

3.16. Контакт металл - полупроводник

Рассмотрим первоначально контакт металла с n -полупроводником, у которых уровни электрохимических потенциалов не совпадают, $\mu_{эм} < \mu_{эп}$ (рис. 3.16.1).

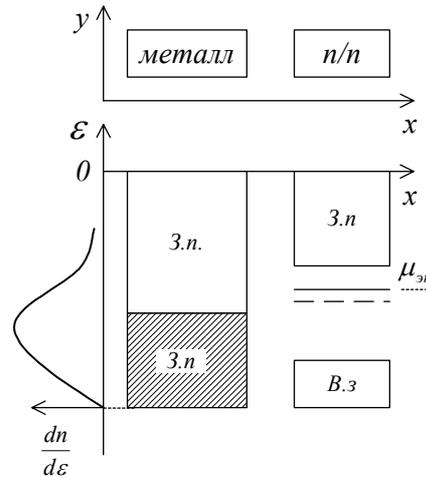


Рис. 3.16.1. Контакт металл-полупроводник (з. п. – зона проводимости; в. з. - валентная зона; $\frac{dn}{d\varepsilon}$ – распределение по энергиям концентрации электронов в з. п. металла)

Так как уровень электрохимического потенциала металла расположен ниже уровня электрохимического n -полупроводника, то в процессе их сближения возникает направленный перенос электронов из n -полупроводника в металл.

Избыточный перенос продолжается до полного выравнивания уровней электрохимических потенциалов, после чего в области контакта у n -полупроводника возникает слой, обедненный свободными носителями заряда. Возникшие там положительные заряды локализованы вблизи атомов примеси и перемещаться не могут (рис. 3.16.2). Из-за незначительной концентрации квазиэлектронов в контактном слое этот слой обладает большим электрическим сопротивлением и называется запиорным слоем.

В области контактного слоя должны существовать два потока квазичастиц: $\Phi_{мс}$ – поток квазиэлектронов из металла в полупроводник и

$\Phi_{пе}$ – поток квазиэлектронов из полупроводника в металл. Они равны между собой. Дырки в валентной зоне n -полупроводника – неосновные носители заряда, их концентрация значительно меньше концентрации квазиэлектронов в зоне проводимости. Поэтому потоком дырок через контактный слой можно пренебречь.

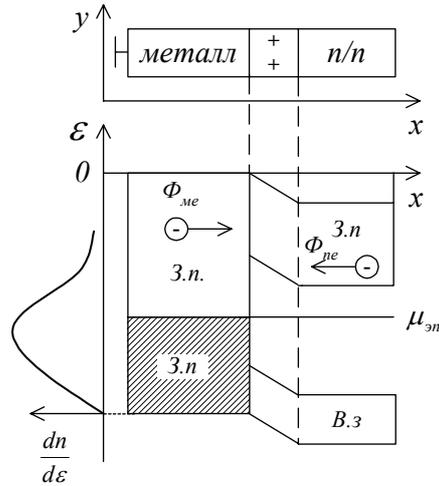


Рис. 3.16.2. Возникновение барьера Шоттки (з. п. – зона проводимости; в. з – валентная зона; $\frac{dn}{d\varepsilon}$ – распределение по энергиям концентрации электронов в з. п. металла)

Возникший в области контакта положительный заряд приводит к искривлению всех уровней полной энергии n -полупроводника, приводит к возникновению потенциального барьера, который называется барьером Шоттки (рис. 3.16.2).

Теперь приложим к образцам внешнюю разность потенциалов (рис. 3.16.3). Так как $\varepsilon_n = e\varphi$, $\varphi > 0$, $e < 0$, то все уровни полной энергии n -полупроводника смещаются вниз, электрохимический потенциал n -полупроводника

$$\mu_{эп} = \mu + e\varphi \quad (3.16.1)$$

также понижается. В такой замкнутой цепи возникает электрический ток. Наибольшее электрическое сопротивление приходится на зазорный слой, поэтому основное падение напряжения происходит в зазорном слое, падение напряжения на остальных участках цепи незначительно.

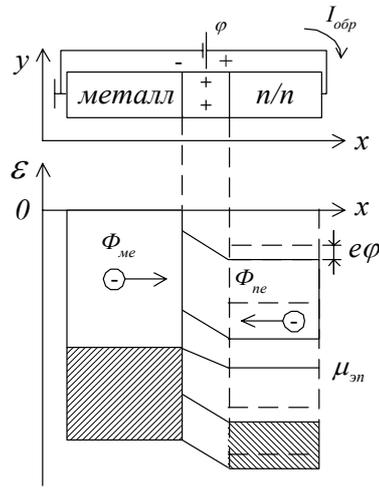


Рис. 3.16.3. Контактный слой во внешнем поле (запорное направление)

Квазиэлектроны металла, попадая в электрическое поле контактного слоя, переносятся в n -полупроводник. В то же время на пути движения квазиэлектронов из n -полупроводника находится потенциальный барьер, высота которого возросла по сравнению с первоначальной. Поэтому поток квазиэлектронов в сторону металла уменьшается. В результате приложенной разности потенциалов с указанной на рис. 3.16.3 полярностью один поток остается неизменным, а второй уменьшается, возникает разностный поток квазичастиц

$$\Delta\Phi_e = \Phi_{me} - \Phi_{ne}. \quad (3.16.2)$$

Значит, в замкнутой цепи будет протекать электрический ток

$$I_{обр} = I_0 \left(\exp\left(-\frac{e\varphi}{kT}\right) - 1 \right), \quad e > 0. \quad (3.16.3)$$

Его физическое происхождение вызвано преобладающим влиянием тока проводимости из металла над диффузионным током из полупроводника.

Приложим теперь внешнюю разность потенциалов обратного знака. Тогда уровень электрохимического потенциала n -полупроводника по сравнению с исходным поднимется вверх, все уровни полной энергии n -полупроводника также поднимутся вверх. В результате высота потенциального барьера понизится (рис. 3.16.4). Но положение уров-

ней полной энергии в металле при этом не изменится, не изменится и величина потока квазиэлектронов из металла в сторону полупроводника. Значит, поток проводимости квазиэлектронов из металла сохраняется прежним. Что же касается диффузионного потока квазиэлектронов из полупроводника в сторону металла, то в результате уменьшения высоты потенциального барьера он возрастет, через контактный слой пройдет результирующий поток квазичастиц

$$\Delta\Phi = \Phi_{пе} - \Phi_{ме}, \quad (3.16.4)$$

будет протекать электрический ток (прямой ток)

$$I_{np} = I_0 \left(\exp\left(\frac{e\varphi}{kT}\right) - 1 \right), e < 0. \quad (3.16.5)$$

Этот ток имеет в основном диффузионное происхождение.

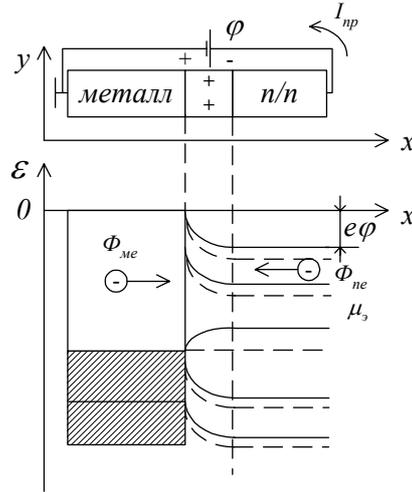


Рис. 3.16.4. Контактный слой во внешнем поле (прямое направление)

Таким образом, омический контакт, запирающий слой должен обладать выпрямляющим для переменного тока действием. На основе этого явления созданы и широко используются диоды Шоттки. Их замечательной особенностью является необычно малое время релаксации – порядка 10^{-11} с, что позволяет с успехом их применять в импульсных радиотехнических схемах, вычислительной технике, автоматике, т.е. там, где требуется высокое быстродействие, способность обрабатывать возможно более короткие и часто следующие друг за другом электрические импульсы.

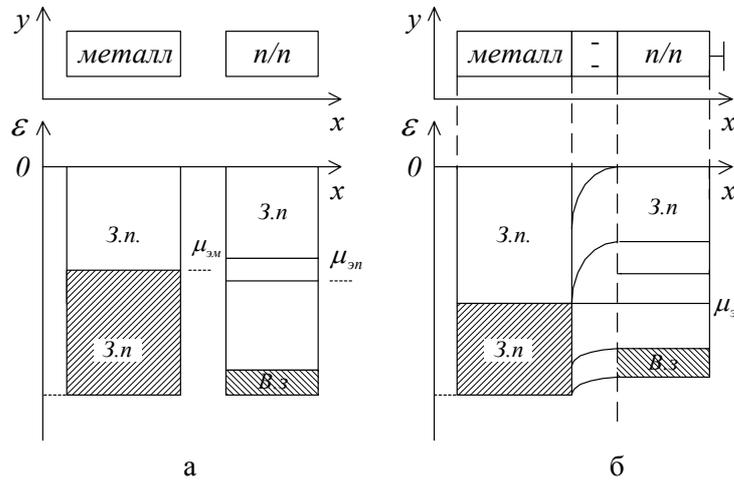


Рис. 3.16.5. Возникновение антизапорного слоя

Рассмотрим теперь контакт металла с n -полупроводником, у которого уровень электрохимического потенциала больше уровня электрохимического потенциала полупроводника (рис. 3.16.5). В результате после контакта в полупроводник переходит некоторое число электронов, в контактном слое полупроводника возникает отрицательно заряженный слой, который называется антиза-порным слоем (рис. 3.16.5, б). Электрическое сопротивление антизапорного слоя очень мало, поэтому такой контакт не обладает выпрямляющим свойством. Тем не менее, он широко применяется для создания омических контактов, с помощью которых полупроводник, полупроводниковые устройства подключаются к электрическим схемам.

3.17. Соотношение Эйнштейна

Предположим, что в объеме образца в направлении координатной оси OX происходит изменение концентрации заряженных частиц идеального газа $n(x)$. Электрический заряд частицы q . Кроме того, в объеме существует электрическое поле E , направление которого противоположно градиенту концентрации $\nabla n(x)$. В результате наличия градиента концентрации в объеме возникает диффузия заряженных частиц, направленная в одну сторону, и поток проводимости (дрейфовый поток), направленный в противоположную сторону.

Тогда полный ток будет складываться из дрейфового тока (тока проводимости) и диффузионного тока, направления которых противоположны. Следовательно, в любой точке полупроводника для произвольного момента времени плотность тока равна сумме диффузионного и дрейфового токов (тока проводимости): $j = j_e + j_d$. В отсутствие внешнего источника тока

$$j_e = -j_d. \quad (3.17.1)$$

Из первого закона Фика следует: плотность тока диффундирующих заряженных частиц

$$j_d = -qD \frac{dn}{dx}, \quad (3.17.2)$$

q – заряд частицы, D – коэффициент диффузии.

Плотность дрейфового тока (тока проводимости)

$$j_e = q \langle v \rangle n, \quad (3.17.3)$$

где n – концентрация, $\langle v \rangle$ – дрейфовая скорость носителей заряда:

$$\langle v \rangle = b E,$$

b – подвижность носителей заряда, E – напряженность электрического поля. После подстановки значения $\langle v \rangle$ в (3.17.3) получаем

$$j_e = qbnE. \quad (3.17.4)$$

Найденные в (3.17.2), (3.17.4) значения j_d , j_e подставляем в (3.17.1):

$$qbnE = qD \frac{dn}{dx}. \quad (3.17.5)$$

Для заряженных частиц идеального газа, находящихся в электрическом поле консервативных сил, справедливо распределение Больцмана, согласно которому концентрация зависит от величины потенциальной энергии частицы в данной точке пространства и температуры. Если принять $q = e > 0$, то

$$n = n_0 e^{-\frac{e\varphi}{kT}}, \quad \frac{dn}{dx} = -\frac{e}{kT} n \frac{d\varphi}{dx}, \quad (3.17.6)$$

φ – потенциал электростатического поля.

Напряженность электрического поля в полупроводнике

$$E = -\frac{d\varphi}{dx}. \quad (3.17.7)$$

Поэтому после подстановки $\frac{dn}{dx}$ и E из (3.17.6) и (3.17.7) в (3.17.5) и сокращения на постоянные множители получаем:

$$\frac{D}{b} = \frac{kT}{e}. \quad (3.17.8)$$

Уравнение (3.17.8) называется соотношением Эйнштейна. Оно справедливо для системы частиц с малой концентрацией, когда взаимодействиями между частицами можно пренебречь. К таким частицам, в частности, относятся квазиэлектроны и дырки в собственных и примесных полупроводниках.

3.18. Выпрямляющее действие p - n перехода

В любом полупроводнике одновременно происходят два противоположно направленных процесса: а) генерация квазиэлектронов и дырок; б) рекомбинация квазиэлектронов и дырок. В результате в объеме между этими процессами устанавливается динамическое равновесие. Основными механизмами рекомбинации являются излучательный механизм, когда энергия рекомбинирующей пары квазиэлектрон–дырка излучается в виде фотона; фононный механизм, когда энергия рекомбинирующей пары передается ионам кристаллической решетки; ударный механизм, когда энергия рекомбинирующей пары передается третьей частице.

Возникающий на границе раздела n - и p -образцов контактный слой обладает рядом физических особенностей. Если, например, к концам p - n образца приложить переменную разность потенциалов, то оказывается, что p - n переход будет обладать преимущественно односторонней проводимостью, а, следовательно, и выпрямляющим действием.

На рис. 3.18.1 изображен контакт n - и p -полупроводников, цифрами 1 и 2 обозначены границы контактного слоя, $\Phi_{не}$, $\Phi_{рд}$ – потоки основных носителей заряда, $\Phi_{ре}$, $\Phi_{нд}$ – потоки неосновных носителей заряда. Как следует из физического смысла рис. 3.18.1, концентрация квазиэлектронов в p -образце в результате рекомбинаций квазиэлектронов и дырок с увеличением расстояния от границы контактного слоя уменьшается. Найдем закон изменения концентрации.

Пусть τ_e – среднее время жизни квазиэлектронов в p -полупроводнике, в котором они являются неосновными носителями заряда. Тогда в отсутствие диффузии из p - в n -образец изменение концентрации по глубине будет происходить только в результате рекомбинации квазиэлектронов зоны проводимости с дырками валентной зоны. Следовательно, изменение концентрации во времени

$$\frac{\partial n}{\partial t} = \frac{n_0 - n}{\tau_e},$$

n_0 – равновесное значение концентрации квазиэлектронов.

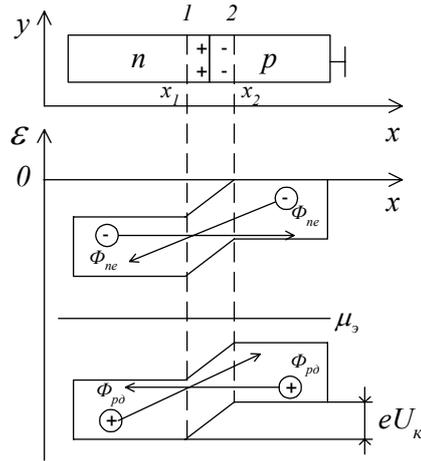


Рис. 3.18.1. Потоки носителей зарядов через p - n переход

С другой стороны, в отсутствие рекомбинаций изменение концентрации во времени определяется 2-м законом Фика:

$$\frac{\partial n}{\partial t} = D_e \frac{\partial^2 n}{\partial x^2},$$

где D_e – коэффициент диффузии квазиэлектронов в p -образце (коэффициент диффузии неосновных носителей заряда).

