

ВВЕДЕНИЕ В ЗОННУЮ ТЕОРИЮ ТВЕРДЫХ ТЕЛ

ЛЕКЦИЯ № 13

Происхождение энергетических зон в кристаллах. Металлы, диэлектрики и полупроводники в зонной теории. Собственная проводимость полупроводников

Зонная теория твердых тел – квантовая теория энергетического спектра электронов в кристалле, согласно которой этот спектр состоит из чередующихся зон разрешенных и запрещенных энергий. Зонная теория объясняет, в частности, различный характер электропроводности *металлов, полупроводников и диэлектриков*.

§ 1. Происхождение энергетических зон в кристаллах. Металлы

Физически происхождение зонной структуры в кристалле связано с образованием кристалла из N атомов, каждый из которых в свободном состоянии обладает дискретным электронным энергетическим спектром (см. лекцию № 9, § 2).

Мы рассмотрим образование энергетических зон на примере воображаемого процесса образования кристалла лития (щелочной металл) путем последовательного добавления атомов.

На рис. 13.1 изображены энергетические схемы двух изолированных атомов.



Рис. 13.1

Если атомы расположены далеко друг от друга (изолированы), то схемы их энергетических уровней будут совершенно одинаковы: два электрона с различной ориентацией спинов на уровнях $1s$ и по одному электрону на уровнях в $2s$ (см. рис. 13.1).

При сближении двух атомов на расстояние, где их взаимодействие уже нельзя пренебречь, энергетическая схема **должна измениться**: иначе мы приходим в противоречие с принципом Паули. Так, при неизменной энергетической схеме на уровне $1s$ было бы уже по **два** электрона в одном квантовом состоянии: **два** со спином вверх и **два** со спином вниз. Принцип Паули приводит к появлению новых состояний: энергетические уровни **расщепляются** на два подуровня.

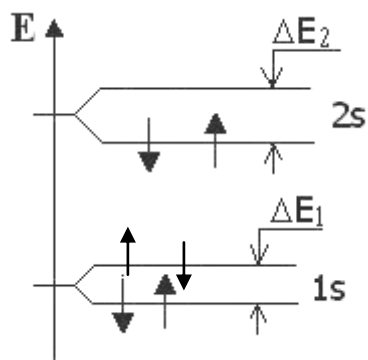


Рис. 13.2

Теперь на подуровнях $1s$, в полном соответствии с принципом Паули, разместились четыре электрона, по одному в каждом квантовом состоянии. На рис. 13.2 изображена энергетическая схема системы из двух атомов.

Обратим внимание на то, что верхний подуровень $2s$ оказался свободным. Величина расщепления уровней ΔE зависит от расстояния между атомами. **При сближении атомов ΔE растет.**

Величина $\Delta E_1 < \Delta E_2$, так как в состоянии $1s$ электроны сильнее связаны с ядром, чем в состоянии $2s$.

Добавим в наш кристалл еще один атом. На рис. 13.3 изображены энергетическая и пространственная схемы системы, состоящей из трёх атомов.

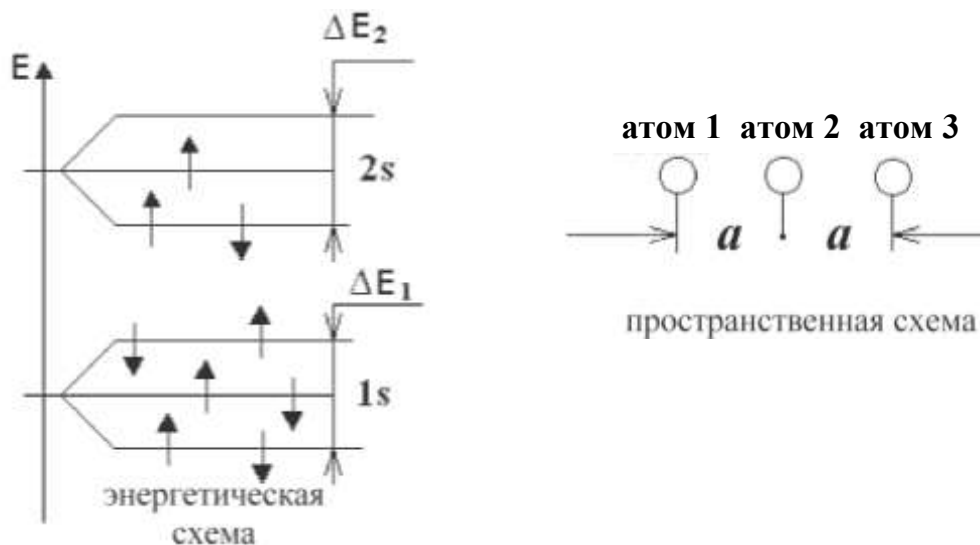


Рис. 13.3

Как видно из пространственной схемы, изображенной на рис. 13.3, минимальное расстояние между атомами – постоянная кристаллической решетки a – осталось неизменным. Значит, величина расщепления (ΔE_1 и ΔE_2) будет *той же самой*, как и для системы из двух атомов (ΔE зависит от *минимального* расстояния между атомами). Третий энергетический уровень расположился *между* двумя крайними.

Продолжая добавлять в нашу систему атомы и рассуждая аналогично, мы придем к выводу, что для системы из N атомов каждый из уровней изолированного атома расщепляется на N подуровней (рис. 13.4). При этом величина расщепления ΔE *не будет зависеть от числа атомов*, так как минимальное расстояние между атомами в кристалле остается неизменным. Следовательно, расстояние между соседними подуровнями будет уменьшаться с ростом N – числа атомов в кристалле. Число атомов N имеет порядок числа Авогадро $N_A = 6,02 \cdot 10^{23}$ 1/моль.

Максимальное расщепление уровней ΔE по порядку величины составляет 1 эВ, значит:

$$\delta E \approx \frac{\Delta E}{N} \sim 10^{-23} \text{ эВ.} \quad (13.1)$$

Энергетическая схема системы, состоящей из N атомов лития, изображена на рис. 13.4.

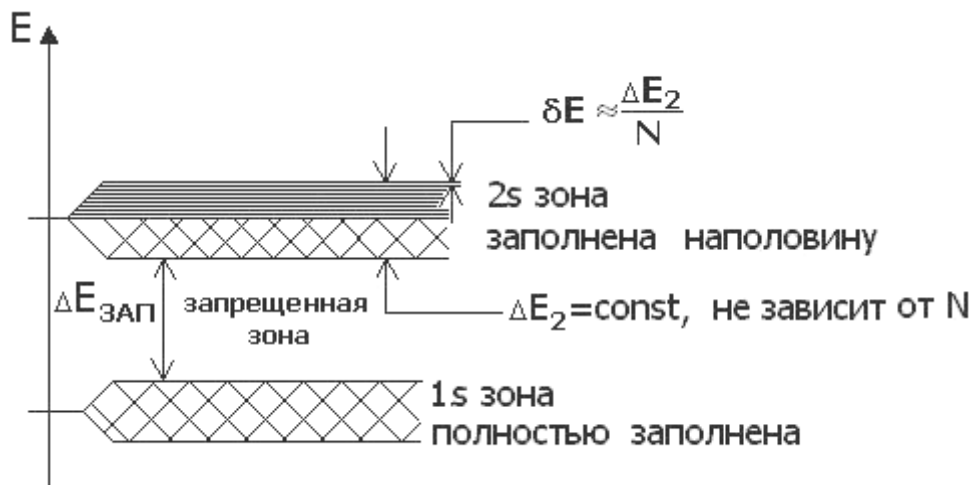


Рис. 13.4

Это очень малая величина по сравнению с величиной kT – добавкой энергии, которую в среднем получает электрон при нагревании. Так, при температуре 1 К $kT \approx 10^{-4}$ эВ, что почти на *два десятка по-*

рядков (!) больше, чем δE . Поэтому можно считать, что энергия в такой системе очень близко расположенных подуровней меняется почти непрерывно (квазинепрерывно). Систему подуровней называют **разрешенной зоной**. В кристалле лития образовалось две разрешенных зоны: из $1s$ и $2s$ уровней изолированного атома. Зона, получившаяся из $1s$ уровня, полностью заполнена. Зона, образовавшаяся из валентного $2s$ уровня, заполнена наполовину (на наших схемах заполнение обозначено штриховкой в клеточку). Между этими зонами находится интервал энергий, запрещенных для электронов: это – **запрещенная зона**, ее ширину мы обозначаем $\Delta E_{\text{зап}}$.

Нетрудно понять, что **кристалл с подобной зонной схемой** (см. рис. 13.4) **будет хорошо проводить электрический ток**: электроны наполовину заполненной зоны могут под действием внешнего электрического поля увеличивать свою энергию квазинепрерывным образом ($\delta E \sim 10^{-23}$ эВ). Увеличение энергии при приложении внешнего электрического поля связано с возникающим упорядоченным движением электронов – электрическим током. Все **металлы** хорошо проводят электрический ток, так как имеют энергетическую схему, подобную только что рассмотренной схеме кристалла лития (см. рис. 13.4).

§ 2. Диэлектрики и полупроводники

Будет ли проводить электрический ток вещество с изображенной на рис. 13.5 зонной схемой? Здесь валентная зона полностью заполнена. Следующая зона свободна.

