text{эВ}] = 2,6 \cdot 10^{-2} \cdot 2,3 \cdot \lg \frac{10^{17}}{10^5} \approx 0,72 \text{ эВ}.$$

Такие же расчеты для германиевого  $p$ - $n$ -перехода с  $N_A = N_D = 10^{16} \text{ см}^{-3}$  дают величину  $e\varphi_k \approx 0,32 \text{ эВ}$ .

Для нахождения контактного электрического поля  $\varepsilon_k$  необходимо решить уравнение Пуассона для конкретного характера распределения заряда и данных граничных условий. Будем считать, что в пределах перехода имеются постоянные концентрации ионов примесей:  $+eN_D$  при изменении координаты от  $-l_n$  до 0 и  $-eN_A$  на отрезке  $[0; l_p]$  (см. рис. 6.3, а).

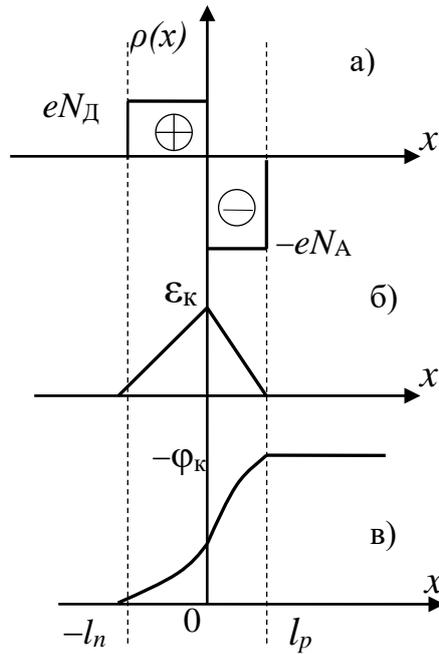


Рис. 6.3. Изменение параметров идеального резкого  $p-n$ -перехода в пределах его ширины: а) заряда  $\rho(x)$ , б) величины контактного электрического поля  $\epsilon_K$ ; в) потенциала  $\phi_K$ .

В такой модели необходимо решить одномерное уравнение Пуассона для двух интервалов:  $-l_n \leq x \leq 0$  и  $0 \leq x \leq l_p$ :

$$\frac{d^2\phi}{dx^2} = -\frac{\rho(x)}{\epsilon\epsilon_0}. \quad (6.8)$$

$$1. -l_n \leq x \leq 0: \frac{d^2\phi}{dx^2} = -\frac{eN_D}{\epsilon\epsilon_0} \qquad 2. 0 \leq x \leq l_p: \frac{d^2\phi}{dx^2} = \frac{eN_A}{\epsilon\epsilon_0}$$

Граничное условие:

$$\left. \frac{d\phi}{dx} \right|_{x=-l_n} = 0$$

(6.9)

$$\frac{d\phi}{dx} = -\frac{eN_D}{\epsilon\epsilon_0}x + C$$

Граничное условие:

$$\left. \frac{d\phi}{dx} \right|_{x=l_p} = 0$$

(6.10)

$$\frac{d\phi}{dx} = \frac{eN_A}{\epsilon\epsilon_0}x + C$$

Константы интегрирования определяются из условий (6.9) и (6.10):

$$\frac{d\phi}{dx} = -\frac{eN_D}{\epsilon\epsilon_0}(x+l_n).$$

$$\frac{d\phi}{dx} = \frac{eN_A}{\epsilon\epsilon_0}(x-l_p).$$

Как известно, электрическое поле равно взятому с обратным знаком градиенту потенциала. В одномерном случае  $\epsilon_k = -\frac{d\phi}{dx}$ , и зависимости величины контактного электрического поля от координаты в рассматриваемых интервалах будут выражаться следующим образом:

$$-l_n \leq x \leq 0: \quad \epsilon_{k1} = \frac{eN_D}{\epsilon\epsilon_0}(x+l_n); \quad (6.11)$$

$$0 \leq x \leq l_p: \quad \epsilon_{k2} = -\frac{eN_A}{\epsilon\epsilon_0}(x-l_p). \quad (6.12)$$

Изменение величины контактного электрического поля  $\epsilon_k$  в пределах перехода показано на рис. 6.3, б. Максимальная величина  $\epsilon_k$  наблюдается на границе между  $p$ - и  $n$ -областями, т. е. при  $x = 0$ :

$$\epsilon_{k \text{ макс}} = \epsilon_{k1}(0) = \epsilon_{k2}(0) = \frac{eN_D}{\epsilon \cdot \epsilon_0} l_n = \frac{eN_A}{\epsilon \cdot \epsilon_0} l_p. \quad (6.13)$$

Отсюда вытекает очень важное соотношение:

$$\frac{l_n}{l_p} = \frac{N_A}{N_D}. \quad (6.14)$$

Видно, что глубина проникновения контактного поля в полупроводник определяется степенью его легирования: область перехода проникает в полупроводник на тем большую глубину, чем меньше в нем концентрация примеси.

Для нахождения зависимости ширины перехода  $l$  от разности потенциалов на переходе и концентрации примесей в смежных областях необходимо вычислить второй интеграл уравнения (6.6):

$$\phi_k = \int_{-l_n}^0 \left( \frac{d\phi}{dx} \right) dx + \int_0^{l_p} \left( \frac{d\phi}{dx} \right) dx.$$

Эта величина равна площади под графиком  $\epsilon_k(x)$  на рис. 6.3, б:

$$\begin{aligned}\varphi_{\text{к}} &= \frac{1}{2} \varepsilon_{\text{к макс}} \cdot l_n + \frac{1}{2} \varepsilon_{\text{к макс}} \cdot l_p = \frac{1}{2} \frac{eN_{\text{Д}}}{\varepsilon\varepsilon_0} l_n^2 + \frac{1}{2} \frac{eN_{\text{А}}}{\varepsilon\varepsilon_0} l_p^2 = \\ &= \frac{e}{2\varepsilon\varepsilon_0} (N_{\text{Д}} l_n^2 + N_{\text{А}} l_p^2).\end{aligned}\quad (6.15)$$

Учитывая соотношение (6.14), выразим  $l_n$  и  $l_p$  через полную ширину перехода  $l$ :

$$l = l_n + l_p = l_n + l_n \frac{N_{\text{Д}}}{N_{\text{А}}} = \frac{N_{\text{А}} + N_{\text{Д}}}{N_{\text{А}}} l_n.$$

Отсюда:

$$l_n = \frac{N_{\text{А}}}{N_{\text{А}} + N_{\text{Д}}} l, \quad (6.16)$$

$$l_p = \frac{N_{\text{Д}}}{N_{\text{А}} + N_{\text{Д}}} l. \quad (6.17)$$

Подставляя (6.16) и (6.17) в выражение (6.15), получим:

$$\varphi_{\text{к}} = \frac{el^2}{2\varepsilon\varepsilon_0} \cdot \frac{N_{\text{Д}} N_{\text{А}}}{N_{\text{Д}} + N_{\text{А}}}. \quad (6.18)$$

Таким образом, ширина равновесного  $p$ - $n$ -перехода выражается формулой:

$$l = \sqrt{\frac{2\varepsilon\varepsilon_0 \varphi_{\text{к}}}{e} \cdot \frac{N_{\text{Д}} + N_{\text{А}}}{N_{\text{Д}} \cdot N_{\text{А}}}}. \quad (6.19)$$

Для односторонних переходов получаются более простые выражения:

$$N_{\text{А}} \gg N_{\text{Д}}: l \simeq l_n \simeq \sqrt{\frac{2\varepsilon\varepsilon_0 \varphi_{\text{к}}}{e N_{\text{Д}}}}; \quad (6.20)$$

$$N_{\text{Д}} \gg N_{\text{А}}: l \simeq l_p \simeq \sqrt{\frac{2\varepsilon\varepsilon_0 \varphi_{\text{к}}}{e N_{\text{А}}}}. \quad (6.21)$$

Из формул (6.19), (6.20) и (6.21) следует, что ширина  $p$ - $n$ -перехода определяется главным образом величинами концентраций примесей в смежных областях: при повышении уровня легирования ширина перехода уменьшается. Величина контактной разности

потенциалов на переходе также оказывает влияние на его ширину. Увеличение  $\varphi_k$  происходит при использовании полупроводников с более широкой запрещенной зоной. Это можно легко увидеть, построив соответствующие энергетические диаграммы  $p$ - $n$ -переходов. Ширина перехода при этом также возрастает. Для плавного  $p$ - $n$ -перехода зависимость его ширины от разности потенциалов на переходе выражена слабее:  $l \sim \sqrt[3]{\varphi_k}$ .

Проведем оценку ширины резкого одностороннего кремниевого перехода. Пусть уровень легирования  $n$ -области, по крайней мере, на порядок меньше, чем в  $p$ -области, и  $N_D = 10^{16} \text{ см}^{-3}$ ,  $\varphi_k \approx 0,7 \text{ В}$ . Относительная диэлектрическая проницаемость кремния имеет величину  $\varepsilon \approx 10$ . Подставляя эти величины в формулу (6.20), получим:

$$l \approx l_n \approx \sqrt{\frac{2 \cdot 10 \cdot 8,85 \cdot 10^{-12} \cdot 7,2 \cdot 10^{-1}}{1,6 \cdot 10^{-19} \cdot 10^{22}}} \approx \frac{2\sqrt{2}}{10^7} \approx 2,8 \cdot 10^{-7} \text{ (м)} = 0,28 \text{ мкм}.$$

Отсюда следует, что контактная разность потенциалов 0,72 В приходится на расстояние 0,28 мкм, т. е. величина напряженности контактного электрического поля на таком (относительно широком) переходе составляет  $\sim 2,6 \cdot 10^6 \text{ В/м}$ .

### 6.1.2. Прямое включение $p$ - $n$ -перехода

Пусть к  $p$ - $n$ -переходу подключен источник напряжения  $U_{\text{пр}}$  плюсом к  $p$ -области, а минусом к  $n$ -области, т. е. полярностью, противоположной контактной разности потенциалов  $\varphi_k$ . Сопротивление перехода гораздо больше сопротивления объемов  $p$ - и  $n$ -областей, поэтому практически все внешнее напряжение оказывается приложенным к переходу. На энергетической диаграмме прямо смещенного перехода (рис. 6.4) уровень Ферми  $p$ -области располагается ниже уровня Ферми  $n$ -области, разность между ними равна  $eU_{\text{пр}}$ .

В результате потенциальный барьер для диффузии основных носителей уменьшается до величины  $e(\varphi_k - U_{\text{пр}})$  (рис. 6.4), равновесное состояние нарушается, и через переход течет диффузионный ток

основных носителей: дырок из  $p$ -области в  $n$ -область и электронов из  $n$ -области в  $p$ -область.

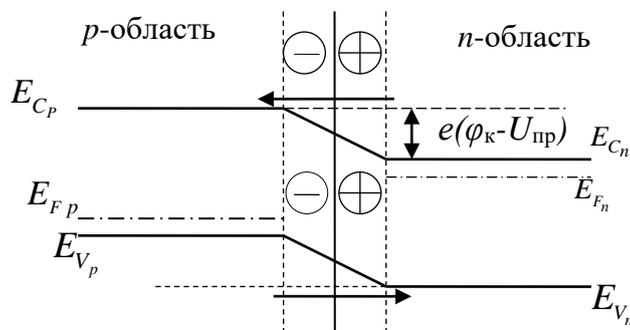


Рис. 6.4. Энергетическая диаграмма  $p$ - $n$ -перехода, смещенного в прямом направлении. Стрелками показаны направления диффузионного движения электронов и дырок

Диффузионное движение основных носителей через  $p$ - $n$ -переход в смежную область, где они становятся неосновными, называется инжекцией неосновных носителей заряда.

С уменьшением барьера уменьшится ширина перехода:

$$l_{пр} = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_k - U_{пр})}{e} \cdot \frac{N_D + N_A}{N_D \cdot N_A}}. \quad (6.22)$$

Соответственно, для односторонних переходов:

$$l_{пр} \approx l_n \approx \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_k - U_{пр})}{eN_D}}; \quad l_{пр} \approx l_p \approx \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_k - U_{пр})}{eN_A}}. \quad (6.23)$$

Введем понятие уровня инжекции:

$$\delta = \frac{\Delta p_n}{n_{n_0}} = \frac{\Delta n_p}{p_{p_0}}. \quad (6.24)$$

Уровень инжекции равен отношению приращения концентрации неосновных носителей к равновесной концентрации основных носителей заряда.

Низким уровнем инжекции ( $\delta \ll 1$ ) будем считать такой, при котором концентрация инжектированных неосновных носителей

намного меньше равновесной концентрации основных носителей. Средний уровень:  $\delta \sim 1$  (разница между  $\Delta n$ ,  $\Delta p_n$  и  $p_{p0}$ ,  $n_{n0}$  мала). Высокий уровень инжекции:  $\delta \gg 1$ ; при этом концентрация инжектированных неосновных носителей намного больше равновесной концентрации основных носителей.

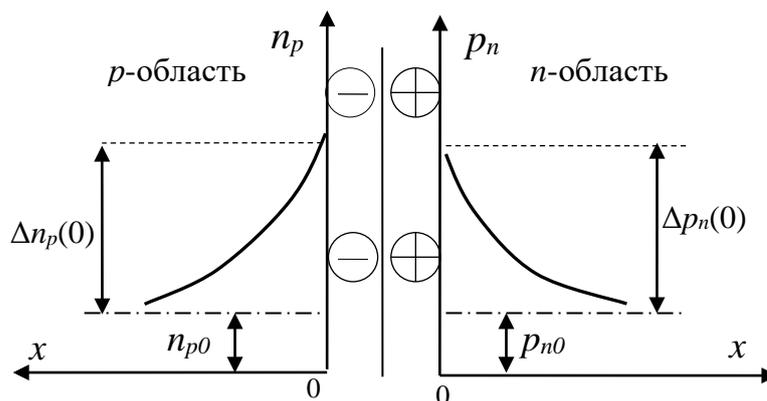


Рис. 6.5. Распределение концентраций инжектированных неосновных носителей по глубине  $p$ - и  $n$ -областей при прямом смещении симметричного перехода

В результате инжекции у границ перехода образуются избыточные концентрации неосновных носителей. Дырки диффундируют от границы вглубь  $n$ -области, постепенно рекомбинируя с электронами; такие же процессы будут происходить с участием электронов в  $p$ -области. На границе возникают следующие условия: за одинаковые промежутки времени число приходящих на границу основных носителей равно числу уходящих от границы вглубь  $p$ - и  $n$ -областей. Их избыточные концентрации на границе поддерживаются постоянными. Поэтому (см. раздел 4.6), если размеры областей значительно больше диффузионных длин для неосновных носителей, то концентрации инжектированных неосновных носителей будут уменьшаться по экспоненциальному закону при удалении от границы вглубь  $p$ - и  $n$ -областей (рис. 6.5).

### 6.1.3. Обратное включение $p$ – $n$ -перехода

Пусть к переходу подключен внешний источник напряжения  $U_{обр}$  минусом к  $p$ -области, а плюсом к  $n$ -области, т. е. полярностью, совпадающей с контактной разностью потенциалов  $\varphi_k$ . При этом равновесие на переходе также нарушается: потенциальный барьер для диффузии основных носителей возрастает до величины  $e(\varphi_k + U_{обр})$  (см. рис. 6.6.). Поскольку величина  $U_{обр}$  составляет обычно десятки-сотни вольт, диффузия основных носителей через переход прекращается, и их ток уменьшается до нуля. Вместе с тем условия для протекания через переход дрейфового тока неосновных носителей становятся еще более благоприятными, так как ускоряющая их разность потенциалов увеличилась на  $U_{обр}$ .

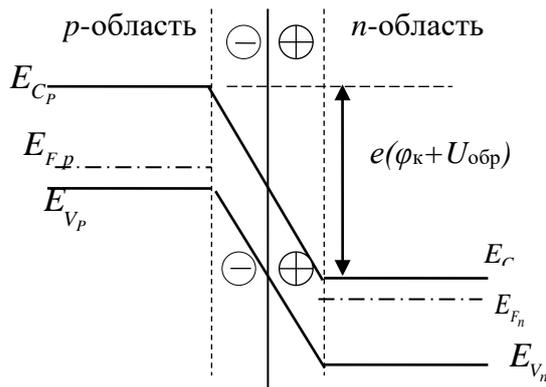


Рис. 6.6. Энергетическая диаграмма  $p$ – $n$ -перехода, смещенного в обратном направлении

Таким образом, при обратном смещении через  $p$ – $n$ -переход протекает дрейфовый ток неосновных носителей заряда: дырок из  $n$ -области в  $p$ -область и электронов из  $p$ -области в  $n$ -область. Этот ток намного меньше прямого диффузионного тока из-за того, что концентрация неосновных носителей очень мала по сравнению с концентрацией основных носителей. Например, для симметричного кремниевого перехода с концентрацией основных носителей в смежных областях  $\sim 10^{16} \text{ см}^{-3}$  концентрация неосновных носителей будет  $\sim 10^4 \text{ см}^{-3}$  (см. формулу (2.18) и табл. 2.2).

При повышении потенциального барьера ширина перехода увеличится:

$$l_{\text{обр}} = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_{\text{к}} + U_{\text{обр}})}{e} \cdot \frac{N_{\text{А}} + N_{\text{Д}}}{N_{\text{А}} \cdot N_{\text{Д}}}}. \quad (6.25)$$

Соответственно, для односторонних переходов при  $N_{\text{А}} \gg N_{\text{Д}}$ :

$$l_{\text{обр}} \approx l_n \approx \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_{\text{к}} + U_{\text{обр}})}{eN_{\text{Д}}}};$$

при  $N_{\text{Д}} \gg N_{\text{А}}$ :

$$l_{\text{обр}} \approx l_p \approx \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_{\text{к}} + U_{\text{обр}})}{eN_{\text{А}}}}. \quad (6.26)$$

Втягивание неосновных носителей полем  $p$ - $n$ -перехода в смежные области, где они становятся основными, называется экстракцией. Из-за экстракции неосновных носителей из областей, прилегающих к границам перехода, их концентрации здесь значительно уменьшаются. Оценим величину граничных концентраций неосновных носителей при обратном смещении перехода.

В случае обратного смещения в формуле (6.6) для контактной разности потенциалов равновесного перехода необходимо заменить  $\varphi_{\text{к}}$  на  $\varphi_{\text{к}} + U_{\text{обр}}$ . Концентрация основных носителей на границе будет такой же, как в объеме ( $n_n$ ), а концентрация неосновных носителей снизится до значения  $n_p(0)$ . Следовательно, для обратного смещения:

$$e(\varphi_{\text{к}} + U_{\text{обр}}) = kT \ln \frac{n_n}{n_p(0)}. \quad (6.27)$$

Отсюда после преобразований получим:  $n_p(0) = n_n \cdot e^{-\frac{e\varphi_{\text{к}}}{kT}} \cdot e^{-\frac{eU_{\text{обр}}}{kT}}$ .

Учитывая, что первые два множителя дают выражение для концентрации неосновных носителей в объеме  $p$ -области (см. формулу (6.5)), имеем:

$$n_p(0) = n_p e^{-\frac{eU_{\text{обр}}}{kT}}. \quad (6.28)$$

Величины обратных напряжений на переходе могут составлять десятки и сотни вольт. Таким образом, граничные концентрации

неосновных носителей при обратном смещении перехода снижаются до значений, близких к нулю; при удалении от границ перехода вглубь  $p$ - и  $n$ -областей значения  $n_p$  и  $p_n$  приближаются к равновесным значениям  $n_{p0}$  и  $p_{n0}$  (рис. 6.7).

