

Глава 31

Элементы физики твердого тела

§ 240. Понятие о зонной теории твердых тел

Используя уравнение Шредингера — основное уравнение динамики в нерелятивистской квантовой механике,— в принципе можно рассмотреть задачу о кристалле, например найти возможные значения его энергии, а также соответствующие энергетические состояния. Однако как в классической, так и в квантовой механике отсутствуют методы точного решения динамической задачи для системы многих частиц. Поэтому эта задача решается приближенно сведением задачи многих частиц к одноэлектронной задаче — задаче об одном электроне, движущемся в заданном внешнем поле. Подобный путь приводит к **зонной теории твердого тела**.

В основе зонной теории лежит так называемое **адиабатическое приближение**. Квантово-механическая система разделяется на тяжелые и легкие частицы — ядра и электроны. Поскольку массы и скорости этих частиц значительно различаются, можно считать, что движение электронов происходит в поле неподвижных ядер, а медленно движущиеся ядра находятся в усредненном поле всех электронов. Считая, что ядра в узлах кристаллической решетки неподвижны, движение электрона рассматривается в *постоянном периодическом поле ядер*.

Далее используется приближение **самосогласованного поля**. Взаимодействие данного электрона со всеми другими электронами заменяется действием на него стационарного электрического поля, обладающего периодичностью кристаллической решетки. Это поле создается усредненным в пространстве зарядом всех других электронов и всех ядер. Таким образом, в рамках зонной теории многоэлектронная задача сводится к задаче о движении одного электрона во внешнем периодическом поле — усредненном и согласованном поле всех ядер и электронов.

Рассмотрим мысленно процесс образования твердого тела из изолированных атомов. Пока атомы изолированы, т. е. находятся друг от друга на макроскопических расстояниях, они имеют совпадающие схемы энергетических уровней (рис. 313). По мере «сжатия» нашей модели до кристаллической решетки, т. е. когда расстояния между атомами станут равными межатомным расстояниям в твердых телах, взаимодействие между атомами приводит к тому, что энергетические уровни атомов *смещаются, расщепляются и расширяются в зоны*, образуется так называемый **зонный энергетический спектр**.

Из рис. 313; на котором показано расщепление уровней как функция расстояния r между атомами, видно, что заметно расщепляются и расширяются лишь уровни внешних, валентных электронов,

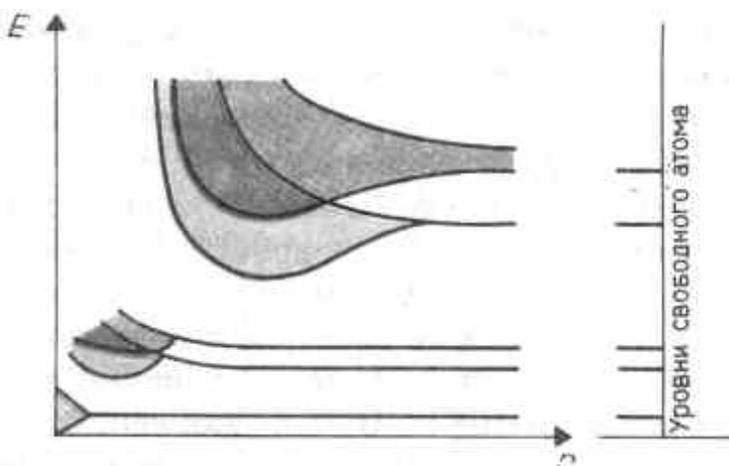


Рис. 313

лее слабо связанных с ядром и имеющих наибольшую энергию, а также более высокие уровни, которые в основном состоянии атома вообще электронами не заняты. Уровни же внутренних электронов либо совсем не расщепляются, либо расщепляются слабо. Таким образом, в твердых телах внутренние электроны ведут себя так же, как в изолированных атомах, валентные же электроны «коллективизированы» — принадлежат всему твердому телу.

Образование зонного энергетического спектра в кристалле является квантово-механическим эффектом и вытекает из соотношения неопределенностей. В кристалле валентные электроны атомов, связанные слабее с ядрами, чем внутренние электроны, могут переходить от атома к атому сквозь потенциальные барьеры, разделяющие атомы, т. е. перемещаться без изменений полной энергии (туннельный эффект, см. § 221). Это приводит к тому, что среднее время жизни τ валентного электрона в данном атоме по сравнению с изолированным атомом существенно уменьшается и составляет примерно 10^{-15} с (для изолированного атома оно примерно 10^{-8} с). Время же жизни электрона в каком-либо состоянии связано с неопределенностью его энергии (шириной уровня) соотношением неопределенностей $\Delta E \sim \hbar/\tau$ (см. (215.5)). Следовательно, если естественная ширина спектральных линий составляет примерно 10^{-7} эВ, то в кристаллах $\Delta E \approx 1 - 10$ эВ, т. е. энергетические уровни валентных электронов расширяются в зону дозволённых значений энергии.

Энергия внешних электронов может принимать значения в пределах закрашенных на рис. 313 областей, называемых **разрешенными энергетическими зонами**. Каждая разрешенная зона «вмещает» в себя столько близлежащих дискретных уровней, сколько атомов содержит кристалл: чем больше в кристалле атомов, тем теснее расположены уровни в зоне. Расстояние между соседними энергетическими уровнями в зоне составляет приблизительно 10^{-22} эВ. Так как оно столь ничтожно, то зоны можно считать практически непрерывными, однако факт конечного числа уровней в зоне играет важную роль для распределения электронов по состояниям.

Разрешенные энергетические зоны разделены зонами запрещённых значений энергии, называемыми **запрещенными энергетическими зонами**. В запрещённых зонах электроны находиться не могут. Ширина зон (разрешенных и запрещённых) не зависит от размера кристалла. Разрешенные зоны тем шире, чем слабее связь валентных электронов с ядрами.

§241. Металлы, диэлектрики и полупроводники по зонной теории

Зонная теория твердых тел позволила с единой точки зрения истолковать существование металлов, диэлектриков и полупроводников, объясняя различие в их электрических свойствах, во-первых, неодинаковым заполнением электронами разрешенных зон и, во-вторых, шириной запрещённых зон.

Степень заполнения электронами энергетических уровней в зоне определяется заполнением соответствующего атомного уровня. Если, например, какой-то уровень атома полностью заполнен электронами в соответствии с принципом Паули, то образующаяся из него зона также полностью заполнена. В общем случае можно говорить о **валентной зоне**, которая полностью заполнена электронами и образована из энергетических уровней внутренних электронов свободных атомов, и о **зоне проводимости (свободной зоне)**, которая либо частично заполнена электронами, либо свободна и образована из энергетических уровней внешних «коллективизированных» электронов изолированных атомов.

В зависимости от степени заполнения зон электронами и ширины запрещённой зоны возможны четыре случая, изображенные на рис. 314. На рис. 314, а самая верхняя зона, содержащая электроны, заполнена лишь частично, т. е. в ней имеются вакантные уровни. В данном случае электрон, получив сколь угодно малую энергетическую «добавку» (например, за счет теплового движения или электриче-

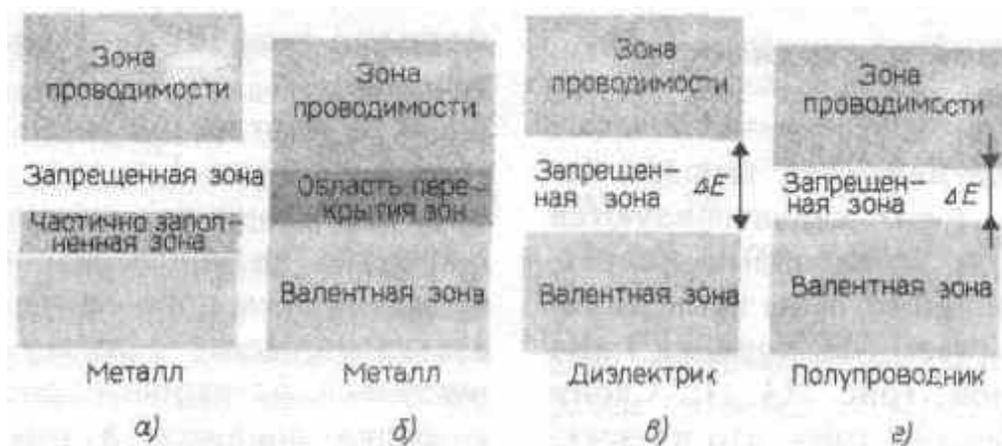


Рис. 314

ского поля), сможет перейти на более высокий энергетический уровень той же зоны, т. е. стать свободным и участвовать в проводимости. Внутризонный переход вполне возможен, так как, например, при 1 К энергия теплового движения $kT \approx 10^{-4}$ эВ, т. е. гораздо больше разности энергий между соседними уровнями зоны (примерно 10^{-22} эВ). Таким образом, если в твердом теле имеется зона, лишь частично заполненная электронами, то это тело всегда будет проводником электрического тока. Именно это свойственно металлам.

Твердое тело является проводником электрического тока и в том случае, когда валентная зона перекрывается свободной зоной, что в конечном счете приводит к не полностью заполненной зоне (рис. 314, б). Это имеет место для щелочноземельных элементов, образующих II группу таблицы Менделеева (Be, Mg, Ca, Zn, ...). В данном случае образуется так называемая «гибридная» зона, которая заполняется валентными электронами лишь частично. Следовательно, в данном случае металлические свойства щелочноземельных элементов обусловлены перекрытием валентной и свободной зон.

Помимо рассмотренного выше перекрытия зон возможно также перераспределение электронов между зонами, возникающими из уровней различных атомов, которое может привести к тому, что вместо двух частично заполненных зон в кристалле окажутся одна целиком заполненная (валентная) зона и одна свободная зона (зона проводимости). Твердые тела, у которых энергетический спектр электронных состояний состоит только из валентной зоны и зоны проводимости, являются диэлектриками или полупроводниками в зависимости от ширины запрещенной зоны ΔE .

Если ширина запрещенной зоны кристалла порядка нескольких электрон-вольт, то тепловое движение не может перебросить электроны из валентной зоны в зону проводимости и кристалл является диэлектриком, оставаясь им при всех реальных температурах (рис. 314, в). Если запрещенная зона достаточно узка (ΔE порядка 1 эВ), то переброс электронов из валентной зоны в зону проводимости может быть осуществлен сравнительно легко либо путем теплового возбуждения, либо за счет внешнего источника, способного передать электронам энергию ΔE , и кристалл является полупроводником (рис. 314, г).

