

2. КОНТАКТНЫЕ ЯВЛЕНИЯ

2.1. Общие сведения об электрических контактах

В полупроводниковых приборах используются различные **контакты**, которые по способности пропускать электрический ток разделяются на две группы. Наиболее наглядно их различие можно оценить с помощью приведенной на рис. 2.1. **вольтамперной характеристики (ВАХ)**, представляющей собой зависимость тока, протекающего через контакт от приложенного к нему напряжения.

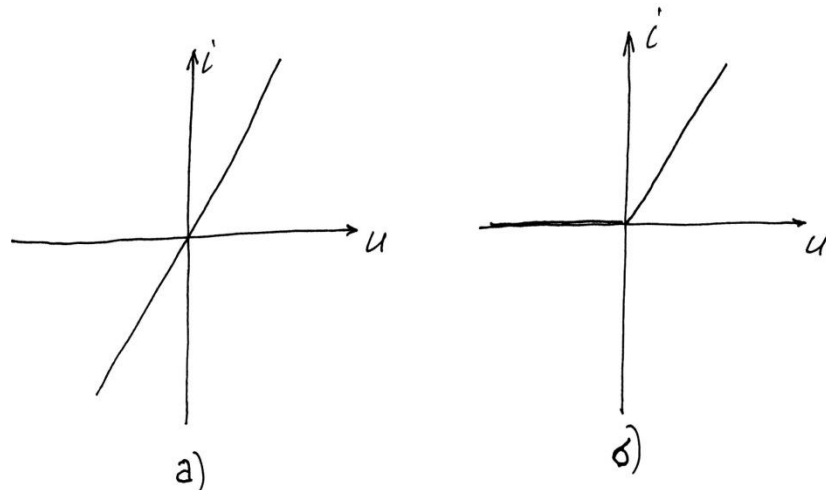


Рис. 2.1. Вольтамперные характеристики невыпрямляющего (а) и выпрямляющего (б) контактов.

К **первой группе** относятся контакты, вольтамперная характеристика которых симметрична относительно начала координат (см. рис. 2.1, а). Такие контакты одинаково пропускают ток в обоих направлениях, поэтому их называют **невыпрямляющими** (переменный ток) или **омическими**, поскольку в электрической цепи они ведут себя как линейное (омическое) сопротивление. **Вторую группу** составляют контакты, вольтамперная характеристика которых резко асимметрична относительно начала координат (см. рис. 2.1, б). Ее **прямая ветвь**, соответствующая положительным напряжениям и токам, идентична рассмотренной выше, а **обратная ветвь** практически совпадает с осью абсцисс. Такие контакты пропускают ток только в одном (прямом) направлении, т.е. обладают **односторонней проводимостью**. Они могут использоваться для **выпрямления** переменного тока (преобразования его в постоянный ток) и поэтому называются **выпрямляющими**.

Выпрямляющие контакты играют исключительно важную роль в работе полупроводниковых приборов, поскольку протекающие в них **физические процессы** (в частности, управляемая напряжением инжекция носителей заряда) **определяют принципы работы многих приборов** (диодов, транзисторов, тиристоров). Наиболее широко в полупроводниковой

электронике используется выпрямляющий контакт двух областей полупроводника, имеющих разный тип электропроводности (электронную и дырочную), т.е. областей полупроводников n- и p-типов. Такой контакт называется **электронно-дырочным переходом** или сокращенно **p-n-переходом**.

Другим контактом, широко распространенным в полупроводниковой электронике, является контакт полупроводника с металлом, называемый **контактом металл-полупроводник**. В зависимости от соотношения физических параметров металла и полупроводника, **контакт металл-полупроводник может быть как выпрямляющим, так и невыпрямляющим**. **Невыпрямляющие контакты используются для создания внешних выводов** полупроводниковых приборов, с помощью которых эти приборы включаются в электрические цепи.

Начнем рассмотрение физических процессов с выпрямляющих контактов и прежде всего, уточним определение самого понятия p-n-переход. Дело в том, что создание p-n-переходов, как и структур приборов в целом, представляет собой очень сложный технологический процесс. Главное требование к этому процессу заключается в том, чтобы на границе n- и p-областей не было значительного количества **дефектов кристаллической структуры** полупроводника. Эта технологическая граница между n- и p-областями получила название **металлургической границы**. При выполнении указанного выше требования по обе стороны от этой границы образуется **тонкий переходный слой** (толщиной в доли микрона), физические свойства которого **существенно отличаются** от свойств образующих его n- и p-областей полупроводника. Именно **этот слой и называется p-n-переходом**. Конкретные особые свойства p-n-перехода, отличающие его от n- и p-областей, будут рассмотрены ниже, в п. 2.2, 2.3 и 2.6.

Остановимся кратко на **классификации p-n-переходов**, это важно для выбора конкретного перехода, на примере которого будут рассматриваться физические процессы. Можно выделить три основных направления, по которым классифицируются переходы.

Во-первых, p-n-переходы разделяются на **гомогенные (гомпереходы)** и **гетерогенные (гетеропереходы)**. Различие между ними определяется шириной запрещенной зоны энергетической диаграммы n- и p-областей, образующих переход. **Гомогенными** называются переходы, образованные областями полупроводника с **одинаковой шириной запрещенной зоны** ($\Delta W_{zn} = \Delta W_{zp}$), **гетерогенными** - переходы, образованные областями с **разной шириной запрещенной зоны** ($\Delta W_{zn} \neq \Delta W_{zp}$), т.е. областями разных полупроводниковых материалов. Традиционные полупроводниковые приборы используют гомпереходы, поэтому для анализа физических процессов выберем именно **гомпереход**. Особенности гетеропереходов рассмотрены в [1,2].

Второе направление классификации p-n-переходов связано с соотношением концентраций донорной и акцепторной примесей в n- и p-областях,

образующих переход. В случае, когда концентрации примесей сопоставимы ($N_D \approx N_A$), переход называется **симметричным**, в случае, когда концентрации примесей различаются на порядки ($N_D \ll N_A$), переход называется **несимметричным**. Выберем для анализа **симметричный** переход, поскольку процессы в нем носят более общий характер.

Наконец, третье направление классификации связано с характером распределения концентраций примесей вблизи металлургической границы. Переход, в котором концентрации донорной примеси в n-области и акцепторной примеси в p-области распределены равномерно, а на металлургической границе резко (скачком) уменьшаются до нуля, называется **резким** переходом. В случае, когда концентрация одной из примесей вблизи металлургической границы изменяется плавно, переход называется **плавным**. Учет плавного распределения концентрации примеси существенно усложняет расчетные выражения, поэтому для анализа выберем **резкий** переход.

Таким образом, физические процессы будем анализировать на примере **резкого симметричного гомоперехода**.

2.2. Электронно-дырочный переход в равновесном состоянии (при отсутствии внешнего напряжения)

На рис. 2.2, а представлена **плоская одномерная структура** p-n-перехода. В одномерной структуре все физические величины изменяются только с одной продольной координатой x , что соответствует бесконечным поперечным размерам структуры. Координата, соответствующая металлургической границе между n- и p-областями, обозначена на рисунке как X_0 , координаты, соответствующие **границам перехода** с n- и p-областями, обозначены соответственно как X_n и X_p , **ширина перехода** обозначена как $\Delta = X_p - X_n$. Также на рисунке толстыми вертикальными линиями показаны металлические пластины, образующие с n- и p-областями невыпрямляющие контакты металл-полупроводник. К этим пластинам присоединяются **внешние выводы**, с помощью которых p-n-переход включается в электрическую цепь. Будем считать, что вывод n-области заземлен, а на вывод p-области относительно земли подается внешнее постоянное напряжение u . **Равновесным состоянием p-n-перехода** называется такое состояние, при котором **внешнее напряжение отсутствует ($u = 0$)**.