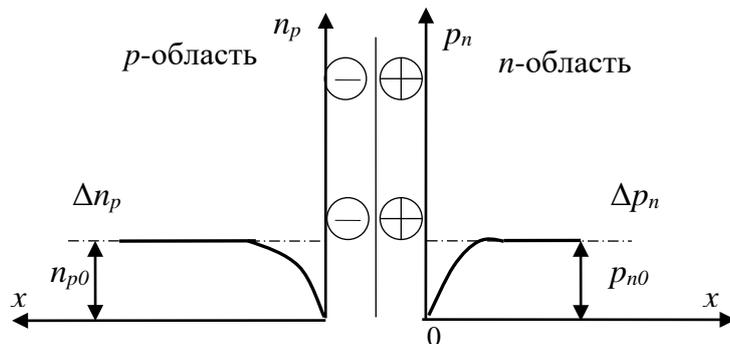
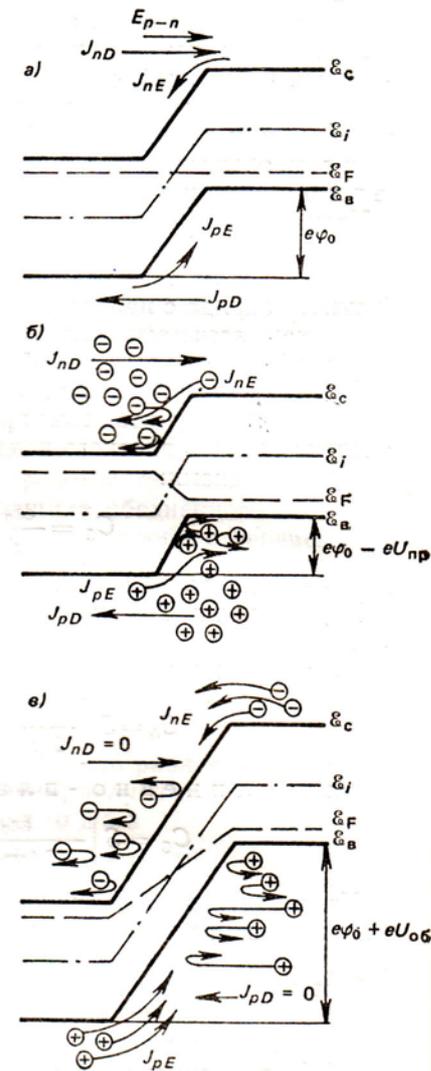


Рис. 6.7. Распределение концентраций неосновных носителей по глубине  $p$ - и  $n$ -областей при обратном смещении симметричного перехода

Таким образом, вблизи границ перехода появляются градиенты концентраций неосновных носителей, и число уходящих в смежные области неосновных носителей за счет экстракции оказывается равным числу приходящих из объема к границе перехода за счет диффузии.

С увеличением обратного напряжения  $p$ - $n$ -переход расширяется, разность потенциалов на нем возрастает, но значение тока через переход остается постоянным. Это объясняется тем, что градиенты концентрации неосновных носителей у границ перехода практически не изменяются, так как параметры диффузии ( $n_{p_0}$ ,  $p_{n_0}$ ,  $L_p$ ,  $L_n$ ) не зависят от приложенного внешнего напряжения. Поэтому ток, текущий через идеальный переход при обратном включении за счет экстракции, называется обратным током насыщения или тепловым током  $I_0$ . Величина  $I_0$  может измениться лишь при изменении температуры перехода за счет увеличения или уменьшения концентраций неосновных носителей.

Рис. 6.8. Потоки носителей заряда через переход в равновесном состоянии (а), при прямом (б) и обратном (в) смещениях



На рис. 6.8 показаны процессы переноса электронов и дырок через переход. В равновесии диффузионные токи основных носителей  $J_{nD}$  и  $J_{pD}$  уравновешены дрейфовыми токами неосновных носителей  $J_{nE}$  и  $J_{pE}$ , и суммарный ток через переход равен нулю. Понижение потенциального барьера для диффузии основных носителей при прямом смещении (рис. 6.8, б) вызывает прямой диффузионный ток основных носителей. Его величина резко возрастает при увеличении прямого напряжения, так как диффузия при снижении барьера все более облегчается. При этом ток неосновных носителей остается практически неизменным благодаря сохраняющейся ускоряющей разности потенциалов на  $p$ - $n$ -переходе.

сохраняющейся ускоряющей разности потенциалов на  $p$ - $n$ -переходе.

#### 6.1.4. Вольтамперная характеристика идеализированного $p$ - $n$ -перехода

Для облегчения расчетов выберем идеальную одномерную модель перехода, что позволит пренебречь падением напряжения в толще  $p$ - и  $n$ -областей и не учитывать процессы генерации и рекомбинации в пределах самого перехода.

В случае прямого смещения через переход протекают диффузионные токи основных носителей. Рассмотрим электронную составляющую прямого тока — диффузию электронов из  $n$ -области в

$p$ -область. Ось  $x$  в пределах  $p$ -области направим от границы перехода внутрь. В обозначениях формул  $e = |e| = 1,6 \cdot 10^{-19}$  Кл. Тогда:

$$I_n = -e \cdot S \cdot D_n \left. \frac{dn_p}{dx} \right|_{x=0}. \quad (6.29)$$

При инжекции в  $p$ -области возникает избыточная концентрация неосновных носителей (см. формулу (4.4)). Поэтому производную от концентрации по координате в (6.29) можно выразить следующим образом:

$$\left. \frac{dn_p}{dx} \right|_{x=0} = -\frac{\Delta n_p(0)}{L_n}.$$

Следовательно, диффузионный ток электронов при прямом смещении равен:

$$I_n = e \cdot S \cdot D_n \frac{\Delta n_p(0)}{L_n}. \quad (6.30)$$

Необходимо выяснить, каким образом избыточная концентрация неосновных носителей  $\Delta n_p(0)$  зависит от приложенного внешнего прямого напряжения. Воспользуемся выведенной выше формулой (6.6) для контактной разности потенциалов на переходе. При прямом смещении перехода концентрация основных носителей на границе в  $n$ -области будет такой же, как в объеме ( $n_{n_0}$ ), а концентрация неосновных носителей увеличится до значения  $n_{p_0} + \Delta n_p(0)$ . Следовательно, при прямом смещении пониженный потенциальный барьер для диффузии электронов равен:

$$e\varphi_k - eU = kT \ln \frac{n_{n_0}}{n_{p_0} + \Delta n_p(0)}. \quad (6.31)$$

Преобразуем это выражение:  $-\frac{e\varphi_k}{kT} + \frac{eU}{kT} = \ln \frac{n_{p_0} + \Delta n_p(0)}{n_{n_0}}$ .

Переходя к соотношению для экспонент, получим:

$$n_{n_0} e^{-\frac{e\varphi_k}{kT}} \cdot e^{\frac{eU}{kT}} = n_{p_0} + \Delta n_p(0).$$

Отсюда следует зависимость избыточной концентрации неосновных носителей, электронов в  $p$ -области, от прямого напряжения:

$$\Delta n_p(0) = n_{p_0} \left( e^{\frac{eU}{kT}} - 1 \right). \quad (6.32)$$

Аналогично, зависимость избыточной концентрации дырок в  $n$ -области от прямого напряжения:

$$\Delta p_n(0) = p_{n_0} \left( e^{\frac{eU}{kT}} - 1 \right). \quad (6.32. a)$$

Подставляя (6.32) в (6.29) получим зависимость электронной составляющей прямого тока  $p$ - $n$ -перехода с площадью  $S$  от приложенного напряжения:

$$I_n = e \cdot S \cdot \frac{D_n}{L_n} n_{p_0} \left( e^{\frac{eU}{kT}} - 1 \right). \quad (6.33)$$

Проведя те же действия для дырок, неосновных носителей в  $p$ -области, можно получить зависимость дырочной составляющей прямого тока от приложенного напряжения:

$$I_p = e \cdot S \cdot \frac{D_p}{L_p} p_{n_0} \left( e^{\frac{eU}{kT}} - 1 \right) \quad (6.34)$$

Прямой ток через переход — это сумма диффузионных токов основных носителей, и формула вольтамперной характеристики для прямого включения  $p$ - $n$ -перехода имеет вид:

$$I = e \cdot S \left( \frac{D_n}{L_n} n_{p_0} + \frac{D_p}{L_p} p_{n_0} \right) \left( e^{\frac{eU}{kT}} - 1 \right) = I_0 \left( e^{\frac{eU}{kT}} - 1 \right). \quad (6.35)$$

При обратном смещении через переход протекает малый дрейфовый ток неосновных носителей в противоположном направлении; его величина поддерживается практически постоянной за счет диффузии неосновных носителей из объема областей к границам перехода (см. раздел 6.4). Поэтому очевидно, что движущей силой, поддерживающей обратный ток, является диффузия, т. е. тот же механизм, на

основе которого выведена формула (6.35). Не выполняя соответствующих весьма объемных математических преобразований, отметим только их конечный результат: формула (6.35) является также и вольтамперной характеристикой для обратного включения  $p$ - $n$ -перехода. Общий вид вольтамперной характеристики показан на рис. 6.9.

Рассмотрим формулу вольтамперной характеристики  $p$ - $n$ -перехода (6.35) и оценим величину показателя экспоненты при прямом смещении. Величины прямых напряжений составляют от десятых долей до единиц Вольт; для комнатных температур величина  $kT \approx 2,6 \cdot 10^{-2}$  эВ. Следовательно, единицей в скобке можно пренебречь, и вольтамперная характеристика для прямого включения перехода хорошо описывается следующим выражением:

$$I_{\text{пр}} = I_0 \cdot e^{\frac{eU}{kT}}. \quad (6.36)$$

При обратном смещении напряжение отрицательно и обычно составляет десятки-сотни Вольт. Поэтому в (6.35) можно пренебречь величиной экспоненты по сравнению с единицей, и обратный ток идеализированного перехода равен по величине тепловому току  $I_0$ :

$$I_{\text{обр}} = -I_0. \quad (6.37)$$

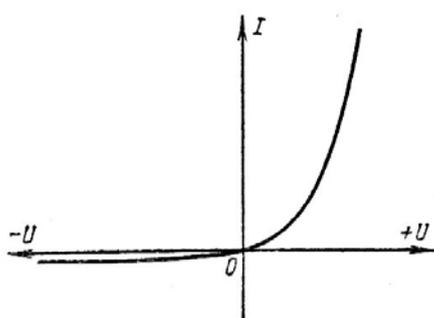


Рис. 6.9. Вольтамперная характеристика идеализированного  $p$ - $n$ -перехода. Ток при обратном смещении очень мал и не зависит от напряжения

Таким образом, обратный ток идеального перехода описывается формулой:

$$I_0 = e \cdot S \left( \frac{D_n}{L_n} n_{p_0} + \frac{D_p}{L_p} p_{n_0} \right). \quad (6.38)$$

Вольтамперная характеристика  $p-n$ -перехода зависит от его температуры и материала, из которого он изготовлен (см. рис. 6.10). Видно, что обратный ток переходов резко возрастает с увеличением температуры, и его величина при данном обратном напряжении для кремниевого перехода значительно меньше, чем для германиевого. Прямой ток при этом также растет, однако в процентном отношении его прирост гораздо меньше, чем прирост обратного тока. Видно также, что прямая ветвь вольтамперной характеристики кремниевого диода сдвинута в область более высоких прямых напряжений.

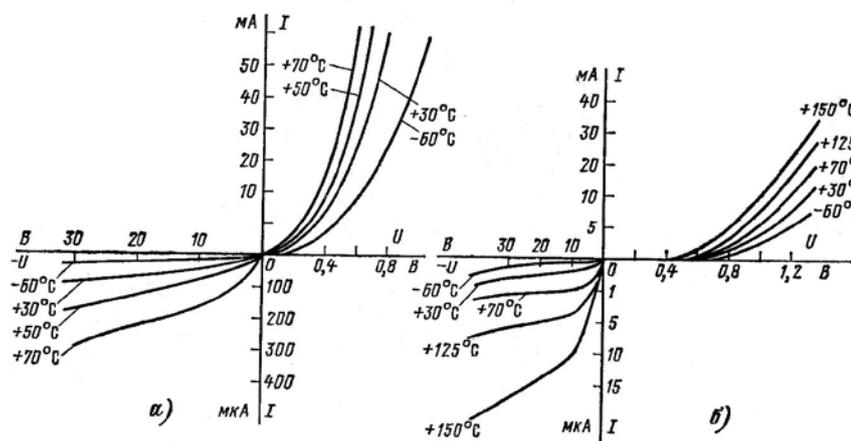


Рис. 6.10. Зависимость вольтамперных характеристик германиевого (а) и кремниевого (б) диодов от температуры

Физическую причину этих зависимостей можно выяснить, преобразовав выражение (6.38) с использованием соотношения (2.18):

$$n_{p_0} = \frac{n_i^2}{p_{p_0}}; \quad p_{n_0} = \frac{n_i^2}{n_{n_0}}.$$

Подставив эти величины в (6.38), и учитывая, что собственная концентрация подвижных носителей  $n_i$  зависит от температуры и ширины запрещенной зоны полупроводника (2.13), получим:

$$I_0 = e \cdot S \cdot N_C N_V \left( \frac{D_n}{L_n p_{p_0}} + \frac{D_p}{L_p n_{n_0}} \right) e^{-\frac{\Delta E_s}{kT}}. \quad (6.39)$$

В полученной формуле (6.39) явным образом прослеживается зависимость обратного тока идеального перехода и от температуры, и от разновидности полупроводникового материала. Наблюдается экспоненциальный рост обратного тока с увеличением температуры и такое же резкое уменьшение его величины при увеличении ширины запрещенной зоны. Это — прямое следствие того, что величина обратного тока определяется тепловой генерацией неосновных носителей заряда в смежных областях  $p$ – $n$ -перехода.

При рассмотрении этих же зависимостей для тока при прямом включении перехода можно впасть в ошибку при использовании формулы (6.36): из нее на первый взгляд следует, что прямой ток должен уменьшаться с ростом температуры, что не соответствует реальному поведению вольтамперных характеристик. Это противоречие возникает из-за того, что не учитывается зависимость  $I_0$  от температуры (6.39).

С учетом этого факта зависимость прямого тока перехода от температуры имеет вид:

$$I_{\text{пр}} = e \cdot S \cdot N_C N_V \left( \frac{D_n}{L_n p_{p_0}} + \frac{D_p}{L_p n_{n_0}} \right) e^{-\frac{\Delta E_3}{kT}} \cdot e^{\frac{eU}{kT}} = C \cdot e^{-\frac{\Delta E_3 - eU}{kT}}. \quad (6.40)$$

Так как величины прямых напряжений таковы, что  $eU$  всегда меньше ширины запрещенной зоны  $\Delta E_3$ , то показатель экспоненты в формуле (6.40) отрицателен, и прямой ток увеличивается с ростом температуры. Физической причиной этого является прирост концентрации основных носителей за счет тепловой генерации. Увеличение прямого тока происходит также из-за уменьшения высоты потенциального барьера  $p$ – $n$ -перехода для их диффузии при повышении температуры. Преобразуем формулу (6.6) для потенциального барьера с учетом соотношений (2.13) и (2.18):

$$e\varphi_k = kT \ln \frac{n_{n_0}}{n_{p_0}} = kT \ln \frac{n_{n_0} p_{p_0}}{n_i^2} = kT \ln(n_{n_0} p_{p_0}) - kT \ln N_C N_V - kT \ln \left( e^{-\frac{\Delta E_3}{kT}} \right)$$

Отсюда получаем зависимость величины потенциального барьера от температуры:

$$e\varphi_{\text{к}} = \Delta E_{\text{з}} - kT \ln \frac{N_{\text{с}} N_{\text{в}}}{n_{\text{н}0} p_{\text{р}0}}, \quad (6.41)$$

при этом необходимо иметь в виду, что при увеличении температуры происходит также и небольшое уменьшение ширины запрещенной зоны полупроводника. Таким образом, потенциальный барьер для диффузии основных носителей уменьшается с ростом температуры, что приводит к росту прямого тока через переход.

При использовании полупроводников с более широкой запрещенной зоной потенциальный барьер перехода  $e\varphi_{\text{к}}$  возрастает. Это приводит к уменьшению прямого тока при данном прямом напряжении. В результате вольтамперная характеристика перехода на основе полупроводника с более широкой запрещенной зоной сдвигается вправо, в область более высоких прямых напряжений. Так, например, резкое увеличение прямого тока для германиевого  $p-n$ -перехода ( $\Delta E_{\text{з}} = 0,66$  эВ) при комнатной температуре начинается при  $U \approx 0,28$  В, а для кремниевого перехода ( $\Delta E_{\text{з}} = 1,12$  эВ) — при  $U \approx 0,7$  В.

## **6.2. КОНТАКТ МЕЖДУ ПОЛУПРОВОДНИКАМИ ОДИНАКОВОГО ТИПА ПРОВОДИМОСТИ С РАЗНОЙ КОНЦЕНТРАЦИЕЙ ПРИМЕСИ**

Для контактов такого типа используют обозначения  $p^+ - p$  и  $n^+ - n$ . Знак «+» относится к полупроводнику с более высокой концентрацией примеси. На рис. 6.11 приведены энергетические диаграммы таких полупроводников и диаграммы соответствующих контактов.

При контакте между полупроводниками одинакового типа проводимости приграничный слой, обедненный подвижными носителями заряда, образуется только у границ областей  $n^+$  и  $p^+$ , а граничный слой областей  $n$  и  $p$  обогащается основными носителями заряда.

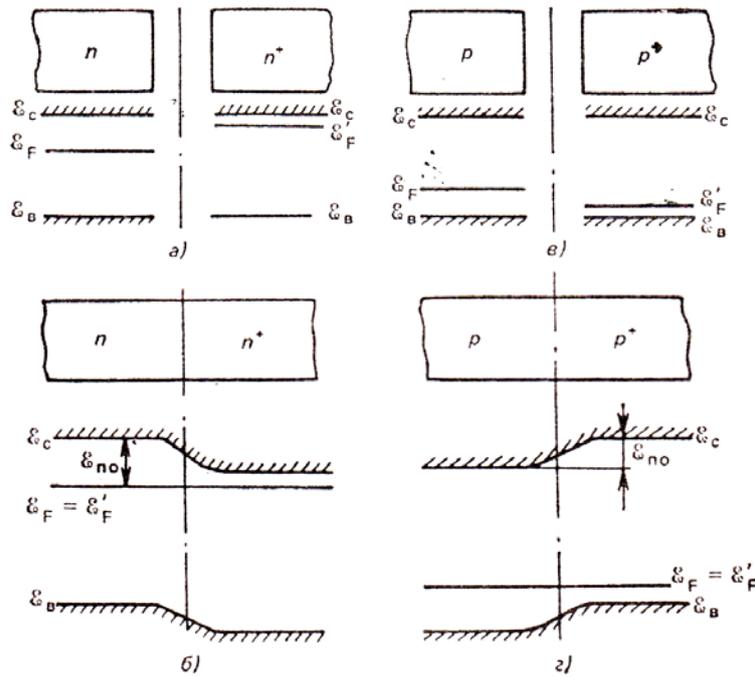


Рис. 6.11. Энергетические диаграммы полупроводниковых контактов одного типа проводимости: *а, в* — полупроводники до установления контакта, *б, г* — контакты  $n^+ - n$  и  $p^+ - p$  соответственно

Высота потенциального барьера в этом случае рассчитывается в той же последовательности, что и для обычного  $p - n$ -перехода с заменой соответствующих концентраций подвижных носителей в смежных областях на  $n_{n^+}$  и  $n_n$  или  $p_{p^+}$  и  $p_p$  (см. формулы (6.2) – (6.7)). Потенциальные барьеры для полупроводниковых контактов одного типа проводимости выражаются формулами:

$$e\varphi_K = kT \ln \frac{n_{n^+}}{n_n} \text{ — для перехода } n^+ - n,$$

$$e\varphi_K = kT \ln \frac{p_{p^+}}{p_p} \text{ — для перехода } p^+ - p.$$

Эти величины в отличие от  $p - n$ -переходов невелики, ( $< 0,1$  эВ), так как обычно  $n_{n^+}/n_n$  и  $p_{p^+}/p_p \approx 10$ .