Различие между металлами и диэлектриками с точки зрения зонной теории состоит в том, что при 0 К в зоне проводимости металлов имеются электроны, а в зоне проводимости диэлектриков они отсутствуют. Различие же между диэлектриками и полупроводниками определяется шириной запрещенных зон: для диэлектриков она довольно широка (например, для NaCl $\Delta E = 6$ эВ), для полупроводников — достаточно узка (например, для германия $\Delta E = 0,72$ эВ). При температурах, близких к 0 К, полупроводники ведут себя как диэлектрики, так как переброса электронов в зону проводимости не происходит. С повышением температуры у полупроводников растет число электронов, которые вследствие теплового возбуждения переходят в зону проводимости, т. е. электрическая проводимость проводников в этом случае увеличивается.

§ 242. Собственная проводимость полупроводников

Полупроводниками являются твердые тела, которые при $T=0$ характеризуются полностью занятой электронами валентной зоной, отделенной от зоны проводимости сравнительно узкой (ΔE порядка 1 эВ) запрещенной зоной (рис. 314, г). Своим названием они обязаны тому, что их электропроводность меньше электропроводности металлов и больше электропроводности диэлектриков.

В природе полупроводники существуют в виде элементов (элементы IV, V и VI групп Периодической системы элементов Менделеева), например Si, Ge, As, Se, Te, и химических соединений, например оксиды, сульфиды, селениды, сплавы элементов различных групп. Различают **собственные и примесные полупроводники**. **Собственными полупроводниками** являются химически чистые полупроводники, а их проводимость называется **собственной проводимостью**. Примером собственных полупроводников могут служить химически чистые Ge, Se, а также многие химические соединения: InSb, GaAs, CdS и др.

При 0 К и отсутствии других внешних факторов собственные полупроводники ведут себя как диэлектрики. При повышении же температуры электроны с верхних уровней валентной зоны I могут быть переброшены на нижние уровни зоны проводимости II (рис.315). При наложении на кристалл электрического поля они перемещаются против поля и создают электрический ток. Таким образом, зона II из-за ее частичного «укомплектования» электронами становится зоной проводимости. Проводимость собственных полупроводников, обусловленная электронами, называется **электронной**

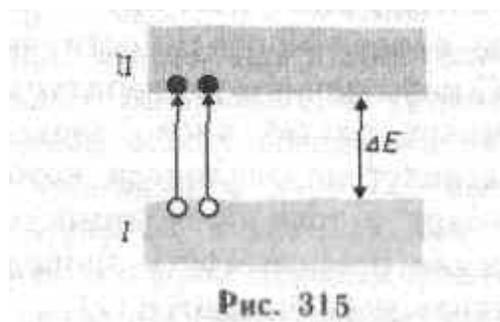


Рис. 315

В результате тепловых забросов электронов из зоны I в зону II в валентной зоне возникают вакантные состояния, получившие название **дырок**. Во внешнем электрическом поле на освободившееся от электрона место — дырку — может переместиться электрон с соседнего уровня, а дырка появится в том месте, откуда ушел электрон, и т. д. Такой процесс заполнения дырок электронами равносителен перемещению дырки в направлении, противоположном движению электрона, так, как если бы дырка обладала положительным зарядом, равным по величине заряду электрона. Проводимость собственных полупроводников, обусловленная квазичастицами — дырками, называется **дырочной проводимостью** или **проводимостью p-типа** (от лат. positive — положительный).

Таким образом, в собственных полупроводниках наблюдаются два механизма проводимости: электронный и дырочный. Число электронов в зоне проводимости равно числу дырок в валентной зоне, так как последние соответствуют электронам, возбужденным в зону проводимости. Следовательно, если концентрации электронов проводимости и дырок обозначить соответственно n_e и n_p , то

$$n_e = n_p. \quad (242.1)$$

Проводимость полупроводников всегда является **возбужденной**, т. е. появляется только под действием внешних факторов (температуры, облучения, сильных электрических полей и т.д.).

В собственном полупроводнике уровень Ферми находится в середине запрещенной зоны (рис.316). Действительно, для переброса электрона с верхнего уров-

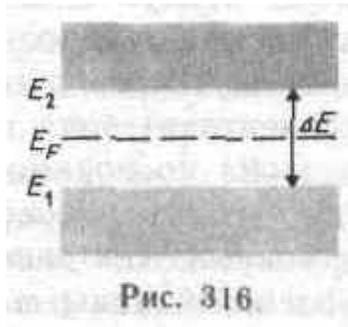


Рис. 316

ня валентной зоны на нижний уровень зоны проводимости затрачивается энергия активации, равная ширине запрещенной зоны ΔE . При появлении же электрона в зоне проводимости в валентной зоне обязательно возникает дырка. Следовательно, энергия, затраченная на образование пары носителей тока, должна делиться на две равные части. Так как энергия, соответствующая половине ширины запрещенной зоны, идет на переброс электрона и такая же энергия затрачивается на образование дырки, то начало отсчета для каждого из этих процессов должно находиться в середине запрещенной зоны. Энергия Ферми в собственном полупроводнике представляет собой энергию, от которой происходит возбуждение электронов и дырок.

Вывод о расположении уровня Ферми в середине запрещенной зоны собственного полупроводника может быть подтвержден математическими выкладками. В физике твердого тела доказывается, что

$$n_e = C_1 e^{-(E_2 - E_F)/(kT)}, \quad (242.2)$$

где E_2 — энергия, соответствующая дну зоны проводимости (рис.316), E_F — энергия Ферми, T — термодинамическая температура, C_1 — постоянная, зависящая от температуры и эффективной массы электрона проводимости. **Эффективная масса** — величина, имеющая размерность массы и характеризующая динамические свойства квазичастиц — электронов проводимости и дырок. Введение в зонную теорию эффективной массы электрона проводимости позволяет, с одной стороны, учитывать действие на электроны проводимости не только внешнего поля, но и внутреннего периодического поля кристалла, а с другой стороны, абстрагируясь от взаимодействия электронов проводимости с решеткой, рассматривать их движение во внешнем поле как движение свободных

$$n_p = C_2 e^{(E_1 - E_F)/(kT)}, \quad (242.3)$$

где C_2 — постоянная, зависящая от температуры и эффективной массы дырки, E_1 — энергия, соответствующая верхней границе валентной зоны. Энергия возбуждения в данном случае отсчитывается вниз от уровня Ферми (рис. 316),

поэтому величины в экспоненциальном множителе (242.3) имеют знак, обратный знаку

$$C_1 e^{-(E_2 - E_F)/(kT)} = C_2 e^{(E_1 - E_F)/(kT)}$$

Если эффективные массы электронов и дырок равны ($m_e^* = m_p^*$), то $C_1 = C_2$ и, следовательно,

$$-(E_2 - E_F) = E_1 - E_F,$$

откуда

$$E_F = (E_1 + E_2)/2 = \Delta E/2,$$

т. е. уровень Ферми в собственном полупроводнике действительно расположен в середине запрещенной зоны.

Так как для собственных полупроводников $\Delta E \gg kT$, то распределение Ферми — Дирака (235.2) переходит в распределение Максвелла — Больцмана. Положив в (236.2) $E - E_F \approx \Delta E/2$, получим

$$\langle N(E) \rangle \approx e^{-\Delta E/(2kT)} \quad (242.4)$$

Количество электронов, перебросившихся в зону проводимости, а следовательно, и количество образовавшихся дырок пропорциональны $\langle N(E) \rangle$. Таким образом, удельная проводимость

$$\gamma = \gamma_0 e^{-\Delta E/(2kT)}, \quad (242.5)$$

где γ_0 — постоянная, характерная для данного полупроводника.

Увеличение проводимости полупроводников с повышением температуры является их характерной особенностью (у металлов с повышением температуры проводимость уменьшается). С точки зрения зонной теории это обстоятельство объяснить довольно просто: с повышением температуры растет число электронов, которые вследствие теплового возбуждения переходят в зону проводимости и участвуют в проводимости. Поэтому удельная проводимость собственных полупроводников с повышением температуры растет.

Если представить зависимость $\ln \gamma$ от $1/T$, то для собственных полупроводников — это прямая (рис.317), по на-

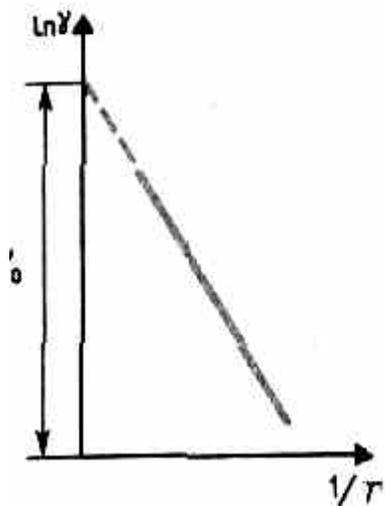


Рис. 317

клону которой можно определить ширину запрещенной зоны ΔE , а по ее продолжению — γ_0 (прямая отсекает на оси ординат отрезок, равный $\ln \gamma_0$).

Одним из наиболее широко распространенных полупроводниковых элементов является германий, имеющий решетку типа алмаза, в которой каждый атом связан ковалентными связями (см. §71) с четырьмя ближайшими соседями. Упрощенная плоская схема расположения атомов в кристалле Ge дана на рис. 318, где каждая черточка обозначает связь, осуществляемую одним электроном. В идеальном кристалле при 0 К такая структура представляет собой диэлектрик, так как все валентные электроны участвуют в образовании связей и, следовательно, не участвуют в проводимости.

При повышении температуры (или под действием других внешних факторов) тепловые колебания решетки могут привести к разрыву некоторых валентных связей, в результате чего часть электронов отщепляется и они становятся свободными. В покинутом электроном месте возникает дырка (она изображена белым кружком), заполнить которую могут электроны из соседней пары. В результате дырка, так

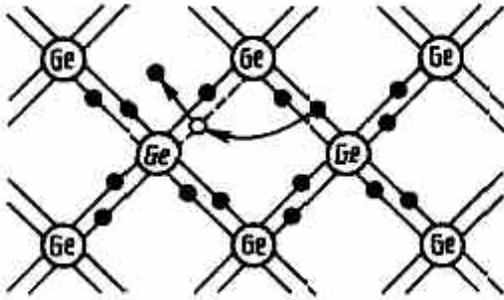


Рис. 318

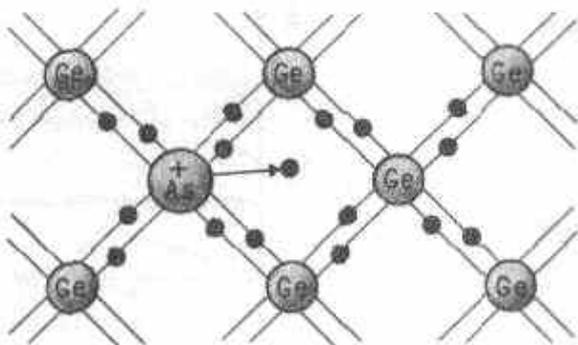
же как и освободившийся электрон, будет двигаться по кристаллу. Движение электронов проводимости и дырок в отсутствие электрического поля является хаотическим. Если же на кристалл наложить электрическое поле, то электроны начнут двигаться против поля, дырки — по полю, что приведет к возникновению собственной проводимости германия, обусловленной как электронами, так и дырками.

