

«Харьковский авиационный институт», 2002 г.
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
Национальный аэрокосмический университет им. Н.Е. Жуковского
«Харьковский авиационный институт»

С.Н. Барсуков

ЭЛЕМЕНТНАЯ БАЗА РАДИОЭЛЕКТРОНИКИ
Часть 1
ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

Учебное пособие

Харьков «ХАИ» 2002

УДК 621.382

Элементная база радиоэлектроники. Ч. 1. Полупроводниковые диоды/С.Н. Барсуков. – Учеб. пособие. – Харьков: Нац. аэрокосмический ун-т «Харьк. авиац. ин-т», 2002. – 88 с.

Рассмотрены физические основы работы различных полупроводниковых диодов. Приведены основные параметры, характеристики и примеры применения приборов.

Для студентов факультета радиотехнических систем летательных аппаратов.

Ил. 95. Табл. 4. Библиогр.: 8 назв.

Рецензенты: д-р. техн. наук В.Д. Карлов,
канд. техн. наук Е.А. Милькевич

©Национальный аэрокосмический университет им. Н.Е. Жуковского

1. Физические основы полупроводниковых приборов

1.1. Электропроводность полупроводников

Полупроводники – вещества, занимающие промежуточное положение между металлами и диэлектриками. Проводящие свойства полупроводников количественно оценивают по параметру удельной проводимости, удовлетворяющему следующему условному неравенству:

$$(10^{-12} \dots 10^{-10}) < \sigma < (10^4 \dots 10^6) \quad (\text{Ом}\cdot\text{см})^{-1}.$$

Этот параметр определяет величину объемного сопротивления

$$R = \frac{1}{\sigma} \cdot \frac{\ell}{S}.$$

В полупроводниковой электронике находит применение широкий класс полупроводниковых материалов:

1. Элементарные полупроводники, относящиеся к IV группе периодической системы элементов: кремний (Si), германий (Ge), т.е. 4-х валентные вещества. Особое место занимает углерод (C) с кристаллической решеткой типа алмаза. Полупроводниковые алмазы являются перспективным материалом для разработки электронных приборов, работающих в экстремальных условиях.

2. Алмазоподобные полупроводники (по типу кристаллической решетки), т.е. вида $A^{III}B^V$, представляющие собой соединения элементов III группы (Al, Ge, In) с элементами V группы (As, P, Sb). Это такие вещества, как арсенид галлия (GaAs), фосфиды галлия и индия (GaP, InP), а также непрерывный ряд твердых растворов вида $Ga_xAl_{1-x}As$, $GaAs_{1-x}P_x$. В этих соединениях на один атом вещества приходится в среднем по 4 валентных электрона, поэтому по структуре кристаллической решетки они являются аналогами германия и кремния.

3. Соединения элементов вида $A^{II}B^{VI}$ (например, CdS, ZnSe, ZnTe) также имеют в среднем по 4 электрона на один атом, но свойства полупроводников у них выражены слабее. В этих веществах более проявляется дополнительная ионная связь.

4. Соединения элементов вида $A^{IV}B^{VI}$ (например, PbS, PbSe, GeTe) имеют в среднем по 5 валентных электронов на атом, используются как приемники инфракрасного излучения.

5. Органические полупроводники в виде монокристаллов, аморфных порошков и пленок находят применение в качестве светочувствительных материалов.

С точки зрения электропроводности полупроводники отличаются от металлов двумя факторами:

- "аномальной" зависимостью проводимости от температуры (увеличение электропроводности при возрастании температуры);
- сильной зависимостью проводимости от концентрации примесей.

Первый фактор является негативным с точки зрения стабильности параметров полупроводниковых приборов. В то же время возможность управлять в широких пределах электропроводностью полупроводника путем технологически регулируемого внедрения примесей является основой разнообразных применений полупроводников.

Электропроводность полупроводника определяется структурой энергетической диаграммы. В соответствии с квантовой теорией строения вещества отдельный атом имеет дискретный энергетический спектр, т.е. каждому электрону соответствует определенная энергия (энергетический уровень). В кристаллической структуре происходит квантово-механическое взаимодействие атомов, при этом в соответствии с принципом Паули дискретные энергетические уровни электронов расщепляются в квазинепрерывные энергетические зоны. На энергетической диаграмме (рис. 1.1) представлена структура энергетических зон, где ЗП – зона проводимости; ВЗ – валентная зона; РЗ – разрешенная зона; ЗЗ – запрещенная зона.

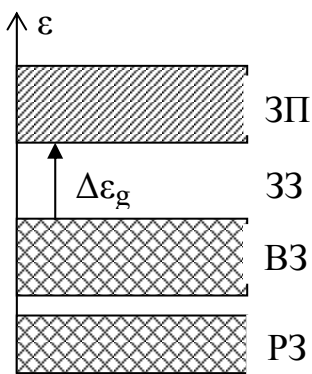


Рис. 1.1

Особенности энергетической диаграммы полупроводника такие:

- 1) при температуре абсолютного нуля ($T = 0$) все энергетические уровни валентной зоны заполнены электронами, а уровни зоны проводимости – свободны;
- 2) зона проводимости и валентная зона разделены запрещенной зоной, соответствующей

диапазону энергий, в котором не могут находиться электроны.

Основной параметр, который определяет электропроводность полупроводника, – это ширина запрещенной зоны $\Delta\epsilon_g$.

Табл. 1.1 дает возможность сравнить типовые полупроводники по ширине запрещенной зоны, выраженной в электронвольтах.

Таблица 1.1

| Полупроводник | InAs | PbS | Ge | Si | GaAs | GaP | CdS |
|---|------|------|------|------|------|------|------|
| Ширина запрещенной зоны $\Delta\epsilon_g$, эВ | 0,36 | 0,37 | 0,66 | 1,12 | 1,43 | 2,24 | 2,59 |

Собственный полупроводник – это химически чистый беспримесный полупроводник с процентным содержанием примесей ($10^{-7} \dots 10^{-8}$)%, т.е. относительное содержание примесей $\frac{\Delta N}{N_0} = 10^{-9}$,

что составляет 1 атом примеси на 10^9 атомов основного вещества. Такую степень очистки обеспечивает метод зонной плавки.

Электропроводность полупроводника определяется концентрацией подвижных (свободных) носителей заряда. В дальнейшем концентрация электронов в собственном полупроводнике обозначается как n_i и измеряется количеством электронов в единице объема вещества, т.е. $1/\text{см}^3$. Здесь индекс i указывает на тип полупроводника (intrinsic – собственный).

Так как движение электрона сопровождается увеличением его энергии, то такой электрон в энергетической диаграмме (см. рис. 1.1) должен перемещаться вверх, занимая более высокие энергетические уровни. Наличие запрещенной зоны и предельная заполненность уровней валентной зоны препятствуют этому процессу. Поэтому пороговый энергетический эффект в виде запрещенной зоны требует предварительного внешнего энергетического воздействия $\Delta \varepsilon > \Delta \varepsilon_g$, вследствие чего электрон переходит из валентной зоны на соответствующий энергетический уровень зоны проводимости (рис. 1.2).

Если таким энергетическим фактором является энергия теплового хаотического движения атомов кристаллической решетки (температура), то рассмотренный процесс называется явлением термогенерации. С переходом электрона в валентной зоне возникает вакантное энергетическое состояние, т.е. свободный энергетический уровень. Вследствие этого появляется возможность для перемещения электронов не только в зоне проводимости, но и внутри валентной зоны.

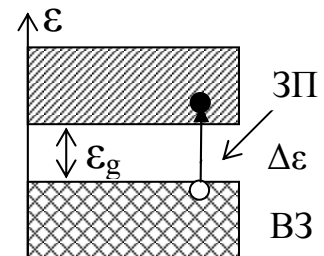


Рис. 1.2

Подвижный заряд в валентной зоне рассматривается как некоторая фиктивная частица (квазичастица), которой приписывается положительный заряд, равный по величине заряду электрона.

Дырка – это положительно заряженная квазичастица, которой соответствует не занятое электроном энергетическое состояние валентной зоны. Эффективная масса дырки больше эффективной массы электрона, т.е. она обладает большей инерционностью.

Таким образом, термогенерация – это процесс образования пар электрон-дырка под воздействием температуры. Отметим, что энергетическая область существования свободных электронов – зона про-

водимости, а дырок – валентная зона. В кристаллической структуре протекает и обратный процесс.

Рекомбинация – это процесс взаимной нейтрализации пар электрон-дырка, сопровождающийся возвратом электрона из зоны проводимости в валентную зону. При данной температуре устанавливается термодинамическое равновесие между процессами термогенерации и рекомбинации.

Очевидно, что в собственном полупроводнике обязательно выполняется условие равенства концентраций дырок и электронов:

$$p_i = n_i.$$

Это равенство является условием электрической нейтральности единицы объема собственного полупроводника.

Процесс термогенерации интерпретируется также на основе кристаллической структуры полупроводника. Собственный полупроводник представляет собой регулярную пространственно-периодическую структуру, плоская модель которой приведена на рис. 1.3. Между четырехвалентными атомами полупроводника, находящимися в узлах

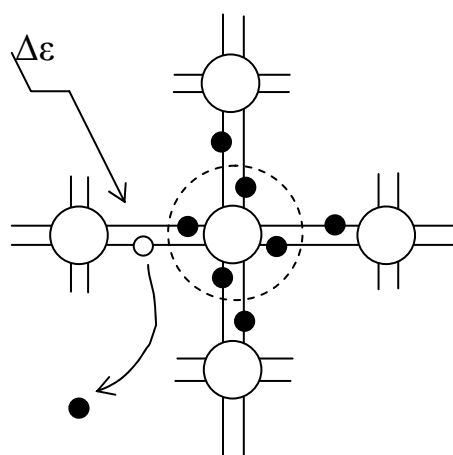


Рис. 1.3

решетки, устанавливаются ковалентные (двойные) связи. Каждая ковалентная связь создается для данного атома собственным электроном и электроном соседнего атома. Уровни энергий таких связанных электронов лежат в валентной зоне (см. рис. 1.1). Под влиянием внешнего энергетического воздействия $\Delta\varepsilon$ разрывается одна из валентных связей, электрон становится свободным, что соответствует его переходу на энергетический уровень проводимости. Разорванная валентная связь представляет собой вакантное место, т.е. дырку, которое может занять любой другой электрон.

Таким образом, с точки зрения кристаллической структуры дырка – это разорванная валентная связь в системе ковалентных связей атомов.

Таким образом, с точки зрения кристаллической структуры дырка – это разорванная валентная связь в системе ковалентных связей атомов.

Если задать в кристалле некоторое приоритетное направление с помощью внешнего электрического поля \vec{E} , то дырка (разорванная валентная связь) будет перемещаться по направлению действия поля. Связанные электроны физически движутся в противоположном направлении, занимая поочередно вакантное место в разорванной связи (рис. 1.4).

Разрыв валентных связей между атомами происходит за счет энергии хаотического теплового движения частиц. Средняя энергия тепловых колебаний рассчитывается по формуле

$$E_{cp} = KT,$$

где K - постоянная Больцмана,
 T - абсолютная температура.

При комнатной температуре ($T = 300 \text{ K}$) $E_{cp} = 0,026 \text{ эВ}$, что значи-

тельно меньше ширины запрещенной зоны. Однако случайный, хаотичный характер теплового движения в кристалле определяет некоторый процент атомов, энергии которых достаточно для разрыва валентных связей. Распределение электронов по энергетическим состояниям носит вероятностный характер и в классическом (невырожденном) полупроводнике подчиняется статистике Максвелла - Больцмана. Вероятность заполнения электроном определенного энергетического уровня зависит от энергии уровня, температуры кристалла и описывается экспоненциальным законом. В результате интегрирования закона распределения определяется концентрация электронов в зоне проводимости при данной температуре:

$$n_i = B_n T^{3/2} e^{-\frac{\Delta \varepsilon_g}{2KT}}, \quad (1.1)$$

где B_n содержит универсальные константы и численные множители.

С точки зрения температурной зависимости определяющим является экспоненциальный множитель. При возрастании температуры или уменьшении ширины запрещенной зоны происходит резкое увеличение концентрации свободных электронов, а следовательно, и проводимости полупроводника.

Несмотря на небольшие отличия по ширине запрещенной зоны для разных полупроводников (см. табл. 1.1), концентрации свободных электронов отличаются на несколько порядков. Для основных полупроводников этот факт иллюстрирует табл. 1.2, в которой данные приведены при $T = 300 \text{ K}$.

Таблица 1.2

| Полупроводник | Ge | Si | GaAs |
|--|---------------------|----------------------|------------------|
| Концентрация свободных электронов, $n_i [1/\text{см}^3]$ | $2,4 \cdot 10^{13}$ | $1,45 \cdot 10^{10}$ | $1,3 \cdot 10^7$ |

Для сравнения концентрация свободных электронов в металлах имеет порядок $10^{22} 1/\text{см}^3$.