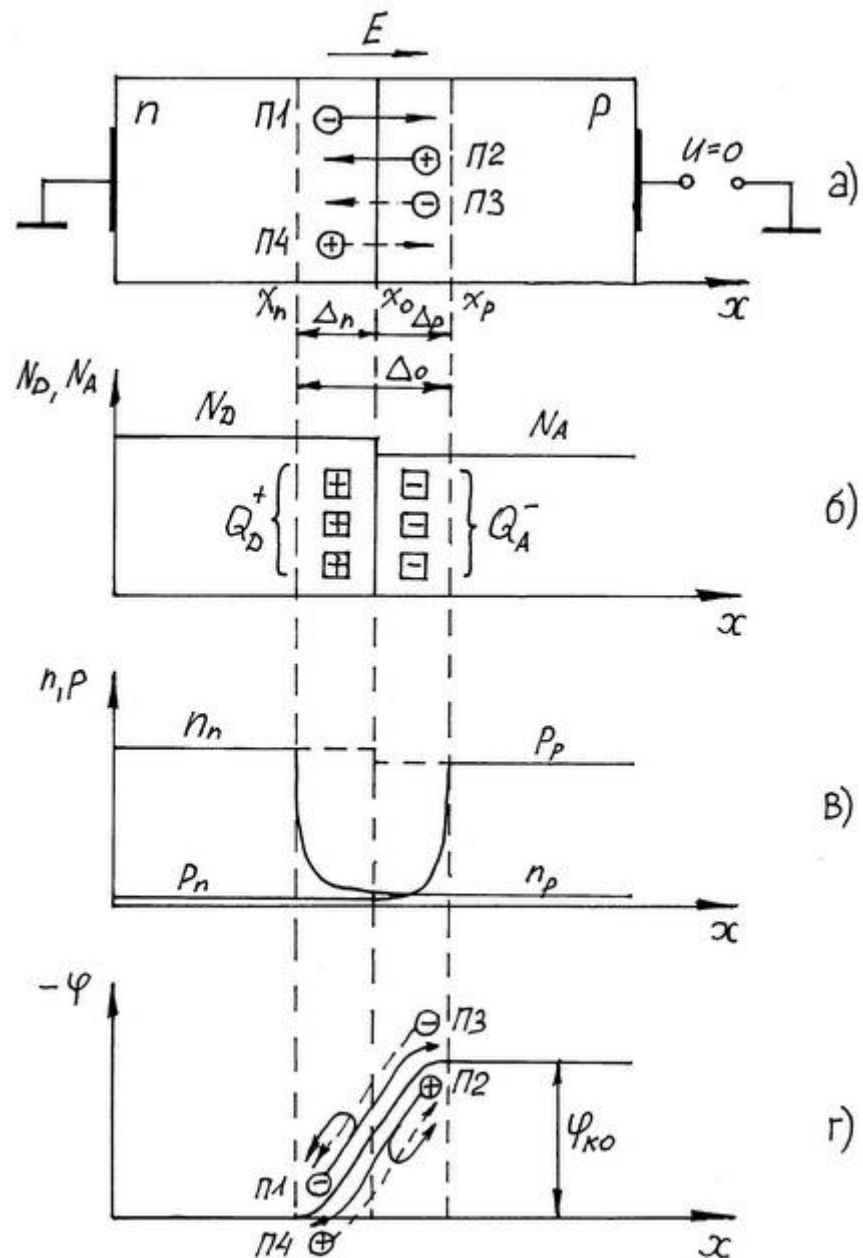


Рис. 2.2. К анализу физических процессов р-п-переходе.

Следует заметить, что поскольку рассматриваемый переход является гомогенным, для которого ширина запрещенной зоны п- и р-областей одинакова ($\Delta W_{zn} = \Delta W_{zp}$), а для разных полупроводниковых материалов ширина запрещенной зоны различна, **обе области, образующие гомопереход, представляют собой единый монокристалл одного материала**, например, кремния.

На рис. 2.2, б приведены пространственные распределения концентраций донорной N_D и акцепторной N_A примесей. Для простоты рассмотрения будем считать, что в соответствующих областях примеси распределены равномерно (в п-области $N_D = \text{const}$, а в р-области $N_A = \text{const}$). Как уже отмечалось, в резком переходе на металлургической границе концентрации

примесей скачком уменьшаются до нуля. Учитывая, что переход симметричный, будем считать $N_D \approx N_A$.

На рис. 2.2, в приведены пространственные распределения свободных носителей заряда. Рассмотрим вначале показанные **пунктиром** распределения **равновесных концентраций** основных и неосновных носителей заряда, соответствующие **отсутствию контакта** между n- и p-областями. Очевидно, что при равномерном распределении концентраций примесей, концентрации основных носителей (электронов в n-области n_n и дырок в p-области p_p) и неосновных носителей (дырок в n-области p_n и электронов в p-области n_p) будут также распределены равномерно.

Рассмотрим теперь ситуацию, при которой технологическим путем **обеспечен контакт** между n- и p-областями, причем, как уже отмечалось, на металлургической границе и вблизи ее отсутствуют дефекты кристаллической структуры. В этом случае свободные носители заряда могут свободно пересекать металлургическую границу. В результате в структуре полупроводника будет происходить **физические процессы, которые приведут к образованию p-n-перехода**, и в то же время будут определять отличие свойств этого достаточно тонкого слоя по сравнению со свойствами образующих его n- и p-областей. Рассмотрим эти процессы.

Как видно из рис. 2.2, в, **слева и справа от металлургической границы равновесная концентрация электронов и дырок различается на порядки, поэтому** при наличии контакта между областями **начнется диффузионное движение** частиц из области, в которой их концентрация велика, в область, в которой она мала. Соответственно электроны начнут двигаться из n-области в p-область (поток **П1** на рис 2.2, а), а дырки - из p-области в n-область (поток **П2** на рис 2.2, а). Заметим, что указанные потоки являются потоками **основных носителей** заряда и имеют **диффузионную** природу.

Попадая в p-область, электроны встречают на своем пути огромное количество дырок и **рекомбинируют** с ними. То же самое происходит с дырками в n-области. В результате рекомбинации концентрации диффундирующих частиц быстро убывают (см. сплошные кривые на рис. 2.2, в), и на расстоянии трех-пяти диффузионных длин их концентрации уменьшаются до **равновесных** значений. Именно эти расстояния и определяют границы перехода с n- и p-областями, в которых концентрации электронов и дырок остаются равновесными.

Как видно из рис. 2.2, в, из-за рекомбинации концентрации электронов и дырок в p-n-переходе оказываются значительно (на порядки) ниже, чем в n- и p-областях. Поэтому **p-n-переход называют слоем, обедненным подвижными носителями заряда**. Соответственно сопротивление перехода оказывается **значительно (на порядки) выше сопротивлений n- и p-областей**.

Уход значительной части свободных носителей заряда из p-n-перехода естественно приводит к нарушению условия **электронейтральности**, и в переходе остаются **нескомпенсированными заряды неподвижных ионов**

доноров (слева от металлургической границы) и **акцепторов** (справа от нее). Эти заряды показаны квадратиками на рис. 2.2, б. Суммарные заряды ионов доноров Q_D^+ и акцепторов Q_A^- одинаковы по величине и противоположны по знаку. Они создаются зарядами частиц, распределенными в конечном объеме (пространстве), потому их называют **объемными** или **пространственными** зарядами, а сам **p-n-переход** называют **слоем объемного** или **пространственного** заряда. Накопление разделенных положительных и отрицательных зарядов в p-n-переходе свидетельствует о том, что он обладает **емкостными свойствами** (подробнее см. ниже в п. 2.6).

Заряды доноров и акцепторов создают в p-n-переходе электрическое поле, вектор напряженности которого E показан на рис. 2.2, а. Это поле называется **внутренним (встроенным)** полем перехода, поскольку оно локализовано в переходе (за пределами перехода поле отсутствует). Часто его называют также **собственным** полем перехода, поскольку оно создается зарядами частиц, находящихся внутри перехода. **Внутреннее электрическое поле препятствует диффузионному движению основных носителей** заряда, **ограничивая количество электронов и дырок**, пересекающих металлургическую границу. Вместе с тем это поле **вызывает дрейфовые потоки неосновных носителей** заряда: поток электронов из p-области в n-область (пунктирный поток ПЗ на рис. 2.2, а) и поток дырок из n-области в p-область (пунктирный поток П4 на рис. 2.2, а). Таким образом, в **симметричном** переходе в равновесном состоянии протекают **четыре потока** носителей заряда: два диффузионных потока основных носителей и два дрейфовых потока неосновных носителей. Заметим, что **соответствующие потоки электронов и дырок направлены навстречу друг другу**.

Наличие электрического поля в p-n-переходе свидетельствует о том, что внутри этого слоя **изменяется электростатический потенциал**. На рис. 2.2, г приведено пространственное распределение потенциала, называемое также **потенциальной диаграммой**. Заметим, что вверх по оси ординат потенциал ϕ отложен минусом. Поскольку n-область заземлена, ее потенциал равен нулю ($\phi_n = 0$), а в p-области поле отсутствует, и ее потенциал не изменяется с координатой. Таким образом из рисунка видно, что в p-n-переходе образуется **потенциальный барьер** высотой ϕ_K .

Потенциальная диаграмма позволяет наглядно проследить влияние внутреннего поля и соответствующего ему потенциального барьера на потоки носителей заряда. На рис. 2.2, г показаны все четыре рассмотренных выше потока носителей заряда. Заметим, что в случае, когда потенциал отложен вверх минусом, **движение электронов** на потенциальной диаграмме можно рассматривать как движение **«шариков»** имеющих конечную **массу**, на которые действует **сила тяжести**. В свою очередь, **движение дырок** можно рассматривать как движение **«воздушных шариков»** в жидкости, на которые действует **выталкивающая сила**. Нетрудно увидеть, что для **основных носителей** (и электронов, и дырок) приведенный на рисунке потенциальный рельеф является **потенциальным барьером**, и лишь **немногие** частицы, обладающие достаточной энергией, способны его преодолеть, остальные

возвращаются назад. Что касается **неосновных носителей**, то для них потенциальный рельеф на рисунке представляет собой не барьер, а **потенциальную горку (ямку)**. Поэтому все неосновные носители, оказывающиеся на границе перехода, **пересекают его**.