Такие переходы не обладают выпрямляющим эффектом: при смене полярности подключения ток меняет направление, но порядок

его величины остается прежним. Кроме того, в них не наблюдается инжекция неосновных носителей заряда. Например, при подаче напряжения плюсом к  $n$ -области и минусом к области  $n^+$  высота потенциального барьера для основных носителей понизится, и электроны будут переходить из области  $n^+$  в область  $n$ . Ток дырок в противоположном направлении будет ничтожно мал, так как в  $n$ -областях они являются неосновными носителями с концентрацией на много порядков меньшей, чем концентрация электронов.

### ***Вопросы к разделам 6.1 – 6.2***

1. Что такое  $p$ - $n$ -переход, каковы его разновидности?
2. Перечислите условия, при которых  $p$ - $n$ -переход можно считать идеальным.
3. Опишите процесс установления равновесия на контакте полупроводников с разным типом проводимости.
4. Каковы физические причины возникновения контактного электрического поля на  $p$ - $n$ -переходе? Какими зарядами оно образовано?
5. Какие встречные токи через переход равны друг другу в равновесном состоянии?
6. От чего и каким образом зависит ширина  $p$ - $n$ -перехода в равновесном состоянии?
7. Чем определяется полярность внешнего напряжения при прямом смещении  $p$ - $n$ -перехода?
8. Какова физическая природа прямого тока через переход?
9. Чем определяется полярность внешнего напряжения при обратном смещении  $p$ - $n$ -перехода?
10. Какова физическая природа обратного тока через переход?
11. Какими физическими параметрами определяется величина обратного тока идеального  $p$ - $n$ -перехода?
12. Какими свойствами обладают переходы между полупроводниками с одинаковым типом проводимости?
13. Поясните, по каким причинам переходы между полупроводниками с одинаковым типом проводимости являются невыпрямляющими.

### 6.3. РЕАЛЬНЫЙ ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

Реальные  $p$ - $n$ -переходы, как правило, несимметричные, и в подавляющем большинстве случаев их можно считать односторонними. Полупроводник с большей концентрацией примеси, и соответственно, с более высокой концентрацией основных носителей называется эмиттером, второй полупроводник с меньшей концентрацией основных носителей, базой. Односторонние переходы обычно обозначают символами  $p^+ - n$  или  $n^+ - p$ .

При рассмотрении физических процессов в реальном  $p$ - $n$ -переходе необходимо учитывать все факторы, которые остались «за бортом» при анализе идеализированного перехода, а именно:

- процессы генерации электронно-дырочных пар и их рекомбинации в пределах ширины перехода;
- падение напряжения на  $p$ - и  $n$ -областях при протекании прямого тока;
- поверхностные явления, в частности, токи утечки при больших обратных напряжениях.

Будем считать, что вблизи омических контактов в  $p$ - и  $n$ -полупроводниках всегда существует равновесная концентрация носителей заряда, и, следовательно, влияние этих контактов на свойства  $p$ - $n$ -перехода несущественно. Выясним, как изменится вольтамперная характеристика перехода при учете указанных выше факторов.

#### **6.3.1. Обратное смещение. Токи генерации и утечки**

Тепловая генерация электронно-дырочных пар происходит не только в  $p$ - и  $n$ -областях, но и в самом  $p$ - $n$ -переходе. Образовавшиеся электроны и дырки разделяются электрическим полем перехода и перебрасываются в области, где они становятся основными носителями. Возникающий при этом дрейфовый ток называется током генерации; он совпадает по направлению с обратным током насыщения  $I_0$  идеализированного перехода.

Следовательно, ток генерации определяется полным числом носителей, образовавшихся в  $p$ - $n$ -переходе в единицу времени. Число подвижных носителей, возникающих в единице объема в единицу времени, т. е. скорость генерации — это отношение  $n_i/\tau$ , где  $\tau$  — время жизни носителей в области перехода,  $n_i$  — собственная концентрация подвижных носителей. Генерация происходит в объеме  $p$ - $n$ -перехода  $S \cdot l$  ( $S$  — площадь перехода,  $l$  — его ширина). Тогда ток генерации можно представить в следующем виде:

$$I_{\Gamma} = eSl \frac{n_i}{\tau}. \quad (6.42)$$

Отсюда следует, что ток генерации увеличивается с ростом обратного напряжения из-за расширения перехода (см. формулы (6.26)).

Оценим долю тока генерации в полном обратном токе реального перехода. Для этого рассмотрим отношение тока генерации к тепловому току  $I_0$  (6.39). Это отношение оказывается обратно пропорциональным собственной концентрации носителей:

$$\frac{I_{\Gamma}}{I_0} \sim \frac{1}{n_i} \sim e^{\frac{\Delta E_g}{2kT}}. \quad (6.43)$$

Видно, что доля тока генерации в полном обратном токе тем больше, чем шире запрещенная зона полупроводника и чем ниже температура. Так, для кремниевого диода средней мощности при комнатной температуре и при  $U_{обр} = 10$  В ток генерации превышает тепловой ток на 5 порядков:  $I_{\Gamma} \approx 10^{-9}$  А,  $I_0 \approx 10^{-14}$  А.

Второй составляющей обратного тока является ток утечки. Этот ток возникает из-за того, что реальные переходы всегда имеют участки, выходящие на поверхность кристалла. Поэтому между смежными областями перехода могут образовываться проводящие поверхностные «дорожки». Ток утечки подчиняется закону Ома и почти не зависит от температуры. В кремниевых  $p$ - $n$ -переходах токи утечки очень малы, так как поверхность кристалла покрыта тонким защитным оксидным слоем.

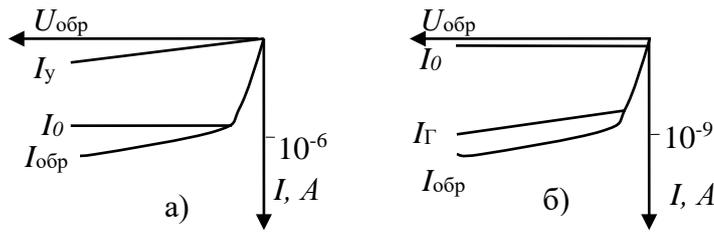


Рис. 6.12. Составляющие обратных токов реальных  $p-n$ -переходов, изготовленных из германия (а) и кремния (б). Указаны порядки величин обратных токов для диодов средней мощности

Из рис. 6.12 видно, что величину обратного тока германиевого перехода определяют, в основном, тепловой ток  $I_0$  и ток утечки  $I_y$ , в то время как в кремниевом переходе обратный ток практически полностью обусловлен током генерации  $I_G$ . Величина полного обратного тока кремниевых переходов на несколько порядков меньше, чем германиевого перехода с такими же параметрами.

### 6.3.2. Прямое смещение. Ток рекомбинации и влияние сопротивления базы

В реальном  $p-n$ -переходе при прямом смещении происходит также и рекомбинация подвижных носителей. Рис. 6.13 поясняет механизм ее возникновения. Электроны  $n$ -области и дырки  $p$ -области с энергиями, меньшими, чем высота барьера для диффузии основных носителей, могут проникать в область обеднения и рекомбинировать друг с другом. Рекомбинировать могут также и носители, проходящие через переход за счет инжекции.

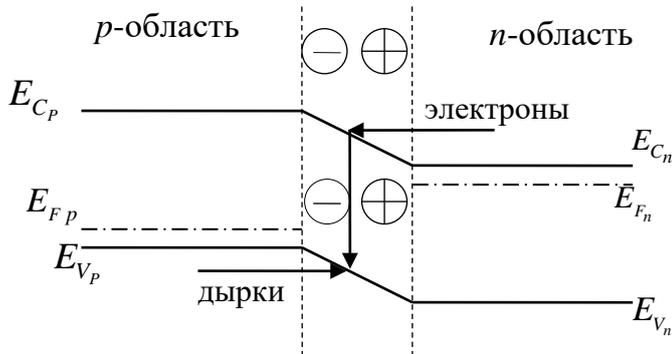


Рис. 6.13. Механизм возникновения тока рекомбинации в реальном  $p-n$ -переходе

При этом электроны уходят из  $n$ -области, а дырки — из  $p$ -области, и направление их движения совпадает с направлением прямого тока через переход. Вследствие этого в реальном  $p$ - $n$ -переходе возникает дополнительный прямой ток — ток рекомбинации.

Ток рекомбинации определяется количеством основных носителей, проникших в область перехода и исчезнувших за счет рекомбинации во всем его объеме в единицу времени. Следовательно, он должен сильно зависеть от высоты потенциального барьера и от температуры. Ток рекомбинации можно определить по следующей упрощенной формуле:

$$I_{\text{рек}} = eSl \frac{n_i}{\tau} e^{\frac{eU}{2kT}}. \quad (6.44)$$

Отсюда видно, что ток рекомбинации, как и диффузионный ток (6.36), экспоненциально возрастает с увеличением прямого напряжения. Вместе с тем, при сравнении формул (6.36) и (6.44) видно, что зависимость  $I_{\text{рек}}$  от прямого напряжения более слабая: показатель экспоненты в (6.44) вдвое меньше. Таким образом, ток рекомбинации является основной составляющей прямого тока при малых прямых напряжениях. Так, например, для кремниевого  $p$ - $n$ -перехода при напряжении менее 0,4 В в прямом токе преобладает ток рекомбинации, а при  $U > 0,4$  В — диффузионный ток основных носителей.

При протекании через реальный переход больших прямых токов необходимо учитывать падение напряжения не только на  $p$ - $n$ -переходе, но и на объемном сопротивлении базы:  $U_{\text{б}} = I \cdot r_{\text{б}}$ . Величина  $r_{\text{б}}$  обычно находится в пределах от десяти до ста Ом. Таким образом, учитывая формулу (6.36), можно записать:

$$U = U_{p-n} + U_{\text{б}} = \frac{kT}{e} \ln \frac{I}{I_0} + I \cdot r_{\text{б}}. \quad (6.45)$$

Это означает, что при достаточно больших прямых токах на вольтамперной характеристике появится почти линейный участок (см. рис. 6.14, кривая 2).

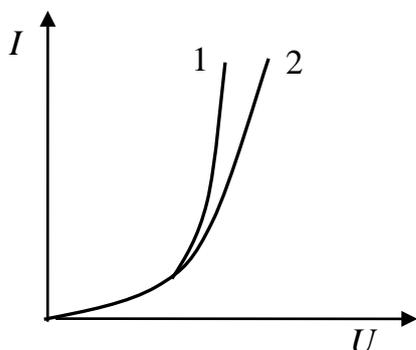


Рис. 6.14. Прямая ветвь вольтамперной характеристики  $p-n$ -перехода.  
1 — для идеального перехода ( $r_B = 0$ ),  
2 — измененная за счет падения напряжения на объемном сопротивлении базы

Проведем оценку степени влияния падения напряжения  $U_B$  на форму вольтамперной характеристики. Пусть  $r_B = 20$  Ом и прямой ток составляет 100 мА. Тогда  $U_B = I \cdot r_B = 10^{-1} \cdot 20 = 2$  В (!). На кривых  $I(U)$  для кремниевых диодов Д210 прямой ток 100 мА наблюдается при напряжении  $U = U_{p-n} + I \cdot r_B \sim 0,85$  В. Такое расхождение объясняется модуляцией сопротивления базы: при больших прямых токах доля инжектированных в базу диода неосновных носителей оказывается очень высокой, и сопротивление базы уменьшается в несколько раз. Этот эффект проявляется как в статическом, так и в импульсном режиме при высоких уровнях инжекции.

#### 6.4. P–N-ПЕРЕХОД НА ОСНОВЕ ВЫРОЖДЕННЫХ ПОЛУПРОВОДНИКОВ

Если концентрация примесей в полупроводнике очень высока ( $\sim 10^{20}$  см<sup>-3</sup>), то примесные атомы располагаются в кристаллической решетке на относительно малых расстояниях друг от друга, и в отличие от обычных полупроводников образуют систему взаимодействующих частиц. Следствием этого является расщепление единого примесного уровня энергии в энергетическую зону, которая при таких высоких концентрациях примеси перекрывается с зоной разрешенных состояний. В полупроводнике  $n$ -типа примесная зона сливается с

зоной проводимости, в  $p$ -полупроводнике с валентной зоной. Таким образом, при очень высоких концентрациях примесей дно зоны проводимости  $n$ -полупроводника совпадает с низшим уровнем примесной донорной зоны, а потолок валентной зоны полупроводника  $p$ -типа — самый верхний уровень примесной зоны акцепторов.

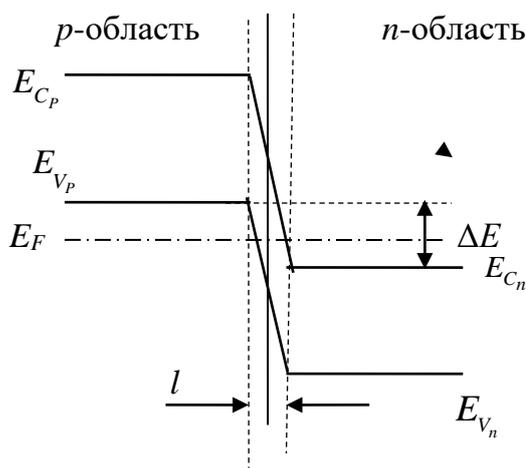


Рис. 6.15. Энергетическая диаграмма  $p$ – $n$ -перехода на основе вырожденных полупроводников в равновесном состоянии

При этом уровень Ферми располагается в зоне разрешенных энергетических состояний (рис. 6.15), и для основных носителей образуется непрерывная энергетическая диаграмма, характерная для металлов. Полупроводники, обладающие такими свойствами, называются вырожденными.  $P$ – $n$ -переход на основе вырожденных полупроводников также обладает особыми свойствами. Энергетическая диаграмма такого перехода приведена на рис. 6.15.

Видно, что на энергетической диаграмме образуется полоса энергий электронов  $\Delta E$ , в пределах которой уровни валентной зоны  $p$ -области находятся напротив уровней зоны проводимости  $n$ -области. Ширина разделяющего их потенциального барьера  $l$  очень мала (десятки нм) из-за высокой концентрации примесей (6.19), и вследствие этого возникает перенос электронов в обоих направлениях без изменения энергии за счет туннельного эффекта (см. раздел 4.3). В состоянии равновесия встречные туннельные токи электронов равны, и суммарный туннельный ток равен нулю.

Подача на переход внешнего напряжения нарушает равенство встречных туннельных токов. При увеличении прямого смещения туннельный ток возрастает, достигает максимума, а затем спадает до нуля. Рассмотрим причину такой зависимости с использованием энергетических диаграмм.

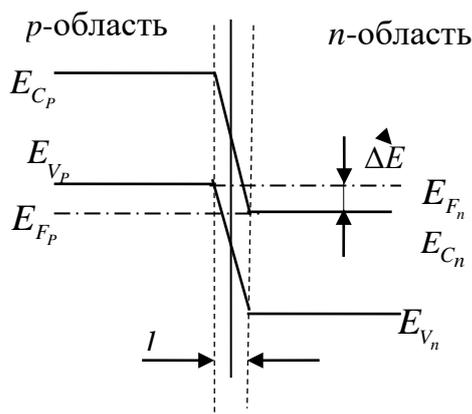


Рис. 6.16. Энергетическая диаграмма прямо смещенного  $p$ - $n$ -перехода на основе вырожденных полупроводников для случая максимального туннельного тока

Для объяснения такого характера изменения прямого туннельного тока следует учесть, что максимум плотности состояний электронов в зоне проводимости находится примерно посередине между ее дном  $E_{C_n}$  и уровнем Ферми. Максимум плотности состояний дырок в валентной зоне располагается примерно посередине между уровнем Ферми и ее верхней границей  $E_{V_p}$ .

При увеличении прямого смещения число туннельных переходов электронов из зоны проводимости  $n$ -полупроводника на уровни дырок в валентной зоне  $p$ -полупроводника растет, так как все большее число свободных состояний в валентной зоне оказывается при той же энергии, что и занятые состояния в зоне проводимости. Число туннельных переходов электронов в обратном направлении соответственно уменьшается. Наибольшее значение прямого туннельного тока соответствует максимальному перекрытию по энергии свободных уровней у потолка валентной зоны  $p$ -полупроводника и занятых электронами уровней зоны проводимости полупроводника  $n$ -типа.

Энергетическая диаграмма на рис. 6.16 соответствует именно такому случаю; при этом уровень Ферми полупроводника  $p$ -типа

совпадает с нижней границей зоны проводимости, а уровень Ферми полупроводника  $n$ -типа находится при той же энергии, что и верхняя граница валентной зоны.

Дальнейшее увеличение прямого смещения приводит к тому, что перекрытие зон по энергиям  $\Delta E$  уменьшается, при совпадении границ  $E_{C_n}$  и  $E_{V_p}$  оно исчезает и прямой туннельный ток становится равным нулю. При последующем возрастании прямого напряжения начинают происходить те же процессы, что и в обычном  $p$ - $n$ -переходе (см. раздел 6.3). Прямой ток опять растет, но только за счет инжекции неосновных носителей.

При увеличении обратного смещения (рис. 6.17) все большее число занятых электронами уровней валентной зоны  $p$ -полупроводника оказывается при тех же значениях энергии, что и свободные уровни зоны проводимости полупроводника  $n$ -типа. Ширина разделяющего их потенциального барьера по-прежнему очень мала, и поэтому наблюдается резкое возрастание туннельного тока электронов, протекающего в обратном направлении: из полупроводника  $p$ -типа в полупроводник  $n$ -типа.

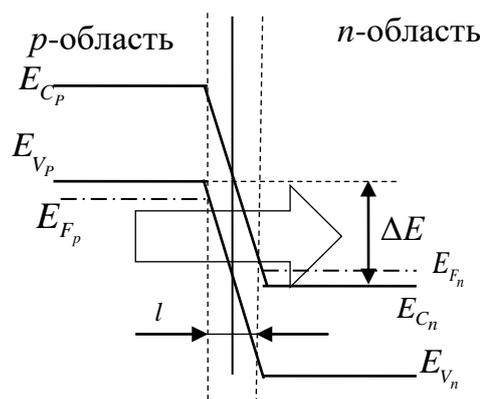


Рис. 6.17. Энергетическая диаграмма  $p$ - $n$ -перехода на основе вырожденных полупроводников при обратном смещении. Стрелка указывает направление перемещения электронов за счет туннельного эффекта

Очевидно, что вольтамперная характеристика (ВАХ)  $p-n$ -перехода на основе вырожденных полупроводников (рис. 6.18, а, кривая 3) должна резко отличаться от таковой для обычного перехода (рис. 6.18). Она получается как результат наложения двух характеристик: для туннельного тока и для обычного  $p-n$ -перехода (кривые 2 и 1 на рис. 6.18, а, соответственно).

