

Министерство образования и науки, молодежи и спорта Украины
Государственное высшее учебное заведение
«Национальный горный университет»

Библиотека иностранного студента

И.П.Гаркуша

Элементы физики полупроводников

Учебное пособие для бакалавров отрасли знаний
«Разработка полезных ископаемых»

Днепропетровск
НГУ
2012

УДК 53(075.4)
ББК 22.379
Г 43

Гаркуша И.П.

Г 43 Элементы физики полупроводников (Текст): учеб пособие:– Д.: Национальный горный университет, 2012.– 74 с. – (Библиотека иностранного студента).

Учебное пособие предназначено для самостоятельной работы студентов направления 6.050301 Горное дело.

Пособие содержит изложение основных представлений физики полупроводников. Рассмотрены элементы зонной теории твердых тел, понятие о статистике электронов и дырок в полупроводниках, электронно-дырочные переходы, выпрямление и усиление переменных токов с помощью *p-n*-переходов. Дается описание принципа действия полупроводниковых устройств – выпрямительного диода, транзистора, фото- и светодиода. Приводятся некоторые применения полупроводниковых устройств.

Учебное пособие может быть использовано в качестве дополнения к существующим учебникам, рассчитано на студентов дневной и заочной формы обучения, может быть также полезно преподавателям кафедр физики технических университетов.

О Г Л А В Л Е Н И Е

Введение Электронное строение металлов, диэлектриков и полупроводников.....	2
§ 1. Энергетические уровни изолированного атома.....	4
§ 2. Расщепление энергетических уровней и образование энергетических зон.....	5
§ 3. Металлы, диэлектрики и полупроводники в модели зонной теории.....	8
§ 4. Понятие о квантовой статистике электронов в твердых телах. Распределение Ферми-Дирака. Энергия Ферми и ее физический смысл для металлов.	12
§ 5. Полупроводники. Собственная проводимость полупроводников	16
§ 6. Собственная проводимость полупроводников в терминах зонной теории. Энергия Ферми для полупроводников.....	19
§ 7. Примесная проводимость полупроводников.....	26
§ 8. Контакт <i>p</i> - и <i>n</i> -полупроводников. Свойства <i>p-n</i> -перехода.....	31
§ 9. Неравновесное состояние <i>p-n</i> -перехода. Прямое и обратное смещение.....	33
§ 10. Контакт <i>p</i> - и <i>n</i> -полупроводников в терминах зонной теории и энергетических барьеров.....	35
§ 11. Выпрямление переменных токов с помощью <i>p-n</i> -перехода. Полупроводниковый выпрямительный диод.....	39
§ 12. Светоизлучающий диод.....	43
§ 13. Усиление переменных токов с помощью <i>p-n</i> -переходов. Транзистор.....	46
§ 14. Понятие о планарной технологии изготовления интегральных схем.....	50
§ 15. Генератор незатухающих электромагнитных колебаний на транзисторе.....	52
§ 16. Фотоэффект в полупроводниках. Фоторезисторы.....	55
§ 17. Фотодиод.....	57
§ 18. Солнечные элементы.....	60
§ 19. Туннельный эффект в полупроводниках. Туннельный диод.....	62
§ 20. Полупроводниковый лазер.....	65
Ответы на контрольные вопросы	70
Список рекомендованной литературы.....	74

Введение. Электронное строение металлов, диэлектриков и полупроводников

Свойства полупроводников можно описать на основе электронного строения кристаллических твердых тел.

Твердое тело представляет собой совокупность отдельных атомов, химические связи между которыми объединяют их в кристаллическую решетку. По типам связи твердые тела делятся на несколько классов, каждый из которых характеризуется определенным пространственным распределением электронов.

При образовании кристаллической решетки **металла** часть валентных

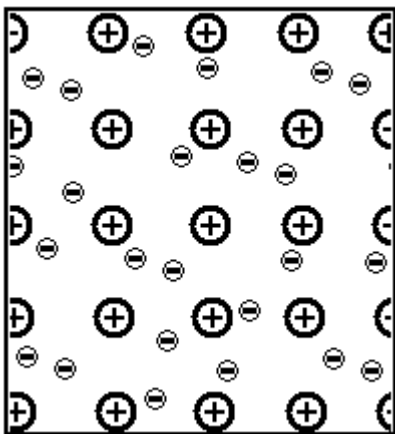


Рис. 1

электронов отрывается от своих атомов. Электроны, как принято говорить, обобществляются, они не принадлежат теперь никакому конкретному иону металла и могут свободно перемещаться под действием внешнего электрического поля, образуя электрический ток.

Кристаллическая решетка металла образована не нейтральными атомами, а положительно заряженными ионами (рис.1). Металл можно представить в виде остова из положительных ионов, погруженного в «электронный газ». Этот газ свободных электронов заполняет кристаллическую решетку и компенсирует силы взаимного электростатического

отталкивания положительных ионов и тем самым связывает их в твердое тело.

Наличие свободных электронов в металлах подтверждается явлениями проводимости электрического тока и электронной эмиссии. В каждом кубическом сантиметре металла содержится примерно 10^{22} свободных электронов.

Употребляя термин «свободный электрон», необходимо иметь в виду, что по-настоящему свободный электрон может быть лишь в вакууме. Если на него не действует электрическое поле, он не испытывает никаких внешних воздействий. Под действием внешнего электрического поля напряженностью E свободный электрон в вакууме движется с постоянным ускорением $a = eE/m$. В кристалле же электрон в результате частых столкновений даже в очень сильных электрических полях движется против поля со скоростью, не превышающей $(1\div 3)\cdot 10^5$ м/с. Так что термин «свободный электрон» употребляется в смысле его способности направленно двигаться под действием электрического поля, обуславливая электрический ток.

Совсем другой характер связей между атомами, образующими кристалл **диэлектрика и полупроводника**. Кристаллическая решетка диэлектрика и полупроводника образована нейтральными атомами. Так как в кристалле атомы сближены на расстояние порядка размеров самого атома, то электронные оболочки атомов сильно перекрываются, и происходит непрерывный переход валентных электронов от одного атома к другому.

В качестве примера рассмотрим кристаллическую решетку *кремния*, схематически изображенную на рис. 2. Четыре электрона из восьми, образующих замкнутую электронную оболочку некоторого атома кремния (произвольно выбранный атом выделен на схеме), принадлежат самому атому, остальные четыре – от соседних по кристаллической решетке атомов. Но атом кремния не отбирает эти электроны у соседей: они проводят возле выделенного нами атома лишь часть времени, другую часть они проводят возле «своего» атома. С другой стороны, валентные электроны рассматриваемого атома теперь проводят возле него только часть времени, другую часть – возле соседних атомов. Такой электронный обмен приводит к образованию прочной *ковалентной связи*.

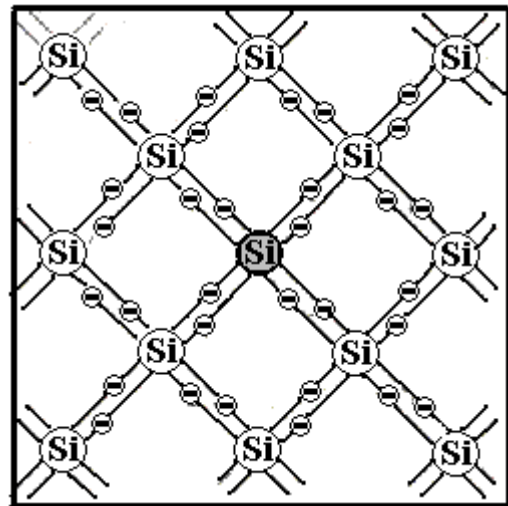


Рис. 2.

Чтобы разорвать электронные связи между атомами и создать таким образом свободные электроны, необходимо затратить энергию. Электронные связи могут быть разорваны за счет тепловых колебаний решетки, за счет облучения фотонами подходящей частоты и другими способами. Но при достаточно низкой температуре такой кристалл является идеальным изолятором (диэлектриком).

Энергию, необходимую для разрыва электронной связи, обозначают символом E_g . Для разных кристаллов она изменяется в пределах от нуля до десятка электрон-вольт.

В *металлах* концентрация электронов не зависит от температуры. Даже при температуре $T = 0$ К все электроны остаются свободными и сохраняют способность проводить ток.

В типичном *диэлектрике* с большим значением E_g (> 3 эВ), например, кварце, алмазе и др., тепловое движение не способно разорвать электронные связи и увеличить концентрацию электронов.

В материалах с небольшим значением E_g , нагревание даже до не слишком высокой температуры приведет к разрыву большого числа электронных связей и появлению свободных электронов. Концентрация свободных носителей тока будет очень резко зависеть от температуры. Этот класс веществ называется *полупроводниками*.

Значение E_g для типичных полупроводников лежит в пределах от нескольких десятых долей электрон-вольта до двух-трех электрон-вольт. Так, для антимонида индия (InSb) значение $E_g = 0,18$ эВ, для германия (Ge) $E_g = 0,67$ эВ, для кремния (Si) $E_g = 1,1$ эВ, арсенида галлия (GaAs) $E_g = 1,4$ эВ. В тройном полупроводниковом соединении GaAlAs, являющемся материалом для светодиодов и лазеров, E_g лежит в пределах от 1,4 эВ до 2,17 эВ.

§ 1. Энергетические уровни изолированного атома

В соответствии с квантовой механикой *свободные* электроны в вакууме могут иметь *любую энергию* — их энергетический спектр непрерывен.

Если же электроны принадлежат *атомам*, то их энергия квантуется — они имеют определённые *дискретные* значения энергии. Эти разделённые конечными промежутками значения энергии E_0, E_1, E_2, \dots ($E_0 < E_1 < E_2, \dots$) называются *уровнями энергии*. Квантовая теория запрещает иметь электрону энергию, лежащую между любыми двумя уровнями энергии.

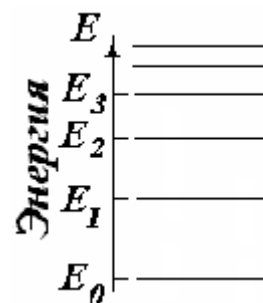


Рис. 3.

Графически состояния с различной энергией можно изобразить по аналогии с потенциальной энергией тела, поднятого на различные высоты (уровни), в виде диаграммы уровней энергии (рис. 3). По вертикали снизу вверх откладываются значения полной энергии электронов.

Каждому значению энергии соответствует горизонтальная линия, проведённая на высоте E_i ($i = 0, 1, 2, \dots$). Такая совокупность дискретных уровней энергии атома образует его *дискретный энергетический спектр*.

Нижний уровень E_0 , соответствующий наименьшей возможной энергии электрона в атоме, называется основным, а все остальные уровни энергии E_1, E_2, \dots — возбуждёнными, т.к. для перехода на них электрона ему необходимо сообщить энергию. Значения энергии принято отсчитывать от основного уровня. Шкала энергий обычно выражается в электронвольтах (эВ).

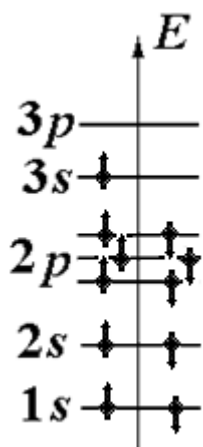


Рис. 4.

Каждый электрон атома обладает одним из разрешённых значений энергии, т. е. занимает один из дозволённых энергетических уровней.

Как известно, любая физическая система в свободном состоянии должна обладать минимумом возможной энергии. Поэтому в основном, невозбужденном состоянии атома суммарная энергия электронов должна быть минимальной. Казалось бы при этом, что все электроны должны находиться на самом низком уровне. Однако электроны подчиняются *принципу запрета Паули*, согласно которому

в каком-либо квантовом состоянии может находиться *только один* электрон.

С учетом спина электрона принцип Паули утверждает, что в любой физической системе (атоме, молекуле, кристалле и т. д.) на каждом энергетическом уровне может находиться *не более двух* электронов, причем собственные моменты импульса (спины) электронов, занимающих одновременно один и тот же уровень, должны иметь противоположные направления. Следовательно, на самом низком уровне атома может разместиться только два электрона, остальные заполняют попарно более высокие уровни.

На рис. 4 показано размещение электронов по уровням в основном состоянии атома натрия, имеющего 11 электронов. В изолированном атоме натрия

имеются две заполненные электронные оболочки, содержащие соответственно 2 и 8 электронов. Эти состояния электронов записываются символической формулой

$$(1s)^2 (2s)^2 (2p)^6 (3s)^1.$$

Такая запись означает, что в $1s$ -состоянии находится два электрона, в $2s$ -тоже два, в $2p$ -состоянии – шесть, в $3s$ состоянии – один. Одиннадцатый валентный электрон атома натрия заполняет лишь наполовину верхний энергетический уровень.

Схема уровней изображена условно, без соблюдения масштаба. Электроны обозначены кружками со стрелкой. Разные направления стрелок соответствуют противоположным направлениям спинов.

Контрольные вопросы

- 1.1. Как определить число электронов в атоме?
- 1.2. Каковы масса и размеры атома? Чем определяются размеры атома – размерами ядра или электронной оболочки?
- 1.3. Какой характер имеет энергетический спектр электронов в атоме?
- 1.4. В чем состоит принцип запрета Паули? Как происходит заполнение электронных оболочек в изолированном атоме?
- 1.5. Могут ли все электроны в атоме занять самый низкий энергетический уровень?
- 1.6. Сколько электронов может разместиться на одном энергетическом уровне?
- 1.7. Может ли на каком-либо уровне энергии находиться один электрон?
- 1.8. Может ли какой-либо уровень энергии оказаться вакантным (пустым)?

§ 2. Расщепление энергетических уровней и образование энергетических зон

Энергия электронов квантуется не только в атомах, но также и в кристаллах.

Рассмотрим мысленный процесс образования кристалла путем сближения отдельных атомов. Пусть первоначально имеется N изолированных атомов какого-либо вещества. Пока атомы изолированы друг от друга, каждый из них имеет одинаковую систему энергетических уровней, заполнение которых электронами осуществляется в каждом атоме независимо от заполнения уровней в других атомах. Принцип Паули, действуя в пределах атома, не накладывает дополнительных ограничений на состояние его электронов в связи с наличием других атомов.

Когда атомы сближаются друг с другом, образуя твердое тело, они представляют собой

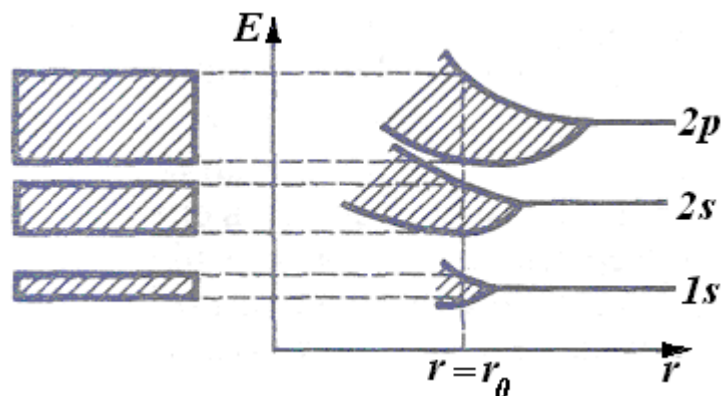


Рис.5.

единую квантовую систему, в пределах которой действует принцип запрета Паули.

Между атомами возникает все усиливающееся взаимодействие. Атомы тесно примыкают друг к другу, и каждый из них оказывает сильное воздействие на электроны соседнего атома. Электроны испытывают притяжение не только к «своему» ядру, но и к ядрам соседних атомов.

Это приводит к изменению энергетических состояний их электронов. Вместо одного одинакового для всех N атомов уровня возникают N очень близких, но не совпадающих (согласно принципу Паули) уровней энергии. Каждое квантовое состояние изолированного атома расщепляется на N квантовых состояний. Таким образом, каждый уровень изолированного атома **расщепляется** в кристалле на N густо расположенных уровней, Системы «разошедшихся» уровней образуют в кристалле разрешенные **энергетические полосы или зоны**.

Отсюда происходит название – зонная теория твердых тел. Решение уравнения Шредингера для электрона, движущегося в периодическом силовом поле кристаллической решетки, очень сложно и рассматривается в курсе теоретической физики. Главное следствие из этого решения – вывод о зонном характере энергетического спектра электронов.

На рис. 5: справа показаны энергетические уровни атома, а слева - расширение уровней в полосы при образовании кристаллической решетки. По горизонтальной оси отложено расстояние между атомами r . Через r_0 обозначено равновесное расстояние между атомами решетки кристалла.

Величина расщепления для разных уровней не одинакова. В квантовой механике движение электронов описывается волновой функцией. Так как атомы в твердом теле находятся на расстоянии порядка размеров самих атомов, то волновые функции наружных электронов соседних атомов перекрываются и распространяются на весь кристалл. Это приводит к тому, что валентные электро-

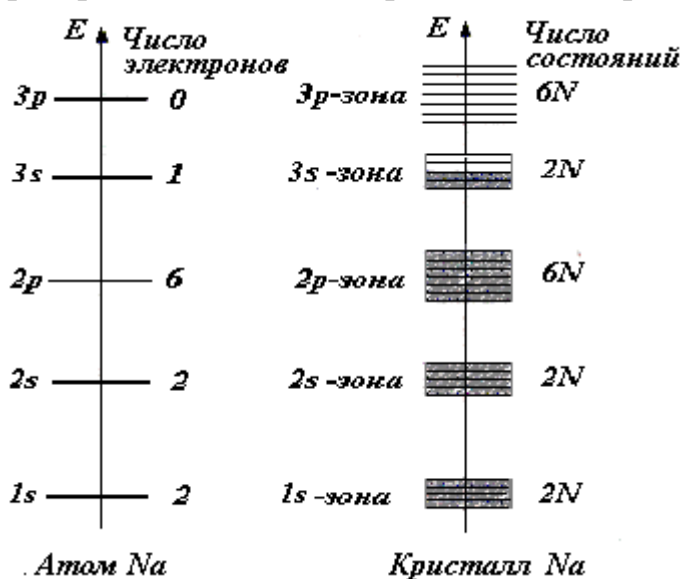


Рис. 6.

ны атомов теряют связь с определенными атомами и могут вести себя как свободные электроны – они теперь принадлежат не отдельным атомам, а всему кристаллу.

Изменения почти не затрагивают наиболее глубоких уровней, образующих внутренние, заполненные оболочки атомов. Волновые функции внутренних электронов друг с другом практически не перекрываются. Положение этих уровней в кристалле мало отличается от их положения

в изолированных атомах.

Расщеплению подвергаются и более высокие уровни, вообще не занятые электронами в основном состоянии атома.

На рис 6 для сравнения показаны энергетические диаграммы свободного атома натрия и кристалла натрия. Каждый из 11 электронов в нейтральном изолированном атоме Na занимает свой энергетический уровень, как это было показано на рис. 4.

Обозначим на рис. 6 в левом столбце числа электронов на соответствующем уровне изолированного атома натрия, они соответствуют данным рис. 4. Когда атомы сближаются на малые расстояния, образуя кристалл натрия, электроны в них приобретают отличающиеся друг от друга значения энергии.

Заполнение энергетических зон электронами происходит по тем же законам, что и заполнение электронных уровней в атоме, т.е. с учетом принципа Паули.

Количество электронов, энергии которых образуют данную зону, будет очевидно в N раз больше, чем количество электронов на соответствующем уровне изолированного атома (рис. 6).

Так в s -состоянии атома может находиться два электрона, а в кристалле, содержащем N атомов, $1s$ -уровень атома натрия превратится в $1s$ -зону, содержащую $2N$ вакантных мест. Все они будут заполнены $2N$ электронами, занимавшими $1s$ -состояния в свободных атомах. $1s$ -зона будет полностью заполнена электронами.

Так же будут заполнены $2s$ - и $2p$ -электронами атомов $2N$ вакантных мест $2s$ -зоны и $6N$ вакантных мест $2p$ -зоны. В $3s$ -зоне будет $2N$ вакантных мест, но у N атомов натрия есть лишь N валентных электронов в состоянии $3s$.

Следовательно, $3s$ -зона будет лишь наполовину заполнена электронами. Зоны, лежащие еще выше, будут совершенно не заполнены электронами. Штриховкой на рис. 6 отмечено заполнение зон электронами.

Энергетические интервалы ΔE между разрешенными зонами представляют собой энергию, которую электроны данных веществ приобрести не могут. Они называются **запрещенными зонами**.

Валентные электроны с противоположными спинами заполняют попарно (согласно принципу Паули) нижние уровни разрешенной зоны, возникшей из того уровня, на котором находятся валентные электроны в основном состоянии атома. Ее называют **валентной зоной**. Более высокие разрешенные зоны будут **пустыми, свободными** от электронов.

Физические свойства кристаллов определяются в основном только последними из занятых электронами зон (валентными зонами) и размещенными над ними запрещенными и разрешенными зонами. Электроны, находящиеся на более низких уровнях, практически не участвуют во взаимодействии атомов.

Контрольные вопросы

2.1. Какими факторами объясняется расщепление дискретных энергетических уровней изолированных атомов при постепенном сближении их и объединении в кристаллическую решетку?

2.2. Какими особенностями отличается движение электронов в периодическом потенциальном поле в кристаллах?

2.3. Оценить среднее расстояние между соседними атомами в кристаллической решетке алюминия, пользуясь следующими данными: молярная масса алюминия $M_{\text{Al}} = 27 \cdot 10^{-3}$ кг/моль, плотность алюминия $\rho = 2700$ кг/м³, число Авогадро $N_A = 6,02 \cdot 10^{23}$ моль⁻¹. Сравнить полученное расстояние с размерами атома ($\sim 10^{-10}$ м). Какие выводы из этого сравнения можно сделать?

2.4. Что представляют собой энергетические зоны в кристалле?

2.5. Опишите механизм образования энергетических зон при соединении отдельных атомов в кристалл.

2.6. Подвергаются ли расщеплению пустые, не занятые электронами энергетические уровни?

§ 3. Металлы, диэлектрики и полупроводники в модели зонной теории

Электрические свойства твердых тел определяются взаимным расположением различных энергетических зон и распределением электронов по этим зонам.

Ширина разрешенных и запрещенных зон определяется силами взаимодействия между атомами и не зависит от числа атомов в кристалле, а число уровней в зоне, естественно, прямо зависит от числа атомов N . Если кристалл состоит из N атомов, то он имеет в N раз больше энергетических уровней, чем отдельный атом. Таким образом, чем больше атомов содержит кристалл, тем теснее располагаются уровни в зоне.

Ширина разрешенных зон имеет величину порядка нескольких электрон-вольт. Как известно, количество атомов в 1 моле вещества равно числу Авогадро $N_A = 6,02 \cdot 10^{23}$ 1/моль. Следовательно, если полагать, что кристалл содержит порядка 10^{23} атомов, то расстояние между соседними уровнями в зоне составит приблизительно 10^{-23} эВ. Эта энергия ничтожно мала, поэтому зоны можно считать практически непрерывными,

Внешние воздействия (повышение температуры, облучение светом, сильные внешние электрические поля и др) могут увеличивать кинетическую энергию электронов. Оценим энергию таких воздействий.

Рассматривая совокупность электронов проводимости как электронный газ и проводя аналогию с классическим идеальным газом, можно оценить энергию теплового движения электрона. Как и для молекул идеального газа, она примерно равна kT . При комнатной температуре это составляет величину порядка 0,025 эВ. Эта величина несравненно больше, чем расстояние между соседними уровнями в зоне (10^{-23} эВ).

Следовательно, при комнатной температуре за счет теплового движения значительная часть электронов приобретает дополнительную кинетическую энергию. На энергетической диаграмме росту энергии электрона соответствует движение электрона по шкале энергий вверх на более высокие энергетические уровни.

Существование электрического тока (направленного потока электронов) также означает, что электроны приобретают дополнительную по отношению к основному состоянию энергию. Дополнительная энергия, вызванная действием

на электрон внешнего электрического поля, составляет примерно $10^{-8} - 10^{-4}$ эВ. И эта энергия также намного больше энергии, которая разделяет два соседних уровня в зоне,

При каких же условиях возможен направленный поток электронов (электрический ток) в твердом теле? Ответ на этот вопрос зависит от заполнения электронами энергетической зоны.

Если число энергетических уровней в зоне больше числа электронов в ней, т.е. в зоне имеются пустые незанятые электронами уровни, то электроны могут легко возбуждаться, обеспечивая тем самым проводимость.

Если же все уровни в зоне заполнены электронами, то проводимость невозможна.

Следовательно, для возникновения электропроводности необходимо чтобы мог осуществляться процесс перехода электронов из состояний с меньшей энергией в состояния с большей энергией, т.е. наличие лежащих выше вакантных (незанятых) уровней энергии

В зависимости от степени заполнения валентной зоны электронами и ширины запрещенной зоны ΔE все вещества делятся на **проводники, полупроводники и диэлектрики.**

На рис. 7 показаны варианты заполнения энергетических зон электронами.

В случае а) слева электроны заполняют валентную зону не полностью и в ней имеются свободные состояния.

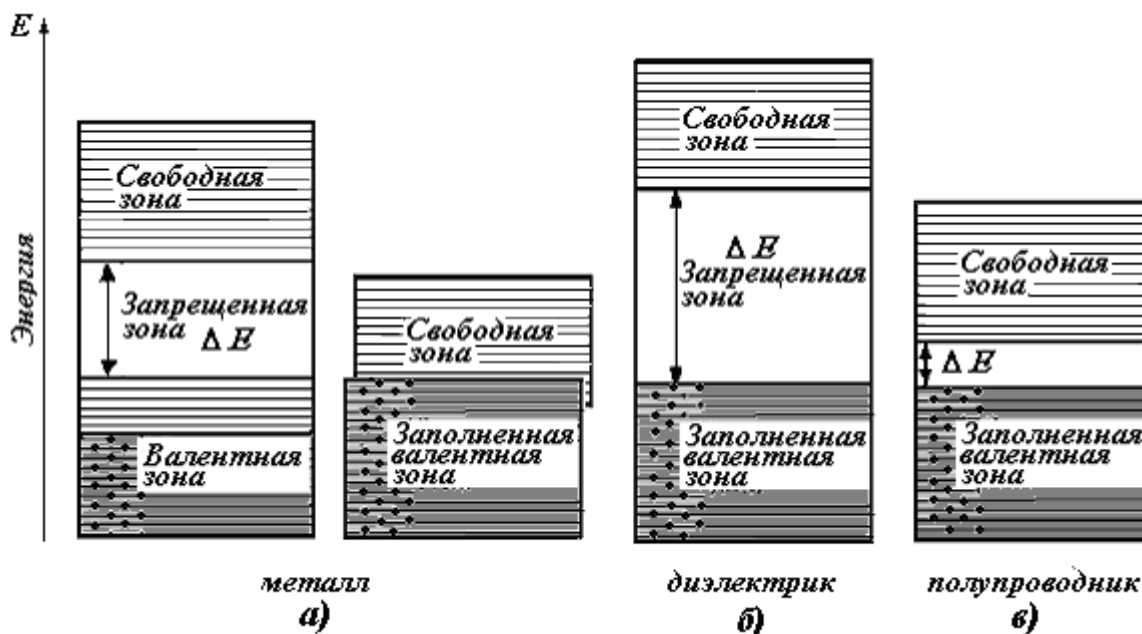


Рис. 7.

Поэтому достаточно сообщить электронам, находящимся на верхних уровнях, совсем небольшую энергию для того, чтобы перевести их на более высокие уровни. Такие электроны могут ускоряться электрическим полем и участвовать в образовании тока. Таким образом, **кристаллы с частично за-**

полненной валентной зоной хорошо проводят электрический ток. Кристалл с такой схемой энергетических уровней представляет собой **металл**. Таким образом, все металлы являются проводниками.

Частичное заполнение валентной зоны (в случае металла ее называют также **зоной проводимости**) может произойти, если кристалл образуют атомы, на последнем занятом уровне которых находится только один электрон (как в атоме Li и других щелочных металлов), или в более общем случае – нечетное число электронов (например, в атомах Al, Ga и некоторых других металлов). В этом случае N электронов проводимости заполняют попарно только половину уровней валентной зоны.

У проводников зона проводимости и валентная зона могут **перекрываться** (рис.7. а) *справа*), образуя одну зону, называемую зоной проводимости. В этом случае кристалл может быть металлом и при четном числе валентных электронов (например, Ca, Zn, Mg и др).

В случаях, изображенных на рис.7, б) и в), **уровни валентной зоны полностью заняты электронами** — зона заполнена. Это **диэлектрики и полупроводники**.

Принцип Паули запрещает электрону перейти в уже занятое другим электроном квантовое состояние. Следовательно, под действием внешнего электрического поля переходы электронов в валентной зоне происходить не могут, электрический ток не возникает. Единственная остающаяся возможность для переходов – это переходы электронов с уровней валентной зоны на уровни зоны проводимости.