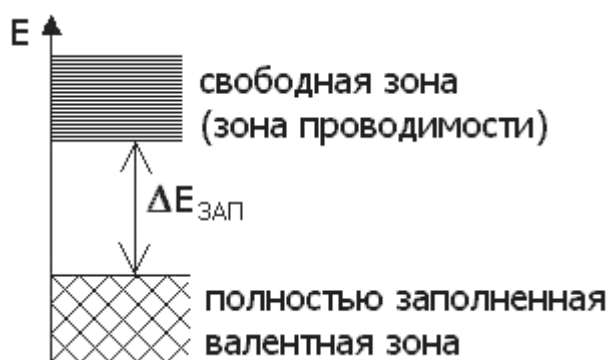


Рис. 13.5

Следующая зона свободна. Проводимость вещества с подобной зонной схемой зависит от ширины запрещенной зоны $\Delta E_{\text{зап}}$ и температуры T .

При $\Delta E_{\text{зап}} \leq 3$ эВ вещество относят к **полупроводникам**, при более широкой запрещенной зоне – к **диэлектрикам** (изоляторам). Резкой границы между этими классами веществ нет. При

$T = 0$ (и отсутствии других внешних воздействий) кристаллы с подобной зонной схемой проводить электрический ток не будут (если $\Delta E_{\text{зап}} \neq 0$). Объясняется это тем, что слабое внешнее электрическое поле не сможет перевести электроны в свободную зону, поэтому, не-

смотря на приложенное внешнее поле, электроны под его воздействием не начнут упорядоченного движения.

У полупроводников $\Delta E_{\text{зап}} \leq 3$ эВ, и при комнатной температуре энергии теплового движения оказывается достаточной, чтобы перевести некоторую малую часть электронов в свободную зону. Там электроны могут увеличивать свою энергию под действием слабого внешнего электрического поля на любую малую величину. Возникает упорядоченное движение зарядов – электрический ток.

Электроны, перебросенные внешним воздействием в свободную зону, называют *электронами проводимости*, а свободная зона называется *зоной проводимости*.

Если $\Delta E_{\text{зап}} = 0$, то мы будем иметь *безщелевой полупроводник*. При больших значениях $\Delta E_{\text{зап}} (\geq 5 \text{ эВ})$ кристалл будет диэлектриком (изолятором). Если полностью заполненная валентная зона частично перекрывается со свободной зоной, то вещество называется *полуметаллом* (рис. 13.6).

В заключение изобразим рядом зонные схемы металла, полупроводника и диэлектрика (рис. 13.7).



Рис. 13.6



Рис. 13.7

§ 3. Собственная проводимость полупроводников

Из элементов таблицы Менделеева типичными полупроводниками являются германий и кремний. Ширина запрещенной зоны у германия 0,66 эВ, у кремния – 1, эВ (при $T = 300\text{ K}$).

Имея по 4 валентных электрона, атомы Ge и Si образуют кристаллические решетки типа алмаза, где каждый атом имеет 4 ближайших соседа, с каждым из которых он связан ковалентной связью.

Условно пространственное расположение атомов в решетке типа алмаза можно представить в виде *плоской* структуры (рис. 13.8, а).

При достаточно высокой температуре тепловое движение способно разорвать некоторые связи, удалив электрон в то место кристалла, где все связи заполнены (рис. 13.8, а). Там электрон будет лишним. Такой электрон в дальнейшем свободно может двигаться по кристаллу, всюду являясь лишним. На рис. 13.8, б) тот же процесс разрыва одной связи изображен на зонной схеме полупроводника: электрон из валентной зоны перешел в свободную зону (зону проводимости для полупроводника). Там, в зоне проводимости, электрон, как мы выше выяснили, может двигаться под действием сколь угодно малого внешнего электрического поля – создавать электрический ток. Отметим, что электрон в зоне проводимости ведет себя как частица с эффективной массой m^* , *не равной* массе электрона ($m^* \neq m_e = 9,1 \cdot 10^{-31}\text{ кг}$).

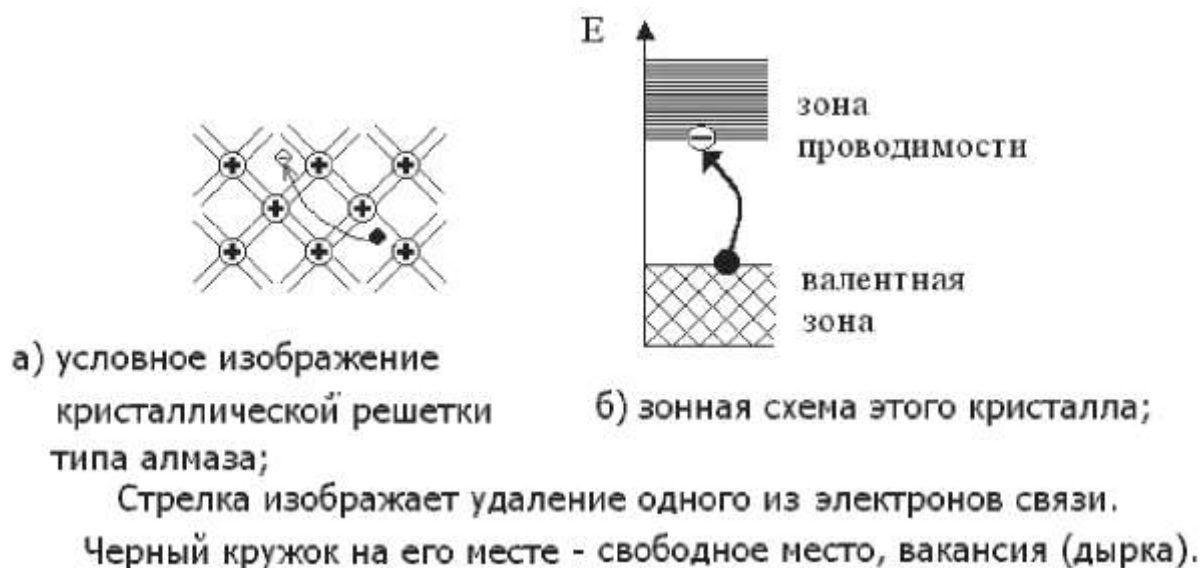


Рис. 13.8

При разрыве одной связи (удаление с нее одного электрона) на месте этой связи останется нескомпенсированный положительный заряд. На зонной схеме эта ситуация изображается освобождением одного состояния в валентной зоне, до этого полностью заполненной. Такие незанятые электронами (вакантные) состояния называют **дырочными** состояниями. Дырки ведут себя как частицы с положительным зарядом, равным заряду электрона.

Во внешнем поле они двигаются в направлении вектора напряженности электрического поля, как частицы с эффективной массой $m^* > 0$.

Если электрон проводимости, блуждая по кристаллу, встретит дырку (частично разорванную связь), то связь заполнится этим электроном. При этом число электронов проводимости уменьшится на единицу, одновременно станет на единицу меньше и число дырок. Этот процесс называется **рекомбинацией носителей**. Он изображен на рис. 13.9.

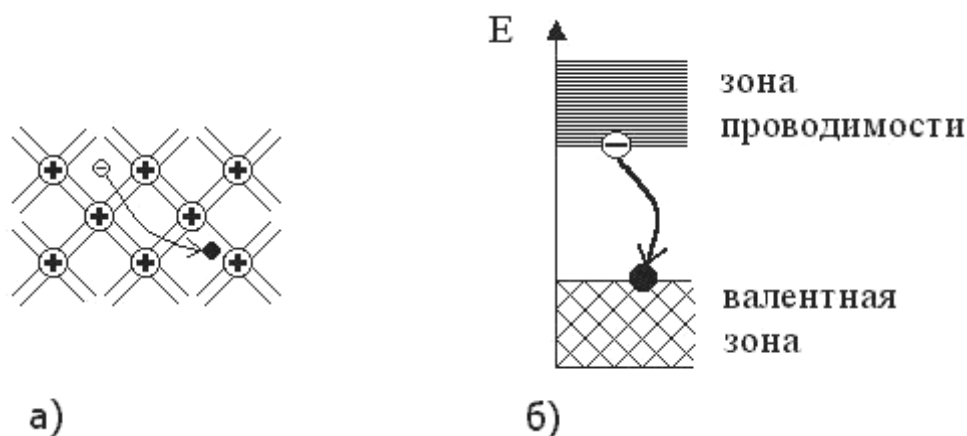


Рис. 13.9

На рис. 13.9, *а* электрон проводимости заполняет незанятое место в ковалентной связи. На зонной схеме (рис. 13.9, *б*) этому процессу соответствует переход электрона из зоны проводимости в вакантное состояние (дырку) валентной зоны.

Как видно из рис. 13.9, *б*, энергия электрона проводимости в процессе рекомбинации уменьшается. Избыток энергии может выделяться в виде излучения (излучательная рекомбинация). Возможна безизлучательная (авторское написание слова. – Прим. ред.) рекомбинация, при которой энергия выделяется в виде колебаний решетки или передается другим электронам проводимости либо дыркам. Излуча-

тельная рекомбинация лежит в основе действия *полупроводниковых лазеров*.

При заданной температуре устанавливается равновесие между процессом образования электронно-дырочных пар и процессом их рекомбинации. Таким образом, устанавливается равновесное для заданной температуры число носителей зарядов.

Проводимость, возникающая за счет переходов под действием температуры электронов идеального кристалла полупроводника из валентной зоны в свободную (зону проводимости), называется **собственной проводимостью** полупроводника.

С ростом температуры растет равновесное число электронов в зоне проводимости и число дырок в валентной зоне. При этом в идеальном кристалле число образовавшихся электронов проводимости равно числу появившихся дырок. Эти электроны и дырки являются носителями тока. Удельная проводимость σ пропорциональна концентрации носителей n (см. (12.2)). Следовательно, удельная проводимость полупроводников будет расти с температурой.