Рассмотренные процессы позволяют заключить, что основные и неосновные носители заряда находятся в **существенно различных условиях**. С одной стороны, концентрации основных носителей на порядки превосходят концентрации неосновных носителей. С другой стороны, лишь очень небольшое количество основных носителей способно преодолеть потенциальный барьер и пересечь переход, в то время как все неосновные носители пересекают его беспрепятственно. В **равновесном состоянии** р-п-перехода, когда внешнее напряжение равно нулю ($u = 0$), **потоки основных и неосновных носителей заряда взаимно компенсируют друг друга**, и ток в переходе и во внешней цепи оказывается **равным нулю** ($i = 0$).

Важно отметить, что взаимная компенсация потоков основных и неосновных носителей заряда **обеспечивается потенциальным барьером**. Действительно **потенциальный барьер уменьшает потоки основных носителей заряда до уровня потоков неосновных носителей**. Это условие обеспечивается при **единственно возможной** высоте потенциального барьера, которая обозначается как Φ_{K0} и называется **равновесной высотой потенциального барьера** или **контактной разностью потенциалов**. Величина Φ_{K0} определяется разностью энергетических уровней Ферми в **n- и p-областях, образующих переход**.

Поскольку **внутреннее электрическое поле и соответствующий ему потенциальный барьер создаются нескомпенсированными зарядами ионов доноров и акцепторов**, концентрация которых фиксирована, конкретное значение контактной разности потенциалов Φ_{K0} обеспечивается при конкретной ширине перехода, которая обозначается как Δ_0 и называется **равновесной шириной перехода**. Величина контактной разности потенциалов невелика: в германии она составляет $\Phi_{K0} \approx 0,3$ В, а в кремнии $\Phi_{K0} \approx 0,7$ В.

Высота потенциального барьера и ширина перехода являются его важнейшими параметрами. Зная, как они изменяются при воздействии различных факторов, можно проследить за изменением тока, протекающего через переход. Выражения для расчета равновесных значений высоты потенциального барьера и ширины перехода приведены в [1,2]

$$\Phi_{K0} = u_T \ln \frac{(N_D N_A)}{n_i^2}; \quad (2.1)$$

$$\Delta_0 = \sqrt{\left(\frac{2\varepsilon\varepsilon_0}{q} \Phi_{K0} \left(\frac{1}{N_D} + \frac{1}{N_A} \right) \right)}, \quad (2.2)$$

где ε - относительная диэлектрическая проницаемость полупроводника; ε_0 - электрическая постоянная.

Как видно из формул (2.1) и (2.2), параметры перехода зависят от температуры и концентраций примесей. С **ростом температуры** линейно увеличивается термический потенциал $\mathbf{u_T}$, но увеличение квадрата собственной концентрации (даже с учетом нивелирующего эффекта логарифма) оказывается сильнее, и **контактная разность потенциалов уменьшается**. Соответственно, **уменьшается с ростом температуры и ширина перехода**. При **увеличении концентраций примесей контактная разность потенциалов растет**. В выражении (2.2) при увеличении N_D и N_A противодействуют две тенденции – рост Φ_{K0} (довольно слабый из-за логарифма) и обратно пропорциональное уменьшение членов в скобках. В результате **ширина перехода с ростом концентраций примесей уменьшается**.

Рассмотрим **особенности несимметричного перехода**. Первая особенность несимметричного перехода связана с тем, что из-за резкого различия концентраций примесей в n- и p-областях ($N_D \ll N_A$), **резко различаются** концентрации свободных носителей заряда, причем как основных, так и неосновных. Рассмотрим случай, при котором n-область легируется сильнее p-области ($N_D \gg N_A$). В этом случае для **основных** носителей заряда справедливо соотношение $n_n \gg p_p$, и при рассмотрении физических процессов **можно пренебречь потоком дырок** из p-области (поток **П2** на рис.2.2, а) по сравнению с потоком электронов из n-области (поток **П1** на рис.2.2, а). Соответственно для **неосновных носителей** справедливо соотношение $p_n \ll n_p$, и **можно пренебречь потоком дырок** из n-области (поток **П4** на рис.2.2, а) по сравнению с потоком электронов из p-области (поток **П3** на рис.2.2, а).

Таким образом, из **четырёх потоков**, которые необходимо рассматривать в **симметричном переходе**, в **несимметричном переходе** достаточно учитывать **только два потока** частиц. При этом важно, что оба эти потока представляют собой потоки **одинаковых частиц** (либо электронов, либо дырок в зависимости от того, какая из областей легирована сильнее). Конкретно необходимо учитывать потоки тех частиц, которые являются **основными в сильнолегированной области**. Заметим, что в реальных приборах используются именно **несимметричные переходы**, причем **сильнолегированная область называется эмиттером** (от английского **emit** – испускать), а **слаболегированная область называется базой** (от английского **base** – основа).

Вторая особенность несимметричного перехода связана с шириной его **участков**, которые располагаются **слева и справа** от металлургической границы. На рис. 2.2, а они обозначены соответственно $\Delta_n = X_0 - X_n$ и $\Delta_p = X_p - X_0$. Учитывая, что объемные заряды неподвижных ионов доноров и акцепторов одинаковы по величине ($Q_D^+ = Q_A^-$), а их концентрации различны, нетрудно получить пропорцию

$$\frac{\Delta_n}{\Delta_p} = \frac{N_A}{N_D}. \quad (2.3)$$

Таким образом, в симметричном переходе $\Delta_n \approx \Delta_p$, т.е. **симметричный переход** формируется относительно металлургической границы **симметрично**. В несимметричном переходе при $N_D \gg N_A$ получаем $\Delta_n \ll \Delta_p$, т.е. **несимметричный переход располагается преимущественно в слаболегированной области (базе)**. В этом случае для упрощения расчетов можно считать $\Delta \approx \Delta_p$.

В дальнейшем, будем рассматривать **только несимметричные переходы** и будем выделять **только их границы** (металлургическая граница не будет представлять интереса).

2.3. Электронно-дырочный переход при подаче внешнего напряжения.

При подаче внешнего напряжения р-п-переход оказывается в неравновесном состоянии, при котором равновесие потоков основных и неосновных носителей нарушается, и в переходе появляется **электрический ток**.

Рассмотрим процессы, протекающие при этом в **несимметричном переходе**, эмиттером которого является **n-область**, а базой - **р-область**. В этом случае $N_D \gg N_A$, и, как показано выше, можно пренебречь дырочными потоками. Для упрощения анализа сделаем еще два важных допущения. **Во-первых**, в связи с тем, что сопротивление перехода значительно выше сопротивлений образующих переход n- и р-областей, пренебрежем падением напряжения на этих областях и будем считать, что **все внешнее напряжение прикладывается к р-п-переходу**. **Во-вторых**, из-за очень малой толщины перехода **пренебрежем эффектами тепловой генерации и рекомбинации носителей заряда в нем** и, следовательно, **изменением тока внутри перехода**. Эти допущения соответствуют широко используемому в учебной литературе понятию **идеализированного перехода**.

Определимся также с такими понятиями, как **прямое и обратное напряжение (смещение перехода относительно равновесного состояния)**. Различие между ними заключается в том, в какую сторону при подаче внешнего напряжения изменяются основные параметры перехода - высота потенциального барьера ϕ_K и ширина перехода Δ . **Прямым называется такое напряжение, при котором высота потенциального барьера и ширина перехода уменьшаются, а обратным - соответственно такое напряжение, при котором они увеличиваются**.

Физические процессы, протекающие при подаче внешнего напряжения в переходе, иллюстрирует рис. 2.3. Левая его колонка соответствует прямому напряжению ($u > 0$), правая - обратному напряжению ($u < 0$). На рис. 2.3, а приведена одномерная структура перехода, на рис. 2,3, б приведена потенциальная диаграмма (пунктиром показана потенциальная диаграмма в равновесном состоянии), на рис. 2.3, в приведено распределение концентрации

электронов (неосновных носителей) в р-области (базе). Равновесное значение концентрации электронов в базе обозначено как n_p .