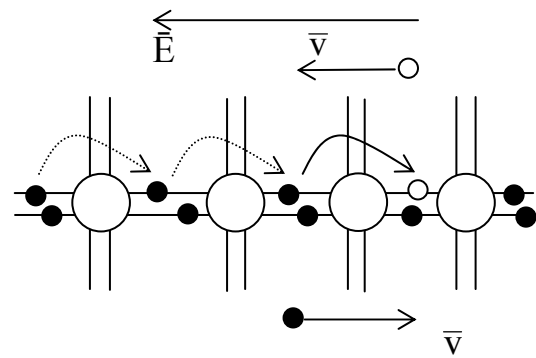


Рис. 1.4

Отметим следующие основные особенности собственных полупроводников:

- электропроводность имеет сильный температурно-зависимый характер;
- полупроводники с широкой запрещенной зоной являются более температуростабильными с точки зрения удельной проводимости;
- полупроводники с малой шириной запрещенной зоны чувствительны к электромагнитному облучению в более длинноволновой области спектра.

Например, ширина запрещенной зоны антимонида индия (InSb) $\Delta\varepsilon_g = 0,17$ эВ соответствует энергии квантов инфракрасного излучения с длиной волны $\lambda \approx 10$ мкм, что представляет интерес при реализации приборов ночного видения.

Токи в полупроводниках. Направленное движение носителей заряда в полупроводниках происходит под действием двух причин: дрейфа и диффузии.

Дрейф – направленное движение носителей заряда под действием электрического поля. В электрическом поле с напряженностью E на носитель заряда q действует электрическая сила $F = qE$. При эффективной массе носителя m он приобретает ускорение $a = F/m$. Тогда за среднее время свободного пробега величина средней скорости дрейфа составляет $U = \tau a$, или с учетом предыдущих соотношений

$$v = \tau \frac{qE}{m} = \mu E. \quad (1.2)$$

Здесь введен параметр, определяющий инерционные свойства носителя заряда, μ – подвижность. Формально подвижность определяется как средняя дрейфовая скорость в поле единичной напряженности $\mu = U / E$. Величина дрейфового тока определяется зарядом, переносимым за 1 с через поперечное сечение площадью ΔS .

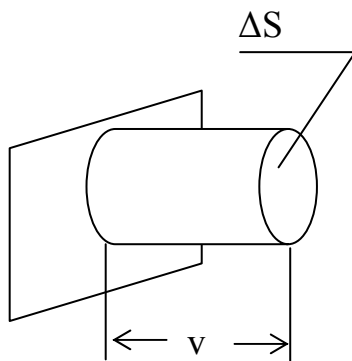


Рис. 1.5

Выделив в объеме вещества цилиндр, высота которого численно равна средней дрейфовой скорости (рис. 1.5), получим

$$\Delta I = qnU\Delta S,$$

откуда плотность дрейфового тока составит

$$j \approx \frac{\Delta I}{\Delta S} = qn\mu E.$$

Если ввести параметр удельной проводимости полупроводника

$$\sigma = qn\mu, \quad (1.3)$$

то

$$j = \sigma E,$$

что соответствует закону Ома в дифференциальной форме. Суммарная плотность дрейфового тока с учетом как электронов, так и дырок определяется по формуле

$$j_{\text{др}} = q(n\mu_n + p\mu_p) \cdot E. \quad (1.4)$$

Вообще подвижность носителей заряда является одним из существенных параметров полупроводника. Значения этого параметра во многом определяют пригодность вещества для изготовления полупроводниковых приборов.

На подвижность влияют два физических фактора: хаотические тепловые колебания атомов кристаллической решетки (рассеивание носителей заряда на тепловых колебаниях атомов) и электрические поля ионизированных примесей (рассеивание на ионах). Следовательно, величину подвижности определяют причины, препятствующие увеличению средней дрейфовой скорости: масса носителя заряда как мера его инертности; частота столкновений с атомами решетки, как следствие – уменьшение средней длины и времени свободного пробега.

Подвижность зависит от вида носителя заряда (электрона, дырки), температуры кристалла, концентрации примесей и особенностей кристаллической решетки.

Большая подвижность электрона, чем дырки, определяется меньшей его эффективной массой. Табл. 1.3 иллюстрирует зависимость подвижности от вида носителя заряда и материала полупроводника при $T = 300$ К.

Таблица 1.3

| Подвижность, $\frac{\text{см}^2}{\text{В} \cdot \text{с}}$ | Si | Ge | GaAs | InAs | InSb |
|--|------|------|------|-------|-------|
| μ_n электрона | 1600 | 3900 | 8000 | 22600 | 78000 |
| μ_p дырки | 500 | 1900 | 400 | 200 | 750 |

Из типовых полупроводниковых материалов наилучшими частотными характеристиками обладает арсенид галлия. Два последних материала приведены для сравнения.

В дрейфовом механизме переноса носителей заряда движущим фактором выступает напряженность электрического поля. В полупроводниковых приборах приходится сталкиваться со значениями напряженности от 0 до 10^9 В/м.

В слабых полях ($E < 3 \cdot 10^3$ В/см) подвижность не зависит от величины поля, так как дрейфовая скорость намного меньше средней скорости теплового движения. В этом случае наличие поля не влияет на

характер столкновений носителей заряда с узлами кристаллической решетки. При постоянной величине подвижности дрейфовая скорость пропорциональна E (1.2), что соответствует выполнению закона Ома.

В сильных полях подвижность уменьшается с ростом напряженности поля, а соотношение $U_{др}(E)$ становится нелинейным, т.е. закон Ома не выполняется. Частота и характер столкновений зависит от энергии частиц, что физически и определяет зависимость подвижности от напряженности поля. В более сильных полях подвижность обратно пропорциональна напряженности поля, что приводит к эффекту насыщения – дрейфовая скорость носителей заряда не зависит от напряженности поля. При этом скорость дрейфа соизмерима со средней тепловой скоростью носителя заряда. Взаимосвязь подвижности и напряженности поля аппроксимируется следующим выражением:

$$\mu = \mu_0 / (1 + \mu_0 E / U_s),$$

где U_s – скорость насыщения.

Повышение температуры также уменьшает подвижность вследствие увеличения рассеяния на колебаниях атомов. При этом увеличивается частота столкновений носителей заряда с атомами кристаллической решетки и уменьшается среднее время свободного пробега.

В примесных полупроводниках с увеличением концентрации примесей подвижность уменьшается, так как увеличивается рассеяние на ионизированных атомах примесей.

Диффузионный ток связан с процессом направленного движения носителей заряда под действием разности концентраций (градиента концентраций). Так как перемещение зарядов происходит в направлении убывания концентрации, то плотности токов для дырок и электронов соответственно будут такими:

$$j_p = qD_p \left(- \frac{dp}{dx} \right) = -qD_p \frac{dp}{dx}, \quad (1.5)$$

$$j_n = (-q)D_n \left(- \frac{dn}{dx} \right) = qD_n \frac{dn}{dx}.$$

Здесь D – коэффициент диффузии [$\text{см}^2/\text{с}$], значение которого определяется соотношением Эйнштейна

$$D = \varphi_T \mu,$$

где $\varphi_T = KT / q$ – тепловой потенциал при комнатной температуре $T = 300 \text{ К}$, $\varphi_T = 0,0258 \text{ В}$.

Выражения для плотностей диффузионного и дрейфового токов (1.4), (1.5) очень близки, если учесть, что напряженность поля опре-

деляется пространственной интенсивностью изменения потенциала φ , т.е.

$$j_p = q p \mu_p \left(-\frac{d\varphi}{dx} \right) = \sigma_p E,$$

$$j_n = (-q) n \mu_n \left(-\frac{d\varphi}{dx} \right) = \sigma_n E.$$

Примесные полупроводники находят основное применение в полупроводниковой электронике. Технологически контролируемое введение в полупроводник незначительного количества примесей существенно изменяет его электрофизические параметры. Во-первых, сильно возрастает концентрация свободных носителей заряда. Во-вторых, их концентрация определяется не случайным внешним фактором – температурой, как у собственного полупроводника, а степенью его легирования. В-третьих, появляется возможность получения структур с избыточной концентрацией электронов (полупроводник *n*-типа) или дырок (*p*-типа).

Чувствительность полупроводника к степени легирования показывает следующий пример. Концентрация свободных электронов в кремнии при комнатной температуре составляет $n_i \sim 10^{10} \text{ 1/см}^3$ (см. табл. 1.2). Введение в кремний примесей с относительным содержанием $\Delta N_n / N_0 = 10^{-5}$, т.е. 0,001%, что соответствует 1 атому примеси на 10^5 атомов кремния. При исходной плотности атомов кремния $N_0 \cong 10^{22} \text{ см}^{-3}$ концентрация атомов примеси составляет $\Delta N_n = 10^{22} / 10^5 = 10^{17} \text{ 1/см}^3$.

Так как каждый атом примеси поставляет один электрон проводимости, то при сравнении с исходным материалом видно, что такое незначительное легирование увеличивает концентрацию свободных электронов в $\Delta N_n / n_i = 10^7$ раз.

Для получения полупроводников с дырочным типом проводимости применяют акцепторную примесь (accеptare - принимать), а для электронного полупроводника – донорную (donare - жертвовать).

Полупроводники *n*-типа (электронный, донорный) получают, используя пятивалентную донорную примесь. При формировании ковалентных связей с атомами основного вещества один электрон примеси имеет малую энергию связи с атомом (рис. 1.6). Поэтому энергетический уровень этого электрона (примесный уровень $\Delta \epsilon_d$) находится вблизи зоны проводимости (рис. 1.7).

Под действием небольшой энергии активации $\Delta \epsilon_d = \epsilon_c - \epsilon_d \cong \underline{\underline{0,05}} \text{ эВ}$ "лишний" электрон теряет связь с атомом примеси и

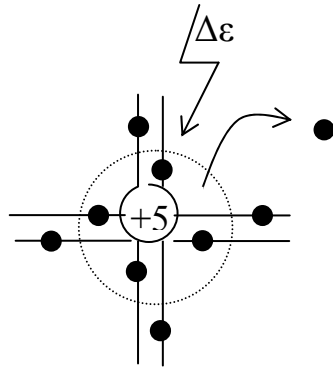


Рис. 1.6

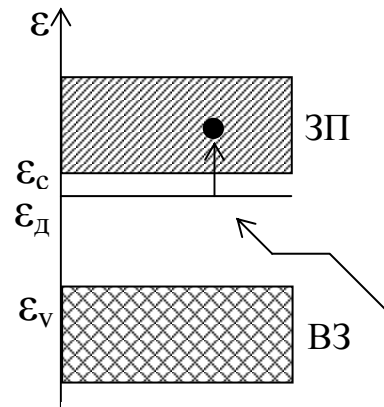


Рис. 1.7

становится свободным электроном проводимости. При этом нейтральный атом примеси превращается в положительный ион, т.е. примесь ионизируется. С энергетической точки зрения этот процесс соответствует переходу электрона с донорного уровня в зону проводимости. Концентрация свободных электронов резко возрастает и практически равна концентрации атомов примеси. Суммарная величина заряда в единице объема вещества должна равняться нулю, что соответствует статическому условию электрической нейтральности полупроводника:

$$n_n = p_n + N_D^+, \quad (1.6)$$

где переменная определяет концентрацию, а индекс – тип полупроводника. Следовательно, концентрация электронов в n -полупроводнике n_n намного превышает концентрацию дырок p_n :

$$n_n \gg p_n, \quad \text{т.е. } n_n \simeq N_D^+,$$

поэтому в рассматриваемом типе полупроводника электроны – это основные носители заряда (ОНЗ), а дырки – неосновные носители заряда (ННЗ).

Полупроводник p -типа (дырочный, акцепторный) формируется при легировании трехвалентной акцепторной примесью. При этом атом примеси забирает еще один недостающий электрон от атома основного вещества (рис. 1.8).

Нейтральный атом примеси превращается в отрицательный ион, а появившаяся разорванная валентная связь представляет собой дырку. Малая энергия активации атомов акцепторной примеси определяет положение примесного энергетического уровня ϵ_a вблизи валентной зоны (рис. 1.9). Вследствие перехода электрона из валентной зоны на примесный уровень появляется дырка в валентной зоне.

Условие электронейтральности дырочного полупроводника имеет следующий вид:

$$p_p = n_n + N_A^-, \quad (1.7)$$

где p_p - концентрация ОНЗ (дырок);
 n_p - концентрация ННЗ (электронов);
 N_A^- - концентрация акцепторных отрицательных ионов,

причем $p_p \gg n_p$, т.е. $p_p \approx N_A^-$.

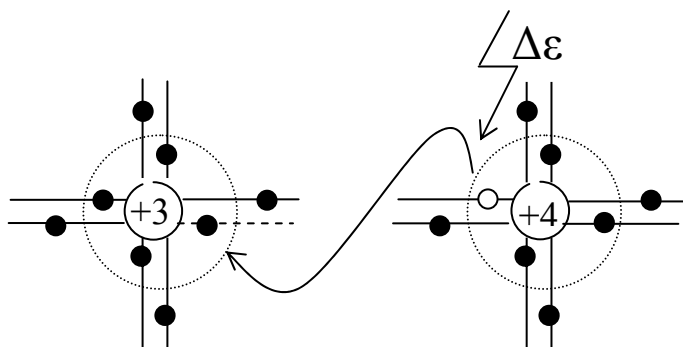


Рис. 1.8

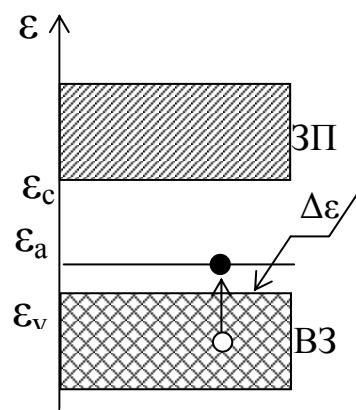


Рис. 1.9

При комнатной температуре все атомы примеси ионизируются. Дальнейший рост концентрации свободных носителей заряда возможен только в результате процесса термогенерации пар электрон-дырка по принципу беспримесного полупроводника.