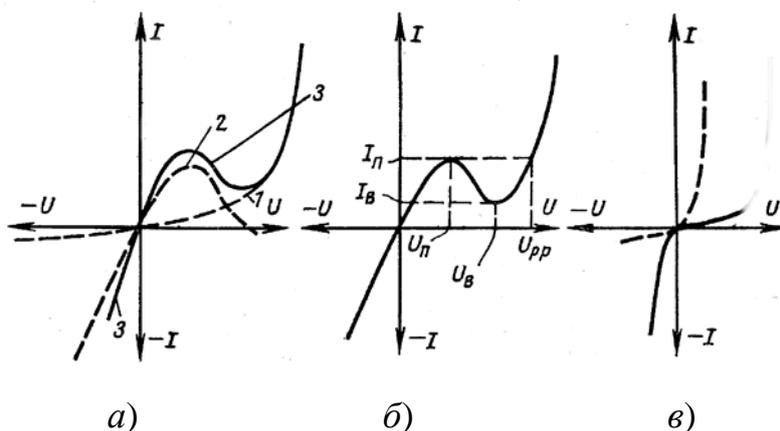


Рис. 6.18. Вольтамперные характеристики  $p-n$ -переходов на основе вырожденных полупроводников:

а, б — туннельного диода, в — обращенного диода (сплошная линия)

Прямая ветвь вольтамперной характеристики содержит область отрицательного дифференциального сопротивления — падающий участок кривой, который отражает уменьшение перекрытия зон по энергиям  $\Delta E$ . Рост прямого тока после прохождения минимального значения имеет ту же причину, что и для обычного перехода на основе невырожденных полупроводников (так называемая инжекционная ветвь вольтамперной характеристики). При обратном смещении, как указано выше, туннельный ток меняет направление и резко возрастает с увеличением напряжения на переходе.

На основе таких переходов изготавливаются туннельные диоды. Туннельный диод — это диод на основе вырожденного полупроводника, в котором туннельный эффект приводит к появлению на вольтамперной характеристике при прямом напряжении участка отрицательного дифференциального сопротивления. Для их изготовления используют германий или арсенид галлия с высокой концентрацией

примесей  $\sim 10^{20} \text{ см}^{-3}$ . Рабочей частью вольтамперной характеристики туннельного диода является ее спадающий участок с отрицательной дифференциальной проводимостью (рис. 6.18, б) в интервале напряжений  $U_{\text{п}} < U < U_{\text{в}}$ .

Характерное время туннелирования  $10^{-13}$ – $10^{-14}$  с; кроме того, туннельные диоды работают при очень малых прямых напряжениях (до появления инжекционного тока), и накопление избыточных неосновных носителей в базе не происходит. Поэтому туннельные диоды практически безынерционны и могут работать на частотах до сотен ГГц (мм-диапазон). В зависимости от варианта схемы устройства с туннельным диодом он может служить либо усилителем, либо переключателем с рекордно высокой граничной частотой.

На рис. 6.18, б показаны параметры туннельного диода:  $I_{\text{п}}$ ,  $U_{\text{п}}$  — ток и напряжение пика,  $I_{\text{в}}$ ,  $U_{\text{в}}$  — ток и напряжение впадины,  $U_{\text{pp}}$  — напряжение раствора вольтамперной характеристики — напряжение, при котором ток на инжекционной ветви равен току пика.

Для вырожденных полупроводников существует понятие «критическая концентрация примесей», при которой уровень Ферми  $n$ -полупроводника совпадает с дном зоны проводимости, а в  $p$ -полупроводнике — с потолком валентной зоны. Если для изготовления  $p$ – $n$ -перехода используются полупроводники с критической концентрацией примесей ( $\sim 10^{19} \text{ см}^{-3}$ ), то на его равновесной энергетической диаграмме не будет наблюдаться начального перекрытия зон  $\Delta E$ . Поэтому в идеальном случае при прямом смещении туннельный ток будет отсутствовать, и будет наблюдаться, как и в обычном переходе, только инжекционный прямой ток. Перекрытие зон возникнет лишь при обратном смещении, и обратная ветвь его вольтамперной характеристики будет такой же, как и у туннельного диода (рис. 6.18, в).

Такие переходы используются в обращенных диодах. Их название объясняется тем, что, в противоположность обычным диодам, обратный ток намного больше прямого вплоть до прямого напряжения,

при котором начинается инжекция неосновных носителей. Благодаря очень высокому быстродействию, эти диоды используются как детекторы и смесители в области сверхвысоких частот.

В реальных обращенных диодах из-за сложности изготовления полупроводников, с критической концентрацией примесей, на прямом участке характеристики наблюдается небольшой «туннельный максимум», однако величины прямых туннельных токов очень малы.

### ***Вопросы к разделам 6.3 – 6.4***

1. Какие факторы вводятся в рассмотрение для реального  $p-n$ -перехода (в отличие от идеального)?
2. Какие дополнительные составляющие обратного тока существуют в реальном переходе? Назовите физические причины их возникновения.
3. Какие дополнительные составляющие прямого тока существуют в реальном переходе? Назовите физические причины их возникновения.
4. Что означает термин «модуляция сопротивления базы»?
5. Какова причина расщепления примесных энергетических уровней в примесные зоны при повышении концентрации доноров и акцепторов до величины  $\sim 10^{20} \text{ см}^{-3}$ ?
6. Где расположен уровень Ферми на энергетических диаграммах вырожденных примесных полупроводников?
7. Какова причина возникновения туннельных токов через  $p-n$ -переход на основе вырожденных полупроводников?
8. Поясните причину наличия максимума и минимума тока на прямой ветви вольтамперной характеристики  $p-n$ -перехода на основе вырожденных полупроводников.
9. Происходит ли в идеальном  $p-n$ -переходе на основе вырожденных полупроводников инжекция неосновных носителей? Почему?
10. В каком случае, и по какой причине в идеальном  $p-n$ -переходе на основе вырожденных полупроводников не наблюдается туннельный ток при прямом смещении?

## 6.5. ПРОБОЙ ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

Пробой  $p$ – $n$ -перехода — это явление резкого увеличения обратного тока при достижении критической величины обратного напряжения. В зависимости от физических явлений, приводящих к пробую, различают лавинный, туннельный и тепловой пробой.

### 6.5.1. Лавинный пробой

Физической причиной лавинного пробоя является ударная генерация электронно-дырочных пар (см. раздел 4.2, рис. 4.2, *a*) в области  $p$ – $n$ -перехода, где приложено электрическое поле, обусловленное контактной разностью потенциалов и внешним напряжением. Образовавшиеся электроны и дырки ускоряются электрическим полем перехода и перебрасываются в те области, где они становятся основными носителями. Направление их тока совпадает с направлением обратного тока через переход. Таким образом, при достижении напряжения лавинного пробоя  $U_{\text{проб. лав.}}$  — наблюдается резкое возрастание обратного тока перехода.

Для возникновения лавинного размножения носителей заряда необходимо, чтобы на ширине перехода укладывалось большое число длин свободного пробега. Поэтому лавинный пробой развивается в относительно широких переходах на основе слаболегированных полупроводников с концентрацией примесей не более  $10^{17} \text{ см}^{-3}$ .

Основные факторы, влияющие на величину  $U_{\text{проб. лав.}}$ : ширина запрещенной зоны полупроводника  $\Delta E_z$ , концентрация примесей в базе и температура перехода.

Ширина запрещенной зоны, как известно, определяется величиной энергии связи валентных электронов. Поэтому в переходах, изготовленных из полупроводников с большей  $\Delta E_z$ , при прочих равных условиях напряжение лавинного пробоя выше.

Концентрация примесей в базе определяет ширину перехода  $l$ , и, следовательно, величину напряженности электрического поля в нем ( $\varepsilon = \varepsilon_k + U_{\text{обр}}/l$ ). При снижении концентрации примесей ширина

перехода увеличивается, что приводит к уменьшению  $\epsilon$  и, соответственно, к увеличению напряжения лавинного пробоя. Влияние вышеуказанных факторов показано на рис. 6.19.

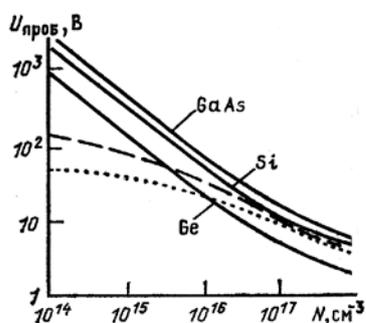


Рис. 6.19. Зависимость напряжения лавинного пробоя  $p$ - $n$ -перехода от концентрации примесей в базе и от материала, из которого изготовлен переход. Наибольшая величина  $\Delta E_3$  (1,4 эВ) — у арсенида галлия, наименьшая (0,66 эВ) — у германия

При увеличении температуры  $p$ - $n$ -перехода напряжение лавинного пробоя возрастает. Физической причиной этого является уменьшение средней длины свободного пробега подвижных носителей в полупроводнике, при этом уменьшается и энергия, которую носитель может приобрести на этой длине в электрическом поле. Поэтому для того, чтобы электрон проводимости, ускоряясь в электрическом поле на меньшей длине пробега, набрал энергию, достаточную для ионизации атома полупроводника, необходимо увеличить обратное напряжение на переходе.

Для оценки влияния температуры на напряжение пробоя перехода используется температурный коэффициент напряжения:

$$\alpha = \frac{\Delta U}{U_{\text{проб}} \cdot \Delta T} \quad (1/\text{K}). \quad (6.46)$$

Для лавинного пробоя этот коэффициент положителен, так как  $\Delta T$  и  $\Delta U$  имеют одинаковые знаки.

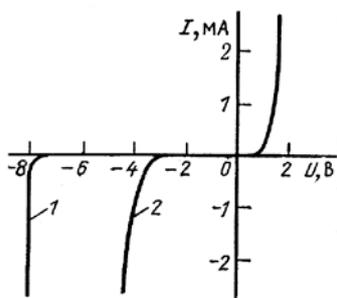


Рис. 6.20. Вольтамперные характеристики кремниевых  $p$ - $n$ -переходов на участках электрического пробоя.  
1 — лавинный пробой,  
2 — туннельный пробой

На рис. 6.20 приведены вольтамперные характеристики двух разновидностей электрического пробоя кремниевого  $p-n$ -перехода. Видно, что для лавинного пробоя (кривая 1) характерен резкий рост обратного тока при практически неизменном обратном напряжении.

### 6.5.2. Туннельный пробой $p-n$ -перехода

Электрический пробой  $p-n$ -перехода, обусловленный явлением полевой генерации электронно-дырочных пар называется туннельным пробоем (см. раздел 4.3, рис. 4.2, б). При подаче обратного напряжения  $U_{обр}$  на переход величина напряженности электрического поля в нем может достигать величин  $\sim 10^6$  В/см и более, при которых становится возможным туннельный перенос электронов с энергетических уровней валентной зоны на уровни зоны проводимости. Очевидно, что для достижения достаточно большой вероятности туннельного переноса электронов сквозь потенциальный барьер перехода необходимо, чтобы его ширина  $l$  была очень малой. Это достигается тем, что для изготовления таких переходов используются высоколегированные невырожденные полупроводники с концентрациями примесей  $\sim 10^{18}$  см $^{-3}$ .

Из рис. 6.21 видно, что в этом случае при обратном смещении перехода на его энергетической диаграмме образуется полоса энергий электронов  $\Delta E_{тун}$ , в пределах которой заняты электронами уровни валентной зоны  $p$ -области находятся напротив свободных уровней зоны проводимости  $n$ -области. Ширина потенциального барьера для туннелирования  $\Delta x$  меньше ширины перехода  $l$  и уменьшается с ростом обратного напряжения. Туннельный пробой начинается при таком значении  $U_{обр}$ , при котором уровень дна зоны проводимости  $n$ -области опускается ниже потолка валентной зоны  $p$ -полупроводника, т. е. появляется энергетическая полоса  $\Delta E_{тун}$ .

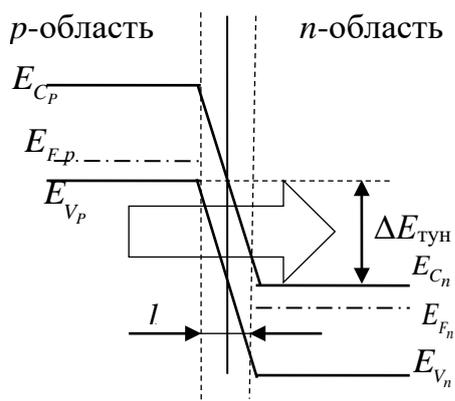


Рис. 6.21. Энергетическая диаграмма, поясняющая причину возникновения туннельного пробоя при обратном смещении  $p$ - $n$ -перехода. Стрелкой показано преимущественное направление перемещения электронов за счет туннельного эффекта

В отличие от лавинного пробоя, туннельный пробой  $p$ - $n$ -перехода происходит при малых обратных напряжениях (3–5 В), кроме того, соответствующая вольтамперная характеристика более пологая (кривая 2 на рис. 6.20).

При увеличении температуры  $p$ - $n$ -перехода напряжение туннельного пробоя уменьшается, т. е. обратный ток при данном напряжении возрастает. Это связано с тем, что при увеличении температуры происходит небольшое уменьшение ширины запрещенной зоны полупроводника. Это приводит к уменьшению ширины барьера  $\Delta x$  и увеличению туннельного тока. Следовательно, температурный коэффициент напряжения пробоя  $\alpha$  (6.46) в этом случае отрицателен. Эффект выражен слабо, поэтому величины изменения напряжения меньше, чем для лавинного пробоя.

### 6.5.3. Тепловой пробой $p$ - $n$ -перехода

Развитие теплового пробоя обусловлено выделением тепла вследствие прохождения через переход обратного тока. Если количество тепловой энергии, выделяемой в переходе в единицу времени, больше тепла, отводимого от него, то его температура повышается, концентрация неосновных носителей и обратный ток возрастают. Это приводит к дальнейшему увеличению температуры за счет выделения дополнительного джоулевого тепла; процесс нарастания температуры

перехода и обратного тока продолжается, и, в конечном счете, это может привести к разрушению перехода.

Рассмотрим особенности вольтамперной характеристики теплового пробоя (рис. 6.22).

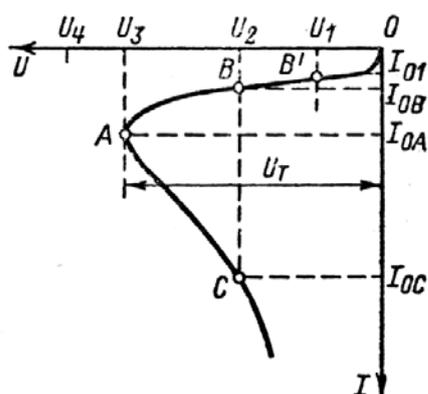


Рис. 6.22. Вольтамперная характеристика теплового пробоя  $p-n$ -перехода

Участок напряжений  $U_1 - U_3$  на вольтамперной характеристике теплового пробоя соответствует обратной ветви характеристики при постоянной температуре. При достижении напряжения теплового пробоя  $U_T$  температура перехода начинает повышаться, и концентрация подвижных носителей в нем будет возрастать по экспоненциальному закону (см. раздел 2.6). Электрическое сопротивление перехода снижается, что вызывает уменьшение падения напряжения на обратно смещенном переходе. Вследствие этого на вольтамперной характеристике обратного тока возникает участок  $AC$  с отрицательным дифференциальным сопротивлением. Если не снять обратное напряжение на переходе, состояние которого соответствует участку  $AC$ ; или не обеспечить его дополнительное охлаждение, то переход расплавится и перестанет существовать.

Напряжение теплового пробоя зависит от многих факторов: от величины теплового тока  $I_0$ , температуры, удельного сопротивления полупроводника, условий теплоотвода, и т. д. Тепловой пробой легко возникает в германиевых переходах из-за относительно

большой величины обратного (теплового) тока (10–100 мкА). В кремниевых переходах, а также в переходах на основе соединений  $A_{III}B_V$  ток  $I_0$  настолько мал, что тепловой пробой практически исключается.

Необходимо отметить, что если обратный ток существенно возрос при туннельном или лавинном пробое перехода, то после этого может наступить и тепловой пробой. Иначе говоря, это может произойти при превышении предельно допустимых значений обратных токов при электрических пробоях.

## 6.6. ЕМКОСТЬ $P-N$ -ПЕРЕХОДА

Наличие равных по величине и противоположных по знаку зарядов по обе стороны границы между  $p$ - и  $n$ -областями можно рассматривать как электрическую емкость  $p-n$ -перехода. Эти заряды могут иметь разную природу. Во-первых, это — ионы доноров и акцепторов, во-вторых, при прямом смещении перехода, инжектированные неосновные носители у границ  $p$ - и  $n$ -областей.

Однако полную аналогию с электростатической емкостью, определяемой как отношение заряда, накопленного на одном из двух проводников к возникшей между ними разности потенциалов, провести нельзя. На обычном конденсаторе при наличии разности потенциалов между обкладками накапливается электрическая энергия, тогда как в равновесном  $p-n$ -переходе этого не происходит. Влияние емкости  $p-n$ -перехода на свойства электрических схем, в которых используются активные полупроводниковые элементы, проявляется только при относительно быстрых изменениях напряжения на  $p-n$ -переходе. В зависимости от того, каково смещение перехода (прямое или обратное), емкостные свойства перехода определяются различными физическими процессами.

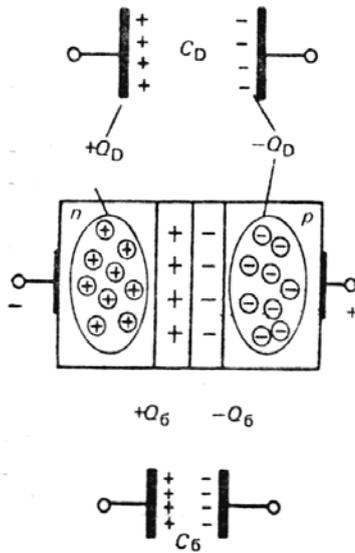


Рис. 6.23. Природа диффузионной и барьерной емкости  $p$ - $n$ -перехода.

$\pm Q_B$  — заряды ионов доноров и акцепторов,

$\pm Q_D$  — заряды инжектированных неосновных носителей.

Полярность подключения перехода соответствует проявлению диффузионной емкости

### 6.6.1. Барьерная емкость

При обратном смещении  $p$ - $n$ -перехода двойной электрический слой на границе между  $p$ - и  $n$ -областями образован положительными ионами доноров и отрицательными ионами акцепторов (рис. 6.23).

Емкость, связанная с перераспределением зарядов неподвижных ионов примесей, называется барьерной емкостью  $p$ - $n$ -перехода. Емкостная реакция перехода проявляется в том, что при изменении обратного напряжения изменяется ширина перехода и величина заряда некомпенсированных ионов примесей. При этом к постоянному обратному току добавляется переменная емкостная составляющая:

$i_{\sim} = C_{\bar{c}} \frac{dU}{dt}$ . Физическая причина возникновения этого тока – изменение объемных зарядов ионов примесей со временем, поэтому

$i_{\sim} = \frac{dQ}{dt} = \frac{dQ}{dU} \cdot \frac{dU}{dt} = C_{\bar{c}} \frac{dU}{dt}$ . Отсюда следует, что барьерная емкость не

определяется формулой электростатики, а является дифференциальной величиной:

$$C_{\bar{c}} = \frac{dQ}{dU}. \quad (6.47)$$

Для резкого одностороннего  $p-n$ -перехода можно получить простую формулу, отражающую зависимость барьерной емкости от обратного напряжения и концентрации примесей. Пусть для данного перехода  $N_A \gg N_D$ , тогда используя формулы (6.47) и (6.26) с учетом того, что  $U_{обр} \gg \phi_k$ , получим:

$$Q = eN_D S \cdot l_{об} = eN_D S \sqrt{\frac{2\varepsilon\varepsilon_0 U_{обр}}{eN_D}};$$

$$C_б = eN_D S \sqrt{\frac{2\varepsilon\varepsilon_0}{eN_D}} \cdot \frac{1}{2\sqrt{U_{обр}}} = \varepsilon\varepsilon_0 S \sqrt{\frac{eN_D}{2\varepsilon\varepsilon_0 U_{обр}}} = \frac{\varepsilon\varepsilon_0 S}{l_{об}}. \quad (6.48)$$

Видно, что барьерная емкость может быть представлена формулой для плоского конденсатора, где роль расстояния между обкладками играет ширина перехода. Соответственно, величина  $C_б$  растет с увеличением концентрации примесей из-за сужения перехода и уменьшается с ростом обратного напряжения, вызывающим противоположный эффект.