Для того чтобы увеличить энергию электрона, необходимо сообщить ему количество энергии, не меньшее, чем ширина запрещенной зоны ΔE . Электрическое поле (несильное, порядка не более 100 В/см, во всяком случае, такой напряженности, при которой не происходит электрический пробой кристалла) сообщить электрону такую энергию не в состоянии. При этих условиях **электрические свойства кристалла определяются шириной запрещенной зоны ΔE** .

Если ΔE невелико – порядка нескольких десятых электронвольта – энергия теплового движения оказывается достаточной для того, чтобы перевести часть электронов в верхнюю свободную зону (случай в)). Эти электроны будут находиться в условиях, аналогичных тем, в которых находятся валентные электроны в металле. Свободная зона окажется для них **зоной проводимости**. Одновременно станет возможным переход электронов валентной зоны на ее освободившиеся верхние уровни. Такое вещество называется **электронным полупроводником**.

Полупроводниками являются материалы со структурой алмаза (тетраэдрической структурой), которые при $T = 0$ характеризуются полностью занятой электронами валентной зоной, отделенной от зоны проводимости сравнительно узкой ($\Delta E \sim 0,1 \div 3$ эВ) запрещенной зоной.

В таблице 1 приведена ширина запрещенной зоны важнейших полупроводниковых материалов.

Таблица 1.

<i>Кристалл</i>	ΔE , эВ	<i>Кристалл</i>	ΔE , эВ
Алмаз	5,47 – 7,02	Si	1,107
ZnSe	2,67	Ge	0,66 – 0,81
Se	1,79	PbTe	0,31
InP	1,35	Ag ₂ Te	0,06
GaP	2,22 – 2,78	PbS	0,41
GaAs	1,428	Sn(<i>серое</i>)	0,08
Cu ₂ O	2,02	InAs	0,356
CdTe	1,52	Mg ₂ Sn	0,14-0,18
CdSe	1,71	PbSe	0,27
CdS	2,41	Ca ₂ Pb	0,46
Ca ₂ Si	1,9	Te	0,33
BeSe	3,61	Ca ₂ Sn	0,9
AlSb	1,62-2,22	InSb	0,18
AgI	2,82	GaSb	0,73

Если же ширина запрещенной зоны велика ($\Delta E \geq 3,5$ эВ) и зоны не перекрываются (случай *б*), то кристалл является *диэлектриком (изолятором)*, тепловое движение не сможет забросить в свободную зону заметное число электронов. Чтобы перевести электрон из валентной зоны в зону проводимости требуется значительная энергия, поэтому диэлектрики ток практически не проводят. В диэлектриках переходы электронов в свободную зону могут происходить только при высоких температурах и очень сильных электрических полях.

Так, у поваренной соли NaCl $\Delta E \approx 6$ эВ, у алмаза $\Delta E \approx 5,5$ эВ, у нитрида бора $\Delta E \approx 4,6$ эВ, у Al₂O₃ $\Delta E \approx 7$ эВ и т.д.

Таким образом, зонная теория объясняет с единой точки зрения существование хороших проводников (металлов), полупроводников и изоляторов.

Контрольные вопросы

- 3.1. Почему участие электронов в электрическом токе в кристалле отражается в зонной теории переходом на более высокий уровень?
- 3.2. Зависит ли ширина энергетических зон и расстояние между уровнями энергии в пределах зоны от количества атомов в кристалле?
- 3.3. Чем объяснить хорошую электропроводность металлов?
- 3.4. Как с точки зрения зонной теории объяснить, что кристаллы с частично заполненной валентной зоной (металлы) хорошо проводят электрический ток.
- 3.5.. Квантуется ли энергия носителей тока в металлах?
- 3.6. Каково различие между металлами, диэлектриками и полупроводниками с точки зрения структуры энергетических зон?
- 3.7. Как на основе зонной теории твердых тел объяснить, что бериллий, в атомах которого внешняя электронная оболочка полностью заполнена, относится к металлам?
- 3.8. Можно ли утверждать, что проводники отличаются от диэлектриков тем, что в проводниках электроны могут двигаться, а в диэлектриках – не могут?

§ 4. Понятие о квантовой статистике электронов в твердых телах. Распределение Ферми-Дирака. Энергия Ферми и ее физический смысл для металлов

В состоянии термодинамического равновесия для данного образца кристалла при заданной температуре существует определенное распределение электронов по энергиям. В результате в кристалле устанавливается определенная концентрация свободных электронов. Вычисление этой концентрации, определяющей электрические свойства, является одной из основных задач статистики электронов в кристаллах.

Электроны в твердом теле подчиняются законам квантовой статистики Ферми-Дирака.

Эта квантовая статистика применяется к системам частиц с полуцелым спином, подчиняющихся принципу запрета Паули, согласно которому одно и то же квантовое состояние не может занимать более одной частицы.

Основой квантовой статистики Ферми-Дирака является формула

$$f(E, T) = \frac{1}{e^{(E-E_F)/kT} + 1}. \quad (1)$$

Функция $f(E, T)$ определяет вероятность того, что электрон будет находиться в квантовом состоянии с энергией E .

Здесь E – энергия уровня, вероятность которого определяется, E_F – некоторая характерная энергия, зависящая от типа кристалла, его состава и состояния, называемая *уровнем (энергией) Ферми*, T – абсолютная температура, k – постоянная Больцмана, равная $1,38 \cdot 10^{-23}$ Дж/К = $0,86 \cdot 10^{-4}$ эВ/К.

Пример 4.1. Пусть на 100 близко расположенных уровней приходится в среднем 10 электронов. Тогда вероятность заполнения этих состояний $f(E) = 10/100 = 0,1$. С учетом принципа Паули число частиц на уровне равно $2f(E)$. Поскольку на каждое состояние приходится в среднем 0,1 частица, то 0,2 можно трактовать как среднее число частиц, которые находятся на энергетическом уровне с энергией E .

Распределение Ферми-Дирака обуславливается необходимостью соблюдения двух требований. Во-первых, полная энергия должна быть минимальна, и, во-вторых, должен соблюдаться принцип Паули. Поэтому электроны начинают заполнять квантовые состояния с самого нижнего энергетического уровня, последовательно занимая квантовые состояния

Физический смысл энергии Ферми E_F можно определить из анализа распределения Ферми-Дирака вблизи абсолютного нуля.

Если тепловое возбуждение электронов кристалла отсутствует ($T = 0$ К), то возможны два случая.

а) Для всех уровней, энергия которых удовлетворяет условию $E < E_F$, имеем

$$e^{(E-E_F)/kT} = e^{-\infty} \rightarrow 0$$

Следовательно, при $E < E_F$ функция Ферми-Дирака равна единице, $f(E) = 1$, т.е. все квантовые состояния электронов с такими энергиями, начиная от уровня с наименьшей энергией, заполнены электронами.

б) Для всех уровней, энергия которых удовлетворяет условию $E > E_F$, имеем

$$e^{(E-E_F)/kT} = e^\infty \rightarrow \infty$$

Тогда при $E > E_F$ функция Ферми-Дирака равна нулю, $f(E) = 0$, что соответствует целиком незанятым уровням, в этих состояниях нет ни одного электрона.

Таким образом, при температуре $T = 0$ К уровень Ферми отделяет состояние (уровень), полностью занятое электронами ($E < E_F$), от состояний (уровней), полностью свободных ($E > E_F$).

График функции (1) распределения Ферми-Дирака приведен на рис. 8.

Если откладывать

по оси ординат вероятность $f(E)$, а по оси абсцисс – энергию уровня, то совокупность точек с ординатой единица образует прямую линию, которая скачком обрывается при некоторой энергии $E = E_F$. Очевидно, E_F для кристалла – это энергия верхнего заполненного при абсолютном нуле уровня.

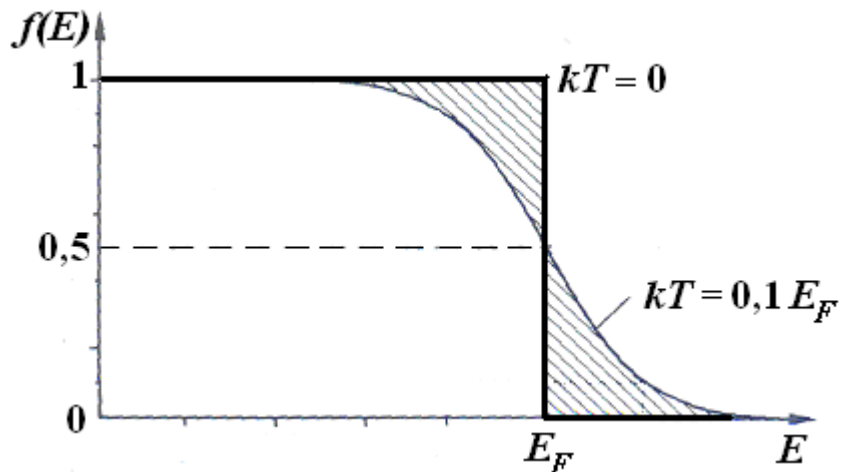


Рис. 8.

Справа от точки с координатой E_F распределение изображается линией с ординатой ноль (уровни совершенно пустые).

Тогда энергия Ферми E_F в металлах, у которых валентная зона заполнена не полностью, имеет наглядное физическое толкование – это максимальное значение энергии, которое может иметь электрон при температуре абсолютного нуля.

Для энергии Ферми в металлах квантовая теория дает значение

$$E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (2)$$

Здесь постоянная Планка $\hbar = 1,05 \cdot 10^{-34}$ Дж·с, m – масса электрона, n – число свободных электронов в единице объема (концентрация свободных электронов). Из этой формулы следует, что энергия Ферми тем больше, чем больше концентрация электронов.

Чтобы оценить концентрацию n свободных электронов в металлическом кристалле, будем полагать, что от каждого атома отщепилось по одному электрону. Тогда число электронов в единице объема n будет равно количеству атомов в единице объема. Если масса единицы объема (плотность) ρ , масса одного моля вещества M , то ρ/M – количество молей в единице объема вещества. Учитывая, что $N_A = 6,02 \cdot 10^{23}$ 1/моль – число атомов в одном моле вещества

(число Авогадро), получим для концентрации атомов (и соответственно электронов)

$$n = \rho N_A / M \quad (3)$$

Пример 4.2. Определить максимальную кинетическую энергию свободного электрона в меди.

Плотность меди $\rho = 8,9 \cdot 10^3 \text{ кг/м}^3$, молярная масса меди $M = 6,4 \cdot 10^{-2} \text{ кг/моль}$. Таким образом, из формулы (2) следует:

$$E_F = \frac{(1,05 \cdot 10^{-34})^2}{2 \cdot 9,11 \cdot 10^{-31}} \left(3 \cdot 3,14^2 \frac{8,9 \cdot 10^3 \cdot 6,02 \cdot 10^{23}}{6,4 \cdot 10^2} \right)^{2/3} = 1,12 \cdot 10^{-18} \text{ Дж} = 7,0 \text{ эВ}.$$

Пример 4.3. Определить максимальную скорость свободного электрона в кристалле меди.

При $T = 0$ заполнены все состояния с энергиями от нуля до энергии Ферми. В этом случае максимальная скорость связана с энергией Ферми соотношением

$$\frac{m_e v_{\max}^2}{2} = E_F, \quad (\text{П. 4.1})$$

т.е.
$$v_{\max} = \sqrt{2E_F / m_e}. \quad (\text{П. 4.2.})$$

Подставляя выражение для энергии Ферми, получим:

$$v_{\max} = \frac{\hbar(3\pi^2 n)^{1/3}}{m_e}. \quad (\text{П. 4.3})$$

Для меди
$$n = \frac{\rho N_A}{M} = \frac{8,9 \cdot 10^3 \cdot 6,02 \cdot 10^{23}}{63,6 \cdot 10^{-3}} \text{ м}^{-3} = 8,4 \cdot 10^{28} \text{ м}^{-3}.$$

Тогда
$$v_{\max} = 1,57 \cdot 10^6 \text{ м/с}.$$

В таблице 2 приведены значения концентрации электронов, энергии Ферми и максимальной скорости электронов для некоторых металлов.

Таблица 2

Металл	Концентрация электронов, см^{-3}	Энергия E_F , эВ	Максимальная скорость электрона, м/с
Li	$4,6 \cdot 10^{22}$	4,7	$1,3 \cdot 10^6$
Ba	$2,5 \cdot 10^{22}$	13,	1,1
Cs	$0,86 \cdot 10^{22}$	1,5	0,73
Cu	$8,5 \cdot 10^{22}$	7,0	1,56
Ag	$5,76 \cdot 10^{22}$	5,5	1,38

При $T = 0 \text{ К}$ вычисления дают для средней энергии электронов

$$\langle E \rangle = \frac{3}{5} E_F. \quad (4)$$

При отличной от нуля температуре

$$\langle E \rangle = \frac{3}{5} E_F \left(1 + \frac{5}{12} \pi^2 \left(\frac{kT}{E_F} \right)^2 + \dots \right). \quad (5)$$

Расчеты дают для максимального значения энергии электрона при $T = 0$ (уровня Ферми) в *металлах* значение $E_F \approx 5 \div 8$ эВ, а для средней энергии электронов $\langle E \rangle$ при абсолютном нуле, соответственно, $3 \div 4,8$ эВ.

Следовательно, даже при $T = 0$ электроны проводимости в *металлах* не находятся в состоянии покоя, как это следовало бы ожидать, исходя из классических представлений, а движутся с различными скоростями. Средняя скорость этого движения очень большая.

Чтобы сообщить классическому электронному газу такую энергию, его нужно нагреть до температуры порядка 400 000 К.

В *металлах* энергия Ферми электронов проводимости (5 ÷ 8 эВ) настолько больше их средней энергии теплового движения при комнатной температуре ($kT \approx 0.026$ эВ), что функция распределения Ферми-Дирака практически не зависит от температуры, и наглядный смысл понятия энергии Ферми как максимальной энергии электрона в зоне проводимости остается справедливым вплоть до температуры плавления и выше.

Для отличной от нуля температуры $T > 0$, для уровня с энергией, соответствующей уровню Ферми $E = E_F$ из формулы (1) вытекает

$$e^{(E-E_F)/kT} = e^0 \rightarrow 1,$$

откуда, $f(E_F) = 1/2$.

В результате получаем важный вывод, который касается не только металлов, но и диэлектриков и полупроводников. **Уровень Ферми совпадает с тем энергетическим уровнем, вероятность заполнения которого при любой температуре ($T > 0$) равна половине.** На таком уровне в среднем находится один электрон, в отличие от уровней с вероятностью заполнения равной единице – на таких уровнях находится по два электрона с противоположными спинами.

При нагревании кристалла происходит возбуждение некоторых электронов, находящихся вблизи уровня Ферми, за счет тепловой энергии. Энергия хаотического теплового движения частиц по порядку величины равняется, как известно, kT , где k – постоянная Больцмана.

В результате теплового возбуждения часть электронов, которые имеют энергию меньше E_F , переходят на уровни с энергией, большей E_F . Устанавливается новое распределение электронов по состояниям. При $T > 0$ функция Ферми изображается непрерывной кривой (рис. 8) и в узкой области энергий, порядка нескольких kT , в окрестности точки $E = E_F$ быстро изменяется от 1 до 0. Размытие функции Ферми тем больше, чем выше температура.

Ордината кривой характеризует среднюю по времени занятость уровня. Например, ордината, равная 0,25, означает, что уровень занят электроном $1/4$ времени.

С понятием энергии Ферми для полупроводников мы познакомимся позже.

Пример 4.4. Энергия Ферми для кристалла натрия равна 3,1 эВ. Какова вероятность найти в натрии электрон с энергией 3,2 эВ при комнатной температуре?

Проведем расчеты. $kT = 0,026$ эВ; $E - E_F = 0,1$ эВ; $(E - E_F) / kT = 4$; $f(E) = (\exp 4 + 1)^{-1} = 0,018$.

Контрольные вопросы

4.1. В чем состоит различие классического и квантового распределений электронов по энергиям?

4.2. Какой физический смысл имеет энергия Ферми в металлах? Какого порядка эта энергия для металлов при абсолютном нуле температуры? Зависит ли энергия Ферми от объема кристаллического образца?

4.3. Полагая, что в кристалле меди на каждый атом приходится по одному свободному электрону, оценить концентрацию свободных электронов в меди. Молярная масса меди $M_{Al} = 64 \cdot 10^{-3}$ кг/моль, плотность меди $\rho = 8290$ кг/м³, число Авогадро $N_A = 6,02 \cdot 10^{23}$ моль⁻¹.

4.4. Какого порядка максимальная скорость свободных электронов в металле при температуре, близкой к абсолютному нулю?

4.5. Какова средняя энергия свободных электронов в металле при абсолютном нуле температуры? До какой температуры нужно нагреть классический электронный газ, чтобы сообщить ему такую энергию?

4.6. Полагая концентрацию свободных электронов в литии $n = 4,6 \cdot 10^{22}$ см⁻³, пользуясь формулой (2), оценить максимальную и среднюю кинетическую энергию электронов при температуре $T = 0$ К

4.7. Определить энергию Ферми для серебра, считая, что его плотность $\rho = 10,5 \cdot 10^3$ кг/м³ и молярная масса $M = 107,9 \cdot 10^{-3}$ кг/моль.

4.8. Энергия Ферми для кристалла меди равна 7 эВ. Какова вероятность найти в меди электрон с энергией 7,1 эВ при комнатной температуре?

§ 5. Полупроводники. Собственная проводимость полупроводников

Полупроводниками называют широкий класс веществ, электропроводность которых (10^5 См/м $> \sigma > 10^{-7}$ См/м) занимает промежуточное положение между электропроводностью металлов $\sigma > 10^5$ (См/м = (Ом·м)⁻¹) и диэлектриков (изоляторов) $\sigma < 10^{-7}$ Ом·м (рис. 9).

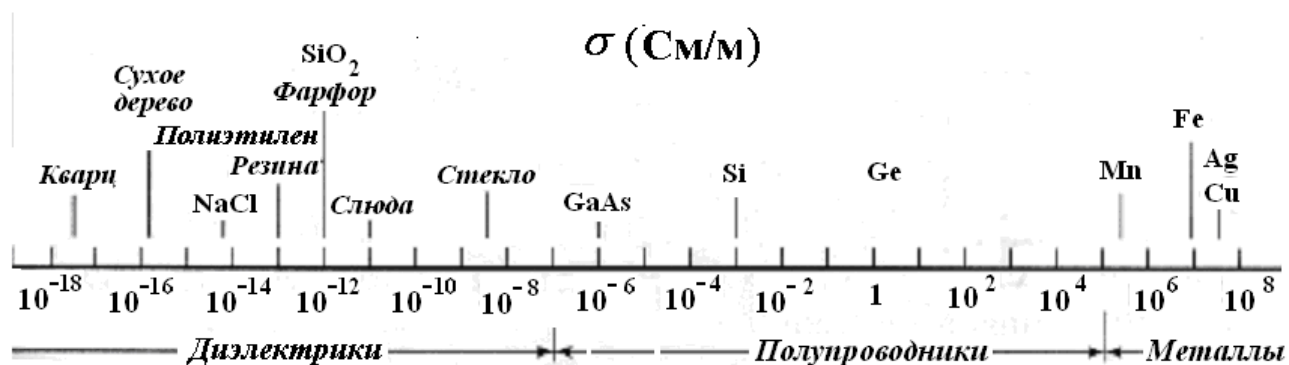


Рис. 9.

Однако характерной особенностью полупроводников является возрастание их электропроводности с ростом температуры. С повышением температуры удельная электропроводность $\sigma = 1/\rho$ резко возрастает по экспоненциальному закону

$$\sigma = \sigma_0 \exp(-\Delta E/(2kT)) \quad (6),$$

где ΔE – ширина запрещенной зоны, σ_0 – коэффициент.

Электропроводность полупроводников сильно изменяется также под действием света, электрических и магнитных полей, добавления малого количества

примесей и других факторов. Возможность в широких пределах управлять электропроводностью полупроводников изменением температуры, введением примесей и т.д. является основой их многочисленных применений.

При высоких температурах полупроводники по электропроводности приближаются к металлам, а при низких температурах являются хорошими изоляторами.

В природе полупроводники существуют *в виде элементов*, например, кремний Si, германий Ge, мышьяк As, галлий Ga, индий In, селен Se, теллур Te и др., и *химических соединений*, например сульфид кадмия CdS, сульфид свинца PbS, арсенид галлия GaAs и др.

Наиболее распространенный полупроводник **кремний** ($^{28}_{14}\text{Si}$) – в природе встречается в виде кремнезема SiO_2 (песок, кремень) – самого распространенного соединения на Земле – и силикатов (глина, шпаты, сланцы). Выделяют кремний из кремнезема – дешевого и легкодоступного сырья. Кремний в чистом виде получают прокаливанием песка с коксом ($\text{SiO}_2 + 2\text{C} = 2\text{CO} + \text{Si}$). Затем в электрических печах из расплава выращивают кристаллы кремния. Плотность кремния $2,33 \cdot 10^3 \text{ кг/м}^3$, температура плавления 1417°C .

Германий $^{74}_{32}\text{Ge}$ рассеян в земной коре. Обычно соединения германия входят в состав других минералов как примесь. Добывается из золы каменных углей и отходов производства цветных металлов. Плотность германия $5,33 \cdot 10^3 \text{ кг/м}^3$, температура плавления $937,5^\circ\text{C}$.

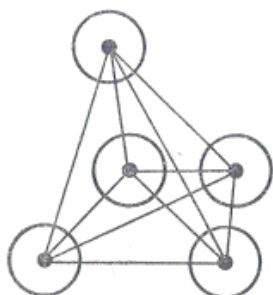


Рис. 9а.

Различают беспримесные полупроводники с собственной проводимостью (*собственные полупроводники*) и легированные полупроводники с примесной проводимостью (*примесные полупроводники*).

Собственными полупроводниками являются химически чистые (концентрация примесей чужеродных атомов порядка 0,001 %) полупроводники. В них концентрация свободных носителей тока определяется только температурой и собственной, присущей только данному полупроводнику ши-

риной запрещенной зоны ΔE .

Рассмотрим для конкретности практически наиболее важный полупроводник – **кремний**. Атомы кремния принадлежат к IV группе периодической системы элементов Менделеева. Они имеют четыре электрона в наполовину заполненной внешней оболочке. В твердом состоянии атомы Si образуют кристаллическую решетку, в которой каждый атом связан ковалентными (парно-электронными) связями с четырьмя равноотстоящими от него соседними атомами (тетрагональная структура алмаза, рис. 9а).

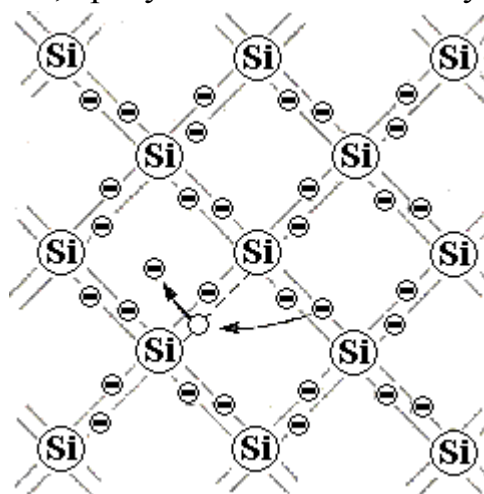


Рис.10.

Взаимное расположение атомов кремния можно представить в виде двумерной структуры, изображенной на рис. 10. Кружки с надписью «Si» обозначают атомы кремния, кружки со знаком «-» обозначают валентные электроны, двойные линии – ковалентные связи.

Вблизи абсолютного нуля температуры в кристалле кремния все электроны заняты в образовании связей. Такой кристалл электрического тока не проводит. При достаточно высокой температуре тепловое движение может разорвать отдельные пары, освободив один электрон.

Тогда в кристалле возникнут *свободные электроны* - электроны проводимости, способные направленно двигаться под действием электрического поля, образуя электрический ток.

Покинутое электроном место перестает быть нейтральным, в его окрестности возникает избыточный положительный заряд $+e$. Такое «пустое» место с отсутствующим электроном связи получило название *«дырки»* – положительной квазичастицы (она изображена на рис. 10 белым кружком). Дырка ведет себя как положительно заряженная частица с зарядом, по величине равным заряду электрона. На место дырки может перескочить электрон одной из соседних пар. В результате дырка перемещается в противоположную сторону и начинает также хаотически перемещаться по кристаллу, как и освобожденный электрон. Если к кристаллу приложить внешнее поле, на хаотическое движение электронов накладывается направленное движение против поля, а на хаотическое движение дырки – вдоль поля. Оба движения приводят к переносу заряда вдоль кристалла, т.е. к возникновению тока.

На рис. 11 изображена схема последовательного (сверху вниз) положения ряда атомов в кристалле. Кружком с плюсом обозначен атом кремния, в окрестности которого возникает положительный заряд вследствие потери электрона. Стрелкой показан переход электрона. Реальное движение электронов против внешнего электрического поля (верх рис. 11) эквивалентно движению дырки в направлении поля (нижняя часть рис. 11). Дырка ведет себя подобно положительно заряженной частице, которая может участвовать в электрической проводимости.

Следует подчеркнуть, что электрон - это реальная *частица*, а дырка - это фиктивная частица (*квазичастица*). Электрон можно удалить из полупроводника наружу (например, с помощью фотоэффекта), дырка же может существовать только внутри полупроводника.

Если свободный электрон встретится с дыркой, они *рекомбинируют* (соединяются). Это означает, что электрон нейтрализует избыточный положительный заряд, имеющийся в окрестности дырки, и теряет свободу передвижения до тех пор, пока снова не получит от кристаллической решетки энергию, достаточную для своего освобождения. *Рекомбинация* приводит к *одновременному*

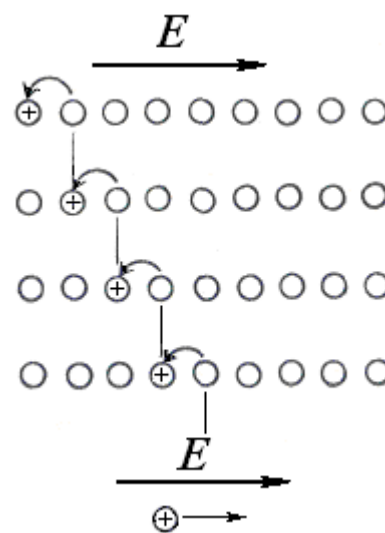


Рис. 11

исчезновению свободного электрона и дырки. При этом избыток энергии может выделяться в виде излучения. Возможна также безызлучательная рекомбинация, при которой энергия передается кристаллической решетке или подвижным носителям заряда.

В полупроводнике идут одновременно два процесса: *рождение* попарно свободных электронов и дырок и *рекомбинация*, приводящая к попарному исчезновению электронов и дырок. Каждой температуре соответствует определенная равновесная концентрация электронов и дырок.

Таким образом, **собственная электропроводность обуславливается носителями зарядов двух знаков – отрицательными электронами и положительными дырками.**

Проводимость собственных полупроводников, **обусловленная электронами**, называется **электронной проводимостью** или **проводимостью n-типа** (от лат. *negative* – отрицательный).

Проводимость собственных полупроводников, **обусловленная квазичастицами – дырками**, называется **дырочной проводимостью** или **проводимостью p-типа** (от лат. *positive* – положительный).

Подводя итоги можно сказать, что **проводимость чистых полупроводников является проводимостью смешанного типа – электронно-дырочной.** Она присуща всем без исключения полупроводникам при достаточно высокой температуре. Собственный полупроводник содержит электроны и дырки в одинаковом количестве.

Контрольные вопросы

5.1. Что такое собственная проводимость полупроводника? Как ведет себя электропроводность полупроводника с ростом температуры?

5.2. В чем смысл понятия “дырка” как носителя тока в полупроводнике? Как образуются дырки в полупроводнике? Каковы их свойства? Существуют ли дырки вне полупроводника? Совпадают ли зоны проводимости для электронов и дырок у полупроводников?

5.3. Чему равняется наименьшая энергия ϵ_{\min} образование пары электрон — дырка в собственном полупроводнике, проводимость которого возрастает в $\mu = 2,01$ раз с повышением температуры от $T_1 = 300$ К до $T_2 = 310$ К?

5.4. Чем отличаются механизмы электропроводности металлов (по классической электронной теории) и полупроводников?

5.5. Одинаково ли ведет себя удельное сопротивление проводников и полупроводников с ростом температуры?

5.6. Сравнить электропроводность германия при $T_1 = 250$ К и $T_2 = 350$ К. Необходимые данные взять из таблицы 1.

5.7. Почему полупроводники могут использоваться как элементы чувствительных электрических термометров?

§ 6. Собственная проводимость полупроводников в терминах зонной теории. Энергия Ферми для полупроводников

Рассмотрим теперь электронные явления в собственных полупроводниках с точки зрения зонной теории (рис. 12).