Для любого типа полупроводника справедливо соотношение, являющееся следствием термодинамического равновесия между процессами генерации носителей заряда и их рекомбинацией:

$$p \cdot n = n_i^2, \quad (1.8,а)$$

т.е. произведение концентраций разноименных свободных носителей заряда есть величина постоянная, равная квадрату концентрации собственных носителей. В частном случае

$$p_p \cdot n_p = n_i^2 \quad - \quad \text{для } p\text{-полупроводника,}$$

$$n_n \cdot p_n = n_i^2 \quad - \quad \text{для } n\text{-полупроводника.} \quad (1.8,б)$$

Это равенство отражает тот факт, что увеличение концентрации носителей одного знака приводит к уменьшению концентрации другого знака за счет роста вероятности рекомбинации.

Если в полупроводнике присутствуют как донорные, так и акцепторные примеси, то возникает эффект компенсации примесей. При этом акцепторные уровни заполняются электронами, перешедшими с донорных уровней. Тогда атомы акцепторов превращаются в отрицательные ионы, а атомы доноров – в положительные, что вызывает

уменьшение подвижности оставшихся свободных зарядов проводимости. При точной компенсации количество зарядов не изменится и будет соответствовать собственному полупроводнику.

Компенсаторную роль примеси используют и для получения высокоомного "полуизолирующего" полупроводника. Так, если исходный GaAs с вредными примесями имел n -тип проводимости, то вводят глубокую акцепторную примесь (хром), которая связывает избыточные электроны. В этом случае не требуется точного соответствия концентраций исходной вредной и компенсаторной примесей (рис. 1.10). Полуизолирующий GaAs ($\rho = 10^7$ Ом·см) применяют при изготовлении подложек для полупроводниковых приборов на его основе.

1.2. Электронно-дырочный переход

1.2.1. Электронно-дырочный переход в равновесном состоянии

Электронно-дырочный переход (p - n -переход) – это переходный слой между двумя областями монокристалла полупроводника с различными типами проводимости. Эта область формируется при легировании полупроводника донорной (n -тип) и акцепторной (p -тип) примесями. Рассмотрим физические процессы при образовании перехода в симметричной структуре, т.е. $N_A = N_D$ (рис. 1.11).

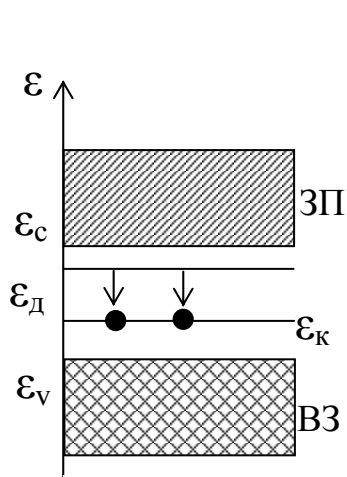


Рис. 1.10

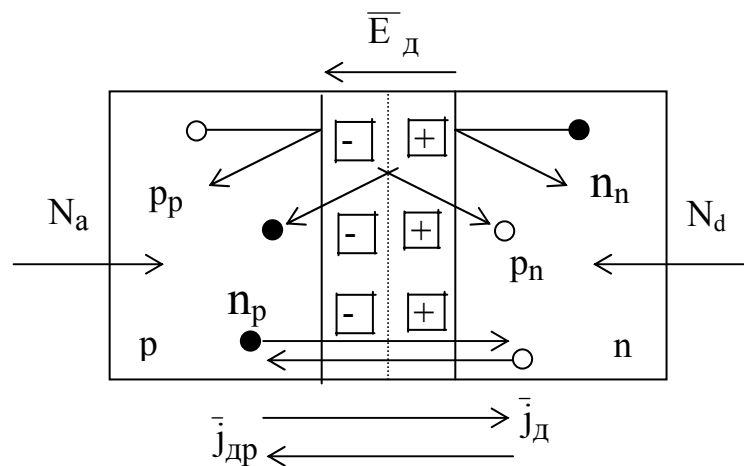


Рис. 1.11

В исходном состоянии по обе стороны от границы раздела соблюдаются условия электрической нейтральности:

$$p_p = N_A^- + n_p \text{ (} p\text{-тип),}$$

$$n_n = N_D^+ + p_n \text{ (} n\text{-тип)}$$

с повышенной концентрацией дырок $p_p \gg n_p$ и электронов $n_n \gg p_n$ в соответствующих областях. Причем объемные заряды основных носителей практически компенсируются противоположными зарядами неподвижных ионов:

$$p_p \approx N_A^-, \quad n_n \approx N_D^+.$$

При таких условиях структура не может находиться в состоянии равновесия из-за резкого различия по концентрации одноименных носителей заряда по обе стороны от границы раздела:

$$p_p \gg p_n, \quad n_n \gg n_p.$$

Поэтому возникают следующие процессы по перераспределению зарядов:

1. Наличие градиента концентраций по одноименным зарядам вызывает процессы диффузионного переноса ОНЗ через границу раздела из приграничных областей. Такому направленному переносу носителей заряда соответствует диффузионный ток \vec{J}_D , совпадающий с направлением движения дырок.

2. Вследствие ухода ОНЗ нарушаются условия электрической нейтральности в приграничных областях: $N_A^- > p_p$, $N_D^+ > n_n$. Обнажаются заряды ионов: отрицательных в p -области, положительных в n -области, т.е. возникает двойной слой ионов – область пространственного заряда.

3. Между пространственно разнесенными разноименными зарядами ионов возникает диффузионное электрическое поле \vec{E} , направленное из области n в область p . Оно обеспечивает динамическую устойчивость, т.е. равновесное состояние данной структуры.

4. Возрастающее поле приводит к уменьшению диффузионного переноса основных носителей заряда. Так как вектор скорости диффузионного переноса направлен навстречу электрической силе $\vec{F} = q\vec{E}$ (см. верхнюю часть рис. 1.11), то диффузионное поле перехода является потенциальным барьером для основных носителей заряда.

5. Диффузионное поле является ускоряющим для неосновных носителей заряда, обеспечивая их дрейф через переход (см. нижнюю часть рис. 1.11), иначе говоря, это поле представляет собой потенциальную яму для ННЗ. Дрейфовому механизму переноса этих зарядов

соответствует дрейфовый ток \bar{j}_{dp} , совпадающий с направлением движения дырок.

6. Процесс формирования электрического перехода завершается при установлении динамического равновесия между величинами дрейфового и диффузионного токов.

В результате контактных явлений образуется *p-n*-переход, состоящий из ионов примеси и нейтральных атомов основного вещества. Эта область представляет собой обедненный слой с пониженной концентрацией подвижных носителей заряда и, как следствие, имеющий повышенное омическое сопротивление. Внутри перехода существует электрическое поле, связанное с величиной контактной разности потенциалов:

$$\varphi_k = \varphi_T \ln \frac{n_n p_p}{n_i^2},$$

или с учетом концентрации собственных носителей заряда (1.8,б)

$$\varphi_k = \varphi_T \ln \frac{p_p}{p_n} = \varphi_T \ln \frac{n_n}{n_p}. \quad (1.9)$$

Именно соотношение между концентрацией одноименных носителей заряда в разных областях проводимости определяет величину контактной разности потенциалов.

В зависимости от материала исходного полупроводника величина контактной разности потенциалов лежит в следующих пределах:

$$\varphi_k = 0,25 \dots 0,45 \text{ В} \quad - \quad \text{для Ge,}$$

$$\varphi_k = 0,5 \dots 0,9 \text{ В} \quad - \quad \text{для Si.}$$

Параметр φ_k увеличивается в полупроводниках с большей шириной запрещенной зоны, с увеличением концентрации легирующих примесей и уменьшается с ростом температуры.

Действительно, первые два фактора увеличивают градиент концентрации носителей, при этом возрастает диффузионный ток и компенсирующий его дрейфовый ток. Следовательно, равновесное состояние наступает при большей напряженности диффузионного поля, что соответствует большей величине контактной разности потенциалов. Увеличение температуры наоборот приводит к увеличению концентрации ННЗ, уменьшению градиента концентраций и, как следствие, к уменьшению φ_k .

Энергетическая диаграмма p - n -перехода приведена на рис. 1.12. Следует отметить, что "дно" зоны проводимости в электронной и дырочной областях (ϵ_{cn} , ϵ_{cp}), как и "потолок" валентных зон (ϵ_{vn} , ϵ_{vp}), соответствует разным значениям энергетических уровней. Однако уровень Ферми ϵ_F в единой кристаллической структуре не зависит от типа проводимости области и соответствует вероятности его заполнения, равной $1/2$.

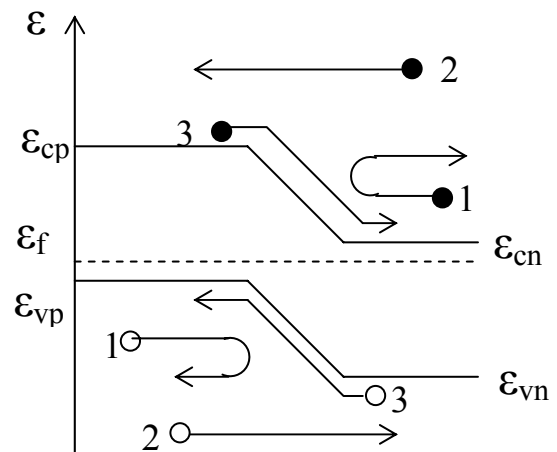


Рис. 1.12

Энергетические зоны для n -области должны быть расположены ниже, чем в p -области, так как \vec{E} внутри перехода направлено из n -области в p -область (см. рис. 1.11). Например, из двух основных носителей зарядов – электронов из зоны проводимости в n -области, энергии электрона (1) недостаточно для преодоления энергетического барьера $\Delta\epsilon_k = q\phi_k$. Только небольшая часть высокоэнергетичных электронов (2) преодолевает энергетический барьер. В то же время неосновной носитель заряда (электрон (3) из зоны проводимости p -области) падает в потенциальную яму, т.е. переносится ускоряющим полем перехода в n -область.

Аналогичным образом ведут себя дырки как основные (1,2) или неосновные (3) носители зарядов. Следует только учитывать тот факт, что дыркам с большей энергией соответствуют более низкие энергетические уровни по оси энергий.

1.2.2. Основное свойство p - n -перехода – свойство односторонней проводимости

Подключение к p - n -переходу внешнего источника напряжения нарушает динамическое равновесие между диффузионной и дрейфовой компонентами токов. В зависимости от полярности приложенного напряжения переход получает прямое смещение (подача прямого напряжения) или обратное. Смещение является прямым, если выполняется правило своего знака: к p -области (positiv) приложен положительный потенциал источника, а к n -области (negativ) – отрицательный потенциал. При обратном смещении нарушается соответствие потенциалов источника типу проводимости областей p - n -структуры, причем изменение полярности внешнего источника приводит к весьма существенному отличию прямого и обратного токов.

Прямое смещение p - n -перехода соответствует рис. 1.13. Внешний источник напряжения U создает внутри структуры электрическое

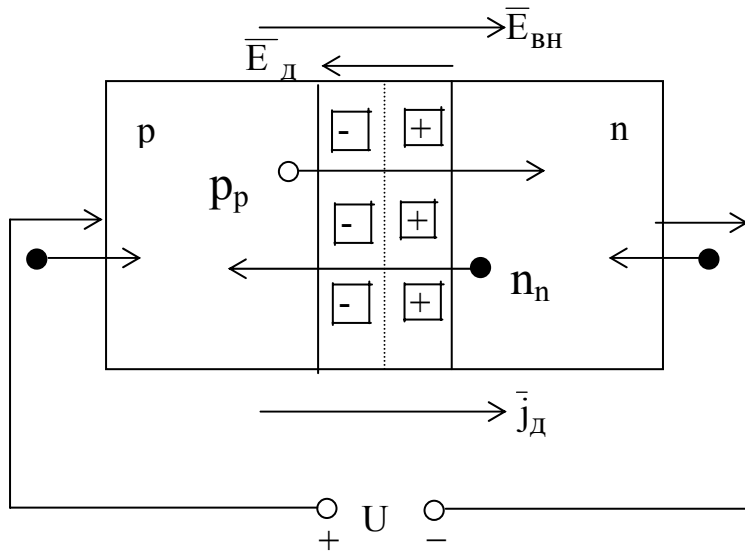


Рис. 1.13

поле $\vec{E}_{вн}$, направленное навстречу диффузионному полю перехода \vec{E}_D . Суммарное поле $E = E_D - E_{вн}$ внутри перехода уменьшается, т.е. падает величина потенциального барьера $\varphi_B = \varphi_K - U$. При этом уменьшается ширина перехода, которая в равновесном состоянии равнялась d . С уменьшением потенциального барьера большая часть

ОНЗ (p_p, n_n) преодолевает этот барьер, что приводит к резкому увеличению диффузионной компоненты тока $\vec{J}_{диф}$ (направление тока совпадает с направлением диффузии дырок).