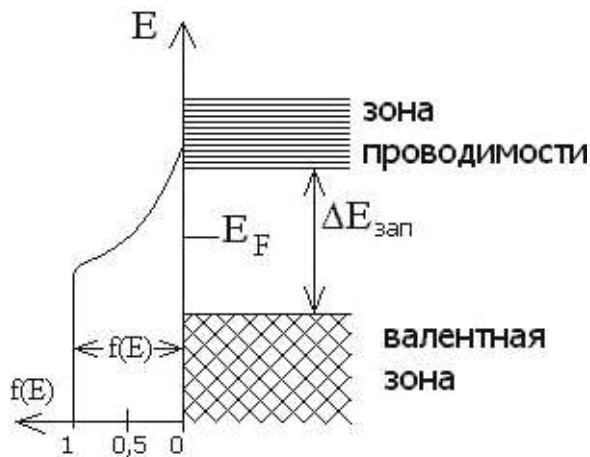


Рис. 13.10

Распределение электронов по уровням зоны проводимости и валентной зоны описывается функцией Ферми – Дирака $f(E)$ (11.4), причем у полупроводников с собственной проводимостью уровень Ферми E_F с большой точностью расположен посередине запрещенной зоны. На рис. 13.10 график $f(E)$ изображен рядом с энергетической схемой полупроводника.

Из рис. 13.10 видно, что для электронов, находящихся у «дна» зоны проводимости, выполняется равенство:

$$E - E_F = \frac{\Delta E_{\text{зап}}}{2}. \quad (13.2)$$

Если $\frac{\Delta E_{\text{зап}}}{2} > kT$, то $e^{(E-E_F)/kT} = e^{\frac{\Delta E_{\text{зап}}}{2kT}} \gg 1$, что обычно и имеет место. Тогда мы имеем дело с «хвостом» функции $f(E)$ (11.5); учитывая (13.2), получим:

$$f(E) \approx e^{-\frac{E-E_F}{kT}} = e^{-\frac{\Delta E_{\text{зап}}}{2kT}}. \quad (13.3)$$

Концентрация электронов проводимости n (и равная ей концентрация дырок) будет пропорциональна $f(E)$, а так как проводимость σ пропорциональна концентрации, то для нее, с учетом (13.3), имеем:

$$\sigma \sim f(E) = e^{-\frac{\Delta E_{\text{зап}}}{2kT}}.$$

Переходя от пропорциональности к равенству, получим формулу зависимости собственной проводимости полупроводников от температуры:

$$\boxed{\sigma(T) = \sigma_0 e^{-\frac{\Delta E_{\text{зап}}}{2kT}}}, \quad (13.4)$$

здесь σ_0 – постоянная величина, имеющая размерность проводимости.

Из полученной формулы видно, что собственная проводимость полупроводников σ экспоненциально быстро растет с температурой.

Изучая на опыте зависимость $\sigma(T)$, можно найти экспериментальное значение ширины запрещенной зоны $\Delta E_{\text{зап}}$.

В полупроводниках и диэлектриках электроны могут перейти из валентной зоны в зону проводимости за счет поглощения фотонов, энергия которых достаточна для обеспечения такого перехода. Энергия фотона ε , как известно, равна $h\nu$. Значит, необходимым условием внутреннего фотоэффекта является неравенство:

$$h\nu \geq \Delta E_{\text{зап}}.$$

В результате внутреннего фотоэффекта возникает **собственная фотопроводимость**. Измеряя граничную частоту $\nu_{\text{кр}}$ (или соответствующую ей длину волны $\lambda_{\text{кр}} = c/\nu_{\text{кр}}$), т. е. определяя **красную границу внутреннего фотоэффекта**, можно найти ширину запрещенной зоны полупроводника или диэлектрика:

$$\Delta E_{\text{зап}} = h\nu_{\text{кр}} = h \frac{c}{\lambda_{\text{кр}}}. \quad (13.5)$$

ИТОГИ ЛЕКЦИИ № 13

1. При объединении атомов в кристалл их энергетические уровни, вследствие принципа Паули, превращаются в систему очень близко расположенных подуровней – *разрешенные энергетические зоны*. Разрешенные зоны могут быть разделены *запрещенными зонами*, электрон не может иметь энергию, лежащую в пределах запрещенной зоны.

2. Если самая верхняя – валентная – зона заполнена наполовину, то она является *зоной проводимости*. Такие кристаллы относятся к *металлам*. Все металлы хорошо проводят электрический ток.

3. Если валентная зона заполнена полностью, а следующая за ней разрешенная зона отделена от валентной широкой запрещенной зоной ($\Delta E_{\text{зап}} \approx 10$ эВ), то такой кристалл будет *диэлектриком*. Диэлектрики почти не проводят электрический ток.

4 *Полупроводники* имеют зонную схему, похожую на зонную схему диэлектриков, но ширина запрещенной зоны у них мала ($\Delta E_{\text{зап}} \leq 3$ эВ). Поэтому при нагревании часть электронов переходит из валентной зоны в свободную зону, которая и является у полупроводников *зоной проводимости*. Проводимость беспримесных полупроводников экспоненциально растет с ростом температуры. Такая проводимость называется *собственной проводимостью* полупроводников.

5. Если электроны в полупроводниках переходят из валентной зоны в зону проводимости за счет поглощения фотонов с энергией $h\nu \geq \Delta E_{\text{зап}}$, то возникает *собственная фотопроводимость*. Это явление называется *внутренним фотоэффектом*.

ЛЕКЦИЯ № 14

*Примесная проводимость полупроводников.
Донорные примеси, полупроводники n-типа.
Акцепторные примеси, полупроводник p-типа.
Электронно-дырочный переход, полупроводниковый диод.
Полупроводниковый триод (транзистор)*

§ 1. Примесная проводимость полупроводников

Некоторые примеси весьма существенно влияют на электрические свойства полупроводников. Так, добавление в кремний (Si) бора (В) в количестве одного атома на 10^5 атомов кремния увеличивает проводимость при комнатной температуре в тысячу раз по сравнению с чистым кремнием.

Донорные примеси, полупроводники n-типа

*Для четырехвалентных полупроводников германия (Ge) и кремния (Si) донорными примесями являются атомы пятивалентных элементов, таких, как фосфор (P), мышьяк (As), сурьма (Sb). Название «донор» происходит от лат. donare – дарить. Каждый атом донорной примеси поставляет один электрон. Примесный полупроводник, в котором носителями заряда являются электроны, заряд которых отрицателен, называется полупроводником **n-типа** (от лат. negativ – отрицательный).*

На рис. 14.1, а изображена схема кристаллической решетки германия (Ge), в которой на месте одного из атомов решетки помещен атом фосфора (P), у которого пять валентных электронов. Четыре из них образуют ковалентные связи с соседними атомами германия, а пятый, донорный, удерживается у положительного иона фосфора слабым кулоновским притяжением, наподобие электрона в атоме водорода.

На рис. 14.1, б изображена энергетическая зонная схема полупроводника с донорной примесью.

На энергетической схеме присутствие донорного электрона изображают, размещая его энергетический уровень на расстоянии E_d от дна зоны проводимости. Для того, чтобы этот электрон перешел в зону проводимости ему, нужно сообщить энергию E_d .

Энергию связи донорного электрона с ионным остатком E_d можно оценить, пользуясь результатами, полученными ранее для атома водорода. Как известно из (4.9), (8.3), энергия связи электрона в основном состоянии атома водорода:

$$E_1 = -\frac{1}{(4\pi\epsilon_0)^2} \cdot \frac{me^4}{2\eta^2}. \quad (14.1)$$

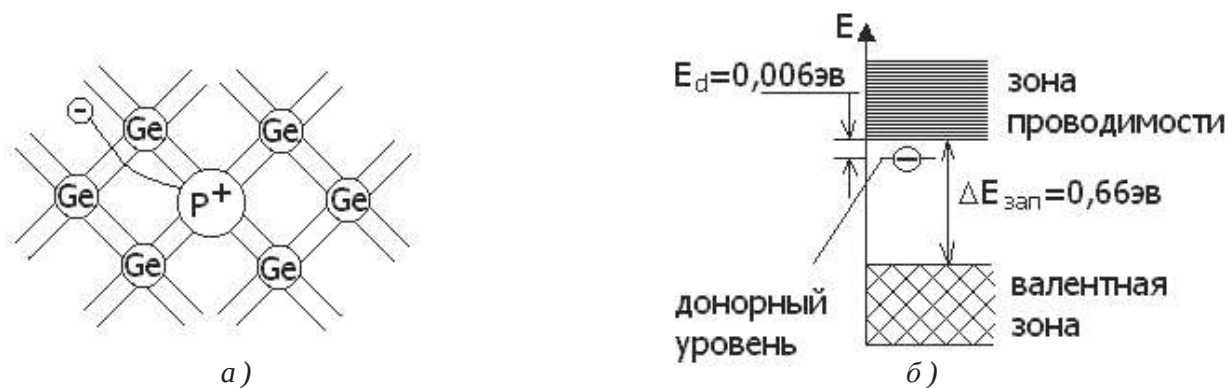


Рис. 14.1

В полупроводнике следует учесть ϵ – диэлектрическую проницаемость среды, для этого надо e^2 заменить на e^2/ϵ . Для германия $\epsilon = 15,8$. Кроме того, масса электрона при его движении в кристалле заменяется на m^* – эффективную массу, для электронов в германии $m^* \approx 0,1 m$. Величина E_1 у атома водорода, как известно, равна $-13,6$ эВ. Следовательно, энергия связи донорного электрона:

$$E_d = \frac{1}{(4\pi\epsilon_0)^2} \cdot \frac{0,1me^4}{\epsilon^2} \cdot \frac{1}{2\eta^2} = |E_1| \frac{0,1}{\epsilon^2} = 0,006 \text{ эВ}. \quad (14.2)$$

Соответствующая оценка для кремния дает $E_d \approx 0,02$ эВ.