В кристаллической решетке кремния или германия четыре внешних электрона атомов связываются ковалентно с четырьмя соседними атомами, в результате чего образуется полностью заполненная валентная зона E_V , а в свободной зоне E_C электроны отсутствуют. Для того, чтобы в собственном полупроводнике возникла электронно-дырочная пара, необходимо затратить энергию ΔE или, другими словами, преодолеть потенциальный барьер высотой ΔE .

Электрическое поле при $T = 0$ К не может перебросить электроны из валентной зоны в зону проводимости, т.е. при $T = 0$ К собственные полупроводники являются диэлектриками.

Для осуществления перехода из валентной зоны в зону проводимости электрон должен получить энергию, достаточную для преодоления запрещенной зоны.

Эту энергию при повышении температуры электрон может получить от ионов решетки, совершающих тепловые колебания. Таким образом, преодоление запрещенной зоны электроном происходит обычно за счет тепловой энергии решетки.

Часть электронов с верхних уровней вблизи потолка валентной зоны переходит под действием теплового возбуждения на нижние уровни вблизи дна свободной зоны (рис. 12). Такие квантовые переходы между уровнями энергии обозначают на диаграммах вертикальными стрелками, соединяющими соответствующие пары уровней энергии.

Теперь при наложении на кристалл электрического поля электроны, попавшие в зону E_C , могут перемещаться на более высокие пустые уровни, т.е. увеличивать свою энергию и участвовать в образовании электрического тока. Поэтому свободная зона получила название *зоны проводимости* E_C .

Когда в зоне проводимости E_C появляется некоторое число электронов, занимающих уровни вблизи дна зоны; то одновременно в валентной зоне освобождается такое же число мест на верхних уровнях.

Электроны валентной зоны E_V также могут теперь изменять свою скорость (энергию) под действием внешнего поля. Во внешнем электрическом поле на освободившееся место может переместиться электрон с нижнего уровня, а вакансия появится в том месте, откуда ушел электрон, и т.д.

Такие свободные от электронов места на уровнях заполненной при абсолютном нуле валентной зоны также называют *дырками*.

На рис. 12 в модели зонной теории показаны квантовые переходы, приводящие к образованию пары носителей заряда – электрона и дырки.

Электрон проводимости обозначен черной точкой, а дырка – белой.

При этом приведены обозначения, используемые в международной практике: E_C – Conduction band – зона проводимости; E_F – Fermi level – уровень Ферми; E_V – Valence band – валентная зона; ΔE (иногда E_g) – Bandgap – промежуток между зонами, щель (запрещенная зона).



Рис.12.

С помощью энергетической диаграммы можно показать также процесс рекомбинации электрона и дырки в собственном полупроводнике. Этот процесс показан на рис.12 штриховой стрелкой. При рекомбинации выделяется та же энергия, которая была затрачена на создание пары носителей. Иногда рекомбинация сопровождается рождением фотона с энергией ΔE , чаще энергия передается решетке кристалла, нагревая ее.

Для диэлектриков и полупроводников при $T = 0$ К характерно полное заполнение электронами энергетических уровней валентной зоны и отсутствие электронов в зоне проводимости. Если сопоставить этот факт с основным свойством уровня Ферми о том, что вероятность его заполнения равна $\frac{1}{2}$ при любых температурах, то становится ясным, что **уровень Ферми** в диэлектриках и собственных **полупроводниках** не может находиться ни в валентной зоне, ни в зоне проводимости. Остается положить, что он **расположен в запрещенной зоне**.

Соответствующий расчет положения уровня Ферми для собственных полупроводников дает:

$$E_F = \frac{E_C + E_V}{2} + \frac{3}{4} kT \ln \frac{m_p^*}{m_n^*} . \quad (7)$$

Здесь E_C - энергия, соответствующая «дну» зоны проводимости, E_V - энергия, соответствующая «потолку» валентной зоны, $(E_C + E_V)/2$ - энергия середины запрещенной зоны, m_n^* и m_p^* - так называемые эффективные массы электронов и дырок. Истинная масса электрона заменяется эффективной массой, чтобы описывать движение электронов и дырок в периодическом поле кристаллической решетки, полагая их свободными.

Поскольку kT даже для достаточно высокой температуры является малой величиной по сравнению с шириной запрещенной зоны, а логарифм - дробная величина, то второе слагаемое в формуле (7) дает малый вклад в значение энергии Ферми. Соответственно, **в диэлектриках и собственных полупроводниках уровень Ферми лежит приблизительно посередине запрещенной зоны** (при абсолютном нуле - точно).

Поскольку энергия Ферми приходится на запрещенную зону, электрон не может обладать такой энергией, т.е. энергия Ферми не соответствует энергии какого-либо реального электрона в полупроводнике.

Однако это обстоятельство не уменьшает значения энергии Ферми для описания статистических свойств электронов в полупроводниках. Так при переходе электрона с верхнего уровня валентной зоны на нижний уровень зоны проводимости затрачивается энергия ΔE . В собственном полупроводнике при появлении электрона в зоне проводимости в валентной зоне обязательно возникает дырка, и поэтому можно считать, что затраченная на весь процесс энергия делится на две равные части - половина тратится на образование дырки, а половина - на перевод электрона в зону проводимости. Иными словами, начало отсчета энергии для каждого из этих процессов должно находиться в центре за-

прещенной зоны, т.е. энергия Ферми представляет такую энергию, от которой происходит возбуждение электронов и дырок (рис. 12).

Если совместить график функции Ферми-Дирака с энергетической диаграммой полупроводника (рис. 13), можно получить дополнительные сведения о заполнении электронами энергетических уровней.

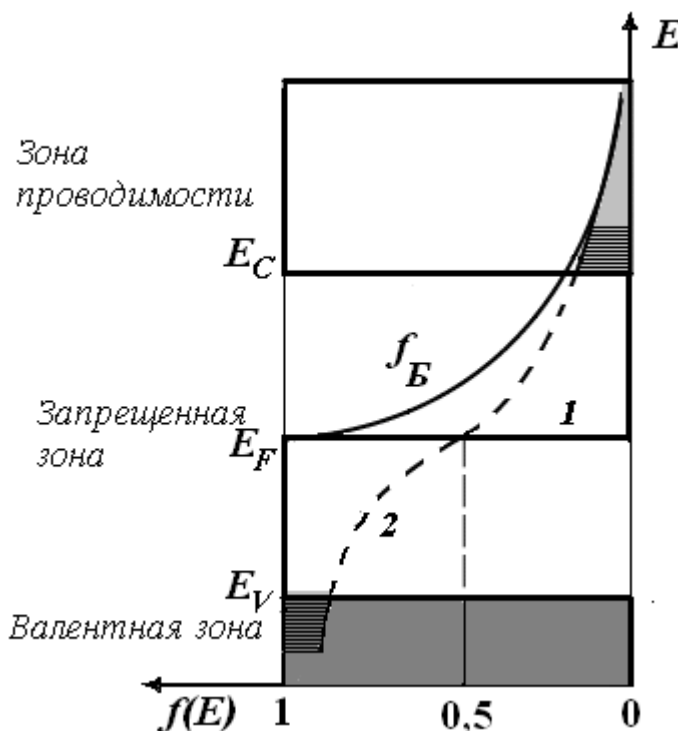


Рис. 13.

казано на рис. 13 ломаной линией 1.

Как уже отмечалось, вероятность заполнения электронами любого уровня с энергией E при любой температуре, как в валентной зоне, так и в зоне проводимости, дается формулой Ферми-Дирака (1).

Поэтому при комнатной температуре часть валентных электронов сможет перейти в зону проводимости. В зоне проводимости появляется некоторое число электронов, занимающих уровни вблизи дна зоны E_C . Одновременно в валентной зоне E_V освобождается такое же число мест на верхних уровнях.

Поэтому вероятность заполнения электронами валентной зоны оказывается уже несколько меньшей единицы, а вероятность заполнения электронами зоны проводимости - большей нуля. Вместо ступенчатой кривой 1 при $T > 0$ имеем непрерывную кривую 2, изменяющуюся от 1 до 0.

Уровни зоны проводимости лежат на «хвосте» функции $f(E)$. Ординаты кривой 2 на участках, отмеченных горизонтальной штриховкой на рис. 13, соответствуют вероятностям нахождения электронов на энергетических уровнях зоны проводимости и дырок в валентной зоне.

Пример 6.1. Оценим вероятность перехода электрона в зону проводимости при тепловом возбуждении для типичных диэлектриков и полупроводников при комнатной температуре. Для этого воспользуемся функцией распределения Ферми-Дирака. (1).

Здесь по оси абсцисс влево отложена вероятность $f(E)$ заполнения электронами соответствующих энергетических уровней, по оси ординат вверх отложена энергия E электрона. Минимальное значение энергии зоны проводимости обозначено E_C , максимальное значение энергии валентной зоны — E_V .

Так как электроны стремятся занять уровни с наименьшей энергией, то при $T = 0$ К в полупроводнике имеется резкая граница между заполненными и пустыми уровнями энергии. При температуре абсолютного нуля ($T = 0$ К) все валентные уровни заполнены с вероятностью, равной $f(E) = 1$, а вероятность заполнения любого уровня зоны проводимости равна нулю. Это по-

Ширина запрещенной зоны типичного диэлектрика – алмаза – $\Delta E \approx 5,5$ эВ. Как уже отмечалось, уровень Ферми в собственных полупроводниках и диэлектриках располагается в середине запрещенной зоны, т.е. $E_F = E_C - \Delta E/2$ и $E_C - E_F = \Delta E/2$.

При комнатной температуре ($kT \approx 0,026$ эВ) вероятность заполнения уровня с энергией E_c вблизи дна зоны проводимости равна

$$f(E_c) = \frac{1}{e^{[E_c - (E_c - \Delta E/2)]/kT} + 1} = \frac{1}{e^{\Delta E/2kT} + 1} = \frac{1}{e^{5,5/0,05} + 1} \approx e^{-5,5/0,05} \approx e^{-100} \approx 5 \cdot 10^{-45}. \quad (\text{П. 6.1.})$$

Смысл полученной вероятности можно трактовать так: только один электрон из примерно 10^{45} электронов в валентной зоне будет иметь энергию для перехода в зону проводимости при комнатной температуре. Учитывая, что в 1 см^3 различных кристаллов содержится порядка 10^{22} атомов (и столько же валентных электронов) можно сделать вывод о том, что при комнатной температуре практически не будет электронов в зоне проводимости, а алмаз является идеальным изолятором.

Теперь возьмем типичный полупроводник – германий. Ширина запрещенной зоны для него составляет 0,67 эВ. Пользуясь формулой (1), найдем

$$f(E_c) = \frac{1}{e^{0,67/0,05} + 1} \approx e^{-0,67/0,05} \approx e^{-13} \approx 10^{-6} \quad (\text{П. 6.2.})$$

В этом случае приблизительно один валентный электрон из миллиона может при возбуждении перейти в зону проводимости на ее дно. Учитывая снова, что в 1 см^3 различных кристаллов содержится порядка 10^{22} атомов (и столько же валентных электронов), получим концентрацию электронов в полупроводнике порядка 10^{16} см^{-3} . Число электронов в зоне проводимости окажется вполне достаточным, чтобы обеспечить электропроводность полупроводника.

Это число, конечно, будет гораздо меньше, чем в случае металлов (проводников), для которых величина $f(E_c)$ порядка единицы, а плотность электронного газа $\sim 10^{22} \text{ см}^{-3}$.

Если теперь провести аналогичный расчет для температуры $T \sim 30 \text{ К}$, то при такой низкой температуре и германий, и кремний становятся диэлектриками.

Из приведенных численных примеров видно, что при ширине запрещенной зоны ΔE полупроводника около $0,5 \div 1,2$ эВ значения $E - E_F$ в показателе экспоненты формулы (1) будут составлять, по крайней мере, несколько десятых эВ. Значения же kT , при комнатной температуре равное 0,026 эВ, даже при температуре $T = 1000 \text{ К}$ составляет всего 0,086 эВ. Поэтому

$$\exp((E - E_F)/(kT)) \gg 1,$$

и единицей в знаменателе формулы (1) можно пренебречь. Формула принимает вид

$$f(E) \approx e^{-(E - E_F)/kT} = e^{-\Delta E/(2kT)} \quad (8)$$

Эта формула выражает распределение по энергиям частиц, подчиняющихся законам классической механики, называемое распределением Максвелла-Больцмана. График функции распределения Больцмана f_B приведен для сравнения на рис. 13.

Таким образом, *в широком интервале температур от комнатной и выше электроны в полупроводниках подчиняются классической статистике Максвелла-Больцмана.*

Положение уровня Ферми относительно краев зон в сильной степени влияет также на концентрацию подвижных электронов. Не имея возможности рассматривать теорию, приведем данные расчетов.

В таблице 3 для некоторых полупроводников приведены значения ширины запрещенной зоны ΔE и собственной концентрации электронов при комнатной температуре.

Таблица 3.

Полупроводник	InSb	Ge	Si	InP	GaAs	GaP	SiC
ΔE , эВ	0,18	0,67	1,14	1,35	1,43	2,26	2,4 – 3,2
Собственная концентрация, см^{-3}	$1,3 \cdot 10^{16}$	$2,4 \cdot 10^{13}$	$1,1 \cdot 10^{10}$	$5,7 \cdot 10^7$	$1,4 \cdot 10^7$	0,8	0,1 – $2 \cdot 10^{-8}$

Из таблицы 3 следует, что при изменении ширины запрещенной зоны в диапазоне от 0,18 эВ для InSb до 2,26 эВ для фосфида галлия, например, собственная концентрация при комнатной температуре изменяется от значения 10^{16} см^{-3} до $0,8 \text{ см}^{-3}$.

Из формулы (6) видно, что вероятность преодоления энергетического барьера шириной ΔE определяется экспоненциальным множителем. Количество электронов, перешедших в зону проводимости, а следовательно и количество образовавшихся дырок будет пропорционально этой вероятности.

Можно полагать, следовательно, что и электропроводность σ собственного полупроводника увеличивается с температурой по экспоненциальному закону

$$\sigma = \sigma_0 e^{-\Delta E/(2kT)}$$

Здесь σ_0 - некоторый коэффициент.

Важно еще раз подчеркнуть, что для полупроводников характерной особенностью является сильная зависимость электропроводности от температуры.

С точки зрения зонной теории эта особенность объясняется так: при повышении температуры растет число электронов, которые вследствие теплового движения могут переходить в зону проводимости и, следовательно, принимать участие в переносе электрического заряда (образовании электрического тока).

При возрастании, например, температуры на 10 К относительно комнатной (т.е. всего на 3 %), вероятность перехода электронов в зону проводимости увеличивается приблизительно на 30 %. Так же сильно влияет на увеличение проводимости полупроводников уменьшение ширины запрещенной зоны.

Пример 6.3. Удельная электропроводность кремниевого образца увеличилась в 4,24 раза при нагревании от температуры $t_1 = 0^\circ\text{C}$ до температуры $t_2 = 18^\circ\text{C}$. Определить ширину запрещенной зоны кремния.

Удельная электропроводность собственных полупроводников зависит от температуры по закону

$$\sigma = \sigma_0 e^{-\Delta E / (2kT)},$$

где σ_0 – постоянная, характерная для данного полупроводника, ΔE – ширина запрещенной зоны.

Тогда

$$\frac{\sigma_2}{\sigma_1} = \frac{\exp(-\Delta E / (2kT_2))}{\exp(-\Delta E / (2kT_1))} = \exp\left(\frac{\Delta E}{2k} \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right). \quad (\text{П. 6.1.})$$

Логарифмируя, получим

$$\ln \frac{\sigma_2}{\sigma_1} = \frac{\Delta E}{2k} \left(\frac{1}{T_1} - \frac{1}{T_2}\right). \quad (\text{П. 6.2.})$$

Откуда искомая ширина запрещенной зоны

$$\Delta E = \frac{2kT_1T_2 \ln(\sigma_2 / \sigma_1)}{T_2 - T_1}. \quad (\text{П. 6.3.})$$

После подстановки чисел

$$\Delta E = 1,76 \cdot 10^{-19} \text{ Дж} = 1,1 \text{ эВ.}$$

Сильная зависимость проводимости полупроводников от температуры используется на практике для создания очень чувствительных термометров. Такие приборы получили название **термисторов (терморезисторов)**.

Процессу рекомбинации с точки зрения зонной теории соответствует обратный переход электрона из зоны проводимости на один из свободных уровней валентной зоны. При этом электрон отдает энергию решетке или испускает квант электромагнитного излучения.

Контрольные вопросы

- 6.1. Оцените энергию теплового движения электрона в полупроводнике при температуре -240°C и 100°C .
- 6.2. Какими электрическими свойствами обладают полупроводники при низких и высоких температурах?
- 6.3. Где расположен уровень Ферми в собственном полупроводнике?
- 6.4. В чем состоит процесс рекомбинации электронов и дырок в полупроводниках? Каков электронный механизм явления и как оно описывается в модели зонной теории?
- 6.5. Как может выделяться избыток энергии при рекомбинации электронов и дырок?
- 6.6. При каких условиях вероятность того, что электрон будет находиться в квантовом состоянии с энергией E будет определяться распределением Максвелла-Больцмана классической статистики?
- 6.7. К каким выводам приводит положение уровня Ферми в собственных полупроводниках внутри запрещенной зоны?
- 6.8. Удельная электропроводность германиевого образца увеличилась в 2,45 раза при нагревании от температуры $t_1 = 0^\circ\text{C}$ до температуры $t_2 = 17^\circ\text{C}$. По этим данным определить ширину запрещенной зоны германия.

§ 7. Примесная проводимость полупроводников

Примесная проводимость возникает, если некоторые атомы данного полупроводника заменить в узлах кристаллической решетки атомами, валентность которых отличается на единицу от валентности основных атомов. Такое контролируемое введение в полупроводник небольших количеств материалов с известными свойствами называют *легированием*. Оно достигается сплавлением, диффузией, ионной имплантацией и другими способами.

Пусть *валентность примеси на единицу больше*, чем валентность основных атомов. Такой примесью в кремнии могут служить, например, фосфор (P), мышьяк (As) или сурьма (Sb).

На рис. 14 условно изображена двумерная решетка кремния, в которой один из атомов 4-валентного кремния замещен атомом 5-валентного фосфора.

Атом фосфора содержит в наружной оболочке лишний по сравнению с кремнием электрон. Четыре из пяти валентных электронов атома фосфора приходятся на связь этого атома с решеткой кремния. Следовательно, пятый валентный электрон оказывается как бы лишним и легко отщепляется от атома за счет энергии теплового движения, образуя странствующий свободный электрон – электрон проводимости. Энергия ΔE_i , которая нужна, чтобы разорвать связь пятого электрона с атомом фосфора и превратить его в свободный электрон составляет примерно 0,01 эВ. Это гораздо меньше, чем энергия $\Delta E = 1,1$ эВ, которую нужно затратить, чтобы разорвать связь между атомами кремния и образовать электронно-дырочную пару. Поэтому практически все пентавалентные примесные атомы ионизируются и отдают «лишний» электрон.

В то же время тепловое движение ионов кристаллической решетки может разорвать одну из ковалентных связей. При этом рождается пара носителей тока – электрон и дырка.

Пользуясь таблицей 3, сравним собственную концентрацию электронов и концентрацию электронов, созданную атомами примеси.

Химически чистым называют вещество, в котором содержится один посторонний атом на 1 000 собственных (т.е. концентрация примесей $\sim 0,1\%$). Напомним, что в одном кубическом сантиметре кристаллического вещества содержится $\sim 10^{22}$ атомов. Тогда даже в случае концентрации примесей 0,001% в каждом кубическом сантиметре вещества будет содержаться порядка 10^{17} посторонних атомов. Поскольку примесные атомы сравнительно легко отдают электроны, то при комнатной температуре чистый германий будет содержать созданных примесью электронов примерно в $4 \cdot 10^3$ раз больше, чем собственных, кремний – в 10^7 раз больше, фосфид галлия – в 10^{17} раз больше.

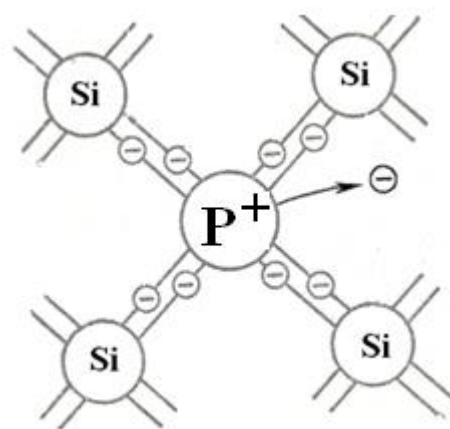


Рис. 14.

Таким образом, число рождающихся под действием теплового движения электронов из электронно-дырочных пар очень мало по сравнению с числом электронов от примесных атомов. Другими словами, *примесная электронная проводимость оказывается доминирующей по сравнению собственной* (электронной и дырочной).

В узле, занимаемом атомом фосфора, образуется результирующий положительный заряд. Образование свободного электрона при этом не сопровождается нарушением ковалентных связей, т.е. образованием дырки. Положительный заряд связан с ионом фосфора и перемещаться по решетке не может.

Благодаря этому заряду атом фосфора может захватить приблизившийся к нему электрон, но связь захваченного электрона с атомом будет непрочной и легко нарушается вновь за счет тепловых колебаний решетки.

Атомы примеси, поставляющие электроны проводимости, называются **донорами** (т.е. отдающими электрон). Полупроводник, в который введена донорная примесь и в котором как носители заряда преобладают электроны, называют **электронным** или **полупроводником n-типа**.

Рождающиеся за счет тепловой генерации дырки в электронном полупроводнике принято называть *неосновными носителями* тока, а электроны в данном случае называются *основными носителями*.

Например, в кремнии *n-типа* при комнатной температуре и концентрации электронов 10^{15} см^{-3} концентрация дырок составляет всего $\sim 10^5 \text{ см}^{-3}$.

С точки зрения зонной теории процесс образования электронов в полупроводнике, содержащем донорную примесь, представлен на рис.15 (масштаб не соблюден). Атомы примеси вызывают на энергетической схеме появление дополнительных уровней, т.н. *донорных уровней*. Донорные уровни расположены в запрещенной зоне на малом расстоянии ΔE_D от дна зоны проводимости (несколько десятых или сотых электронвольта от нижнего края свободной зоны, см. таблицу 4).



Рис. 15.

Поскольку энергетическое расстояние ΔE_D мало, энергия теплового движения даже при обычных температурах оказывается достаточной для того, чтобы перевести электрон с донорного уровня в зону проводимости, и электроны уходят с них в зону проводимости.

Обратный переход электрона из зоны проводимости на донорный уровень соответствует захвату свободного электрона атомом примеси.

Поскольку уровень Ферми соответствует средней энергии всех возбужденных электронов и дырок в кристалле, то в этом случае он лежит посередине между донорным уровнем и дном зоны проводимости (рис. 15).

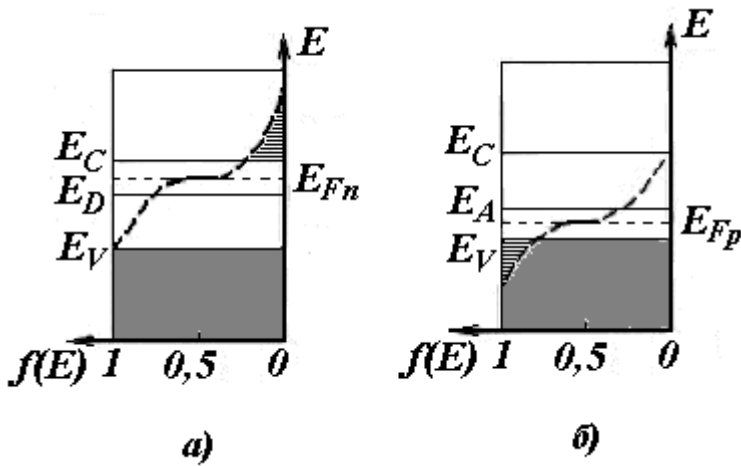


Рис. 16.

Энергетическая диаграмма полупроводника *n*-типа показана на рис. 16, а.

В отличие от собственного полупроводника у полупроводника *n*-типа кривая распределения Ферми — Дирака и уровень Ферми смещаются вверх. Это объясняется тем, что пятивалентные атомы примеси имеют энергетические уровни валентных электронов вблизи зоны проводимости собственного полупроводника.

Величина

$$\Delta E_D = E_C - E_D$$

мала (около 0,05 эВ), поэтому даже при комнатной температуре почти все электроны с примесного уровня переходят в зону проводимости.

Предположим теперь, что в решетку кремния (Si) введен примесный атом, *валентность* которого *на единицу меньше*, чем валентность основных атомов. Это могут быть атомы бора (B), алюминия (Al), галлия (Ga) или индия (In) с тремя валентными электронами (рис. 17).

Трех валентных электронов, например, атома бора недостаточно для образования связей со всеми четырьмя соседями. Поэтому одна из связей окажется неуплотненной и будет представлять собой место, способное захватить электрон.

Чтобы электрон от соседнего атома кремния перешел к бору, необходимо ему сообщить энергию — энергию активации, которая для атома B в Si составляет ~ 0,045 эВ. Если тепловые колебания решетки сообщат электрону эту энергию, то электрон с соседнего атома перейдет к бору.

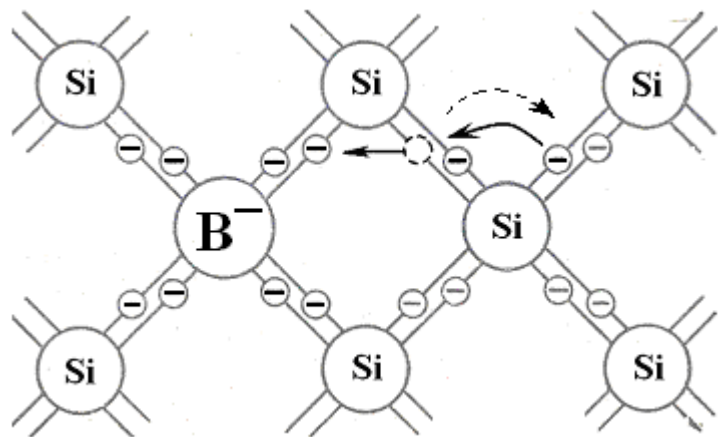


Рис.17.

При этом образуется пустая связь у атома кремния и возникает дырка — свободный носитель положительного заряда.

На рис. 17 сплошными стрелками показано реальное движение электрона, а штриховой стрелкой — движение дырки.

В то же время избыточный отрицательный заряд, который возникает вблизи атома бора, присоединившего электрон (отрицательного иона бора), связан с ионом бора и по решетке перемещаться не может.

Атомы примеси, вызывающие возникновение дырок, называются **акцепторами** (т.е. присоединяющими электрон).

Полупроводник с акцепторной примесью обладает **дырочной проводимостью** или является **полупроводником *p*-типа**

Так же, как и в случае полупроводников *n*-типа, за счет разрыва ковалентной связи могут рождаться пары носителей тока. Поэтому в полупроводнике *p*-типа появляются также свободные электроны. Дырки, концентрация которых в данном случае больше, являются **основными носителями** тока, а электроны – **неосновными**.

На рис. 18 показано, как акцепторы влияют на зонную структуру. Созданные акцепторами локальные уровни E_A располагаются в запрещенной зоне ближе к верхнему краю валентной зоны (несколько десятых или сотых электронвольта от верхней границы валентной зоны). Эти уровни свободны и заполняются электронами из валентной зоны.



Рис. 18.

Образованию дырки соответствует переход электрона из валентной зоны на акцепторный уровень E_A .

Обратный переход соответствует разрыву одной из четырех ковалентных связей атома примеси с его соседями и рекомбинации образовавшегося при этом электрона и дырки.

Уровень Ферми в полупроводнике *p*-типа лежит между потолком валентной зоны и акцепторными уровнями.

У полупроводника *p*-типа кривая распределения Ферми — Дирака и уровень Ферми смещаются вниз (см. рис. 16, б).

Величина $\Delta E_A = E_A - E_V$ также мала (около 0,05 эВ), поэтому электроны валентной зоны легко переходят на примесный уровень. При этом в валентной зоне появляется большое число дырок. Они заполняются другими электронами валентной зоны, что сопровождается образованием новых дырок. Следовательно, появляется возможность перемещения электронов в валентной зоне и повышения электропроводности, называемой дырочной.

Поскольку энергия, необходимая для образования дырки, мала (акцепторные уровни лежат в запрещенной зоне очень близко к потолку валентной зоны), при комнатной температуре на каждый примесный атом приходится по одной дырке. Собственная (электронная и дырочная) проводимость при этом значительно меньше, чем примесная, так что **доминирующей проводимостью является примесная дырочная проводимость**.

С ростом температуры увеличивается **собственная проводимость** полупроводника, которая при достаточно **высокой температуре становится преобладающей**. Это происходит потому, что относительная концентрация примесных атомов очень мала ($\sim 10^{-6}$) и при некоторой температуре наступает так называемое примесное истощение – все атомы примеси оказываются ионизованными.