В результате взаимной диффузии электронов и дырок нарушаются электрические нейтральности областей p и n . Область p приобретает избыточный заряд отрицательных электронов, а область n – положительных дырок. При нарушении электронейтральности действует динамический закон электронейтральности (ЗЭН). Всякое нарушение электрической нейтральности каких-либо областей полупроводника вызывает такие потоки электрических зарядов, которые стремятся восстановить нарушенную нейтральность. В результате действия этого закона избыточные электроны из p -области выходят из нее, дрейфуя к положительному потенциалу источника. В то же время дополнительные электроны от минуса источника поступают в n -область, компенсируя избыточный заряд дырок. Токи во внешних выводах направлены навстречу потокам электронов. Таким образом, ток во внешних выводах направлен от высокого потенциала источника (плюса) к низкому. Очевидно, что из-за большой концентрации ОНЗ внутри структуры (при понижении высоты потенциального барьера) возникает большая величина прямого тока во внешних выводах.

Итак, в прямосмещенном p - n -переходе протекает достаточно большой прямой ток, созданный основными носителями зарядов, который имеет диффузионный характер.

При увеличении внешнего напряжения структура приобретает все более однородный характер (так как ширина перехода уменьшается), приближаясь к характеристикам полупроводникового резистора. При нарушении энергетического равновесия теплового обмена с внешней средой, с увеличением прямого тока возможно тепловое разрушение структуры.

Обратное смещение перехода соответствует нарушению правила своего знака (рис. 1.14). В этом режиме приоритетные условия создаются для дрейфа ННЗ,

а диффузионный ток ОНЗ практически отсутствует. Малая величина обратного тока обусловлена небольшой концентрацией ННЗ, которые и создают этот ток. Поле, созданное внешним источником $\vec{E}_{вн}$, совпадает по направлению с полем перехода $\vec{E}_д$, что приводит к увели-

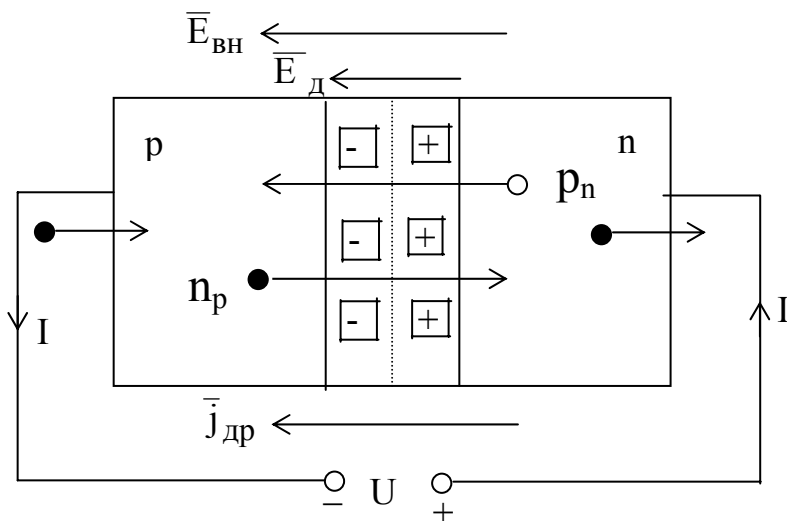


Рис. 1.14

чению суммарного поля $E = E_д + E_{вн}$, а следовательно, и высоты потенциального барьера для ОНЗ, поэтому диффузионный ток стремится к нулю. Дрейф ННЗ под действием ускоряющего суммарного поля перехода создает обратный ток перехода $\vec{j}_{др}$. Под действием ЗЭН появляется ток во внешних выводах I , направленный по-прежнему от высокого потенциала к низкому.

Следовательно, обратный ток перехода имеет дрейфовый характер, является током ННЗ и тепловым током, так как концентрация ННЗ, а значит, и ток зависят от температуры.

Характерная особенность обратного тока – его независимость от величины обратного напряжения, что связано с малой концентрацией ННЗ. Поэтому этот ток называется также током насыщения.

Таким образом, свойство односторонней проводимости перехода проявляется в большой величине прямого тока и малом значении обратного. Это связано с наличием двух типов зарядов: ОНЗ (с высокой концентрацией) и ННЗ (с низкой концентрацией), а также фактора, регулирующего потоки этих зарядов, – диффузионного поля перехода. Причем $E_д$ действует ускоряюще на ННЗ и оказывает тормозящее

действие на ОНЗ. Усиливает или ослабляет этот регулирующий фактор соответствующая полярность внешнего источника смещения.

1.2.3. Вольт-амперная характеристика идеализированного *p-n*-перехода

Электронно-дырочный переход является существенно нелинейным элементом. Если для резистора ток изменяется пропорционально приложенному напряжению, т.е. подчиняется линейному закону Ома, то рассматриваемая структура имеет нелинейную вольт-амперную характеристику (ВАХ).

Идеализированный переход является упрощенной моделью реального *p-n*-перехода. Условия идеализации следующие:

- рассматривается переход с плоскими границами, заряды движутся в направлении, перпендикулярном границам, не учитываются краевые эффекты;
- сопротивления областей, прилегающих к переходу, пренебрежимо малы, т.е. все внешнее напряжение приложено непосредственно к переходу;

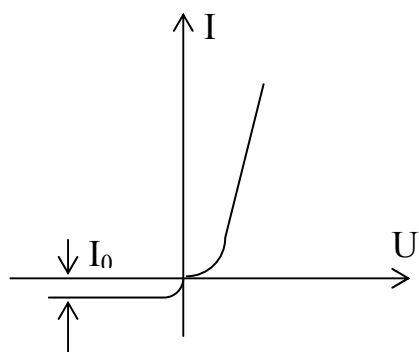


Рис. 1.15

- процессы генерации и рекомбинации внутри области объемного заряда отсутствуют;

- рассматривается статический режим работы при малой интенсивности диффузионного переноса ОНЗ.

Вольт-амперная характеристика определяет зависимость между током, протекающим через переход, и напряжением, приложенным к нему: $I = f(U)$. Эта характеристика показана на рис. 1.15. Прямая

ветвь соответствует положительным напряжениям (прямое смещение перехода) и формируется диффузионным током ОНЗ. Обратная ветвь определяет ток насыщения I_0 , который является дрейфовым током ННЗ. Прямая и обратная ветви приведены в различных масштабах токов, что вызвало излом характеристики в начале координат.

Аналитическое представление ВАХ соответствует следующей экспоненциальной зависимости:

$$I = I_0 \left(e^{\frac{U}{\Phi T}} - 1 \right), \quad (1.10)$$

где I_0 - обратный ток перехода,

φ_T - тепловой потенциал.

Тепловая константа φ_T определяет динамику изменений как прямого, так и обратного токов. При комнатной температуре ($T = 300$ К, $\varphi_T = 26$ мВ) и превышении обратным ($U < 0$) напряжением величины $|U| > 3\varphi_T$, т.е. $|U| \gtrsim 0,1$ В, обратный ток достигает своего значения насыщения $I \simeq -I_0$ (1.10). Изменение прямого напряжения всего лишь на величину $\Delta U = 2,3\varphi_T \simeq 60$ мВ вызывает увеличение тока в 10 раз. Рабочие токи прямой ветви ВАХ соответствуют $I > (10^3 \div 10^4)I_0$.

Для нелинейной ВАХ вводятся два понятия сопротивления. Пусть к p - n -переходу приложено напряжение $U(t) = U_\rho + u(t)$, содержащее как постоянную составляющую U_ρ , так и малую переменную составляющую $u(t)$. Тогда статическое сопротивление

$$R_{cm} = \frac{U_\rho}{I_\rho} \quad (1.11)$$

определяет сопротивление перехода постоянному току, а динамическое (дифференциальное)

$$R_d = \frac{dU}{dI} \approx \frac{\Delta U}{\Delta I} \quad (1.12)$$

соответствует сопротивлению по переменной составляющей. Для аналитического расчета R_d нужно взять производную от ВАХ (1.10):

$$\frac{1}{R_d} = \frac{dI}{dU} = \frac{I_0}{\varphi_T} e^{\frac{U}{\varphi_T}} = \frac{I + I_0}{\varphi_T}.$$

Пренебрегая малой величиной обратного тока, получаем расчетное соотношение для динамического сопротивления

$$R_d \approx \frac{\varphi_T}{I} \quad (1.13)$$

или при комнатной температуре

$$R_d \simeq \frac{26}{I(\text{mA})} (\text{Ом}).$$

При приближенном графическом расчете динамического сопротивления в окрестности рабочей точки строится характеристический треугольник, катеты которого определяют нужные для расчета (1.12) приращения ΔU и ΔI (рис. 1.16). Очевидно, что значения сопротивлений зависят от положения рабочей точки, а также уменьшаются с увеличением прямого напряжения. На ВАХ перехода влияет также выбор материала, из которого выполнена данная структура (рис. 1.17). Материал с более широкой запрещенной зоной $\Delta \epsilon_{g2}$ обеспечивает большую высоту потенциального барьера и, как следствие, меньшее значение прямого тока $I_2 < I_1$ при одной и той же величине напряжения.

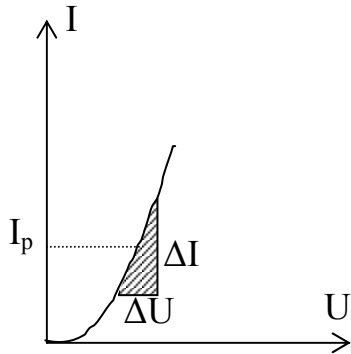


Рис. 1.16

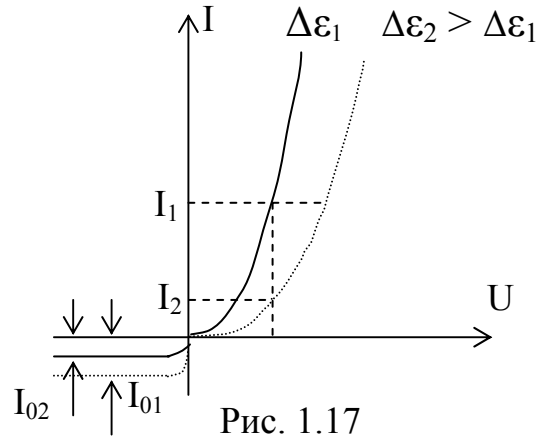


Рис. 1.17

Обратная ветвь характеристики для такого материала также соответствует меньшему значению тока насыщения, так как это ток неосновных носителей заряда. Их концентрация уменьшается с увеличением ширины запрещенной зоны. Так, учитывая в выражении для концентрации собственных носителей заряда (1.1) только наиболее значимый экспоненциальный множитель, можно отметить следующую зависимость:

$$I_0 \sim n_i^2 \sim e^{-\frac{\Delta \epsilon_g}{kT}} \quad (1.14)$$

Тогда при сравнении обратных токов Ge- и Si-переходов имеем

$$\frac{I_0(\text{Ge})}{I_0(\text{Si})} = e^{\frac{\Delta \epsilon_g(\text{Si}) - \Delta \epsilon_g(\text{Ge})}{kT}} \approx 10^6.$$

Численное значение приведено при комнатной температуре.

На ВАХ перехода также влияет температура (рис. 1.18). Смещение характеристики при изменении температуры оценивают числовыми параметрами

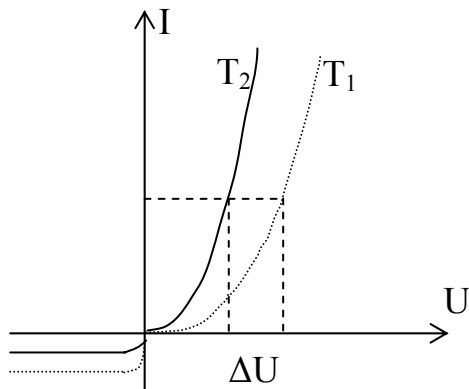


Рис. 1.18

рами.

Так, обратная ветвь при повышении температуры смещается в область больших токов, что связано с увеличением концентрации ННЗ. Числовой параметр, определяющий взаимосвязь изменений обратного тока и температуры, называется температурой удвоения T_y . Для обоснования этого параметра воспользуемся вышеприведенным выражением (1.14) и сравним обратные токи для двух значений температуры T_0 и $T = T_0 + \Delta T$:

$$\frac{I_0(T)}{I_0(T_0)} = e^{\Delta \varepsilon_g \frac{T - T_0}{K T T_0}}.$$

Учитывая, что $T - T_0 = \Delta T$, введем приближенную константу

$$a = \frac{\Delta \varepsilon_g}{K(T_0 + \Delta T)} \approx \frac{\Delta \varepsilon_g}{K T_0^2}$$

и перейдем к новому основанию степени $e^{a \Delta T} = 2^{\frac{\Delta T}{T_y}}$, где введен параметр $T_y = (\ln 2)/a$. Тогда соотношение между обратными токами определит выражение

$$I_0(T) = I_0(T_0) \cdot 2^{\frac{\Delta T}{T_y}}. \quad (1.15)$$

Следовательно, температура удвоения соответствует такому изменению температуры, при котором обратный ток увеличивается вдвое. Зная величину T_y , при известном токе и номинальной температуре $I_0(T_0)$ можно находить новое значение обратного тока для заданного изменения температуры. Значение температуры удвоения для типовых материалов составляет:

для Ge - $T_y = 8^{\circ} \text{C}$, для Si - $T_y = 5^{\circ} \text{C}$.