На основе барьерной емкости перехода изготавливаются специальные диоды — варикапы. Это диоды с электрически изменяемой емкостью. Как правило, для изготовления варикапов используются переходы с более сложным характером распределения примесей, что позволяет получить большую крутизну вольт-фарадной характеристики (кривая 3 рис. 6.24).

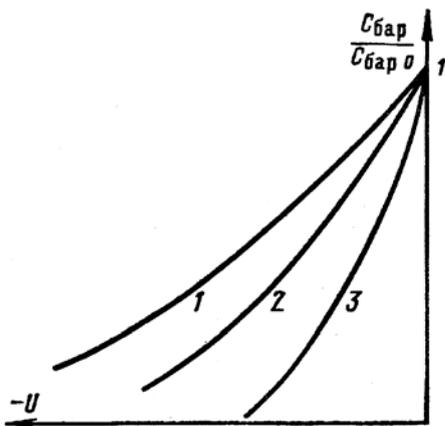


Рис. 6.24. Зависимость барьерной емкости  $p-n$ -перехода от обратного напряжения.

- 1 — для плавного перехода,
- 2 — для резкого перехода,
- 3 — для перехода с особым распределением примесей

Скорость изменения барьерной емкости при увеличении обратного напряжения  $U_{обр}$  зависит от характера распределения примесей в пределах ширины перехода. Как следует из (6.48) и рис. 6.24, для резкого одностороннего перехода  $C_{\bar{\sigma}} \sim \frac{1}{\sqrt{U_{обр}}}$ ; для плавного перехода, как показывают расчеты, эта зависимость более слабая:  $C_{\bar{\sigma}} \sim \frac{1}{\sqrt[3]{U_{обр}}}$ .

### 6.6.2. Диффузионная емкость

Прямое смещение перехода вызывает инжекцию неосновных носителей, и в прилегающих к переходу частях  $p$ - и  $n$ -областей возникает двойной электрический слой, обусловленный избыточными неосновными носителями: электронами  $p$ -области и дырками  $n$ -области.

Емкость, связанная с перераспределением зарядов инжектированных неосновных носителей при изменении прямого напряжения называется диффузионной емкостью  $p$ - $n$ -перехода (рис. 6.23). Изменение этих зарядов определяется зависимостью избыточной концентрации неосновных носителей от приложенного напряжения (см. формулы (6.32) и (6.32. а)), поэтому диффузионная емкость при невысоких частотах изменения напряжения пропорциональна прямому току:  $C_{диф} \sim I_{пр}$ . Вольт-фарадная характеристика для диффузионной емкости  $C_{диф} = f(U_{пр})$  имеет ту же форму, что и вольтамперная характеристика прямо смещенного перехода.

На высоких частотах, когда период переменного прямого напряжения становится меньше времени жизни неосновных носителей, заряд избыточных неосновных носителей в течение малого периода не успевает измениться. Следовательно, диффузионная емкость с ростом частоты будет уменьшаться практически до нуля. Поэтому, на очень высоких частотах, а также при резких «ступенчатых» изменениях напряжения на переходе ее влияние будет не заметно.

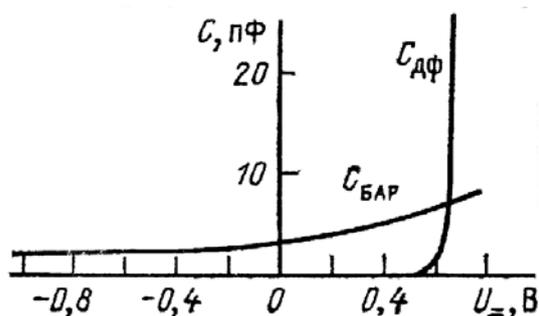


Рис. 6.25. Вольт-фарадные характеристики  $p-n$ -перехода при прямом и обратном смещении

Из вольт-фарадных характеристик барьерной и диффузионной емкости, изображенных в едином масштабе (рис. 6.25), следует, что барьерная емкость проявляется также и при прямом смещении; она играет определяющую роль до тех значений прямого напряжения, при которых начинается инжекция и, соответственно, резкий рост прямого тока. Например, для  $p-n$ -перехода на основе арсенида галлия преобладающее влияние барьерной емкости при прямом смещении наблюдается вплоть до напряжений 0,8–0,9 В.

## 6.7. ПЕРЕХОДНЫЕ ПРОЦЕССЫ В $P-N$ -ПЕРЕХОДЕ

### 6.7.1. Установление прямого напряжения при подаче на переход скачка прямого тока

Если на переход подается не постоянное, а быстро меняющееся переменное напряжение, то при увеличении его частоты выпрямляющие свойства перехода ухудшаются: величина обратного тока в течение части отрицательного полупериода становится соизмеримой с током во время положительного полупериода. Наиболее ярко инерционность  $p-n$ -перехода проявляется при подаче на него ступенчатых (импульсных) сигналов, переключающих переход скачком с прямого смещения на обратное и наоборот.

Переходными называются процессы установления постоянных токов и напряжений на  $p-n$ -переходе при подаче на него ступенчато меняющихся сигналов. Характер этих процессов зависит от амплитуды импульсного сигнала, т. е. от уровня инжекции, а также от свойств

источника сигнала. Если внутреннее сопротивление источника значительно больше сопротивления перехода, то источник может поддерживать постоянный ток при любых изменениях параметров внешней цепи. Такой тип источника называется генератором тока. При обратном соотношении указанных сопротивлений источник при тех же условиях может поддерживать постоянным напряжение на выходе и поэтому называется источником напряжения.

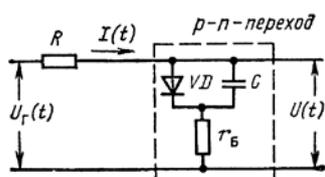


Рис. 6.26. Эквивалентная схема  $p-n$ -перехода

Для рассмотрения переходных процессов используем эквивалентную схему на рис. 6.26. Напряжение от генератора импульсов  $U_T(t)$  через резистор  $R$  подается на цепь, состоящую из «идеального диода», емкости перехода  $C$ , включенной параллельно диоду, и последовательно включенного сопротивления базы  $r_B$ . Из множества вариантов выберем два случая: установление прямого напряжения на переходе и установление постоянного обратного тока перехода.

Пусть в некоторый момент  $t^{(1)}$  ток перехода скачком меняется от нуля до  $I_{пр}$  (рис. 6.27, а). Влияние диффузионной емкости на продолжительность переходного процесса будет пренебрежимо малым (см. раздел 6.3). При низком уровне инжекции ( $\delta \ll 1$ ) влияние падения напряжения на объемном сопротивлении базы, т. е. низколегированной области, можно не учитывать, поэтому наблюдается плавное нарастание прямого напряжения с последующим выходом на постоянное значение  $U_{пр}$  (рис. 6.27, б). Запоздывание прямого напряжения на переходе связано с конечностью времени установления новой, меньшей ширины перехода, т. е. со временем перезарядки барьерной емкости.

При большом прямом токе  $\delta \gg 1$ , и характер переходного процесса меняется (рис. 6.27, в). Во-первых, из-за падения напряжения на объемном сопротивлении базы наблюдается мгновенный скачок прямого напряжения  $\Delta U_1$ . Далее процесс идет так же, как и при

низком уровне инжекции: прямое напряжение нарастает и достигает наибольшего значения  $U_{\text{макс}}$ . Скорость нарастания напряжения на переходе гораздо выше, чем в случае низкого уровня инжекции, так как при большой величине броска прямого тока время перезарядки барьерной емкости становится очень малым. Например, при скачке тока  $\Delta I = I_{\text{пр}} = 10$  мА и емкости 5 пФ время перезарядки  $C_b \Delta U_{\text{пер}} / I_{\text{пр}}$  окажется равным 0,4 нс. Здесь  $\Delta U_{\text{пер}}$  соответствует прямому напряжению на участке резкого роста прямого тока; для кремниевого перехода эта величина составляет  $\sim 0,8$  В.

На последней стадии переходного процесса прямое напряжение на переходе снижается до стационарного значения  $U_{\text{пр}}$ . Это связано с тем, что из-за высокого уровня инжекции концентрация подвижных носителей в базе резко возрастает, что приводит к уменьшению ее объемного сопротивления. Этот эффект называется модуляцией сопротивления базы. Следствием этого явления является уменьшение падения напряжения на сопротивлении базы и установление на переходе прямого напряжения  $U_{\text{пр}} < U_{\text{макс}}$ .

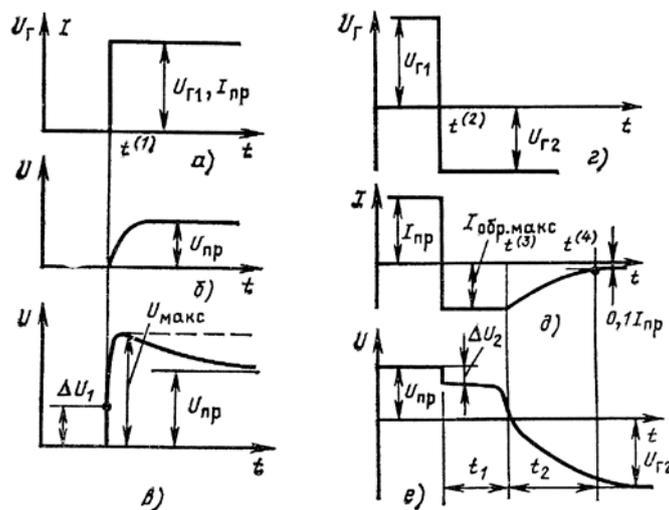


Рис. 6.27. Процесс установления прямого напряжения (а–в) и восстановления стационарного обратного тока и напряжения (г–е) на  $p-n$ -переходе

Поскольку диффузионная ёмкость не проявляется при установлении прямого напряжения на переходе, инерционность переключения определяется только временем перезарядки барьерной емкости. Вследствие этого время этого переходного процесса при низком и высоком уровнях инжекции оказывается очень малым, порядка единиц–десятков наносекунд.

### ***6.7.2. Установление постоянного обратного тока перехода при ступенчатом переключении напряжения с прямого на обратное***

Пусть в момент времени  $t^{(2)}$  выходное напряжение генератора импульсов скачком меняется с прямого  $U_{Г1}$  на обратное  $U_{Г2}$  (рис. 6.27, *з*). Поскольку в данном случае источник является генератором напряжения, то при постоянном уровне  $U_{Г2}$  протекает переходный процесс: обратный ток и связанное с ним падение напряжение на переходе с течением времени приходят к своему стационарному значению.

Скачок напряжения вызывает резкую смену полярности тока через переход, величина броска тока  $\Delta I = (U_{Г1} + U_{Г2})/R$ . Через переход начинает протекать обратный ток  $I_{обр. макс}$ , намного превышающий величину обратного тока при постоянном обратном смещении (рис. 6.27, *д*). Этот большой ток обусловлен тем, что в момент переключения в базе очень высока концентрация избыточных неосновных носителей, накопленных при прямом смещении, и при смене полярности напряжения происходит их экстракция через *p-n*-переход. В момент переключения падение напряжения на переходе скачком уменьшается на величину  $\Delta U_2 = \Delta I \cdot r_b$ , но он остается смещенным в прямом направлении. Причина такого необычного явления в том, что на переходе по обе стороны от границы имеется двойной электрический слой избыточных неосновных носителей, и распределение их

концентраций соответствует прямому смещению перехода (см. кривую для  $t \leq t^{(2)}$  на рис. 6.28).

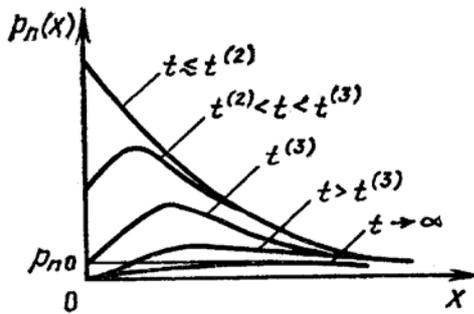


Рис. 6.28. Изменение распределения концентрации неосновных носителей в базе со временем после переключения

Время переходного процесса принято разделять на два этапа:  $t_1$  — участок высокой обратной проводимости и  $t_2$  — стадию восстановления постоянного значения обратного тока (рис. 6.27, e). В течение всего промежутка времени  $t_1$  концентрация неосновных носителей в базе превышает равновесную (рис. 6.28). Такое состояние является причиной того, что величина обратного тока соизмерима с величиной прямого тока, при этом обратный ток почти постоянен. Это явление обусловлено диффузионным механизмом движения накопленных неосновных носителей к переходу. Видно, что градиент их концентрации (наклон кривых вблизи перехода на рис. 6.28) приблизительно одинаков на всем протяжении отрезка времени  $t_1$ . Это обеспечивает постоянную скорость поступления неосновных носителей базы на границу перехода и, соответственно, постоянство обратного тока за счет экстракции на стадии высокой обратной проводимости.

Данную стадию можно рассматривать как разряд диффузионной емкости перехода, происходящий из-за рассасывания накопленных неосновных носителей. На рис. 6.27, e, видно, что при этом прямое напряжение на переходе, которое можно считать напряжением на диффузионной емкости, постепенно уменьшается.

Как видно из рис. 6.28, в момент времени  $t^{(3)}$  избыточные неосновные носители на границе базовой области с  $p$ - $n$ -переходом исче-

зают ( $\Delta p_n(0) = 0$ ), при этом падение напряжения на переходе становится равным нулю (рис. 6.27, *e*). Далее, на всем протяжении участка  $t_2$  продолжается экстракция оставшихся у границ перехода неосновных носителей, т. е. заканчивается разряд диффузионной емкости. Одновременно происходит перестройка обедненного слоя к состоянию обратного смещения, что соответствует перезарядке барьерной емкости перехода. При этом напряжение на переходе и обратный ток стремятся к стационарным значениям  $U_{Г2}$  и  $I_{обр}$ .

Основной параметр, характеризующий инерционность перехода при ступенчатом изменении напряжения с прямого на обратное — время восстановления обратного сопротивления. Наименьшие значения времени восстановления составляют единицы наносекунд. Оно отсчитывается от момента переключения до момента, когда обратный ток, уменьшаясь, достигнет значения  $0,1 I_{пр}$ . Этот промежуток времени равен  $t_1 + t_2$  (рис. 6.27, *d, e*).

Время восстановления не является универсальным параметром. Его значение зависит от характеристик цепи, в которую включается диод. На стадии высокой обратной проводимости сопротивление  $p-n$ -перехода мало, и ток определяется практически полностью номиналом резистора  $R$ , включенного последовательно с переходом:  $I_{обр. макс} \approx U_{Г2}/R$  (см. рис. 6.26). При уменьшении величины  $R$  обратный ток возрастает, и время, необходимое для рассасывания накопленного заряда избыточных неосновных носителей, также уменьшается.

Для увеличения диапазона частот, в котором могут работать диоды на основе  $p-n$ -переходов, необходимо обеспечивать как можно меньшее время восстановления. Для этого используют переходы малой площади, низкий уровень инжекции, а также диоды Шоттки (диоды на основе выпрямляющего контакта металл-полупроводник, см. раздел 6.8.1).

### Вопросы к разделам 6.6 – 6.7

1. Каков физический механизм лавинного пробоя  $p-n$ -перехода? При каких условиях наблюдается это явление?
2. От каких факторов зависит напряжение лавинного пробоя? Каковы физические причины этих зависимостей?
3. Каков физический механизм туннельного пробоя  $p-n$ -перехода? При каких условиях наблюдается это явление?
4. Как и по какой причине изменяется напряжение туннельного пробоя при увеличении температуры?
5. Опишите физический механизм развития теплового пробоя  $p-n$ -перехода.
6. По каким причинам  $p-n$ -переход обладает емкостными свойствами?
7. В чем отличие емкости  $p-n$ -перехода от обычной электростатической емкости?
8. Дайте определение термина «барьерная емкость»  $p-n$ -перехода. Как и по какой причине, ее величина изменяется при возрастании обратного напряжения?
9. Дайте определение термина «диффузионная емкость»  $p-n$ -перехода. Как и по какой причине, ее величина изменяется при возрастании прямого напряжения?
10. Поясните разницу в переходных процессах установления прямого напряжения на переходе при низком и высоком уровне инжекции.
11. Какова физическая причина большого обратного тока перехода в момент ступенчатого переключения напряжения с прямого на обратное?
12. Что такое «время восстановления обратного сопротивления перехода»? Каковы способы его уменьшения?

## 6.8. ПЕРЕХОДЫ МЕТАЛЛ – ПОЛУПРОВОДНИК

### 6.8.1. Выпрямляющий контакт (барьер Шоттки)

Контакты металл-полупроводник широко применяются в полупроводниковой электронике. Во-первых, они обеспечивают соединение полупроводникового прибора с элементами внешней цепи. Такие контакты должны обладать линейной вольтамперной характеристикой. Во-вторых, на их основе созданы быстродействующие диоды, обладающие очень малой инерционностью и работающие на высоких частотах. Существуют две разновидности таких контактов:

- выпрямляющий контакт с «односторонней» проводимостью;
- омический контакт; ток и падение напряжения на нем приблизительно соответствуют закону Ома.

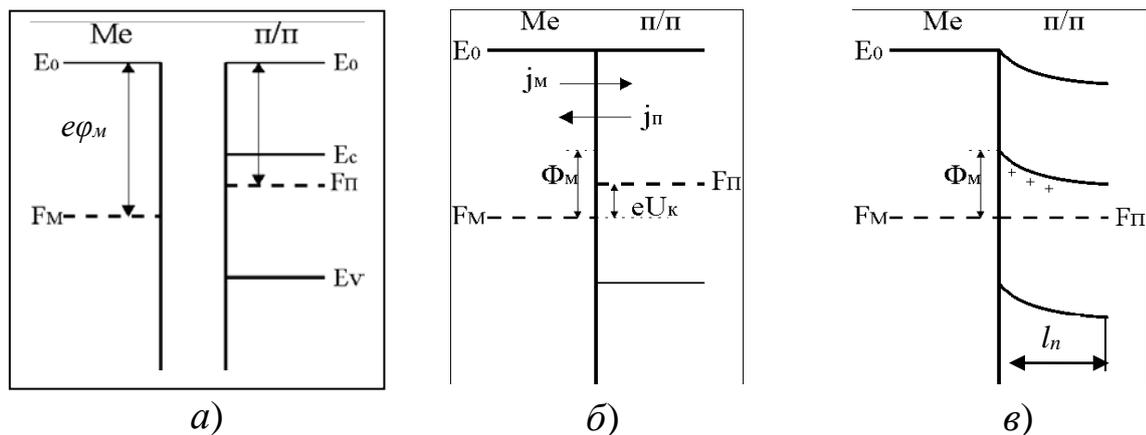


Рис. 6.29. Энергетические диаграммы металла и  $n$ -полупроводника:  
 а) до контакта, б) непосредственно после соприкосновения,  
 в) после установления равновесия на контакте.