Таким образом, проводимость примесного полупроводника складывается из примесной и собственной проводимостей. При низких температурах преобладает примесная, а при высоких – собственная проводимость.

В качестве донорных примесей для германия и кремния используют элементы: P, As, Sb. В качестве акцепторных примесей — B, Al, Ga, In. *Минимальная энергия возбуждения примесного атома, необходимая для создания примесной электропроводности полупроводника, называется энергией активации примесей (ΔE_D и ΔE_A на рисунках 15 и 18)*. Значения энергии активации примесей для германия и кремния приведены в табл. XX. *Энергия активации доноров — это энергия, необходимая для перехода электрона с донорного уровня в зону проводимости, а энергия активации акцептора — энергия, необходимая для перехода электрона из валентной зоны, на акцепторный уровень.* Для германия энергия активации всех примесей примерно одинакова, порядка 0,01 эВ, поэтому в германии атомы примеси оказываются ионизированными при очень низких температурах. В кремнии энергия ионизации примесей несколько выше, но и в кремнии примеси при комнатных температурах уже ионизированы.

Таблица 4

Элемент	Ge	Si	GaAs	Вид примесей
B	0,01	0,045		Акцептор
Al	0,011	0,068		Акцептор
Ga	0,011	0,07		Акцептор
In	0,011	0,16		Акцептор
P	0,0127	0,045		Донор
As	0,014	0,05		Донор
Sb	0,01	0,04		Донор
Te	0,1		0,01	Донор
Zn		0,55	0,01	Акцептор

Контрольные вопросы

7.1. Для чего в полупроводник добавляют примеси? Каким является главное свойство примесного полупроводника?

7.2.. В кристалл германия введены примеси. Электропроводность какого типа (*p*- или *n*-) будет иметь полученный кристалл, если введены примеси: а) индия; б) сурьмы?

7.3. Как изменяется вклад собственной и примесной электропроводности в полную электропроводность полупроводника с повышением температуры?

7.4. Какие элементы добавляют в полупроводник для повышения концентрации электронов? Дырок?

7.5. Возможны ли переходы электронов с донорных уровней на акцепторные?

7.6. Где находится уровень Ферми в примесном полупроводнике донорного типа? Акцепторного типа?

§ 8. Контакт p - и n -полупроводников. Свойства p - n -перехода

Контакт двух примесных полупроводников с проводимостью разных типов называют *электронно-дырочным переходом* или p – n -переходом. Это основная структурная ячейка полупроводниковых приборов. На свойствах таких переходов основан принцип действия многочисленных полупроводниковых приборов, которые широко применяются в вычислительной, электро- и радиотехнике и электронике.

Рассмотрим в общих чертах физические процессы, которые происходят в p – n -переходе.

Пусть два участка полупроводника с проводимостями разного типа разделяет плоская граница (рис. 19): слева от нее расположен полупроводник p -типа, справа – полупроводник n -типа.

В p -области ионы акцепторной примеси обозначены кружком со знаком минус, рядом белой точкой обозначена дырка. Соответственно, в n -области ионы донорной примеси обозначены кружком со знаком плюс, рядом черной точкой обозначен электрон.

Ионы примесей находятся в узлах кристаллической решетки (фиксированы), а дырки и электроны могут перемещаться по кристаллу.

Носители заряда, концентрация которых в данном полупроводнике больше, называют *основными*, а носители, концентрация которых меньше, – *неосновными*. Концентрация основных носителей гораздо больше, чем концентрация неосновных носителей (примерно в 10^6 раз).

В p -области основными носителями тока являются дырки. Однако, в этой области имеется также небольшое число неосновных носителей – электронов.

Соответственно, в n -области основные носители заряда – электроны, а неосновные носители – небольшое число дырок.

Вследствие значительного различия в концентрации электронов и дырок по разные стороны от перехода *происходит диффузия* – поток носителей из области с большей концентрацией в область с меньшей концентрацией.

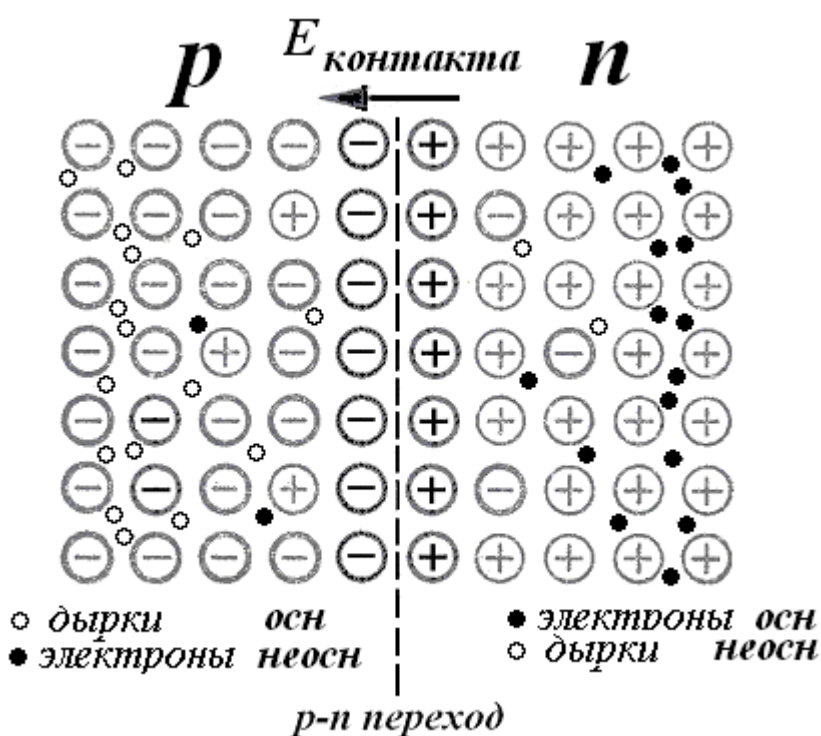


Рис.19.

Дырки из p - области, где их концентрация выше, диффундируют (инжектируются) в n - область, где концентрация дырок ниже. Электроны диффундируют в противоположном направлении – в направлении $n \rightarrow p$.

Диффундируя во встречных направлениях через пограничный слой, дырки и электроны **рекомбинируют** (соединяются) друг с другом в приконтактной области.

Дырки интенсивно рекомбинируют с электронами, концентрация которых в n -области велика. Вследствие этого новые дырки могут входить из p - в n -область. То же самое имеет место для электронов, инжектируемых из n -области в p - область.

Рекомбинация приводит к исчезновению пары электрон проводимости – дырка.

Такой встречный процесс диффузии заряженных частиц эквивалентен электрическому току $I_{\text{осн}}$ через p - n -переход, который течет из p -области в n -область. Этот **ток основных носителей** заряда называется **диффузионным**.

Диффузионный поток дырок приводит к обеднению дырками области p -полупроводника, прилегающей к переходу (рис. 19). В p - области после ухода дырок вблизи границы раздела остаются неподвижные отрицательные ионы акцепторной примеси, заряд которых теперь не компенсируется дырками. Они образуют слой отрицательно заряженных ионов (рис. 19).

Диффузионный поток электронов приводит к обеднению электронами области n - полупроводника, прилегающей к переходу. В n - области вследствие ухода электронов вблизи границы остаются неподвижные положительные донорные ионы, заряд которых теперь не компенсируется электронами. Они образуют прилегающий к первому слой положительно заряженных ионов (рис. 19).

Таким образом, в p – n -переходе возникает **двойной электрический слой пространственного заряда**. Электрическое поле перехода направлено от положительных зарядов к отрицательным, т.е. из слоя n в слой p . Контактное электрическое поле создает потенциальный барьер - **контактную разность потенциалов** между смежными слоями.

Контактная разность потенциалов $\Delta\phi$ обычно составляет десятые доли вольта. Для германиевого перехода $\Delta\phi \approx 0,4$ В, для кремниевого перехода $\Delta\phi \approx 0,6$ В

Напряженность электрического поля двойного слоя $E_{\text{контакта}}$ направлена от n - к p – полупроводнику (рис. 19). Это поле будет отталкивать дырки p - области влево от границы раздела полупроводников, а электроны n -области - вправо от этой границы. Контактное поле препятствует дальнейшему переходу дырок – направо, а электронов – налево, т.е. **препятствует диффузионному току основных носителей тока**.

Зато двойной электрический слой в p - n -переходе **способствует** встречному **движению неосновных** носителей заряда – дырок из n -области и электронов из p -области. Такой «дрейф» заряженных частиц через переход представляет собой электрический **ток неосновных носителей** $I_{\text{неосн}}$, который направлен противоположно диффузионному току и называется **дрейфовым**.

В результате двух процессов, действующих в противоположные стороны: диффузии и движению зарядов в электрическом поле, устанавливается стационарное (не зависящее от времени) состояние.

В условиях термодинамического равновесия, которое устанавливается в p - n -переходе, если к нему не приложена внешняя разность потенциалов, диффузионный ток $I_{\text{осн}}$ по величине точно равен дрейфовому $I_{\text{неосн}}$, оба тока компенсируют друг друга и **полный ток через переход равен нулю**.

Область p - n -перехода обеднена подвижными носителями заряда, т.к. здесь благодаря встречному потоку электронов и дырок происходит их интенсивная рекомбинация. По этой причине область p - n -перехода обладает гораздо большим удельным сопротивлением, чем весь кристалл полупроводника, и называется **«запирающим слоем»**.

Толщина слоя p - n -перехода составляет порядка одного микрометра, а контактная разность потенциалов – десятые доли вольта.

Поскольку подавляющая часть электронов в зоне проводимости (и дырок в валентной зоне) имеют энергию порядка kT (т.е. 0,026 эВ при комнатной температуре), большинство носителей не в состоянии преодолеть такой энергетический барьер. Носители заряда могут преодолеть такую разность потенциалов только при температуре в десятки тысяч градусов, т.е. при обычных температурах контактный слой является запирающим.

Контрольные вопросы

- 8.1. Что называют основными и неосновными носителями заряда?
- 8.2. Как действует электрическое поле контакта на основные и неосновные носители заряда?
- 8.3. Какие два конкурирующих процесса компенсируют действие друг на друга в p - n -переходе в состоянии равновесия?
- 8.4. Как образуется диффузионный ток? Дрейфовый ток?
- 8.5. Почему p - n -переход называют запирающим слоем?

§ 9. Неравновесное состояние p - n -перехода. Прямое и обратное смещение

Сопротивление запирающего слоя можно изменить с помощью внешнего электрического поля. При этом **возможны два варианта**.

1. **Прямое смещение**. Приложим к p - n -переходу внешнее напряжение так, чтобы положительный полюс источника напряжения (+) был подан на p -область, а отрицательный полюс источника (-) был подан на n -область (**режим прямого смещения**, рис. 20).

При этом внешнее электрическое поле в кристалле $E_{\text{внеш}}$ будет направлено противоположно полю контактного запирающего слоя $E_{\text{конт}}$. Внешнее поле вызовет движение дырок из области p -полупроводника и электронов из области n -полупроводника к границе p - n -перехода. (рис. 20).

Как только основные носители минуют электронно-дырочный переход, они становятся неосновными. Поэтому концентрация неосновных носителей по

обе стороны перехода увеличивается. Происходит так называемая **инжекция неосновных носителей** от латинского *injection* (вбрасываю)

Одновременно в p – и n – области через боковые металлические контакты от источника тока входят равные количества основных носителей, которые **рекомбинируют** с инжектированными носителями. В результате появляется отличный от нуля ток

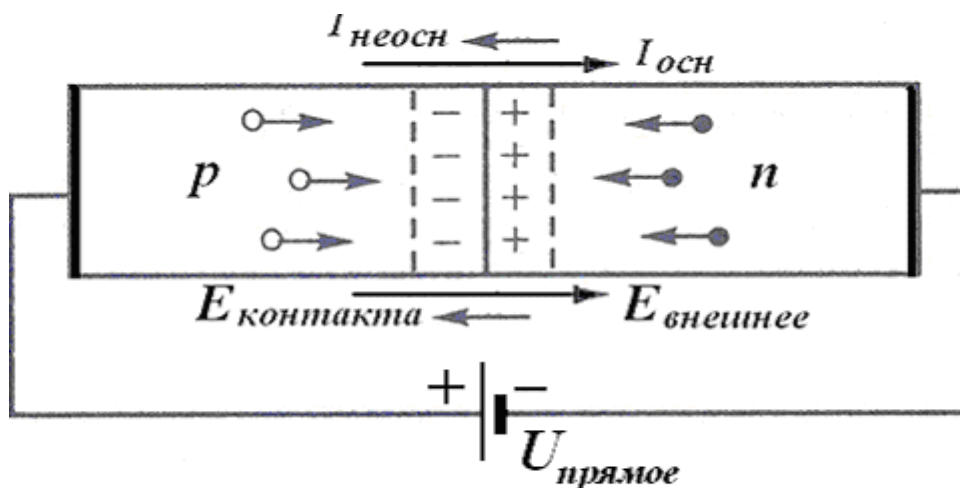


Рис. 20.

через переход, который с ростом напряжения экспоненциально возрастает.

Ток на всех участках цепи обеспечивается основными носителями, сам p – n –переход обогащен носителями тока. Электрическое поле «поджимает» основные носители к границе между областями, вследствие чего ширина переходного слоя, обедненного носителями, уменьшается. Соответственно уменьшается и сопротивление перехода, причем тем сильнее, чем больше напряжение.

Ток же неосновных носителей (дрейфовый) останется практически без изменения.

Уменьшение контактной разности потенциалов (понижение потенциального барьера) пропорционально приложенному напряжению.

Таким образом в направлении от p – области к n – области p – n –переход пропускает ток, сила которого **быстро нарастает** при увеличении приложенного напряжения. Это напряжение называется **прямым** (или пропускным).

2. **Обратное смещение.** Теперь приложим к p - n - переходу напряжение противоположного направления, чтобы «+» был подключен к n -области, а «-» был подключен к p -области (**режим обратного смещения**, рис. 21).

При этом направление внешнего электрического поля

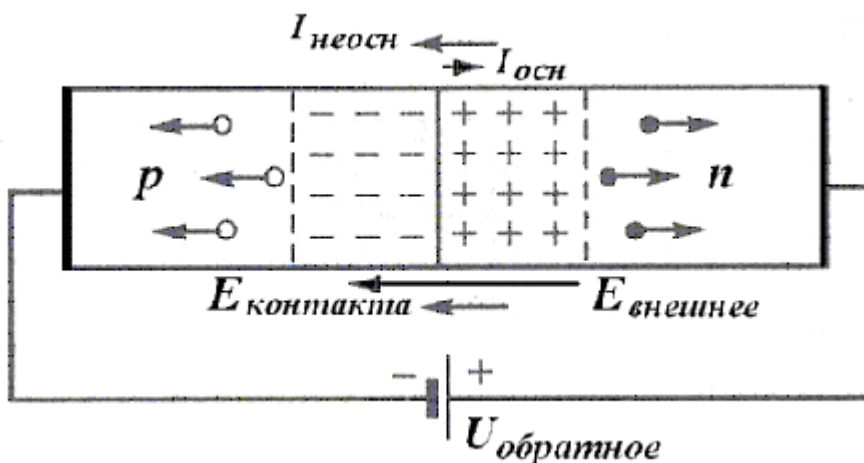


Рис. 21.

$E_{\text{внеш}}$ будет совпадать с направлением контактного поля $E_{\text{конт}}$ (рис. 21).

В этом случае внешнее поле будет усиливать поле контактного слоя и обусловит движение электронов и дырок от границы $p - n$ -перехода в противоположных направлениях. Поле, возникающее в кристалле при наложении обратного напряжения «оттягивает» основные носители от границы между областями. На границе раздела областей в итоге не останется основных носителей тока.

Это приводит к возрастанию ширины переходного слоя, обедненного носителями. Соответственно увеличивается и сопротивление перехода. Следовательно, $p - n$ -переход обладает в обратном направлении гораздо большим сопротивлением, чем в прямом.

В данном случае через $p - n$ -переход протекает только небольшой ток (он называется *обратным*), целиком обусловленный неосновными носителями. Ток неосновных носителей не зависит от величины контактного поля, так как для них потенциального барьера не существует. Потoki неосновных носителей определяются скоростью тепловой генерации электронно-дырочных пар. Эти пары диффундируют к переходу и разделяются его полем. В результате через переход течет ток (ток насыщения), который обычно мал и почти не зависит от напряжения.

Контрольные вопросы

- 9.1. Чем отличаются режимы прямого и обратного смещения?
- 9.2.. Что означает термин «инжекция неосновных носителей»?
- 9.3. Каков механизм пополнения количества дырок в p -области прямосмещенного $p - n$ -перехода?

§ 10. Контакт p - и n -полупроводников в терминах зонной теории и энергетических барьеров

Рассмотрим теперь механизм возникновения в $p-n$ -переходе контактной разности потенциалов с точки зрения зонной теории.

На рис. 22 изображены схемы энергетических уровней до контакта.

Обозначения на рис 22:

E_C – нижний уровень (дно) зоны проводимости, E_V – верхний уровень (потолок) валентной зоны, E_F – уровень Ферми, E_A – акцепторный уровень (пустой), E_D – донорный уровень (занят электронами).

В левой части рис. 22 в области полупроводника p -типа уровень Ферми E_F находится у акцепторного уровня E_A , т. е. вблизи потолка валентной зоны E_V , а в правой части рис. 21 в области полупроводника n -типа — вблизи донорного уровня E_D , т. е. у дна зоны проводи-

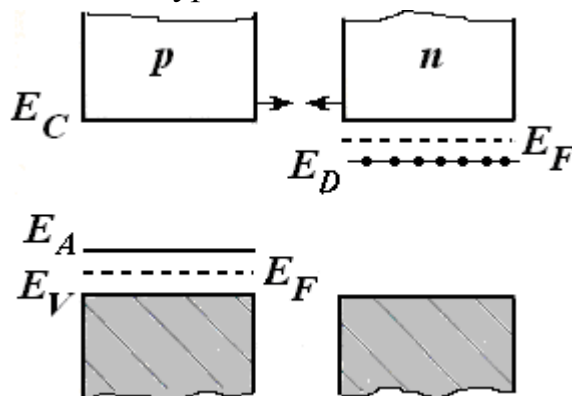


Рис. 22.

мости E_C .

При контакте двух полупроводников электронам энергетически выгодно переходить из n - в p - область, т.к. в p - области имеются вакантные состояния с более *низкой* энергией.

Диффузия электронов приводит к уменьшению полной энергии электронов в n -области и увеличению их полной энергии в p -области. Этот процесс продолжается до тех пор, пока в соответствии с законами термодинамики энергии Ферми не сравняются в обоих материалах (рис. 23).

Один из принципов термодинамики состоит в том, что *при контакте двух тел, способных обмениваться друг с другом частицами, в состоянии термодинамического равновесия их электрохимические потенциалы, т.е. уровни Ферми выравниваются.*

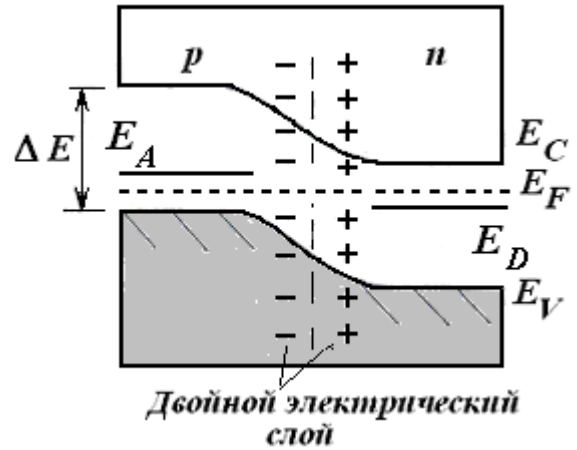


Рис. 23.

Выравнивание уровней Ферми обеспечивает *баланс* энергетических распределений электронов в зоне проводимости (и дырок в валентной зоне) по обе стороны границы раздела и прекращение результирующего потока зарядов через эту границу. Оба полупроводника приходят в *равновесие*.

С классической точки зрения данная ситуация аналогична смешиванию двух электронных «газов» с различными «температурами» или средними кинетическими энергиями электронов, т. е. с различными энергиями Ферми. Следствием этого явилось бы выравнивание энергий Ферми холодного и горячего газов; система пришла бы к тепловому равновесию.

Установление равновесия приводит к образованию *контактной разности потенциалов* между двумя различными материалами.

В каждой из смежных областей сохраняется при этом прежнее расположение энергетических зон и уровней относительно химического потенциала (уровня Ферми), за исключением узкого переходного слоя. Поскольку уровни Ферми в смежных областях выравниваются (рис.23), происходит *искривление зон*.

Изгибание энергетических зон в области перехода вызвано следующими обстоятельствами.

С электрическим полем $E_{\text{контакта}}$ можно связать потенциальную энергию дырки E_{ph} и электрона E_{pe} в смежных областях (рис. 24).

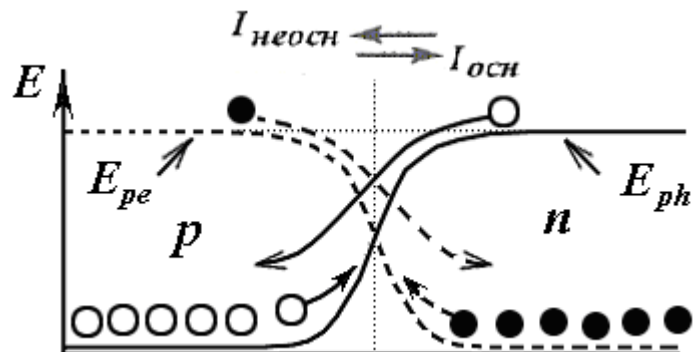


Рис. 24.

p -Область в месте контакта заряжается отрицательно (рис. 19) и ее **потенциал** в состоянии равновесия **ниже**, чем потенциал n -области, **зато потенциальная энергия электрона** (его заряд отрицательный!) **в p -области наоборот больше**, чем в n -области. Это вызывает **повышение всех уровней p -области**, в том числе и уровня Ферми.

Соответственно зарядение в месте контакта n -области положительно вызывает понижение всех ее уровней, в том числе и уровня Ферми.

В равновесии уровни Ферми располагаются на одной высоте.

Поэтому граница валентной зоны E_V описывает ход потенциальной энергии электрона E_{pe} в направлении, перпендикулярном к переходу (рис. 24, штриховая линия, электроны – черные кружки). Поскольку заряд дырок противоположен заряду электронов, потенциальная энергия дырок E_{ph} (сплошная линия, дырки – белые кружки) больше там, где меньше потенциальная энергия электронов E_{pe} и наоборот.

При этом основным носителям приходится преодолевать контактное поле. Область, в которой повышается потенциальная энергия, соответствующая силам отталкивания, называется **потенциальным барьером**.

Получается, что дырка для перехода из p -области в n -область должна "забраться" на потенциальный барьер высоты E_{ph} (рис. 24.) На аналогичный барьер должен "забраться" электрон для перехода из n -области в p -область.

Некоторому количеству основных носителей (кружки на нижних уровнях) удастся преодолеть потенциальный барьер, вследствие чего через переход течет небольшой **ток основных носителей**. Неосновных носителей гораздо меньше, но они легко проходят через границу областей, «скатываясь» с потенциального барьера Они образуют **ток неосновных носителей**, который не зависит от высоты барьера.

Равновесие наступает при такой высоте барьера, когда оба тока компенсируют друг друга.

Если теперь к p - n -переходу приложить внешнюю разность потенциалов U , чтобы произошло **прямое включение** перехода, то внешнее поле $E_{внеш}$ уменьшит существующее в кристалле поле $E_{конт}$, и высота барьеров уменьшится (рис. 25).

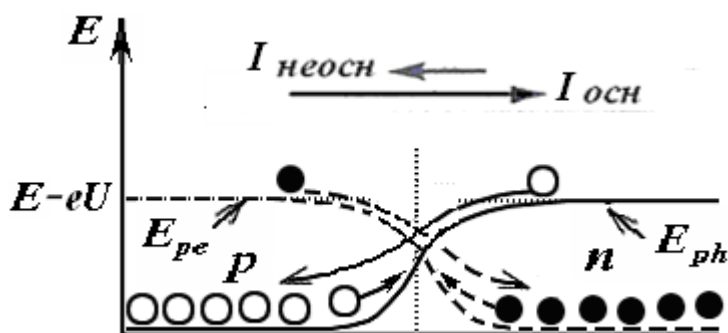


Рис. 25.

Понижение потенциального барьера на величину eU пропорционально приложенному напряжению U . Не удерживаемые больше полем барьера, электроны инжектируются в p -область, а дырки – в n -область.

Тогда ток основных носителей возрастет экспоненциально с ростом приложенного напряжения U :

$$I_{осн} = I_0 \exp(eU/kT). \quad (9)$$

Ток же неосновных носителей ($-I_0$) при этом практически не изменится. Следовательно, результирующий ток через контакт будет отличен от нуля.

Если к p - n -переходу приложить **обратное смещение**, то внешнее поле $E_{\text{внеш}}$ увеличит существующее на границе поле $E_{\text{конт}}$, и высота барьеров увеличится (рис. 26). Ток основных носителей от этого уменьшится, а ток неосновных носителей практически не изменится, так как он определяется только числом неосновных носителей.

Тогда полный ток через p - n -переход определится выражением

$$I = I_0 (\exp(eU/kT) - 1) \quad (10)$$

Здесь I_0 – максимальный обратный ток, U – приложенное внешнее напряжение, I – результирующий дырочный ток, текущий в n -область. Там дырки рекомбинируют с электронами проводимости, а убыль электронов при рекомбинации пополняется за счет тока электронов из внешнего источника напряжения.

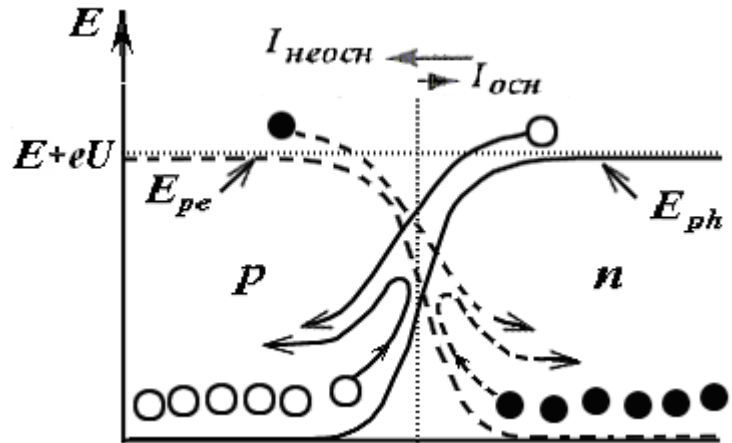


Рис. 26.

Из формулы (10) видно, что в отсутствие внешнего поля ($U = 0$) ток I равен нулю. Прямой ток ($U > 0$) на несколько порядков превышает обратный ток ($U < 0$). При увеличении отрицательного напряжения (так, что $|eU|$ составляет несколько величин kT) ток через переход насыщается: $\exp(-eU/kT) \rightarrow 0, I \approx -I_0$.

На рис. 27 изображены зависимости тока от внешнего напряжения U при прямом и обратном включении, и построена вольтамперная характеристика p - n -перехода при $U \geq 0$ и при $U \leq 0$.

Обратный ток быстро достигает насыщения, т.е. перестает зависеть от U . Только при очень большом обратном напряжении сила тока начинает резко возрастать, что обусловлено электрическим пробоем перехода (см. штриховую часть на левой ветви на рис. 27)

При обратном смещении ток практически можно считать равным нулю, поэтому говорят, что p – n -переход пропускает ток практически только в одном направлении, т.е. обладает односторонней проводимостью и может использоваться для **выпрямления тока и детектирования сигналов**.



Рис. 27.

Пример 10.1. Сопротивление p - n -перехода при температуре 0°C и при прямом напряжении $U = 0,1$ В равно $R_{\text{прям}} = 10$ Ом. Каким будет сопротивление перехода $R_{\text{обрат}}$, если изменить полярность напряжения.

В соответствии с законом Ома и формулой (8), сопротивление перехода обратно пропорционально величине $(\exp(eU/kT) - 1)$. Поэтому отношение сопротивлений в обратном и прямом направлении равно (учтем, что при изменении напряжения на обратное, ток через переход меняет знак)

$$\frac{R_{\text{обр}}}{R_{\text{прям}}} = - \frac{(\exp(eU/kT) - 1)}{(\exp(eU/kT) - 1)} \quad (\text{П. 10.1})$$

Отдельно вычислим показатель экспоненты:

$$\frac{eU}{kT} = \frac{1,6 \cdot 10^{-19} \cdot 0,1}{1,38 \cdot 10^{-23} \cdot 273} = 4,25.$$

Тогда $\frac{R_{\text{обр}}}{R_{\text{прям}}} = - \frac{(\exp(4,247) - 1)}{(\exp(-4,247) - 1)} = 69,873$. Отсюда $R_{\text{обрат}} \approx 699$ Ом.