В полупроводниках наряду с процессом генерации электронов и дырок идет процесс **рекомбинации**: электроны переходят из зоны проводимости в валентную зону, отдавая энергию решетке и испуская кванты электромагнитного излучения. В результате для каждой температуры устанавливается определенная равновесная концентрация электронов и дырок, изменяющаяся с температурой пропорционально выражению (242.4).

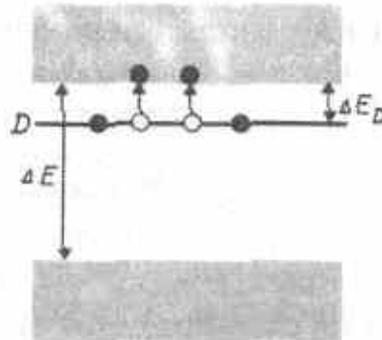
§ 243. Примесная проводимость полупроводников

Проводимость полупроводников, обусловленная примесями, называется **примесной проводимостью**, а сами полупроводники — **примесными полупроводниками**. Примесная проводимость обусловлена примесями (атомы посторонних элементов), а также дефектами типа избыточных атомов (по сравнению со стехиометрическим составом), тепловыми (пустые узлы или атомы в междоузлиях) и механическими (трещины, дислокации и т. д.) дефектами. Наличие в полупроводнике примеси существенно изменяет его проводимость. Например, при введении в кремний примерно 0,001 ат. % бора его проводимость увеличивается примерно в 10^6 раз.

Примесную проводимость полупроводников рассмотрим на примере Ge и Si, в которые вводятся атомы с валентностью, отличной от валентности основных атомов на единицу. Например, при замещении атома германия пятивалентным атомом мышьяка (рис. 319, а) один электрон не может образовать ковалентной связи, он оказывается лишним и может быть легко при тепловых колебаниях решетки отщеплен от атома, т. е. стать свободным. Обра-



а)



б)

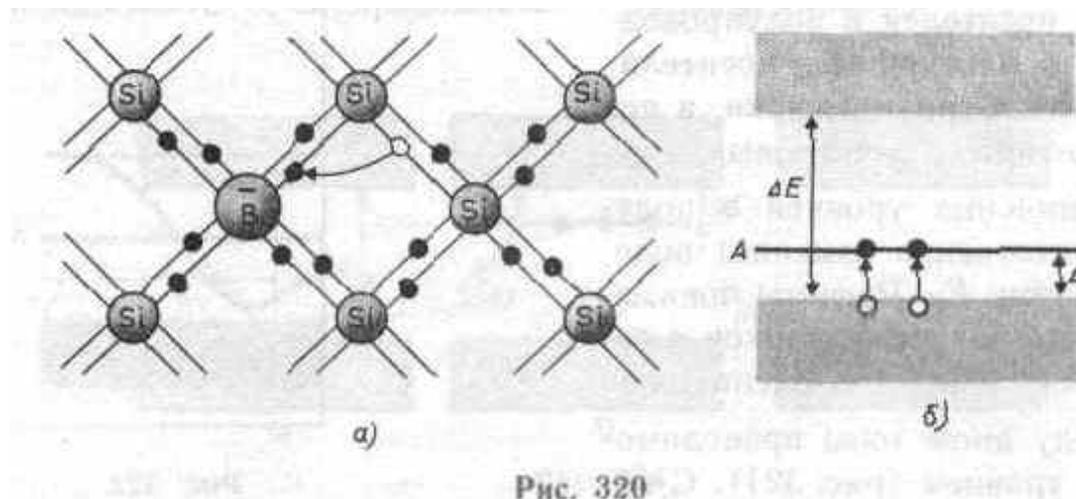
Рис. 319

зование свободного электрона не сопровождается нарушением ковалентной связи; следовательно, в отличие от случая, рассмотренного в § 242, дырка не возникает. Избыточный положительный заряд, возникающий вблизи атома примеси, связан с атомом примеси и поэтому перемещаться по решетке не может.

С точки зрения зонной теории рассмотренный процесс можно представить следующим образом (рис. 319, б). Введение примеси искажает поле решетки, что приводит к возникновению в запрещенной зоне энергетического уровня D валентных электронов мышьяка, называемого **примесным уровнем**. В случае германия с примесью мышьяка этот уровень располагается от дна зоны проводимости на расстоянии $\Delta E_0 = 0,013$ эВ. Так как $\Delta E_D < kT$, то уже при обычных температурах энергия теплового движения достаточна для того, чтобы перебросить электроны примесного уровня в зону проводимости; образующиеся при этом положительные заряды локализуются на неподвижных атомах мышьяка и в проводимости не участвуют.

Таким образом, в полупроводниках с примесью, валентность которой *на единицу больше валентности основных атомов*, носителями тока являются электроны; возникает **электронная примесная проводимость (проводимость n-типа)**. Полупроводники с такой проводимостью называются **электронными (или полупроводниками n-типа)**. Примеси, являющиеся источником электронов, называются **донорами**, а энергетические уровни этих примесей — **донорными уровнями**.

Предположим, что в решетку кремния введен примесный атом с тремя валентными электронами, например бор (рис. 320, а). Для образования связей с четырьмя ближайшими соседями у атома бора не хватает одного электрона, одна из связей остается неупакованной и четвертый электрон может быть захвачен от соседнего атома основного вещества, где соответственно образуется дырка. Последовательное заполнение образующихся дырок электронами эквивалентно движению дырок в полупроводнике, т. е. дырки не остаются локализованными, а перемещаются в решетке кремния как свободные положительные заряды. Избыточный же отрицательный заряд, возникающий вблизи



мом примеси и по решетке перемещаться не может.

По зонной теории, введение трехвалентной примеси в решетку кремния приводит к возникновению в запрещенной зоне примесного энергетического уровня A , не занятого электронами. В случае кремния с примесью бора этот уровень располагается выше верхнего края валентной зоны на расстоянии $\Delta E_A = 0,08$ эВ (рис. 320, б). Близость этих уровней к валентной зоне приводит к тому, что уже при сравнительно низких температурах электроны из валентной зоны переходят на примесные уровни и, связываясь с атомами бора, теряют способность перемещаться по решетке кремния, т. е. в проводимости не участвуют. Носителями тока являются лишь дырки, возникающие в валентной зоне.

Таким образом, в полупроводниках с примесью, валентность которой на единицу меньше валентности основных атомов, носителями тока являются дырки; возникает **дырочная проводимость (проводимость р-типа)**. Полупроводники с такой проводимостью называются **дырочными (или полупроводниками р-типа)**. Примеси, захватывающие электроны из валентной зоны полупроводника, называются **акцепторами**, а энергетические уровни этих примесей — **акцепторными уровнями**.

В отличие от собственной проводимости, осуществляющейся одновременно электронами и дырками, примесная проводимость полупроводников обусловлена в основном носителями одного знака: электронами—в случае донорной примеси, дырками — в случае акцепторной. Эти носители тока называются **основными**. Кроме основных носителей в полупроводнике имеются и неосновные носители: в полупроводниках n-типа — дырки, в полупроводниках р-типа — электроны.

Наличие примесных уровней в полупроводниках существенно изменяет положение уровня Ферми E_F . Расчеты показывают, что в случае полупроводников n-типа уровень Ферми E_{F0} при 0 К расположен посередине между дном зоны проводимости и донорным уровнем (рис. 321). С повышением температуры все большее число электронов переходит из донорных состояний в зону проводимости, но, помимо этого, возрастает и число тепловых флуктуации, способных возбуждать электроны из валентной зоны и перебрасывать их через запрещенную зону энергий. Поэтому при высоких температурах уровень Ферми имеет тенденцию смещаться вниз (сплошная кривая) к своему предельному положению в центре запрещенной зоны, характерному для собственного полупроводника.

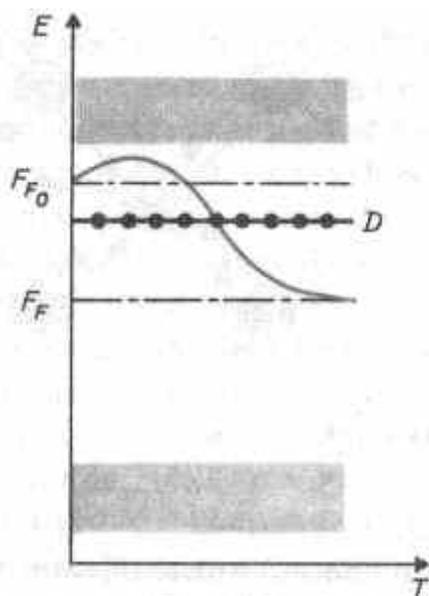


Рис. 321

Уровень Ферми в полупроводниках р-типа при 0 К E_{F0} располагается посередине между потолком валентной зоны и акцепторным уровнем (рис.322). Сплошная кривая опять-таки показывает его смещение с температурой. При температурах, при которых примесные атомы

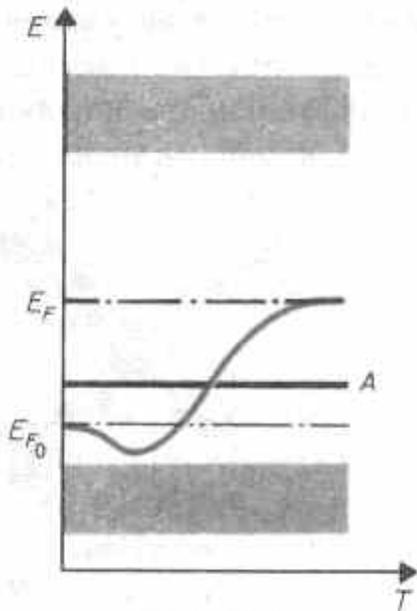


Рис. 322



Рис. 323

оказываются полностью истощенными и увеличение концентрации носителей происходит за счет возбуждения собственных носителей, уровень Ферми располагается посередине запрещенной зоны, как в собственном полупроводнике.