$e\phi_M$  — работа выхода металла,  $e\phi_{\Pi}$  — работа выхода полупроводника,  
 $eU_K$  — контактная разность потенциалов, т. е. потенциальный барьер  
 для электронов полупроводника,  $\Phi_M$  — потенциальный барьер  
 для электронов металла

Вид контакта определяется следующими факторами: типом проводимости полупроводника; соотношением работ выхода металла и

полупроводника; концентрацией примеси в полупроводнике; знаком и величиной поверхностного заряда.

Осуществим идеальный контакт между металлом и полупроводником  $n$ -типа, работа выхода которого меньше работы выхода электронов из металла:  $e\varphi_{\text{м}} > e\varphi_{\text{п}}$  (рис. 6.29, *а, б*). При этом предполагается, что заряд на поверхности полупроводника отсутствует:  $Q_{\text{пов}} = 0$ . При таком контакте энергетически выгодным оказывается переход электронов с уровней зоны проводимости полупроводника в металл (ток  $j_{\text{п}}$ ). Они занимают свободные энергетические уровни в металле выше уровня Ферми и сообщают ему отрицательный заряд. Из-за ухода электронов в металл, в приконтактной области полупроводника образуется слой, обедненный основными носителями шириной  $l_n$ , в котором остается некомпенсированный положительный заряд ионов доноров.

Образующееся на контакте электрическое поле препятствует дальнейшему переходу электронов из  $n$ -полупроводника в металл и одновременно ускоряет электроны, движущиеся в противоположном направлении (ток  $j_{\text{м}}$ ). В результате на переходе устанавливается равновесие, при котором встречные токи электронов  $j_{\text{м}}$  и  $j_{\text{п}}$  уравниваются и суммарный ток равен нулю. При равновесии в системе образуется единый уровень Ферми (рис. 6.29, *в*). Обедненный слой  $l_n$  лежит целиком в приграничной области полупроводника, его удельное сопротивление намного больше, чем в объеме.

При подключении прямого внешнего напряжения  $U$  («плюс» на металл, «минус» на полупроводник) высота потенциального барьера для электронов зоны проводимости полупроводника снижается, ширина области обеднения уменьшается ( $L_{\text{п}} < l_n$ ), и через переход течет ток основных носителей — электронов проводимости полупроводника — в металл (рис. 6.30, *а*). Важнейшим преимуществом барьера Шоттки перед  $p$ - $n$ -переходом является то, что при прямом включении

в нем не происходит накопление избыточных неосновных носителей — электронов за счет инжекции, так как «базой» в этом случае является металл, в котором концентрация электронов на несколько порядков выше, чем в полупроводнике.

При подключении обратного напряжения  $U_{обр}$  («плюс» на полупроводник, «минус» на металл) область обеднения расширяется ( $L_3 > l_n$ ), потенциальный барьер для основных носителей повышается пропорционально величине  $U_{обр}$ , и прямой ток электронов уменьшается практически до нуля.

Малый обратный ток  $I_0$  противоположного направления образуется электронами, переходящими из металла в полупроводник. Их энергия должна быть достаточной для преодоления потенциального барьера для электронов, движущихся в этом направлении (рис. 6.30, б). Этот потенциальный барьер  $\Phi_M$  равен разности энергий,

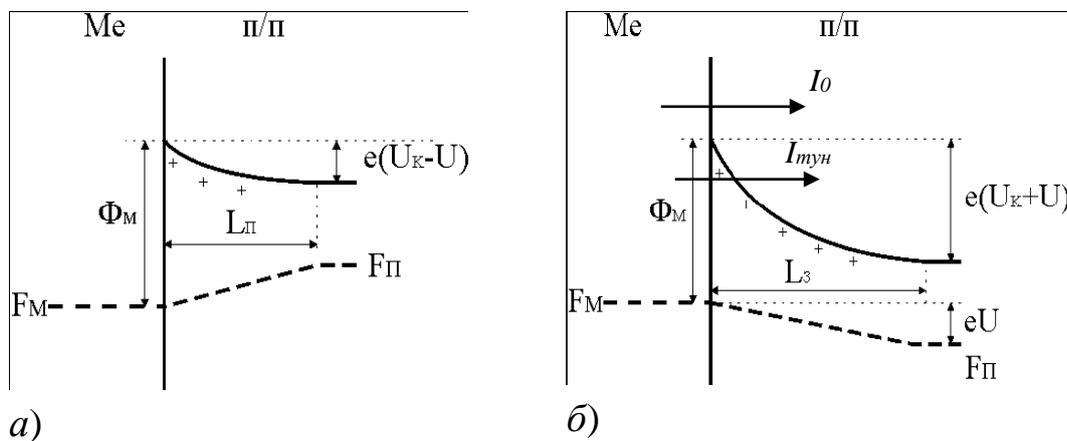


Рис. 6.30. Энергетические диаграммы барьера Шоттки при прямом (а) и обратном (б) включении

соответствующих дну зоны проводимости полупроводника на границе раздела и уровню Ферми в металле (рис. 6.30, а, б); его высота в первом приближении не зависит от величины и полярности приложенного к переходу напряжения. Величина этого «надбарьерного»

тока  $I_0$  выражается (подобно току термоэлектронной эмиссии) следующей формулой:

$$I_0 = S \cdot A \cdot T^2 e^{-\frac{e\phi_{МП}}{kT}}, \quad (6.49)$$

где  $S$  — площадь перехода,  $A$  — размерная константа, ( $\sim 120 \text{ А}/(\text{см}^2 \cdot \text{К})$ ). Этот ток на несколько порядков меньше тока при прямом включении.

Сравним величины обратного тока для барьера Шоттки (металл–полупроводник  $n$ -типа) и  $p$ - $n$ -перехода, изготовленных на основе кремния. Пусть площадь каждого из переходов равна  $10^{-4} \text{ см}^2$ , концентрация донорной примеси  $N_{\text{д}} = 10^{15} \text{ см}^{-3}$ ,  $T = 300 \text{ К}$ . Тогда обратный ток для барьера Шоттки будет составлять  $\sim 10^{-9} \text{ А}$ , а та же величина для  $p$ - $n$ -перехода будет равна  $10^{-13} \text{ А}$ . Видно, что обратный ток выпрямляющего контакта металл-полупроводник на много порядков превышает тепловой ток  $p$ - $n$ -перехода.

Формула, описывающая вольтамперную характеристику барьера Шоттки, а также вид его вольтамперной характеристики совпадают с данными для  $p$ - $n$ -перехода (см. формулу (6.35) и рис. 6.9). Однако их сравнение при указанных выше условиях показывает также, что помимо большего обратного тока наблюдается различие и при прямом смещении (рис. 6.31).

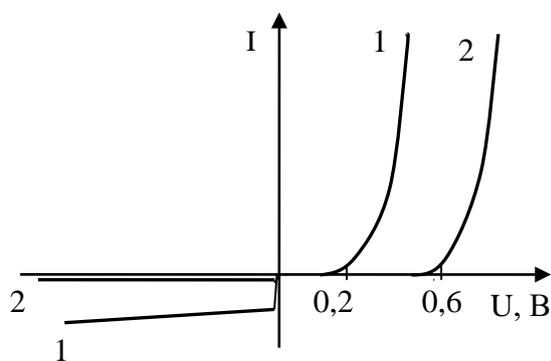


Рис. 6.31. Вольтамперные характеристики барьера Шоттки алюминий– $n$ -кремний (1) и кремниевого  $p$ - $n$ -перехода (2) при одинаковых площадях контактов и концентрациях донорной примеси

Прямая ветвь вольтамперной характеристики барьера Шоттки располагается в области меньших прямых напряжений. Это обусловлено большей по сравнению с  $p$ - $n$ -переходом величиной его обратного тока  $I_0$ , а также меньшей величиной потенциального барьера для электронов полупроводника  $n$ -типа на границе с металлом.

Кроме того, поскольку прямой ток выпрямляющего контакта — это ток электронов из полупроводника  $n$ -типа в металл, то при больших прямых токах можно пренебречь падением напряжения на малом сопротивлении металла. Вследствие этого вольтамперная характеристика барьера Шоттки в отличие от  $p$ - $n$ -перехода описывается экспонентой в более широком интервале прямых напряжений. Аналогичная картина «односторонней проводимости» получается и при контакте металла с полупроводником  $p$ -типа, у которого работа выхода больше, чем у металла.

Приведенное выше описание физических процессов на контакте при подаче обратного напряжения справедливо для концентраций примесей в полупроводнике не выше  $10^{17}$  см<sup>-3</sup>. При этом потенциальный барьер для электронов, движущихся из металла в полупроводник, достаточно широк, и можно пренебречь туннельным переносом электронов в том же направлении (на рис. 6.30, б, ток  $I_{\text{тун}}$  пренебрежимо мал).

Предположение об отсутствии поверхностного заряда в действительности не выполняется. В полупроводнике  $n$ -типа всегда существует отрицательный поверхностный заряд; его плотность зависит от вещества, а также от состояния поверхности. Так, для GaAs  $n$ -типа среднее значение плотности поверхностного заряда, отнесенное к заряду электрона, составляет  $10^{13}$  см<sup>-2</sup>, для кремния  $n$ -типа эта величина равна  $10^{10}$ – $10^{11}$  см<sup>-2</sup>. Отрицательный поверхностный заряд вытесняет электроны из приграничной области вглубь полупроводника. В результате, у его поверхности даже в отсутствие контакта с металлом образуется область обеднения, и энергетические зоны у по-

верхности изгибаются вверх. В этом случае, особенно при высокой концентрации отрицательного поверхностного заряда, может оказаться, что тип контакта не будет зависеть от свойств металла, так как его роль как приемника электронов из полупроводника при установлении равновесия не будет определяющим фактором.

Таким образом, при наличии в полупроводнике  $n$ -типа высокой плотности отрицательного поверхностного заряда независимо от соотношения работ выхода металла и полупроводника при их контакте может образоваться выпрямляющий контакт — барьер Шоттки.

### **6.8.2. Омический контакт металл – полупроводник**

Существуют две разновидности контакта, вольтамперная характеристика которых, близка к закону Ома.

Во-первых, в случае контакта металла с  $n$ -полупроводником, когда работа выхода металла меньше, чем работа выхода полупроводника, энергетически выгодным оказывается переход электронов металла на свободные энергетические уровни зоны проводимости полупроводника. При этом в полупроводнике вблизи границы с металлом образуется слой с повышенной концентрацией основных носителей. Такой контакт металла с полупроводником называется омическим, так как он обладает двухсторонней, примерно одинаковой проводимостью.

Во-вторых, барьер Шоттки при концентрации примеси в полупроводнике, превышающей  $10^{17} \text{ см}^{-3}$  превращается также в омический контакт. При таких концентрациях примеси ширина области обеднения на границе с металлом значительно уменьшается, и при подключении обратного напряжения образуется острый и узкий потенциальный барьер для электронов металла. Это приводит к тому, что величина туннельного тока электронов  $I_{\text{тун}}$  из металла в полупроводник при обратном смещении (см. рис. 6.30, б) становится такой же, как и величина тока при прямом смещении.

## 6.9. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называется переход, образованный при контакте двух различных по химическому составу полупроводников с неодинаковой шириной запрещенной зоны. Для стабильности гетероперехода необходимо, чтобы у этих материалов был один и тот же тип кристаллической решетки. При этом значения постоянной решетки и температурного коэффициента расширения должны быть практически одинаковыми. При несоблюдении этих условий на контакте возникнут большие механические напряжения, которые приведут к его разрушению. Очевидно, что количество материалов для гетеропереходов ограничено. Наиболее распространенными из них являются германий и арсенид галлия (переход Ge–GaAs), арсенид галлия и трехкомпонентный твердый раствор GaAlAs (переход GaAs–GaAlAs), фосфид индия и трехкомпонентный раствор InGaAs (переход InP–In<sub>0,53</sub> Ga<sub>0,47</sub> As). Такие полупроводники образуют так называемые изопериодические системы. Сюда же можно отнести селениды ртути и кадмия, а также систему GaSb–GaAlSb с изменяемой долей алюминия.

На основе гетеропереходов изготавливаются высокоэффективные полупроводниковые излучатели и фотоприемники (фотодиоды, светодиоды и лазеры). Гетероструктуры с последовательным чередованием слоев разных полупроводников очень малой толщины, не более 10–15 нм, (сверхрешетки) находят еще более широкое применение в полупроводниковой наноэлектронике.

Гетеропереходы, образованные полупроводниками с одинаковым типом проводимости ( $p^+$ - $p$  или  $n^+$ - $n$ ), называются изотипными. При использовании полупроводников с разным типом проводимости формируются анизотипные гетеропереходы.

## Выпрямляющие гетеропереходы

Рассмотрим анизотипный гетеропереход, у которого ширина запрещенной зоны  $n$ -полупроводника больше, чем у  $p$ -полупроводника, и энергетические зоны полностью перекрываются. Такой переход образуют, например,  $n$ -GaAs и  $p$ -Ge (рис. 6.32).

При идеальном контакте полупроводников с различной шириной запрещенной зоны  $\Delta E_z$  происходит установление равновесного состояния с единым уровнем Ферми. Физические процессы, протекающие при этом те же, что и для случая обычного  $p$ - $n$ -перехода. Однако из-за разницы в величинах  $\Delta E_z$  на линии раздела  $p$ - и  $n$ -областей возникают разрывы границ энергетических зон. В случае контакта  $p$ -Ge –  $n$ -GaAs (рис. 6.32, а) при разрыве границы дна зоны проводимости образуется маленький узкий пик («пичок»), высота которого равна разности энергий  $E_{C \text{ GaAs}} - E_{C \text{ Ge}}$ ; при разрыве границы валентной зоны образуется дополнительная ступенька. Если в контакт приводятся те же полупроводники с противоположным соотношением типа проводимости, то характер разрыва зон на границе изменяется на противоположный.

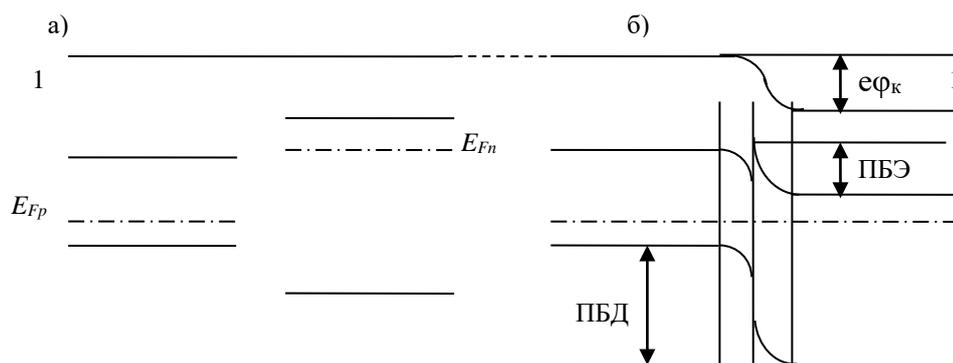


Рис. 6.32. Выпрямляющий анизотипный  $p$ - $n$  гетеропереход.

Линия 1 – уровень вакуума. а) взаимное расположение энергетических зон до образования контакта, б) равновесная энергетическая диаграмма

Появление разрывов границ зон приводит к тому, что потенциальные барьеры для диффузии основных носителей в смежные области оказываются не одинаковыми. Для гетероперехода, у которого ширина запрещенной зоны  $n$ -полупроводника больше, чем у  $p$ -полупроводника (рис. 6.32, б), потенциальный барьер для диффузии электронов  $n$ -GaAs в  $p$ -Ge (ПБЭ) гораздо меньше, чем барьер (ПБД) для диффузии дырок в обратном направлении. Поэтому при приложении прямого напряжения будет преобладать инжекция электронов, даже если  $p$ -область имеет одинаковую или даже большую концентрацию примесей. В случае гетероперехода  $n$ -Ge –  $p$ -GaAs будет преобладать инжекция дырок в  $n$ -Ge.

Таким образом, в отличие от обычного  $p$ – $n$ -перехода, гетеропереход всегда обладает важным свойством: односторонней инжекцией из широкозонного полупроводника в узкозонный, практически не зависящей от уровня легирования  $p$ - и  $n$ -областей. Именно это свойство используется для создания эффективных источников когерентного (лазеры) и некогерентного (светодиоды) излучения. Расчеты показывают, что для обеспечения односторонней инжекции достаточно разницы в ширине запрещенной зоны  $\sim 0,1$  эВ.

Вольтамперная характеристика идеального гетероперехода описывается той же формулой (6.35), что и для обычного  $p$ – $n$ -перехода. В реальных гетеропереходах вблизи границы раздела сосредоточены дефекты решетки различной природы, которые являются центрами генерации или рекомбинации. Вследствие этого обратный ток реального анизотипного гетероперехода, в основном, является током генерации, а доля тока рекомбинации в прямом токе оказывается значительно большей, чем в обычном переходе.

Изотипные гетеропереходы в отличие от таких переходов между полупроводниками одной природы (см. раздел 6.2) могут быть как «омическими», так и выпрямляющими.

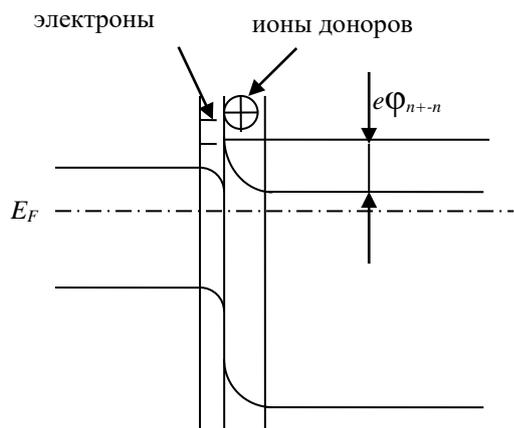


Рис. 6.33. Энергетическая диаграмма выпрямляющего изотипного  $n-n^+$  гетероперехода

На энергетической диаграмме равновесного  $n-n^+$  Ge-GaAs выпрямляющего гетероперехода (рис...)

видно, что на границе между полупроводниками образуется точно такой же энергетический пик, как и у  $p-n$  гетероперехода (рис. 6.32). Однако двойной электрический слой на границе раздела имеет иную природу: со стороны области  $n^+$  находятся положительные ионы доноров, а с другой стороны – узкая область, обогащенная электронами.

Прямое смещение такого перехода понижает потенциальный барьер для электронов зоны проводимости широкозонного  $n^+$ -полупроводника; он уменьшается до величины  $e\phi_{n+-n} - eU_{пр}$ . Прямой ток обусловлен интенсивным переносом электронов проводимости  $n^+$ -полупроводника над пониженным потенциальным барьером в узкозонный полупроводник  $n$ -типа.

При обратном смещении этот потенциальный барьер возрастает до значения  $e\phi_{n+-n} + eU_{обр}$ , и ток электронов в  $n$ -полупроводник прекращается. На пути встречного (обратного) тока электронов проводимости  $n$ -области в  $n^+$ -полупроводник находится потенциальный барьер  $e\phi_{n-n+}$ ; его высота практически не изменилась по сравнению с равновесным состоянием. При малых обратных напряжениях небольшой обратный ток обусловлен надбарьерным переносом электронов проводимости в  $n^+$ -полупроводник.