Так как оба процесса в образце существуют одновременно, то

$$\frac{\partial n}{\partial t} = \frac{n_0 - n}{\tau_e} + D_e \frac{\partial^2 n}{\partial x^2}.$$

При наличии динамического равновесия $\frac{\partial n}{\partial t} = 0$, поэтому

$$\frac{\partial^2 n}{\partial x^2} = \frac{n_0 - n}{D_e \tau_e}.$$

Обозначим

$$\lambda_e = \sqrt{D_e \tau_e}, \quad (3.18.1)$$

λ_e – диффузионная длина квазиэлектронов (неосновных носителей заряда) в p -образце. Тогда

$$\frac{d^2(n - n_0)}{dx^2} = \frac{n - n_0}{\lambda_e^2}. \quad (3.18.2)$$

Непосредственной подстановкой можно показать, что решением дифференциального уравнения (3.18.2) является

$$n(x) = n_0 + c_1 e^{-\frac{x+\lambda_e}{\lambda_e}}, \quad (3.18.3)$$

где

$$c_1 = -n_0 e^{-\frac{\lambda_e+\lambda_e}{\lambda_e}}, \quad n(\lambda_e) = 0.$$

Тогда после подстановки найденного значения c_1 в уравнение (3.18.3) получаем:

$$n(x) = n_0 \left(1 - \exp\left(-\frac{x-\lambda_e}{\lambda_e}\right) \right), \quad (3.18.4)$$

$$x \leq \lambda_e$$

Уравнение (3.18.4) позволяет дать физически обоснованное определение диффузионной длины квазиэлектронов в p -образце: диффузионная длина ℓ_e – это то расстояние, на котором в результате рекомбинаций квазиэлектронов и дырок неосновные носители заряда полностью исчезают: квазиэлектроны – в зоне проводимости p -полупроводника, дырки – в валентной зоне n -полупроводника.

Оценим теперь величину диффузионной длины квазиэлектронов ℓ_e и дырок ℓ_d в примесных полупроводниках n - и p -типа на основе Si и сравним ее с толщиной контактного слоя L . Для уменьшения объема вычислений предположим, что $n_{\text{дон}} = n_{\text{акц}}$, а температура образца $T = 300\text{K}$. Из опыта известны подвижности квазиэлектронов b_e и дырок b_d в этих полупроводниках и времена жизни. Необходимые для вычисления диффузионной длины по формуле (3.18.1) значения коэффициента диффузии вычислим по формуле соотношения Эйнштейна. В табл. 3.18.1 представлены результаты таких вычислений для двух концентраций неосновных носителей заряда.

$n_{\text{дон}} = n_{\text{акц}}$	b_e	b_d	τ_e	τ_d	D_e	D_d	ℓ_e	ℓ_d
М^3	$\text{М}^2/(\text{В}\cdot\text{с})$		с		$\text{М}^2/\text{с}$		см	
$1 \cdot 10^{22}$	0,074	0,038	0,72	10,5	$1,9 \cdot 10^{-3}$	$9,9 \cdot 10^{-4}$	3,7	10,2
$1 \cdot 10^{21}$	0,140	0,046	50	167	$3,6 \cdot 10^{-3}$	$1,2 \cdot 10^{-4}$	42,4	44,7

Таблица 3.18.1. Динамические характеристики неосновных носителей заряда в Si

Из таблицы 3.18.1 следует, что диффузионная длина ℓ_e неосновных носителей заряда в десятки тысяч раз превышает толщину контактно-

го слоя. Поэтому при движении в сторону контактного слоя последний не может оказать на них заметного влияния.

Квазиэлектроны в p -образце появляются из n -образца в результате прохождения через контактный слой $L = L_e + L_d$. Как следует из табл. 3.18.1, в полупроводниках $\ell_e \gg L$, а так как концентрация квазиэлектронов в p -образце всегда мала, то квазиэлектроны в полупроводнике образуют ЭФГ, обладающий, как известно, свойствами идеального газа. В результате контакта n - и p -полупроводников в области контактного слоя толщиной L возникает электрическое поле, замедляющее движение квазиэлектронов из n - в p -образец. Поэтому в пределах контактного слоя происходит постепенное изменение потенциальной энергии квазиэлектронов, что в соответствии с распределением Больцмана приводит к изменению их концентрации. В отсутствие приложенной внешней разности потенциалов концентрация квазиэлектронов на правой границе контактного слоя в точке x_2

$$n(x_2)_0 = n(x_1) e^{-\frac{eU_k}{kT}}, \quad (3.18.5)$$

где $n(x_1)$ – концентрация квазиэлектронов на левой границе x_1 контактного слоя.

Так как контактный слой обеднен основными носителями заряда, то он обладает электрическим сопротивлением, во много раз превышающим электрическое сопротивление остального p - n полупроводника. Поэтому, если к p - n -полупроводнику приложена внешняя разность потенциалов, то практически все падение напряжения U будет в области контактного слоя L . Тогда концентрация квазиэлектронов на правой границе изменится и окажется равной

$$n(x_2) = n(x_2)_0 e^{\frac{eU}{kT}}. \quad (3.18.6)$$

Возникшее изменение концентрации квазиэлектронов в точке x_2

$$n_e = n(x_2)_0 \left(e^{\frac{eU}{kT}} - 1 \right). \quad (3.18.7)$$

После подстановки в (3.18.7) значение (3.18.3) получаем

$$n_e = \left(n(x_2)_0 + c_1 e^{-\frac{x+x_2}{\lambda_e}} \right) \left(e^{\frac{eU}{kT}} - 1 \right). \quad (3.18.8)$$

Так как плотность тока квазиэлектронов в p -образце

$$j_e = eD_e \frac{dn_e}{dx}, \quad (3.18.9)$$

то с учетом (3.18.3) и (3.18.8) получаем:

$$j_e = e \frac{D_e}{\lambda_e} n_e(x_2) \left(e^{-\frac{eU}{kT}} - 1 \right). \quad (3.18.10)$$

Проведя аналогичные преобразования для дырок, получаем

$$j_\delta = e \frac{D_\delta}{\lambda_\delta} n_\delta(x_2) \left(e^{-\frac{eU}{kT}} - 1 \right). \quad (3.18.11)$$

Тогда результирующая плотность тока квазиэлектронов и дырок равна

$$j = j_S \left(e^{-\frac{eU}{kT}} - 1 \right), \quad (3.18.12)$$

где

$$j_S = e \left(\frac{D_e}{\lambda_e} n_e + \frac{D_\delta}{\lambda_\delta} n_\delta \right), \quad (3.18.13)$$

j_S – плотность тока насыщения.

Найденное из (3.18.1) выражение для коэффициента диффузии подставим в (3.18.13):

$$j_S = e \left(\frac{\lambda_e}{\tau_e} n_e + \frac{\lambda_\delta}{\tau_\delta} n_\delta \right). \quad (3.18.14)$$

Так как рекомбинация квазиэлектронов и дырок или их генерация вызваны переходами из зоны проводимости в случае рекомбинации и, наоборот, из валентной зоны в зону проводимости – в случае генерации, то времена жизни квазиэлектронов τ_e и дырок τ_δ обычно достаточно велики. Что же касается обычных столкновений, при которых число неосновных носителей заряда сохраняется неизменным, то они происходят соответственно в пределах одной и той же зоны, такие столкновения приводят лишь к внутризонным переходам неосновных носителей заряда. Экспериментальные оценки показывают, что время жизни носителей заряда в результате межзонных переходов находится в интервале $10^{-8} \text{ с} \leq \tau_d \leq 10^{-3} \text{ с}$, а время жизни неосновных носителей заряда в результате внутризонных столкновений составляет $10^{-13} \text{ с} \leq \tau_e \leq 10^{-12} \text{ с}$.

Подставим в (3.18.13) значения коэффициента диффузии, найденного из соотношения Эйнштейна. После сокращения на постоянные множители имеем:

$$j_s = kT \left(\frac{b_e}{\lambda_e} + \frac{b_o}{\lambda_o} \right). \quad (3.18.15)$$

Пусть в уравнении (3.18.12) $U > 0$. Тогда добавочная потенциальная энергия квазиэлектрона $eU < 0$. Это приводит к тому, что все зоны n -полупроводника будут смещаться вниз (рис. 3.18.1). И теперь на пути движения квазиэлектронов встретится потенциальный барьер, величина которого возросла на $|eU|$. Поэтому теперь не все квазиэлектроны из диффузионного потока n -образца смогут преодолеть потенциальный барьер. Что же касается тока проводимости, по-прежнему существующего в p -полупроводнике, то его величина практически не изменится, так как квазиэлектроны p -полупроводника, достигнув границы контактного слоя, попадают в ускоряющее электрическое поле и перебрасываются в n -полупроводник. В результате через контактный слой L будет протекать разностный ток, состоящий из дрейфового и диффузионного токов, направление тока будет определяться током проводимости (дрейфовым током), а сам ток имеет малую величину. Указанный разностный ток называется обратным током:

$$j_{обр} = j_s \left(e^{\frac{|eU|}{kT}} - 1 \right). \quad (3.18.16)$$

Если же $U < 0$, то $eU > 0$. Поэтому зона проводимости смещается вверх, высота потенциального барьера на пути движения квазиэлектронов в p -образец уменьшается, через контактный слой будет протекать ток большей величины по сравнению с первоначальной. Теперь диффузионный поток больше потока проводимости, а направление разностного тока будет совпадать с направлением диффузионного тока. Этот ток называется прямым током:

$$j_{прям} = j_s \left(e^{\frac{eU}{kT}} - 1 \right). \quad (3.18.17)$$

Как видно из уравнений (3.18.16), (3.18.17), прямой и обратный токи сильно зависят от температуры, с ростом температуры выпрямляющее действие p - n перехода резко убывает.

Для количественной характеристики выпрямляющего действия p - n перехода вводится коэффициент выпрямления

$$K = \frac{j_{прям}}{j_{обр}}. \quad (3.18.18)$$

T, K	300	350	400	450
U, B	$K = j_{\text{прям}} / j_{\text{обр}}$			
0,1	47,7	27,5	18,1	13,2
0,2	2274	754	329	113
0,3	$1,1 \cdot 10^5$	$2,1 \cdot 10^4$	$5,8 \cdot 10^3$	$2,3 \cdot 10^3$
0,4	$5,1 \cdot 10^6$	$5,7 \cdot 10^5$	$1,1 \cdot 10^5$	$3,0 \cdot 10^4$

Таблица 3.18.2. Значения коэффициентов выпрямления при различных значениях разности потенциалов U в интервале температур 300–450 К

В табл. 3.18.2 приведены значения коэффициентов выпрямления при различных значениях разности потенциалов U в интервале температур 300–450 К. Из табл. 3.18.2 видно, что коэффициент выпрямления K p - n перехода сильно зависит от величины внешней разности потенциалов U и температуры. Так как полупроводник с p - n переходом по отношению к переменному электрическому току обладает односторонней проводимостью, т.е. выпрямляющим действием, то он называется полупроводниковым диодом. Если сравнивать выпрямляющее действие полупроводникового диода с выпрямляющим действием двухэлектродной лампы, состоящей только из катода и анода, то у вакуумного диода коэффициент выпрямления не зависит ни от температуры катода, ни от разности потенциалов между анодом и катодом, что является несомненным их преимуществом перед полупроводниковым диодом. Однако большим достоинством полупроводниковых диодов являются их малые габариты и отсутствие дополнительного источника тока для питания катода, что всегда имеет место в вакуумном диоде.

3.19. Транзисторы

Транзистор – это устройство, состоящее из трех полупроводников и двух p - n переходов, используемое для управления электрическим током, протекающим через него. Транзисторы применяются для усиления, генерации или переключения токов.

Когда к полупроводниковому диоду добавляется третий слой полупроводника, то получается устройство, которое может усиливать мощность или напряжение. Это устройство и называется транзистором. Транзистор состоит из трех областей с чередующимся типом проводимости: n - p - n или p - n - p .

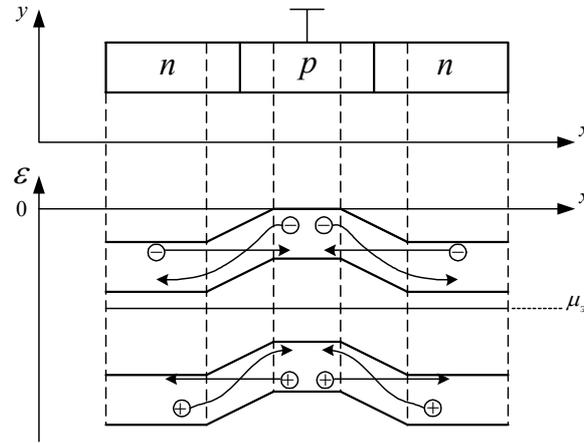


Рис. 3.19.1. Энергетические зоны в n - p - n транзисторе

На рис. 3.19.1 изображена принципиальная схема транзистора n - p - n , а также расположение валентной зоны и зоны проводимости. Так как к транзистору внешняя разность потенциалов не приложена, то уровень электрохимического потенциала μ , всюду один и тот же.

Из условия динамического равновесия вытекает равенство потоков, протекающих через контактные слои в прямом и обратном направлениях:

$$\Phi_{ne} - \Phi_{pe} = \Phi_{pd} - \Phi_{nd}. \quad (3.19.1)$$

Отсюда следует, что в областях n - p , p - n переходов должно иметь место равенство токов: диффузионный ток $I_{\text{диф}}$ равен дрейфовому току (току проводимости) $I_{\text{пров}}$:

$$I_{pe} + I_{ng} = I_{pg} + I_{ne}. \quad (3.19.2)$$

Однако если к транзистору подключить внешние источники тока, как это показано на рис. 3.19.2, то расположение уровней электрохимического потенциала в n -полупроводниках изменится. А так как в данном случае p -полупроводник заземлен, то в нем уровень электрохимического потенциала останется прежним. Тогда в соответствии с изменением положения уровня электрохимического потенциала изменятся и положения зон проводимости и валентных зон в полупроводниках. На рис. 3.19.2 левый n -полупроводник называется эмиттером, правый n -полупроводник – коллектором, а средний p -полупроводник – базой. Если сравнивать транзистор с аналогичной по своему действию электронной лампой-триодом, то происхождение названий эмит-

тер, база, коллектор можно объяснить следующим образом. В транзисторе эмиттер является источником носителей заряда, его роль эквивалентна катоду триода, коллектор является приемником носителей заряда, его роль эквивалентна аноду триода, что же касается p - n перехода, то он выполняет роль управляющей сетки электронной лампы.