Проводимость примесного полупроводника, как и проводимость любого проводника, определяется концентрацией носителей и их подвижностью. С изменением температуры подвижность носителей меняется по сравнительно слабому степенному закону, а концентрация носителей — по очень сильному экспоненциальному закону, поэтому зависимость проводимости примесных полупроводников от температуры определяется в основном температурной зависимостью концентрации носителей тока в нем. На рис. 323 дан примерный график зависимости $\ln \chi$ от $1/T$ для примесных полупроводников. Участок AB описывает примесную проводимость полупроводника. Рост примесной проводимости полупроводника с повышением температуры обусловлен в основном ростом концентрации примесных носителей. Участок BC соответствует области истощения примесей (это подтверждают и эксперименты), участок CD описывает собственную проводимость полупроводника.

§ 244. Фотопроводимость полупроводников

Фотопроводимость (см. §202) **полупроводников** — увеличение электропроводности полупроводников под действием электромагнитного излучения — может быть связана со свойствами

как основного вещества, так и содержащихся в нем примесей. В первом случае при поглощении фотонов, соответствующих собственной полосе поглощения полупроводника, т. е. когда энергия фотонов равна или больше ширины запрещенной зоны ($h\nu \geq \Delta E$), могут совершаться перебросы электронов из валентной зоны в зону проводимости (рис. 324, а), что приведет к появлению добавочных (неравновесных) электронов (в зоне проводимости) и дырок (в валентной зоне). В результате возникает **собственная фотопроводимость**, обусловленная как электронами, так и дырками.

Если полупроводник содержит примеси, то фотопроводимость может возникать и при $h\nu < \Delta E$: для полупроводников с до-норной примесью фотон должен обладать энергией $h\nu \gg \Delta E_D$, а для полупроводников с акцепторной примесью — $h\nu \gg \Delta E_A$. При поглощении света примесными центрами происходит переход электронов с донорных уровней в зону проводимости в случае полупроводника n-типа (рис. 324, б) или из валентной зоны на акцепторные уровни в случае полупроводника p-типа (рис. 324, в). В результате возникает **примесная фотопроводимость**, являющаяся чисто электронной для полупроводников n-типа и чисто дырочной для полупроводников

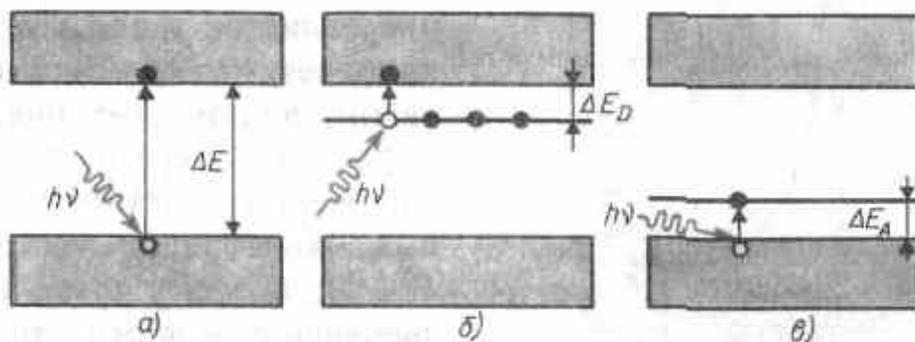


Рис. 324

Таким образом, если

$h\nu \gg \Delta E$ для собственных полупроводников,
(244.1)

$h\nu \gg \Delta E_n$ для примесных полупроводников

(ΔE_n — в общем случае энергия активации примесных атомов), то в полупроводнике возбуждается фотопроводимость. Из (244.1) можно определить **красную границу фотопроводимости** — максимальную длину волны, при которой еще фотопроводимость возбуждается:

$\lambda_0 = ch/\Delta E$ для собственных полупроводников,

$\lambda_0 = ch/\Delta E_n$ для примесных полупроводников.

Учитывая значения ΔE и ΔE_n для конкретных полупроводников, можно показать, что красная граница фотопроводимости для собственных полупроводников приходится на видимую область спектра, для примесных же полупроводников — на инфракрасную.

На рис. 325 представлена типичная зависимость фотопроводимости j и коэффициента поглощения χ от длины волны K падающего на полупроводник света. Из рисунка следует, что при $\lambda > \lambda_0$ фотопроводимость действительно не возбуждается. Спад фотопроводимости в коротковолновой части полосы поглощения объясняется большой скоростью рекомбинации в условиях сильного поглощения в тонком

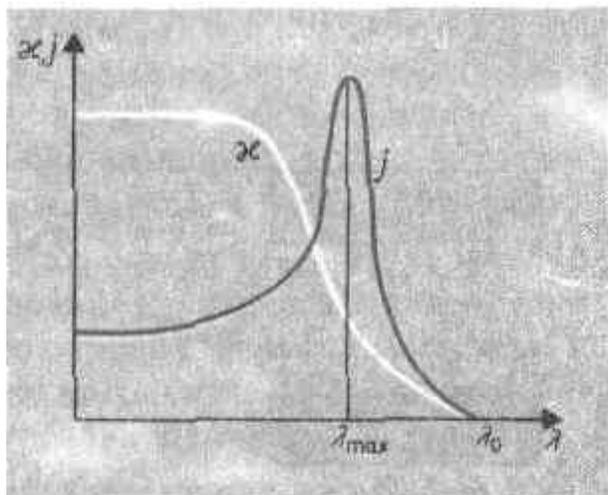


Рис. 325

поверхностном слое толщиной $\chi \approx 1$ мкм (коэффициент поглощения $\approx 10^6 \text{ м}^{-1}$).

Наряду с поглощением, приводящим к появлению фотопроводимости, может иметь место экситонный механизм поглощения. **Экситоны** представляют собой квазичастицы — электрически нейтральные связанные состояния электрона и дырки, образующиеся в случае возбуждения с энергией, меньшей ширины запрещенной зоны. Уровни энергии экситонов располагаются у дна зоны проводимости. Так как экситоны электрически нейтральны, то их возникновение в полупроводнике не приводит к появлению дополнительных носителей тока, вследствие чего экситонное поглощение света не сопровождается увеличением фотопроводимости.

§ 245. Люминесценция твердых тел

В природе давно известно излучение, отличное по своему характеру от всех известных видов излучения (теплового излучения, отражения, рассеяния света и т.д.). Этим излучением является люминесцентное излучение, примерами которого может служить свечение тел при облучении их видимым, ультрафиолетовым и рентгеновским излучением, γ -излучением и т. д. Вещества, способные под действием различного рода возбуждений светиться, получили название **люминофоров**.

Люминесценция — неравновесное излучение, избыточное при данной температуре над тепловым излучением тела и имеющее длительность, большую периода световых колебаний. Первая часть этого определения приводит к выводу, что люминесценция не является тепловым излучением (см. § 197), поскольку любое тело при температуре выше 0 К излучает электромагнитные волны, а такое излучение является тепловым. Вторая часть показывает, что люминесценция не является таким видом свечения, как отражение и рассеяние света, тормозное излучение заряженных частиц и т. д. Период световых колебаний составляет примерно 10^{-15} с, поэтому длительность, по которой свечение можно отнести к люминесценции, больше — примерно 10^{-10} с. Признак

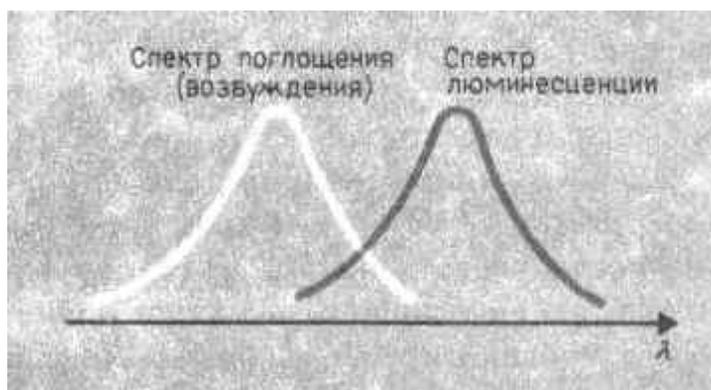


Рис. 326

длительности свечения дает возможность отличить люминесценцию от других неравновесных процессов. Так, по этому признаку удалось установить, что излучение Вавилова—Черенкова (см. §189) нельзя отнести к люминесценции.

В зависимости от способов возбуждения различают: **фотолюминесценцию** (под действием света), **рентгенолюминесценцию** (под действием рентгеновского излучения), **катодолюминесценцию** (под действием электронов), **электролюминесценцию** (под действием электрического поля), **радиолюминесценцию** (при возбуждении ядерным излучением, например γ -излучением, нейтронами, протонами), **хемилюминесценцию** (при химических превращениях), **триболлюминесценцию** (при растирании и раскалывании некоторых кристаллов, например сахара).

По длительности свечения условно различают: **флуоресценцию** ($t \leq 10^{-8}$ с) и **фосфоресценцию** — свечение, продолжающееся заметный промежуток времени после прекращения возбуждения.

Первое количественное исследование люминесценции проведено более ста лет назад [Дж. Стоксом](#), сформулировавшим в 1852 г. следующее правило: длина волны люминесцентного излучения всегда больше длины волны света, возбудившего его (рис. 326). С квантовой точки зрения правило Стокса означает, что энергия $h\nu$ падающего фотона частично расходуется на какие-то неоптические процессы, т. е.

$$h\nu = h\nu_{\text{люм}} + \Delta E,$$

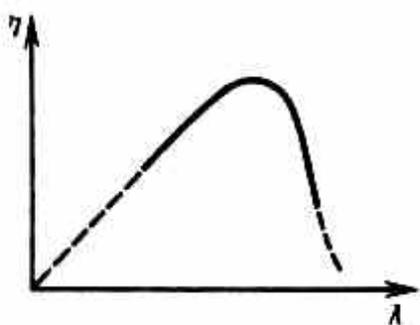


Рис. 327

откуда $\nu_{\text{люм}} < \nu$ или $\lambda_{\text{люм}} > \lambda$, что и следует из сформулированного правила.

Основной энергетической характеристикой люминесценции является **энергетический выход**, введенный С. И. Вавиловым в 1924 г. — отношение энергии, излученной люминофором при полном высвечивании, к энергии, поглощенной им. Типичная для органических люминофоров (на примере раствора флуоресцина) зависимость энергетического выхода η от длины волны λ , возбуждающего света представлена на рис. 327. Из рисунка следует, что вначале η растет пропорционально λ , а затем, достигая максимального значения, быстро спадает до нуля при дальнейшем увеличении λ (**закон Вавилова**). Величина энергетического выхода для различных люминофоров колеблется в довольно широких пределах, максимальное ее значение может достигать примерно 80%.