Контрольные вопросы

- 10.1. Как изменяется высота потенциального барьера при прямом и обратном смещении?
- 10.2. Как располагаются уровни Ферми в p - и n -полупроводниках при их контакте в состоянии равновесия?
- 10.3. Чем вызвано изгибание зон в месте контакта p - и n -полупроводников?
- 10.4. Как зависит сила тока неосновных носителей от знака смещения?
- 10.5. Как объяснить режим насыщения обратного тока?
- 10.6. Какое напряжение U приложено к p - n -переходу, если при температуре $T = 300$ К коэффициент выпрямления $n = I_{\text{пр}}/I_{\text{обр}} = 1200$?
- 10.7. При некотором напряжении сила прямого тока сквозь p - n -переход $I_{\text{пр}} = 2,02$ мА, а обратного $I_{\text{обр}} = 20$ мкА. Найти по этим данным силу тока насыщения I_0 для p - n -перехода.
- 10.8. При некоторой температуре коэффициент выпрямления для p - n -перехода $n_1 = 1600$. Чему будет равен коэффициент выпрямления n_2 , если температура полупроводника будет повышаться на $\eta = 0,1$ от начальной при постоянном напряжении?
- 10.9. Почему контакт p - и n -полупроводников обладает выпрямляющими свойствами?

§ 11. Выпрямление переменных токов с помощью p - n -перехода. Полупроводниковый выпрямительный диод

Электронно-дырочный переход находит многообразные применения, являясь основой различных полупроводниковых приборов.

Для преобразования переменного тока в постоянный (выпрямления тока) используют **полупроводниковый диод**.

Как известно, генераторы электростанций вырабатывают электроэнергию в форме переменного напряжения так называемой промышленной частоты (50 Гц). В то же время многие приборы, машины и промышленные установки нуждаются в питании постоянным током. Например, постоянным током питаются все виды транспорта на электрической тяге: трамваи, троллейбусы, электропоезда. Чтобы обеспечить работу электротранспорта, необходимо преобразовать переменное напряжение в постоянное. Необходимость в выпрямлении

тока существует и в радиотехнике (питание радиоприемников и телевизоров), технике СВЧ и т.д.

Полупроводниковый **диод** - это полупроводниковый прибор с одним *p-n*-переходом и двумя выводами (электродами). Конструкция прибора определила его название. Диод в переводе с греческого – «два пути». Этот электронный прибор обладает различной проводимостью в зависимости от направления электрического тока

На схемах полупроводниковый диод обозначается, как показано на рис. 28 (буквы *p* и *n* на схемах, конечно, не ставятся, здесь они проставлены для пояснения). Стрелкой показано направление, в котором диод проводит ток. **Прямое** (пропускное) направление тока получается, если «+» подавать на *p*-область, а «-» подавать на *n*-область. Противоположное подключение диода (показано в скобках) называется **обратным** (запорным).



Рис. 28.

Практически используемые *p - n*-переходы являются резко несимметричными. Одна из областей, образующих переход, легируется значительно сильнее, чем другая. В результате концентрация носителей, которые сильнолегированная область способна инжектировать в слаболегированную, может стать во много раз больше, чем исходная концентрация носителей в слаболегированной области.

Поэтому сильнолегированную область диода называют **эмиттером** (от латинского *emitto* - выпускаю). Слаболегированную область диода называют **базой**.

Полупроводниковые диоды при соответствующем охлаждении используют даже при очень больших токах порядка 1000 А.

Самый простой выпрямитель на диоде – **однополупериодный**.

Воспользуемся схемой подключения диода, изображенной на рис. 29. Если на вход подать синусоидальный сигнал $U_{\text{вход}}$, то через переход будет проходить ток (прямой) только в те половины периода, когда на *p*-область подается «плюс» от источника напряжения. Во вторую половину периода, когда на *p*-область подается «минус», ток падает практически до нуля, диод закрыт.

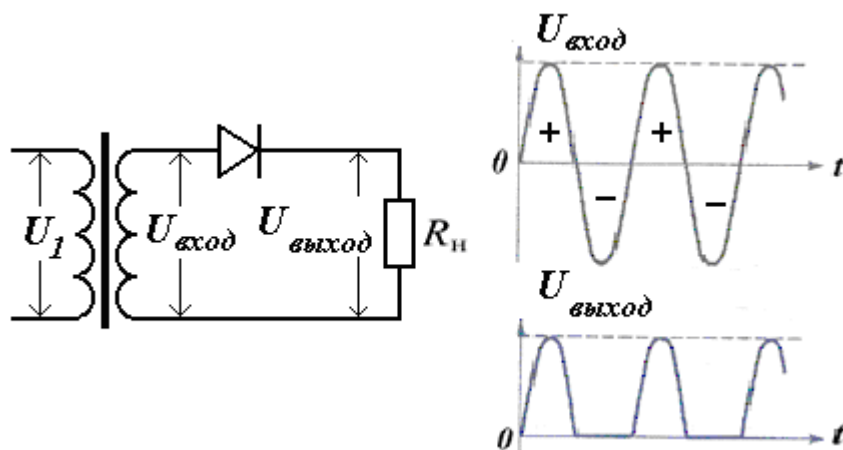


Рис. 29.

Таким образом, диод "пропустит" только положительные полу-волны синусоиды. На сопротивлении нагрузки R_n (на выходе) сигнал будет иметь вид пульсирующего напряжения $U_{\text{выход}}$ (рис. 29).

Чтобы сгладить пульсации и получить огибающую сигнала, используют дополнительный конденсатор C , включенный параллельно нагрузке. В течение времени, когда через диод проходит ток, конденсатор заряжается, а в остальное время разряжается через нагрузку. Напряжение в нагрузку подается только с заряженного в предыдущий полупериод конденсатора. Заряжаясь и разряжаясь, такой фильтр сгладит острые полуволны (рис. 30).

По такой схеме работают простейшие выпрямители напряжения — устройства, преобразующие переменный ток в постоянный.

Из сравнения осциллограмм видно, что пульсации

напряжения на нагрузке сглаживаются. Чем больше нагрузочное сопротивление, тем меньше ток разряда конденсатора, тем медленнее падает на нем напряжение, и тем лучше сглаживаются пульсации.

Четыре диода, собранные как указано на рис. 31, образуют так называемый *мост*. Принципиальная схема выпрямителя и осциллограммы напряжения

приведены на рис. 31.

Пусть в *первую половину периода* на верхний вывод вторичной обмотки трансформатора поступает «+», на нижний конец поступает «-». Ток прохо-

дит по следующей цепи: верхний вывод вторичной обмотки — диод D2 — верхний вывод нагрузки — нагрузка — нижний вывод нагрузки — диод D3 — нижний вывод вторичной обмотки — обмотка.

Во *вторую половину периода* на верхний вывод вторичной обмотки трансформатора поступает «-», на нижний конец поступает «+». Ток проходит по следующей цепи: нижний вывод вторичной обмотки — диод D4 — верхний

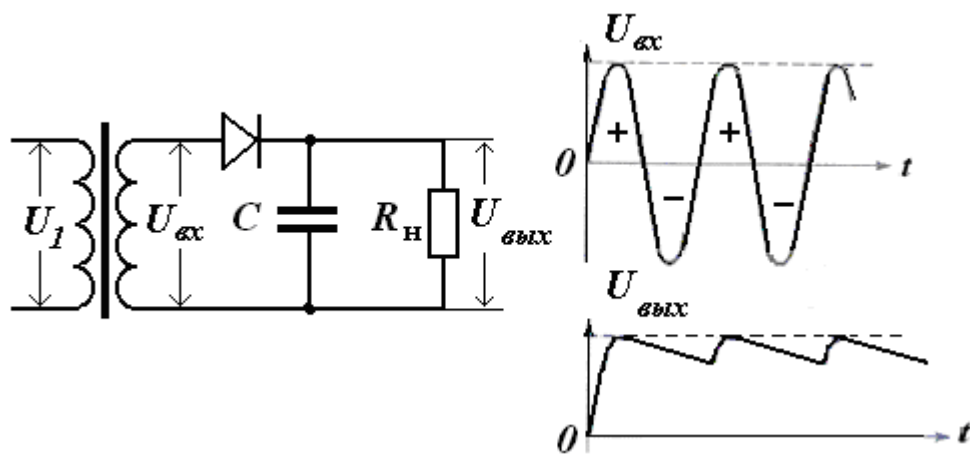


Рис. 30.

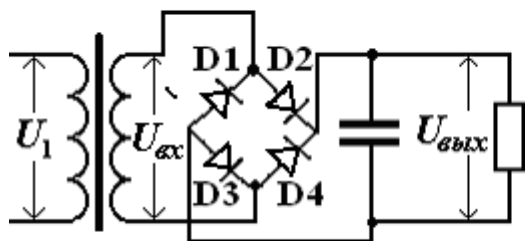


Рис. 31.

вывод нагрузки - нагрузка – нижний вывод нагрузки – диод D1 – верхний вывод вторичной обмотки – обмотка.

Как видим, в обоих случаях направление тока через нагрузку одинаково.

По сравнению с однополупериодной схемой мостовая схема имеет в 2 раза меньший уровень пульсаций.

Диодный выпрямитель или диодный мост – основной компонент блоков питания практически всех электронных устройств.

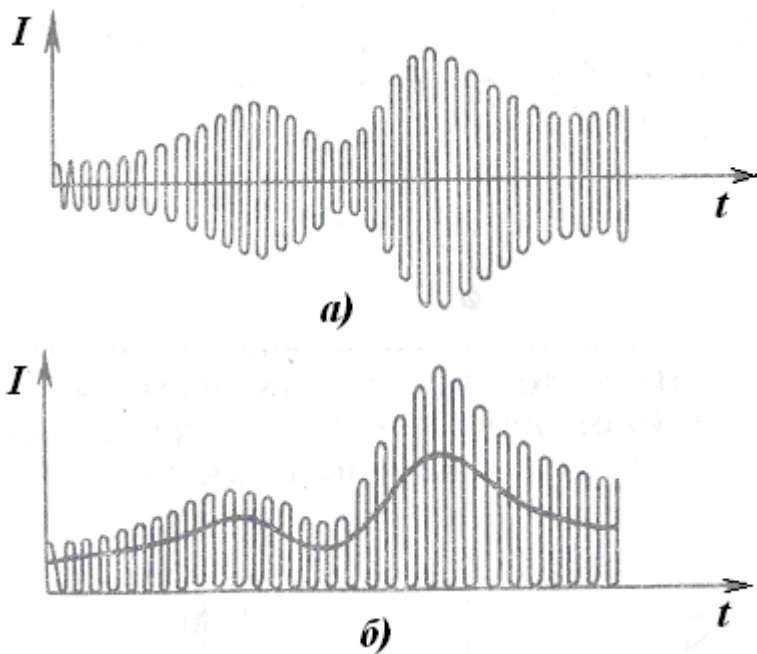


Рис. 32.

Другим важным практическим применением диода является **детектирование (демодуляция)**.

Для передачи информации с помощью радиоволн применяется модуляция колебаний. Переносчиком сигнала являются синусоидальные колебания высокой частоты. Их амплитуда (при амплитудной модуляции) модулируется передаваемым сигналом. Колебания силы тока, модулированные по амплитуде показаны на рис.32 а).

ны на рис.32 а).

В радиоприемном устройстве для получения колебаний звуковой частоты необходимо произвести демодуляцию. Такое детектирование в большинстве случаев осуществляется с помощью диодов. График колебаний силы тока в цепи диода имеет вид, изображенный на рис. 32 б).

Далее ток звуковой частоты (показан жирной линией) отделяется от высокочастотного и пропускается через устройство, воспроизводящее звук.

Контрольные вопросы

- 11.1. На чем основано действие выпрямительного диода?
- 11.2. С какой целью в выпрямителе включают параллельно нагрузке конденсатор?
- 11.3. Какой вид будет иметь напряжение на резисторе в схеме рис.28, если диод подключить наоборот?
- 11.4.. Можно ли принятые радиоприемным устройством электромагнитные колебания после их усиления подать непосредственно на динамик?
- 11.5. Почему необходимо отделять высокочастотные колебания от низкочастотных при радиоприеме? Как это осуществляется?
- 11.6. Где применяются выпрямительные полупроводниковые диоды?

§ 12. Светоизлучающий диод

Если включить p - n -переход в прямом направлении (рис. 33), то внешнее напряжение U понижает потенциальный барьер. В n -область инжектируются избыточные дырки, а в p -область — электроны, которые затем рекомбинируют с носителями заряда противоположного знака.

Часть носителей рекомбинирует излучательно, т.е. с рождением фотона. Так возникает свет, рожденный протеканием через p - n -переход прямого тока.

Это явление получило название *инжекционной электролюминесценции*. На основе явления инжекционной электролюминесценции создан полупроводниковый прибор, преобразующий электрическую энергию в энергию оптического излучения, - *светодиод (светоизлучающий диод)*.

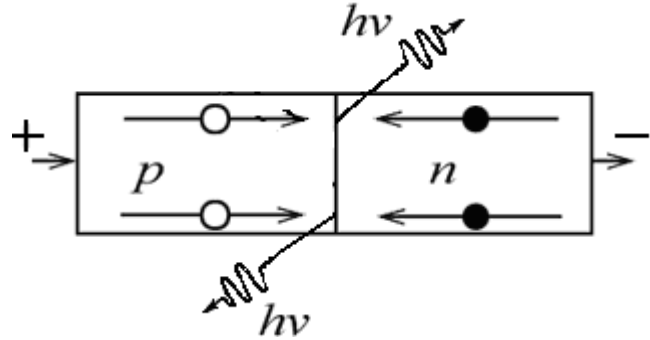


Рис. 33.

Конечно, процесс рекомбинации не обязательно должен сопровождаться излучением. Часть энергии, которая выделяется при рекомбинации, может расходоваться на нагревание кристалла, а остальная может излучаться в виде квантов света. Полупроводниковый кристалл должен содержать мало дефектов, из-за которых рекомбинация происходит без излучения.

В терминах зонной теории можно представить себе три варианта таких процессов:

- переход электрона из зоны проводимости к дырке валентной зоны;
- переход электрона из зоны проводимости к дырке донорного уровня;
- переход электрона с акцепторного уровня к дырке валентной зоны (рис.. 34).

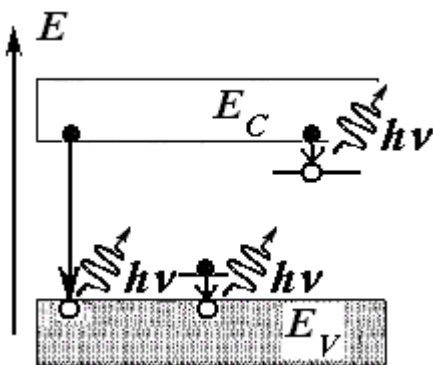


Рис. 34.

В зависимости от ширины запрещенной зоны в полупроводнике испускаются кванты электромагнитного излучения той или иной частоты от инфракрасного до ультрафиолетового излучения. Длина волны λ излучения светодиода определяется энергией фотона, испускаемого при рекомбинации

$$\varepsilon_{\phi} = hc/\lambda.$$

В большинстве случаев эта энергия близка к ширине запрещенной зоны полупроводника ΔE .

Тогда

$$\lambda = hc/\Delta E. \quad (11)$$

В таблице 4 представлены значения ширины запрещенной зоны некоторых светодиодов и длина волны излучаемого ими света.

Таблица 4.

Полупроводник	Ширина запрещенной зоны ΔE , эВ	Длина волны λ , мкм, свет
AgI	2,8	0,44, синий
ZnSe	2,7	0,46, синий
CdS	2,4	0,52, зеленый
GaP	2,3	0,54, зеленый
CdSe	1,7	0,73, красный
GaAsP	1,6	0,78, красный
GaAs	1,43	0,87, инфракрасный

Для изготовления светодиодов пригодны только материалы с достаточно высокой вероятностью излучательной рекомбинации.

Германий и кремний этому требованию не отвечают. В них подавляющая доля электронов и дырок рекомбинирует безызлучательно.

Зато в арсениде галлия GaAs и тройных полупроводниковых соединениях на его основе GaAsP и GaAlAs вероятность излучательной рекомбинации близка к единице.

Цвет (длина волны) свечения при инжекционной электролюминесценции зависит от материала основы и природы примесей. Широкое распространение получили светодиоды с инжекционной электролюминесценцией на *гетероструктурах* (гетероструктура — выращенная на подложке слоистая структура из различных полупроводников, позволяющая конструировать нужную ширину запрещенной зоны).

Светодиоды на основе твердых растворов AlGaInN дают свечение в зеленом, синем, фиолетовом и ближнем ультрафиолетовом диапазонах спектра.

На основе InGaN созданы светодиоды высокой яркости, дающие синее, зеленое и ультрафиолетовое излучение, на основе AlGaInP - желтое, оранжевое и красное, на основе AlGaAs - красное и инфракрасное.

Основой светодиодов белого цвета свечения является структура InGaN, излучающая на длине волны 470 нм (синий цвет) и нанесенный сверху на нее люминофор — желтый фосфор, излучающий в широком диапазоне видимого спектра и имеющий максимум в его желтой части. Человеческий глаз комбинацию такого рода воспринимает как белый цвет.

Четверное полупроводниковое соединение GaInAsP позволяет создать инфракрасные светодиоды с длиной волны $\lambda \sim 1 \div 1.3$ мкм, которые широко применяются в волоконно-оптических линиях связи.

Самая распространенная *конструкция светодиода* включает полупроводниковый кристалл с *p-n*-переходом, рамку с внешними выводами, на которой размещен кристалл, и герметизирующую эпоксидную смолу, которая окружает

и защищает кристалл, а также рассеивает свет (формирует диаграмму направленности) (рис. 35).

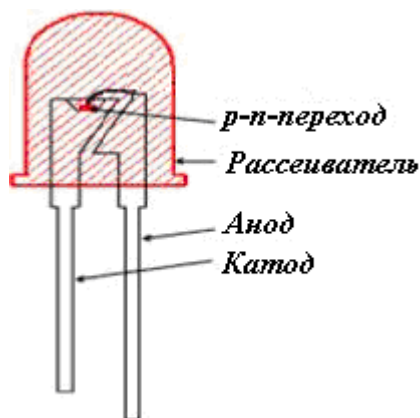


Рис. 35.

В светодиодах наиболее значительны потери энергии на полное внутреннее отражение излучения. В связи с большим различием показателей преломления полупроводника n_p и воздуха n_v доля выходящего излучения определяется значением предельного угла полного внутреннего отражения. — угла $i_{пр}$ между направлением светового луча и нормалью к поверхности. Излучение, падающее на поверхность раздела полупроводник - воздух под углом, меньшим критического, выводится из кристалла, а под углом, большим критического, испытывает полное внутреннее отражение.

Для полупроводников GaAs и GaP значения показателя преломления составляют соответственно 3,54 и 3,3, а значения предельного угла $i_{пр}$ равны примерно 16° и 18° . С целью снижения потерь на полное внутреннее отражение для рассеивателя выбирают полусферическую форму.

Светодиоды обладают высоким КПД, достигающим 80 %; КПД лучших ламп накаливания на порядок ниже.

Светодиоды долговечны и надежны, так как не содержат нитей накаливания, катодов и других быстро изнашивающихся узлов, как, например, лампы накаливания или газоразрядные лампы.

Из других преимуществ светодиодных источников света следует отметить такие:

- излучение светодиодов близко к монохроматическому. Это позволяет их использовать без светофильтров; например, в светофорах, автомобильных сигнальных огнях;
- светодиод устойчив к механическим воздействиям и работает в широком диапазоне температур;
- светодиод не имеет экологически вредных компонентов (например, ртути в люминесцентных лампах);
- не имеет разбивающихся стеклянных баллонов, необходимых в лампах накаливания, выделяет очень мало тепла и имеет безопасное электропитание;
- светодиодные светильники не боятся бросков напряжения и не обладают стробоскопическим эффектом.

Сейчас практически нет областей промышленности, где бы ни использовались светодиоды. Они используются в качестве различных индикаторов в бытовых и промышленных устройствах, цифробуквенных индикаторов, в устройствах для подсветки ЖК экранов и мониторов, для освещения помещений, для наружной рекламы, в игрушках и т.д.

В заключение приведем условное обозначение светодиода на электрических схемах (рис. 36).



Рис. 36.

Контрольные вопросы

12. 1. На чем основано действие светодиодов? Каковы их преимущества по сравнению с другими источниками света? Какие полупроводниковые материалы используют для изготовления светодиодов?

12. 2. От чего зависит цвет излучения светодиода?

12. 3. Какой должна быть ширина запрещенной зоны полупроводника, из которого изготовлен светодиод, светящийся зеленым светом ($\lambda = 500$ нм)?

12. 4. Пользуясь таблицей 1, определите длину волны излучения светодиода, изготовленного из арсенида галлия GaAs, фосфида галлия GaP (без примесей),

12. 5. Какого цвета излучение светодиода, изготовленного из карбоната кремния α -SiC с шириной запрещенной зоны 3,12 эВ?

12. 6. Чем отличаются светодиод от тепловых источников света? Каковы их преимущества по сравнению с другими источниками света?

12. 7. Чем обусловлена полусферическая конструкция светодиода?

12. 8. При выполнении каких условий рекомбинация электронов и дырок сопровождается испусканием квантов излучения?

§ 13. Усиление переменных токов с помощью *p-n*-переходов.

Транзистор

Одно из самых важных применений полупроводников заключается в их использовании для усиления и генерации электрических колебаний. Полупроводниковый *транзистор* – прибор, способный усиливать сигналы, токи и напряжения.

Одним из типов полупроводниковых усилителей, использующих явление инжекции, является *биполярный транзистор* (рис. 37). Он состоит из пластинки монокристаллического полупроводника, в которой путем соответствующего распределения примесей созданы *два* близко расположенных *p-n-перехода*.

Область между обоими переходами называют *базой*, оконечные области – *эмиттером* и *коллектором*. К каждой из областей подведены проводящие контакты.

В зависимости от типа проводимости этих зон различают *p-n-p* (эмиттер – *p*-область, база – *n*-область, коллектор – *p*-область) и *n-p-n* транзисторы.

Для определенности рассмотрим *p-n-p* транзистор, все физические процессы аналогичны для случая транзистора *n-p-n*.

Транзистор имеет три электрода, из которых в схеме включения один – входной, другой – выходной, а третий – общий для цепей входа и выхода. В зависимости от того, какая часть транзистора соединяется с общей точкой схемы, его можно включить с цепь тремя способами.

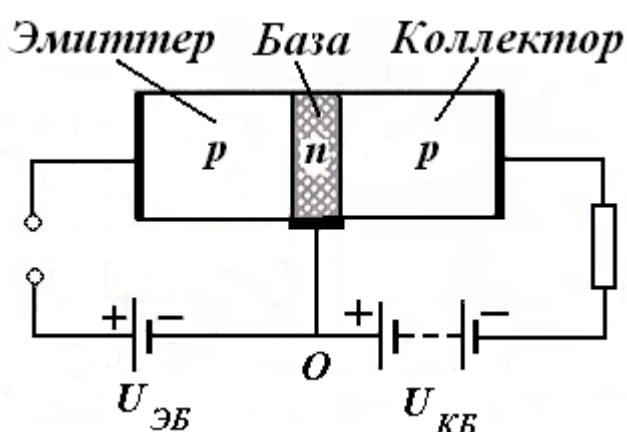


Рис. 37.

Одна из возможных схем включения, в которой базовый электрод является общим для обеих цепей – эмиттера и коллектора (*схема с общей базой*), показана на рис 37.

Между эмиттером и базой приложено небольшое прямое напряжение $U_{ЭБ}$, а между базой и коллектором – постоянное смещающее напряжение $U_{КБ}$ в обратном направлении.

Дырки как основные носители тока в p -области легко проходят через левый открытый p - n -переход. Происходит **инжекция дырок** в область базы.

Часть этих дырок рекомбинирует с основными носителями заряда в базе (электронами). Однако, базу делают очень тонкой и сравнительно слабо легированной.

Толщину базы w оценивают по ее отношению к диффузионной длине дырок в базе L_p . Диффузионной длиной L_p называют расстояние, на котором избыточная концентрация неравновесных носителей заряда в процессе диффузии уменьшается вследствие рекомбинации в e раз.

В транзисторах базу делают очень тонкой:

$$w \ll L_p. \quad (12)$$

При таком условии каждая дырка, инжектированная в базу, будет находиться от коллекторного перехода на расстоянии, меньшем, чем их диффузионная длина.

Поэтому большая часть дырок, инжектированных из эмиттера, не успевают рекомбинировать с электронами, «проскакивают» в область коллектора.

Для этих же целей концентрацию носителей в эмиттере и коллекторе, т. е. в дырочной области, подбирают гораздо большей, чем концентрацию носителей в пределах базы, т. е. в электронной области.

Сильное электрическое поле обратного смещенного коллекторного перехода захватывает дырки (напомним, что в базе они — неосновные носители, поэтому для них переход открыт), и затягивает их в коллекторную область.

Обе цепи связаны друг с другом с помощью инжекции дырок.

Всякое изменение тока в цепи эмиттера приводит к изменению количества дырок, проникающих в коллектор, и, следовательно, к почти такому же изменению тока в цепи коллектора. Ток коллектора, таким образом, практически равен току эмиттера (рис. 38), за исключением небольшой потери на рекомбинацию в базе, которая и образует

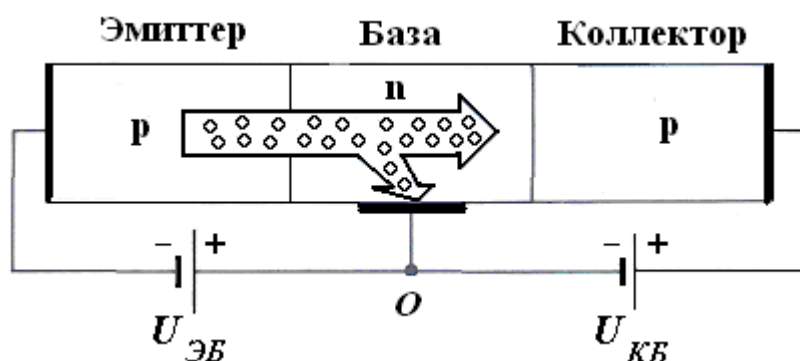


Рис. 38.

ток базы:

$$I_{Э} = I_{Б} + I_{К}. \quad (13)$$

Чем тоньше база, тем большую долю входящих в базу дырок захватывает коллектор. Долю дырок, которые переносятся благодаря процессу диффузии через базу от эмиттера к коллектору, характеризуют коэффициентом переноса α :

$$\alpha = \left| \frac{I_K}{I_E} \right| \quad (14)$$

Для различных транзисторов α может лежать в пределах от $\sim 0,9$ до $\sim 0,999$.

Казалось бы, описанное устройство как усилитель бесполезно. Действительно, изменения силы тока в цепи коллектора примерно равны изменениям сил тока в цепи эмиттера

$$I_K \approx \alpha I_E \quad (15)$$

Однако надо учесть, что правый p - n -переход включен в обратном направлении и поэтому *имеет гораздо большее сопротивление*, чем левый (несколько миллионов Ом).

Поэтому при одинаковых изменениях токов изменения напряжения в цепи коллектора будут во много раз больше, чем в цепи эмиттера. После прохода через большое нагрузочное сопротивление получается значительное усиление по напряжению и мощности.

Следовательно, транзистор, включенный по схеме с общей базой, *усиливает напряжения и мощности*. Снимаемая с транзистора повышенная мощность появляется за счет использования энергии источника постоянного напряжения $U_{КБ}$, включенного в цепь коллектора. Усиление по напряжению может достигать 10 000.

Описанный процесс можно иллюстрировать (рис. 39) *кривыми потенциальной энергии* электронов (сплошная линия) и дырок (штриховая линия). Верхняя часть рисунка изображает ход потенциальной энергии до подачи напряжения на транзистор.

Затем на переход эмиттер - база подается напряжение в прямом направлении порядка 0,3 В так, чтобы переход эмиттер - база – был слегка открыт. На переход база - коллектор подается большее напряжение (порядка 3– 30 В) в обратном направлении. Это приводит к понижению потенциального барьера на левом переходе и повышению барьера на правом (низ рис. 39).

Протекание тока в цепи эмиттера сопровождается проникновением дырок в область базы (встречный поток электронов мал вследствие того, что их концентрация невелика). Если длина диффузии дырок в области базы больше толщины базы, то значительная часть инжектированных дырок достигает коллектора. Здесь положительные дырки как неосновные носители в базе легко скатываются с потенциального

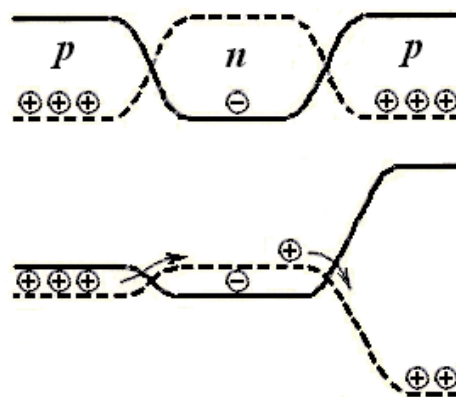


Рис. 39.