При дальнейшем увеличении  $U_{обр}$  ширина энергетического пика, разделяющего зоны проводимости, становится все меньше. Об-

ратное напряжение величиной в несколько Вольт оказывается достаточным для того, чтобы электроны зоны проводимости  $n$ -полупроводника, обладающие энергиями ниже  $e\varphi_{n-n^+}$ , начали переходить в смежную  $n^+$ -область за счет туннельного эффекта, давая дополнительный вклад в обратный ток. Еще большее возрастание обратного напряжения может вызвать большой обратный туннельный ток электронов из валентной зоны  $n$ -полупроводника на свободные уровни зоны проводимости  $n^+$ -области – так же, как и при туннельном пробое обычного  $p$ – $n$ -перехода.

Главным преимуществом выпрямляющих изотипных переходов по сравнению с анизотипными является малая инерционность, что позволяет использовать их в качестве детекторов и смесителей в области сверхвысоких частот. Это объясняется тем, что они «работают» на основных носителях. Избыточные концентрации неосновных носителей в смежных областях при этом не образуются, и поэтому не требуется время для их рассасывания, которое является основным фактором, ограничивающим частотный диапазон применимости приборов, основанных на  $p$ – $n$ -переходах.

Прямая ветвь вольтамперной характеристики выпрямляющих гетеропереходов практически такая же, как и у обычных однородных  $p$ – $n$ -переходов. Рабочие обратные напряжения оказываются небольшими (меньше 10 В) из-за возникновения туннельных токов, что можно считать разновидностью туннельного пробоя.

### **Невыпрямляющие («омические») гетеропереходы**

Такие переходы могут образоваться в тех случаях, когда энергетические зоны составляющих полупроводников перекрываются не полностью, как в случае выпрямляющих гетеропереходов, а лишь частично. Наибольший интерес представляет невыпрямляющий изотипный  $n$ - $n^+$  гетеропереход, в котором ток переносится электронами.

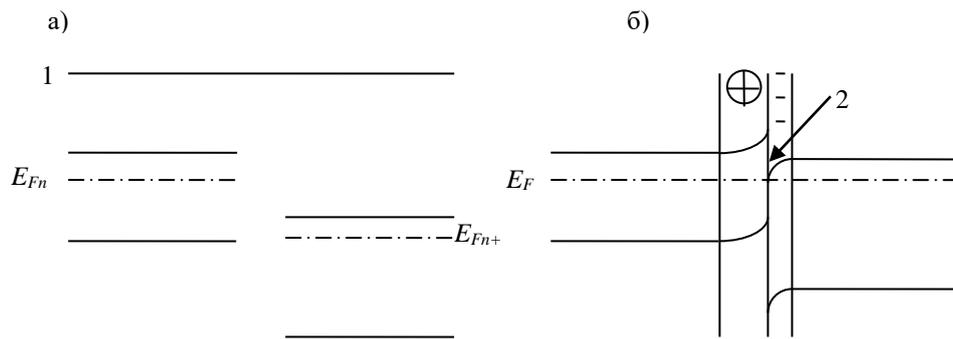


Рис. 6.34. Невыпрямляющий изотипный  $n-n^+$  гетеропереход.

- а) взаимное расположение энергетических зон до образования контакта, 1 – уровень вакуума, б) равновесная энергетическая диаграмма, 2 – область, обогащенная электронами

В этом случае при образовании контакта в  $n^+$ -области у границы с  $n$ -полупроводником образуется потенциальная яма (2), в пределах которой концентрация электронов значительно выше, чем в объеме. Когда  $n^+$ -область положительна по отношению к  $n$ -области, потенциальный барьер для электронов проводимости  $n$ -полупроводника понижается ( $e\phi_{n-n^+} - eU$ ), и через переход течет ток электронов в направлении  $n^+$ -области.

При обратной полярности приложенного напряжения ток течет в противоположном направлении. Его величина практически такая же за счет очень большой концентрации электронов у границы раздела в области обогащения. По мере возрастания напряжения к току электронов с энергиями выше потенциального барьера добавляется возрастающий туннельный ток из области обогащения (2) через все более узкий разделительный пикок.

### **Вопросы к разделам 6.8 – 6.9**

1. Какими факторами определяются разновидности контакта металл-полупроводник?
2. Опишите механизм установления равновесия при образовании выпрямляющего контакта металл-полупроводник.
3. Какова физическая природа прямого и обратного тока барьера Шоттки на основе контакта металла с полупроводником  $n$ -типа?
4. При каких условиях выпрямляющий контакт металл-полупроводник может стать омическим?
5. По какой причине диоды на основе выпрямляющего контакта металл-полупроводник (диоды Шоттки) обладают высоким быстродействием (малой инерционностью)?
6. Каким образом проявляется влияние поверхностного заряда на свойства контакта металл-полупроводник?
7. Каковы разновидности омического контакта металл-полупроводник, и какие факторы обеспечивают в каждом случае его двустороннюю проводимость?
8. Дайте определение термина «гетеропереход». Перечислите его разновидности и необходимые условия, при которых этот переход может существовать.
9. Какова причина наличия «скачков» потенциала на энергетической диаграмме гетероперехода?
10. Что такое «односторонняя инжекция» на гетеропереходе и каковы физические причины ее возникновения?

## **7. ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ И СВЕТОИЗЛУЧЕНИЕ**

### **7.1. ВЗАИМОДЕЙСТВИЕ ОПТИЧЕСКОГО ИЗЛУЧЕНИЯ С ПОЛУПРОВОДНИКАМИ. ФОТОПРОВОДИМОСТЬ**

Характер взаимодействия света с полупроводниками может быть различным. Он определяется свойствами материала полупроводника и длиной волны оптического излучения, т. е. энергией фотонов. В разделе 4.4 была рассмотрена генерация электронно-дырочных пар при облучении полупроводника светом (биполярная световая генерация). Это явление находит очень широкое применение в полупроводниковой электронике. На его основе изготавливаются различные фотоэлектрические приемники: фоторезисторы, фотодиоды, фототранзисторы и фототиристоры.

При облучении полупроводников светом с энергией квантов, меньшей, чем ширина запрещенной зоны (как правило, в инфракрасной части спектра), генерация электронно-дырочных пар не происходит. Концентрация подвижных носителей в этом случае увеличивается только за счет однополярной генерации подвижных электронов или дырок при ионизации примесных атомов. Такой механизм фотопроводимости может использоваться в фотоприемниках инфракрасного излучения, но при низких рабочих температурах ( $\sim 80$  К), при которых мала скорость тепловой генерации.

При таких же температурах работают болометры, датчики теплового излучения. Однако, физический механизм фотопроводимости в этом случае принципиально иной: кванты света поглощаются подвижными носителями заряда, например, электронами, и они занимают более высоко лежащие свободные энергетические уровни в зоне проводимости. В результате увеличивается подвижность электронов  $\mu$  (см. раздел 3.2), и ток через полупроводник возрастает.

Существуют также разновидности поглощения квантов света полупроводниками, не приводящие к изменению их проводимости, например, решеточное поглощение, когда энергия света передается на возбуждение колебаний кристаллической решетки, или экситонное поглощение, когда энергия кванта расходуется на образование связанной электрически нейтральной пары электрон-дырка.

Рассмотрим более подробно физические процессы, при которых фотопроводимость возникает за счет биполярной световой генерации.

## 7.2. ФОТОРЕЗИСТИВНЫЙ ЭФФЕКТ. ФОТОРЕЗИСТОР

Фоторезистивным эффектом называется изменение сопротивления полупроводника под действием света. В отсутствие освещения полупроводник обладает темновой проводимостью, обусловленной равновесными концентрациями подвижных носителей:

$$\sigma_0 = e(n_0\mu_n + p_0\mu_p). \quad (7.1)$$

При облучении полупроводника светом с энергией квантов, превышающей ширину запрещенной зоны, появляются неравновесные электронно-дырочные пары. Концентрация подвижных носителей изменяется до величин  $n_0 + \Delta n$  и  $p_0 + \Delta p$ , при этом проводимость полупроводника оказывается равной

$$\sigma = e[(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p]. \quad (7.2)$$

Таким образом, фотопроводимость равна:

$$\sigma_{\phi n} = \sigma - \sigma_0 = e(\Delta n\mu_n + \Delta p\mu_p). \quad (7.3)$$

Увеличение концентрации подвижных носителей при облучении светом вызывает уменьшение электрического сопротивления и, соответственно, рост тока через полупроводник. На основе этого явления изготавливаются простейшие фотоэлектрические приемники — фоторезисторы.

Чувствительным элементом фоторезистора является кристаллическая пластинка полупроводника или полупроводниковая пленка на

изолирующей подложке. Во внешнюю цепь элемент подключается через омические контакты.

В качестве светочувствительных материалов для фоторезисторов применяют, как правило, полупроводниковые соединения типа  $A^{II}B^{VI}$  (сульфид и селенид кадмия  $CdS$ ,  $CdSe$  — видимый диапазон) или  $A^{IV}B^{VI}$  (сульфид и селенид свинца  $PbS$ ,  $PbSe$  — инфракрасный диапазон).

Вольтамперная характеристика (рис. 7.1, а) отображает зависимость тока через фоторезистор от приложенного напряжения; параметром семейства является величина светового потока. Строго говоря, эта зависимость не линейна. В области малых напряжений сказывается влияние неидеальных омических контактов. При больших напряжениях из-за разогрева фоторезистора заметно увеличивается темновая проводимость  $\sigma_0$  (см. формулу (7.1)) и, соответственно, темновой ток, и скорость нарастания фототока уменьшается.

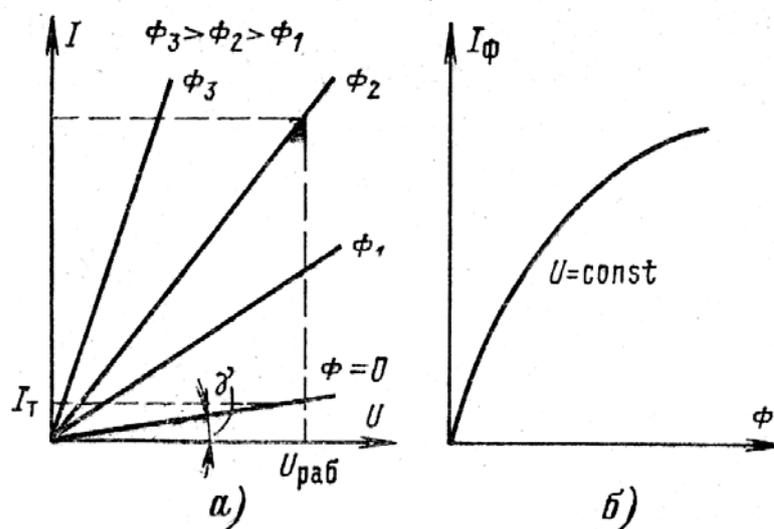


Рис. 7.1. Вольтамперная (а) и энергетическая (б) характеристики фоторезистора

Однако в области рабочих напряжений (рис. 7.1, а) характеристики практически линейны, их наклон определяет сопротивление

фоторезистора при данном световом потоке. В частности, котангенс угла наклона ВАХ при  $\Phi = 0$  равен темновому сопротивлению фоторезистора, его величина обычно находится в пределах  $10^6 - 10^8$  Ом. При облучении светом сопротивление фоторезистора снижается на 3–4 порядка.

Энергетическая характеристика фоторезистора (рис. 7.1, б) линейна для невысоких значений светового потока  $\Phi$ . При увеличении  $\Phi$  рост фототока замедляется: значительный рост концентрации подвижных носителей приводит к увеличению скорости рекомбинации.

Важнейшим параметром фоточувствительного прибора является токовая чувствительность — отношение фототока к некоторой величине, характеризующей падающий световой поток. Для фоторезисторов в качестве такого параметра используют удельную интегральную чувствительность  $S_{\Phi \text{ инт.уд.}}$ , (мА/В·лм): отношение фототока к световому потоку, приходящееся на 1 Вольт рабочего напряжения:

$$S_{\Phi \text{ инт.уд.}} = \frac{I_{\Phi}}{\Phi \cdot U}. \quad (7.4)$$

Как следует из семейства ВАХ (рис. 7.1, а), чувствительность фоторезистора зависит от того, при каком напряжении она измеряется. Поэтому, для исключения влияния напряжения используют удельную чувствительность  $S_{\Phi \text{ инт.уд.}}$ , приведенную к 1 вольту напряжения. Ее величина для различных типов фоторезисторов колеблется в пределах от десятых долей до сотен единиц мА/В·лм.

При включении фоторезистора в измерительную цепь как составной части делителя (рис. 7.2, а) выходным сигналом является изменение падения напряжения на нагрузочном резисторе  $R_n$ . Если световой поток сформирован в виде импульсов, то на сопротивлении нагрузки  $R_n$  появится импульс напряжения с амплитудой  $\Delta u_{\text{вых}} = \Delta I_{\Phi} \cdot R_n$ , повторяющий форму импульса тока фоторезистора.

В этом случае наиболее удобным параметром является вольт-овая чувствительность  $S_U$  (В/лм) при данном рабочем напряжении:

$$S_U = \frac{\Delta u_{\text{ВЫХ}}}{\Phi} \quad (7.5)$$

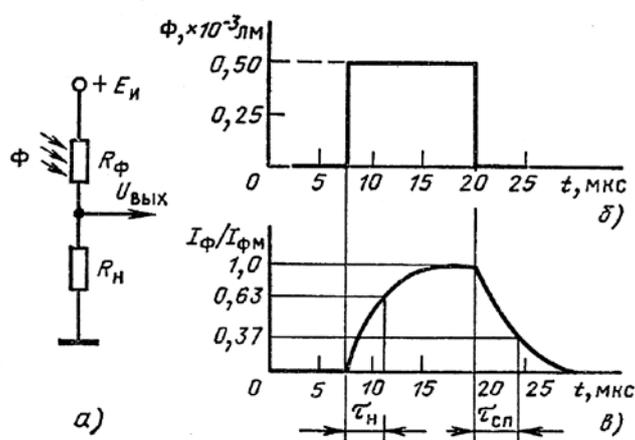


Рис. 7.2. Измерительная цепь фоторезистора (а) и искажения, вносимые при преобразовании импульсного светового сигнала (б)

При облучении фоторезистора прямоугольными световыми импульсами форма выходного импульса искажается (рис. 7.2, б): ток нарастает и уменьшается с запаздыванием относительно светового импульса. Инерционность фоторезистора связана со временем жизни неравновесных носителей и определяется двумя параметрами: временем нарастания  $\tau_n$  и временем спада  $\tau_{сп}$ . Эти параметры тесно связаны с граничной частотой  $f_{гр}$ . Она обратно пропорциональна времени жизни носителей, ее средние значения составляют  $10^4$  Гц. Таким образом, чем больше время жизни неравновесных носителей, тем более инерционен фоторезистор.

### 7.3. ФОТОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ В P–N-ПЕРЕХОДЕ

При облучении *p–n*-перехода светом с энергией квантов, превышающей ширину запрещенной зоны, во всей области поглощения происходит биполярная световая генерация подвижных носителей. Дальнейшая «судьба» этих электронов и дырок зависит от того, в каком месте они появились.

1. Если электронно-дырочные пары возникли непосредственно в переходе с шириной  $l$ , то они разделяются контактным электрическим

полем и переходят в те области, где становятся основными носителями.

2. Если световая генерация происходит в частях  $p$ - и  $n$ -областей, прилегающих к границам перехода шириной порядка диффузионной длины неосновных носителей ( $L_p$  и  $L_n$ , см. раздел 4.3), то электронно-дырочные пары также разделяются. Основные носители остаются в месте возникновения электронно-дырочной пары, а неосновные диффундируют в область  $p$ - $n$ -перехода и затем за счет экстракции переходят в смежную область, где также становятся основными носителями.

3. Электронно-дырочные пары, образовавшиеся в глубине  $p$ - и  $n$ -областей, в течение времени жизни не попадают в контактное электрическое поле перехода и рекомбинируют.

Таким образом, при постоянном уровне облучения в области шириной  $l + L_p + L_n$  происходят следующие процессы:

Ток, обусловленный электронами и дырками из разделенных полем перехода электронно-дырочных пар (фототок  $I_\phi$ ) совпадает по направлению с дрейфовым током неосновных носителей  $I_0$ . Следовательно, через переход протекает суммарный дрейфовый ток  $I_\phi + I_0$ . Концентрации электронов в  $n$ -области и дырок в  $p$ -области повышаются, происходит частичная компенсация объемного заряда неподвижных ионов примесей. Вследствие этого потенциальный барьер перехода понижается, и возникает диффузионный ток основных носителей  $I_{дф}$ , направленный навстречу дрейфовому току. Увеличиваясь по мере снижения барьера, диффузионный ток становится равным встречному дрейфовому току. В результате на переходе при постоянном уровне облучения устанавливается равновесие, при котором суммарный ток  $I = I_{дф} - I_\phi - I_0$  равен нулю, и на переходе существует фото-ЭДС или напряжение холостого хода  $U_{хх}$  (рис. 7.3, б) обусловленное накопленными основными носителями в  $p$ - и  $n$ -областях.

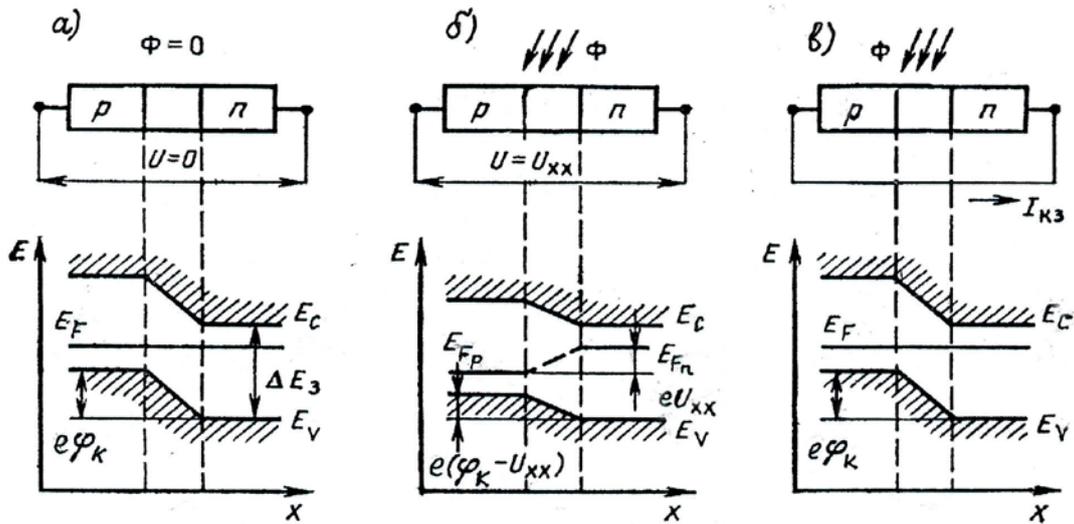


Рис. 7.3. Энергетические диаграммы  $p-n$ -перехода:  
 а) в отсутствие освещения;  
 б) при освещении в режиме холостого хода,  
 в) при освещении в режиме короткого замыкания

Как было установлено ранее (см. раздел 6.5), диффузионный ток основных носителей при данном прямом напряжении (в нашем случае  $U_{xx}$ ) равен  $I_{дф} = I_0 \exp\left(\frac{eU_{xx}}{kT}\right)$ . Для равновесного состояния  $I_{\phi} = I_{дф} - I_0$ , следовательно  $I_{\phi} = I_0 \left[ \exp\left(\frac{eU_{xx}}{kT}\right) - 1 \right]$ , и напряжение холостого хода:

$$U_{xx} = \frac{kT}{e} \ln \left( 1 + \frac{I_{\phi}}{I_0} \right). \quad (7.6)$$

На энергетической диаграмме освещенного перехода в режиме холостого хода (рис. 7.3, б) видно, что световой поток смещает уровни Ферми  $p$ - и  $n$ -областей на величину  $eU_{xx}$  из-за повышения концентраций основных носителей, что приводит к снижению потенциального барьера на ту же величину.