При таких значениях энергии связи для перевода электрона с донорного уровня в зону проводимости достаточно энергии теплового движения kT при комнатных и даже более низких температурах. Так, при $E_d = 0,006$ эВ достаточно уже температуры $T \approx 70$ К (или около -200 °С), чтобы kT сравнялось с E_d . Это значит, что при комнатных температурах все электроны с донорных уровней перейдут в зону проводимости: произойдет полная ионизация доноров. *Вследствие полной ионизации доноров, примесная проводимость не будет зави-*

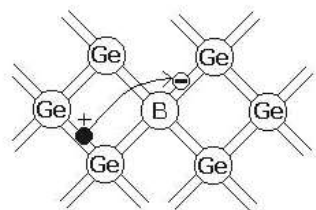
сеть от температуры, а будет определяться только концентрацией примесных атомов.

§ 2. Акцепторные примеси. Полупроводники р-типа

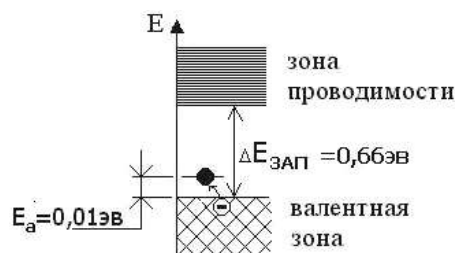
Акцепторными *примесями для германия и кремния являются атомы трехвалентных элементов*, таких, как бор (В), алюминий (Al), галлий (Ga), индий (In).

Название «акцептор» происходит от лат. acceptor – приемник. Каждый атом акцептора забирает из валентной зоны один электрон, создавая в валентной зоне носитель заряда – дырку. Такой примесный полупроводник, в котором носителями заряда являются положительные дырки, называется полупроводником р-типа (от лат. positiv – положительный).

На рис. 14.2, а изображена схема кристаллической решетки германия (Ge) в которой на месте одного из атомов германия помещен атом бора (В), у которого три валентных электрона. На рис. 14.2, б изображена энергетическая зонная схема полупроводника с акцепторной примесью.



а) электрон из объема кристалла германия переходит к примесному трехвалентному атому бора.



б) электрон из валентной зоны переходит на акцепторный уровень.

Черный кружок на рисунках – дырка

Рис. 14.2

Трех валентных электронов, которые имеет атом бора, окажется недостаточно для образования ковалентных связей с четырьмя соседями: одна из связей окажется лишь с одним электроном, полученным от атома германия. На эту незаполненную связь от соседних атомов германия переходит электрон, образуя положительно заряженную дырку на своем прежнем месте, и атом бора, в результате, становится отрицательным ионом. Эта дырка будет связана с отрицательным ионом бора примерно так же, как донорный электрон связан с положительным ионом.

На энергетической схеме (см. рис. 14.2, б) вакантный уровень (с дыркой на нем) мы должны разместить недалеко от «потолка» валентной зоны, его энергия выше «потолка» валентной зоны на величину E_a . За счет теплового движения электрон из валентной зоны может перейти на акцепторный уровень, создав свободную дырку в валентной зоне. На пространственной схеме (см. рис. 14.2, а) этому процессу соответствует возможность удаления положительной дырки от отрицательного иона бора на сколь угодно большое расстояние: происходит ионизация акцептора и переход дырки из связанного состояния в свободное.

Для оценки *энергии связи дырки* E_a (она же – энергия ионизации акцептора) можно использовать те же соображения, что применялись для оценки энергии связи донорного электрона в предыдущем параграфе. Для бора в германии величина $E_a \approx 0,01$ эВ. Так как энергия E_a невелика, то при комнатной температуре $kT > E_a$ и все акцепторы будут ионизированы. Таким образом, как и в случае электронной проводимости, *дырочная проводимость вследствие ионизации акцепторов не будет зависеть от температуры, а будет определяться только концентрацией примесных атомов.*

§ 3. Электронно-дырочный переход. Полупроводниковый диод

Создадим контакт из двух полупроводников, один из которых р-типа, а другой п-типа, как это изображено на рис. 14.3. Такой контакт называют электронно-дырочным переходом, или р-п переходом.



Схематическое изображение р-п перехода.

Рис. 14.3

Предположим для удобства рассмотрения, что контакт создан приведением в соприкосновение двух образцов полупроводника: р- и п-типа. В первый момент обе части созданного перехода будут электрически нейтральны.

В материале р-типа имеются свободные дырки, причем их концентрация равна концентрации отрицательно ионизированных примесных акцепторных атомов. В материале п-типа, справа от перехода, имеются свободные электроны. Их концентрация равна концентрации положительно заряженных примесных донорных атомов.

Кроме примесных носителей, в полупроводнике всегда присутствует некоторое количество собственных носителей. Их концентрация при комнатной температуре мала по сравнению с концентрацией примесных носителей, поэтому их называют неосновными носителями.

Таким образом, в р-области концентрация дырок велика, а в n-области мала. С электронами дело обстоит наоборот, их концентрация велика в n-области, а в р-области мала. За счет различия концентраций возникают диффузионные потоки (см. ч. 4, лекция № 6, § 3).

Дырки из р-области будут двигаться в n-область, одновременно электроны из n-области будут диффундировать в область р.

Возникшие потоки зарядов приведут к нарушению электрической нейтральности. В р-области останутся нескомпенсированные отрицательно заряженные ионы акцепторных атомов. В n-области будет избыток положительно заряженных ионов донорных атомов. В результате образуется двойной слой разноименных зарядов, которые создадут электрическое поле, направленное от n-области к р-области, как это изображено на рис. 14.4.

Возникшее поле будет препятствовать диффузионным потокам. Установится равновесное распределение носителей в области р-п перехода. В области двойного электрического слоя электроны и дырки, двигаясь навстречу друг другу, рекомбинируют, в результате р-п переход оказывается обедненным носителями, проводимость его становится маленькой.

Полупроводниковый диод — прибор, обладающий способностью хорошо пропускать через себя ток одного направления и плохо — противоположного направления. Полупроводниковый диод представляет собой полупроводниковую пластину с двумя областями различной проводимости: электронной (n-типа) и дырочной (р-типа). Между ними возникает р-п переход, который и обладает односторонней проводимостью.

Подадим на р-п переход разность потенциалов $\Delta\phi$ (рис. 14.5).

На рис. 14.5, а р-п переход, смещенный в обратном направлении (к области р подан отрицательный потенциал, к области n — положи-

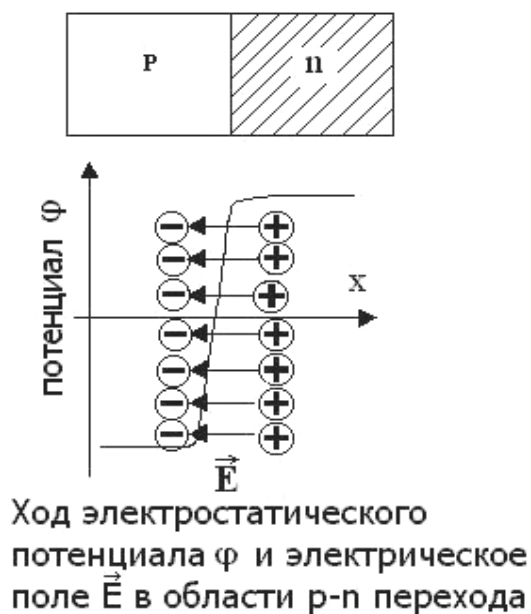


Рис. 14.4

тельный), ток через переход почти отсутствует. На рис. 14.5, б р-п переход смещен в прямом направлении (к области р подан положительный потенциал, к области п – отрицательный), в этом случае ток резко растет с ростом разности потенциалов на р-п переходе. Происходит это по следующим причинам.