На рис. 3.19.2 приведена схема включения с общей базой. Рассмотрим принцип действия транзистора как усилителя электрических сигналов в схеме с общей базой. Как видно из рис. 3.19.2, на эмиттер подается потенциал $U_э$, понижающий потенциальный барьер на границе n - и p -полупроводников. В результате концентрация квазиэлектронов в p - области и дырок в n -области резко возрастает по сравнению с равновесной. Происходит инжекция носителей зарядов. Квазиэлектроны, инжектируемые из эмиттера $э$ в базу $б$, затем диффундируют к коллектору $к$. Если база достаточно тонкая, то практически все квазиэлектроны достигают коллектора. Действительно, попадая в электрическое поле контактного слоя, существующего между базой и коллектором, квазиэлектроны ускоряются электрическим полем p - n перехода и перебрасываются в коллекторную цепь транзистора. Геометрия транзистора выбирается такой, чтобы через базовый вывод протекал незначительный по сравнению с коллекторным ток.

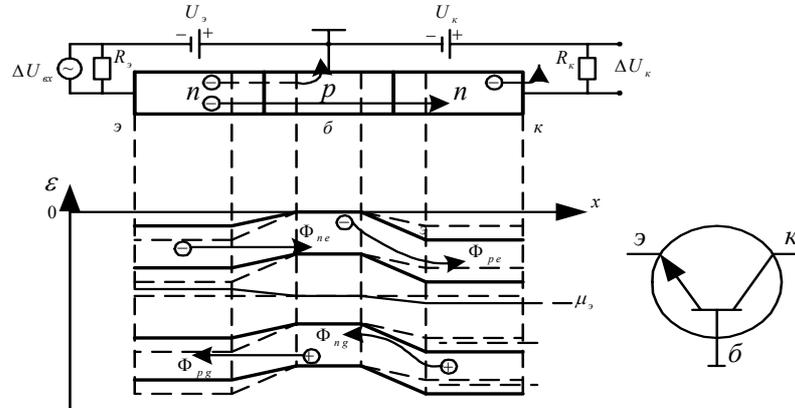


Рис. 3.19.2. Включение транзистора с общей базой

Предположим, что на входное сопротивление $R_э$ подается напряжение $\Delta U_{эx}$. Это приводит к изменению тока эмиттера на величину

$$\Delta I_э = \frac{\Delta U_{эx}}{R_э}. \quad (3.19.3)$$

А так как $I_3 \approx I_k$, то практически на ту же величину изменится и ток в цепи коллектора: $\Delta I_k \approx \Delta I_3$. Значит, в результате этого на сопротивлении нагрузки R_k произойдет изменение падения напряжения на величину

$$\Delta U_{ex} = \Delta I_k R_k. \quad (3.19.4)$$

Тогда коэффициент усиления по напряжению в цепи транзистора с общей базой будет

$$K_U = \frac{\Delta U_{вых}}{\Delta U_{ex}} = \frac{R_k}{R_3}. \quad (3.19.5)$$

Из рис. 3.19.2 следует, что эмиттерный переход включен в прямом направлении, следовательно, он обладает малым электрическим сопротивлением, в то же время коллекторный переход включен в запорном направлении, следовательно, он обладает большим электрическим сопротивлением, и к нему можно подключить сопротивление нагрузки R_k большой величины.

Из опыта следует, что коэффициент усиления по напряжению находится в интервале $10 \leq K_U \leq 50$.

Итак, используя транзистор с общей базой, можно получить даже в одном каскаде значительное усиление как по напряжению, так и по мощности.

В настоящее время существует несколько типов транзисторов. Одним из них, кроме рассмотренного выше, является полевой транзистор, в котором рабочий ток протекает по узкому участку полупроводника. Параллельно ему расположен другой электрод (затвор), напряжение на котором определяет ширину участка с протекающим током. Например, если на затвор подается отрицательное напряжение, то оно отталкивает основные носители заряда в *n*-полупроводнике и, наоборот, притягивает их в *p*-полупроводнике. В результате участок полупроводника, в котором протекает рабочий ток, в первом случае сужается, его электрическое сопротивление возрастает, электрический ток уменьшается, а во втором случае электрический ток соответственно возрастает. Таким образом, в полевом транзисторе управление током осуществляется напряжением U , приложенным перпендикулярно направлению тока.

Любой транзистор представляет собой твердотельный цилиндр. Его изготавливают с помощью специальных технологических операций, создающих чередование *p-n* переходов, а также металлические контакты на полупроводниках. Такие транзисторы обладают большой механической прочностью, поэтому их можно применять в условиях боль-

ших вибраций и ускорений (автомобиль, ракета и т.д.). Для работы транзистора используют низковольтные источники питания (2–3 В).

В 60-е годы 20-го века были разработаны интегральные микросхемы. Они изготавливаются следующим образом. На диэлектрической пластине размещаются транзисторы, конденсаторы, резисторы, а также проводники, соединяющие их в определенной последовательности. В результате получаются очень компактные и взаимозаменяемые микросхемы. Уже существуют микросхемы с плотностью упаковки элементов порядка 10^6 мм². Они заменили электронные блоки из отдельных транзисторов, что привело к появлению ЭВМ третьего поколения.

Микросхемы применяют на заводах в станках с числовым программным управлением, в вычислительных комплексах, устройствах автоматики, связи, электронной разведки и др.

3.20. Внутренний фотоэффект в полупроводниках

Внутренним эффектом называется перераспределение электронов по энергетическим состояниям, происходящее в результате поглощения электромагнитного излучения. В полупроводниках внутренний фотоэффект проявляется в изменении величины электропроводности и называется фотопроводимостью, в изменении под действием падающего излучения относительной диэлектрической проницаемости и называется фотодиэлектрическим эффектом, в возникновении под действием излучения ЭДС и называется фотоЭДС.

Физическая причина перечисленных выше эффектов заключается в особенностях расположения зон у полупроводников. Самым характерным для полупроводников процессом поглощения света является собственное поглощение, при котором электрон валентной зоны, поглотив фотон, переходит либо на примесный уровень в *n*-полупроводнике, либо в зону проводимости. При этом энергия фотона

$$h\nu = \varepsilon_K - \varepsilon_H,$$

где ε_K , ε_H – энергии электрона в конечном и начальном состояниях.

Собственное поглощение света возможно при $h\nu \geq \varepsilon$, где $\varepsilon = \Delta\varepsilon_0$, $\Delta\varepsilon_a$, $\Delta\varepsilon_g$, $\Delta\varepsilon_0$ – ширина запрещенной зоны, $\Delta\varepsilon_a$ – интервал энергии между верхней границей валентной зоны и акцепторным уровнем в *p*-полупроводнике, $\Delta\varepsilon_g$ – интервал энергии между донорным уровнем и нижней границей зоны проводимости.

При таких переходах наряду с законом сохранения энергии должен выполняться закон сохранения квазиимпульса:

$$p = p_0 + \eta q,$$

где p_0 – начальный квазиимпульс электрона, p – конечный квазиимпульс электрона, ηq – изменение квазиимпульса электрона, возникающее в результате таких переходов. Для видимого и длинноволнового излучения $\eta q \ll p_0$, поэтому $p \approx p_0$.

Различают прямые и непрямые переходы. Прямые переходы возникают при условии $p \approx p_0$. Непрямые переходы возникают, когда $p \neq p_0$.

На основе использования внутреннего фотоэффекта созданы многочисленные виды приемников электромагнитного излучения, к их числу, в частности, относятся фоторезисторы, фотодиоды, фотоэлементы и другие, нашедшие в настоящее время самое широкое применение. Рассмотрим физические основы работы некоторых видов приемников более подробно.

3.20.1 Фоторезисторы (фотосопротивления)

Фотосопротивления изготавливают из светочувствительных материалов, таких как сульфид кадмия (CdS), селенид кадмия (CdSe), являющиеся полупроводниками n -типа, а также и из других полупроводников.

Светочувствительный материал наносится на изолирующую подложку из стекла или керамики. В зависимости от величины и спектрального состава падающего электромагнитного излучения электрическое сопротивление фоторезистора меняется от 100 Ом до 10 МОм. Чувствительный элемент изготавливается таким, чтобы в нем поглощалась значительная часть падающего потока электромагнитной волны:

$$\Phi_{\text{погл}} = (1 - r) \Phi_0, \quad (3.20.1.1)$$

где r – коэффициент отражения, Φ_0 – величина падающего потока.

Под действием поглощенного потока в полупроводнике возникают дополнительные носители заряда двух знаков. Число носителей заряда, генерируемых в чувствительном элементе за 1с, равно

$$N = \eta \frac{\Phi_{\text{погл}}}{h\nu}, \quad (3.20.1.2)$$

где η – квантовый выход внутреннего фотоэффекта. Квантовым выходом η называется число пар носителей заряда (квазиэлектронов и дырок), рождаемых в среднем каждым поглощенным фотоном.

Предположим, что фоторезист состоит из собственного полупроводника. Если теперь к концам фоторезиста приложить разность потенциалов U , то в резисте возникнет электрическое поле напряженностью E . Под действием этого поля носители заряда будут перемещаться в определенном направлении, в полупроводнике возникает дрейфовый ток (ток проводимости), он называется фототоком I_ϕ .

Пусть v_e – дрейфовая скорость носителей заряда одного знака, допустим квазиэлектрона, ℓ – длина резистора. Тогда время дрейфа носителя заряда через резистор

$$t = \frac{\lambda}{v_e}. \quad (3.20.1.3)$$

Так как $v_e = b_e E = b_e \frac{U}{\lambda}$, b_e – подвижность квазиэлектронов, то после подстановки v_e в (3.20.1.3) имеем

$$t = \frac{\lambda^2}{b_e U}. \quad (3.20.1.4)$$

Пусть τ_e – время жизни квазиэлектронов, тогда с учетом (3.20.1.1), (3.20.1.2), (3.20.1.4) получаем выражение для фототока квазиэлектронов

$$I_{\phi e} = eN \frac{\tau_e}{t}$$

или

$$I_{\phi e} = \frac{e\eta\lambda}{hc\lambda^2} (1-r) \tau_e b_e \Phi_0 U. \quad (3.20.1.5)$$

Аналогичным образом получается выражение для дырочного фототока

$$I_{\phi d} = \frac{e\eta\lambda}{hc\lambda^2} (1-r) \tau_d b_d \Phi_0 U, \quad (3.20.1.6)$$

где τ_d – время жизни дырок, b_d – подвижность дырок.

Тогда полный фототок носителей заряда двух знаков будет

$$I_\phi = \frac{e\eta\lambda}{hc\lambda^2} (1-r) (\tau_e b_e + \tau_d b_d) \Phi_0 U. \quad (3.20.1.7)$$

Время жизни носителей заряда определяет не только величину фототока, но и инерционность фоторезистора. Действительно, в результате рекомбинации квазиэлектронов и дырок концентрация носителей заряда, квазиэлектронов, появившихся в зоне проводимости, и дырок

– в валентной зоне, как это следует из 2-го закона Фика, убывает по закону

$$\Delta n = \Delta n_0 \exp\left(-\frac{t-t_0}{\tau}\right), \quad (3.20.1.8)$$

где t_0 – момент включения света.

Поэтому по такому же закону будет меняться и фототок I_ϕ . Из табл. 3.17.1, приведенной в разделе 3.17, следует, что время жизни носителей заряда порядка 0,1–10 с. Следовательно, фоторезисторы должны обладать значительной инерционностью. Так как время установления стационарного состояния у них достаточно велико, то для регистрации быстро протекающих процессов фоторезисторы непригодны.

3.20.2. Фотодиоды, фотоэлементы

Физический принцип работы фотодиодов и фотоэлементов основан на использовании фотогальванического эффекта. Суть его заключается в следующем.

При освещении потоком электромагнитного излучения p - n перехода и прилегающих к нему областей полупроводника происходят переходы электронов из валентной зоны в зону проводимости. В результате в полупроводнике возникает дополнительное число возбужденных электронов и дырок. Следовательно, как в валентной зоне, так и в зоне проводимости появляется дополнительное число носителей заряда. Процесс появления дополнительного числа носителей заряда соответственно в зонах проводимости и валентной под действием излучения называется фотогальваническим эффектом.

Пусть имеется образец, состоящий из контактирующих между собой полупроводников n - и p -типа (рис. 3.20.1). Поверхность левого полупроводника облучается потоком электромагнитной волны.

Предположим, что толщина n -полупроводника $x_n \ll \ell_n$, ℓ_n – диффузионная длина квазиэлектронов в n -полупроводнике. Кроме того, $x_n \ll x_p$, x_p – толщина p -полупроводника, $\ell_n > L$, L – толщина n - p перехода.

Пусть Φ_0 – падающий поток излучения, r – коэффициент отражения потока от n - p полупроводника. Тогда поглощенный в полупроводниках поток

$$\Phi = (1-r)\Phi_0 = \Phi_n + \Phi_p,$$

где Φ_n , Φ_p – поглощенные потоки излучения соответственно в n - и p -полупроводниках.

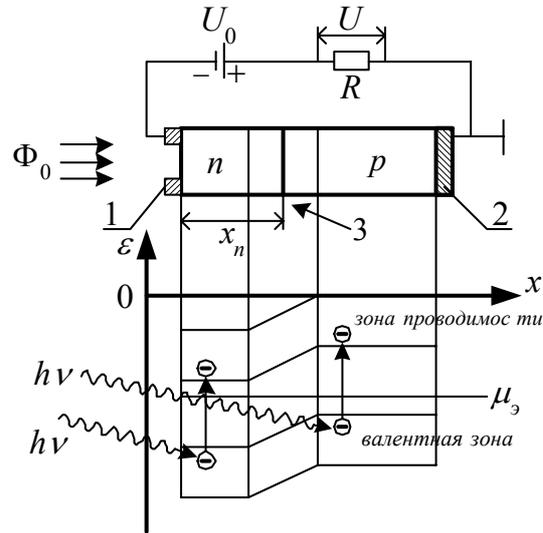


Рис. 3.20.1. Фотоэлемент: 1 - контактное металлическое кольцо, 2 - металлическое основание, 3 - область n - p перехода; $x_n \ll l_n$

Следовательно, в результате облучения значительная часть падающего первичного потока проникает в p -образец и в нем поглощается. Если в p -полупроводнике произошло поглощение кванта энергии излучения валентным электроном, то при величине энергии кванта больше ширины запрещенной зоны электрон может перейти из валентной зоны в зону проводимости, но одновременно с этим в валентной зоне образуется дырка. Возникшие в зоне проводимости p -полупроводника квазиэлектроны будут неосновными носителями заряда. Что же касается n -образца, в котором также происходит поглощение излучения, то поглощенный там поток Φ_0 будет возбуждать электроны валентной зоны, в результате некоторое их число перейдет в зону проводимости с одновременным возникновением в валентной зоне дырок. Но так как $x_n \ll x_p$, x_p - толщина p -образца, то $\Phi_n \ll \Phi_p$. Следовательно, концентрация вновь возникших квазиэлектронов и дырок в p -образце должна значительно превышать аналогичные концентрации в n -образце.