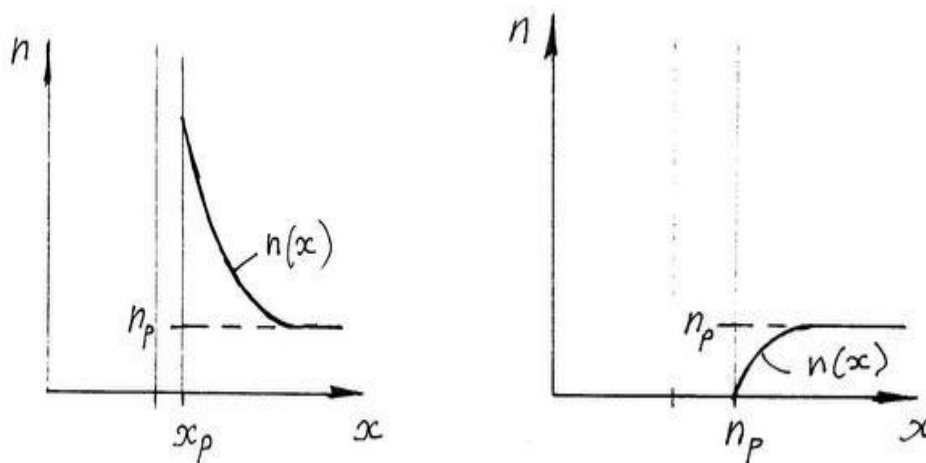
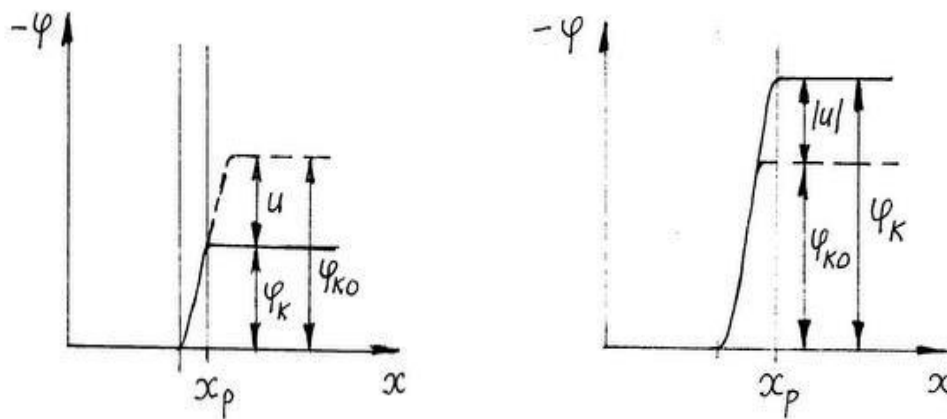
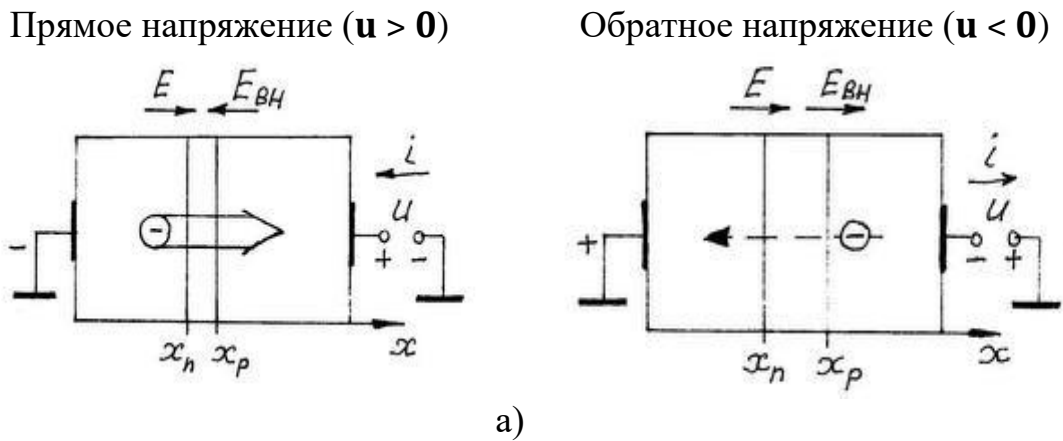


Рис. 2.3. К иллюстрации процессов, протекающих в р-п-переходе в неравновесном состоянии.

В связи с тем, что все внешнее напряжение u прикладывается к переходу, высота потенциального барьера в неравновесном состоянии Φ_K будет определяться формулой

$$\Phi_K = \Phi_{K0} - u. \quad (2.4)$$

Заметим, что поскольку прямое напряжение $u > 0$, потенциальный барьер будет ниже равновесного ($\Phi_K < \Phi_{K0}$), а поскольку обратное напряжение $u < 0$, потенциальный барьер будет выше равновесного ($\Phi_K = \Phi_{K0} + |u|$, т.е. $\Phi_K > \Phi_{K0}$).

Уменьшение потенциального барьера при прямом смещении приведет к тому, что больше электронов (основных носителей заряда в эмиттере) сможет его преодолеть, и будет происходить **инжекция** электронов через переход из эмиттера в базу. **Этот диффузионный поток электронов** (показан на рис. 2.3, а) **резко увеличивается с ростом прямого напряжения**. Что касается встречного дрейфового потока электронов (неосновных носителей) из базы в эмиттер, то поскольку они не преодолевают потенциального барьера, а, напротив, все «скатываются» с потенциальной горки, **этот поток не будет зависеть от приложенного напряжения**. Реально **этот поток очень мал, и им при прямом смещении можно пренебречь** по сравнению с потоком основных носителей (электронов из эмиттера в базу).

Вследствие **инжекции** неравновесная концентрация электронов на границе базы и перехода окажется значительно выше равновесной концентрации ($n(x_p) \gg n_p$) – см. рис. 3.3, в. В результате будет происходить **диффузия** электронов вглубь базы, где в начальный момент времени концентрация электронов остается равновесной. По мере диффузионного движения электронов через базу происходит их **рекомбинация** с дырками (основными носителями), в результате чего концентрация электронов экспоненциально уменьшается – см. рис. 3.3, в. Как показано в п.1.6, на расстоянии всего **(3-5) L_n** от границы x_p (здесь L_n - диффузионная длина электронов) концентрация электронов уменьшается до равновесного значения. Потери дырок на их рекомбинацию с электронами восполняются их поступлением из внешней цепи через вывод базы. Таким образом, три процесса: **инжекция, диффузия и рекомбинация определяют движение электронов** и, соответственно, **прямой ток через переход**.

Увеличение потенциального барьера при обратном смещении приводит к тому, что электроны (основные носители) в эмиттере не смогут его преодолеть, и их **поток уменьшается до нуля (запирается)**. Через переход будет протекать **слабый дрейфовый поток электронов (неосновных носителей заряда в базе)**, которые будут **экстрагироваться** из базы в эмиттер. Как отмечалось выше, этот поток не будет зависеть от величины приложенного напряжения.

Вследствие **экстракции** концентрация электронов на границе базы и перехода оказывается близкой к нулю ($n(x_p) \approx 0$), в то время, как в глубине базы она останется равновесной. Неравномерное распределение концентрации электронов приведет к их **диффузии** из глубины базы к переходу. В области

базы, примыкающей к границе x_p , где неравновесная концентрация электронов оказывается ниже равновесной ($n(x_p) < n_p$) - см. рис. 2.3, в, скорость **тепловой генерации** превосходит скорость рекомбинации. Образовавшиеся в результате генерации электроны экстрагируются в эмиттер, а дырки выталкиваются во внешнюю цепь. Таким образом три процесса: **экстракция, диффузия и тепловая генерация определяют движение электронов** и соответственно **обратный ток через переход**.

Важно определить **полярности напряжения, которые прикладываются к эмиттеру и базе, при прямом и обратном смещении перехода**. При подаче внешнего напряжения электрическое поле в переходе представляет собой суперпозицию внутреннего поля E и внешнего поля $E_{вн}$, создаваемого в переходе внешним напряжением. Поскольку при прямом смещении высота потенциального барьера уменьшается, суммарная напряженность поля в переходе также должна уменьшаться. Очевидно, что в этом случае вектор напряженности внешнего поля $E_{вн}$ должен быть направлен навстречу вектору напряженности внутреннего поля E - см. рис. 2.3, а, что соответствует подаче **прямого напряжения плюсом на р-область, а минусом на п-область**. Аналогичные рассуждения показывают, что **обратное напряжение подается плюсом на п-область, а минусом на р-область**.

Отметим несколько важных различий, проявляющихся при прямом и обратном смещении перехода. Во-первых, в связи с тем, что **количество электронов, инжектируемых из эмиттера в базу резко увеличивается при уменьшении высоты потенциального барьера, прямой ток, создаваемый этим потоком, очень эффективно управляется прямым напряжением и может достигать большой величины**. С другой стороны, поскольку **неосновные носители не преодолевают потенциального барьера и все пересекают переход, создаваемый ими обратный ток от величины обратного напряжения не зависит, т.е. оказывается неуправляемым**. **Величина обратного тока** определяется концентрацией неосновных носителей заряда и **оказывается очень низкой по сравнению с прямым током**.

Во-вторых, **величина прямого напряжения, которое может быть приложено к переходу, невелика и не превышает десятых долей вольта**. Нетрудно видеть, что она ограничена равновесной высотой потенциального барьера ($u < \Phi_{к0}$), поскольку при $u = \Phi_{к0}$ потенциальный барьер исчезнет и все электроны из эмиттера будут пересекать переход, в результате чего прямой ток станет очень большим и неуправляемым. Следует также заметить, что при $u \approx \Phi_{к0}$ резко увеличивается падение напряжения на «теле» базы, и формула (2.4) оказывается несправедливой. **Величина обратного напряжения никак не связана с величиной $\Phi_{к0}$, она может ее превосходить на порядки**. Ограничение обратного напряжения связано с эффектом пробоя р-п-перехода (см. ниже, п. 2.5). Таким образом, **маленькое прямое напряжение очень эффективно управляет достаточно большим прямым током, в то время как достаточно большое обратное напряжение не влияет на маленький обратный ток**.

Наконец, в-третьих, **при прямом смещении** за счет инжекции переход насыщается пересекающими его электронами, в результате чего **сопротивление перехода резко** (на порядки) **уменьшается** по сравнению с равновесным состоянием. **При обратном смещении** маленький обратный ток не может насытить переход электронами, и **сопротивление перехода остается очень большим**.

В завершении раздела два важных определения.