При замыкании  $p$ - и  $n$ -областей (рис. 7.3, в) через переход течет ток короткого замыкания:

$$I_{\phi} = I_{\text{кз}} = eSG(l + L_n + L_p), \quad (7.7)$$

где  $G$  — скорость биполярной световой генерации,  $S$  — площадь перехода. При коротком замыкании перехода генерируемые светом носители не накапливаются в смежных областях и не компенсируют объемный заряд ионов примесей. Они образуют дрейфовый ток короткого замыкания во внешней цепи. В силу этих причин энергетическая диаграмма для данного случая полностью подобна равновесной диаграмме  $p$ - $n$ -перехода при отсутствии освещения (см. рис. 7.3, а, в).

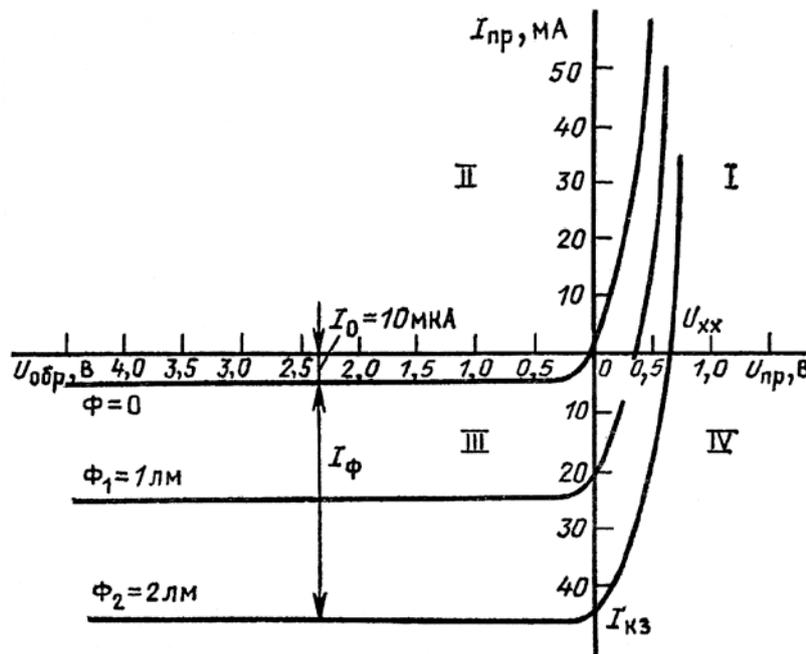


Рис. 7.4. Пример вольтамперной характеристики  $p$ - $n$ -перехода при разных уровнях освещенности

Вольтамперная характеристика  $p$ - $n$ -перехода, облученного светом, (рис. 7.4) пересекает ось прямых напряжений в точке  $U = U_{\text{хх}}$  и ось токов при  $I = I_{\text{кз}}$ , т. е. она смещена вниз относительно ВАХ в отсутствие освещения тем больше, чем выше значение светового потока. На характеристиках можно выделить две области, в которых возможно практическое применение фотоэффекта в  $p$ - $n$ -переходе:

область обратных напряжений (III) — режим фотодиода, область IV — режим фотоэлемента.

В режиме фотоэлемента (IV)  $p$ - $n$ -переход является источником электрической энергии. При включении нагрузки параллельно переходу протекающий в цепи фототок обеспечивает полезную мощность, выделяющуюся за счет падения напряжения на нагрузочном сопротивлении. Такие элементы на основе кремния собираются в последовательно-параллельно соединенные блоки и служат основой для изготовления солнечных батарей.

При обратном смещении перехода и освещении его светом работают широко распространенные фотоприемники — фотодиоды. Их вольтамперная характеристика представляет собой часть ВАХ перехода, облучаемого светом, лежащую в области III (рис. 7.4). Ток фотодиода образован неосновными носителями, концентрация которых увеличивается пропорционально световому потоку при освещении. Так же, как и для фоторезистора, фототок  $I_{\Phi}$  равен общему току  $I$  за вычетом темнового тока  $I_0$ . Основным параметром фотодиода — интегральная токовая чувствительность  $S_I$ , определяемая отношением фототока либо к световому потоку (мА/лм), либо к мощности светового облучения (мА/Вт):

$$S_{I\Phi} = I_{\Phi} / \Phi; S_{IP} = I_{\Phi} / P \quad . \quad (7.8)$$

В отличие от фоторезистора,  $S_{I\Phi}$  и  $S_{IP}$  не зависят от напряжения, так как фототок при изменении напряжения на фотодиоде остается постоянным.

## 7.4. ИЗЛУЧЕНИЕ СВЕТА ПОЛУПРОВОДНИКАМИ

В основе принципа действия полупроводниковых излучателей света лежит явление электролюминесценции — излучения света полупроводниками под действием внешнего электрического поля.

Практически все переходы электронов, при которых их энергия уменьшается, могут сопровождаться излучением в той или иной спектральной области. Используя полупроводники с различной шириной запрещенной зоны и различные примеси, можно получить излучение во всем видимом и ближнем инфракрасном диапазоне.

Излучающие полупроводниковые приборы преобразуют электрическую энергию в энергию оптического излучения в узком диапазоне длин волн. Электролюминесценция в этом случае происходит за счет излучательной рекомбинации. Для того чтобы поддерживать излучение света из полупроводника, необходимо создавать условия для создания высокой избыточной концентрации подвижных носителей. Наиболее простой и очевидный способ – использование явления инжекции неосновных носителей в прямо смещенном  $p-n$ -переходе.

Полупроводниковые переходы могут быть источниками некогерентного и когерентного излучения. В первом случае это — светодиоды, во втором — инжекционные лазеры.

## **7.5. ИНЖЕКЦИОННЫЕ СВЕТОДИОДЫ С $P-N$ -ПЕРЕХОДАМИ**

В данном разделе будут рассмотрены физические механизмы некогерентного светового излучения, обусловленного самопроизвольной излучательной рекомбинацией носителей заряда, инжектированных через прямо смещенный  $p-n$ -переход.

Наибольший интерес представляет излучение света в видимой области, поэтому необходимо использовать полупроводники с широкой запрещенной зоной (1,8–3,4 эВ). Основные полупроводниковые материалы для светодиодов: нитрид галлия GaN ( $\Delta E_z = 3,39$  эВ), фосфид галлия GaP ( $\Delta E_z = 2,27$  эВ), арсенид галлия ( $\Delta E_z = 1,4$  эВ), карбид кремния ( $\Delta E_z = 2,36$ – $3,05$  эВ, в зависимости от типа кристаллической

решетки). Используются также тройные соединения, например  $\text{GaAs}_{1-x}\text{P}_x$ , в которых путем изменения соотношения атомов мышьяка и фосфора в интервале  $0 \leq x \leq 0,45$  можно изменять ширину запрещенной зоны для прямых излучательных переходов от 1,4 эВ до 2,0 эВ.

Поскольку вероятность такой рекомбинации невелика (см. раздел 4.5), важнейшей задачей является увеличение доли излучательных переходов между уровнями или энергетическими зонами. Эта задача решается, прежде всего, путем использования полупроводников, в которых происходит межзонная излучательная рекомбинация.

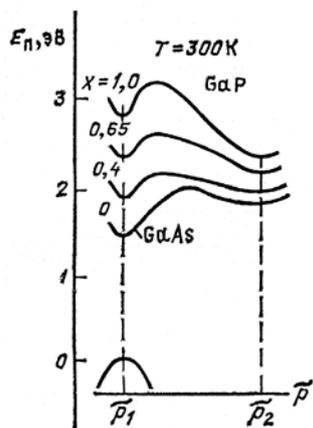


Рис. 7.5. Зависимость энергии электронов от квазиимпульса для арсенида галлия, фосфида галлия и тройного соединения  $\text{GaAs}_{1-x}\text{P}_x$

Такие полупроводники называются прямозонными; в них рекомбинация электрона и дырки с испусканием кванта света происходит без участия кристаллической решетки.

Модель плоских энергетических зон не пригодна для отображения физической картины процессов излучательной и безызлучательной рекомбинации. Для этого необходимо учесть зависимость энергии, соответствующей краям зон, от квазиимпульса частиц.

Так, из рис. 7.5 следует, что арсенид галлия  $\text{GaAs}$  — прямозонный полупроводник: абсолютный минимум энергии дна зоны проводимости и максимум энергии потолка валентной зоны находятся при

одном и том же значении квазиимпульса  $\tilde{p}_1$ . Напротив, фосфид галлия GaP не прямозонный полупроводник, так как абсолютные экстремумы краев зон находятся при разных значениях квазиимпульса  $\tilde{p}_1$  и  $\tilde{p}_2$ . Здесь же хорошо видно, что трехкомпонентное вещество GaAs<sub>1-x</sub>P<sub>x</sub> остается прямозонным полупроводником только при увеличении доли фосфора  $x$  от 0 до 0,45.

Несмотря на то, что в GaP невозможны прямые переходы, он широко используется как материал для изготовления светодиодов видимой области спектра. Для этого в полупроводник вводят чужеродные атомы для образования ловушек — центров излучательной рекомбинации. Чужеродные атомы образуют ловушки донорного (S, O<sub>2</sub>) и акцепторного типов (Cd, Zn, Cu). Такие же примеси вводятся в трехкомпонентный полупроводник GaAs<sub>1-x</sub>P<sub>x</sub> при  $x > 0,45$ .

Такое легирование открывает возможность излучательной рекомбинации с участием ловушек за счет прямых переходов на конечной стадии ступенчатой рекомбинации (см. рис. 4.3) и за счет таких же переходов между уровнями ловушек донорного и акцепторного типов. В результате на основе фосфида галлия удалось создать светодиоды, дающие излучение в зеленой и красной части спектра.

Современные светодиоды излучают красный, желтый, зеленый, оранжевый, синий и белый свет. Наиболее широкое применение светодиодов — устройства отображения цифр и знаков в световых табло и экранах различного назначения; светофоры, габаритные огни легковых автомобилей. В последнее время светодиоды широко используются как источники белого света.

Для изготовления светодиодов часто используют гетеропереходы. Их основное преимущество — односторонняя инжекция из широкогозонного полупроводника в узкозонный, что позволяет существенно увеличить интенсивность излучательной рекомбинации.

## 7.6. ПОЛУПРОВОДНИКОВЫЕ ИНЖЕКЦИОННЫЕ ЛАЗЕРЫ

В разделе 7.5 было показано, что инжекция электронов или дырок на  $p$ - $n$ -переходе приводит к спонтанной рекомбинации и испусканию некогерентного света. Полупроводниковые структуры на основе обычных  $p$ - $n$ -переходов или с использованием гетеропереходов могут служить также источниками когерентного оптического излучения. Для этого в активном, излучающем слое структуры необходимо создать инверсную заселенность энергетических уровней, или, иначе говоря, обеспечить как можно большее заполнение электронами возбужденных уровней. Создание инверсной заселенности достигается путем интенсивной инжекции неосновных носителей (как правило, электронов в  $p$ -область).

Излучение, возникающее при первых актах рекомбинации электронов и дырок, вызывает вынужденные переходы других электронов из возбужденных энергетических состояний на нижние энергетические уровни. При этом возникающее излучение имеет строго ту же частоту, фазу и плоскость поляризации, что и первоначальное.

В лазерах любого типа обязательно используется удержание фотонов в активной среде с помощью оптических резонаторов — плоскопараллельных зеркал. В полупроводниковом лазере роль зеркал играют строго параллельные отполированные грани монокристалла. В результате свет многократно проходит область инверсной заселенности, и возникает интенсивное вынужденное излучение.

Для изготовления полупроводниковых лазеров используют прямозонные вырожденные полупроводники с концентрацией подвижных носителей  $\sim 10^{20} \text{ см}^{-3}$ , например, арсенид галлия или GaAlAs  $n$ - и  $p$ -типа. В таких полупроводниках возможна межзонная излучательная рекомбинация без участия кристаллической решетки как третьего тела, необходимого для соблюдения закона сохранения импульса.

Наиболее эффективные полупроводниковые лазеры изготавливают с использованием гетеропереходов, обладающих свойствами односторонней инжекции и разными величинами потенциальных барьеров для электронов и дырок. Такие переходы применяются для удержания инжектированных носителей в активной зоне высоким потенциальным барьером, а также для обеспечения интенсивной инжекции носителей одного типа в активную зону полупроводникового лазера.

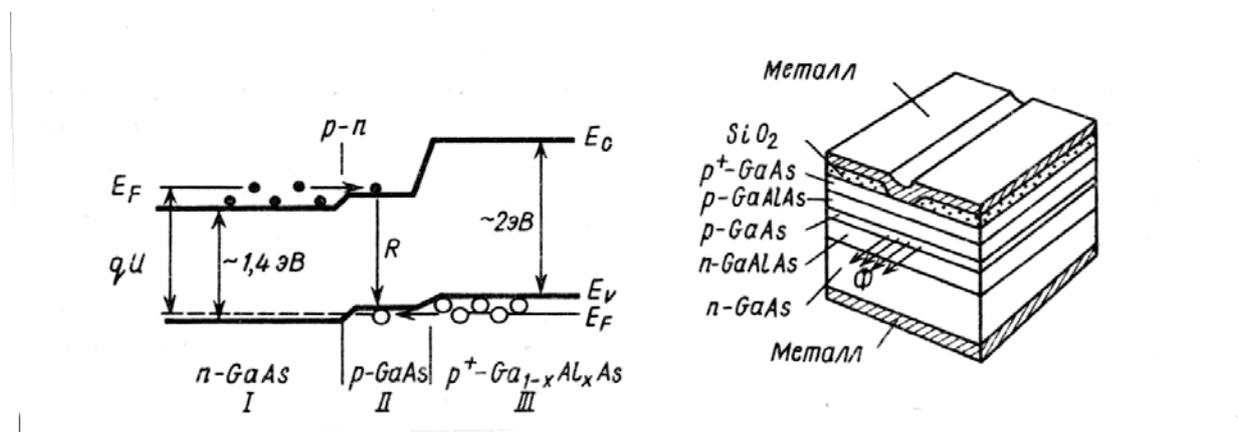


Рис. 7.6. а) — энергетическая схема двухпереходного инжекционного лазера с  $p-n$ -переходом на основе вырожденного арсенида галлия и гетеропереходом  $p-p^+$  GaAs – Ga<sub>1-x</sub>Al<sub>x</sub>As. Подано внешнее напряжение  $U$ . б) — один из вариантов структуры полупроводникового лазера

Рассмотрим принцип работы полупроводникового лазера с использованием энергетической диаграммы на рис. 7.6. Через  $p-n$ -переход инжектируются электроны из области I в область II. Установлено, что для создания инверсии заселенности в полупроводниковом лазере на основе вырожденного  $p-n$ -перехода на него необходимо подавать прямое напряжение такой величины, чтобы соблюдалось условие  $qU > h\nu = \Delta E_z$ . (здесь  $q$  — заряд электрона). При этом в активной области II создается инверсная населенность:

концентрация электронов у дна зоны проводимости  $p$ - GaAs превышает концентрацию дырок в валентной зоне. Переход  $p-p^+$  между областями  $II$  и  $III$  создает потенциальный барьер для электронов, удерживая их в активной области и способствуя их накоплению.

Когерентное излучение возникает при достижении порогового тока через переход  $I - II$  (от нескольких десятков до нескольких сотен миллиампер). Для уменьшения рабочих токов и ослабления нагрева используется тонкий активный слой толщиной 5–20 мкм, расположенный между отражающими поверхностями.

В качестве материала, инжектирующего электроны в активную область, используют также более широкозонный полупроводник GaAlAs (рис. 7.6, б). В этом случае активный слой  $II$  располагается между двумя широкозонными полупроводниками, обладающими меньшим коэффициентом преломления. Это приводит к усилению отражения света от боковых границ, т. е. к уменьшению потерь.

Мощность излучения полупроводниковых лазеров, работающих в непрерывном режиме, составляет десятые доли Ватта; в импульсном режиме мощность оказывается значительно выше, так как уменьшаются потери на нагрев устройства. Его размеры составляют несколько миллиметров, коэффициент полезного действия превышает 50 %. Инжекционные лазеры обладают высоким быстродействием: инерционность составляет  $\sim 10^{-9}$  с. Световой поток легко модулируется по интенсивности при изменении напряжения.

### **Вопросы к разделу 7**

1. Какие физические процессы происходят при взаимодействии света с полупроводниками?
2. Опишите механизм возникновения фотопроводимости полупроводников.
3. Каков принцип действия простейшего фотоприемника — фоторезистора? Каковы его вольтамперные характеристики и параметры?
4. Какие процессы происходят в  $p-n$ -переходе и прилегающих к нему областях при облучении светом в условиях биполярной генерации?
5. Объясните причину существования напряжения холостого хода и тока короткого замыкания перехода, облучаемого светом.
6. Как изменяется вольтамперная характеристика  $p-n$ -перехода при облучении светом?
7. Каким образом можно использовать  $p-n$ -переход, облучаемый светом, в качестве фотоприемника?
8. Каким образом можно использовать  $p-n$ -переход, облучаемый светом, в качестве источника напряжения?
9. Какое физическое явление обуславливает излучение некогерентного света  $p-n$ -переходом? Как можно обеспечить излучение света в видимой области спектра?
10. Перечислите способы увеличения доли излучательной рекомбинации в переходе. Дайте соответствующие обоснования.
11. Каковы необходимые условия для возникновения когерентного излучения в  $p-n$ -переходе?
12. Поясните, каким образом можно удерживать инжектированные в активную  $p$ -область лазера электроны.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Росадо Л.* Физическая электроника и микроэлектроника [пер. с испанского] / Л. Росадо — М.: Высшая школа, 1991. — 351 с.
2. *Жеребцов И. П.* Основы электроники / И. П. Жеребцов — Л.: Энергоатомиздат, 1989. — 352 с.
3. *Степаненко И. П.* Основы микроэлектроники / И. П. Степаненко — М.: Лаб. базовых знаний, 2004. — 488 с.
4. *Митрофанов О. В.* Физические основы функционирования изделий микроэлектроники / О. В. Митрофанов, Б. М. Симонов, Л. А. Коледов — М.: Высш. шк., 1987. — 168 с.
5. *Ефимов И. Е.* Микроэлектроника: физические и технологические основы, надежность: / И. Е. Ефимов, И. Я. Козырь, Ю. И. Горбунов.— М.: Высшая школа, 1986 .— 463 с.
6. *Морозова И. Г.* Физика электронных приборов: / И. Г. Морозова. — Москва: Атомиздат, 1980 .— 392 с.
7. *Зи С.* Физика полупроводниковых приборов. Кн. 1 / С. Зи. — Москва: Мир, 1984. — 455 с.
8. *Соболев В. Д.* Физические основы электронной техники: / В. Д. Соболев. — М.: Высш. шк., 1979. — 447 с.
9. *Дулин В. Н.* Электронные приборы / [В. Н. Дулин [и др.]]; под ред. Г. Г. Шишкина. — М.: Энергоатомиздат, 1989. — 494 с.
10. *Пасынков В. В.* Полупроводниковые приборы / В. В. Пасынков, Л. К. Чиркин. — СПб.: Лань, 2009.— 480 с.

### Web-сайты

11. <http://avnsite.narod.ru/physic/pp/>
12. <http://www.club155.ru/semiconductors-physics>