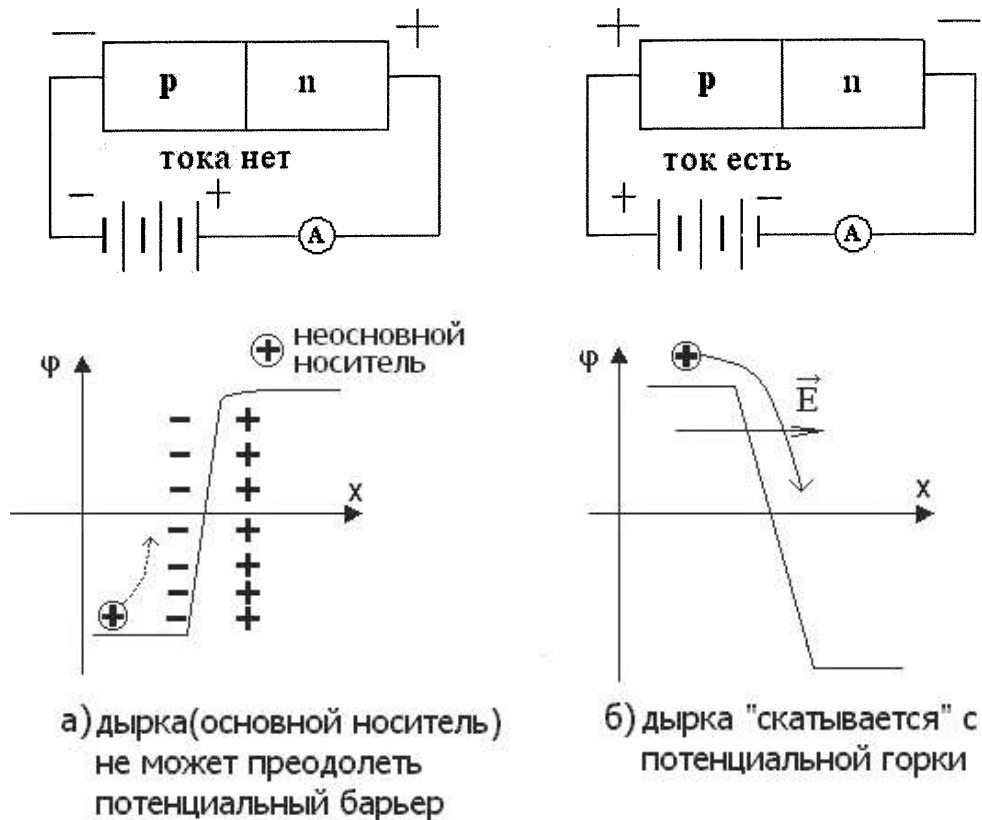


Рис. 14.5

Если отрицательный полюс источника напряжения соединен с р-областью, а положительный – с п-областью (см. рис. 14.5, а), то высота потенциального барьера для основных носителей возрастет. Иными словами, усилится электрическое поле, препятствующее движению основных носителей через р-п переход. В этом случае под действием внешнего поля через переход смогут двигаться только неосновные носители (на рис. 14.5, а) в п-области изображена дырка, которая может «скатиться» с «потенциальной горки»). Следовательно, через р-п переход при обратном смещении будет течь только слабый ток неосновных носителей.

Теперь соединим положительный полюс источника с р-областью, а отрицательный – с п-областью.

В этом случае внешнее поле будет направлено в сторону, противоположную полю двойного электрического слоя. Величина потен-

циального барьера будет меньше, чем при отсутствии внешнего поля. При достаточно большой положительной внешней разности потенциалов барьер превратится в «горку» для основных носителей. Дырки из р-области будут под действием внешнего поля переходить в область n, а электроны из n-области – в область р. Возникает ток основных носителей через р-п переход, он будет экспоненциально возрастать с ростом положительной разности потенциалов.

Зависимость тока от напряжения (разности потенциалов) называют **вольт-амперной характеристикой** для полупроводникового диода. Вольт-амперная характеристика изображена на рис. 14.6.

При отрицательном напряжении течет очень маленький ток неосновных носителей. Если отрицательное напряжение больше чем $U_{пр}$ – возникает электрический пробой, через переход течет большой отрицательный ток.

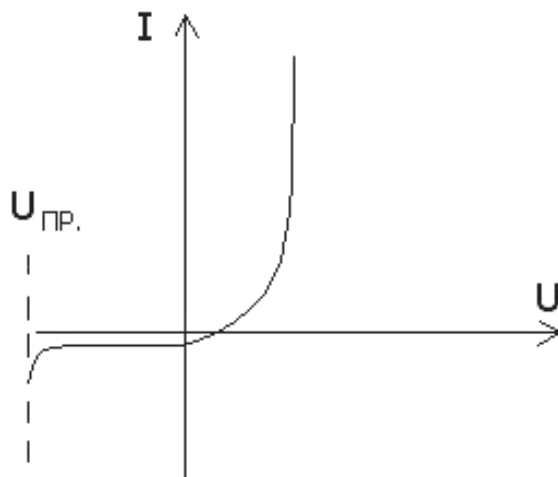


Рис. 14.6

§ 4. Полупроводниковый триод (транзистор)

Полупроводниковый триод, или транзистор, – это электронный прибор, предназначенный для усиления, генерирования и преобразования электрических сигналов. Состоит он из двух р-п переходов, созданных в одном кристалле.

В зависимости от чередования переходов, различают р-п-р и n-р-п транзисторы. Средняя часть триода называется базой. Толщина ее должна быть, по возможности, меньше. Области с противоположным типом проводимости, прилегающие к базе, называют эмиттером и коллектором. Конструктивно коллектор имеет больший объем, чем эмиттер.

Рассмотрим **принцип работы транзистора** на примере схемы, изо-

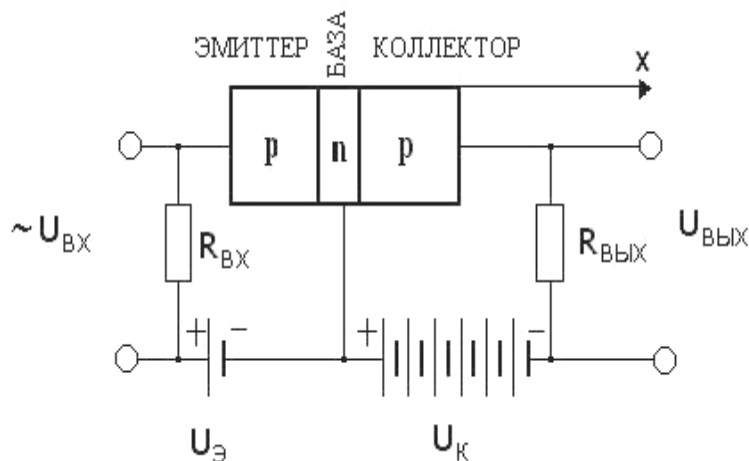


Рис. 14.7

браженной на рис. 14.7 (схема с общей базой).

На переход «эмиттер – база» подаются небольшое постоянное смещение $U_э$ в прямом направлении и усиливаемый переменный сигнал. Переход «база – коллектор» смещается в обратном направлении, значительно большем, чем $U_э$ напряжением $U_к$. При таких смещениях сопротивление перехода «эмиттер – база» невелико, сопротивление перехода «база – коллектор» велико. Это позволяет взять в качестве нагрузки большое сопротивление $R_{вых}$.

На рис. 14.8 изображены графики потенциала в зависимости от координаты x в направлении, перпендикулярном плоскостям p-n и n-p переходов (см. рис. 14.7).

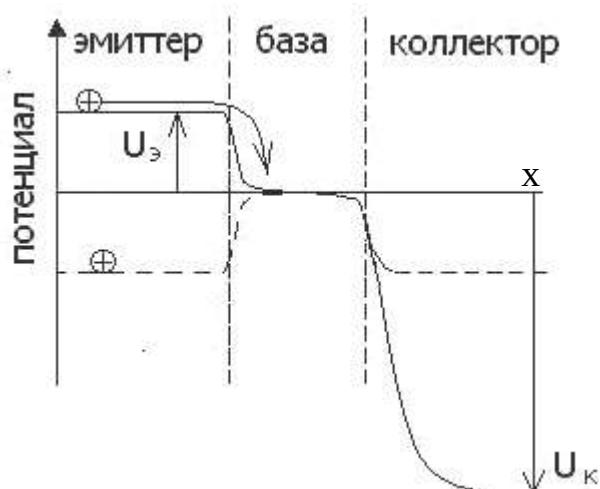


Рис. 14.8

В случае отсутствия смещения двойной электрический слой, как мы узнали выше, препятствует движению основных носителей через p-n переход. При прямом смещении перехода «эмиттер – база» величина «барьера» уменьшается и «барьер» может превратиться в «горку», с которой будут «скатываться» основные носители (см. рис. 14.5, б).

Так, дырки из эмиттера (у нас – p-область) будут в большом количестве переходить в область базы (n-область в нашем случае). Если база достаточно тонкая, то большая часть пришедших из эмиттера дырок за счет диффузии дойдет до перехода «база – коллектор», не успев рекомбинировать. А здесь для них, дырок, приготовлена потенциальная «горка», с которой они «скатываются» в область коллектора. У хорошего транзистора до 99 % (и больше) основных носителей, вышедших из эмиттера, доходят до области коллектора. Можно считать, что ток коллектора $I_к$ примерно равен току эмиттера $I_э$. При изменении тока эмиттера, вызванном входным сигналом, настолько же изменится и ток коллектора. При этом мощность выходного сигнала будет больше, чем у входного, так как разность потенциалов на переходе «база – коллектор» больше, чем на переходе «эмиттер – база», а электрическая мощность, как известно, равна произведению тока на напряжение:

$$P = IU.$$

Таким образом, рассмотренная нами схема с общей базой усиливает сигнал по мощности.

ИТОГИ ЛЕКЦИИ № 14

1. Атомы пятивалентных элементов, таких, как фосфор (P), мышьяк (As), сурьма (Sb), добавленные в кристаллическую решетку четырехвалентных полупроводников германия (Ge) или кремния (Si), называются *донорными примесями*.

2. Каждый атом донорной примеси может поставить в зону проводимости один электрон. Полупроводник с донорной примесью называется полупроводником n-типа, так как носителями заряда в этом случае являются электроны, заряд которых отрицателен (от лат. negativ – отрицательный).

3. Энергия связи донорного электрона с ионным остатком $\sim 10^{-2}$ эВ, поэтому при комнатных температурах все донорные электроны переходят в зону проводимости (полная ионизация доноров). Вследствие этого примесная электронная проводимость не зависит от температуры, а определяется только концентрацией доноров.

4. Атомы трехвалентных элементов, таких, как бор (B), алюминий (Al), галлий (Ga), индий (In), добавленные в кристаллическую решетку четырехвалентных полупроводников германия (Ge) и кремния (Si), называются *акцепторными примесями*.

5. Каждый атом акцептора может забрать из валентной зоны один электрон, создавая в ней носитель положительного заряда – дырку. Такой примесный полупроводник называется полупроводником p-типа (от лат. positiv – положительный).