Состояние перехода, при котором к нему приложено прямое напряжение, через него протекает достаточно большой управляемый прямой ток, а сопротивление перехода оказывается низким, называется открытым состоянием перехода.

Состояние перехода, при котором к нему приложено обратное напряжение, через него протекает очень маленький неуправляемый обратный ток, а сопротивление перехода оказывается очень высоким называется закрытым (запертым) состоянием перехода.

2.4. Вольтамперная характеристика идеализированного перехода.

Выражение для вольтамперной характеристики (ВАХ) идеализированного перехода приведено в [1,2]

$$i = I_0 [\exp (u / u_T) - 1], \quad (2.5)$$

где I_0 – обратный тепловой ток, называемый также током насыщения перехода; u_T – термический потенциал.

В симметричном переходе ($N_D \approx N_A$) необходимо учитывать обе составляющие обратного тока, создаваемые неосновными носителями заряда (и электронами и дырками). В этом случае выражение для теплового тока имеет вид

$$I_0 = qS (\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p}), \quad (2.6)$$

где n_p и p_n – концентрации неосновных электронов и дырок, D_n и L_n , D_p и L_p – коэффициенты диффузии и диффузионные длины электронов и дырок соответственно, q – заряд электрона, S – площадь поперечного сечения перехода. В несимметричном переходе при $N_D \gg N_A$ дырочную составляющую можно не учитывать. В этом случае

$$I_0 = qS (\frac{D_n n_p}{L_n}). \quad (2.7)$$

Тепловой ток I_0 получил свое название в связи с тем, что неосновные носители появляются в примесном полупроводнике за счет тепловой генерации, и их концентрация и, соответственно, величина этого тока

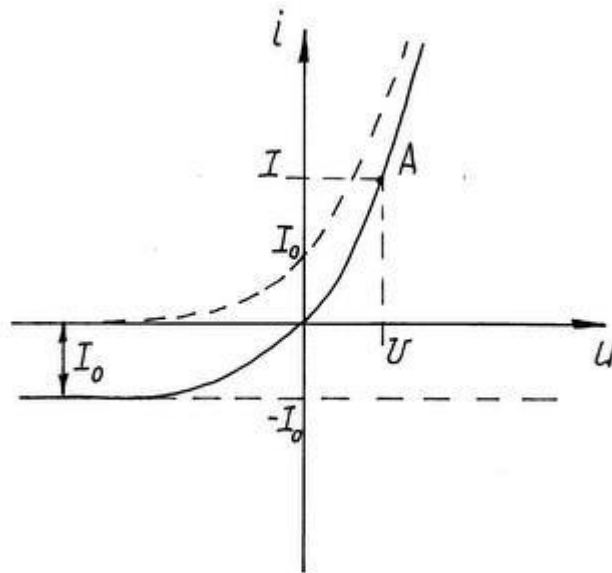


Рис. 2.4. Вольтамперная характеристика идеализированного перехода.

экспоненциально увеличиваются с ростом температуры. Термин **ток насыщения** подчеркивает тот факт, что этот ток не зависит от приложенного к переходу напряжения.

Для построения, приведенного на рис. 2.4., графика ВАХ идеализированного перехода удобно раскрыть скобки в выражении (2.5):

$$i = I_0 \exp(u/u_T) - I_0. \quad (2.8)$$

Здесь первое слагаемое отражает ток, создаваемый основными носителями заряда, и он экспоненциально зависит от напряжения. Второе слагаемое представляет собой ток, создаваемый неосновными носителями, и он от напряжения не зависит. На рис. 2.4. указанные составляющие тока приведены штриховыми линиями, а сплошной линией показана их сумма, т.е. ВАХ идеализированного перехода. Следует отметить, что поскольку масштаб по оси токов одинаков на прямой и обратной ветви, на графике представлен только самый начальный участок прямой ветви, на котором прямой ток соизмерим с обратным тепловым током. Такой режим часто называют **режимом микротоков**.

Как видно из рисунка, **ВАХ идеализированного перехода нелинейна**. Для нелинейных элементов особый интерес представляет вопрос об их электрическом сопротивлении. Во-первых, **сопротивление нелинейного двухполюсника постоянному току**, называемое также **статическим сопротивлением**, не является постоянным, а **зависит от того, в какой конкретной точке на ВАХ оно определяется** (эта точка обычно называется **рабочей точкой**). Для иллюстрации сказанного на рис. 2.4 на прямой ветви указана точка **A** и показаны величины напряжения **U** и тока **I**, соответствующие этой точке. В этом случае сопротивление перехода постоянному току в соответствии с законом Ома будет равно

$$r_0 = U/I. \quad (2.9)$$

Если рабочая точка будет смещаться по ВАХ вверх от точки А, то в силу нелинейности ВАХ напряжение при этом будет увеличиваться все меньше, а ток, напротив, будет увеличиваться все сильнее. Нетрудно увидеть, что при этом сопротивление r_0 будет уменьшаться.

Во-вторых, для описания работы нелинейных двухполюсников на переменном токе вводится параметр, называемый **дифференциальным сопротивлением**

$$r_d = du/di. \quad (2.10)$$

Выражение для дифференциального сопротивления идеализированного перехода получается дифференцированием выражения (2.5)

$$r_d = u_T / (i + I_0). \quad (2.11)$$

Графически величина дифференциального сопротивления определяется как правило через малые приращения напряжения Δu и тока Δi , которые выбираются в окрестности рабочей точки

$$r_d \approx \Delta u / \Delta i. \quad (2.12)$$

При этом необходимо, чтобы участок ВАХ, соответствующий выбранным приращениям, можно было считать линейным (с удовлетворительной точностью). Это условие получило название **критерия** так называемого **малосигнального приближения**, широко используемого в радиотехнике (подробнее см. ниже, в п. 2.7). При подаче на нелинейный двухполюсник **малого переменного напряжения дифференциальное сопротивление связывает его амплитуду с амплитудой переменного тока, протекающего через двухполюсник**

$$r_d = U_M / I_M. \quad (2.13)$$

Нетрудно убедиться, что **сопротивление постоянному току и дифференциальное сопротивление перехода различаются по величине**. На прямой ветви дифференциальное сопротивление меньше сопротивления постоянному току. Важно также отметить, что строго **вертикальному участку ВАХ** (если такой возможен) **соответствует дифференциальное сопротивление равное нулю**, а **горизонтальному**, соответственно, **равное бесконечности**. Таким образом, как видно из рис. 2.4, **дифференциальное сопротивление идеализированного перехода на обратной ветви равно бесконечности**.

2.5. Вольтамперная характеристика реального перехода (полупроводникового диода)

Электронно-дырочный переход лежит в основе структуры одного из полупроводниковых приборов – диода. Простейшая одномерная структура полупроводникового диода практически совпадает со структурой, приведенной на рис. 2.2, а. Основные свойства диодов определяются свойствами перехода, среди них **односторонняя проводимость, нелинейность ВАХ, емкостные свойства**. Эти свойства в свою очередь определяют конкретные применения полупроводниковых диодов.

В п. 2.4 рассмотрена ВАХ идеализированного перехода – см. рис. 2.4. Для получения ВАХ реального перехода (полупроводникового диода) необходимо снять принятые в п. 2.4 допущения. Во-первых, **необходимо учесть падение напряжения на п- и р-областях**, образующих переход, т.е. **на эмиттере и базе диода**. Поскольку эмиттер легируется очень сильно, его сопротивление не превышает нескольких единиц Ом, и им реально можно пренебречь. В силу более слабого легирования базы ее объемное сопротивление r'_B может достигать десятков Ом, и его необходимо учитывать. Во-вторых, **необходимо учесть тепловую генерацию и рекомбинацию электронов и дырок в самом переходе**.

Учет объемного сопротивления базы r'_B приводит к тому, что напряжение, прикладываемое к самому переходу уменьшается на величину падения напряжения на «теле» базы. Поскольку база представляет собой **линейный резистор**, падение напряжения на ней линейно связано с протекающим через переход током. В этом случае выражение для характеристики реального перехода имеет вид

$$i = I_0 \left[\exp \left(\frac{u - ir'_B}{u_T} \right) - 1 \right], \quad (2.14)$$

где ir'_B - падение напряжения на «теле» базы. Использовать для расчетов неявное выражение (2.14) неудобно, поскольку в нем ток присутствует и в левой и правой части. Для получения явного выражения удобно выразить напряжение через ток. Логарифмируя (2.14) получаем **выражение для ВАХ реального перехода** в виде

$$u = u_T \ln \left(\frac{i}{I_0} + 1 \right) + ir'_B, \quad (2.15)$$

здесь первое слагаемое отражает падение напряжения на самом (идеализированном) переходе, а второе, соответственно, на базе.

Для построения графика ВАХ реального перехода следует сделать **важное замечание**. На практике работа полупроводниковых диодов в режиме микротоков не представляет интереса. Реально даже маломощные диоды работают при прямых токах в единицы - десятки миллиампер, на несколько порядков превосходящих обратные токи. Поэтому **построить совмещенный график**, подобный приведенному на рис. 2.4, **в едином масштабе по оси токов для прямой и обратной ветвей не удастся**. Аналогичным образом не

удастся использовать единый масштаб по оси напряжений, поскольку прямое напряжение невелико, с учетом падения напряжения на базе оно не превышает одного-двух вольт, в то время, как обратное напряжение может достигать нескольких сотен вольт.