уступа. Вступая внутрь коллектора в качестве основных носителей, дырки притягиваются к отрицательно заряженному коллектору и изменяют ток коллектора.

Прикладывая между эмиттером и базой переменное напряжение, получим в цепи коллектора переменный ток, а на выходном сопротивлении — переменное напряжение. При этом инжекция дырок происходит в такт изменению входного напряжения, следовательно, форма усиленного сигнала на выходе будет сохраняться.

Условное обозначение транзистора в электронных схемах (рис. 40) — три черточки, заключенные в кружок. Черточка, в которую упираются две других, обозначает базу. Наклонная черточка без стрелки — коллектор, наклонная со стрелкой — эмиттер. Стрелка эмиттера указывает направление тока, который проходит через него. Если стрелка обращена в сторону базы — транзистор типа *p-n-p*, если от базы — *n-p-n*.

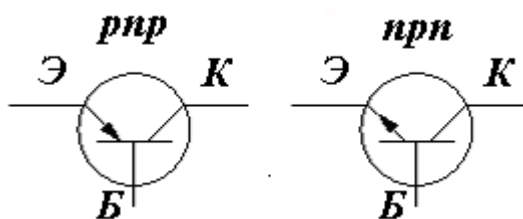


Рис. 40.

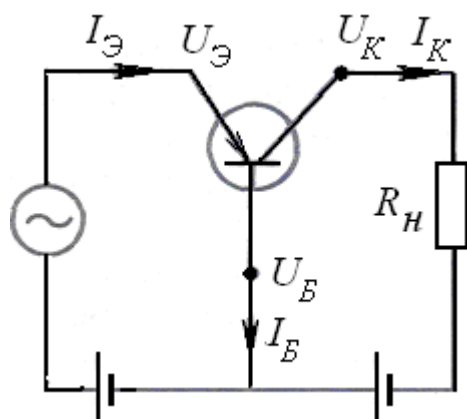


Рис. 41.

Рассмотренный нами случай транзистора с общей базой в условных обозначениях будет выглядеть как показано на рис. 41.

Другой вариант включения (схема с общим эмиттером) показан на рис. 42.

Поскольку рекомбинирует с электронами только малая часть ($1 - \alpha$) дырок, поступающих в базу (от 0,1 до 0,001), то ток, поступающий в базу, вызывает появление в эмиттерной и коллекторной цепях тока, в десятки, сотни и тысячи раз большего. Малые изменения силы тока ΔI_B в цепи базы сопровождаются значительными из-

менениями ΔI_K силы тока коллектора.

Усиление по току характеризуется коэффициентом

$$\beta = \frac{\Delta I_K}{\Delta I_B} \quad (16)$$

Из формул (13) и (15) следует

$$\beta = \frac{\alpha}{1 - \alpha} \quad (17)$$

При $\alpha = 0,9$ усиление по току $\beta = 9$, $\alpha = 0,999$ усиление по току $\beta \approx 1000$.

Таким образом, при включении транзистора по схеме с общим эмиттером он действует как **усилитель тока**.

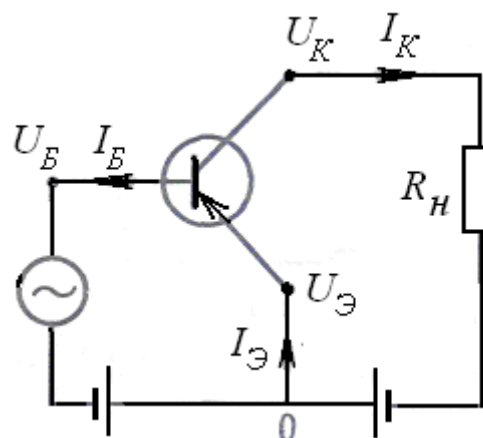


Рис. 42.

Наконец, при включении по *схеме с общим коллектором* (рис. 43) коллектор включен последовательно с входным сигналом. Поэтому входное сопротивление оказывается высоким, выходное - низким. Кроме того, получается значительное усиление по току.

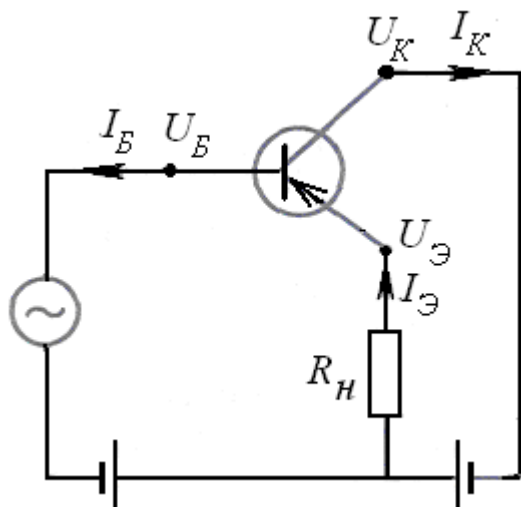


Рис. 43.

Становится понятным термин «биполярный» транзистор. В его работе одинаково важную роль играют носители обоих знаков – и электроны, и дырки. Электроны поступают в *n*- базу *p-n-p* транзистора и вызывают в эмиттерной и коллекторной цепях появление гораздо большего по величине дырочного тока.

В отличие от рассмотренного типа биполярного транзистора так называемый *полевой транзистор* является монополярным. Ток в нем осуществляется либо электронами, либо дырками. Описание полевого транзистора выходит за рамки данного учебного пособия.

стора выходит за рамки данного учебного пособия.

Контрольные вопросы

- 13.1. Почему часть электронов, которые инжектируются в *n-p-n*-транзисторе из эмиттера в базу, приходят к *p-n*-переходу база-коллектор, не успев рекомбинировать с дырками? Какое физическое явление обуславливает движение электронов в базе?
13. 2. Какие требования предъявляются к базе в *p- n- p-* транзисторе?
13. 3. Как изменяется высота потенциальных барьеров при подаче на эмиттер плюса, а на коллектор минуса в *p- n- p-* транзисторе?
13. 4. По аналогии с приведенным описанием работы транзистора *p-n-p*-типа опишите работу транзистора *n-p-n* –типа.
13. 5. За счет какого источника энергии происходит увеличение мощности сигнала в транзисторе?
- 13.6. Какая схема соединения транзистора позволяет получить усиление тока? Усиление напряжения?

§ 14. Понятие о планарной технологии изготовления интегральных схем

Современные средства электронной техники в качестве элементной базы содержат не отдельные диоды и транзисторы, а интегральные схемы (ИС) – совокупность электрически связанных между собой транзисторов, диодов, конденсаторов и резисторов, изготовленных одновременно в едином технологическом цикле.

Все элементы ИС выполняются в тонком (5 – 10 мкм) приповерхностном слое полированной полупроводниковой пластины в результате комбинации процессов легирования, травления, оксидирования, металлизации и др., проводимых с использованием методов литографии.

Один из вариантов создания кремниевого биполярного p-n-p-транзистора в процессе изготовления ИС в упрощенном виде показан на рис. 44.

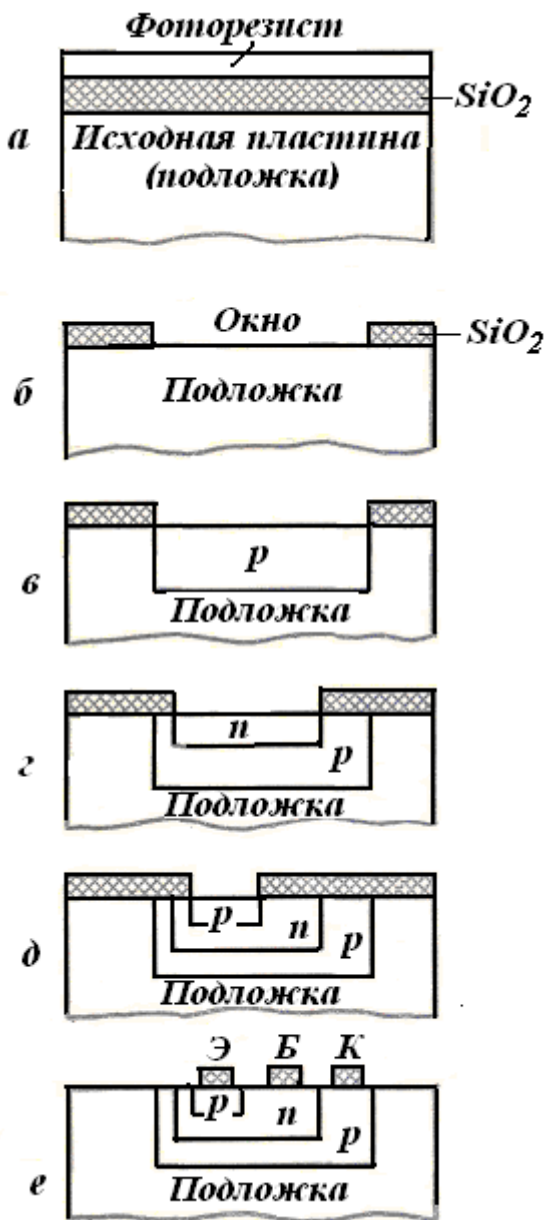


Рис. 44.

защищенными слоем прочной, устойчивой к высокой температуре двуокиси кремния — так называемой *маской*.

в) Теперь через маску в пластину из нагретого газа проводят диффузию бора (В) или какого-либо другого вещества, образующего в Si акцепторную примесь. Там, где поверхность пластины защищена маской SiO_2 , атомы примеси не достигают кремния. Таким образом, примесь попадает в кремний только в окна, т. е. селективно (избирательно).

Так формируются сразу все будущие коллекторы транзисторов.

г) Маска SiO_2 удаляется путем химического травления, и на пластине создается новая маска с окнами, соответствующими расположению базовых областей транзистора. Через эти окна проводится диффузия мышьяка (As) или сурьмы (Sb), являющихся донорами. Концентрация доноров выбирается

а) Поверхность кремниевой пластины окисляется. На ней образуется слой двуокиси кремния SiO_2 . Затем поверх окисла наносится тонкий слой светочувствительного полимера — фоторезиста. Фоторезист обладает способностью изменять свои химические свойства под действием света.

К покрытой фоторезистом полупроводниковой пластине прижимается стеклянная пластинка с нанесенным на ней рисунком, составленным из прозрачных и непрозрачных участков — фотошаблон. После этого фоторезист облучается УФ лучами через фотошаблон. При этом те участки фоторезиста, на которые попадает свет, приобретают способность растворяться в определенных травителях. Эта операция называется *фотолитографией*.

б) Пластина с засвеченным фоторезистом погружается в травитель. При этом происходит растворение участков фоторезиста, на которые попадал свет, а также находящихся под ними участков двуокиси кремния. На полупроводниковой пластине в этих местах открываются области чистой поверхности кремния — так называемые *окна*. Затем в другом травителе, растворяющем незасвеченный фоторезист, но не растворяющем SiO_2 , удаляется оставшийся фоторезист. В результате поверхность пластины представляет собой сочетание окон с участками, защищенными слоем прочной, устойчивой к высокой температуре двуокиси кремния — так называемой *маской*.

большой, чем концентрация акцепторов, так что те. области полупроводника, куда попадает As и Sb, будут областями n -типа.

Так формируются сразу все базы будущих транзисторов,

д) На поверхности пластины формируется новая маска, и новой диффузией акцепторной примеси формируются эмиттерные p -области.

е) На поверхности полученных $p - n - p$ структур также с помощью соответствующих масок и одновременно на всей пластине формируются металлические контакты (чаще всего из алюминия).

Совокупность способов изготовления полупроводниковых приборов и интегральных схем путем формирования их структур только с одной стороны пластины (подложки), вырезанной из монокристалла, получила название *планарной технологии*.

При разработке интегральных схем (ИС), когда на одном кристалле изготавливается готовая схема, содержащая до нескольких сотен тысяч транзисторов размером несколько микрон, планарная технология является единственно возможной.

Контрольные вопросы

14. 1. Каким образом формируется маска из пленки кремнезема?
14. 2. Для каких целей создаются окна в слое SiO_2 ?
14. 3. Позволяет ли планарная технология изготавливать $n - p - n$ транзисторы?
14. 4. Каковы преимущества интегральной схемы по сравнению с дискретными элементами?

§ 15. Генератор незатухающих электромагнитных колебаний на транзисторе.

Появление полупроводникового транзистора произвело самую большую техническую революцию в двадцатом веке, поскольку появилась возможность создавать компактные и надежные элементы компьютеров и другой электронной техники.

Важнейшее устройство электроники – генератор электрических колебаний. Генерирование электрических колебаний осуществляется обычно путем преобразования энергии источников постоянного тока при помощи электронных приборов. Одно из замечательных применений транзистора – генератор высокочастотных синусоидальных (гармонических) колебаний.

Упрощенная схема транзисторного генератора гармонических колебаний с индуктивной обратной связью показана на рис. 45.

Соединяются колебательный контур LC , источник постоянного напряжения U и транзистор. Катушку колебательного контура индуктивно связывают с катушкой $L_{св}$ в цепи эмиттер — база.

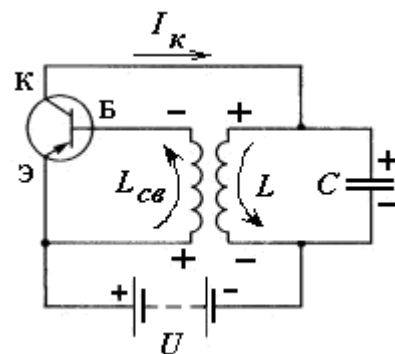


Рис. 45.

Чтобы колебания в колебательном контуре не затухали, необходимо компенсировать потери энергии за каждый период. Пополнять энергию в контуре можно, подзаряжая конденсатор. Для этого надо периодически подключать контур к источнику постоянного напряжения. Конденсатор должен подключаться к источнику только в те промежутки времени, когда присоединенная к положительному полюсу источника пластина конденсатора заряжена положительно. Только в этом случае источник будет подзаряжать конденсатор, пополняя его энергию (рис. 46, верх).

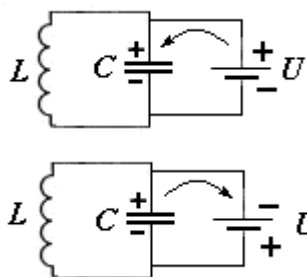


Рис. 46.

Если присоединенная к положительному полюсу источника пластина имеет отрицательный заряд, а присоединенная к отрицательному полюсу — положительный, то конденсатор будет разряжаться через источник (рис. 46, низ). Энергия конденсатора при этом будет убывать.

Как видно из рис. 45, колебательный контур соединен последовательно с источником напряжения и транзистором таким образом, что на эмиттер всегда подается положительный потенциал, а на коллектор — то положительный, то отрицательный. При этом переход эмиттер — база (эмиттерный переход) *всегда является прямым*, а переход база — коллектор (коллекторный переход) оказывается попеременно прямым и обратным.

Чтобы в цепи коллектора возникал ток I_k и подзаряжал конденсатор контура в ходе колебаний, нужно, чтобы оба $p-n$ -перехода были прямыми. Для этого необходимо сообщать базе отрицательный относительно эмиттера потенциал, причем в те интервалы времени, когда верхняя на рисунке пластина конденсатора заряжена

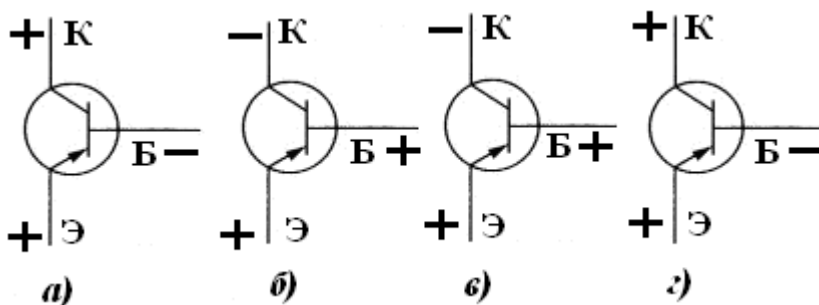


Рис. 47.

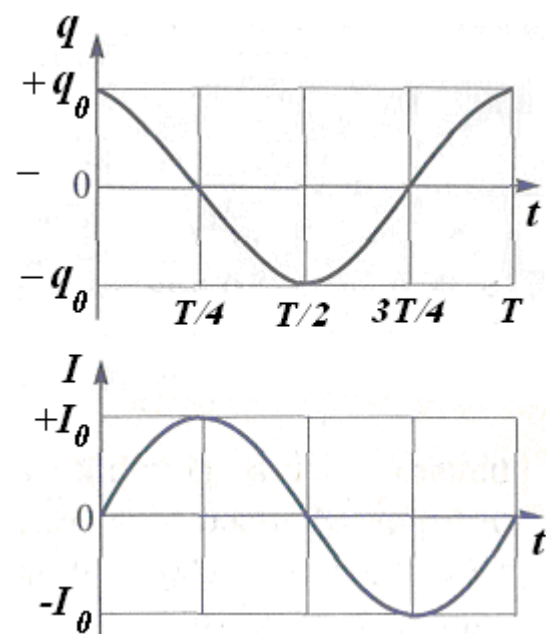


Рис. 48.

положительно, а нижняя — отрицательно. При этом на коллектор подается плюс, и оба перехода открыты (рис. 47, а)).

Как известно, заряд q на обкладках конденсатора и сила тока I , текущего через индуктивность колебательного контура, изменяются по гармоническому закону. При этом в момент времени, когда ток достигает максимального значения, заряд обращается в нуль и наоборот (см. рис. 48).

Рассмотрим последовательные ситуации в системе в каждую четверть периода колебаний.

В течение первой четверти периода положительно заряженная пластина конденсатора, соединенная с коллектором, разряжается. Ток в колебательном контуре возрастает до максимального значения. В катушке связи возникает **индукционный ток** указанного на рис. 35 направления, так что база имеет отрицательный потенциал относительно эмиттера. Переходы база — коллектор и эмиттер — база прямые (рис. 47, а)). Транзистор открыт. Энергия от источника поступает через транзистор в колебательный контур.

Во вторую четверть периода ток в контуре убывает.

Тогда в катушке связи $L_{св}$ ток меняет направление, и на базу подается положительный потенциал. Осуществляется ситуация, изображенная на рис. 47, б). Переход база – коллектор обратный. Тока в цепи нет.

В третью четверть периода конденсатор перезаряжается. Ток растет до максимального значения, направлен от нижней пластины к верхней. В катушке связи ток направлен так, что база получает положительный потенциал. Переход база — коллектор обратный. Тока в цепи нет (рис. 47, в)).

Четвертая четверть периода. Ток в контуре, не меняя направления от нижней пластины к верхней, **убывает**. В катушке связи ток меняется по направлению. Заряд на базе становится отрицательным. Переходы база — коллектор и эмиттер — база прямые (рис. 47 г)). Энергия поступает от источника в колебательный контур.

Таким образом, осуществляются незатухающие электромагнитные колебания за счет поступления энергии от источника в колебательный контур в течение первой и четвертой четвертей периода.

Система, которая сама регулирует поступление энергии в колебательный контур для компенсации потерь энергии, называется автоколебательной. Генератор на транзисторе — пример автоколебательной системы. Необходимым элементом автоколебательной системы является *обратная связь*.

Так называют воздействие результатов какого-либо процесса на его протекание. Колебательный контур управляет источником энергии через цепь обратной связи, содержащую индуктивно связанную с ним катушку $L_{св}$, включенную в цепь эмиттера. Транзистор здесь играет роль вентиля, открывающего доступ энергии источника в колебательный контур. Отпирание транзистора происходит в нужные моменты времени, чтобы импульс тока от источника пополнял энергию колебаний, компенсируя потери на сопротивление катушки и проводов.

Транзисторные генераторы электрических колебаний позволяют получать колебания мощностью от долей мВт до сотен Вт в диапазоне частот от нескольких кГц до десятков ГГц.

Контрольные вопросы

- 15.1. Какую роль в радиотехнике играет генератор незатухающих электромагнитных колебаний? Почему колебания называются почти, а не точно гармоническими?
- 15.2. Каким образом регулируется поступление энергии в колебательный контур?
- 15.3. Какая система называется автоколебательной? Что означает термин «обратная связь»?
- 15.4. Поясните принцип передачи и приема амплитудно-модулированных радиосигналов.

§ 16. Фотоэффект в полупроводниках. Фоторезисторы

Вследствие поглощения света в полупроводниках наблюдается генерация свободных носителей заряда – электронов и дырок, от чего увеличивается электропроводность вещества. Это явление называется *внутренним фотоэффектом или фотопроводимостью*.

Для электрона, получившего дополнительную энергию в результате поглощения фотона, закон сохранения энергии может быть записан в виде

$E_2 - E_1 = h\nu$, где E_1 и E_2 – энергия электрона до и после возбуждения, h – постоянная Планка, ν – частота электромагнитных колебаний.

Если энергия фотона $h\nu$ превышает ширину запрещенной зоны чистого полупроводника ($h\nu > \Delta E$), то поглотивший фотон электрон переходит из валентной зоны в зону проводимости (*собственная фотопроводимость*). В результате появляется дополнительная пара носителей заряда — электрон и дырка (рис. 49).

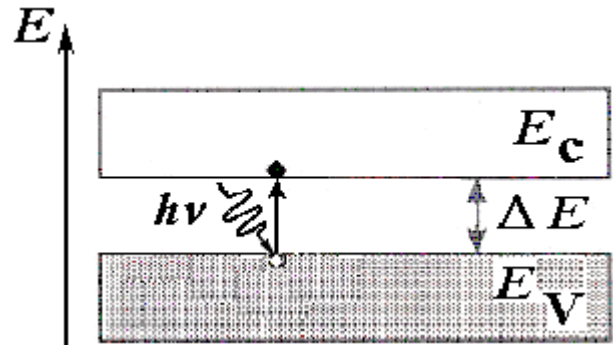


Рис. 49.

В примесных полупроводниках энергия фотона расходуется на перевод электрона с донорного уровня в зону проводимости ($h\nu > \Delta E_D$, рис. 50, а),

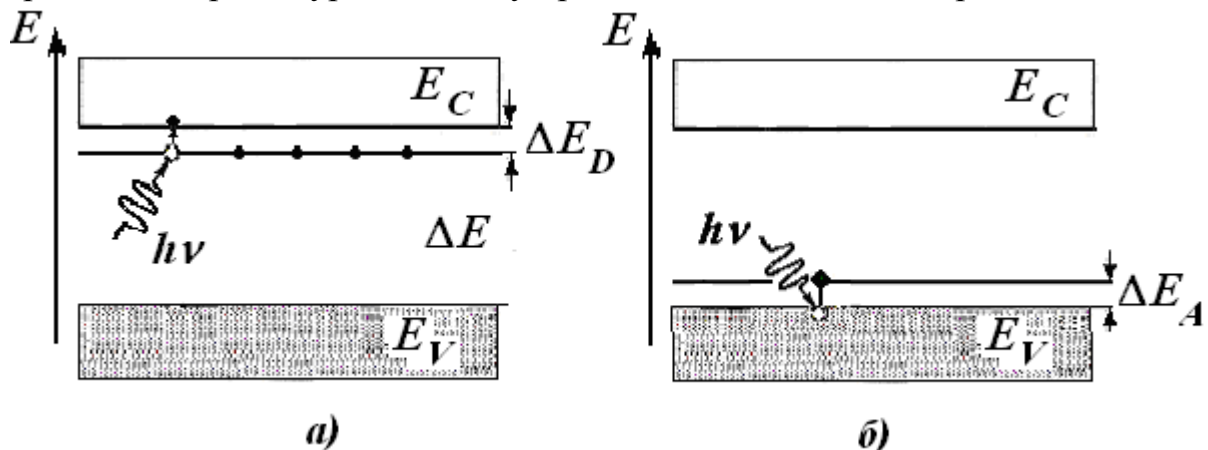


Рис. 50.

э

ле

электронная примесная фотопроводимость), или на перевод электрона из валентной зоны на свободный акцепторный уровень ($h\nu > \Delta E_A$, рис. 50, б), *дырочная примесная фотопроводимость*).

Поскольку количество образующихся носителей заряда пропорционально падающему световому потоку, то явление внутреннего фотоэффекта можно использовать для регистрации оптического излучения.

На внутреннем фотоэффекте основано действие *фотосопротивлений (фоторезисторов)*. Фоторезистор — полупроводниковый прибор, изменяющий величину своего сопротивления при облучении светом.

Обычно полупроводник наносят в виде тонкого слоя на стеклянную или кварцевую подложку или вырезают в виде тонкой пластинки из монокристалла.

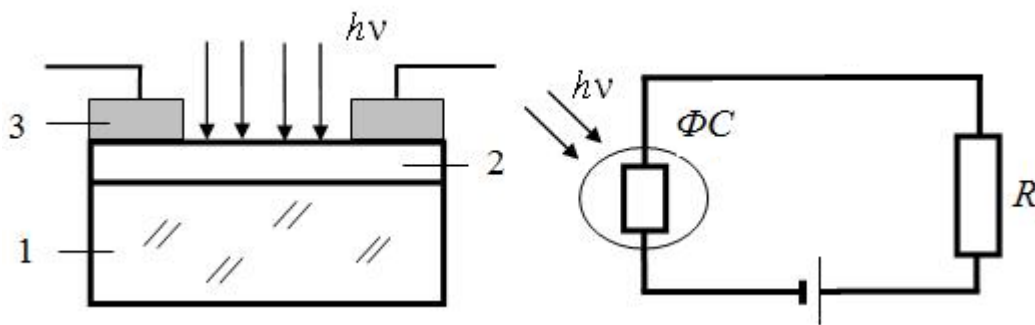


Рис. 51.

Слой или пластинку полупроводника снабжают двумя электродами и помещают в защитный корпус.

Схема

устройства фоторезистора, его условное обозначение на схемах и схема включения в электрическую цепь показана на рис. 51. Здесь 1 - подложка, 2 - пленка полупроводника, 3 - контакты.

Следует отметить, что энергия квантов света, необходимая для возникновения внутреннего фотоэффекта, для примесных полупроводников значительно меньше, чем для чистых, так как $\Delta E_D \ll \Delta E$ и $\Delta E_A \ll \Delta E$ (рис. 50). Это позволяет увеличить максимальную длину волны света, который еще вызывает фотопроводимость:

$$\lambda_0 = ch/\Delta E_D, \quad (18)$$

и использовать собственные полупроводники для регистрации видимого, а примесные полупроводники - инфракрасного излучения, что применяется, например, в приборах ночного видения. В видимой части спектра применяются фотосопротивления, изготовленные из сульфида кадмия (CdS). Фотосопротивления из полупроводников PbS, PbSe, PbTe InSb используются в качестве детекторов инфракрасного излучения.

Пример 16. 1. В передающей телевизионной камере (видиконе) объектив формирует оптическое изображение объекта на поверхности пластинки, покрытой фотопроводящим слоем. Электронный луч "просматривает" поверхность, перемещаясь по ней (сканирует). Вследствие непрерывного сканирования электронным лучом поверхность всегда заряжена. За промежуток времени, который проходит между двумя последовательными попаданиями луча на некоторый участок пластинки, она теряет заряд — разряжается. Скорость разряда зависит от освещенности участка. Это фиксируется электронным лучом, который таким образом считывает потенциальный рельеф поверхности и создает видеосигнал. Электропроводность участков поверхности меняется благодаря явлению внутреннего фотоэффекта. Чем больше освещенность участка изображения, тем меньше сопротивление фотопроводника и тем быстрее происходит его разряд.

Контрольные вопросы

16. 1 Чем отличается внутренний фотоэффект от внешнего (фотоэлектронной эмиссии)?

16. 2. Какие физические явления или свойства полупроводников могут быть использованы в фоторезисторе и фототранзисторе?

16. 3. Указать интервалы длин волн электромагнитного излучения, для которых имеет место внутренний фотоэффект в германии, селене и сером олове. Необходимые данные взять из таблицы 1. Какие из этих интервалов принадлежат видимому свету, инфракрасному и рентгеновскому излучению?

16. 4. Может ли кристалл германия поглощать зеленый свет? Ширина запрещенной зоны германия 0,75 эВ.

16. 5. На фоторезистор на основе селена падает электромагнитное излучение, частота которого: а) $\omega_1 = 3,0 \cdot 10^{15} \text{ с}^{-1}$; б) $\omega_2 = 2,5 \cdot 10^{15} \text{ с}^{-1}$. Изменяется ли при этом его сопротивление?

16. 6. На фоторезистор на основе германия падает электромагнитное излучение, частота которого $\omega = 1,0 \cdot 10^{15} \text{ с}^{-1}$. Изменяется ли при этом его сопротивление? Как изменится ответ, если германий содержит примесь мышьяка?