6. Энергия, необходимая для ионизации акцептора невелика ($\sim 10^{-2}$ эВ), поэтому уже при комнатных температурах все акцепторы будут ионизированы. Вследствие этого, дырочная проводимость не будет зависеть от температуры, а определяется только концентрацией акцепторов.

7. Контакт из двух примесных полупроводников с разным типом проводимости называется *p-n переходом*. Такой переход обладает односторонней проводимостью. На основе свойств p-n перехода работает *полупроводниковый диод*.

8. Прибор, состоящий из двух p-n переходов, созданных в одном кристалле, называется *полупроводниковым триодом, или транзистором*. Транзистор используется для усиления, генерирования и преобразования электрических сигналов.

ОСНОВЫ ФИЗИКИ ЛАЗЕРОВ

ЛЕКЦИЯ № 15

§ 1. Вводные сведения

Лазер (оптический квантовый генератор) – *устройство, генерирующее когерентные электромагнитные волны за счет вынужденного испускания света активной средой, находящейся в оптическом резонаторе.*

Термин «лазер» происходит от первых букв английского названия этого устройства: Light Amplification by Stimulated Emission of Radiation – усиление света за счет вынужденного испускания излучения.

Первые квантовые генераторы были созданы в 1953 г. советскими физиками Н.Г. Басовым и А.М. Прохоровым и независимо от них американским ученым Таунсом. Всем троим в 1964 г. за эти работы присуждена Нобелевская премия по физике. Квантовые генераторы, созданные Басовым, Прохоровым и Таунсом, работали в микроволновом диапазоне, и их английское название «мазер» образовано по тому же принципу, что и термин «лазер», только вместо слова Light (свет) используется слово Microwave (микроволновое излучение). Первый квантовый генератор, работающий в оптическом диапазоне, – рубиновый лазер – был создан в 1960 г. Т. Мейманом (США).

Лазер содержит **три основных компонента**:

- 1) **активная среда**, в которой создают *инверсию населенности*;
- 2) **система накачки** – устройство для создания инверсии населенности;
- 3) устройство положительной обратной связи – **оптический резонатор**.

Главными процессами, приводящими к лазерному излучению, являются:

- 1) **вынужденное излучение**;
- 2) **положительная обратная связь**.

§ 2. Вынужденное (стимулированное) излучение

В соответствии со вторым постулатом Бора, излучение испускается или поглощается в виде квантов энергии $\eta\omega$ при переходе электрона в атоме из одного стационарного состояния в другое. При этом:

$$\eta\omega = E_n - E_m.$$

Если $E_n < E_m$ атом поглощает фотон $\eta\omega$, при $E_n > E_m$ атом излучает фотон.

А. Эйнштейн в 1916 г. показал, что этих двух процессов недостаточно для установления состояния равновесия между излучением и веществом. Он высказал идею о существовании вынужденного излучения.

Вынужденное излучение возникает, если на атом, находящийся в возбужденном состоянии, например, с энергией E_2 , воздействует фотон с частотой, удовлетворяющей условию: $\eta\omega = E_2 - E_1$ (E_1 – энергия основного состояния). В этом случае с определенной вероятностью атом переходит из состояния E_2 в основное состояние E_1 и излучает *еще один фотон* с энергией $\eta\omega = E_2 - E_1$. При одинаковой интенсивности излучения вероятность единичного акта вынужденного излучения равна вероятности перехода атома из состояния E_1 в состояние E_2 , происходящего за счет поглощения фотона $\eta\omega$. На рис. 15.1 схематически изображены процессы: а) спонтанного излучения, б) вынужденного излучения и в) поглощения фотонов.

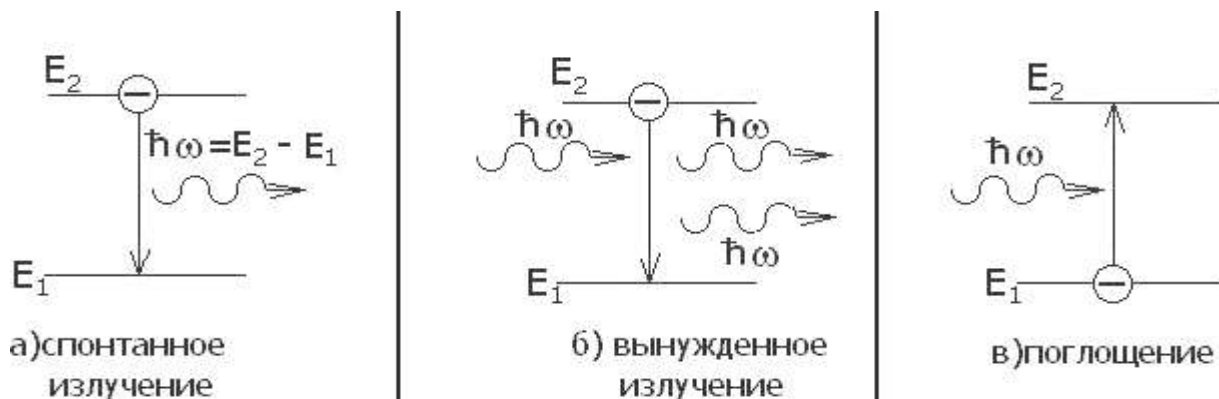


Рис. 15.1

Фотон вынужденного излучения по своим характеристикам *тождественен* фотону, вызывающему вынужденное излучение: он имеет те же самые частоту, фазу, поляризацию и направление распространения. На волновом языке можно сказать, что вынужденное излучение приводит к увеличению амплитуды электромагнитной волны без изменения ее частоты, фазы, поляризации и направления распространения. Следовательно, *вынужденное излучение строго когерентно с вынуждающим излучением*.

§ 3. Состояние с инверсией населенности

Пусть, для простоты рассуждений, атомы среды могут находиться только в двух энергетических состояниях: E_1 и E_2 . Числа атомов N_1 и N_2 , находящихся в этих состояниях, в условиях термодинамического равновесия определяются распределением Максвелла – Больцмана (см. ч. 4, (2.14)). Их отношение дается следующей формулой:

$$\frac{N_2}{N_1} = e^{-\frac{E_2 - E_1}{kT}} < 1.$$

Как отмечено выше, вероятность единичного акта вынужденного испускания фотона равна вероятности поглощения такого же фотона. Поэтому, при одинаковой интенсивности излучения, **полная вероятность поглощения будет пропорциональна N_1 , а полная вероятность вынужденного испускания – пропорциональна N_2** . В равновесных условиях $N_2 < N_1$, и поэтому процесс поглощения будет преобладать над процессом вынужденного испускания: вещество будет поглощать электромагнитные волны.

Если создать каким-либо образом в **активной среде** неравновесное состояние с $N_2 > N_1$ – **состояние с инверсией населенности** – то процесс вынужденного испускания будет преобладать над процессом поглощения.

Схематически этот процесс изображен на рис. 15.2.

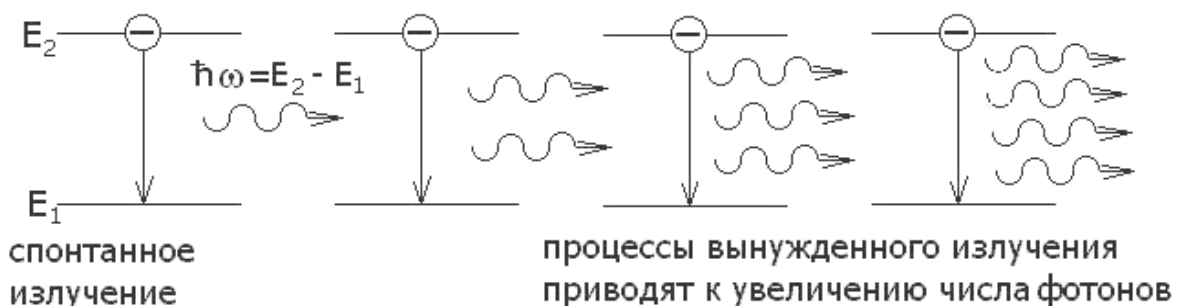


Рис. 15.2

Появившийся за счет спонтанного излучения фотон с энергией $\hbar\omega = E_2 - E_1$ «размножается» за счет последовательных процессов вынужденного излучения. Возникает вспышка света, интенсивность которой зависит от размеров активной среды вдоль направления движения фотонов. Эта вспышка света называется **сверхлюминесценцией**.

§ 4. Оптический резонатор

Для превращения сверхлюминесценции в генерацию лазерного излучения необходимо наличие **положительной обратной связи**, осуществляемой за счет **оптического резонатора**. Простейший оптический резонатор состоит из двух плоских зеркал, расположенных параллельно друг другу (рис. 15.3).

Между этими зеркалами расположена активная среда, находящаяся в состоянии с инверсией населенности.



Рис. 15.3

Фотоны, испущенные под углом к оси резонатора, порождают лишь короткие вспышки сверхлюминесценции. Фотоны, испущенные вдоль оси резонатора, многократно проходят через активную среду за счет отражений от зеркал. При этом они вызывают вынужденное излучение, которое увеличивает интенсивность световой волны. В наиболее благоприятных условиях находятся те фотоны, для которых выполняются условия возникновения *стоячих волн*.

Стоячие электромагнитные волны в оптическом резонаторе возникают при выполнении условия, аналогичного условию возникновения стоячих волн в струне, закрепленной с двух концов (ч. 3, лекция № 6, § 5), т. е.:

$$l = n \frac{\lambda_n}{2},$$

здесь l – расстояние между зеркалами; n – целое число; λ_n – резонансная длина волны.