Возникшие в зоне проводимости p -образца квазиэлектроны движутся хаотически, и некоторое их число проникает в область контактного слоя p - n перехода, там они ускоряются электрическим полем контакта и перебрасываются в n -образец. Дырки валентной зоны n -

полупроводника также движутся хаотически, и некоторое их число попадает в область контактного слоя и электрическим полем перебра- сывается в p -полупроводник. Из физической постановки задачи сле- дует, что потоки квазичастиц образованы неосновными носителями заряда. Заметим, основными носителями заряда в n -полупроводнике являются квазиэлектроны, а в p -полупроводнике – дырки. При диффу- зии в сторону контактного слоя основные носители заряда на своем пути встречают потенциальный барьер, поэтому значительное их чис- ло не в состоянии преодолеть его (рис. 3.20.1), в то же время неоснов- ные носители легко переносятся в противоположные полупроводники ускоряющим для них электрическим полем.

Итак, квазиэлектроны из p -полупроводника и дырки из n - полупроводника, неосновные носители заряда, переносятся в избы- точном количестве соответственно в n - и p -полупроводники. В ре- зультате n -полупроводник относительно p -полупроводника заряжает- ся отрицательно. А так как потенциал полупроводников первоначаль- но был положительным, то потенциальная энергия заряженных частиц уменьшается на величину $\Delta\varepsilon = -e\Delta U < 0$, e – электрический заряд ква- зичастиц. Если теперь p -полупроводник заземлить, то все уровни пол- ной энергии n -полупроводника в процессе переходов квазичастиц че- рез контактный слой будут постепенно опускаться до уровня динамиче- ского равновесия. При наличии динамического равновесия между потоками основных и неосновных носителей заряда величина потен- циального барьера (величина внутренней контактной разности потен- циалов) возрастает на ΔU .

Плотность фототока на границе контактного слоя, созданного не- основными носителями заряда,

$$j_{\varphi} = e n_{\text{отн}} \cdot G, \quad (3.20.2.1)$$

где $n_{\text{отн}}$ – относительное число неосновных носителей заряда, дости- гающих p - n перехода в процессе диффузии без учета рекомбинации,

$$G = \eta \frac{\Phi_p}{h\nu}, \quad (3.20.2.2)$$

η – квантовый выход фотоэффекта.

Как следует из физического смысла формулы (3.17.5), изменение концентрации неосновных носителей заряда в пределах контактного слоя

$$\Delta n = n_p \exp\left(-\frac{q \Delta U}{kT}\right), \quad (3.20.2.3)$$

где ΔU – изменение внутренней контактной разности потенциалов в результате появления в зоне проводимости p -полупроводника квазиэлектронов, n_p – концентрация добавочного числа квазиэлектронов в p -полупроводнике в результате облучения его потоком электромагнитного излучения.

Тогда плотность фототока в области контактного слоя равна току проводимости:

$$j_\phi = j_s \left(\exp\left(-\frac{q\Delta U}{kT}\right) - 1 \right), \quad (3.20.2.4)$$

где j_ϕ определяется из (3.20.2.1), j_s – плотность тока насыщения, тока, возникающего в результате теплового (термического) возбуждения, его величина может быть определена по формуле (3.17.13):

$$j_s = e \left(\frac{n_e D_e}{\lambda_e} + \frac{n_d D_d}{\lambda_d} \right).$$

Из формулы (3.20.2.4) находим величину изменения внутренней контактной разности потенциалов

$$\Delta U = \frac{kT}{e} \ln \left(\frac{j_\phi}{j_s} + 1 \right). \quad (3.20.2.5)$$

Так как по физическому смыслу ΔU – фотоЭДС, то $U_\phi = \Delta U$.

Итак, если на рис. 3.20.1 $U_0 = 0$, то полупроводник работает в режиме фотоЭДС, а сам полупроводник является фотоэлементом. Если же к p - n полупроводнику подключен источник тока и сопротивление нагрузки R , как это изображено на рис. 3.20.1, то такая электрическая схема работает в режиме фотодиода, а сам полупроводник называется фотодиодом.

Фотодиоды и фотоэлементы в настоящее время нашли широкое применение. В частности, фотодиоды применяют для оптического управления током, для измерения интенсивности электромагнитного излучения. Поскольку интенсивность излучения зависит от прозрачности среды, то такие датчики могут быть индикаторами запыленности воздуха, наличия газов, поглощающих данное излучение, протекания реакций в прозрачных растворах. Их применяют также в турникетах метро, в станках-автоматах для определения количества изготовленных на них деталей и т. д.

Существуют фотодиоды для различных участков спектра длин волн. Например, фотодиод на основе полупроводника $InSb$ обладает чувствительностью к инфракрасному излучению с длиной волны 10

мкм, поэтому его используют в приборах ночного видения, с помощью которых можно наблюдать различные объекты в темноте.

Фотодиоды обладают большим быстродействием, их время срабатывания достигает величины порядка 10^{-11} с. Поэтому их широко используют для передачи оптической информации, в волоконно-оптической связи, а также в ЭВМ.

В отсутствие внешнего источника тока фотодиод является фотоэлементом. Обычное значение фотоЭДС – примерно 0,5 В. Если отдельные фотоэлементы соединить последовательно, то получается солнечная батарея. Для повышения мощности солнечные батареи соединяют параллельно, их коэффициент полезного действия может достигать 60%. Солнечные батареи являются источниками тока бортовых систем космических аппаратов.

В некоторых случаях рекомбинация носителей заряда в диоде приводит к возникновению электромагнитного излучения. Такие диоды называются светодиодами. Принцип их работы следующий. При пропускании через светодиод прямого тока концентрация неосновных носителей заряда на границе *p-n* перехода оказывается больше равновесной. Диффундируя через *p-n* переход в глубь полупроводника *p*-типа, неосновные носители заряда при этом постепенно рекомбинируют с дырками. Этот процесс сопровождается высвечиванием квантов энергии, т.е. излучением потока электромагнитной волны. Для светодиодов геометрические размеры полупроводника подбираются такими, чтобы излучением квантов света происходило наружу. Величина свечения зависит от двух условий:

- от соотношения между излучательными и безизлучательными рекомбинациями неосновных носителей заряда;
- от величины угла полного внутреннего отражения.

Дело здесь в том, что при малом угле полного внутреннего отражения выход излучения за пределы поверхности светодиода будет очень малым, значительная часть потока излучения будет отражаться на внешней границе раздела. Так, у полупроводника на основе *GaAs* угол полного внутреннего отражения примерно равен 17° . В результате лишь 2% испускаемого излучения выходит наружу.

В светодиодах, созданных на основе полупроводника *GaAs*, ширина запрещенной зоны $\Delta\varepsilon = 1,4$ эВ. Поэтому в результате рекомбинаций квазиэлектронов и дырок возникает инфракрасное излучение с длиной волны 0,89 мкм.

В тройном полупроводнике *GaAsP* можно изменять ширину запрещенной зоны и соответственно длину волны излучения. Так, на-

пример, из светодиода с красным свечением с помощью других примесей можно получить зеленый, синий и другие цвета свечения.

На рис. 3.20.2 изображена схема оптронной пары, с помощью которой можно осуществлять преобразование напряжения в электрических схемах.

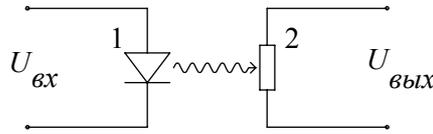


Рис. 3.20.2. Оптронная пара: 1 – светодиод, 2 – фотоприемник (фотодиод, фоторезист и др.)

Из рис. 3.20.2 следует, что оптронная пара является своеобразным трансформатором напряжения, управляемого с помощью светового сигнала.

Светодиод как излучатель и фотодиод как приемник закрепляются в один блок. Такой блок называется оптроном.

Подчеркнем еще раз, различие между фотодиодом и светодиодом состоит в том, что у фотодиодов в процессе работы имеют место безизлучательные переходы неосновных носителей заряда, а у светодиодов имеют место излучательные переходы, когда в результате переходов квазиэлектронов из зоны проводимости в валентную излучается квант энергии электромагнитной волны. Что же касается различия между фотоэлементом и фотодиодом, то, помимо своего назначения, они еще различаются особенностями подключения к внешней электрической цепи, а также геометрическими размерами. А так как с помощью фотоэлементов, как правило, необходимо получить максимальную величину фототока, то освещаемая поверхность фотоэлемента должна быть достаточно велика.

3.21. Физические основы работы лазеров

3.21.1 *Время жизни атома в возбужденном состоянии*

Ранее мы ввели в рассмотрение систему уровней полной энергии, которые соответствуют стационарным состояниям электрона в атоме. Это означало следующее. Представим себе атом как замкнутую систему не взаимодействующих с окружающей средой частиц. Тогда электрон атома на стационарном уровне в отсутствие внешних воз-

действий может находиться неограниченное время, при этом, двигаясь вокруг ядра, электрон не излучает электромагнитную энергию. Однако такой идеальной замкнутой системы в природе не существует. Во-первых, даже у простейшего из атомов – атома водорода – в основном состоянии уровень $1s$ оказывается несколько размытым. Это происходит из-за взаимодействия электрона с ядром в процессе движения относительно центра инерции атома. Величина такого размытия экспериментально наблюдается и называется лэмбовским сдвигом. Стало быть, строго говоря, в природе не существует дискретных, строго определенных уровней полной энергии. Во-вторых, возбужденный атом всегда находится во взаимодействии с окружающей средой, и это взаимодействие имеет случайный характер, в результате такого взаимодействия атом в какой-то момент времени самопроизвольно переходит в менее возбужденное состояние, излучая при этом в окружающее пространство квант электромагнитной энергии. Такое излучение, происходящее в отсутствие заметных электромагнитных воздействий со стороны окружающей среды, называется самопроизвольным, или спонтанным.

Определим время жизни атома в возбужденном состоянии на примере двухуровневой системы полных энергий. Пусть электрон атома в момент времени t обладает энергией ε_2 . Тогда под действием случайных причин, механизм которых детально проследить невозможно, электрон самопроизвольно переходит в состояние $\varepsilon_1 < \varepsilon_2$. Такой переход осуществляется с некоторой вероятностью ω .

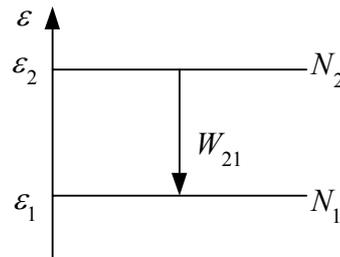


Рис. 3.21.1. Система двухуровневых атомов

Обозначим вероятность спонтанного перехода за 1 с W_{21} , или, что то же самое, скорость спонтанного распада, буквой W_{21} (рис. 3.21.1). Если в веществе в момент времени t в состоянии ε_2 находилось N_2

атомов, то за время dt вероятность перехода в состояние ε_1 будет $d\omega = W_{21} dt$.

Вероятность спонтанного перехода атома, находящегося в возбужденном состоянии, за среднее время жизни τ_{cn} в возбужденном состоянии

$$\int_0^{\tau_{cn}} W_{21} dt = 1. \quad (3.21.1)$$

Следовательно, среднее время жизни в возбужденном состоянии при спонтанных переходах ($W_{21} = \text{const}$)

$$\tau_{cn} = (W_{21})^{-1}.$$

За время dt в состояние ε_1 перейдет dN_2 атомов:

$$dN_2 = -N_2 W_{21} dt,$$

причем знак (-) указывает на убыль числа атомов на уровне ε_2 . Тогда после интегрирования получаем

$$N_2 = N_{20} e^{-W_{21} t}, \quad (3.21.2)$$

где N_{20} – число атомов на уровне ε_2 в начальный момент времени $t=0$. Каждый такой переход сопровождается излучением кванта электромагнитной энергии. Поэтому

$$\varepsilon_2 - \varepsilon_1 = h\nu.$$

Энергия электромагнитного излучения атомов dN_2 за время dt будет

$$d\varepsilon = h\nu W_{21} N_2 dt,$$

а поток электромагнитной энергии, возникающий в результате таких переходов, с учетом (3.21.2) будет

$$\Phi = \frac{d\varepsilon}{dt} = h\nu W_{21} N_2,$$

или

$$\Phi = \Phi_0 e^{-W_{21} t}. \quad (3.21.3)$$

Таким образом, поток электромагнитной энергии, возникающий в результате спонтанных переходов атомов из возбужденного в невозбужденное состояние, с течением времени убывает по экспоненте.

Подставим найденной значение $t = \tau_{cn}$ в уравнение (3.21.3): $\Phi(\tau_{cn}) = \Phi_0 e^{-1}$. Это означает, что за время, равное средней продолжительности жизни атома в возбужденном состоянии на уровне ε_2 , поток квантов электромагнитного излучения уменьшается в e раз, следовательно, число атомов, находящихся в состоянии ε_2 , также уменьшает-

ся в e раз. Опыт показывает, что в видимой области спектра $\tau_{\text{сп}} \approx 10^{-8}$ с, стало быть, $W_{21} \approx 10^8 \text{ с}^{-1}$.

3.21.2. Принцип детального равновесия

В квантовой системе, обладающей дискретными уровнями энергии, существуют три типа переходов электронов между энергетическими состояниями: индуцированные электромагнитным полем переходы или, что тоже самое, вынужденные переходы, спонтанные переходы и безизлучательные переходы. При индуцированных переходах квантовая система может переходить из одного энергетического состояния в другое (рис. 1) как с поглощением энергии электромагнитного поля (переход с нижнего энергетического уровня на верхний), так и с излучением электромагнитной энергии (переход с верхнего энергетического уровня на нижний), при этом исходный квант сохраняется, а возникший квант имеет те же физические свойства, что и исходный.

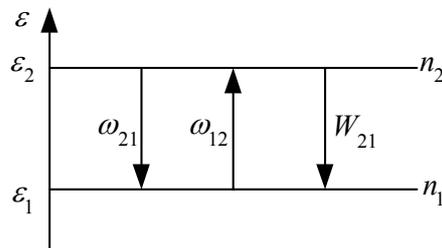


Рис. 3.21.2. Переходы в двухуровневой системе: n_1, n_2 – населенности нижнего и верхнего уровней энергии; ω_{21}, ω_{12} – скорости индуцированных переходов (число переходов за 1 с); W_{21} – вероятность спонтанного перехода за 1с.

Индуцированные переходы обладают следующими важными свойствами.

1. Скорости индуцированных переходов ω_{21}, ω_{12} отличны от нуля только для внешнего электромагнитного поля резонансной частоты. Следовательно,

$$h\nu = \varepsilon_2 - \varepsilon_1.$$

2. Возникающие при индуцированных переходах кванты электромагнитного поля полностью тождественны квантам поля, вызвавшего

эти переходы. Значит, внешнее электромагнитное поле и поле, возникшее в результате индуцированных переходов, имеют одинаковую частоту, фазу колебаний, поляризацию и направление распространения, т.е. они физически неразличимы, когерентны.