Прямая ветвь диода с увеличением температуры смещается в область меньших напряжений (см. рис. 1.18), т.е. падение напряжения на переходе уменьшается при заданной величине тока. Физически это связано с изменением распределения носителей заряда по энергиям, т.е. увеличивается плотность заполнения электронами более высоких энергетических уровней в зоне проводимости (см. рис. 1.12). Поэтому большая часть электронов имеет возможность преодолеть потенциальный барьер, увеличивая тем самым прямой диффузионный ток.

Зависимость прямого тока от температуры определяет уравнение

$$\text{ВАХ} \quad I = I_0(T) \left(e^{\frac{U}{\Phi T}} - 1 \right),$$

или, учитывая температурную зависимость тока насыщения (1.14), имеем

$$I = A \exp\left(-\frac{\Delta\varepsilon_g - qU}{KT}\right).$$

Логарифмируя это выражение, получим

$$U = \frac{\Delta\varepsilon_g}{q} - \frac{KT}{q} \ln \frac{A}{I},$$

т.е. при фиксированном токе напряжение на переходе уменьшается с ростом температуры. Температурный коэффициент напряжения (ТКН) определяет, на сколько уменьшается прямое напряжение на переходе при увеличении температуры на 1 градус:

$$\text{ТКН} = \frac{dU}{dT} = -\frac{k}{q} \ln \frac{A}{I}. \quad (1.16)$$

Зная величину ТКН ≈ -2 мВ/град, можно вычислить падение напряжения на переходе для заданного изменения температуры $\Delta T = T - T_0$:

$$U(T) = U(T_0) + \text{ТКН} \cdot \Delta T. \quad (1.17)$$

1.2.4. ВАХ реальных *p-n*-переходов

Реальные переходы являются несимметричными, т.е. дырочная и электронная области имеют разную степень легирования: слаболегированная область – база, высоколегированная – эмиттер.

Так, на рис. 1.19 при условии $N_A \gg N_D$ *p*-область представляет собой эмиттер, а *n*-область – базу. Вследствие этого прямой диффузионный ток \bar{j}_p

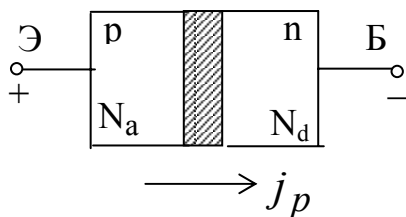


Рис. 1.19

определяется в основном дырочной компонентой. Геометрически переход также несимметричен. На рис. 1.19 пунктиром показана «нейтральная» граница раздела.

Наличие более высокоомной базовой области приводит к тому, что внешнее напряжение действует не непосредственно на *p-n*-переход, а перераспределяется между базой и самим переходом. Тогда в масштабе реальных рабочих токов прямая ветвь ВАХ имеет ярко выраженный пороговый эффект (рис. 1.20). Начальный участок прямой ветви, соответствующий малым токам, в данном масштабе не виден на графике, причем материалу с более широкой запрещенной зоной соответствует большее значение порогового напряжения. Если для Ge-переходов $U_{п1} \approx 0,3$ В, то для Si - $U_{п1} \approx 0,5$ В, а для GaAs - $U_{п2} > 0,7$ В.

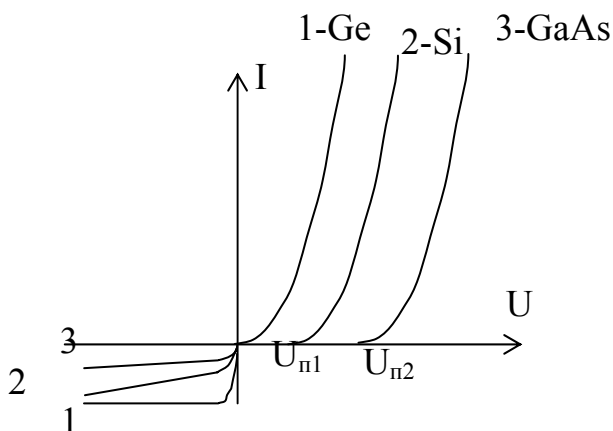


Рис. 1.20

Обратная ветвь для материалов с большей шириной запрещенной зоны (Si, GaAs) не имеет выраженного участка насыщения (см. рис. 1.20). Это связано с наличием дополнительной компоненты обратного тока в таких переходах – тока термогенерации I_e , т.е.

$$I_{обр} = I_0 + I_e.$$

Эта компонента тока образуется за счет термогенерации пар электрон-дырка внутри перехода, причем при увеличении обратного напряжения растет ширина перехода и процесс термогенерации происходит более интенсивно. Например, для кремниевого перехода компоненты обратных токов отличаются на несколько порядков. В связи с этим в реальных переходах соотношения между обратными токами для германия и кремния существенно ниже, чем это было в случае идеализированных переходов:

$$I_0(\text{Ge})/I_0(\text{Si}) \approx 10^2.$$

1.2.5. Свойства *p-n*-перехода

1. Основное свойство – односторонняя проводимость, обосновано выше и заключается в существенном отличии в величинах прямого и обратного токов. Столь же сильно отличаются статические и динамические сопротивления перехода в зависимости от направления смещения.

2. Инжекция – это явление одностороннего внедрения зарядов через переход из одной области полупроводника, где они являются основными, в другую область путем диффузии при условии снижения

высоты потенциального барьера. Иначе говоря, инжекция – это односторонняя диффузия.

Условия реализации этого явления такие:

1) наличие несимметричного перехода, так как $N_A \gg N_D$, то $p_p \gg n_n$;

2) подача прямого смещения на переход с целью снижения высоты потенциального барьера для основных носителей заряда.

Следствие инжекции – появление в базовой области избыточных (неравновесных) носителей заряда (рис. 1.21). В этом случае нарушается равенство концентраций, соответствующее условию термодина-

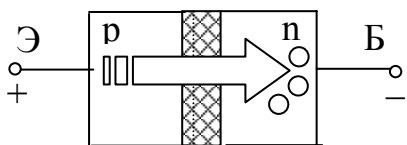


Рис. 1.21

мического равновесия $n_n \cdot p_n \neq n_i^2$.

Качество инжекции, т.е. степень односторонности диффузии, оценивают коэффициентом инжекции

$$\gamma_p = \frac{j_p}{j_p + j_n}.$$

Эта величина составляет $\gamma = 0,95 \dots 0,99$.

3. Рекомбинация – процесс взаимной нейтрализации пар электрон-дырка. В основном это явление протекает в базовой области. Динамика процесса характеризуется временем жизни T неосновных неравновесных носителей заряда (временем рекомбинации). Так, если в результате импульсной инжекции концентрация избыточных зарядов составила Δp_0 , то в результате рекомбинации она уменьшается по закону экспоненты

$$\Delta p(t) = \Delta p_0 e^{-\frac{t}{\tau}}. \quad (1.18)$$

Следовательно, T – это время, в течение которого избыточная концентрация уменьшается в e раз. С этим параметром рекомбинации связан другой – диффузионная длина

$$L = \sqrt{D\tau}. \quad (1.19)$$

Инжектированные в базу заряды в результате диффузии распространяются вглубь структуры, при этом часть их рекомбинирует. Поэтому диффузионная длина – это расстояние, на котором концентрация избыточных зарядов уменьшается в e раз. Иначе говоря, L – это расстояние, проходимое зарядами за среднее время их существования.

В процессе рекомбинации должны выполняться законы сохранения энергии и сохранения импульса. Так как при рекомбинации энер-

гия электрона уменьшается, то она выделяется в виде кванта тепловой энергии (фонона), т.е. передается кристаллической решетке или другим электронам и дыркам (Оже-рекомбинация). Такая рекомбинация называется безизлучательной. Если энергия выделяется в виде кванта электромагнитной энергии (фотона), то такая рекомбинация является излучательной.

С энергетической точки зрения возможны два варианта процесса рекомбинации: непосредственной рекомбинации электрона и дырки, которая сопровождается переходом электрона из зоны проводимости в валентную зону, или двухступенчатой рекомбинации через энергетический уровень ловушки ϵ_l (рис. 1.22). Роль ловушек (центров рекомбинации) выполняют атомы некоторых элементов (золото, платина, медь, серебро). Так легирование кремния золотом с относительным содержанием $10^{-6}\%$ уменьшает время жизни избыточных носителей в тысячу раз. Атом рекомбинационной примеси удерживает захваченный электрон (1) до появления в его окрестности дырки, что приводит к их окончательной рекомбинации (2).

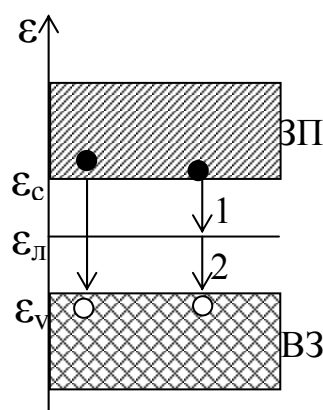


Рис. 1.22

В германии и кремнии основным механизмом рекомбинации является двухступенчатая рекомбинация через ловушки с выделением квантов тепловой энергии. Межзонная рекомбинация маловероятна из-за несоизмеримости средней энергии фотона (KT) с шириной запрещенной зоны. Для осуществления Оже-рекомбинации встреча трех подвижных носителей заряда также маловероятна.

4. Излучательная рекомбинация – процесс выделения квантов электромагнитной энергии вследствие рекомбинации инжектированных в базу носителей заряда.

Условия реализации этого процесса заключаются в следующем:

1) прямое смещение на переходе, которое обеспечивает высокий уровень инжекции, т.е. большую концентрацию избыточных зарядов в базовой области структуры;

2) соответствие выделяемой энергии при переходе электрона с высокого на более низкий энергетический уровень $\Delta\epsilon = \epsilon_2 - \epsilon_1$ (эВ) необходимой частоте электромагнитного излучения $\nu = \Delta\epsilon/h$ или соответствующей длине волны

$$\lambda = \frac{1,24}{\Delta\epsilon} [\text{мкм}]; \quad (1.20)$$

3) соответствие суммарного импульса рекомбинирующего электрона и дырки импульсу излучаемого фотона $h\nu/c$. Так как импульс фотона на несколько порядков меньше импульса электрона или дырки, то межзонная излучательная рекомбинация возможна, если рекомбинирующие заряды имеют приблизительно одинаковые и противоположно направленные импульсы. Это условие выполняется для арсенида галлия (GaAs), фосфида галлия (GaP), фосфида индия (InP).

Эффективность излучения характеризуется квантовым выходом – отношением числа излученных фотонов к числу рекомбинированных пар носителей заряда.

5. Диффузионная емкость – это емкость, образованная объемными зарядами инжектированных в базу неосновных и неравновесных основных носителей. Появление диффузионной емкости связано с процессом накопления избыточных носителей заряда в базовой области структуры (рис. 1.23). Прямое смещение перехода обеспечивает

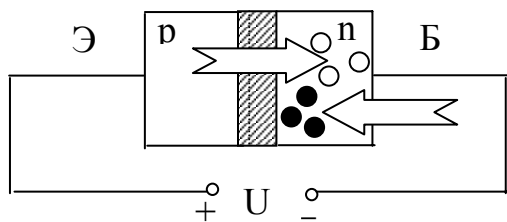


Рис. 1.23

большую величину инжекционного тока (тока диффузии), который нарушает электрическую нейтральность базы (избыточный заряд дырок). В результате действия ЗЭН восстанавливается нейтральность этой области за счет эквивалентного потока дополнительных

электронов от внешнего источника смещения. При конечном времени жизни неравновесных носителей устанавливается динамическое равновесие между процессами инжекции и рекомбинации в базовой области, что поддерживает наличие избыточного заряда. Таким образом, условиями реализации диффузионной емкости являются: наличие прямого смещения (инжекционного тока) и конечного времени жизни неравновесных носителей заряда.

Количественно величина емкости определяется приращением накопленного заряда при изменении напряжения:

$$C_{\text{д}} = \frac{dQ}{dU}. \quad (1.21)$$

Этот процесс описывается следующей логической цепочкой:

$$\Delta U \uparrow \rightarrow \Delta j_{\text{диф}} \uparrow \rightarrow \Delta p_n \uparrow (\Delta Q \uparrow).$$

Для получения расчетной формулы выражение (1.21) записывается в эквивалентной форме:

$$C_{\text{д}} = \frac{dQ}{dI} \cdot \frac{dI}{dU}.$$

Так как $dl/dU = 1/R_d$ определяется через динамическое сопротивление, а накопленный заряд пропорционален времени жизни и величине прямого инжекционного тока ($Q = It$, и $dQ/dl = \tau$), то

$$C_d = \frac{\tau}{R_d} = \frac{\tau I}{\varphi_T} \quad (1.22)$$

Заметим, что произведение $C_d R_d = \tau$ всегда имеет размерность времени. При более точном расчете значение емкости, определяемой по формуле (1.22), следует уменьшить вдвое.