Твердые тела, представляющие собой эффективно люминесцирующие искусственно приготовленные кристаллы с чужеродными примесями, получили название **кристаллофосфоров**. На примере кристаллофосфоров рассмотрим механизмы возникновения люминесценции с точки зрения зонной теории твердых тел. Между валентной зоной и зоной проводимости кристаллофосфора

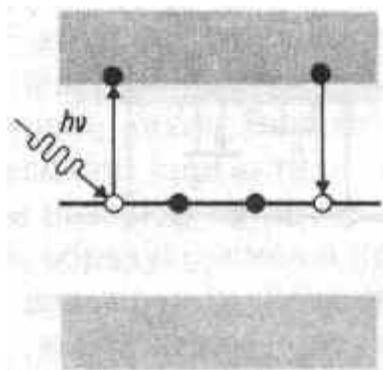


Рис. 328

поглощении атомом активатора фотона с энергией $h\nu$ электрон с примесного уровня переводится в зону проводимости, свободно перемещается по кристаллу до тех пор, пока не встретится с ионом активатора и не рекомбинирует с ним, перейдя вновь на примесный уровень. Рекомбинация сопровождается излучением кванта люминесцентного свечения. Время высвечивания люминофора определяется временем жизни возбужденного состояния атомов активатора, которое обычно не превышает миллиардных долей секунды. Поэтому свечение является кратковременным и исчезает почти вслед за прекращением облучения.

Для возникновения длительного свечения (фосфоресценции) кристаллофосфор должен содержать также **центры захвата, или ловушки** для электронов, представляющие собой незаполненные локальные уровни (например, L_1 и L_2), лежащие вблизи дна зоны проводимости (рис. 329). Они могут быть образованы атомами примесей, атомами в междоузлиях и т. д. Под действием света атомы активатора возбуждаются, т. е. электроны с примесного уровня переходят в зону проводимости и становятся свободными. Однако они захватываются ловушками, в результате чего теряют свою подвижность, а следовательно, и способность рекомбинировать с ионом активатора. Освобождение электрона из ловушки требует затраты определенной энергии, которую электроны могут получить, например, от тепловых колебаний решетки. Освобожденный из ловушки электрон попадает в зону проводимости и движется по кристаллу до тех пор, пока или не будет снова захвачен

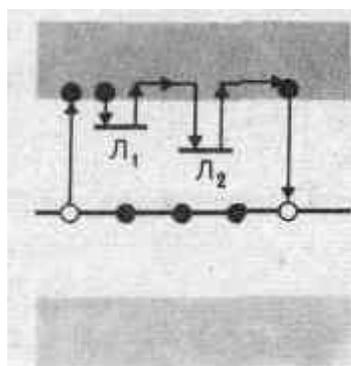


Рис. 329

В последнем случае возникает квант люминесцентного излучения. Длительность этого процесса определяется временем пребывания электронов в ловушках.

Явление люминесценции получило широкое применение в практике, например **люминесцентный анализ** — метод определения состава вещества по характерному его свечению. Этот метод, являясь весьма чувствительным (примерно 10^{-10} г/см³), позволяет обнаруживать наличие ничтожных примесей и применяется при тончайших исследованиях в биологии, медицине, пищевой промышленности и т. д. **Люминесцентная дефектоскопия** позволяет обнаружить тончайшие трещины на поверхности деталей машин и других изделий (исследуемая поверхность покрывается для этого люминесцентным раствором, который после удаления остается в трещинах).

Люминофоры используются в люминесцентных лампах, являются активной средой оптических квантовых генераторов (см. § 233) и сцинтилляторов (будут рассмотрены ниже), применяются в электронно-оптических преобразователях (см. § 169), употребляются для создания аварийного и маскировочного освещения и для изготовления светящихся указателей различных приборов.

§ 246. Контакт двух металлов по зонной теории

Если два различных металла привести в соприкосновение, то между ними возникает разность потенциалов, называемая **контактной разностью потенциалов**. Итальянский физик А. Вольта (1745— 1827) установил, что если металлы Al, Zn, Sn, Pb, Sb, Bi, Hg, Fe, Cu, Ag, Au, Pt, Pd привести в контакт в указанной последовательности, то каждый предыдущий при соприкосновении с одним из следующих зарядится положительно. Этот ряд называется **рядом Вольта**. Контактная разность потенциалов для различных металлов составляет от десятых до целых вольт.

Вольта экспериментально установил два закона:

1. Контактная разность потенциалов зависит лишь от химического состава

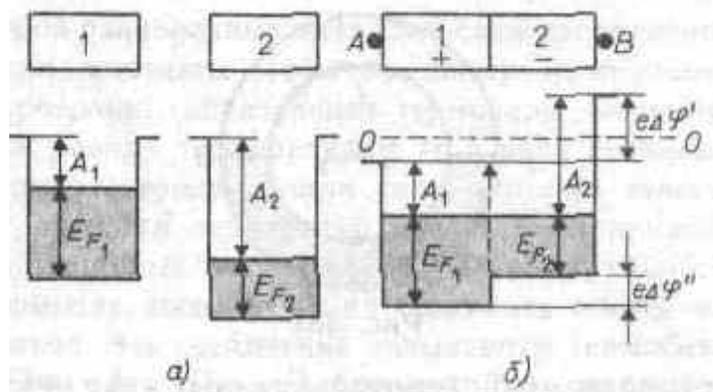


Рис. 330

и температуры соприкасающихся металлов.

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

Для объяснения возникновения контактной разности потенциалов воспользуемся представлениями зонной теории. Рассмотрим контакт двух металлов с различными работами выхода A_1 и A_2 , т.е. с различными положениями уровня Ферми (верхнего заполненного электронами энергетического уровня). Если $A_1 < A_2$ (этот случай изображен на рис. 330, а), то уровень Ферми располагается в металле 1 выше, чем в металле 2. Следовательно, при контакте металлов электроны с более высоких уровней металла 1 будут переходить на более низкие уровни металла 2, что приведет к тому, что металл 1 зарядится положительно, а металл 2 — отрицательно. Одновременно происходит относительное смещение энергетических уровней: в металле, заряжающемся положительно, все уровни смещаются вниз, а в металле, заряжающемся отрицательно, — вверх. Этот процесс будет происходить до тех пор, пока между соприкасающимися металлами не установится равновесие, кото-

рое, как доказывалось в статистической физике, характеризуется совпадением уровней Ферми в обоих металлах (рис. 330, б).

Так как для соприкасающихся металлов уровни Ферми совпадают, а работы выхода A_1 и A_2 не изменяются (они являются константами металлов и не зависят от того, находятся металлы в контакте или нет), то потенциальная энергия электронов в точках, лежащих вне металлов в непосредственной близости к их поверхности (точки Л и В на рис. 330, б), будет различной. Следовательно, между точками А и В устанавливается разность потенциалов, которая, как следует из рисунка, равна $\Delta\varphi'=(A_2-A_1)/e$. (246.1)

Разность потенциалов (246.1), обусловленная различием работ выхода контактирующих металлов, называется **внешней контактной разностью потенциалов**. Чаще говорят просто о контактной разности потенциалов, подразумевая под ней внешнюю.

Если уровни Ферми для двух контактирующих металлов не одинаковы, то между внутренними точками металлов наблюдается **внутренняя контактная разность потенциалов**, которая, как следует из рисунка, равна

$$\Delta\varphi''=(E_{F1}-E_{F2})/e. \quad (246.2)$$

В квантовой теории доказывалось, что причиной возникновения внутренней контактной разности потенциалов является различие концентраций электронов в контактирующих металлах. $\Delta\varphi''$ зависит от температуры T контакта металлов (поскольку наблюдается зависимость E_F от T), обуславливая термоэлектрические явления. Как правило, $\Delta\varphi'' \ll \Delta\varphi'$.

Если, например, привести в соприкосновение три разнородных проводника, имеющих одинаковую температуру, то разность потенциалов между концами разомкнутой цепи равна алгебраической сумме скачков потенциала во всех контактах. Она, как можно показать (предоставляем это сделать читателю), не зависит от природы промежуточных проводников (второй закон Вольта).

Внутренняя контактная разность потенциалов возникает в двойном электрическом слое, образующемся в приконтактной области, называемом **контактным слоем**. Толщина контактного слоя в металлах составляет примерно 10^{-10} м, т. е. соизмерима с междоузельными расстояниями в решетке металла. Число электронов, участвующих в диффузии через контактный слой, составляет примерно 2 % от общего числа электронов, находящихся на поверхности металла. Столь незначительное изменение концентрации электронов в контактном слое, с одной стороны, и малая по сравнению с длиной свободного пробега электрона его толщина — с другой, не могут привести к заметному изменению проводимости контактного слоя по сравнению с остальной частью металла. Следовательно, электрический ток через контакт двух металлов проходит так же легко, как и через сами металлы, т. е. контактный слой проводит электрический ток в обоих направлениях ($1 \rightarrow 2$ и $2 \rightarrow 1$) одинаково и не дает эффекта выпрямления, который всегда связан с односторонней проводимостью.

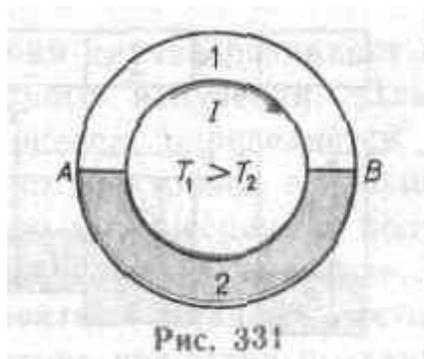
§ 247. Термоэлектрические явления и их применение

Согласно второму закону Вольта, в замкнутой цепи, состоящей из нескольких металлов, находящихся при одинаковой температуре, э.д.с. не возникает, т. е. не происходит возбуждения электрического тока. Однако если температура контактов не одинакова, то в цепи возникает электрический ток, называемый **термоэлектрическим**. Явление возбуждения термоэлектрического тока (**явление Зеебека**), а также тесно связанные с ним **явления Пельтье и Томсона** называются **термоэлектрическими явлениями**.

1. Явление Зеебека (1821). Немецкий физик Т. Зеебек (1770—1831) обнаружил, что в замкнутой цепи, состоящей из последовательно соединенных разнородных проводников, контакты между которыми имеют различную температуру, возникает электрический ток.

Рассмотрим замкнутую цепь, состоящую из двух металлических проводников 1 и 2 с температурами спаев T_1 (контакт А) и T_2 (контакт В), причем $T_1 > T_2$ (рис. 331).

Не вдаваясь в подробности, отметим, что в замкнутой цепи для многих пар



металлов (например, Cu— Bi, Ag—Cu, Au — Cu) электродвижущая сила прямо пропорциональна разности температур в контактах:

$$\xi = \alpha(T_1 - T_2).$$

Эта э.д.с. называется **термоэлектродвижущей силой**. Направление тока при $T_1 > T_2$ на рис.331 показано стрелкой. Термоэлектродвижущая сила, например для пары металлов медь — константан, для разности температур 100 К составляет всего 4,25 мВ.