3. Число индуцированных переходов за 1 с пропорционально плотности энергии внешнего поля $\rho_{\lambda,T}$ в единичном спектральном интервале длин волн:

$$\omega_{12} = g_1 B_{12} \rho_{\lambda,T}, \quad \omega_{21} = g_2 B_{21} \rho_{\lambda,T}, \quad (3.21.4)$$

где B_{12} , B_{21} – коэффициенты Эйнштейна для индуцированного поглощения и излучения; g_1 , g_2 – кратность вырождения уровней ε_1 и ε_2 ;

$$\rho_{\lambda,T} = \frac{\Phi_{\lambda,T}}{V},$$

$\Phi_{\lambda,T}$ – спектральный поток электромагнитной энергии, V – объем вещества, в пределах которого распространяется этот поток. Благодаря наличию вынужденного излучения в веществе всегда происходит усиление первичного излучения без изменения его спектрального состава.

Но, кроме индуцированного излучения, еще существует спонтанное излучение, вероятность которого не зависит от величины внешнего электромагнитного поля. Оно некогерентно по отношению к внешнему полю. Своим действием спонтанное излучение лишь опустошает верхний энергетический уровень и способствует возвращению атомов в менее возбужденное состояние.

Предположим, что в объеме, заполненном излучением абсолютно черного тела, находится газ, концентрация молекул которого настолько мала, что взаимодействием между ними на расстоянии можно пренебречь. Распределение по энергиям молекул такого газа описывается статистикой Максвелла-Больцмана. Так как спонтанное излучение не зависит от плотности энергии излучения $\rho_{\lambda,T}$, то число спонтанных переходов W_{21} пропорционально числу атомов, находящихся на уровне ε_2 : $N_{cn} = W_{21} N_2$. Число индуцированных переходов зависит от числа атомов, соответственно находящихся в состояниях ε_1 и ε_2 .

Тогда с учетом (3.21.4) можно записать:

$$N_{ind1} = g_1 B_{12} N_1 \rho_{\lambda,T}, \quad N_{ind2} = g_2 B_{21} N_2 \rho_{\lambda,T}.$$

В состоянии динамического равновесия имеет место равенство между числом актов испускания и числом актов поглощения квантов энергии. А так как уровни ε_1 и ε_2 выбраны произвольно, то говорят, что в объеме имеет место детальное равновесие для переходов между уровнями ε_1 и ε_2 . При наличии детального равновесия

$$N_{инд1} = N_{сн} + N_{инд2},$$

т. е.

$$g_1 B_{12} N_1 \rho_{\lambda,T} = W_{21} + g_2 B_{21} \rho_{\lambda,T}. \quad (3.21.5)$$

Отсюда

$$\rho_{\lambda,T} = W_{21} \left(g_1 B_{12} \frac{N_1}{N_2} - g_2 B_{21} \right)^{-1}.$$

В стационарных условиях в веществе имеет место термодинамическое равновесие. Поэтому атомы N_1 и N_2 распределяются по энергетическим уровням в соответствии с распределением Максвелла-Больцмана

$$N_1 = N_2 \exp\left(\frac{hc}{\lambda kT}\right).$$

Тогда

$$\rho_{\lambda,T} = W_{21} \left(g_1 B_{12} e^{\frac{hc}{\lambda kT}} - g_2 B_{21} \right)^{-1}. \quad (3.21.6)$$

Как известно, плотность энергии равновесного излучения абсолютно черного тела

$$\rho_{\lambda,T} = \frac{8\pi hc}{\lambda^5} \left(e^{\frac{hc}{\lambda kT}} - 1 \right)^{-1}. \quad (3.21.7)$$

Выражения (3.21.6) и (3.21.7) будут идентичны, если выполняются следующие соотношения Эйнштейна:

$$g_1 B_{12} = g_2 B_{21} = B; \quad \frac{W_{21}}{B} = \frac{8\pi hc}{\lambda^5}. \quad (3.21.8)$$

Так как по физическому смыслу $B = \frac{c_1}{\tau_{инд}}$, $W_{21} = \frac{c_1}{\tau_{сн}}$, то

$$\tau_{инд} = \frac{8\pi hc}{\lambda^5} \tau_{сн} \quad (3.21.9)$$

Таким образом, время жизни атома при наличии индуцированного излучения зависит от длины волны, для коротких длин волн оно больше, чем для длинных. Вычисления $\tau_{сн}$ и $\tau_{инд}$ по формуле (3.21.9) для двух длин волн сведены в таблицу значений, причем $\tau_{сн}$ определялось по формуле: $\tau_{сн} = 4,7 \cdot 10^4 \lambda^2$, $[\lambda] = \text{м}$.

λ	$\tau_{\text{сп}}$	$\tau_{\text{инд}}$	$\tau_{\text{инд}}/\tau_{\text{сп}}$
нм	с		
400	$7,5 \cdot 10^{-9}$	0,37	$4,9 \cdot 10^7$
800	$3 \cdot 10^{-8}$	0,046	$1,5 \cdot 10^6$

Из таблицы видно, что время жизни атомов при индуцированных переходах во много раз превышает время жизни при спонтанных переходах. Поэтому состояния атомов, находящихся на уровнях, с которых происходят индуцированные переходы, считают метастабильными (долгоживущими).

Итак, условию детального равновесия соответствует тождество (3.21.5), которое выполняется, когда на уровне ε_1 находится N_1 атомов, а на уровне ε_2 находится N_2 атомов. Если подобные рассуждения провести не для газа, а для твердого тела, в котором присутствует незначительная концентрация атомов примеси, то приближенно также можно применять распределение Максвелла-Больцмана для числа атомов на отдельных примесных уровнях полной энергии. Следовательно, и в этом случае будет справедливо тождество (3.21.5), а также все физические выводы, установленные ранее для газов.

3.21.3. Основные принципы работы лазеров

Термин "лазер" образован из начальных букв английского выражения Light Amplification by Stimulated Emission of Radiation, что в переводе означает "усиление света вынужденным излучением". В 1954 г. советским физикам Н.Г. Басову и А.М. Прохорову и независимо от них американцу Ч. Таунсу удалось создать такой источник электромагнитного излучения. Первый лазер действовал в радиодиапазоне, а первая генерация в оптическом диапазоне была получена в 1960 г. американскими физиками Т. Мейманом на рубине и А. Джаваном с сотрудниками в гелий-неоновой смеси. За пионерские работы по созданию лазера в 1964 г. Н.Г. Басову, А.М. Прохорову и Ч. Таунсу была присуждена Нобелевская премия по физике.

Основными свойствами лазерного излучения, отличающими его от излучения других источников света, являются:

- 1) высокая когерентность;
- 2) возможность генерации импульсов предельно малой длительности $\approx 10^{-14}$ с;
- 3) высокая мощность излучения в импульсном режиме генерации, до 10^{14} Вт.

Лазеры различаются:

1) по типу активной среды – газовые (атомные, ионные, молекулярные), твердотельные (на примесных кристаллах и стеклах, полупроводниковые), жидкостные;

2) по режиму генерации – непрерывные, импульсные.

Из всех перечисленных выше типов лазеров мы более подробно рассмотрим принцип работы только газового гелий-неонового и полупроводникового лазеров.

Газовые лазеры работают в очень широком интервале длин волн – от вакуумного ультрафиолета до инфракрасной области спектра, как в импульсном, так и в непрерывном режимах. Малая плотность и высокая однородность газообразной активной среды позволяет получить высокую направленность и моноэнергетичность излучения. Вместе с тем малая плотность возбужденных частиц не позволяет получить высокую концентрацию возбуждаемых частиц.

Для создания инверсной населенности отдельных уровней полной энергии в подавляющем большинстве газовых лазеров используется электрический разряд, в процессе которого электроны разряда возбуждают атомы газа, введенного в рабочий объем в качестве примеси к атомам основного газа. Далее передача энергии возбуждения от атомов примеси атомам основного газа осуществляется в результате неупругих столкновений между ними. Такая передача энергии тем более эффективна, чем более точно совпадают уровни полной энергии сталкивающихся между собой атомов. Процесс возбуждения протекает в два этапа. Сначала возбуждаются атомы вспомогательного газа, носители избыточной энергии и вместе с тем являющиеся донором энергии возбуждения, затем в результате неупругих столкновений атомов примеси с атомами основного газа энергия возбуждения передается атомам рабочего газа. В результате верхние уровни полной энергии населяются электронами атомов рабочего газа, при этом электрон атома вспомогательного газа на верхнем уровне должен обладать достаточно большим временем жизни, чтобы атомы примеси могли бы накапливать достаточно большую совокупную энергию для последующей передачи ее в процессе неупругих столкновений.

3.21.4. Гелий-неоновый лазер

В He-Ne лазере рабочей средой является смесь газов *He* и *Ne*, причем рабочим газом является *Ne*, а примесью – газ *He*. В объем атомы *He* вводятся для создания инверсной населенности. При правильно выбранных давлениях можно добиться заселения одного или обоих

уровней атомов Ne в смеси газов, значительно превышающих заселенность атомов в случае чистого Ne , и, следовательно, создать инверсию населенности на этих уровнях по сравнению с населенностью на уровне ε_3 (см. рис. 3.21.3). Заметим, на рис. 3.21.3 представлена лишь упрощенная схема уровней полных энергий атомов Ne . На самом деле каждое энергетическое состояние $\varepsilon_2 - \varepsilon_2$ состоит из нескольких близко расположенных состояний, условно они изображены в виде зоны энергий. Опустошение нижних уровней ε_2 происходит в столкновительных процессах как между атомами Ne , так и при столкновениях атомов Ne с поверхностью газоразрядной трубки.

Столкновительные процессы происходят следующим образом (см. рис. 3.21.3).

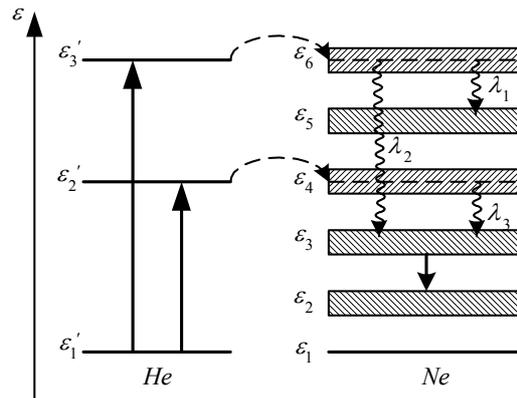


Рис. 3.21.3. Схема действия $He-Ne$ лазера

Уровни $\varepsilon_2, \varepsilon_4, \varepsilon_6$ – метастабильные; переход из состояния ε_2 в состояние ε_1 возникает в результате столкновений между атомами Ne , а также столкновений атомов Ne со стенкой газоразрядной трубки (безизлучательный переход); $\lambda_1 = 3,39$ мкм, $\lambda_2 = 0,6328$ мкм, $\lambda_3 = 1,15$ мкм – длины волн индуцированного (вынужденного) излучения.

В процессе столкновений атомов He со свободными электронами атомы He переходят из основного состояния ε_1' в возбужденные состояния ε_2' и ε_3' , соответствующие метастабильным уровням энергии со временем жизни $\approx 10^{-3}$ с.

Верхние уровни атомов Ne при столкновениях со свободными электронами также заселяются электронами, но время жизни на них

невелико. Поэтому за время порядка 10^{-8} с в обратном направлении происходит спонтанный переход. Наряду с этим, при столкновениях атомов *He* с невозбужденными атомами *Ne* интенсивно происходят безизлучательные переходы атомов *He* в невозбужденное состояние с резонансной передачей энергии атомам *Ne*. Этот процесс на рис. 3.21.3 условно изображен пунктирными линиями. В результате концентрация возбужденных атомов *Ne* на уровнях ε_4 и ε_6 сильно возрастает, возникает инверсная заселенность уровней ε_4 и ε_6 относительно заселенности уровней ε_3 и ε_5 . Если в объеме, допустим, возникнет квант вынужденного излучения, то при столкновении с инверсно заселенным атомом *Ne* в объеме возникнет новый квант излучения, тождественный первоначальному. В результате в объеме будут существовать два кванта излучения, и т.д. Таким образом, в объеме будет происходить лавинообразное нарастание квантов.

Но в веществе наряду с увеличением числа квантов всегда происходят необратимые потери энергии излучения в результате их резонансного поглощения атомами рабочего вещества, поглощения материалом стенки газоразрядной трубки. Пусть α – коэффициент квантового усиления, β – коэффициент потерь, тогда величина наибольшего потока энергии на выходе резонатора можно вычислить по формуле:

$$\Phi = \Phi_0 \exp((\alpha - \beta) \cdot \lambda),$$

где ℓ – расстояние между зеркалами резонатора.

Опустошение верхних уровней *Ne* ε_3 , ε_5 происходит за счет спонтанного излучения. Опустошение нижнего метастабильного уровня *Ne* ε_2 происходит за счет соударений атомов *Ne* со стенками газоразрядной трубки. Возбуждение атомов *He* и *Ne* осуществляется либо высокочастотным генератором, либо разрядным током, возникающим в объеме газоразрядной трубки при разности потенциалов между анодом и катодом порядка 1–2 кВ.

При газовом разряде частота, фаза и направление колебаний вектора \vec{E} в электромагнитной волне распределены хаотически, также хаотически в различных направлениях распространяется и электромагнитная волна. Поэтому для усиления интенсивности лазерного излучения применяют резонаторы – устройства, с помощью которого выделяют и усиливают излучение только одного направления и фазы. Простейший резонатор представляет собой два зеркала с общей оптической осью, в котором между зеркалами находится активное вещество, в данном случае смесь газов *Ne* и *He*. Одно зеркало имеет почти 100%-ный коэффициент отражения, другое – примерно 99%. В такой системе зеркал будут усили-

ваются только те потоки электромагнитных волн, которые распространяются строго вдоль оптической оси. Усиление происходит в результате многократного прохождения квантов индуцированного излучения через активную среду после отражения от зеркал, но на каждом акте отражения от зеркала с коэффициентом отражения порядка 99% некоторый поток излучения выходит за пределы резонатора. Поэтому, если случайно возникший пучок индуцированного излучения составляет с оптической осью даже незначительный угол, то после отражения он уже выбывает из активной среды. В результате за пределы резонатора выходит монохроматическое излучение, обладающее строгой направленностью и когерентностью.

В настоящее время наибольшее распространение получила конструкция лазера, у которого газоразрядная трубка имеет окна, расположенные под углом Брюстера к оптической оси. Она помещается внутрь резонатора. Это позволяет исключить повреждение поверхности зеркал резонатора газовым разрядом. Кроме того, при таком расположении проще юстировать зеркала резонатора, увеличивается срок службы газоразрядной трубки и зеркал, облегчается смена зеркал.

Минимальная расходимость лучей в пучке достигает $5 \cdot 10^{-4}$ рад. Когерентность в лазерном пучке позволяет получать цуги волн длиной в десятки и сотни метров, что позволяет значительно увеличить вероятность индуцированного излучения. На рис. 3.21.3 указаны длины волн, соответствующие вынужденным (индуцированным) излучениям. Изменением расстояния между зеркалами резонатора до величины, кратной половине длины волны индуцированного излучения, можно выделить излучение необходимой длины волны.