Физически емкость проявляется в инерционности процесса установления обратного тока через переход при действии на него переменного напряжения (рис. 1.24). Появление выброса обратного тока во время отрицательного полупериода вызвано экстракцией накопленных в базе дырок обратно через p - n -переход (см. рис. 1.23). При повышенных частотах это приводит к ухудшению свойства односторонней проводимости. Для уменьшения диффузионной емкости базу легируют рекомбинационными примесями, которые снижают время жизни неосновных носителей.

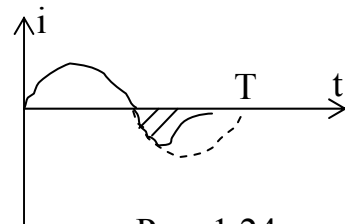


Рис. 1.24

6. Экстракция – это процесс извлечения неравновесных носителей заряда из области, где они являются неосновными, через p - n -переход при помощи ускоряющего диффузионного поля (рис. 1.25). Условия реализации этого явления: наличие в базовой области избыточных носителей заряда и обратное смещение перехода, которое обеспечи-

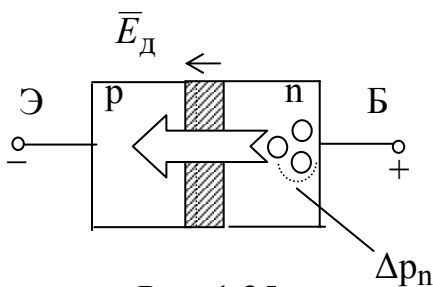


Рис. 1.25

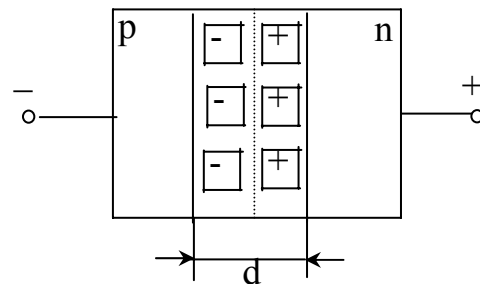


Рис. 1.26

вает большое значение напряженности диффузионного поля.

7. Барьерная (зарядная) емкость – это емкость, образованная ионами примеси внутри самого p - n -перехода (рис. 1.26). Емкостный эффект проявляется в изменении объемного заряда ионов при изменении обратного напряжения в соответствии со следующей логической цепочкой:

$$\Delta U \uparrow \rightarrow \Delta N_D^+, \Delta N_A^- \uparrow \rightarrow \Delta d \uparrow \rightarrow \Delta Q \uparrow.$$

Наличие емкости приводит к появлению дополнительной компоненты – емкостного тока

$$I_c(t) = \frac{dQ}{dt} = \frac{dQ}{dU} \cdot \frac{dU}{dt} = C_B \frac{dU}{dt}.$$

Если провести аналогию с плоским конденсатором, обладающим емкостью

$$C = \varepsilon\varepsilon_0 \frac{S}{d},$$

то расстояние между обкладками равно ширине p - n -перехода, причем ширина перехода зависит от величины обратного напряжения

$$d = a \left(\varphi_k + |U_{\text{обр}}| \right)^m,$$

где a – размерная константа, а m – зависит от закона распределения примесей при формировании перехода. В таком случае величина барьерной емкости определяется нелинейной зависимостью от обратного напряжения:

$$C_B = \frac{C_0}{\left(1 + \frac{|U_{\text{обр}}|}{\varphi_k} \right)^m}, \quad (1.23)$$

где C_0 – номинальная емкость при $U = 0$, а $m = 0,5 - 0,3$ (0,5 – резкий переход, 0,3 – плавный).

Очевидно, что увеличение обратного напряжения приводит к расширению перехода, а следовательно, к уменьшению емкости. В то же время повышение концентрации примесей увеличивает плотность объемного заряда ионов, а значит, и барьерную емкость.

8. Пробой – это явление резкого увеличения обратного тока перехода при превышении обратным напряжением некоторого порогового значения – напряжения пробоя.

Этот процесс сопровождается резким уменьшением динамического сопротивления, при этом напряжение практически не изменяется. Различают следующие основные виды пробоев: электрический (обратимый) и тепловой (необратимый). В свою очередь электрический пробой подразделяют в зависимости от механизма пробоя на лавинный и туннельный.

Туннельный пробой обусловлен эффектом квантово-механического просачивания электронов сквозь потенциальный барьер без изменения их энергии. В основном энергия электронов меньше высоты потенциального барьера, но вследствие волновых свойств

этих частиц они, с некоторой вероятностью, могут оказаться по другую сторону барьера. Механизм туннелирования показан на рис. 1.27, где приведена энергетическая диаграмма перехода, в котором возможен данный тип пробоя. Электроны из валентной зоны p -области туннелируют через потенциальный барьер на свободные уровни в зону проводимости n -области. Этот поток электронов создает большую величину обратного туннельного тока в отличие от небольшого обратного тока неосновных носителей заряда.

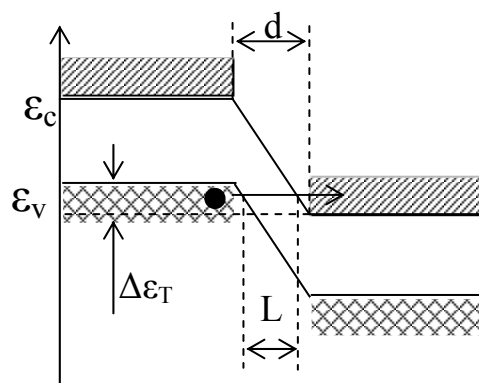


Рис. 1.27

Условия реализации туннельного пробоя следующие: 1) узкий обратнорасширенный p - n -переход, чтобы длина волны электрона была соизмерима с шириной перехода; 2) наличие свободных разрешенных энергетических уровней в зоне туннелирования $\Delta\varepsilon_T$ (взаимное соответствие энергетических уровней в обеих областях структуры, так как туннелирование происходит без изменения энергии); 3) напряженность электрического поля внутри перехода должна превосходить некоторое критическое значение.

Ширина области объемного заряда уменьшается с ростом концентрации примесей, а напряженность поля в переходе возрастает, поэтому туннельный пробой происходит в структурах с сильно легированными p - и n -областями. С ростом напряженности поля увеличивается также наклон энергетических зон (см. рис. 1.27). В свою очередь ширина потенциального барьера L определяется электрическим полем, т.е. наклоном энергетических уровней. Следовательно, условия туннелирования возникают только при определенной напряженности поля, т.е. определенной величине обратного напряжения.

Лавинный пробой определяется эффектом лавинного размножения носителей заряда в широком p - n -переходе в результате ударной ионизации. Если ширина области объемного заряда превышает длину свободного пробега, то неосновные носители заряда, попадая внутрь перехода, ускоряются, набирают достаточную энергию для ионизации атома. В результате ударного взаимодействия с нейтральным атомом происходит разрыв валентных связей с образованием дополнительной пары электрон-дырка. Новые носители заряда участвуют в дальнейшей ионизации, т.е. процесс носит лавинообразный характер (рис. 1.28). Для материалов с большей шириной запре-

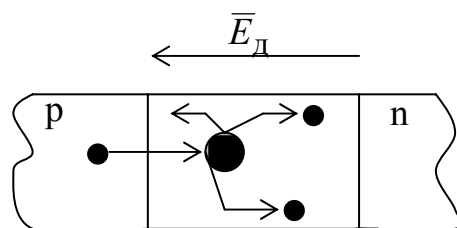


Рис. 1.28

щенной зоны требуется большая энергия ионизации, что соответствует большему значению напряженности пробоя. Таким образом, условиями лавинного пробоя являются: 1) обратное смещение перехода с величиной напряжения, превышающего некоторое критическое значение (энергия носителя заряда, приобретенная на длине свободного пробега, должна превышать ширину запрещенной зоны – энергию ионизации); 2) ширина перехода должна значительно превышать длину свободного пробега для создания условий лавинного размножения носителей заряда.

На рис. 1.29 приведены вольт-амперные характеристики пробоев, соответственно лавинного – 1 и туннельного – 2, для различных значений температуры.

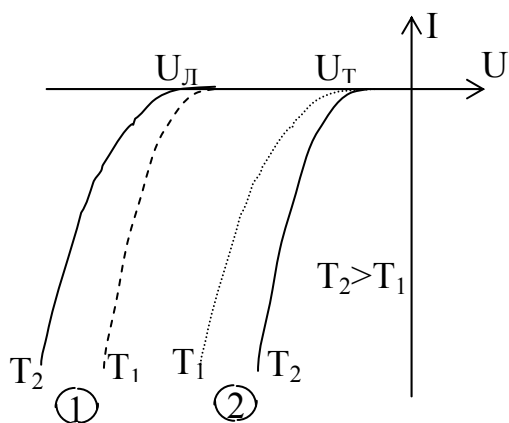


Рис. 1.29

Основными отличия в параметрах следующие. Напряжение лавинного пробоя превышает соответствующий параметр туннельного $U_L > U_T$, а именно: $U_T = 3...5$ В, $U_L > 7$ В. По величинам динамических сопротивлений $R_{дл} < R_{дт}$, т.е. ВАХ лавинного пробоя имеет большую крутизну, так как процесс лавинного размножения зарядов более интенсивный. Два вида пробоев

имеют также разные знаки температурного коэффициента напряжения (ТКН), т.е. положительный – для лавинного и отрицательный – для туннельного. При изменении температуры характеристики смещаются в разные стороны от номинальных. Это связано с тем, что для лавинного пробоя уменьшается длина свободного пробега при повышении температуры, а значит, и энергия, которую приобретают заряды на этом пространственном интервале. Поэтому при меньшей $I_{св}$ нужна большая напряженность поля (большие обратные напряжения), чтобы носитель заряда приобрел энергию, достаточную для ионизации.

Для туннельного пробоя с увеличением температуры напряжение пробоя уменьшается (или увеличивается ток). Это вызвано уменьшением ширины запрещенной зоны с ростом температуры, в результате чего уменьшается и толщина потенциального барьера, т.е. при той же напряженности электрического поля увеличивается вероятность туннелирования (рост туннельного тока).

Электрические пробои многократно повторимы, т.е. носят обратимый характер в отличие от необратимого теплового пробоя.

Тепловой пробой определяется неуправляемым ростом обратного тока вследствие саморазогрева перехода, что приводит к разрушению структуры. Этот вид пробоя связан с нарушением энергетического равновесия между мощностью, выделяемой в переходе, и мощностью, отдаваемой во внешнюю среду, причем в результате действия внутренней положительной обратной связи (ПОС) процесс увеличения температуры развивается лавинообразно. Логическую цепочку действия ПОС можно отобразить следующим образом:

$$P \uparrow \rightarrow I_{\text{обр}} \uparrow U_{\text{обр}} \rightarrow T \uparrow \rightarrow \varepsilon \uparrow = KT \rightarrow p, n \uparrow.$$

Границу между электрическим и тепловым пробоями определяет на ВАХ гипербола критического режима: $P_{\text{max}} = I_{\text{max}} U_{\text{max}} = \text{const}$ (рис. 1.30). Ниже этой кривой развивается тепловой пробой, для которого рост тока сопровождается уменьшением напряжения. Наличие «падающего» участка ВАХ с отрицательным динамическим сопротивлением характерно при действии ПОС. Предотвратить переход лавинного пробоя в тепловой можно, если ограничить величину обратного тока. Это можно сделать, включив последовательно со структурой перехода ограничительный резистор.

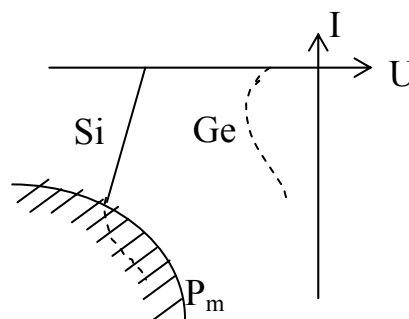


Рис. 1.30

Следует отметить, что в переходах, выполненных на основе германия, практически отсутствует выраженный участок лавинного пробоя. Вследствие большой величины начального обратного тока перехода тепловой пробой возникает раньше лавинного (см. рис. 1.30). Поэтому германиевые переходы критичны даже относительно импульсных (кратковременных) перегрузок.

9. Фотогальванический эффект – это процесс генерации электронно-дырочных пар, возникающих при облучении *p-n*-перехода, с последующим их разделением полем перехода. В результате действия этого эффекта увеличивается ток обратносмещенного перехода или появляется фото-ЭДС. Особенности этого явления рассматриваются в разделе оптоэлектронных приборов.

2. Полупроводниковые диоды

Диод – это электропреобразовательный прибор, который имеет два внешних вывода. Этот элемент является нелинейным двухполюсником, т.е. его вольт-амперная характеристика описывается нелинейной функцией.

Диоды выполняют различные функциональные преобразования сигналов: выпрямление, переключение, умножение и преобразование частоты, преобразование электромагнитного сигнала в электрический и т.д.

Работа большинства диодов основана на использовании физических явлений в электрическом переходе. Чаще всего применяют электронно-дырочный переход, переход металл-полупроводник (Шотки), гетеропереход. Существуют диоды, которые не имеют электрических переходов (диоды Ганна) или содержат несколько переходов (динистор, *p-i-n*-диод).