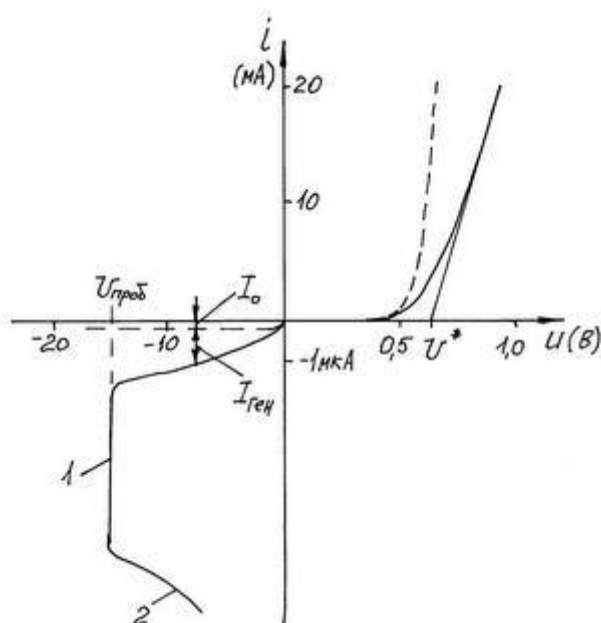


Рис. 2.5. Вольтамперная характеристика реального перехода.

Поэтому на графике ВАХ реального перехода, приведенном на рис. 2.5, масштабы по осям токов и напряжений для прямой и обратной ветвей выбраны разными. Они соответствуют характерным для маломощных диодов значениям токов и напряжений. Штриховой линией на рисунке показана ВАХ идеализированного перехода. Следует обратить внимание на характерный излом, который имеет ВАХ в начале координат. Этот излом связан с выбором разного масштаба по осям для прямой и обратной ветвей и никакой физической причины не имеет (в режиме микротоков, как видно из рис. 2.4, ВАХ представляет собой гладкую кривую).

Прямая ветвь ВАХ идеализированного перехода, отражающая первое слагаемое в выражении (2.15), характеризуется очень резким экспоненциальным ростом тока при увеличении напряжения. Нетрудно подсчитать, что при увеличении напряжения всего на 57 милливольт ток увеличивается на порядок. Учет падения напряжения на теле базы приводит к замедлению роста тока (см. сплошную линию на рис. 2.5.), причем поскольку падение напряжения на базе линейно увеличивается с ростом тока, ВАХ реального перехода все больше отстает от экспоненты и постепенно приобретает линейный характер. При достаточно больших прямых напряжениях (1...2 вольта) падение напряжения на базе оказывается больше падения напряжения на самом переходе, и ВАХ становится практически линейной.

Тепловая генерация и рекомбинация носителей заряда в переходе, приводит к тому, что к току, создаваемому пересекающими переход носителями, **добавляются дополнительные составляющие тока,** создаваемые носителями, возникающими в переходе за счет генерации и исчезающими в переходе за счет рекомбинации. Эти составляющие направлены навстречу друг другу и **называются соответственно генерационной $i_{ГЕН}$ и рекомбинационной $i_{РЕК}$.** В равновесном состоянии перехода ($u = 0$) они взаимно компенсируют друг друга. При прямом смещении рекомбинация превалирует над генерацией, и рекомбинационная составляющая тока, превышает генерационную, при обратном смещении, напротив, генерационная составляющая превышает рекомбинационную.

Рекомбинационная составляющая тока обычно очень мала по сравнению с прямым током, создаваемым носителями, инжектируемыми из эмиттера в базу, **и ей можно пренебречь.** Поэтому на рис. 2.5 вклад рекомбинационной составляющей в прямой ток не показан. **Что касается генерационной составляющей тока,** то она может не только быть сравнимой с обратным тепловым током, но и значительно его превосходить, поэтому **ее необходимо учитывать.** Приведенная на рис. 2.5 сплошной линией ВАХ реального перехода построена с учетом генерационной составляющей. Как видно из рисунка **$i_{ГЕН}$ увеличивается с ростом обратного напряжения,** что объясняется соответствующим увеличением ширины перехода.

При некотором обратном напряжении, которое называется пробивным напряжением, происходит очень резкое увеличение обратного тока, называемое пробоем перехода - см. рис. 2.5. Увеличение обратного тока связано с резким увеличением концентрации носителей заряда в переходе. Существуют **три причины увеличения концентрации носителей** и, соответственно, различают **три вида пробоя перехода.**

Первой из отмеченных причин является так называемый **эффект ударной ионизации атомов кристаллической структуры полупроводника** ускоренными электрическим полем носителями заряда. Как отмечалось в п. 1.2, при соударении с атомами подвижные носители заряда отдают им свою кинетическую энергию. При относительно слабом поле носители заряда получают от него относительно небольшую энергию, которая может только несколько увеличить амплитуду тепловых колебаний атомов. Но **при некоторой критической напряженности поля, соответствующей пробивному напряжению, энергия, получаемая носителями от поля, оказывается достаточной для ионизации атома при соударении с ним.** Этот эффект и называется **ударной ионизацией.** Нетрудно видеть, что энергия, которую для этого должны получать носители должна быть не меньше ширины запрещенной зоны энергетической диаграммы полупроводника. Появившиеся в результате ударной ионизации электроны и дырки ускоряются полем и также участвуют в ионизации атомов, в результате чего **концентрация носителей лавинно увеличивается.** Соответственно

пробой, вызываемый ударной ионизацией, называется лавинным пробоем.

Другой причиной резкого увеличения концентрации носителей заряда является так называемый **туннельный эффект**. Соответствующий вид пробоя перехода получил название **туннельного пробоя**. Туннельный эффект возникает в переходе при очень высокой концентрации примесей ($N_D, N_A > 5 \cdot 10^{18} \text{ см}^{-3}$) и очень высокой напряженности поля ($E > 10^6 \text{ В/см}$). При высокой концентрации примесей толщина перехода оказывается очень маленькой ($\Delta \approx 10^{-2} \text{ мкм}$), сравнимой с длиной свободного пробега электронов. При очень высокой напряженности поля потенциальный и, соответствующий ему энергетический барьеры, оказываются настолько высокими, что **потолок валентной зоны р-области оказывается выше дна зоны проводимости п-области**. В этом случае электроны из валентной зоны р-области могут переходить в зону проводимости п-области сквозь тонкий переход не изменяя своей энергии. Такой переход электронов и называется **туннельным эффектом**.

Лавинный и туннельный пробой объединяет то, что они происходят под действием электрического поля, поэтому их оба **называют полевыми пробоями**. Вольтамперная характеристика перехода при возникновении лавинного и туннельного пробоев также ведет себя одинаково – см. участок 1 на рис 2.5. Обратный ток в режиме пробоя увеличивается очень сильно и может достигать величины, сравнимой с прямым током, поэтому масштаб по оси токов на этом участке ВАХ следует считать условным.

Другая особенность, объединяющая лавинный и туннельный пробой, заключается в том, что **оба они оказываются обратимыми**, т.е. не вызывают выхода из строя полупроводниковых диодов. Обратимые эффекты, как правило, используются в различных приборах. Так **лавинный и туннельный пробой используются в специальных диодах, называемых стабилитронами**. Поскольку в режимах лавинного и туннельного пробоя напряжение на диоде изменяется в очень небольших пределах, **стабилитроны**, как следует из их названия, **используются для стабилизации постоянного напряжения в электрических схемах**. Следует отметить, что при относительно низких пробивных напряжениях (до 6...7 вольт) проявляется туннельный пробой, при напряжениях выше 9...10 вольт имеет место лавинный пробой, при напряжениях 6...10 вольт могут происходить оба пробоя одновременно.

В СВЧ диапазоне широко используются специальные диоды, принцип работы которых основан на использовании лавинного и туннельного пробоя, называемые, соответственно, **лавинно-пролетным диодом (ЛПД) и туннельным диодом (ТД)**. Эти диоды будут рассматриваться в ходе изучения дисциплины «Электроника».

Третий вид пробоя перехода получил название **теплового пробоя**, поскольку в этом случае **причиной увеличения концентрации носителей заряда является ускорение тепловой генерации в переходе, связанное с увеличением температуры перехода** из-за выделения в нем тепла

протекающим током. Здесь следует отметить, что при увеличении концентрации носителей увеличивается ток, протекающий через переход, и, соответственно, мощность, выделяемая в переходе. В результате увеличивается температура и еще больше ускоряется тепловая генерация. Таким образом замыкается **цепь положительной обратной связи**, приводящей к неограниченному росту температуры и разрушению кристаллической структуры полупроводника. **Тепловой пробой оказывается необратимым, при его возникновении полупроводниковые приборы выходят из строя.** Участок ВАХ, соответствующий тепловому пробую (обозначен цифрой 2 на рис. 2.5) характеризуется уменьшением напряжения при увеличении тока.