16. 7. С помощью полупроводникового приемника излучения обнаруживают нагретые тела, расположенные на значительном расстоянии от наблюдателя. В чем состоит принцип действия такого приемника?

§ 17. Фотодиод

Фотодиодом называется полупроводниковый фотоэлектрический приемник оптического излучения, обладающий односторонней проводимостью, которая возникает под действием света. Фотодиод обладает электронно-дырочным переходом ($p-n$ -переходом), смещенном в обратном направлении.

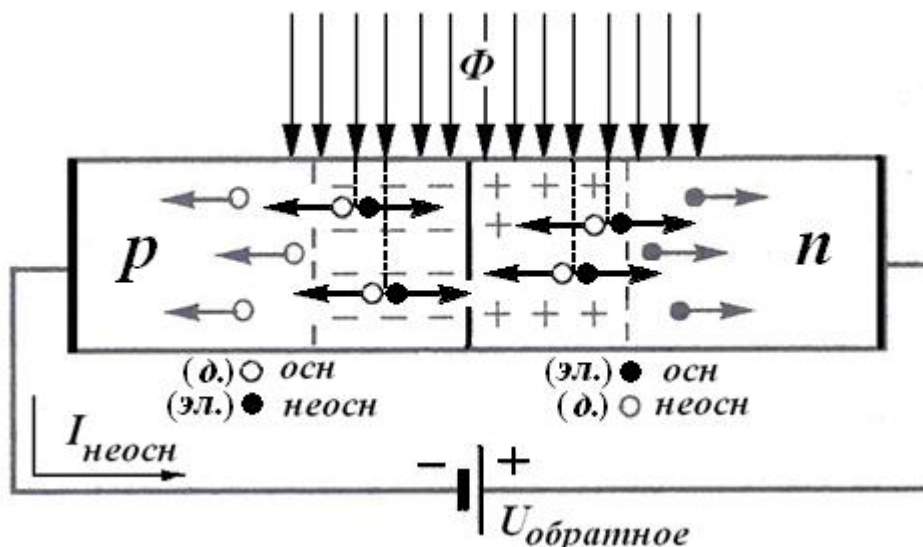
Светочувствительным элементом является переходная область между p - и n -областями (область объемного заряда). Это слой практически полностью обедненный свободными носителями – электронами и дырками.

Различают **два режима** работы фотодиода: **фотодиодный**, когда во внешней цепи фотодиода содержится источник постоянного тока, создающий на $p-n$ -переходе обратное запирающее смещение, и **вентильный**, когда источник тока отсутствует.

В фотодиодном режиме фотодиод, как и фоторезистор, используется для управления электрическим током в цепи в соответствии с изменением интенсивности падающего излучения.

Поскольку внешний источник включен в обратном, запирающем направлении, то при отсутствии освещения ток в цепи практически отсутствует (через фотодиод протекает незначительный обратный ток – ток неосновных носителей $I_{\text{неосн}}$).

Пусть под действием света (Φ – световой поток) в области объемного заряда обратно-смещенного $p-n$ -перехода генерируются новые носители заряда – возникает электрон и дырка (рис. 52). Оба они будут подхвачены электрическим полем и унесены: электрон в n -область, дырка – в p -область.



электрическим полем и унесены: электрон в n -область, дырка – в p -область.

В темноте, без освещения ток через $p-n$ -переход (так называемый темновой ток)

Рис. 52.

определяется числом электронов и дырок, возникающих в области объемного заряда за счет тепловой генерации.

При освещении обратно-смещенного перехода светом с энергией квантов $h\nu > \Delta E$ обратный ток через переход увеличивается. Возникший под действием света фототок во столько раз больше темнового, во сколько раз число носителей, создаваемых в области объемного заряда светом, будет больше, чем возникающих за счет тепловой генерации.

Свойства фотодиода можно охарактеризовать *вольтамперной характеристикой*. Она представляет собой зависимость обратного тока от напряжения при неизменном световом потоке. Вольтамперные характеристики фотодиода приведены на рис 53.

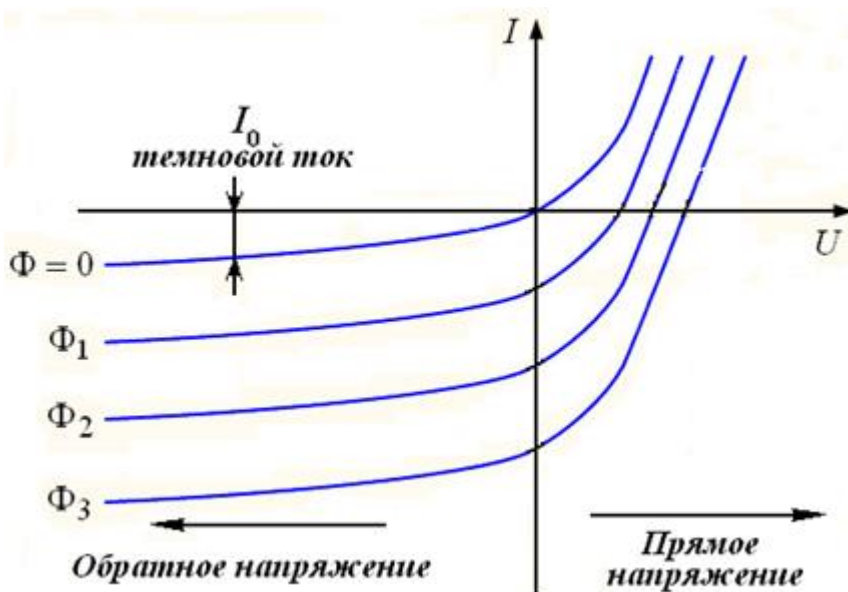


Рис. 53

Если светового потока нет ($\Phi = 0$), то через фотодиод протекает обычный обратный ток (темновой ток I_0).

Под действием светового потока (Φ_1, Φ_2, Φ_3 и т.д.) обратный ток в диоде возрастает, и характеристика располагается ниже. **Чем больше световой поток, тем больше обратный ток.** Таким образом, световой поток управляет обратным током фотодиода.

Следует отметить такую особенность фотодиода. Из образовавшихся в результате поглощения световой энергии носителей не все будут участвовать в образовании обратного тока.

Часть неосновных носителей успевают рекомбинировать, снижая эффективность использования световой энергии. Поэтому целесообразно освещать не обе области полупроводника, а только одну, зато очень тонкую, например, n -область (рис. 54). Ширину n -области делают такой (несколько микрометров), чтобы дырки не успевали рекомбинировать до перехода в p -область. Обычно плоскость $p-n$ -перехода располагается в фотодиодах на расстоянии в несколько десятых долей микрометра от поверхности.

Тогда практически все образовавшиеся под действием освещения неосновные

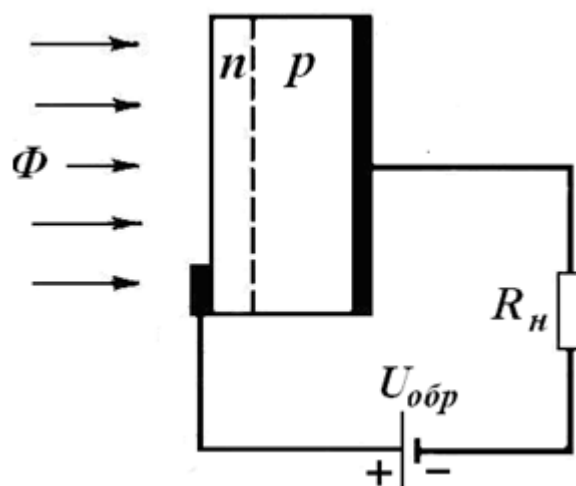


Рис. 54.

носители будут разделены $p-n$ -переходом.

Фотодиод, изготовленный из определенного материала, способен реагировать на свет с длиной волны, лежащей только в определенном диапазоне. Это свойство фотодиодов называют селективностью.

Большим достоинством фотодиодов является их высокая скорость срабатывания (быстродействие). Это позволяет использовать фотодиоды в системах передачи и обработки информации.

Важным и широко применяемым устройством, в котором используются фотодиоды, является матрица.

Светочувствительная матрица – это специализированная аналоговая или цифро-аналоговая интегральная микросхема, массив полупроводниковых светочувствительных элементов – фотодиодов.

С помощью линз объектива цифрового фотоаппарата или видеокамеры изображение объекта фокусируется на матрицу. Матрица преобразует проецированное на неё оптическое изображение в аналоговый электрический сигнал или в поток цифровых данных.

Матрица является основным элементом цифровых фотоаппаратов, современных видео- и телевизионных камер, фотокамер, встроенных в мобильный телефон, камер систем видеонаблюдения и многих других устройств. Она применяется также в оптических детекторах перемещения компьютерных мышей, сканерах штрих-кодов, планшетных и проекционных сканерах, системах астро- и солнечной навигации и т.д..

Контрольные вопросы

17. 1. Какие физические явления или свойства $p-n$ -перехода используются в таких приборах: а) полупроводниковом диоде; б) транзисторе; в) солнечной батарее; г) фотодиоде?
17. 2. Почему $p-n$ -переход фотодиода смещен в обратном направлении?
17. 3. Какой ток – прямой или обратный – возрастает в фотодиоде под действием света?
17. 4. Где применяется светочувствительная матрица?

§ 18. Солнечные элементы

В вентильном режиме, т.е. без приложенного внешнего напряжения, фотодиод (см. § 17) используют в качестве генератора **фотоздс**.

Принцип работы так называемого полупроводникового **солнечного элемента** поясняет схема на рис. 55.

На ней изображен $p-n$ – переход, в котором формируется электриче-

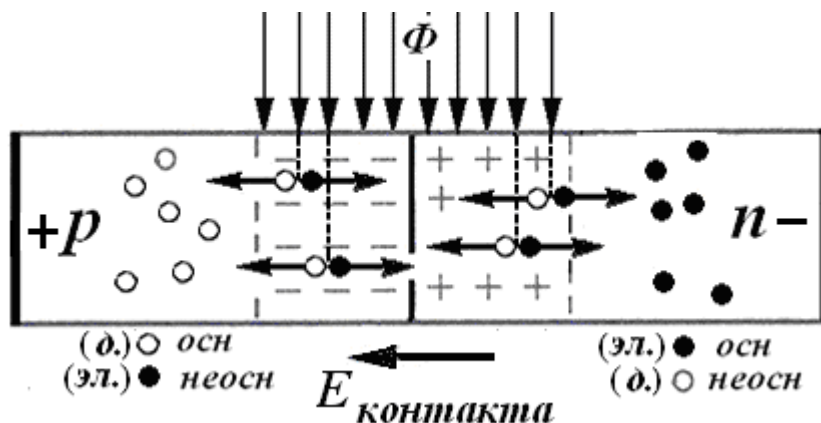


Рис. 55.

ское поле контакта. p -Кристалл заряжается в месте контакта отрицательно, n -кристалл – положительно (см. § 8).

Если на область p - n – перехода направить свет, то электроны будут поглощать фотоны. Каждый такой фотон создаст пару носителей тока – свободный электрон и свободную дырку.

Дырки в p -области являются основными носителями, поэтому поле p - n -перехода будет их отталкивать от границы раздела, и они остаются в своей области. А вот образовавшиеся свободные электроны, являясь в p -области неосновными носителями, легко проходят через переход. Электроны перебрасываются полем через границу раздела в n -область, где они являются основными.

Аналогично, в n -области из образовавшихся носителей «электрон – дырка» только дырки, являясь неосновными носителями, будут переброшены через границу раздела в p -область, а образовавшиеся свободные электроны только пополнят количество основных носителей в n -области, увеличив их концентрацию.

На рис. 56 этот процесс рассмотрен с энергетической точки зрения. Потенциальная энергия дырок в области объемного заряда изображена сплошной линией, потенциальная энергия электронов – штриховой линией.

Неосновные для данной области носители (электроны в p -области и дырки в n -области), возникшие под действием света, легко «скатываются» с потенциального уступа, т.е. беспрепятственно проходят через переход. В результате в p -области накапливается, избыточный положительный заряд, а в n -области— избыточный отрицательный заряд. Это приводит к возникновению разности потенциалов между p -областью и n -областью. Эту разность потенциалов называют **фотоЭДС**.

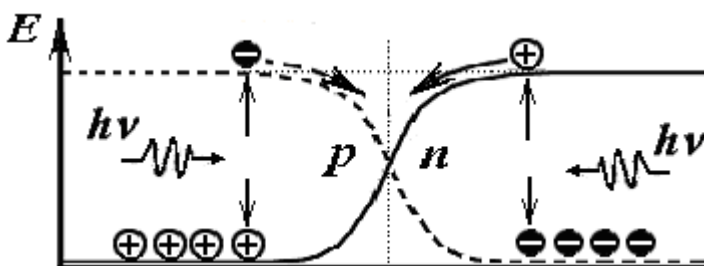


Рис. 56.

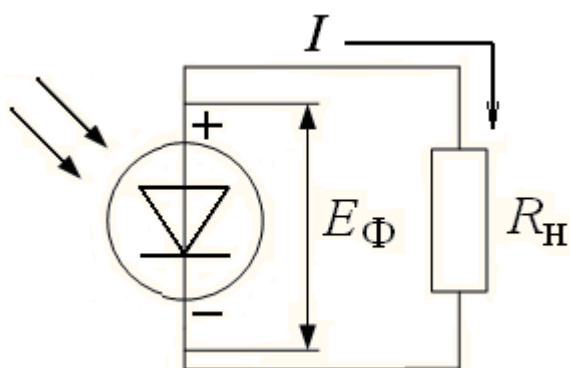


Рис. 57.

Если теперь подключить такой солнечный элемент к внешней нагрузке, то в ней будет течь фототок даже в отсутствие приложенного напряжения (рис. 57).

Полупроводниковые солнечные элементы обычно изготавливают в виде пластины полупроводника p -типа, на которую нанесен тонкий прозрачный слой металла, который можно считать полупроводником n -типа; затем на слой металла наносят прозрачные защитные покрытия.

Световые кванты, пройдя эти покрытия и тонкий слой металла, поглощаются в области p - n – перехода. Ток отводят от полупроводниковой пластины и от тонкого металлического покрытия (рис. 54). Такой элемент обеспечивает

напряжение порядка долей вольта и ток порядка нескольких миллиампер. Обычно элементы соединяют в батарею (солнечная батарея), используя последовательное и параллельное соединение элементов.

Солнечные батареи используются в качестве источников энергии на космических кораблях и спутниках. Солнечные элементы в комплекте с конденсаторами сверхбольшой емкости и яркими светодиодами используются для освещения территорий. В дневное время суток конденсатор накапливает энергию, в темное время суток – отдает.

В настоящее время все большее значение приобретает промышленная технология выработки экологически чистой электроэнергии. Солнечные батареи — один из экологически чистых источников энергии со сравнительно высоким КПД.

В Крыму в 2011 году запустили самую большую в Европе солнечную электростанцию. Она является четвертым в мире фотоэлектрическим парком по показателю мощности. Мощность этой солнечной электростанции составляет пока около 100 МВт (для сравнения – мощность Приднепровской тепловой электростанции 5 000 МВт). Планируется до 2014 года увеличить мощность до 300 МВт.

Станция расположена более чем на 160 га (размер ее как двести футбольных полей!), и состоит из примерно 360 тыс. солнечных модулей.

Несмотря на то, что цена производимой солнечной электростанцией энергии пока в несколько раз выше той, что дают тепловые и атомные электростанции, солнечная электростанция имеет большое будущее. Солнечный свет можно отнести к числу неисчерпаемых ресурсов энергии. И главное достоинство альтернативных источников энергии, начиная от ветряной электростанции и заканчивая солнечной, состоит в том, что они являются экологически чистыми. Например, введенная в строй Крымская солнечная электростанция позволит Украине сократить выбросы углекислого газа до 80 тысяч тонн в год.

Контрольные вопросы

18. 1. Какой режим фотодиода называется вентильным?
18. 2. Поясните механизм генерации электронно-дырочных пар под действием света.
18. 3. Чем обусловлено пространственное разделение возбужденных пар носителей заряда?
18. 4. Интенсивность солнечного излучения на поверхности Земли равна $1,37 \text{ кВт/ м}^2$. Какой должна быть площадь солнечной батареи с КПД = 15 %, чтобы создать ток мощностью 100 Вт?

§ 19. Туннельный эффект в полупроводниках. Туннельный диод

При обсуждении проводимости полупроводников электроны в зоне проводимости рассматривались, в сущности, как «классические» частицы.

Однако не всегда можно пренебрегать квантовыми свойствами электронов. Важную роль играет такой неклассический эффект, как туннелирование электронов, или туннельный эффект.

Полупроводниковый диод, содержащий p - n -переход с очень малой толщиной запирающего слоя, действие которого основано на прохождении электронов сквозь узкий потенциальный барьер, называют *туннельным диодом*.

Как было показано в § 8, контактное поле p - n -перехода создает потенциальный барьер, который препятствует диффузии носителей заряда через переход.

Если частица, имеющая энергию E , движется по оси x по направлению к потенциальному барьеру высотой eU (рис. 58, а), то по классической теории при

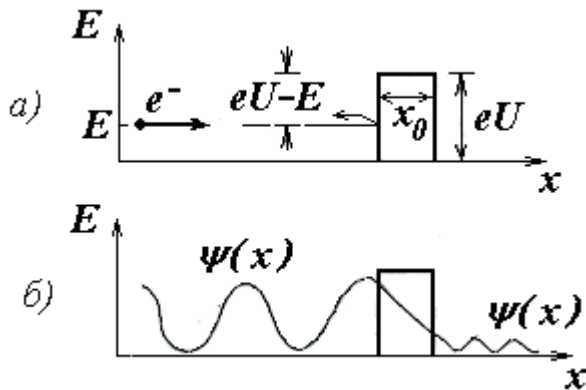


Рис. 58.

$E < eU$ частица не сможет преодолеть потенциальный барьер, поскольку ее энергии недостаточно для этого. Переход через барьер невозможен, так как при этом частица будет иметь отрицательную кинетическую энергию и мнимую скорость, что не имеет смысла.

Однако если учесть, что электроны являются квантовыми частицами и обладают волновыми свойствами, то квантовая механика показывает, что существует определенная вероятность

проникновения частицы через потенциальный барьер.

Это явление существенно квантовой природы, не имеющее аналога в классической механике, называют *туннельным эффектом*.

Состояние электрона описывается волновой функцией. Слева от барьера (рис. 58, б) показана волновая функция электрона, падающего на барьер. Некоторая порция волны с амплитудой, меньшей, чем падающей, появляется справа от барьера.

Квантовая теория позволяет оценить вероятность прохождения через барьер. Эта вероятность оказывается пропорциональной

$$\exp\left(-\frac{x_0}{\hbar} \sqrt{8m(eU - E)}\right). \quad (19)$$

Вероятность тем больше, чем меньше масса микрочастицы, чем уже потенциальный барьер и чем меньше разность между высотой потенциального барьера и энергией частицы.

Важно отметить, что при туннельном эффекте полная энергия туннелировавшей частицы остается неизменной.

Для электрона, туннелирующего через барьер высотой 1 эВ, характерная длина туннелирования имеет порядок 0,1 нм.

Если в p - и n -областях p - n -перехода увеличивать концентрацию примесей (осуществить сильное легирование), то при концентрации электронов и, соответственно, дырок $\sim 10^{19} \div 10^{21} \text{ см}^{-3}$, т. е. приблизительно в 10^3 раз выше, чем в обычных полупроводниковых материалах происходит расширение примес-

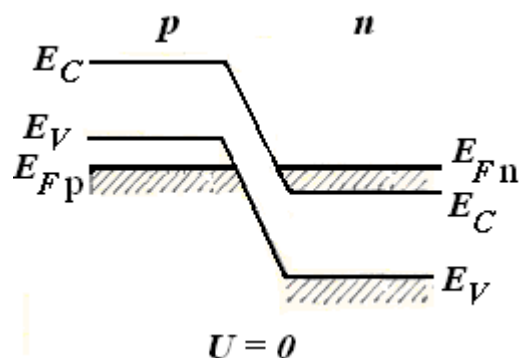


Рис. 59.

ных уровней, и они перекрывают границу между зонами.

В результате этого уровни Ферми в p - области попадают внутрь валентной зоны, а уровни Ферми в n -области попадают внутрь зоны проводимости (рис. 59).

При этом получается резкий p - n -переход с толщиной запирающего слоя 5 – 15 нм (обычная толщина p - n -перехода порядка одного микрометра).

В отсутствие внешнего напряжения ($U = 0$) энергии Ферми по разные стороны перехода одинаковы. Будем считать, что все разрешенные энергетические уровни, расположенные ниже уровня Ферми, заняты, а расположенные выше него – свободны (на рис. 59 это показано штриховкой).

Тогда при $U = 0$ туннельный переход невозможен.

При наложении небольшого прямого напряжения происходит уменьшение высоты потенциального барьера или смещение энергетических уровней p -области относительно энергетических уровней n -области. Все энергии n -области повышаются относительно p -области (рис. 60).

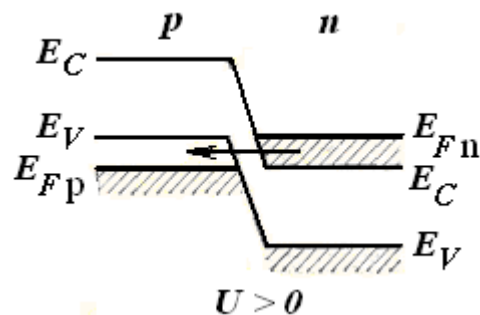


Рис. 60.

В прямом направлении возникает обычный диодный ток. Однако вследствие того, что по разные стороны перехода, разделенного потенциальным барьером, энергии носителей одинаковы, возникает туннельный эффект. В результате туннельного эффекта носители проникают через потенциальный барьер на другую сторону от перехода без изменения энергии. Благодаря этому через переход течет больший, чем диффузионный ток.

При некотором напряжении U_0 при наиболее полном перекрытии зон ток достигает максимума и затем убывает по мере того, как уменьшается степень перекрытия зоны проводимости n -области и валентной зоны p -области. Начиная с некоторого значения U_1 эти зоны не перекрываются, и туннельный ток прекращается. Зоны сдвинуты друг относительно друга настолько, что каждой из них на другой стороне перехода противостоит запрещенная зона (рис. 61). Через p - n -переход течет только диффузионный ток и далее как в обычном диоде проводимость целиком обеспечивается тепловым возбуждением электронов.

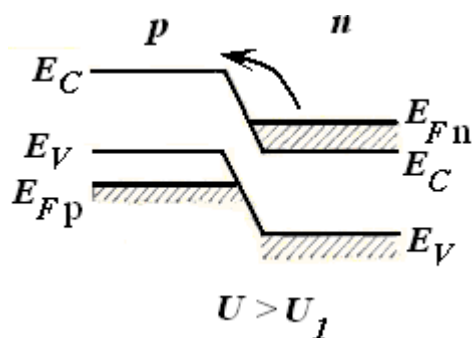


Рис. 61.

Вольтамперная характеристика туннельного диода имеет N -образный вид, показанный на рис. 62.

Обычные диоды при увеличении прямого напряжения монотонно увеличивают пропускаемый ток. В туннельном диоде квантово-механическое туннелирование электронов добавляет горб в вольтамперную характеристику, при этом,

из-за высокой степени легирования p - и n - областей, напряжение пробоя уменьшается практически до нуля.

В интервале напряжений от первого максимума кривой до следующего за ним минимума туннельный диод представляет собой т.н. отрицательное сопротивление. Увеличение напряжения приводит к уменьшению силы тока.

Если использовать туннельный диод в резонансном колебательном контуре генератора электромагнитных колебаний, то его отрицательное сопротивление может компенсировать положительное сопротивление остальных элементов контура, и процессы будут происходить так, как будто контур не имеет сопротивления. Благодаря этому он будет осуществлять колебания тока точно на резонансной частоте, что и используется в высокочастотных усилителях и генераторах.

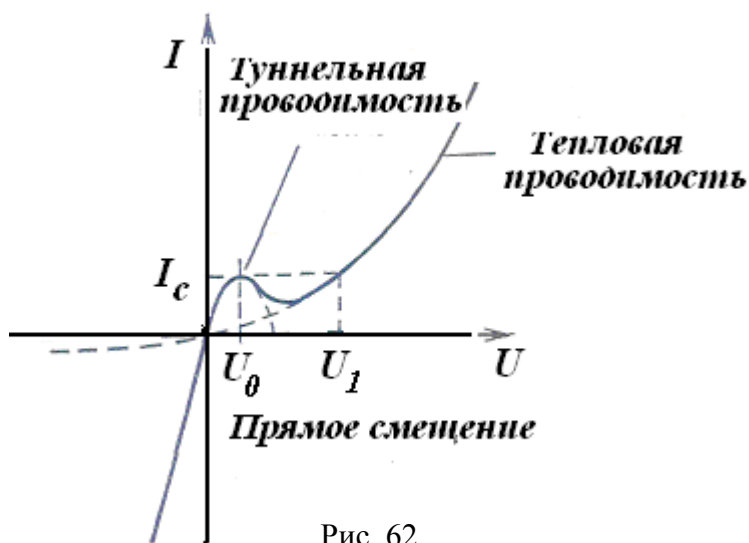
Туннельный диод можно использовать в качестве быстродействующего **триггера** – электронного устройства, обладающего способностью длительно находиться в одном из двух устойчивых состояний и чередовать их под воздействием внешних сигналов.

Даже очень слабый импульс тока через диод будет вызывать мгновенное (время переключения ≈ 1 нс) изменение напряжения на диоде от U_0 до U_1 . Приняв одно из состояний за «1», а другое за «0», можно считать, что триггер хранит (помнит) один разряд числа, записанного в двоичном коде.

Другим примером практического применения туннельного эффекта являются получившие в последнее время большое распространение полупроводниковые (твердотельные) устройства хранения информации, т.н. **флэш-память**.

Флэш-память является разновидностью чипа электрически стираемого программируемого запоминающего устройства. В таком чипе имеется сетка из строк и столбцов, а на каждом пересечении строки и столбца располагается ячейка из двух транзисторов.

Эти два транзистора отделены друг от друга тонким оксидным слоем. Один транзистор называют плавающим затвором, а другой – управляющим затвором. Плавающий затвор может связаться со строкой, или числовой шиной, только через управляющий затвор. Если эта связь имеется, ячейка имеет значение 1. Чтобы поменять значение на 0, нужно чтобы произошло туннелирование электронов.



- 19.1 Что называется потенциальным барьером? Чем обусловлено его появление в p - n -переходе? Как можно регулировать высоту потенциального барьера?
- 19.2. Каким образом достигается уменьшение толщины p - n -перехода?
- 19.3. В чем состоит туннельный эффект в p - n -переходе?
- 19.4. Как изменяется полная энергия электрона при туннелировании?
- 19.5. В чем особенность вольтамперной характеристики туннельного диода?
- 19.6. Поясните с помощью энергетических диаграмм, почему при некотором напряжении ток достигает максимума и затем уменьшается.

§ 20. Полупроводниковый лазер

В отличие от лазеров других типов, в полупроводниковом лазере используются квантовые переходы между разрешенными энергетическими зонами.

Как известно, для генерации лазерного излучения необходимо иметь: 1) активную среду, т.е. вещество с *инверсной населенностью* уровней энергии; 2) оптический резонатор, т.е. систему зеркал, которая создает так называемую *положительную обратную связь*.

Рассмотрим, каким образом можно создать инверсию населенностей, при которой в состоянии с большей энергией находится большее число электронов, чем в состоянии с меньшей энергией. Соответствующий процесс называется *накачкой*.

Наиболее удобным методом накачки полупроводниковых лазеров является использование полупроводникового *диода*, в котором возбуждение происходит за счет тока, протекающего в *прямом направлении*.

Известно, что примесные атомы в полупроводнике вызывают появление дополнительных уровней в запрещенной зоне энергетического спектра полупроводника. При этом атомы, которые легко отдают электроны (доноры), образуют уровни в запрещенной зоне близ дна зоны проводимости E_C , а атомы, которые легко принимают дополнительные электроны (акцепторы), образуют уровни близ верхнего края валентной зоны E_V .

Если примесных атомов очень много, их уровни расширяются в зоны, которые могут сливаться с зоной проводимости или валентной зоной.

В случае сильного легирования (добавление примесей) полупроводников p -типа акцепторные уровни окажутся незаполненными и появляются дырки в валентной зоне, а в полупроводниках n -типа заполнены и донорные уровни, и частично зона проводимости (рис. 63).