Причина возникновения термоэлектродвижущей э.д.с. ясна уже из формулы (246.2), определяющей внутреннюю контактную разность потенциалов на границе двух металлов. Дело в том, что положение уровня Ферми зависит от температуры. Поэтому если температуры контактов разные, то разными будут и внутренние контактные разности потенциалов. Таким образом, сумма скачков потенциала отлична от нуля, что и приводит к возникновению термоэлектрического тока. Отметим также, что при градиенте температуры происходит и диффузия электронов, которая тоже обуславливает термо-э.д.с.

Явление Зеебека не противоречит второму началу термодинамики, так как в данном случае внутренняя энергия преобразуется в электрическую, для чего используется два источника теплоты (два контакта). Следовательно, для поддержания постоянного тока в рассматриваемой цепи необходимо поддерживать постоянство разности температур контактов: к более нагретому контакту непрерывно подводить теплоту, а от холодного — непрерывно ее отводить.

Явление Зеебека используется для измерения температуры. Для этого применяются **термоэлементы, или термопары** — датчики температур, состоящие из двух соединенных между собой разнородных металлических проводников. Если контакты (обычно спаи) проводников (проволок), образующих термопару, находятся при разных температурах, то в цепи возникает термоэлектродвижущая сила, которая зависит от разности температур контактов и природы применяемых материалов. Чувствительность термопар выше, если их соединять последовательно. Эти соединения называются **термобатареями (или термостолбиками)**. Термопары применяются как для измерения ничтожно малых разностей температур, так и для измерения очень высоких и очень низких температур (например, внутри доменных печей или жидких газов). Точность определения температуры с помощью термопар составляет, как правило, несколько кельвин, а у некоторых термопар достигает $\approx 0,01$ К. Термопары обладают рядом преимуществ перед обычными термометрами: имеют большую чувствительность и малую инерционность, позволяют проводить измерения в широком интервале температур и допускают дистанционные измерения.

Явление Зеебека в принципе может быть использовано для генерации электрического тока. Так, уже сейчас к. п. д. полупроводниковых термобатарей достигает $\approx 18\%$. Следовательно, совершенствуя полупроводниковые термоэлектрогенераторы, можно добиться эффективного прямого преобразования солнечной энергии в электрическую.

2. Явление Пельтье (1834). Французский физик Ж. Пельтье (1785—1845) обнаружил, что при прохождении через контакт двух различных проводников электрического тока в зависимости от его направления помимо джоулевой теплоты выделяется или поглощается дополнительная теплота. Таким образом, явление Пельтье является обратным по отношению к явлению Зеебека. В отличие от

джоулевой теплоты, которая пропорциональна квадрату силы тока, теплота Пельтье пропорциональна первой степени силы тока и меняет знак при изменении направления тока.

Рассмотрим замкнутую цепь, состоящую из двух разнородных металлических проводников 1 и 2 (рис. 332), по которым пропускается ток I' (его направление в данном случае выбрано совпадающим с направлением термотока (на рис.331 при условии $T_1 > T_2$)). Согласно наблюдениям Пельтье, спай A ,

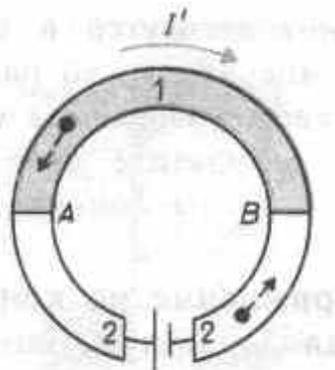


Рис. 332

при более высокой температуре, будет теперь охлаждаться, а спай B — нагреваться. При изменении направления тока I' спай A будет нагреваться, спай B — охлаждаться.

Объяснить явление Пельтье можно следующим образом. Электроны по разную сторону спаев обладают различной средней энергией (полной — кинетической плюс потенциальной). Если электроны (направление их движения задано на рис. 332 пунктирными стрелками) пройдут через спай B и попадут в область с меньшей энергией, то избыток своей энергии они отдадут кристаллической решетке и спай будет нагреваться. В спае A электроны переходят в область с большей энергией, забирая теперь недостающую энергию у кристаллической решетки, и спай будет охлаждаться.

Явление Пельтье используется в термоэлектрических полупроводниковых холодильниках, созданных впервые в 1954 г. под руководством А. Ф. Иоффе, и в некоторых электронных приборах.

3. Явление Томсона (1856). Вильям Томсон (Кельвин), исследуя термоэлектрические явления, пришел к заключению, подтвердив его экспериментально, что при прохождении тока по *неравномерно* нагретому проводнику должно происходить дополнительное выделение (поглощение) теплоты, аналогичной теплоте Пельтье. Это явление получило название явления Томсона. Его можно объяснить следующим образом. Так как в более нагретой части проводника электроны имеют большую среднюю энергию, чем в менее нагретой, то, двигаясь в направлении убывания температуры, они отдают часть своей энергии решетке, в результате чего происходит выделение теплоты Томсона. Если же электроны движутся в сторону возрастания температуры, то они, наоборот, пополняют свою энергию за счет энергии решетки, в результате чего происходит поглощение теплоты Томсона.

§ 248. Выпрямление на контакте металл — полупроводник

Рассмотрим некоторые особенности механизма процессов, происходящих при приведении в контакт металла с полупроводником. Для этого возьмем полупроводник n -типа с работой выхода A , меньшей работы выхода A_m из металла. Соответствующие энергетические диаграммы до и после вступления в контакт показаны на рис. 333, *а, б*.

Если $A_m > A$, то при контакте электроны из полупроводника будут переходить в металл, в результате чего контактный слой полупроводника обеднится электронами и зарядится положительно, а металл — отрицательно. Этот процесс будет происходить до достижения равновесного состояния, характеризуемого, как и при контакте двух металлов, выравниванием уровней Ферми для металла и полупроводника. На контакте образуется двойной электрический слой d , поле которого (контактная

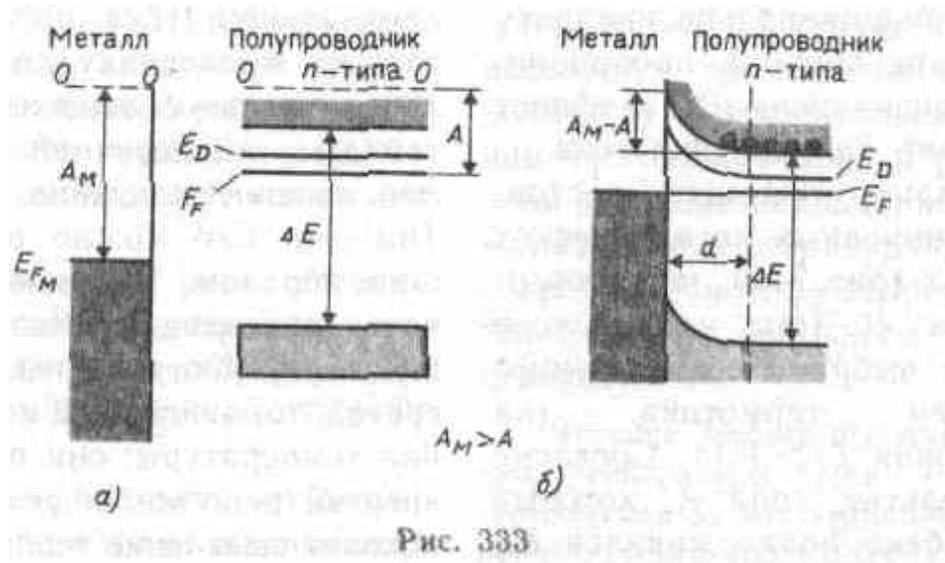
разность потенциалов) препятствует дальнейшему переходу электронов. Вследствие малой концентрации электронов проводимости в полупроводнике (порядка 10^{15}см^{-3} вместо 10^{22}см^{-3} в металлах) толщина контактного слоя в полупроводнике достигает примерно 10^{-6}см , т.е. примерно в 10000 раз больше, чем

в металле. Контактный слой полупроводника обеднен основными носителями тока — электронами в зоне проводимости, и его сопротивление значительно больше, чем в остальном объеме полупроводника. Такой контактный слой называется запирающим.

При $d = 10^{-6}\text{см}$ и $\Delta\phi \approx 1\text{В}$ напряженность электрического поля контактного слоя $E = \Delta\phi/d \approx 10^8\text{В/м}$. Такое контактное поле не может сильно повлиять на структуру спектра (например, на ширину запрещенной зоны, на энергию активации примесей и т. д.) и его действие сводится лишь к параллельному искривлению всех энергетических уровней полупроводника в области контакта (рис. 333, б). Так как в случае контакта уровни Ферми выравниваются, а работы выхода — величины постоянные, то при $A_M > A$ энергия электронов в контактном слое полупроводника больше, чем в остальном объеме. Поэтому в контактном слое дно зоны проводимости поднимается вверх, удаляясь от уровня Ферми. Соответственно происходит и искривление верхнего края валентной зоны, а также донорного уровня.

Помимо рассмотренного выше примера возможны еще следующие три случая контакта металла с примесными полупроводниками: а) $A_M < A$, полупроводник n-типа; б) $A_M > A$, полупроводник p-типа; в) $A_M < A$, полупроводник p-типа. Соответствующие зонные схемы показаны на рис. 334.

Если $A_M < A$, то при контакте металла с полупроводником n-типа электроны из металла переходят в



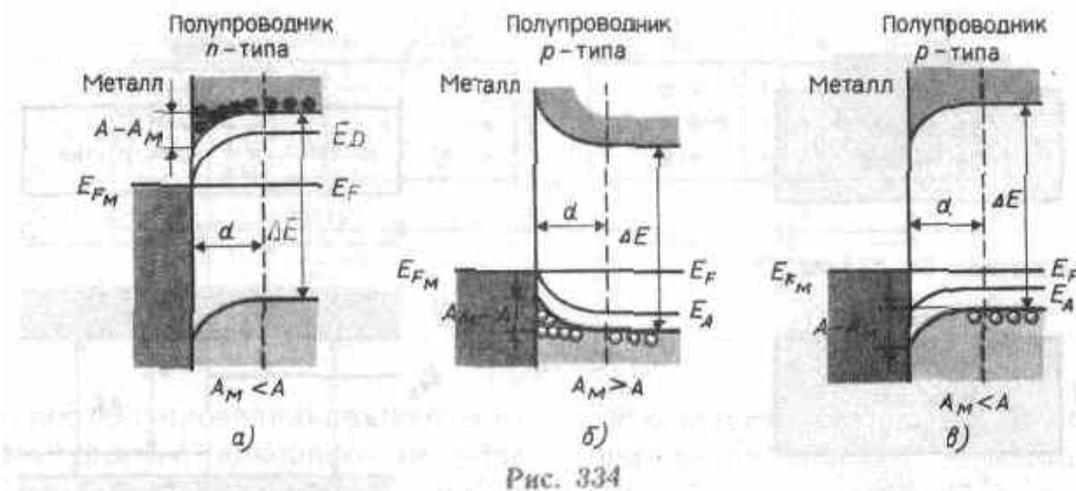


Рис. 334

разуют в контактном слое полупроводника отрицательный объемный заряд (рис. 334, а). Следовательно, контактный слой полупроводника обладает повышенной проводимостью, т. е. не является запирающим. Рассуждая аналогично, можно показать, что искривление энергетических уровней по сравнению с контактом металл— полупроводник n-типа ($A_M > A$) происходит в обратную сторону.