Несмотря на не очень большую мощность (от 1 мВт до 50 мВт) и низкий коэффициент полезного действия (от 0,1% до 0,3%), He-Ne лазеры вследствие хорошей направленности и монохроматичности лазерного луча нашли достаточно широкое применение.

3.21.5. Полупроводниковые лазеры

Ранее мы показали, что полупроводниками являются не только кристаллические *Si* и *Ge* (собственные полупроводники), но и те же кристаллические вещества с внедренными в их кристаллическую решетку атомами примесей из соседних групп периодической системы Менделеева. Так, полупроводники на основе *Si* с примесью атомов *P*, *As* можно условно обозначить как полупроводники группы A^4B^5 , а с примесью атомов *Ga*, *In* – как полупроводники группы A^3B^4 . В настоящее время установлено, что полупроводниковыми свойствами

обладают полупроводники групп A^3B^5 , такие как *AsGa*, *AsIn* и другие, а также групп A^4B^6 , такие как *SPb*, *SePb* и другие.

В инжекционных полупроводниковых лазерах активным веществом служит кристаллическая матрица собственного полупроводника, а примеси являются источником носителей заряда. При создании в полупроводнике с помощью накачки избыточной по сравнению с равновесной концентрации квазиэлектронов и дырок возникает оптический межзонный переход – фоторекомбинация квазиэлектронов и дырок. Оказывается, что вероятность фоторекомбинации велика лишь для прямозонных полупроводников, т.е. таких, у которых максимум полной энергии в валентной зоне и минимум полной энергии в зоне проводимости основных носителей заряда соответствуют одному и тому же значению квазиимпульса $p = \eta k$, $k = \frac{2\pi}{\lambda}$, k – величина волнового

вектора (волновое число). По этой причине все инжекционные полупроводники, на которых получена генерация лазерного излучения, являются прямозонными.

Перечислим важнейшие достоинства инжекционных полупроводниковых лазеров.

1. Большой коэффициент усиления при малых геометрических размерах.
2. Большой КПД, достигающий до 40%.
3. Широкий интервал длин волн, в котором возможна генерация лазерного излучения (от 300 нм до 30 мкм).
4. Возможность перестройки длины волны генерации путем изменения ширины запрещенной зоны в результате изменения давления, температуры, магнитного поля.
5. Малая инерционность, позволяющая модулировать излучение изменением частоты накачки до 10 ГГц.
6. Простота конструкции, поскольку процесс инверсии населенности осуществляется постоянным током.

Недостатками инжекционных полупроводниковых лазеров являются:

1. низкая выходная мощность,
2. значительная чувствительность к перегрузкам.

Для создания инверсной населенности в основном применяют либо инжекцию носителей заряда, либо электронную накачку. Резонаторами обычно служат параллельно расположенные кристаллографические плоскости на поверхностях полупроводника. В этом случае коэффициент отражения от таких кристаллографических плоскостей

оказывается достаточным для возбуждения генерации лазерного излучения без использования дополнительных зеркал.

На рис. 3.21.4, а изображены соответственно графики нижней и верхней границ зон проводимости и валентной для прямозонного полупроводника, а на рис. 3.21.4, б изображены те же графики для полупроводников A^3B^5 , A^4B^6 , например, соответственно для $AsGa$ и SPb .

Так как элементарные полупроводники A^3B^5 или A^4B^6 не являются прямозонными, то они и не применяются в инжекционных полупроводниковых лазерах.

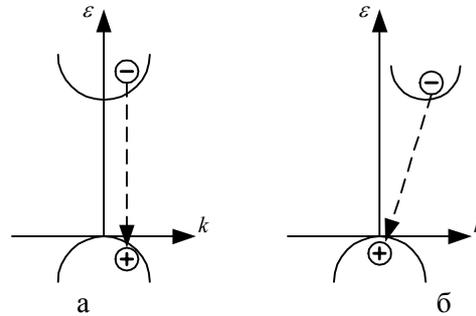


Рис. 3.21.4. Энергетические зоны в полупроводниковых лазерах: ε – полная энергия носителей заряда, k – волновое число.

При переходе электрона из зоны проводимости в валентную выполняется закон сохранения импульса:

$$\vec{p}_{e1} = \vec{p}_{e2} + \vec{p}_{\phi},$$

где \vec{p}_{e1} , \vec{p}_{e2} – импульсы электрона соответственно в зоне проводимости и в валентной зоне, \vec{p}_{ϕ} – импульс фотона. Так как импульс фотона значительно меньше импульса электрона в валентной зоне, то переходы, связанные с излучением фотона, можно изобразить в виде вертикальной линии. В случае же Si и Ge места на графике $\varepsilon(k)$, занятые квазиэлектроном и дыркой, сдвинуты относительно друг друга на некоторую величину Δk (рис. 3.21.4, б). В этом случае возможен только непрямой переход, при котором изменение импульса электрона компенсируется кристаллической решеткой. Это означает, что добавочное изменение импульса принимает на себя фотон кристаллической решетки. Следовательно, такие переходы являются безизлучательными.

Итак, переход из зоны проводимости в валентную у прямозонных полупроводников сопровождается испусканием кванта электромаг-

нитной энергии, а обратный процесс возможен только при резонансном поглощении квантов. Поэтому, чтобы излучение преобладало над поглощением, необходимо создать инверсную населенность. Идеальный случай инверсной населенности возникает, если верхний край валентной зоны оказывается полностью занятым дырками. Тогда вероятность перехода с излучением кванта энергии будет достаточно велика. В случае же непрямозонного полупроводника носители заряда накапливаются в состояниях с различными значениями квазиимпульса, различными значениями волнового числа. В результате прямые переходы оказываются запрещенными или маловероятными, и рекомбинации квазичастиц не возникает.

Соединения A^3V^5 образуются в результате взаимодействия атомов V^3 группы периодической таблицы Менделеева (*B, Al, Ga, Ta*) с атомами V^5 группы (*N, P, As, Sb*). *Tl* и *Bi* не образуют соединений рассматриваемого ряда.

Соединения A^3V^5 принято классифицировать по металлоидному элементу, а именно, *N* (нитриды), *P* (фосфиды), *As* (арсениды), *Sb* (антимониды).

Для соединений A^3V^5 характерен донорно-акцепторный тип связи, при котором из четырех ковалентных связей три связи образуются в результате обобществления валентных электронов атомов A^3 и V^5 , а четвертая связь осуществляется двумя валентными электронами атомов V^5 . Образование этой связи соответствует энергетически выгодному переходу электронов от атомов V^5 в энергетическое состояние, общее для атомов V^5 (доноры) и атомов A^3 (акцепторы). В каждой ковалентной связи максимум электронной плотности смещен в сторону атома V^5 . Благодаря такой поляризации связей атомы A^3 приобретают эффективный положительный заряд, а атомы V^5 – отрицательный.

За исключением антимонидов, все соединения A^3V^5 при нагревании разлагаются по схеме:



Примеси замещения в кристаллической решетке A^3V^5 должны быть распределены таким образом, чтобы в ней не возникало бы центров с большим избыточным зарядом. Поэтому примеси II группы (*Be, Mg, Zn, Cd*) всегда занимают в кристаллической решетке A^3V^5 узлы металлического компонента и являются акцепторами благодаря меньшей валентности по сравнению с замещаемыми атомами. Приме-

си VI группы (*S*, *Se*, *Te*) всегда располагаются в узлах кристаллической решетки, замещая V^5 , и являются донорами.

Ценным свойством многих полупроводников типа $A^3 B^5$ является высокая эффективность излучения рекомбинации неравновесных носителей заряда. Для генерации излучения в видимой области спектра ширина запрещенной зоны должна быть больше 1,7 эВ. Такому условию удовлетворяют *GaP* и *GaN*. Материалы с более узкой запрещенной зоной способны эффективно излучать в инфракрасной области спектра. К их числу относится *GaAs* с шириной запрещенной зоны 1,43 эВ. Излучение фотонов в *GaAs* происходит в результате прямой межзонной рекомбинации электронов и дырок. При легировании *GaN* атомами *Zn* в зависимости от концентрации примесей *Zn* можно получить рекомбинационное излучение в видимой области спектра (желтое, зеленое, голубое излучение).

Первый инжекционный полупроводниковый лазер был создан в 1962 году. В нем генерация когерентного излучения возникала в кристалле арсенида галлия *GaAs*. Но лазерный эффект проявляется здесь в том случае, если плотность тока через *p-n* переход превышает некоторое пороговое значение, в противном случае это устройство будет работать в режиме светодиода.

Известно, что кристалл *GaAs* представляет собой твердый раствор, образованный атомами *Ga* и *As*. Поскольку понятие твердого раствора в общем курсе физики не объясняется, то приведем некоторые основные сведения о типах и физическом строении твердых растворов. Твердые растворы – это твердые фазы сплавов, в которых соотношения между концентрациями компонентов могут быть изменены без нарушения однородности вещества. В природе существуют три основных типа твердых растворов: замещения, внедрения, вычитания. В твердых растворах замещения атомы растворенных компонентов замещают атомы растворителя. Если же твердый раствор – химическое соединение, то реализуется схема, когда кристаллическая решетка образуется молекулами АВ. В твердых растворах внедрения атомы растворенного компонента располагаются в промежутках между атомами решетки растворителя. Твердый раствор вычитания возникает только на основе химических соединений, причем в этом случае всегда некоторое число узлов кристаллической решетки оказывается вакантным.

В настоящее время твердые растворы широко используются для изготовления инжекционных полупроводниковых приборов. Необходимыми условиями образования таких растворов являются кристаллохимическое подобие решеток компонентов раствора и близость их

кристаллографических индексов. Наиболее хорошо изучены тройные твердые растворы, в которых замещение происходит лишь по узлам одной из подрешеток металлического или металлоидного бинарного соединения. Заметим, что под словом "подрешетка" понимается система периодически расположенных в твердом теле одинаковых атомов или ионов, а под словами "бинарное соединение" – соединения типа АВ, например, *GaAs* или *InP*, т.е. соединения без примесных атомов. Состав таких твердых растворов принято характеризовать символами $A_xB_{1-x}C$ и AC_yD_{1-y} , где символами А и В обозначаются элементы III группы, а символами С и D – элементы V группы. В формуле $A_xB_{1-x}C$ индексом x определяют мольную долю соединения АВ в твердом растворе. Если твердые растворы существуют во всем интервале изменения концентраций, то x может меняться от 0 до 1. Как и в бинарных соединениях типа A^3B^5 , в твердых растворах не наблюдается существенных отклонений от стехиометрии, поэтому они просты по механизму легирования, т.е. у них теми же методами, что и в бинарных соединениях, может быть получен *p-n* переход.

Особый интерес к твердым растворам, в частности, обусловлен возможностью плавного управления шириной запрещенной зоны полупроводника путем изменения их компонентного состава. Так, зависимость ширины запрещенной зоны от состава твердых растворов $Ga_xIn_{1-x}As$ или InP_yAs_{1-y} очень близка к линейной, что весьма удобно при изготовлении полупроводников с заданными спектральными характеристиками.

По мере развития технологии полупроводниковых материалов появилась возможность создания в полупроводнике неоднородность состава. Переходный слой между двумя полупроводниками с разным составом называют гетеропереходом. В гетеропереходах потенциальные барьеры для электронов и дырок отличаются друг от друга по величине, у этих полупроводников различна ширина запрещенной зоны. При наличии внешней разности потенциалов протекающий через границу раздела электрический ток создается носителями заряда только одного знака, например, квазиэлектронами.

Для получения гетероперехода со свойствами идеального контакта необходимо выполнить ряд условий совместимости материалов по механическим, кристаллографическим и термическим свойствам. Решающим критерием при выборе материалов контактной пары является соответствие постоянных кристаллических решеток и температурных коэффициентов расширения. Среди полупроводников типа A^3B^5 лучшими парами материалов для создания идеальных гетероперехо-

дов являются $GaAs - Al_x Ga_{1-x} As$ и $GaSb - Al_x Ga_{1-x} Sb$. Преимущества этих гетеропар состоят в том, что величина постоянной кристаллической решетки твердых растворов как у $Al_x Ga_{1-x} As$, так и у $Al_x Ga_{1-x} Sb$ слабо зависит от состава и близка к постоянной кристаллической решетки бинарного соединения как $GaAs$, так и $GaSb$ соответственно.

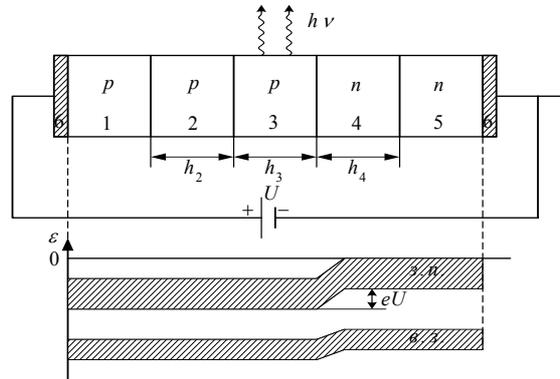


Рис. 3.21.5. Лазер с двойной гетероструктурой на $GaAs - Al_x Ga_{1-x} As$:
 1. $n-GaAs$; 2. $p-Al_x Ga_{1-x} As$; 3. $p-GaAs$; 4. $n-Al_x Ga_{1-x} As$;
 5. $n-GaAs$; 6. омические (электрические) контакты;
 $0,2 \text{ мкм} \leq h_3 \leq 1 \text{ мкм}$; $h_2 = h_4 = 1 \text{ мкм}$

На рис. 3.21.5 приведена схема инжекционного лазера с двойной гетероструктурой. Область рекомбинации и электромагнитного излучения сосредоточена в узкозонном активном слое $p-GaAs$, заключенном между двумя широкозонными эмиттерами $Al_x Ga_{1-x} As$. При прямой внешней разности потенциалов в такой структуре имеет место инжекция носителей заряда в активный слой. В результате удается достигнуть очень высокого квантового выхода и существенно снизить пороговую мощность тока, необходимую для генерации когерентного излучения. Снижение порогового тока позволяет осуществить непрерывный режим генерации даже при комнатной температуре, которую не удастся реализовать в лазерах на гомогенных структурах с $p-n$ переходом.

Так как толщина переходного слоя, в котором генерируется лазерное излучение, достаточно мала, то мощность лазерного излучения всегда оказывается незначительной. Так, выходная мощность инжекционного полупроводникового лазера на гетеропереходах в режиме непрерывной генерации при комнатной температуре не превышает 100 мВт.

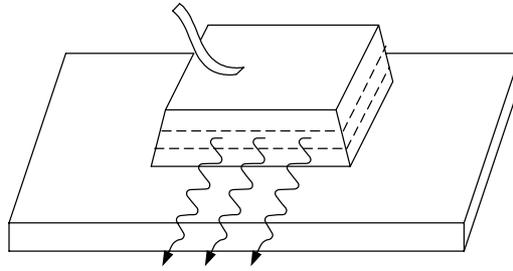


Рис. 3.21.6. Схематический вид конструкции полупроводникового инжекционного лазера

Конструкция инжекционного лазера следующая. Кристаллу полупроводника с линейными размерами порядка десятых долей мм придают форму усеченной четырехгранной пирамиды (рис. 3.21.6). Две боковые грани строго параллельны, их поверхности отполированы, и поэтому они представляют собой оптический резонатор. Для исключения оптического резонанса между двумя другими гранями их сшивают под углом к основанию.