В оптическом резонаторе $l \gg \lambda$, поэтому целое число n принимает большие значения. При этом в пределах естественной ширины спектральной линии оказываются несколько резонансных длин волн – несколько **мод оптического резонатора**. Эту ситуацию иллюстрирует рис. 15.4, где изображена зависимость интенсивности излучения I от частоты ω для активной среды, находящейся в оптическом резонаторе.

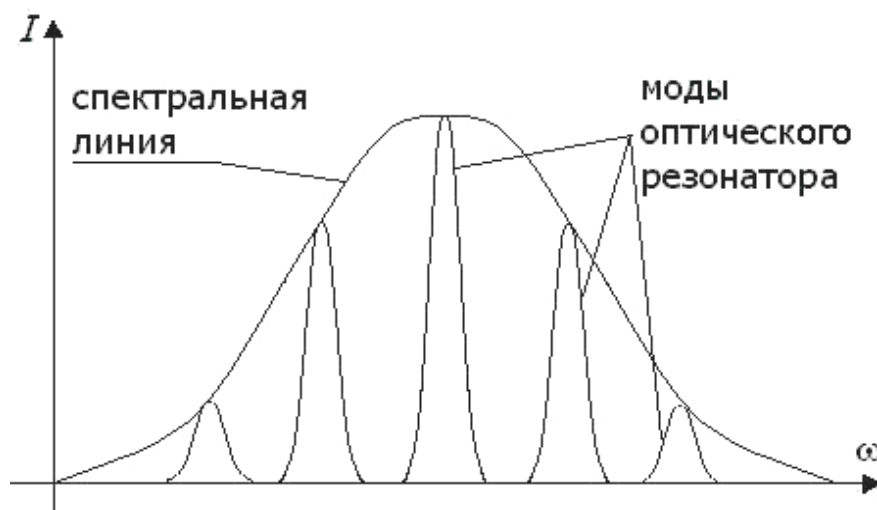


Рис. 15.4

Существуют способы выделения одной из мод и подавления остальных. Такой *одномодовый режим* генерации позволяет достичь наивысшей когерентности излучения.

Таким образом, оптический резонатор формирует лазерное излучение,

направленное строго по оси резонатора и обладающее высокой степенью когерентности.

Для того, чтобы выпустить лазерное излучение из резонатора, одно из зеркал делают полупрозрачным.

§ 5. Способы создания инверсии населенности

Процесс создания инверсии населенности называется **накачкой**. В зависимости от структуры активной среды, используются различные виды накачки.

В твердых телах и жидкостях используют **оптическую накачку**. В этом случае для создания инверсии населенности активной среды необходимо, по крайней мере, *три энергетических уровня* атомов или молекул активной среды. Такая трехуровневая схема накачки была реализована в первом твердотельном лазере, созданном в 1960 г. Т. Мейманом (США). Активной средой в этом лазере являлся кристалл рубина, отполированные торцы которого служили зеркалами

оптического резонатора. Один торец покрывался непрозрачным слоем серебра, слой серебра на другом торце пропускал 8 % упавшей на него энергии.

Рубин представляет собой окись алюминия (Al_2O_3), в которой 0,03–0,05 % атомов алюминия заменены трехвалентными ионами хрома Cr^{3+} .

На рис. 15.5 изображена энергетическая схема иона хрома, который и является основным элементом активной среды.

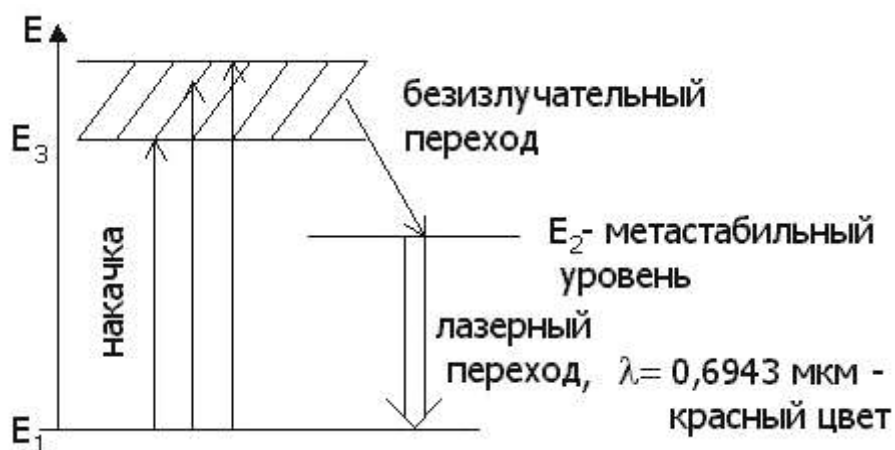


Рис. 15.5

Перевод электронов с уровня E_1 на уровни широкой полосы E_3 (накачка) осуществляется за счет интенсивного облучения рубина некогерентным светом мощной импульсной лампы. В возбужденном состоянии E_3 ионы хрома проводят около 10^{-7} с, а затем отдают часть энергии колебаниям решетки, и электроны ионов хрома без излучения света переходят на уровень E_2 .

Этот уровень **метастабильный**, так как время жизни иона хрома в состоянии E_2 порядка 10^{-3} с, что на четыре порядка (т. е. в десять тысяч раз) больше времени жизни в состоянии E_3 . Большое время жизни электрона на уровне E_2 позволяет перевести достаточное число ионов хрома в это состояние. Для создания инверсии населенности необходимо, чтобы число ионов хрома в состоянии E_2 было больше, чем в основном состоянии E_1 , т. е. требуется возбудить больше половины ионов хрома.

На рис. 15.6 изображена **четырёхуровневая система**.

Ее преимущество перед трехуровневой системой состоит в том, что E_3 – нижний рабочий уровень лазерного перехода – расположен

выше основного уровня E_1 . По этой причине в условиях термодинамического равновесия он может быть населен, в соответствии с

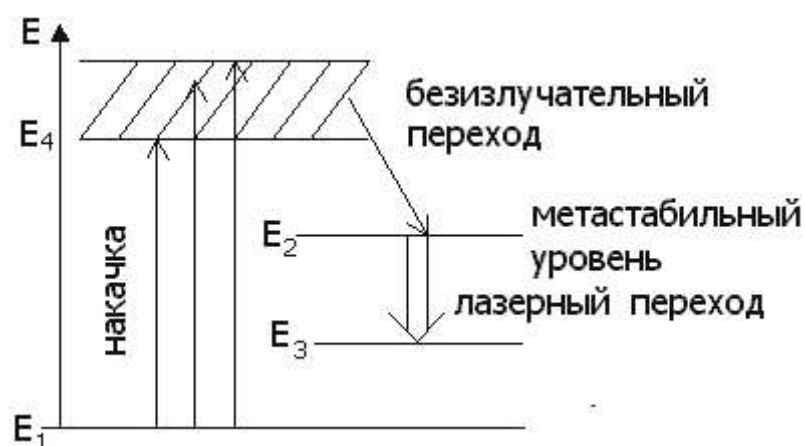


Рис. 15.6

распределением Максвелла – Больцмана, *значительно слабее*, чем уровень E_1 . Здесь состояние с инверсией населенности достигается, когда населенность метастабильного уровня E_2 больше населенности нижнего рабочего уровня E_3 .

Таким образом, на основном уровне может

остаться больше половины атомов.

Наиболее эффективным четырехуровневым ионом является трехвалентный ион неодима Nd^{3+} , вводимый в состав специальных сортов стекла.

Оптическую накачку применяют, главным образом, в твердотельных лазерах и лазерах на стеклах, активированных неодимом.

В газовых лазерах более эффективны другие методы накачки: электрический разряд, химические реакции, тепловая накачка в газодинамических лазерах.

В **полупроводниковых лазерах** важнейшим способом накачки является **инжекция носителей через р-п переход**. В отличие от лазеров других типов, в полупроводниковом лазере используются излучательные квантовые переходы между энергетическими зонами: зоной проводимости и валентной зоной полупроводникового диода.

Генерация лазерного излучения возникает в слое, примыкающем к р-п переходу при пропускании через диод большого прямого тока. Плотность тока, соответствующего началу генерации, обычно порядка $10^3 \text{ А / см}^2 = 10^7 \text{ А / м}^2$.

Оптическим резонатором в полупроводниковом лазере служат две плоскопараллельные грани, расположенные перпендикулярно плоскости р-п перехода. Коэффициент отражения этих граней около 30 %. Схема такого лазера изображена на рис. 15.7.

В полупроводниковых лазерах с электронной накачкой для создания инверсии населенности используют пучок электронов с энергией $10^4 - 10^5$ эВ.

При замедлении этих электронов за счет ионизации в полупроводнике образуются избыточные носители заряда.

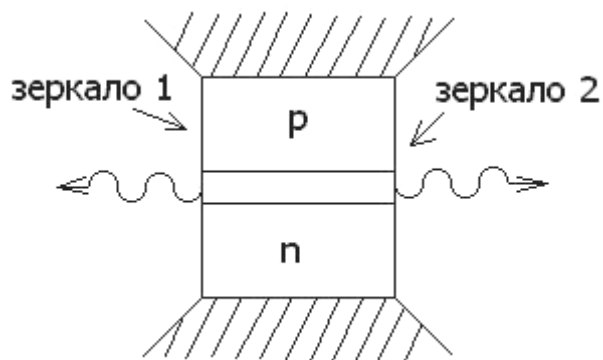


Рис. 15.7

§ 6. Виды лазеров и их применение

По режиму работы лазеры можно разделить на **импульсные и непрерывного действия**. По виду активной среды лазеры делятся на **газовые, жидкостные, полупроводниковые и твердотельные**. По способу накачки различают: лазеры с оптической накачкой, газоразрядные лазеры, химические лазеры, инжекционные лазеры и лазеры с электронной накачкой.