При контакте металла с полупроводником p-типа запирающий слой образуется при $A_M < A$ (рис. 334, в), так как в контактном слое полупроводника наблюдается избыток отрицательных ионов акцепторных примесей и недостаток основных носителей тока — дырок в валентной зоне. Если же $A_M > A$ (рис. 334, б), то в контактном слое полупроводника p-типа наблюдается избыток основных носителей тока — дырок в валентной зоне, контактный слой обладает повышенной проводимостью.

Исходя из приведенных рассуждений, видим, что запирающий контактный слой возникает при контакте донорного полупроводника с меньшей работой выхода, чем у металла (см. рис. 333, б), и у акцепторного — с большей работой выхода, чем у металла (рис. 333, в).

Запирающий контактный слой обладает **односторонней (вентильной) проводимостью**, т. е. при приложении к контакту внешнего электрического поля он пропускает ток практически только в одном направлении: либо из металла в полупроводник, либо из полупроводника в металл. Это важнейшее свойство запирающего

слоя объясняется зависимостью его сопротивления от направления внешнего поля. Если направления внешнего и контактного полей противоположны, то основные носители тока втягиваются в контактный слой из объема полупроводника; толщина контактного слоя, обедненного основными носителями тока, и его сопротивление уменьшаются. В этом направлении, называемом **пропускным**, электрический ток может проходить через контакт металл — полупроводник. Если внешнее поле совпадает по знаку с контактным, то основные носители тока будут перемещаться от границы с металлом; толщина обедненного слоя возрастает, возрастает и его сопротивление. Очевидно, что в этом случае ток через контакт отсутствует, выпрямитель заперт — это **запорное** направление. Для запирающего слоя на границе металла с полупроводником n-типа ($A_M > A$) пропускным является направление тока из металла в полупроводник, а для запирающего слоя на границе металла с полупроводником p-типа ($A_M < A$) — из полупроводника в металл.

§ 249. Контакт электронного и дырочного полупроводников (p-n-переход)

Граница соприкосновения двух полупроводников, один из которых имеет электронную, а другой — дырочную проводимость, называется **электронно-дырочным переходом** (или **p-n-переходом**). Эти переходы имеют большое практическое значение, являясь основой работы многих полу-



Рис. 335

проводниковых приборов. p-n-Переход нельзя осуществить просто механическим соединением двух полупроводников. Обычно области различной проводимости создают либо при выращивании кристаллов, либо при соответствующей обработке кристаллов. Например, на кристалл германия n-типа накладывается индиевая «таблетка» (рис. 335, а). Эта система нагревается примерно при $500\text{ }^{\circ}\text{C}$ в вакууме или в атмосфере инертного газа; атомы индия диффундируют на некоторую глубину в германий. Затем расплав медленно охлаждают. Так как германий, содержащий индий, обладает дырочной проводимостью, то на границе закристаллизовавшегося расплава и германия n-типа образуется p-n-переход (рис. 335, б).

Рассмотрим физические процессы, происходящие в p-n-переходе (рис.336). Пусть донорный полупроводник (работа выхода — A_n , уровень Ферми — E_F) приводится в контакт (рис. 336, б) с акцепторным полупроводником (работа выхода — A_p , уровень Ферми — E_F). Электроны из n-полупроводника, где их концентрация выше, будут диффундировать в p-полупроводник, где их концентрация ниже. Диффузия же дырок происходит в обратном направлении — в направлении $p \rightarrow n$.

В n-полупроводнике из-за ухода электронов вблизи границы остается нескомпенсированный положительный объемный заряд неподвижных ионизованных донорных атомов. В p-полупроводнике из-за ухода дырок вблизи границы образуется отрицательный объемный заряд неподвижных ионизованных акцепторов (рис. 336, а).

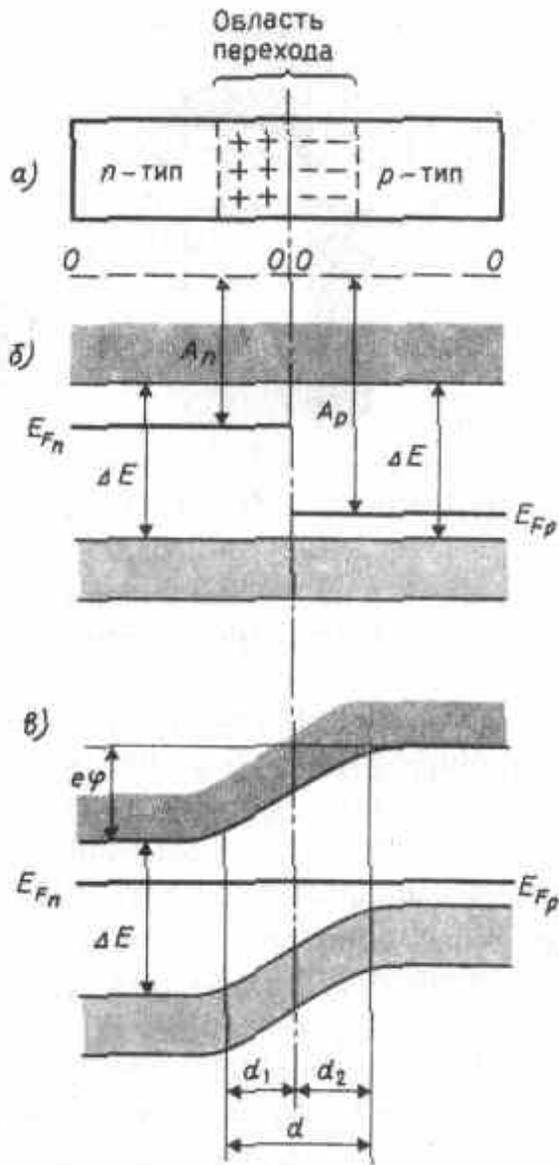


Рис. 336

Эти объемные заряды образуют у границы двойной электрический слой, поле которого, направленное от n-области к p-области, препятствует дальнейшему переходу электронов в направлении $n \rightarrow p$ и дырок в направлении $p \rightarrow n$. Если концентрации доноров и акцепторов в полупроводниках n- и p-типа одинаковы, то толщины слоев d_1 и d_2 (рис. 336, в), в которых локализируются неподвижные заряды, равны ($d_1 = d_2$).

При определенной толщине p-n-перехода наступает равновесное состояние, характеризуемое выравниванием уровней Ферми для обоих полупроводников (рис. 336, в). В области p-n-перехода энергетические зоны искривляются, в результате чего возникают потенциальные барьеры как для электронов, так и для дырок. Высота потенциального барьера $e\varphi$ определяется первоначальной разностью положений уровня Ферми в обоих полупроводниках. Все энергетические уровни

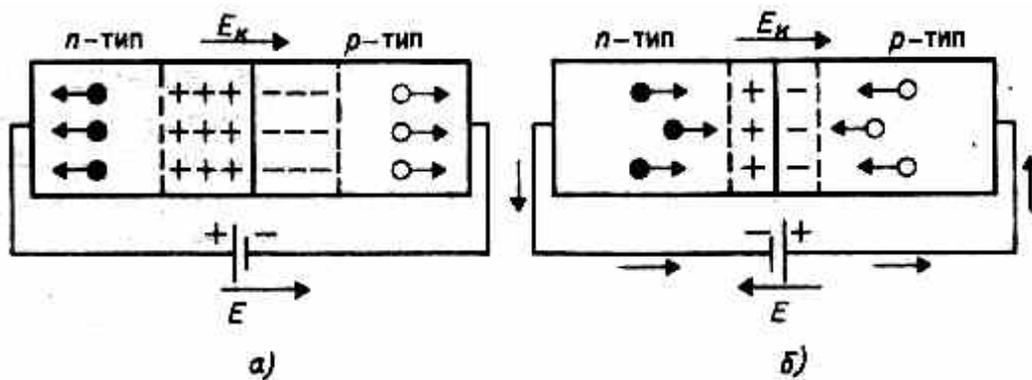


Рис. 337

акцепторного полупроводника подняты относительно уровней донорного полупроводника на высоту, равную $e\phi$, причем подъем происходит на толщине двойного слоя d .

Толщина d слоя р-п-перехода в полупроводниках составляет примерно $10^{-6} — 10^{-7}$ м, а контактная разность потенциалов — десятые доли вольт. Носители тока способны преодолеть такую разность потенциалов лишь при температуре в несколько тысяч градусов, т. е. при обычных температурах равновесный контактный слой является **запирающим** (характеризуется повышенным сопротивлением).

Сопротивление запирающего слоя можно изменить с помощью внешнего электрического поля. Если приложенное к р-п-переходу внешнее электрическое поле направлено от п-полупроводника к р-полупроводнику (рис. 337, а), т. е. совпадает с полем контактного слоя, то оно вызывает движение электронов в п-полупроводнике и дырок в р-полупроводнике от границы р-п-перехода в противоположные стороны. В результате запирающий слой расширится и его сопротивление возрастет. Направление внешнего поля, расширяющего запирающий слой, называется **запирающим (обратным)**. В этом направлении электрический ток через р-п-переход практически не проходит. Ток в запирающем слое в запирающем направлении образуется лишь за счет неосновных носителей тока (электронов в р-полупроводнике и дырок в п-полупроводнике).

Если приложенное к р-п-переходу внешнее электрическое поле направлено противоположно полю контактного слоя (рис. 337, б), то оно вызывает движение электронов в п-полупроводнике и дырок в р-полупроводнике к границе р-п-перехода навстречу друг другу. В этой области они рекомбинируют, толщина контактного слоя и его сопротивление уменьшаются. Следовательно, в этом направлении электрический ток проходит сквозь р-п-переход в направлении от р-полупроводника к п-полупроводнику; оно называется **пропускным (прямым)**.