Основные области применения:

1. передача информации на большие расстояния;
2. измерение расстояний, линейных и угловых скоростей, деформации поверхности и т.д. на основе использования интерференции лазерного излучения;
3. использование в качестве стандарта частоты;
4. изготовление логических элементов высокого быстродействия с временем быстродействия до 10^{-10} с для сверхбыстродействующих ЭВМ.

Первые патенты на гетеропереходы были получены В. Шокли и Г. Кремером (оба из США), которые исследовали гетеропереходы теоретически. В начале шестидесятых годов в Физико-техническом институте (г. Ленинград) была выпущена монография Н.Н. Губанова по теории гетеропереходов. Тогда многие полагали, что физика гетеропереходов остается на уровне идей, что реализовать гетеропереходы не удастся, потому что на границе двух полупроводников с разным составом будет сильно искажаться кристаллическая решетка, будут накапливаться примеси, не будет одностороннего протекания носителей заряда через гетеропереход. Тем не менее, Ж.И. Алферов в Ленинградском ФТИ начал исследование гетеропереходов системы *GaAs–GaAlAs*. Оказалось, что в этой системе полупроводников в переходном

слое возникает очень мало искажений кристаллической решетки, которые сильно влияют на эффективность работы приборов. На основе работ группы Ж.И. Алферова были созданы полупроводниковые лазеры, работающие при комнатной температуре в режиме постоянного тока, сначала в инфракрасном, а затем в видимом диапазонах длин волн. В этих полупроводниках при рекомбинации возбужденных электронно-дырочных пар энергия возбуждения передается в основном квантам света, т.е. происходит излучательная рекомбинация. Поэтому развитие физики и технологии гетеропереходов оказалось наиболее важным для приборов оптоэлектроники – для светодиодов, полупроводниковых лазеров и фотоприемников.

В начале семидесятых годов ФТИ направил Ж.И. Алферова на стажировку в США в лабораторию профессора Ника Холоньяка. Н. Холоньяк происходил из семьи, эмигрировавшей в Америку из Украины в двадцатых годах, хорошо знал русский язык и был знаком с работами Ж.И. Алферова по гетеропереходам. В это время ученые из фирмы "Белл" также сделали гетеропереходы на основе арсенида галлия, и их работы были представлены на золотую медаль Американского физического общества. Заявка попала на отзыв Н. Холоньяка, который в своем отзыве написал, что такие переходы на 2–3 года раньше были сделаны Ж.И. Алферовым. В результате золотая медаль Американского физического общества была присуждена Ж.И. Алферову.

В декабре 2000 г. за разработку и практическое использование гетеропереходов академику Ж.И. Алферову вместе с двумя американскими учеными Г. Кремером и Дж. Килби была присуждена Нобелевская премия по физике, причем Г. Кремеру – за теоретические исследования в области гетеропереходов, а Дж. Килби – за разработку методики и технологии микросхем с гетеропереходом на одной пластине, которая называется "чипом".

3.22. Эмиссионные явления

В зависимости от особенностей внешних физических воздействий различают следующие виды эмиссии заряженных частиц с поверхности твердого тела: термоэлектронная, термоионная, фотоэлектронная, вторичная электронная, вторичная ионная, полевая (туннельная). Во всех видах эмиссий, кроме полевой, роль внешних воздействий сводится к увеличению энергии у некоторого числа электронов или ионов твердого тела до значения, позволяющего преодолеть действие сил,

которое связывает их с телом, и вылететь в вакуум или в другую среду, т. е. преодолеть существующий на границе раздела двух сред потенциальный барьер.

Для наблюдения и использования электронной эмиссии необходимо создать у поверхности эмиттера электрическое поле, удаляющее от вещества эмиттированные частицы. В случае полевой эмиссии внешнее электрическое поле на границе твердого тела создает потенциальный барьер конечной толщины, благодаря чему оказывается возможным преодоление барьера туннельным эффектом. Однако для возникновения полевой эмиссии необходимо вблизи поверхности твердого тела создать сильное электрическое поле напряженностью порядка 10^7 В/см, в этом случае плотность тока может достигать величины порядка 10^7 А/см². В импульсных электрических полях еще большей величины отдельные участки эмиттера (выступы, заострения) сильно разогреваются и взрываются, в результате часть вещества эмиттера из конденсированной фазы непосредственно переходит в плотную плазму. Такой процесс сопровождается испусканием интенсивного электронного потока – возникает взрывная электронная эмиссия.

Важнейшей эмиссионной характеристикой твердого тела является работа выхода, равная минимальной энергии, которая необходима для перемещения электрона с поверхности Ферми (с уровня Ферми) в твердом теле в вакуум, т.е. в ту точку пространства, где напряженность электрического поля практически равна нулю. Согласно современным представлениям, в поверхностный потенциальный барьер, на преодоление которого и совершается работа выхода, основной вклад вносят обменные и корреляционные эффекты, а также, в меньшей степени, электрический двойной слой, существующий у поверхности твердого тела.

3.22.1. Термоэлектронная эмиссия из металлов

Опыт показывает, что все твердые тела при достаточно высокой температуре испускают электроны, число которых быстро возрастает с температурой. Это явление называется термоэлектронной эмиссией.

Рассмотрим более подробно физические основы термоэлектронной эмиссии на примере металлов. Известно, что в металлах при любой температуре всегда имеются свободные электроны с кинетической энергией, превышающей величину потенциального барьера на границе металл–вакуум. Чем выше температура металла, тем больше будет в металле электронов проводимости с кинетической энергией, достаточной для преодоления потенциального барьера на границе металл–

вакуум, тем интенсивней будет происходить эмиссия электронов (термоэлектронная эмиссия). В результате вблизи поверхности металла возникает электронное облако, в котором вследствие малой концентрации свободные электроны обладают свойствами частиц идеального газа. А так как этот газ описывается статистикой Ферми–Дирака, то он называется электронным Ферми-газом (ЭФГ).

В отсутствие вблизи поверхности ускоряющего электрического поля между свободными электронами металла и электронами ЭФГ существует динамическое равновесие, согласно которому число вылетающих из металла электронов равно числу проникающих в металл электронов ЭФГ. Наличие ускоряющего внешнего электрического поля приводит к некоторому увеличению потока электронов из металла. Но так как число свободных электронов в металле очень велико, то уменьшение их числа практически не сказывается на концентрации свободных электронов в металле. Если же электрическая цепь при этом замкнута, то убыль числа электронов из металла пополняется из соседних участков цепи. Поэтому полное число свободных электронов, а также их концентрация все время останутся неизменными.

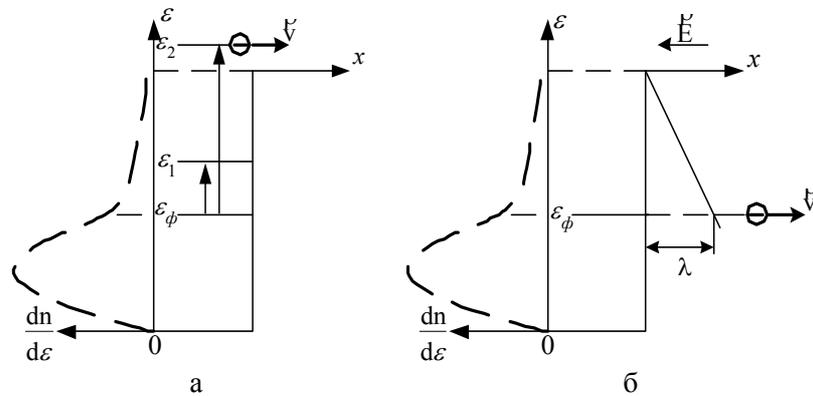


Рис.3.22.1. Потенциальная энергия электронов вблизи поверхности

На рис. 3.22.1, а схематически изображена зона проводимости, уровни полной энергии в отсутствие действия внешнего электрического поля, а также распределение квазиэлектронов по энергиям в зоне проводимости. На рис. 3.22.1, б показано изменение потенциальной энергии вблизи поверхности металла при наличии ускоряющего для

электронов напряженности электрического поля \vec{E} . В этом случае вблизи поверхности металла возникает потенциальный барьер, который для различных энергий имеет различную толщину. Для простоты дальнейших рассуждений предполагается, что потенциальная яма прямоугольная. На рис. 3.22.1, а указаны возможные переходы свободных электронов с уровня Ферми на один из свободных уровней полной энергии: в первом случае переход происходит в пределах металла, ему соответствует энергия ε_1 , во втором случае переход сопровождается вылетом электрона с поверхности металла – ему соответствует энергия ε_2 . Как следует из рис. 3.22.1, наибольшая концентрация электронов n имеет место вблизи уровня ε_f . Поэтому в наибольшем количестве поверхность металла будут покидать электроны с уровня ε_f . Следовательно, условие вылета электронов из металла должно определяться неравенством

$$\varepsilon_2 - \mu \geq A, \quad (3.22.1)$$

где A – работа выхода электрона из металла, μ – химический потенциал.

Согласно квантовой статистике, распределение квазиэлектронов по энергиям в металле и ЭФГ имеет вид

$$\frac{dn}{d\varepsilon} = \frac{4\pi(2m)^{\frac{3}{2}}}{h^3} \left(e^{\frac{\varepsilon - \mu}{kT}} + 1 \right)^{-1}, \quad (3.22.2)$$

где m – эффективная масса электрона (квазиэлектрона), она принимается здесь постоянной; h – постоянная Планка; k – постоянная Больцмана.

Практически у всех металлов температура плавления меньше 4000 К, поэтому у твердых тел температура всегда меньше 4000 К, $kT \leq 0,25$ эВ. Так как химический потенциал μ порядка единиц эВ, $\varepsilon > \mu$, то $\varepsilon - \mu > kT$.

Тогда для электронов ЭФГ металла можно записать:

$$\frac{dn}{d\varepsilon} = \frac{4\pi(2m)^{\frac{3}{2}}}{h^3} e^{\frac{\mu}{kT_e}} \varepsilon^{\frac{1}{2}} e^{-\frac{\varepsilon}{kT_e}}, \quad (3.22.3)$$

причем здесь T_e – термодинамическая температура ЭФГ.

Из условия экстремума функции (3.22.3) можно определить величину наиболее вероятной энергии $\varepsilon_{н.в.}$, т.е. той энергии, которая соответствует наибольшей концентрации электронов в объеме ЭФГ:

$$\left(\frac{1}{2\varepsilon_{н.в.}^{0.5}} - \frac{\varepsilon_{н.в.}^{0.5}}{kT_e} \right) e^{-\frac{\varepsilon_{н.в.}}{kT_e}} = 0.$$

Отсюда

$$\varepsilon_{н.в.} = \frac{kT_e}{2}.$$

С другой стороны, при выполнении неравенства $\varepsilon - \mu > kT$ функцию Ферми-Дирака можно представить в виде произведения двух множителей:

$$f_{\phi} = \left(e^{\frac{\varepsilon - \mu}{kT}} + 1 \right)^{-1} = e^{-\frac{\mu}{kT}} e^{-\frac{\varepsilon}{kT}}.$$

Из полученного выражения следует, что в принятом приближении эта функция достаточно хорошо соответствует распределению Максвелла-Больцмана.

Так как полная энергия электрона в выбранной системе отсчета

$$\varepsilon = \frac{mv^2}{2},$$

то $d\varepsilon = mv dv$. После подстановки найденных значений энергии в (3.22.3) получаем:

$$\frac{dn}{dv} = \frac{4\pi(2m)^{\frac{3}{2}}}{2^{\frac{1}{2}}h^3} m^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2kT_e}}. \quad (3.22.4)$$

По определению, функция распределения частиц по скоростям имеет вид:

$$f(v) = \frac{dn}{n dv}, \quad (3.22.5)$$

где dn – концентрация частиц идеального газа в интервале скоростей от v до $v+dv$.

Тогда в ЭФГ функция распределения электронов по скоростям будет:

$$f_{\phi e} = f_{0e} v^2 e^{-\frac{mv^2}{2kT_e}}, \quad (3.22.6)$$

где

$$f_{0e} = \frac{8\pi m^3}{n h^3} e^{-\frac{\mu}{kT_e}},$$

тогда как распределение молекул идеального газа по скоростям (распределение Максвелла) имеет вид:

$$f_M = f_{0M} v^2 e^{-\frac{mv^2}{2kT}}, \quad (3.22.7)$$

где

$$f_{0M} = 4\pi \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}}.$$

Сравнивая между собой формулы (3.22.6) и (3.22.7), мы видим, что с точностью до постоянных множителей они совпадают, следовательно, электроны объемного заряда вблизи поверхности металла обладают свойствами идеального газа, их распределение по скоростям совпадает с распределением Максвелла.

При достаточно высокой температуре вблизи поверхности металла всегда существует электронное облако отрицательного заряда. Однако если вблизи поверхности возникает ускоряющее для электронов электрическое поле, то с его возрастанием объемный заряд постепенно будет рассасываться. Моменту исчезновения объемного заряда будет соответствовать ток насыщения, величина которого слабо зависит от величины ускоряющей напряженности электрического поля и сильно зависит от температуры катода. Зависимость тока насыщения от температуры хорошо описывается формулой Ричардсона–Дэшмана:

$$I_a = S B T^2 e^{-\frac{A}{kT}}, \quad (3.22.8)$$

где

$$B = D B_0, \quad B_0 = 4\pi k^2 e m h^{-3}, \quad B_0 = 1,2 \cdot 10^{-6} \text{ A} \cdot (\text{мк})^{-2},$$

$D = 1 - r$, D – коэффициент прозрачности потенциального барьера на границе металл–вакуум; r – коэффициент отражения электронов от потенциального барьера; B_0 – эмиссионная постоянная; S – площадь поверхности того участка металла, с которого измеряется ток насыщения.

На рис. 3.22.2 приводится график зависимости анодного тока от величины ускоряющей разности потенциалов, на котором можно выделить три физически различных участка:

$0 \leq U_a \leq U_1$ – интервал значений U_a , в котором вблизи поверхности металла существует объемный заряд с постоянной концентрацией

электронов, в этом и только в этом интервале значений U_a выполняется закон $3/2$; на графике $I_a = f(U_a)$ U_1 соответствует точке перегиба;

$U_2 \geq U_a \geq U_1$ – интервал значений U_a , в котором концентрация объемного заряда постепенно уменьшается;

$U_a > U_2$ – в этом интервале значений U_a объемный заряд вообще отсутствует, этот интервал соответствует режиму насыщения.