Диоды классифицируются по ряду признаков, важнейшими из которых являются физические эффекты и явления, положенные в основу их работы. В частности, в справочной литературе диоды подразделяются по функциональному назначению и применению.

С точки зрения геометрических особенностей переходов диоды делятся на плоскостные или точечные. В зависимости от технологии изготовления электрических переходов – на сплавные (микросплавные), диффузионные, эпитаксиальные и т.д.

Диоды выпускаются в различном конструктивном исполнении: стеклянном, керамическом, пластмассовом, металлокерамическом корпусах.

В основу маркировки диодов положен буквенно-цифровой код, содержащий информацию об исходном материале, их назначении, основных электрических параметрах, порядковом номере разработки.

В малогабаритных диодах для обозначения типа наносят на корпус прибора условные элементы: цветные кольца, точки и т.д.

Для изготовления диодов используют германий, кремний, арсенид галлия, причем основная номенклатура приборов выполняется на основе кремния.

2.1. Выпрямительные диоды

Выпрямительный диод предназначен для преобразования переменного тока в однополярный пульсирующий. Работа диода основана на свойстве односторонней проводимости. Процесс преобразования приложенного к диоду переменного напряжения приведен на

рис. 2.1. В преобразованном законе изменения тока содержится постоянная составляющая

$$I_{cp} = \frac{1}{T} \int_0^{T/2} I_m \sin(\omega t) dt = \frac{I_m}{\pi}, \quad (2.1)$$

так как циклическая частота $\omega = 2\pi/T$. Поэтому приведенное выше преобразование называется выпрямлением. Условное графическое обозначение данного диода показано на рис. 2.2. Наличие стрелки отражает свойство односторонней проводимости. Диод проводит ток, если его направление совпадает со стрелкой ($I_{пр}$ – прямой ток). При этом следует помнить, что ток протекает от высокого потенциала к низкому.

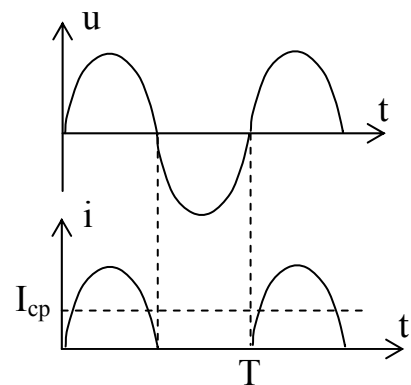


Рис. 2.1

Пример маркировки выпрямительного диода КД109А. Первый элемент (буква или цифра) определяет полупроводниковый материал, на основе которого изготовлен диод (Г или 1 – германий, К или 2 – кремний, А или 3 – арсенид галлия). Второй элемент (буква Д) – подкласс диода (выпрямительный, импульсный, универсальный). Третий элемент (цифра) определяет функциональное назначение диода (101-299 – выпрямительные, причем две последние цифры – это порядковый номер разработки). Шестой элемент (буква А) определяет разновидность по некоторым параметрам.

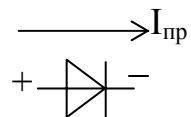


Рис. 2.2

Для диодов, разработанных до 1964 года, использовалась упрощенная маркировка: Д226 – кремниевый выпрямительный диод (201-299 – выпрямительные кремниевые, 301-399 – германиевые).

Основные параметры выпрямительных диодов объединяются в такие группы:

1. Предельно эксплуатационные – это максимальные параметры, превышение которых приводит к выходу прибора из строя. Эти значения ограничивают возможный рабочий диапазон ВАХ как на прямой, так и на обратной ветви (рис. 2.3). Очевидно, что ограничению подлежат наиболее динамичные параметры: а) максимальное значение прямого тока $I_{пр.макс}$ (или средний прямой ток, учитывая эффект выпрямления, см. рис. 2.2); б) макси-

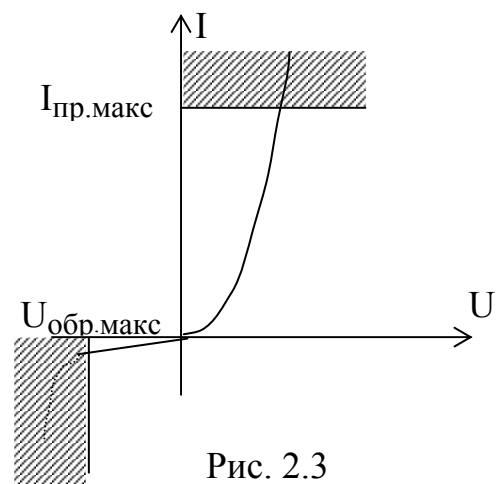


Рис. 2.3

мальное обратное напряжение $U_{обр. макс}$ (постоянное или импульсное). Превышение прямого тока приводит к недопустимому разогреву структуры, а увеличенное значение обратного напряжения выводит на участок пробоя ВАХ, что нарушает свойство односторонней проводимости. Обычно это напряжение берется с некоторым запасом $U_{обр. макс} = (0,5 - 0,8)U_{проб.}$ По величине среднего выпрямленного тока различают диоды малой $I_{пр.ср.макс} < 0,3$ А (параметр маркировки – 101-199) и средней мощности $0,3 А < I_{пр.ср.макс} < 10$ А (201-299). Диапазоны указанных предельных параметров составляют: по обратному напряжению – от десятков до тысяч вольт, по прямому току – от сотен миллиампер до десятков (сотен) ампер.

Температурный диапазон работы диодов ограничен снизу $t = -60^{\circ}\text{C}$, а сверху, в зависимости от материала, Ge – 85°C , Si – 125°C , GaAs – 150°C .

2. Статические электрические параметры определяют степень близости реального выпрямительного диода к идеальному. Идеальный ключевой элемент имеет ВАХ, представленную на рис. 2.4. Он

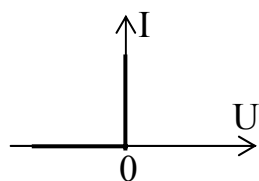


Рис. 2.4

характеризуется идеальным свойством односторонней проводимости: нулевым значением обратного тока и нулевым падением напряжения в проводящем состоянии. Бесконечное значение крутизны в начале координат указывает на нулевое значение статического и динамического сопротивлений.

Реальный диод имеет следующие (конечные) значения соответствующего тока и напряжения (см. рис. 2.3):

а) $U_{пр}$ – прямое падение напряжения при заданном прямом токе ($I_{пр.макс}$), или среднее значение этого напряжения (десятые доли вольта);

б) $I_{обр}$ – обратный ток при заданном обратном напряжении ($U_{обр. макс}$) (от долей микроампер до единиц миллиампер).

3. Динамические параметры характеризуют выпрямительный эффект диода при действии на него переменного напряжения. К ним относятся:

а) $I_{вп.ср}$ – среднее за период значение выпрямленного тока;

б) $I_{обр.ср}$ – среднее за период значение обратного тока;

г) $f_{гр}$ – граничная частота, на которой выпрямленный ток уменьшается до заданного уровня.

Операция усреднения соответствует выражению (2.1). При повышении частоты выпрямленный ток уменьшается, так как увеличивает-

ся импульс обратного тока в результате проявления эффекта накопления носителей заряда.

Выпрямительные диоды имеют следующие особенности. Малое сопротивление в прямом направлении позволяет пропускать большие прямые токи. Высокие значения прямых токов требуют электрических переходов с большой площадью. Поэтому диоды имеют большие значения барьерных и диффузионных емкостей, т.е. работают на низких частотах. Частотный диапазон составляет от 50 Гц для сети переменного тока и достигает нескольких десятков килогерц в транзисторных преобразователях напряжения.

Диоды имеют сравнительно большие размеры и предусматривают конструктивные особенности для возможности рассеивания больших тепловых мощностей.

В основном эти диоды изготавливаются на основе кремния, меньшая часть – на основе германия, кроме того, выпускаются диоды из арсенида галлия. Используются электрические переходы в виде *p-n*-структур, металл-полупроводник. Выпрямляющие переходы с чередованием *p-i-n*-областей характеризуются повышенными значениями обратных напряжений.

Кремниевые диоды в сравнении с германиевыми обладают целым рядом преимуществ: они имеют большие значения прямых токов и обратных напряжений, меньшие величины обратных токов

$$I_{обр}(Ge) = I_{обр}(Si)10^{1,5...2},$$

а также рассеивают большие мощности. К числу их недостатков следует отнести более высокие падения напряжений: $U_{пр}(Si) = 0,7...1,2$ В, $U_{пр}(Ge) = 0,3...0,5$ В. Это связано с большей величиной контактной разности потенциалов, а также более высоким сопротивлением базы из-за меньшей подвижности носителей заряда.

Диоды из арсенида галлия могут работать при больших температурах (до 250⁰С) и имеют лучшие частотные характеристики, но при этом более высокие падения напряжения (АД112А $U_{пр} = 3$ В, АД110А $U_{пр} = 1,5$ В) и небольшие допустимые обратные напряжения ($U_{обр.макс} = 30 ... 50$ В).

В высоковольтных источниках применяют выпрямительные столбы, которые работают при повышенных обратных напряжениях. Этот элемент представляет собой совокупность из последовательных диодов с близкими параметрами, объединенных в одном корпусе. Практическая идентичность диодов в этой конструкции обеспечивает равномерное распределение напряжения между диодами и не вызывает пробоя отдельных диодов. Так, например, выпрямительный столб 2Ц101А рассчитан на предельное $U_{обр} = 700$ В, а 2Ц202Е – на им-

пульсное $U_{обр} = 10$ кВ. Элемент маркировки Ц и числовые значения (101-299) определяют выпрямительные столбы. Представитель столба более ранней разработки Д1008 обеспечивает $U_{обр. имп.} = 10$ кВ.

Объединение диодов в последовательную схему в целях повышения обратного напряжения можно реализовать из отдельных диодов. Однако в этом случае нужно шунтировать каждый диод параллельно включенным резистором с сопротивлением несколько меньшим обратного сопротивления диода. При неидентичности диодов это исключает неравномерное распределение напряжения между ними.

Для повышения допустимого прямого тока отдельные диоды соединяют параллельно. Для выравнивания токов последовательно с каждым диодом включают резистор.

При реализации стандартных выпрямительных схем применяют выпрямительные блоки. Это совокупность диодов, соединенных по определенной схеме (например мостовой) и оформленных в единую конструкцию. Например, КЦ405А – выпрямительный блок, содержащий четыре диода, соединенных по схеме однофазного моста, и предназначенный для монтажа на печатную плату (рис. 2.5). Элемент маркировки Ц совместно с числовым параметром (401-499) определяют выпрямительные блоки.

В схемах выпрямления используются также диодные сборки – набор однотипных диодов, не соединенных электрически или с объединением одноименных выводов. Пример диодной сборки из двух кремниевых диодов (КДС111А) показан на рис. 2.6.

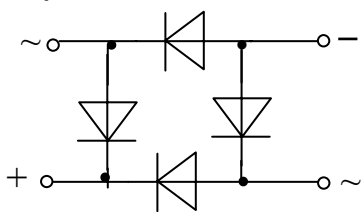


Рис. 2.5

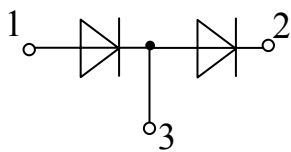


Рис. 2.6

динением одноименных выводов. Пример диодной сборки из двух кремниевых диодов (КДС111А) показан на рис. 2.6.

2.1.1. Применение выпрямительных диодов

1. Однополупериодная схема выпрямления. Схема содержит выпрямительный диод VD, резистор нагрузки R и емкость фильтра C (рис. 2.7). Диод преобразует входное переменное напряжение в пульсирующее U_R . При отсутствии емкости выпрямленное напряжение выделяется на резисторе нагрузки. Емкость осуществляет сглаживание выходного напряжения U_{RC} (отфильтровыв-

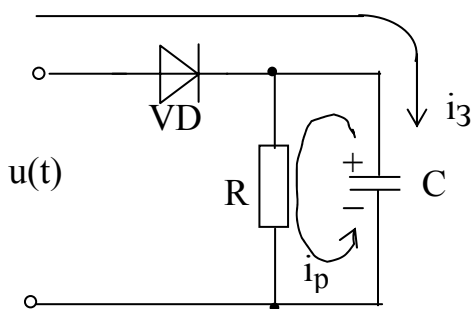


Рис. 2.7

вает, выделяет постоянную составляющую U_0). В течение интервала времени t напряжение на емкости превосходит мгновенное значение входного напряжения $U_C > U(t)$, поэтому диод закрыт. Емкость медленно разряжается через резистор. Как только входное напряжение превысит емкостное $U(t) > U_C$, диод открывается и происходит быстрый заряд емкости в течение t_3 через диод. При этом он пропускает короткие импульсы зарядного тока i_3 длительностью t_3 , причем импульсный ток заряда намного больше среднего выпрямленного тока, так как из условия равенства зарядов $I_3 t_3 = I_{CP} T$ (рис. 2.8). Для эффективного сглаживания выпрямленного напряжения резистор нагрузки должен шунтироваться (закорачиваться) емкостью по переменной составляющей, т.е. соотношение между емкостным и активным сопротивлением должно удовлетворять следующему неравенству:

$$X = \frac{1}{\omega C} \ll R.$$

Учитывая, что $\omega = 2\pi/T$, и полагая, что десятикратное отличие в сопротивлениях достаточно, получаем эквивалентное равенство

$$\tau_p = RC = 10T/2\pi > T.$$

Следовательно, постоянная времени разряда емкости должна превышать период входного напряжения T . Заметим, что емкость разряжается по закону экспоненты

$$u_c(t) = U_{C0} e^{-\frac{t}{\tau_p}},$$

т.е. в течение $t = \tau_p$ напряжение на емкости уменьшается в e раз.