Таким образом, р-п-переход (подобно контакту металла с полупроводником) обладает **односторонней (вентильной) проводимостью**.

На рис. 338 представлена вольт-амперная характеристика р-п-перехода. Как уже указывалось, при пропускном (прямом) напряжении внешнее электрическое поле способствует движению основных носителей тока к границе р-п-перехода (см. рис. 337, б). В результате толщина контактного слоя уменьшается. Соответственно уменьшается и сопротивление перехода (тем сильнее, чем больше напряжение), а сила тока становится большой (правая ветвь на рис. 338). Это направление тока называется **прямым**.

При запирающем (обратном) напряжении внешнее электрическое поле пре-

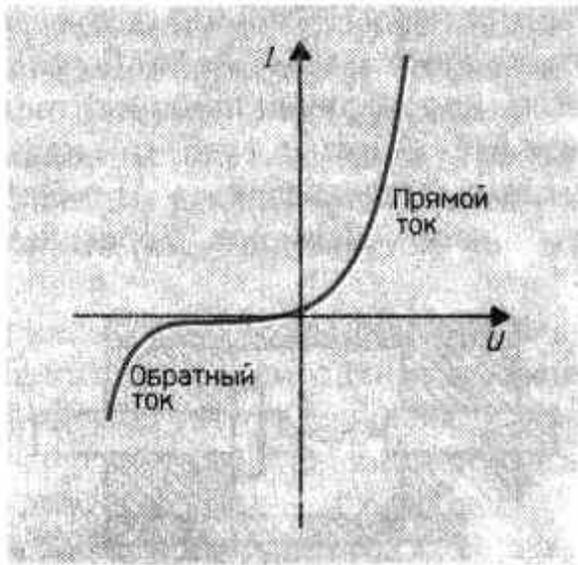


Рис. 338

пятствует движению основных носителей тока к границе p-n-перехода (см. рис. 337, а) и способствует движению неосновных носителей тока, концентрация которых в полупроводниках невелика. Это приводит к увеличению толщины контактного слоя, обедненного основными носителями тока. Соответственно увеличивается и сопротивление перехода. Поэтому в данном случае через p-n-переход протекает только небольшой ток (он называется **обратным**), полностью обусловленный неосновными носителями тока (левая ветвь рис. 338). Быстрое возрастание этого тока означает пробой контактного слоя и его разрушение. При включении в цепь переменного тока p-n-переходы действуют как выпрямители.

§ 250. Полупроводниковые диоды и триоды (транзисторы)

Односторонняя проводимость контактов двух полупроводников (или металла с полупроводником) используется для выпрямления и преобразования переменных токов. Если имеется один электронно-дырочный переход, то его действие аналогично действию двухэлектродной лампы — диода (см. § 105). Поэтому полупроводниковое устройство, содержащее один p-n-переход, называется **полупроводниковым (кристаллическим) диодом**. Полупроводниковые диоды по конструкции делятся на **точечные и плоскостные**.

В качестве примера рассмотрим точечный германиевый диод (рис.339), в котором тонкая вольфрамовая проволока 1 прижимается к n-германию 2 острием, покрытым алюминием. Если через диод в прямом направлении пропустить кратковременный импульс тока, то при этом резко повышается диффузия Al в Ge и образуется слой германия, обогащенный алюминием и обладающий

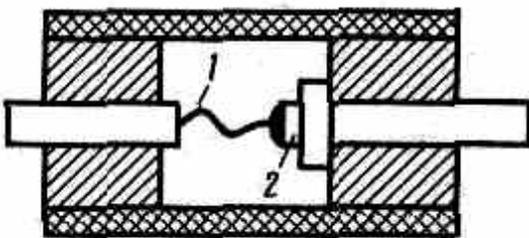
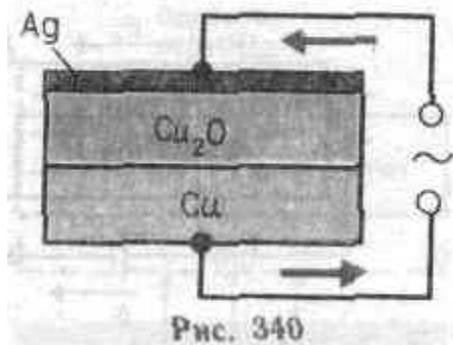


Рис. 339



На границе этого слоя образуется p - n -переход, обладающий высоким коэффициентом выпрямления. Благодаря малой емкости контактного слоя точечные диоды применяются в качестве детекторов (выпрямителей) высокочастотных колебаний вплоть до сантиметрового диапазона длин волн.

Принципиальная схема плоскостного меднозакисного (купроксного) выпрямителя дана на рис. 340. На медную пластину с помощью химической обработки наращивается слой окиси меди Cu_2O , который покрывается слоем серебра. Серебряный электрод служит только для включения выпрямителя в цепь. Часть слоя Cu_2O , прилегающая к Cu обогащенная ею, обладает электронной проводимостью, а часть слоя Cu_2O , прилегающая к Ag и обогащенная (в процессе изготовления выпрямителя) кислородом,— дырочной проводимостью. Таким образом, в толще окиси меди образуется запирающий слой с пропускным направлением тока от Cu_2O к Cu ($p \rightarrow n$).

Технология изготовления германиевого плоскостного диода описана в §249 (см. рис. 325). Распространенными являются также селеновые диоды и диоды на основе арсенида галлия и карбида кремния. Рассмотренные диоды обладают целым рядом преимуществ по сравнению с электронными лампами (малые габаритные размеры, высокие к. п. д. и срок службы, постоянная готовность к работе и т. д.), но они очень чувствительны к температуре, поэтому интервал их рабочих температур ограничен (от -70 до $+120^\circ\text{C}$). p - n -Переходы обладают не только прекрасными выпрямляющими свойствами, но могут быть использованы также для усиления, а если в схему ввести обратную связь, то и для генерирования электрических колебаний. Приборы, предназначенные для этих целей, получили название **полупроводниковых триодов** или **транзисторов** (первый транзистор создан в 1949 г. американскими физиками Д. Бардином, У. Браттейном и У. Шокли; Нобелевская премия 1956 г.).

Для изготовления транзисторов используются германий и кремний, так как они характеризуются большой механической прочностью, химической устойчивостью и большей, чем в других полупроводниках, подвижностью носителей тока. Полупроводниковые триоды делятся на **точечные и плоскостные**. Первые значительно усиливают напряжение, но их выходные мощности малы из-за опасности перегрева (например, верхний предел рабочей температуры точечного германиевого триода лежит в пределах 50 — 80°C). Плоскостные триоды являются более мощными. Они могут быть типа p - n - p и типа n - p - n в зависимости от чередования областей с различной проводимостью.

Для примера рассмотрим принцип работы плоскостного триода p - n - p , т. е. триода на основе p -полупроводника (рис. 341). Рабочие «электроды» триода, которыми являются **база** (средняя часть транзистора), **эмиттер и коллектор** (прилегающие к базе с обеих сторон области с иным типом проводимости), включаются в схему с помощью невыпрямляющих контактов — металлических проводников. Между эмиттером и базой прикладывается постоянное смещающее напряжение в прямом направлении, а между базой и коллектором — постоянное смещающее напряжение в обратном направлении. Усиливаемое переменное напряжение подается на входное сопротивление $R_{\text{вх}}$, а усиленное — снимается с выходного сопротивления $R_{\text{вых}}$.

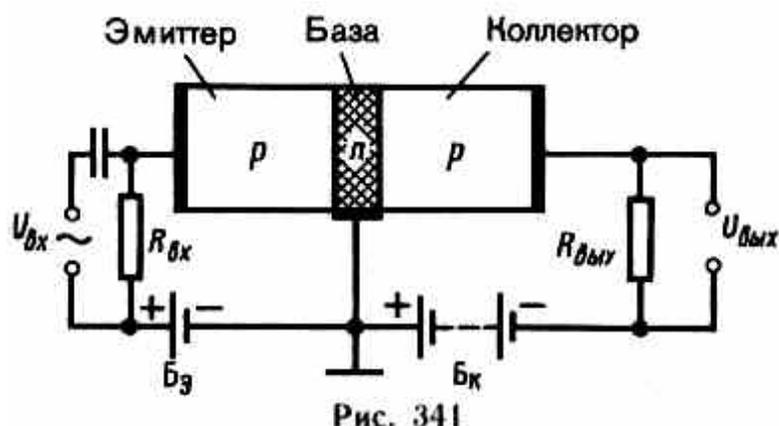


Рис. 341

Протекание тока в цепи эмиттера обусловлено в основном движением дырок (они являются основными носителями тока) и сопровождается их «впрыскиванием» — **инжекцией** — в область базы. Проникшие в базу дырки диффундируют по направлению к коллектору, причем при небольшой толщине базы значительная часть инжектированных дырок достигает коллектора. Здесь дырки захватываются полем, действующим внутри перехода (притягиваются к отрицательно заряженному коллектору), и изменяют ток коллектора. Следовательно, всякое изменение тока в цепи эмиттера вызывает изменение тока в цепи коллектора.

Прикладывая между эмиттером и базой переменное напряжение, получим в цепи коллектора переменный ток, а на выходном сопротивлении — переменное напряжение. Величина усиления зависит от свойств p-n-переходов, нагрузочных сопротивлений и напряжения батареи B_k . Обычно $R_{\text{вых}} \gg R_{\text{вх}}$, поэтому $U_{\text{вых}}$ значительно превышает входное напряжение $U_{\text{вх}}$ (усиление может достигать 10000). Так как мощность переменного тока, выделяемая в $R_{\text{вых}}$, может быть больше, чем расходуемая в цепи эмиттера, то транзистор дает и усиление мощности. Эта усиленная мощность появляется за счет источника тока, включенного в цепь коллектора.

Из рассмотренного следует, что транзистор, подобно электронной лампе, дает усиление и напряжения и мощности. Если в лампе анодный ток управляется напряжением на сетке, то в транзисторе ток коллектора, соответствующий анодному току лампы, управляется напряжением на базе.

Принцип работы транзистора *n-p-n*-типа аналогичен рассмотренному выше, но роль дырок играют электроны. Существуют и другие типы транзисторов, так же как и другие схемы их включения. Благодаря своим преимуществам перед электронными лампами (малые габаритные размеры, высокие к. п. д. и срок службы, отсутствие накаливаемого катода (поэтому потребление меньшей мощности), отсутствие необходимости в вакууме и т. д.)

транзистор совершил революцию в области электронных средств связи и обеспечил создание быстродействующих ЭВМ с большим объемом памяти.