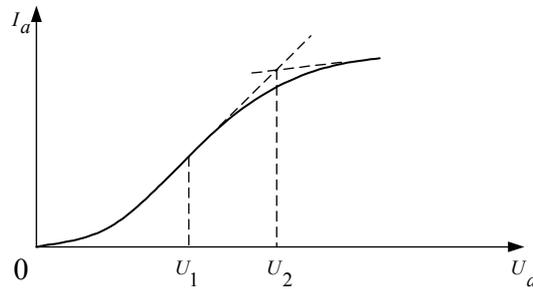


Рис. 3.22.2. Вольтамперная характеристика катода

Внешнее электрическое поле, проникая в металл, понижает высоту потенциального барьера на границе металл–вакуум, в результате на границе раздела возникает потенциальный барьер конечной толщины. Поэтому всегда существует определенная вероятность свободным электронам металла туннельным эффектом преодолеть возникший потенциальный барьер.

Из квантовой механики известно следующее выражение для коэффициента прозрачности потенциального барьера конечной толщины:

$$D = \frac{4 \left(\frac{a_2}{a_1} \right)^2}{4 \left(\frac{a_2}{a_1} \right)^2 + \left(\left(\frac{a_2}{a_1} \right)^2 - 1 \right)^2 \operatorname{sh}^2(a_2 \lambda)}, \quad (3.22.9)$$

где

$$a_2^2 = \frac{2m}{\eta^2} (U - \varepsilon), \quad a_1^2 = \frac{2m}{\eta^2} \varepsilon, \quad \left(\frac{a_2}{a_1} \right)^2 = \frac{U}{\varepsilon} - 1,$$

m – масса электрона, U – потенциальная энергия электрона в металле, в первом приближении ее график совпадает с дном потенциальной ямы, ε – полная энергия электрона на границе металл–вакуум.

Пусть ℓ – толщина потенциального барьера на уровне Ферми. Тогда, если отсчет энергии вести от дна потенциальной ямы, то $\varepsilon_f = \mu$, μ – химический потенциал. Тогда $U = A + \mu$, A – работа выхода электрона из металла.

Из квантовой статистики следует, что химический потенциал ЭФГ с большой точностью можно вычислить по формуле:

$$\mu = \frac{\eta^2}{8m} \left(\frac{3n}{\pi} \right)^{\frac{2}{3}}, \quad (3.22.10)$$

n – концентрация электронов в ЭФГ.

Таким образом, используя экспериментальные значения тока насыщения при различных температурах T и U_a , можно вычислить величину коэффициента прозрачности потенциального барьера для различных значений полной энергии, а также толщину потенциального барьера, что другими методами определить достаточно сложно.

3.22.2. Полевая электронная эмиссия

Полевая (туннельная, автоэлектронная) эмиссия возникает в результате испускания твердыми телами электронов под действием сильного электрического поля у поверхности. Если внешнее электрическое поле достаточно велико для того, чтобы потенциальный порог на границе тела превратился в потенциальный барьер малой толщины, то становится возможным квантовомеханическое туннелирование электронов и выход их в вакуум. При этом электроны непосредственно после прохождения сквозь барьер имеют ту же энергию, что и внутри тела.

Токи полевой эмиссии с большой плотностью удается получать с эмиттеров, имеющих форму острия. Предельная плотность тока, еще не разрушающего острия, возрастает с увеличением угла при вершине эмитирующего конуса, так как с увеличением этого угла улучшается отвод теплоты от острия. В очень сильных электрических полях, когда плотность тока полевой эмиссии достигает величины (10^{12} – 10^{13}) А/м², локальные участки на поверхности острия, из которых происходит эмиссия, из-за сильного разогрева взрываются, образуя плотную плазму, расширяющуюся со скоростью порядка 10^4 м/с. Этот процесс сопровождается возникновением интенсивной электронной эмиссии и называется взрывной электронной эмиссией.

Взрывная эмиссия впервые была обнаружена при исследовании вакуумного пробоя в наносекундном диапазоне длительностей импульсов высокого напряжения, а также в экспериментах по автоэлектронной

эмиссии при предельно высоких плотностях тока в микросекундном диапазоне. Открытие взрывной эмиссии было сделано в 1971 г. Авторами открытия являются Г.Н. Фурсей (г. Ленинград) и Г.А. Месяц (г. Новосибирск). Ими было установлено, что под действием сильного электрического поля собственный автоэмиссионный ток с автоэлектронного эмиттера разрушает микроучасток поверхности острия катода. Этот процесс сопровождается резким возрастанием тока.

Ток взрывной электронной эмиссии с одиночного острия

$$I = 3,7 \cdot 10^{-7} U^{\frac{3}{2}} vt (d - vt)^{-1}, \quad (3.22.11)$$

где U – разность потенциалов между катодом и анодом в процессе взрывной эмиссии; d – расстояние между катодом и анодом; t – время от момента приложения импульса напряжения до момента взрыва, $t < d/v$, v – скорость расширяющейся плазмы; $[I] = \text{А}$, $[U] = \text{В}$, $[d] = \text{м}$, $[t] = \text{с}$.

В результате взаимодействия автоэлектронов со слоем плотной плазмы вещества эмиттера вблизи поверхности острия энергетический спектр электронов при взрывной эмиссии шире, чем при автоэлектронной эмиссии. Источником взрывной эмиссии могут быть не только металлические, но и полупроводниковые острия, а также жидкие металлические катоды.

Так как для деионизации плазмы и образования новых микровыступов требуется некоторое время, то катоды со взрывной эмиссией могут работать только в режиме однократных включений или малой частоты повторения.

В настоящее время полевая эмиссия используется в электронной и ионной микроскопии (автоэлектронная, автоионная эмиссии). Взрывная электронная эмиссия используется в сильноточных ускорителях электронов, в импульсных источниках рентгеновского излучения высокой интенсивности, в мощных газовых лазерах, в установках для управляемого термоядерного синтеза. При этом для получения больших токов применяются многоострийные катоды.

3.22.3. Фотозлектронная эмиссия

Фотозлектронная эмиссия – это эмиссия электронов с поверхности твердого или жидкого тела, возникающая в результате облучения поверхности электромагнитным излучением.

Фотозлектронная эмиссия впервые была обнаружена Г. Герцем в 1887 г. при изучении воздействия ультрафиолетового излучения на электрический разряд. Однако Герц не занимался систематическим исследованием обнаруженного им явления. Основные первоначаль-

ные результаты по изучению фотоэлектронной эмиссии были получены прежде всего в работах А.Г. Столетова, а затем Эльстера и Гейттеля, Ленарда и А. Эйнштейна.

А.Г. Столетовым в 1888 г. было установлено, что фототок пропорционален первой степени интенсивности падающего света. Однако в дальнейшем оказалось, что эта зависимость характерна только для слабых световых потоков. С появлением лазеров было экспериментально показано, что фототок пропорционален n -ой степени интенсивности света.

В приближении однофотонных процессов основные законы фотоэлектронной эмиссии – следующие.

1. Фототок в режиме насыщения прямо пропорционален потоку падающей электромагнитной волны.

2. Максимальная кинетическая энергия фотоэлектронов линейно возрастает с частотой падающего излучения ν и не зависит от величины падающего на поверхность потока электромагнитной волны:

$$\frac{m_e v_{\max}^2}{2} = h \nu - A, \quad (3.22.12)$$

где m_e – масса покоя электрона, A – работа выхода электрона.

Формула (3.22.12) впервые была получена А. Эйнштейном в 1905 г.

При очень больших плотностях потока электромагнитной волны ($\Phi > 10^4$ Вт/м²) существенными становятся многофотонные процессы, когда электрон, прежде чем покинуть поверхность освещаемого твердого тела, в процессе движения может столкнуться не с одним, а последовательно с несколькими фотонами. В этом случае вместо уравнения (3.22.12) следует записать уравнение

$$\frac{m_e v_{\max}^2}{2} = N h \nu - A, \quad (3.22.13)$$

где N – число столкновений электрона с фотонами, сопровождающимися поглощением фотонов. Но и в этом случае существует длинноволновая граница фотоэффекта, однако граничная частота теперь определяется выражением

$$\nu_0 = \frac{A}{N h}, \quad (3.22.14)$$

т. е. уменьшается в N раз.

3. Для каждого вещества существует длинноволновая граница падающего излучения λ_0 , выше которой фотоэффект не наблюдается, и, следовательно, существует соответствующая ей пороговая частота ν_0 .

4. Фотоэлектронная эмиссия практически безинерционна.

Количественной характеристикой фотоэлектронной эмиссии является квантовый выход, т.е. число вылетевших электронов, приходящихся на один фотон, падающий на поверхность тела. Величина квантового выхода зависит от состояния поверхности и энергии фотонов и оказывается разной у различных тел.

Фотоэлектронная эмиссия из металлов. В случае однофотонных процессов фотоэлектронная эмиссия из металлов возникает, если энергия фотона превышает работу выхода металла. Вблизи порога фотоэлектронной эмиссии для большинства металлов квантовый выход имеет величину порядка 10^{-4} электрон/фотон. Столь малая величина квантового выхода объясняется тем, что падающий на поверхность металла поток фотонов поглощается в нем вблизи поверхности на глубине порядка 10^{-7} м.

Возникшие фотоэлектроны сильно взаимодействуют с электронами проводимости, быстро теряя полученную от фотонов энергию. В результате энергию, достаточную для преодоления работы выхода, сохраняют только те фотоэлектроны, которые возникли вблизи поверхности металлов на глубине, не превышающей 10^{-9} м.

При $\nu > \nu_0$ квантовый выход для всех металлов возрастает с ростом частоты колебаний в падающей электромагнитной волне. Вблизи красной границы величина фототока, как следует из опыта,

$$I_{\phi} = I_0 \cdot (\nu - \nu_0)^2. \quad (3.22.15)$$

При анализе взаимодействия потока электромагнитных волн с металлами можно выделить три области спектра частот.

1. $\nu < \tau^{-1}$, τ – время релаксации возбужденных состояний электронов в зоне проводимости, $\tau \approx 10^{-13}$ с. В этой области металл сильно отражает падающее электромагнитное излучение, коэффициент поглощения практически постоянный и не зависит от частоты ν . Область соответствует дальнему инфракрасному и СВЧ - диапазонам частот.

2. $\tau^{-1} < \nu < \nu_{pv}$,

$$\nu_{pv} = \left(\frac{e n_e}{2 \pi \epsilon_0 m_e} \right)^{\frac{1}{2}}, \quad (3.22.16)$$

ν_{pv} – плазменная частота колебаний электронов в ЭФЖ; n_e – концентрация электронов проводимости, ϵ_0 – электрическая постоянная.

В области 2 коэффициент отражения имеет значительную величину, а коэффициент поглощения убывает с ростом частоты пропорционально ν^2 . Неравенство $\nu > \nu_{pv}$ (ν_{pv} – частота объемного плазмона) соответствует области ультрафиолетовой прозрачности металлов. Крас-

ная граница однофотонной фотоэлектронной эмиссии обычно находится в области, которая охватывает оптический интервал частот.

3. $\nu \geq \nu_{ps}$ (ν_{ps} – частота поверхностного плазмона) соответствует резкому росту фототока. Величина $\nu = \nu_{ps}$ соответствует второй красной границе фотоэлектронной эмиссии.

Первая теория фотоэлектронной эмиссии была разработана Р. Фаулером в 1931 г. В основу теории были положены следующие допущения.

1. Электроны в металле являются свободными, т.е. находятся в потенциальной яме с плоским дном.

2. В фотоэлектронной эмиссии принимают участие электроны с энергией близкой к энергии Ферми.

3. Коэффициент отражения на границе раздела для электронов с полной энергией больше потенциальной равен нулю и равен единице для всех электронов, находящихся в потенциальной яме.

4. Неупругие столкновения электронов проводимости приводят к такой потере энергии, при которой электрон не в состоянии вылететь за пределы металла и остается в пределах потенциальной ямы.

При использовании таких предположений ЭФГ в металле можно рассматривать как смесь двух газов – "холодного" (невозбужденного) и "горячего" (возбужденного). Энергетический спектр у "горячих" электронов предполагается таким же, как и у "холодных", со смещением вверх по шкале энергий на величину энергии фотона $h\nu$, что эквивалентно снижению на ту же величину высоты потенциального барьера.

Выполненный расчет показал, что при $T = 0$ К фототок фотоэлектронной эмиссии

$$I_{\phi} = \frac{2\pi e m_e \alpha S}{h} (\nu - \nu_0)^2, \quad (3.22.17)$$

$\nu > \nu_0$; α – относительное число горячих электронов по отношению к холодным (постоянная определяется из опыта); S – площадь освещаемой поверхности металла. При $T > 0$ К и $\nu < \nu_0$ фототок асимптотически стремится к нулю, и, следовательно, красная граница должна отсутствовать. Однако в случае $\nu_0 \leq \nu \leq 1,5 \nu_0$ уравнение (3.22.17) все-таки выполняется.

Из сравнения формул (3.22.16) и (3.22.15) следует, что теория Фаулера хорошо соответствует опыту лишь вблизи "красной" границы спектра. Следовательно, теория Фаулера имеет ограниченную применимость и недостаточное познавательное значение.

В 1964 г. В. Спайсером была предложена теория, согласно которой процесс выхода фотоэлектронов из металлов следует разделить на три последовательных процесса:

1. поглощение фотона и появление в твердом (или жидком) теле фотоэлектрона с энергией более высокой по сравнению со средней энергией электронов проводимости;
2. движения фотоэлектрона к поверхности, сопровождающегося рассеянием энергии;
3. выход фотоэлектрона в другую среду через поверхность раздела.

Полученное в этой теории выражение для фототока имеет весьма сложный вид и содержит несколько подгоночных параметров, значения которых можно определить только из эксперимента. Наличие определяемых из опыта параметров делают теорию неудобной как для практического использования, так и для анализа физических процессов в фотоэлектронной эмиссии.

Таким образом, строгая теория фотоэлектронной эмиссии до сих пор еще не создана. Но очевидно, что эта теория должна основываться на квантовомеханическом рассмотрении возбуждения фотонами электронов металла с последующим использованием уравнения кинетической теории переноса, в которое следует ввести микроскопические параметры взаимодействия как фотонов с электронами проводимости, так и фотоэлектронов с электронами проводимости металла.

Фотоэлектронная эмиссия из полупроводников. В полупроводниках фотоэлектронная эмиссия может быть вызвана возбуждением электронов из валентной зоны, с уровней примесей, из зоны проводимости и др. Для каждого из этих случаев пороговая частота ν имеет свое значение.

Если полупроводник собственный, то минимальная энергия квантов, достаточная для перехода электронов из валентной в зону проводимости, равна ширине запрещенной зоны. Отсюда следует, что эффективность фотоэлектронной эмиссии при заданной ширине запрещенной зоны $\Delta\varepsilon_0$ определяется электронным сродством χ . Тогда красная граница фотоэлектронной эмиссии будет определяться из уравнения

$$h\nu_0 = \Delta\varepsilon_0 + \chi. \quad (3.22.18)$$

Следовательно, в вакуум могут перейти лишь те электроны, которые достигают поверхности эмиттера с энергией, достаточной для преодоления потенциального барьера высотой χ . Создание на поверхности отрицательного электронного сродства позволяет существенно повысить квантовый выход фотоэлектронной эмиссии. Для соз-

дания отрицательного электронного средства используются полупроводники p -типа, у которых при сильном легировании уровень Ферми почти совпадает с верхней границей валентной зоны.