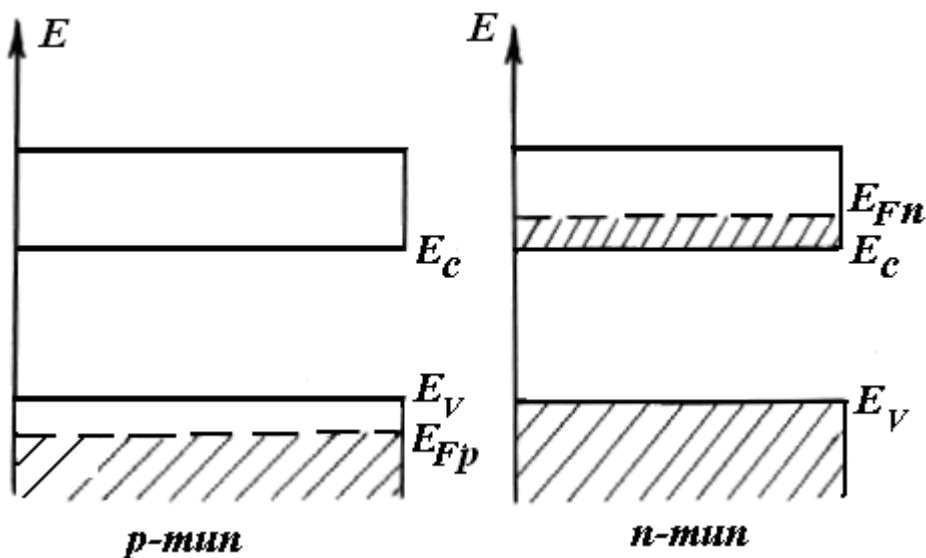


Рис. 63.

На рис. 63 показано, что в сильно легированном полупроводнике уровень Ферми для электронов E_{Fn} лежит внутри зоны проводимости. Соответственно E_{Fp} - уровень Ферми для дырок расположен теперь в валентной зоне. Чем больше электронов попало в зону проводимости, тем выше расположен уровень Ферми. Аналогичную картину наблюдаем для дырок, только в этом случае уровень Ферми для дырок расположен в валентной зоне и с увеличением количества дырок опускается вниз.

Возьмем два таких **сильно легированных** полупроводника p - и n - типов, в которых большинство уровней в верхней части валентной зоны занято дырками, а в нижней части зоны проводимости занято электронами (рис. 63) и соединим их.

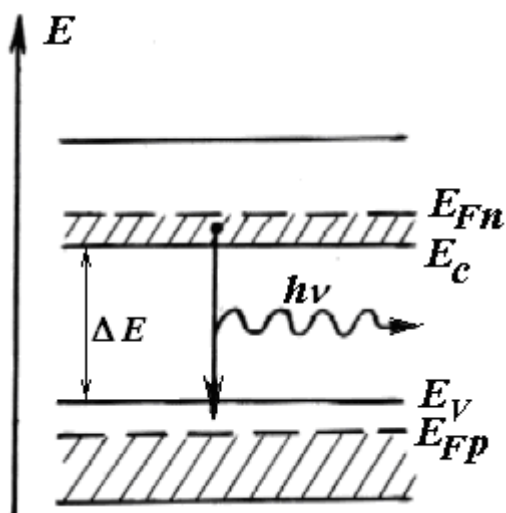


Рис. 64.

На рис. 64 показана энергетическая схема p - n -перехода в момент соединения полупроводников p - и n -типа. Из рисунка видно, что электроны из зоны проводимости могут “упасть” в валентную зону только на свободные уровни, которые лежат в интервале от E_V до E_{Fp} .

Если пропустить через p - n -переход свет, энергия квантов которого лежит в интервале от

$$h\nu_{\max} = E_{Fn} - E_{Fp}$$

до

$$h\nu_{\min} = E_C - E_V = \Delta E,$$

это такие фотоны не могут поглощаться, из-за

того, что не могут вызвать электронных переходов из валентной зоны в зону проводимости (в этом можно убедиться, анализируя возможные квантовые переходы в схеме энергетических уровней, изображенных на рис. 64).

В то же время эти фотоны могут инициировать рекомбинацию (**вынужденную рекомбинацию**). Поэтому условие

$$E_{Fn} - E_{Fp} > \Delta E \tag{20}$$

означает, что расстояние между уровнями Ферми для электронов E_{Fn} и дырок E_{Fp} больше, чем ширина запрещенной зоны ΔE . Оно называется **условием инверсии населенностей**.

Однако, это условие будет выполняться в области перехода только в пер-

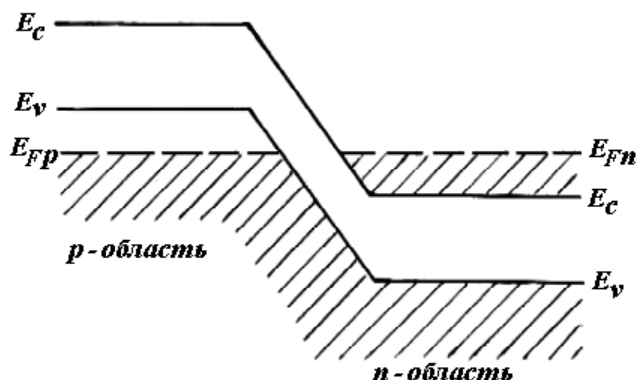


Рис. 65.

вый момент соединения полупроводников. Через очень короткое время в результате рекомбинации электронов и дырок наступит равновесие и уровни Ферми в p - и n -областях совьются (рис. 65), а в области p - n перехода исчезнет инверсия населенностей.

Для того, чтобы снова создать инверсию населенностей в $p-n$ –переходе, то есть добиться, чтобы уровни Ферми в p - и n – областях снова разошлись на расстояние, которое превышает ширину запрещенной зоны, нужно **приложить к переходу напряжение в прямом направлении** (прямое смещение). Теперь уровень Ферми для электронов в n области лежит выше «дна» зоны проводимости, а уровень Ферми для дырок в этой же точке полупроводника лежит ниже «потолка» валентной зоны. Выполняется условие инверсии (20).

На рис. 66 показана энергетическая диаграмма $p-n$ перехода, смещенного в прямом направлении. Как известно, при этом высота потенциального барьера уменьшается и через $p-n$ -переход потечет электрический ток, который состоит из электронов и дырок, движущихся навстречу друг другу. Вертикальными линиями показанная зона инверсной заселенности d – тонкий слой перехода.

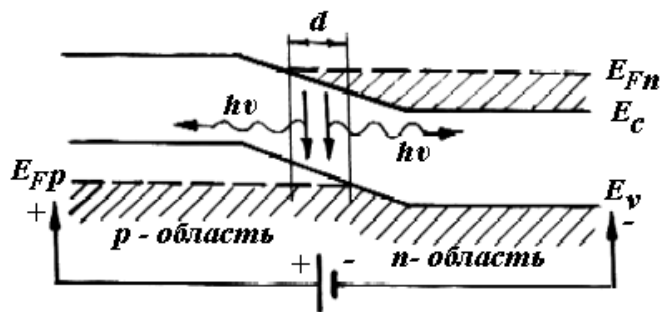


Рис. 66.

Электроны из зоны проводимости n -области могут переходить на лежащие ниже незанятые уровни валентной зоны p - области, т.е. самопроизвольно (спонтанно) «соскакивать» вниз на незанятые уровни валентной зоны. При этом они заполняют пустые места (дырки) в валентной зоне, дырка и электрон проводимости исчезают – рекомбинируют. Потерянная электроном порция энергии может излучаться в виде фотона (рис. 66).

При малых значениях тока, который проходит через переход, рекомбинационное излучение является спонтанным. Такое устройство представляет собой уже рассмотренный **светодиод**.

Кроме спонтанной рекомбинации электронов и дырок в полупроводниках может происходить и **вынужденная рекомбинация под действием света**.

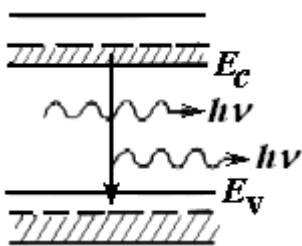


Рис. 67.

На рис. 67 изображено два фотона – тот, который инициировал переход, и тот, что родился вследствие квантового перехода. При этом вторичный фотон является тождественным с фотоном, вызвавшим рекомбинацию. Таким образом, создаются условия для лазерного эффекта усиления и генерации света.

Если теперь выполнить условия, необходимые для получения генерации, светодиод начнет генерировать лазерное излучение.

С увеличением тока через $p-n$ -переход встречные потоки электронов и дырок будут поддерживать в переходном слое концентрацию, которая будет достаточной для выполнения условия инверсии.

Для осуществления положительной обратной связи необходимо также поместить $p-n$ -переход **между параллельными зеркалами**. В полупроводниках можно обойтись без внешних зеркал. Роль зеркал обычно выполняют две отполированные грани самого полупроводникового кристалла, которые перпенди-

кулярны плоскости p - n -перехода и вдоль которых кристалл естественно скалывается. Эти поверхности не нуждаются в дальнейшей обработке.

На рис. 68 приведена схема работы полупроводникового лазера.

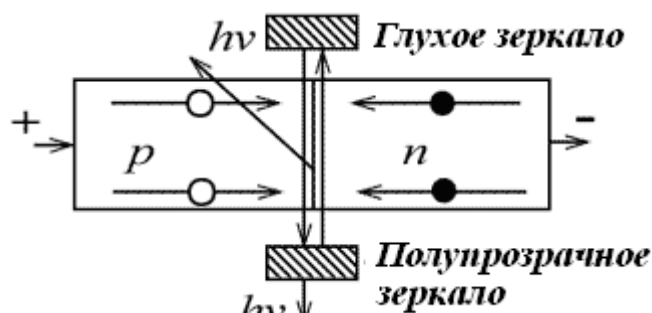


Рис. 68.

Если p - и n -области изготовлены из одного и того же полупроводникового материала (например, арсенида галлия GaAs), такой лазер носит название лазера на *гомопереходе*. Его недостатком является невозможность работы в непрерывном режиме при комнатной температуре, такой лазер нуждается в значительном охлаждении.

Этих недостатков лишены лазеры на так называемом *гетеропереходе*. Гетеропереход формируется между двумя различными по химическому составу полупроводниками. Например, Ge-Si, GaAlAs – GaAs, GaAs – Ge, InGaAs – InP.

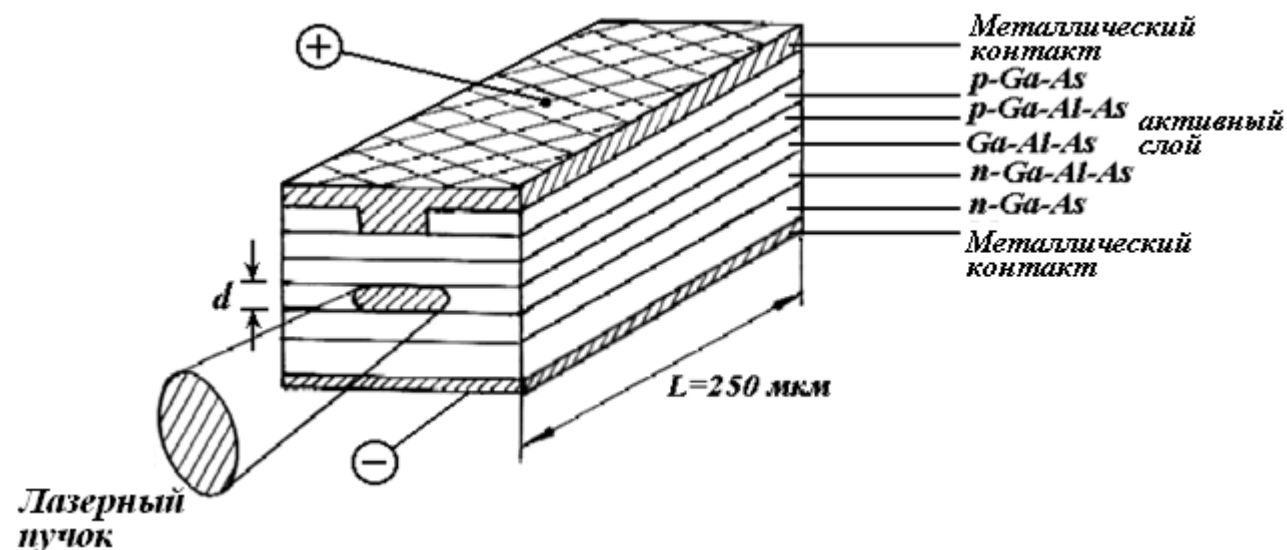


Рис. 69.

На рис. 69 показана одна из возможных конструкций лазера на двойном гетеропереходе.

Полупроводниковый лазер имеет небольшие размеры, они не превышают $1 \div 2$ мм. Переход, из которого выходит излучение, является тонким слоем толщиной несколько микрометров. Лазерный пучок имеет эллиптическое сечение (~ 1 мкм \times 5 мкм). Зеркалами для обеспечения обратной связи и выхода излучения служат грани скола без покрытия.

Излучение полупроводникового лазера легко может быть промодулировано с высокой частотой путем модулирования силы тока, который проходит через переход.

Кроме компактности, особенностями полупроводниковых лазеров является малая инерционность ($\sim 10^{-9}$ с), высокий КПД (до 50 %) и большой выбор веществ для генерации в широком диапазоне длин волн от 0,3 мкм до 30 мкм.

Малые размеры, простота способа возбуждения, высокая надежность сделали полупроводниковые лазеры массовыми приборами. Они широко используются в оптических устройствах записи и чтения информации.

Полупроводниковые лазеры обеспечивают сильный световой поток, достаточный для оплавления полимерных пленок, поэтому используются как оптические головки в компакт-дисковых системах при записи информации. Лазерные диоды используются в лазерных принтерах, системах передачи информации по стекловолоконным кабелям, в лазерных геодезических приборах и т.д.

Контрольные вопросы

20. 1. Чем отличаются полупроводниковые лазеры от лазеров других типов?
20. 2. Чем отличается вынужденная излучательная рекомбинация от спонтанной?
20. 3. Что такое инверсная населенность уровней? Как она создается в полупроводниковом лазере?
20. 4. Какие квантовые переходы обуславливают генерацию вынужденного излучения в полупроводниковом лазере?
20. 5. Что такое положительная обратная связь? Как в кристаллах полупроводников создается система оптического резонатора?
20. 6. Чем отличается полупроводниковый светодиод от полупроводникового лазера?
21. 7. Перечислите особенности полупроводниковых лазеров.

Ответы на контрольные вопросы

2.1. Перекрытием электронных оболочек соседних атомов решетки, между которыми возникает обменное взаимодействие. Вследствие принципа Паули это взаимодействие вызывает расщепление энергетических уровней валентных электронов изолированных атомов и образование энергетических зон.

Кроме того, в квантовой механике показывается, что «запереть» электрон в потенциальной яме конечной глубины и ширины невозможно. Такой потенциальной ямой является для электрона атом.

При сближении атомов для валентных электронов уменьшается ширина энергетического барьера, который разделяет атомы, вследствие чего благодаря туннельному эффекту увеличивается вероятность перехода электрона к другому атому. Электрон уже не связан с каким-либо отдельным атомом решетки, а обобществляется между ними. Время жизни электрона в атоме становится конечным и согласно соотношению неопределенностей для энергии и времени это вызывает расширение уровня и, следовательно, образование энергетической зоны.

2.2. Наличие соседних атомов в кристаллической решетке изменяет потенциальную яму отдельного атома, результирующее кулоновское поле уже не будет удерживать валентные электроны около каких-то конкретных атомов, так что они смогут находиться в кристалле где угодно. Это эквивалентно квантовомеханическому утверждению о перекрытии и распространении на весь кристалл волновых функций валентных электронов.

$$2.3. \quad d \approx \sqrt{\frac{M}{\rho N_A}} = 2,6 \cdot 10^{-10} \text{ м.}$$

3.3. Электропроводность металлов обусловлена наличием в них «свободных» электронов, т.е. электронов, не связанных с конкретным атомом. Эти электроны как бы обобществлены и принадлежат всем атомам металла.

3.4. Чтобы вызвать электрический ток в металлическом образце, помещенном во внешнее электрическое поле, оно должно сообщить электронам ускорение, т.е. произвести над ними некоторую работу. Другими словами, поле должно перевести электроны на более высокие энергетические уровни. Эти уровни должны быть свободными, иначе по принципу Паули переход электронов на них окажется невозможным.

3.7. Электронная конфигурация атома бериллия имеет вид $1s^2 2s^2$. При образовании кристалла бериллия уровни обоих валентных электронов расщепляются и образуется валентная зона, которая содержит $2N$ уровней. Эта зона является полностью заполненной, что соответствует случаю диэлектрика. Однако, в бериллии над валентной зоной существует зона, содержащая вакантные (пустые) уровни, которые образуются при расщеплении уровней $2p$ -электронов. Эта зона перекрывается с валентной, что и предопределяет принадлежность бериллия к металлам.

3.8. Электроны могут двигаться как в проводниках, так и в диэлектриках. Даже при температуре $T = 0$ К в твердых кристаллах валентные электроны, которые могут перемещаться вдоль всего кристалла, обладают огромной кинетической энергией. Различие между проводниками и диэлектриками состоит в характере заполнения зон.

4.1. В классической теории вероятность заполнения энергетических состояний электронами определяется функцией Максвелла-Больцмана $f(E) = \exp(-\frac{E - E_F}{kT})$. При этом в каждом энергетическом состоянии может находиться любое число электронов.

В квантовой теории вероятность заполнения энергетических состояний электронами определяется функцией Ферми $f(E) = (\exp(\frac{E - E_F}{kT}) + 1)^{-1}$. Квантовая статистика базируется на принципе Паули, согласно которому в каждом энергетическом состоянии может находиться только один электрон.

С классической точки зрения энергия всех электронов при температуре абсолютного нуля должна равняться нулю. А по принципу Паули даже при абсолютном нуле число электронов на каждом уровне не может превышать двух. И если общее число свободных электронов в кристалле равно n , то при 0 К они займут $n/2$ наиболее низких энергетических уровней.

4.2. Энергия Ферми представляет собой максимальную энергию, которую могут иметь электроны в металле при абсолютном нуле температуры. $E_F(0) \approx 5$ эВ. Энергия Ферми тем больше, чем больше концентрация электронов.

$$4.3. n = \frac{\rho N_A}{M} = 7,8 \cdot 10^{22} \text{ см}^{-3}$$

4.4. Скорость порядка 10^6 м/с.

4.5. $\langle E \rangle \approx 3$ эВ. До температуры 25 000 К.

$$4.6. E_F = \frac{\pi^2 \hbar^2}{2m_e} \left(\frac{3n}{\pi} \right)^{2/3} = 4,6 \text{ эВ. } \langle E \rangle = 3E_F / 5 = 2,76 \text{ эВ.}$$

4.7. 5,47 эВ.

4.8. 0,02.

5.2. Под дыркой понимают квазичастицу, которая означает, что электрона на данном энергетическом уровне в валентной зоне нет (то есть имеется электронная вакансия). Вне полупроводника дырки не существуют. Зоны проводимости для электронов и дырок не совпадают: для электронов зоной проводимости является свободная зона, а для дырок - валентная зона.

$$5.3. \varepsilon_{\min} = \frac{2kT_1 T_2}{T_2 - T_1} \ln \mu = 1,12 \text{ эВ.}$$

5.4. Электрический ток в металлах представляет собой упорядоченное движение электронов проводимости. В кристаллическом полупроводнике навстречу друг другу под действием электрического поля движутся электроны и дырки.

5.5. У полупроводников удельное сопротивление уменьшается, а у проводников — увеличивается.

$$5.6. \frac{\sigma_1}{\sigma_2} = e^{\frac{\Delta E(T_2 - T_1)}{2kT_1 T_2}} = 84,68.$$

6.1. 0,003 эВ и 0,03 эВ.

6.2. При низких температурах полупроводники становятся диэлектриками, так как в них отсутствуют электроны в зоне проводимости, а энергия теплового движения при таких низких температурах недостаточна, чтобы «перебросить» электроны с уровней валентной зоны на уровни зоны проводимости.

6.7. Если уровень Ферми лежит в запрещенной зоне энергий и удален от края зоны E_C хотя бы на $(2 \div 3) kT$, то в распределении Ферми-Дирака $f(E) = \frac{1}{\exp\left(\frac{E - E_F}{kT}\right) + 1}$ показатель

экспоненты $\frac{E - E_F}{kT}$ принимает значения $\approx 2 \div 3$, а сама экспонента принимает значения 7,39 – 20,09, которые позволяют пренебречь единицей в знаменателе. Тогда распределение Ферми-Дирака переходит в распределение Максвелла-Больцмана $f(E) = C \exp(-E/(kT))$, где C – постоянная. Это значительно упрощает вычисление различных статистических величин.

6.8. $\Delta E = 0,72$ эВ.

7.1. Для увеличения концентрации носителей заряда в собственный полупроводник добавляют специальные примеси. Главное свойство примесного полупроводника состоит в том, что концентрации электронов и дырок в нём резко отличаются друг от друга. Полупроводник, в котором преобладают электроны (основные носители заряда), называется полупроводником n -типа, а полупроводник, в котором преобладают дырки, – полупроводник p -типа. В первом случае преобладают донорные примеси, во втором – акцепторные.

7.2. а) p -типа; б) n -типа.

7.3. При низких температурах основной вклад в электропроводность полупроводника составляют примесные носители заряда. С ростом температуры главную роль играет концентрация носителей собственной проводимости.

7.5. Если в германии или кремний вводят примеси и третьей, и пятой групп Периодической системы, то в запрещенной зоне образуются и донорные, и акцепторные уровни. Поскольку доноры обладают способностью отдавать, а акцепторы — принимать электроны, то будут происходить переходы электронов с донорных уровней на акцепторные. При равной концентрации примесей третьей и пятой групп они взаимно компенсируют друг друга, и такой полупроводник будет обладать собственной электропроводностью.

9.3. За счет притяжения со стороны «плюса» источника электроны покидают p -область и оставляют в ней такое же количество дырок. Возникшие дырки движутся слева направо от металлического контакта в сторону $p - n$ -перехода и инжектируются в n -область. Попав в n -область, они благодаря диффузии движутся вглубь нее и по пути рекомбинируют с проходящими им навстречу электронами.

$$10.6. U = \frac{kT}{e} \ln n = 0,18 \text{ В.}$$

$$10.7. I_0 = I_{обp} \frac{I_{np}}{I_{np} - I_{обp}} = 20,2 \text{ мкА.}$$

$$10.8. n_2 = n_1^{1/(1+n)} = 818,16.$$

12.1. Ток, который проходит сквозь включенный в прямом направлении p - n -переход, при определенных условиях вызывает излучение света. . Вследствие спонтанной рекомбинации инжектированных под действием приложенного напряжения в область p - n -перехода электронов и дырок излучаются кванты В этом состоит явление инжекционной электролюминесценции. Для осуществления этого процесса необходимо создать инверсную населенность уровней энергии в полупроводнике. Тогда частота перехода электронов из зоны проводимости в валентную зону будет большей, чем в обратном направлении.

В отличие от тепловых источников света светодиоды имеют узкий спектр излучения, поэтому воспринимаются как одноцветные (красные, зеленые). Другое преимущество — быстрое действие, которое обусловлено малым временем жизни неосновных носителей .Чаще всего используют GaAs, GaP, SiC с примесями цинка, кислорода (красные) или азота (зеленые).

12.2. Цвет излучения светодиода определяется шириной запрещенной зоны полупроводника и положением донорных и акцепторных уровней примесей.

$$12.3. \Delta E = 2,48 \text{ эВ}$$

$$12.4. \text{ Ближнее ИК- излучение, } \lambda = 871 \text{ нм; зеленый свет, } \lambda = 552 \text{ нм.}$$

$$12.5. \lambda = 398 \text{ нм, фиолетового.}$$

12.7. Явлением полного внутреннего отражения света вот границы раздела кристалл-воздух. Излучение, идущее изнутри наружу и падающее на поверхность раздела полупроводник — воздух под углом, меньшим предельного, выводится из кристалла, а под углом, большим предельного, - испытывает полное внутреннее отражение и идет снова внутрь кристалла.

.Для большинства светодиодов предельный угол довольно мал. Так, для полупроводников GaAs и GaP значения показателя преломления составляют соответственно 3,54 и 3,3, а значения предельного угла $i_{\text{пр}}$ равны примерно 16° и 18° . С целью снижения потерь на полное внутреннее отражение для рассеивателя (обычно из прозрачной эпоксидной смолы) выбирают полусферическую форму.

13.1. Ширина базы делается меньшей диффузионной длины носителей заряда; диффузия.

16.2. Благодаря внутреннему фотоэффекту фоторезистор меняет свое электрическое сопротивление под действием излучения; в фототранзисторе осуществляется управления коллекторным током, при этом световые сигналы превращаются в электрические с одновременным их усилением.

16.3. $\lambda \leq hc/\Delta\varepsilon$; для германия: $\lambda \leq 1,85 \cdot 10^{-6}$ м, для селена $\lambda \leq 6,88 \cdot 10^{-7}$ м, для серого олова $\lambda \leq 1,24 \cdot 10^{-5}$ м. Для германия и серого олова установленный интервал содержит инфракрасные лучи, видимое и рентгеновское излучения, а для селена — видимое и рентгеновское излучение.

16.5. а) изменится; б) нет.

16.6. В первом случае сопротивление не изменяется, во втором — изменяется.

16.7. В использовании внутреннего фотоэффекта: нагретые тела испускают тепловое излучение, в спектре которого есть частоты, которые вызывают в полупроводнике внутренний фотоэффект.

17.1. a, b — нелинейный характер вольт-амперной характеристики p - n -перехода; v — возникновение фотоэдс, обусловленное пространственным разделением пар носителей заряда в электрическом поле p - n -перехода (вентильный фотоэффект); z — рождение избыточных электронов и дырок под действием света, который приводит к росту обратного тока фотодиода.

$$18.4. \approx 0,5 \text{ м}^2.$$

19.4. В проникновении носителей заряда через потенциальный барьер перехода на другую сторону от перехода без изменения энергии.

19.5. Туннельный диод проявляет эффект отрицательного сопротивления.

20.6. Полупроводниковый лазер, как и светодиод, изготавливают из специальных легированных полупроводниковых материалов (т.е. с наличием примесей), которые имеют меньшую по сравнению с обычными полупроводниками запрещенную энергетическую зону. Пример таких материалов – арсенид галлия GaAs и другие. При рекомбинации электронов из зоны проводимости с дырками валентной зоны рождаются фотоны, в большинстве же полупроводников (обычные диоды) выделяемая энергия уходит в кристаллическую решетку (выделение тепла).

Чтобы создать из $p-n$ -перехода лазер, необходимо обеспечить условие инверсной населенности, т.е. ситуацию, при которой свободных носителей заряда – электронов больше в зоне проводимости, чем в валентной зоне. Обычно электроны в зоне проводимости находятся очень недолго и возвращаются в валентную зону с излучением света (самопроизвольная рекомбинация). Таков принцип работы светодиода.

Однако в некоторых полупроводниках существуют метастабильные энергетические уровни в зоне проводимости, где электроны могут находиться достаточно долго. Посредством накачки – т.е. передачи внешней энергии электроном полупроводника (например, путем прямого включением $p-n$ -перехода) возможно накопить электроны на метастабильных энергетических уровнях. Эти электроны затем почти одновременно перейдут в валентную зону (вынужденная рекомбинация), что вызовет когерентное (с одинаковой частотой и с одинаковой фазой) излучение многих фотонов, которые многократно отражаясь от зеркал вызовут лавину фотонов – возникнет излучение лазера. Свет же от светодиода не когерентен и во много раз слабее.

Список рекомендованной литературы

1. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников. Учеб. пособие для вузов– 2-е изд., перераб. и доп. – М.: Наука. Гл. ред. физ.-мат. лит. 1990.– 688с.
2. Савельев, И.В. Курс общей физики Учеб. пособие для вузов. В 5 кн. Кн. 5: Квантовая оптика. Атомная физика. Физика твердого тела. Физика атомного ядра и элементарных частиц / И.В. Савельев. – М.: Астрель; АСТ, 2002. – 368 с.
3. Матвеев А.Н. Атомная физика: Учеб. пособие для студентов вузов.– М.:Высш. шк., 1989.–439 с.
4. Готра З.Ю., Лопатинський І.Є., Лукіянець Б.А. та ін. Фізичні основи електронної техніки: Підручник – Львів. Вид. «Бескид Біт», 2004.– 880 с.
5. Калашников. Н.П. Основы физики: Учеб. для вузов. В 2 т. / Н.П.Калашников, М.А. Смондырев.– 2-е изд., перераб.– М. :Дрофа, 2004.– Т.2 .– 432 с.
6. Робертсон Б. Современная физика в прикладных науках: Пер. с англ. – М. :Мир, 1985.– 272 с.
7. Офир Дж. Физика: Пер. с англ., в 2 т. М. :Мир, 1981.– 288 с. Т. 2.
8. Трофимова Т.И. Курс физики: Учеб. пособие для вузов / Т.И. Трофимова. – 8-е изд. – М. : Высш. шк., 2004. – 544 с.
9. Электроника. Энциклопедический словарь. Гл. ред. В.Г.Колесников. – М. Сов. Энциклопедия, 1991. – 688с