Обычные полупроводниковые диоды, например, выпрямительные **не предназначены для работы в режиме пробоя**, поэтому их **обратное напряжение ограничивается** указываемым в справочниках **максимальным (допустимым) значением**, которое на 20...30% ниже пробивного напряжения.

Для расчетов удобно использовать **кусочно-линейную аппроксимацию ВАХ** реального перехода (диода). Для ее построения линейный участок прямой ветви ВАХ продлевается до пересечения с осью напряжений – см. прямую линию на рис. 2.5. Напряжение, соответствующее пересечению этой линии с осью напряжений U^* , **называется пороговым напряжением**. В результате получаем аппроксимацию ВАХ в виде

$$i = \begin{cases} 0 & \text{при } u < U^*, \\ \frac{u - U^*}{r_d} & \text{при } u > U^*, \end{cases} \quad (2.16)$$

где r_d – дифференциальное сопротивление перехода в области рабочих токов. **Пороговое напряжение определяется равновесной высотой потенциального барьера** в переходе ($U^* \approx \phi_{K0}$), в германиевых диодах $U^* \approx 0,3$ В, в кремниевых диодах $U^* \approx 0,7$ В. Таким образом, полупроводниковый диод **открывается при прямом напряжении, превышающем пороговое напряжение. При напряжениях, меньших порогового, диод остается практически запертым.**

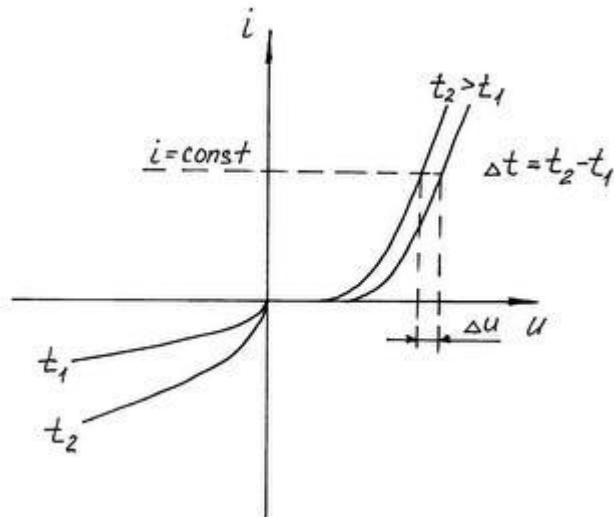


Рис. 2.6. К оценке влияния температуры на вольтамперную характеристику реального перехода.

Влияние температуры на ВАХ диода иллюстрирует рис. 2.6, на котором представлены две характеристики, различающиеся температурой. Заметим, что здесь температура выражается в $^{\circ}\text{C}$. Уменьшение высоты потенциального барьера с ростом температуры объясняет увеличение прямого тока при заданном напряжении. Если рассматривать ситуацию при постоянном токе, то при увеличении температуры прямое напряжение уменьшается, т.е. характеристика несколько смещается в область более низких напряжений – см. рис. 2.6. Это смещение можно оценить с помощью температурного коэффициента напряжения

$$\varepsilon_u = \left. \frac{\Delta u}{\Delta t} \right|_{i = \text{const}}. \quad (2.17)$$

Для германиевых диодов $\varepsilon_u \approx - 2 \text{ мВ}/^{\circ}\text{C}$; для кремниевых диодов $\varepsilon_u \approx - 3 \text{ мВ}/^{\circ}\text{C}$.

Увеличение обратного тока с ростом температуры объясняется возрастанием скорости тепловой генерации и соответствующим увеличением концентрации неосновных носителей заряда. Если известен обратный ток при некоторой температуре t_0 , то для оценки величины обратного тока при изменении температуры на величину Δt используется выражение

$$i(t_0 + \Delta t) = i(t_0) 2^{\Delta t/t^*}, \quad (2.18)$$

где t^* - температура удвоения обратного тока. Для германиевых диодов $t^* = 10 \text{ }^{\circ}\text{C}$; для кремниевых диодов $t^* = 8 \text{ }^{\circ}\text{C}$.

2.6. Емкостные свойства перехода

Как отмечалось в п. 2.2., **p-n-переход обладает емкостными свойствами**, которые выражаются в способности накапливать положительные и отрицательные заряды и изменять их количество при изменении приложенного напряжения. Поскольку процесс накопления заряженных частиц происходит в двух характерных областях структуры перехода, **емкость перехода складывается из двух составляющих**.

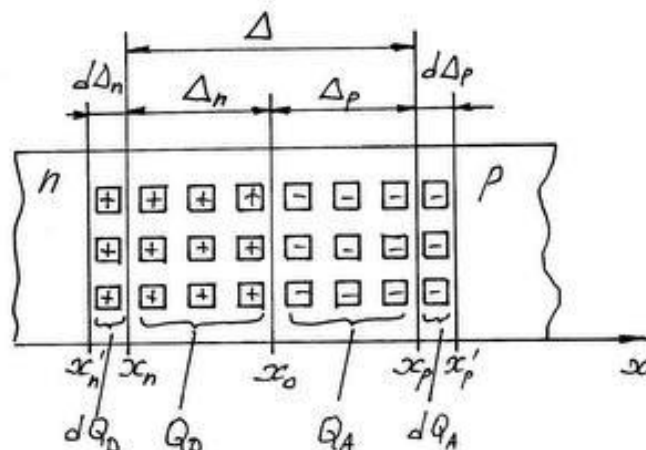


Рис.2.7. К формированию барьерной емкости перехода.

Первая из отмеченных составляющих емкости определяется **накоплением в переходе положительных и отрицательных объемных зарядов неподвижных ионов доноров и акцепторов**. Эта составляющая называется **барьерной емкостью перехода C_B** . Название связано с тем, что барьерная емкость создается теми же зарядами, которые создают в переходе электрическое поле и потенциальный барьер. Формирование барьерной емкости иллюстрирует рис. 2.7, на котором представлен фрагмент одномерной структуры симметричного перехода при подаче на него небольшого обратного напряжения.

Как видно из рисунка, при подаче на переход обратного напряжения его ширина увеличивается, и к зарядам ионов доноров Q_D и акцепторов Q_A добавляются заряды dQ_D и dQ_A двух тонких слоев толщиной $d\Delta_n$ и $d\Delta_p$, расположенных на расстоянии, равном ширине перехода Δ . Поэтому **барьерная емкость может рассматриваться как емкость плоского конденсатора с расстоянием между обкладками, равным ширине перехода**. Следует заметить, что роль изолятора между обкладками играет сам переход, поскольку при обратном смещении его сопротивление очень велико. Выражение для барьерной емкости в этом случае имеет вид

$$C_B = \frac{\epsilon \epsilon_0 S}{\Delta}, \quad (2.19)$$

где ϵ - относительная диэлектрическая проницаемость полупроводника; ϵ_0 - электрическая постоянная; S - площадь поперечного сечения перехода.

При увеличении обратного напряжения ширина перехода увеличивается, и **барьерная емкость уменьшается** согласно выражению

$$C_B = C_{B0} / \sqrt{1 - (u/\phi_{K0})}, \quad (2.20)$$

где C_{B0} – барьерная емкость в равновесном состоянии ($u = 0$). Зависимость барьерной емкости от напряжения на переходе иллюстрирует рис. 2.8.

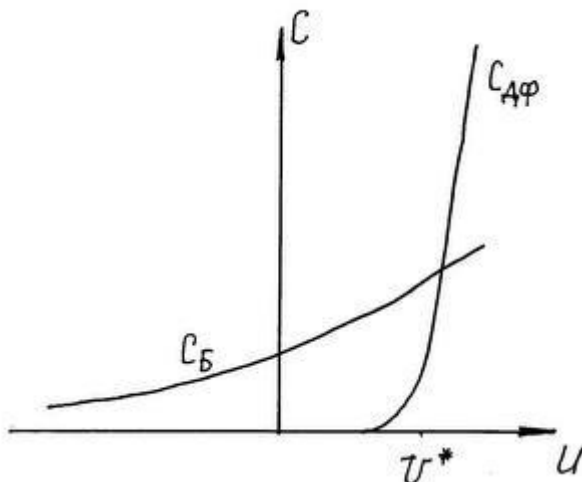


Рис. 2.8. Вольтфарадная характеристика перехода.

Вторая составляющая емкости перехода определяется **накоплением избыточных зарядов подвижных носителей в базе перехода**. Как видно из рис. 2.3, в, при подаче прямого напряжения вследствие инжекции в прилегающей к переходу области базы избыточная концентрация электронов может значительно превышает равновесную концентрацию. В силу диэлектрической релаксации в этой области возникает такая же по величине избыточная концентрация дырок ($\Delta p \approx \Delta n$) – см. выше, п. 1.6.. **Накопление в базе перехода избыточных объемных зарядов электронов и дырок определяет появление диффузионной емкости перехода**. Ее название связано с диффузионным движением электронов в базе. Поскольку избыточная концентрация электронов в базе увеличивается с ростом прямого тока, **диффузионная емкость также растет с увеличением прямого тока** в соответствии с выражением

$$C_{ДФ} = i\tau_n/u_T, \quad (2.21)$$

где τ_n - время жизни электронов в базе.