Для всех лазеров характерны следующие особенности излучения:

1) большая временная и пространственная когерентность. Время когерентности τ составляет 10^{-3} с, что соответствует длине когерентности $l = c \cdot \tau = 10^5$ м;

2) строгая монохроматичность: $\Delta\lambda \sim 10^{-11}$ м;

3) большая плотность потока энергии;

4) очень малое угловое расхождение в пучке (от $5 \cdot 10^{-4}$ радиан до $4 \cdot 10^{-2}$ радиан).

Коэффициент полезного действия лазеров изменяется от 0,01 % (для гелий-неонового лазера) до 75 % (для лазера на стекле с неодимом).

Мощность непрерывного излучения лазеров изменяется от 10^{-3} Вт (гелий-неоновый лазер) до 10^5 Вт (газодинамический лазер на CO_2). Мощность импульсного излучения изменяется от 10 Вт (полупроводниковые лазеры) до 10^{13} Вт (лазеры на стекле с неодимом).

Особенности лазерного излучения находят самое разнообразное применение. Способность лазера концентрировать световую энергию в пространстве, времени и узком спектральном интервале может быть использована двояко:

1) **нерезонансное воздействие** мощных световых потоков на вещество в непрерывном и импульсном режимах (лазерная обработка материалов), использование мощных лазеров для решения проблемы термоядерного синтеза;

2) **резонансное воздействие** на атомы, молекулы и молекулярные комплексы, вызывающие процессы фотодиссоциации, фотоионизации, фотохимические реакции.

Нерезонансное, тепловое воздействие лазерного излучения, используемое в **лазерной технологии** обработки материалов, упрощает операцию получения отверстий в твердых, хрупких, тугоплавких материалах. Например, лазерная технология эффективна при изготовлении алмазных фильер – рабочего инструмента машин для волочения проволоки: через отверстие в фильере протягивается обрабатываемый материал. Лазерная технология используется для резки материала, нанесения рисунка на его поверхность, образования нужного микро рельефа на ней. Лазерная сварка позволяет соединить металлы и сплавы, не свариваемые обычным способом.

В частности, в медицине (хирургии) лазерный луч в ряде случаев с успехом используется в качестве хирургического скальпеля. В офтальмологии лазерным лучом прикрепляют отслоившуюся сетчатку глаза. Отметим, что в медицине используют и резонансное воздействие лазерного луча на ткани организма, в частности, маломощное излучение гелий-неонового лазера. Механизмы такого воздействия пока в деталях не изучены, предполагается, что его необычно высокая эффективность при очень малой мощности излучения (десятки милливатт) объясняется цепными фотохимическими реакциями, возникающими под воздействием лазерного излучения.

Применение лазеров в **спектроскопии** резко повысило возможность традиционных методов, кроме того, позволило создать методы, основанные на принципиально новых физических принципах. Чувствительность спектроскопических методов доведена до предельного уровня, ограниченного регистрацией единичных атомов и молекул. Методы лазерной спектроскопии используются в **лазерной химии**, лазерном разделении изотопов.

Лазеры широко применяют в **измерительной технике**. Например, лазерные интерферометры на гелий-неоновых лазерах позволяют с большой точностью производить юстировочные и нивелировочные работы. Широко используются лазерные светодальномеры и даже лазерные рулетки на портативных полупроводниковых лазерах.

Применения лазеров столь обширны, что здесь невозможно даже их простое перечисление, кроме того, область применения лазеров постоянно расширяется.

С появлением лазеров связано рождение таких новых разделов физики, как нелинейная оптика и голография.

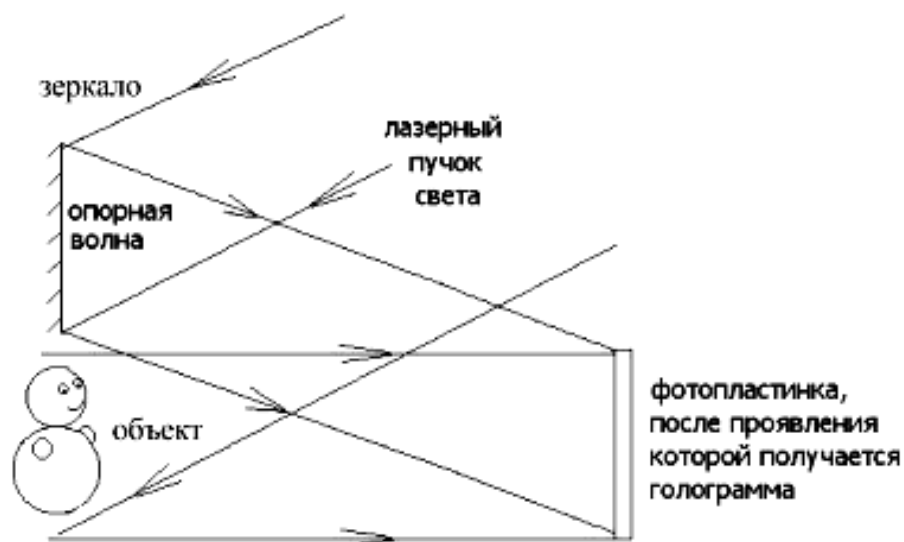
Нелинейная оптика исследует распространение мощных световых пучков в твердых телах, жидкостях и газах и их взаимодействия с веществом. Напряженности электрического поля в мощных лазерных пучках сравнимы или даже превышают внутриатомные поля. Это приводит к возникновению новых оптических эффектов и существенно меняет характер уже известных явлений. В частности, в 1969 г. была обнаружена *самофокусировка* света: мощный световой пучок, распространяясь в среде, не испытывает дифракционной расходимости, а, напротив, самопроизвольно сжимается.

Голография (от греческого *holos* – весь, полный, *grapho* – пишу) – способ записи и восстановления волнового поля, основанный на регистрации интерференционной картины, которая образована волной, отраженной предметом, освещаемым источником света (*предметная волна*), и когерентной с ней волной, идущей непосредственно от источника света (*опорная волна*). Зарегистрированная интерференционная картина называется **голограммой**.

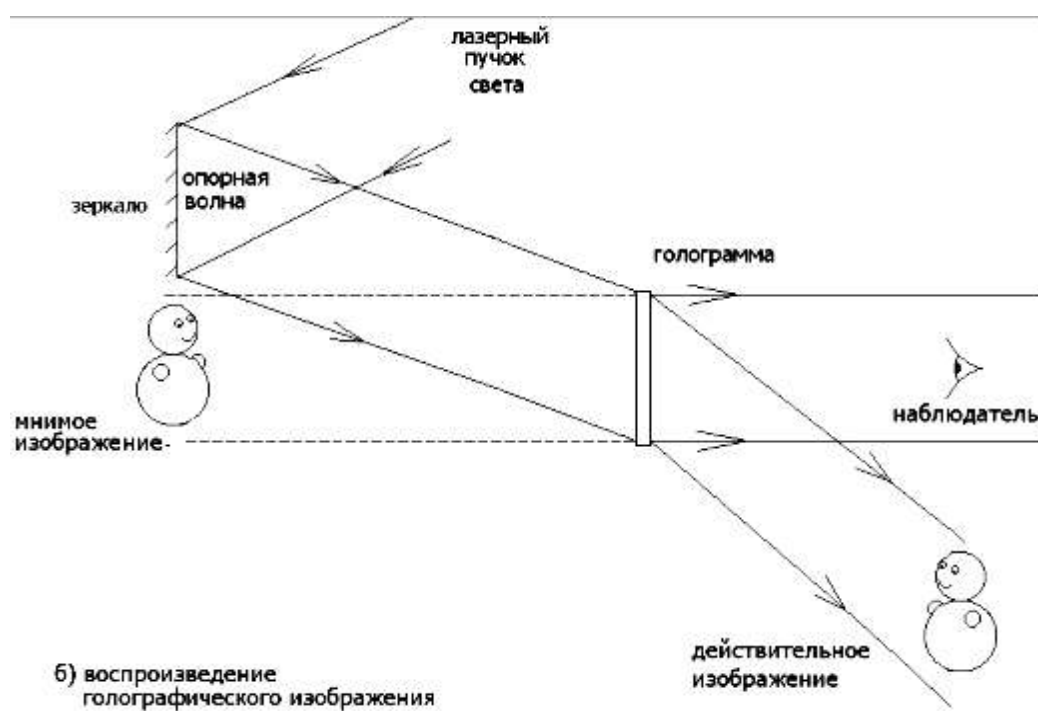
Голограмма, освещенная опорной волной, создает такое же амплитудно-фазовое пространственное распределение волнового поля, которое создавала при записи предметная волна. Таким образом, голограмма, за счет дифракции опорной волны на записанной в ней интерференционной картине, преобразует опорную волну в копию предметной.

Основы голографии были заложены в 1948 г. английским физиком Д. Габором, венгром по происхождению. Экспериментальное воплощение и дальнейшая разработка этого способа стали возможными лишь после появления источников света высокой степени когерентности – лазеров.

Схемы записи и воспроизведения голографического изображения показаны на рис. 15.8.



а) запись голограммы



б) воспроизведение голографического изображения

Рис. 15.8

ИТОГИ ЛЕКЦИИ № 15

1. *Лазер*, или оптический квантовый генератор, – это устройство, генерирующее когерентные электромагнитные волны за счет вынужденного испускания света активной средой, находящейся в оптическом резонаторе.

2. *Вынужденное излучение* возникает, если на атом, находящийся в возбужденном состоянии с энергией E_2 , воздействует фотон с частотой,