Зависимость диффузионной емкости от напряжения на переходе также приведена на вольтфарадной характеристике – см. рис. 2.8. Как видно из рисунка, вид этой зависимости аналогичен виду прямой ветви вольтамперной характеристики перехода. **При обратном смещении диффузионная емкость практически равна нулю, и при оценке емкости перехода достаточно учитывать его барьерную емкость. При прямом напряжении, превышающем пороговое напряжение, когда переход отпирается, и через него течет большой прямой ток, диффузионная емкость резко увеличивается и может на порядки превосходить барьерную, величиной которой в этом**

случае **можно пренебречь**. При небольших прямых напряжениях необходимо учитывать обе составляющие емкости перехода

$$C = C_B + C_{дф}. \quad (2.22)$$

Следует иметь в виду, что **заряд и разряд емкости** (накопление и рассасывание избыточных зарядов) **представляет собой инерционные процессы**. Поэтому **наличие емкости переходов ограничивает быстродействие полупроводниковых приборов**. Поскольку время заряда и разряда пропорционально величине емкости, влияние диффузионной емкости, которая может принимать достаточно большие величины, сказывается сильнее, чем влияние барьерной емкости. Для того, **чтобы уменьшить диффузионную емкость необходимо** в соответствии с выражением (2.21) **уменьшать время жизни электронов в базе**. С этой целью базы диодов и биполярных транзисторов легируют специальными примесями, создающими разрешенные энергетические уровни вблизи середины запрещенной зоны – ловушки электронов – см. выше, п. 1.6.

Емкостные свойства перехода находят практическое применение в специальном диоде, который называется варикапом. Он используется в качестве **нелинейной емкости, изменяемой приложенным к нему напряжением**. Варикап работает при обратном смещении, и в основе его работы лежит зависимость барьерной емкости от напряжения.

2.7. Математические модели и эквивалентные схемы полупроводникового диода

Для расчета различных схем используются **математические модели полупроводниковых диодов**. Они **представляют собой совокупность эквивалентной схемы диода и математических выражений, описывающих связь внешних параметров (токов и напряжений) с элементами этой схемы**. Следует подчеркнуть, что математическая модель должна быть адекватна режиму, в котором используется диод. **Конкретный вид эквивалентной схемы диода и, соответственно, выражений, описывающих ее работу, очень сильно зависят от параметров подаваемого на схему сигнала, в частности, его амплитуды и частоты**.

Здесь следует отметить характерные режимы, которые выделяются в зависимости от параметров входного сигнала. В частности, различают **режимы малого сигнала (малых амплитуд) и большого сигнала (больших амплитуд)**. Критерий режима малого сигнала отмечался в п. 2.4. В соответствии с ним **режим может считаться малосигнальным в случае, если участок ВАХ, в пределах которого перемещается рабочая точка под воздействием переменного напряжения, можно считать линейным**. В противном случае имеет место **режим большого сигнала**. **Режим, при котором переменный сигнал отсутствует, и на диод подается постоянное напряжение, называется статическим**.

Физические процессы, определяющие работу диода (движение носителей заряда и перезаряд емкостей перехода) являются инерционными, т.е. занимают конечное время. Если время задержки распространения сигнала в диоде составляет ничтожно малую часть периода сигнала, то режим называют режимом низкой частоты (квазистатическим). В этом режиме можно пренебречь инерционностью упомянутых процессов и считать время задержки равным нулю. В случае, когда время задержки сопоставимо с периодом, инерционностью процессов пренебречь нельзя, и режим называется высокочастотным (динамическим).

На рис. 2.9, а приведена эквивалентная схема диода, соответствующая наиболее общему случаю - режиму большого сигнала и высокой частоты. Приведенный на схеме диод соответствует идеализированному переходу и отражает нелинейную связь тока и напряжения. Параллельно переходу включены конденсаторы, соответствующие барьерной и диффузионной емкостям перехода и отражающие инерционность протекающих в переходе процессов. Подчеркнем, что эти емкости также оказываются нелинейными, поскольку их величины зависят от приложенного напряжения – см. выше, п. 2.6. При наличии емкостей к току i , создаваемому носителями заряда, пересекающими переход, добавляется емкостной ток i_C . Линейный резистор r'_B представляет собой объемное сопротивление («тела») базы.

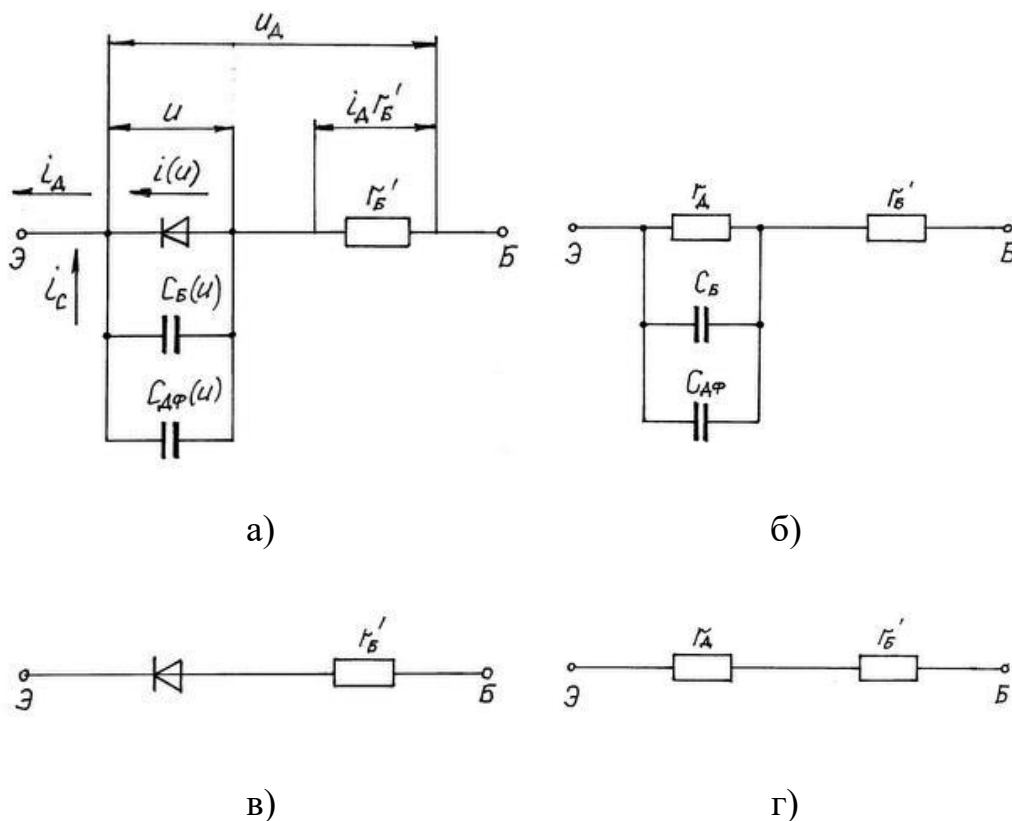


Рис. 2.9. Эквивалентные схемы полупроводникового диода в режимах: большого сигнала и высокой частоты (а); малого сигнала и высокой частоты (б); большого сигнала и низкой частоты (в); малого сигнала и низкой частоты (г).

Процессы в приведенной на рис. 2.9, а эквивалентной схеме диода описываются системой уравнений

$$\begin{aligned} u_D &= u + i_D r'_B, \\ i_D &= i + i_C, \\ i &= I_0 \left[\exp\left(\frac{u}{u_T} - 1\right) \right], \\ i_C &= [C_B(u) + C_{ДФ}(u)] \frac{du}{dt}, \end{aligned} \quad (2.23)$$

где $C_B(u)$ и $C_{ДФ}(u)$ определяются выражениями (2.20) и (2.21). Заметим, что **рассмотренная модель** является простейшей и **не учитывает многих факторов**, например, эффект пробоя перехода, ток утечки и других. Вместе с тем аналитически решить систему (2.23) не удастся, и поэтому данная модель **используется для численных расчетов в компьютерных программах математического моделирования процессов в полупроводниковых диодах.**

На рис. 2.9, б приведена эквивалентная схема диода, соответствующая более простому случаю – **режиму малого сигнала и высокой частоты**. В малосигнальном режиме **переход можно заменить его дифференциальным сопротивлением**, рассчитанным в рабочей точке по постоянному току. **Барьерная и диффузионная емкости оказываются линейными**, они также рассчитываются в рабочей точке по постоянному току. Таким образом **в малосигнальном режиме эквивалентная схема диода оказывается линейной**, и для расчета протекающих в ней процессов можно использовать известные методы теории линейных электрических цепей.

На рис. 2.9, в приведена эквивалентная схема диода, соответствующая режиму **большого сигнала и низкой частоты**. В квазистатическом режиме можно пренебречь инерционностью процессов, протекающих в диоде, и не включать емкости в схему. Нелинейность связи тока с напряжением отражает, как и в схеме, приведенной на рис. 2.9, а, диод, соответствующий идеализированному переходу.

Наконец, на рис. 2.9, г приведена эквивалентная схема диода для самого простого случая – **режима малого сигнала и низкой частоты**. Эта линейная схема содержит два резистора, соответствующие дифференциальному сопротивлению перехода и объемному сопротивлению базы.