Параметр, который определяет качество выпрямительной схемы, – это коэффициент пульсаций $K_{\Gamma} = \Delta U / U_0$, где ΔU – диапазон напряжения пульсаций. Так как ток через диод протекает в течение небольшого интервала времени, то пульсации выходного напряжения аппроксимируются пилообразным законом (рис. 2.9). Изменение заряда на емкости $\Delta Q = \Delta U C = I_0 T$, где средний ток в нагрузке

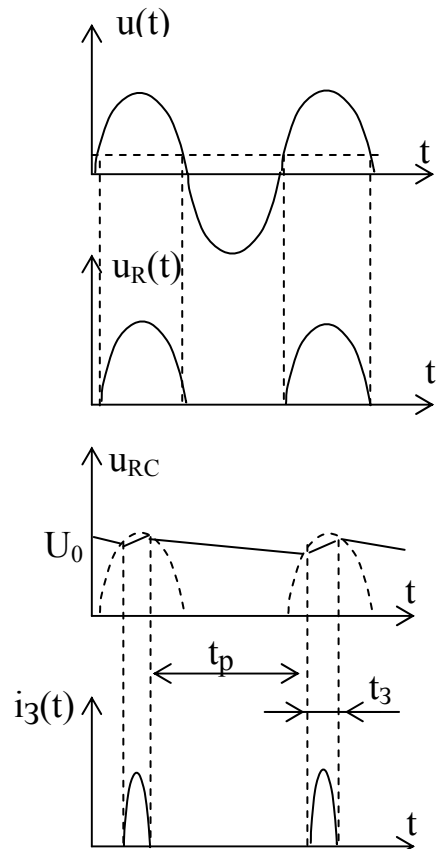


Рис. 2.8

$I_0 = U_0/R$, откуда $\Delta U = U_0 T/(RC)$, следовательно, коэффициент пульсаций

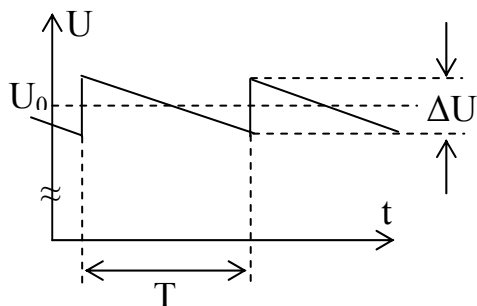


Рис. 2.9

$$K_{\Pi} = \frac{T}{\tau_p} = \frac{1}{fRC} \quad (2.2)$$

Недостатки рассмотренной схемы следующие:

1) неполное использование входного напряжения, так как при формировании постоянной составляющей используется только положительная полуволна; 2) диод в закрытом состоянии должен выдерживать двойное амплитудное значение входного напряжения, так как на диод действует сумма напряжений: отрицательная полуволна входного напряжения и напряжение на емкости.

2. Двухполупериодная схема выпрямления. В этой схеме в процессе выпрямления используется как положительная, так и отрицательная полуволна напряжения. Каждая из них должна формировать на нагрузке положительный импульс, т.е. направления токов в резисторе должны совпадать для двух тактов работы. При данном условии используем вначале две схемы, каждая из которых работает по

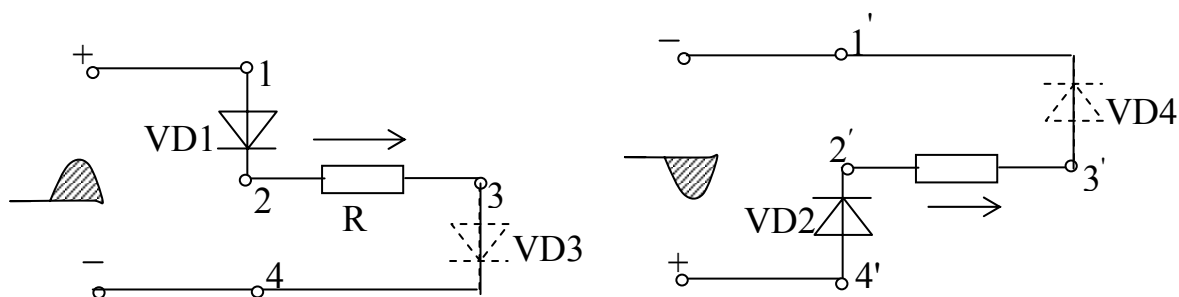


Рис. 2.10

соответствующей полуволне напряжения (рис. 2.10).

Объединим две схемы в одну, наложив одну на другую. В этом случае нужно включить два диода VD3, VD4, которые устраняют короткие замыкания, возникающие при наложении схем. Окончательный вариант схемы приведен на рис. 2.11. Выпрямленное напряжение показано на рис. 2.12. В этой схеме период следования импульсов уменьшается вдвое, что облегчает фильтрацию постоянной составляющей. Кроме того, каждый из диодов в закрытом состоянии удерживает только амплитудное напряжение.

3. Схемы удвоения напряжений. В схемах выпрямления отфильтрованная постоянная составляющая не может превышать амплитудного значения входного напряжения. Отрицательную полуволну на-

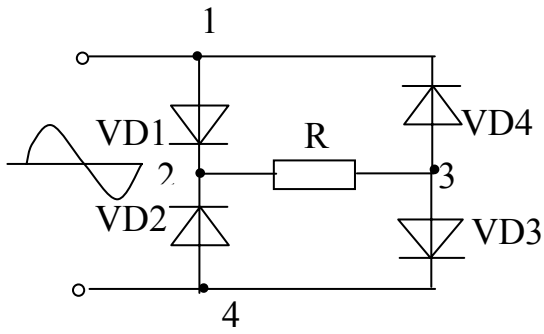


Рис. 2.11

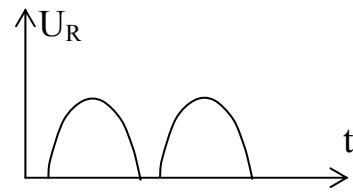


Рис. 2.12

пряжения также можно использовать для увеличения амплитуды в процессе выпрямления. Один из вариантов такой схемы приведен на рис. 2.13. Каждый из диодов работает на свою нагрузку, причем общее выходное напряжение равно сумме напряжений на резисторах. Схема работает в двухтактном режиме. В течение положительной полуволны происходит подзаряд емкости C_1 через диод VD_1 , при отрицательной полуволне работает диод VD_2 и емкость C_2 .

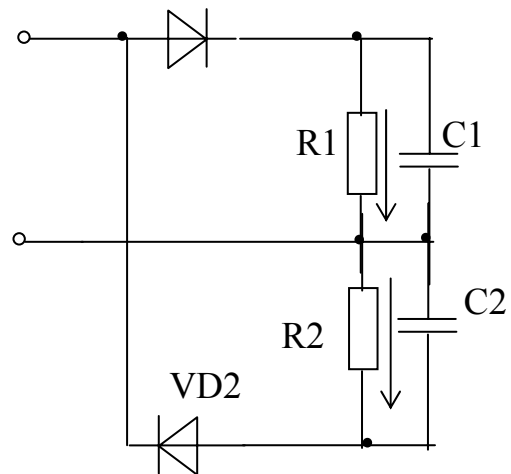


Рис. 2.13

Второй вариант схемы показан на рис. 2.14. Во время отрицательной полуволны через диод VD_1 подзаряжается емкость C_1 . Под действием положительной полуволны (полярность входного напряжения показана в скобках) через диод VD_2 подзаряжается емкость C_2 . Так как источниками заряда C_2 являются первичный входной и вторичный источники (напряжение на емкости C_1), включенные согласовано, то C_2 заряжается практически до удвоенного амплитудного значения.

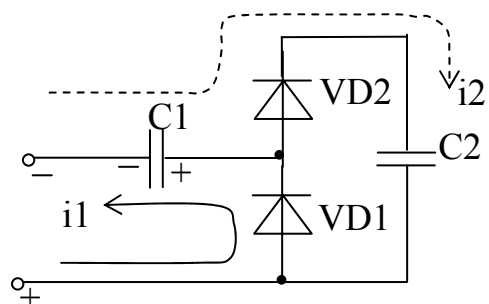


Рис. 2.14

В рассмотренных выше схемах используется ключевое свойство диодов, которые коммутируют различные части схемы в зависимости от мгновенной полярности входного напряжения.

2.1.2. Особенности расчета простейших схем с диодами

1. Диод в цепи постоянного тока. Простейшая схема содержит источник

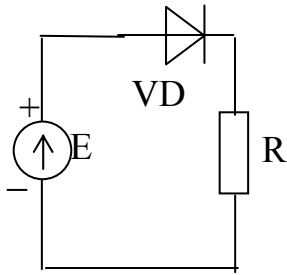


Рис. 2.15

источник постоянного напряжения E , диод VD и резистор нагрузки R (рис. 2.15). Для заданных E, R и известной характеристики диода определим величину тока в цепи I и падение напряжения на диоде U . Особенность схемы заключается в наличии нелинейного элемента – диода, в котором зависимость тока от напряжения не подчиняется закону Ома, а описывается нелинейной ВАХ: $I = f(U)$. Анализ схемы можно провести двумя методами.

тодами.

Графоаналитический метод. Анализируемая схема содержит два пассивных элемента – диод и резистор, которые описываются следующими зависимостями: $I = f(U)$, $U_R = IR$. Заменим второе соотношение на уравнение, записанное по второму закону Кирхгофа, чтобы получить систему из двух уравнений с двумя неизвестными I, U : $I = f(U)$, $E = U + IR$. Так как одно из уравнений нелинейное, то решать систему точным аналитическим методом не представляется возможным. Поэтому используется графический метод решения. Для этого находим точку пересечения нелинейной ВАХ с прямой, которая описывается вторым уравнением системы. Прямую строим по двум точкам, взятым на осях координат (рис. 2.16). Координаты точки (1), лежащей на оси напряжений, $I = 0$, $U = E$, для точки (2), лежащей на оси

токов, $U = 0$, $I = E/R$. Ненулевые координаты определяются подстановкой соответствующих нулевых токов и напряжений в уравнение прямой. Эти же значения тока и напряжения можно определить схемотехнически. Точка (1) соответствует режиму холостого хода, когда цепь разомкнута, т.е. $I = 0$. Точка (2) – режиму короткого замыкания, когда диод замкнут накоротко, так как $U = 0$. Тогда длины отрезков, отсекаемых прямой на осях координат, численно равны напряжению холостого хода и току короткого замыкания: $U_x = E$, $I_{кз} = E/R$, что

отмечается соответствующими индексами. Построенная прямая называется нагрузочной прямой. Нагрузочная прямая – это ВАХ линейного элемента (резистора нагрузки), построенная в системе координат нелинейного элемента (диода). Решением системы уравнений является точка A – рабочая точка. Ее координаты определяют ток в цепи I_0

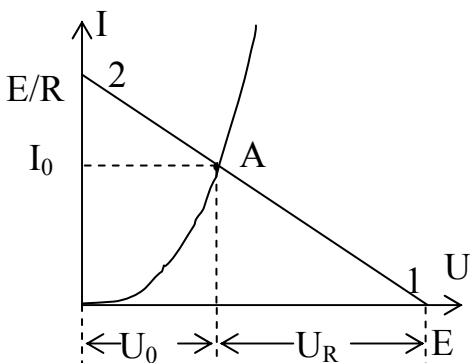


Рис. 2.16

и падение напряжения на диоде U_0 . ЭДС источника E перераспределяется между диодом и резистором, т.е. $E = U_0 + U_R$.

Приближенный аналитический метод. В этом методе задаются средней величиной падения напряжения на диоде, например, $U = 0,7$ В – для кремниевого диода, $U = 0,3$ В – для германиевого. Основанием для такого приближения является большая крутизна ВАХ диодов особенно при повышенных токах. Поэтому ток в цепи определяется через падение напряжения на резисторе:

$$I_0 = \frac{U_R}{R} = \frac{E - U}{R}. \quad (2.3)$$

2. Диод в цепи переменного тока. В этой схеме действует источник с гармонической ЭДС $e(t) = E_m \sin \omega t$ (рис. 2.17).

Тогда нагрузочная прямая перемещается параллельно самой себе при изменении мгновенного значения напряжения на входе (рис. 2.18). Импульс выходного тока формирует в нагрузке импульс напряжения $u_R(t) = i(t)R$. Изменения входного и выходного напряжений сопоставлены на рис. 2.19.

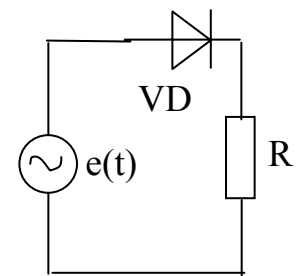


Рис. 2.17

Амплитудное значение напряжения на нагрузке U_R

меньше амплитуды входного сигнала E_m на величину падения напря-

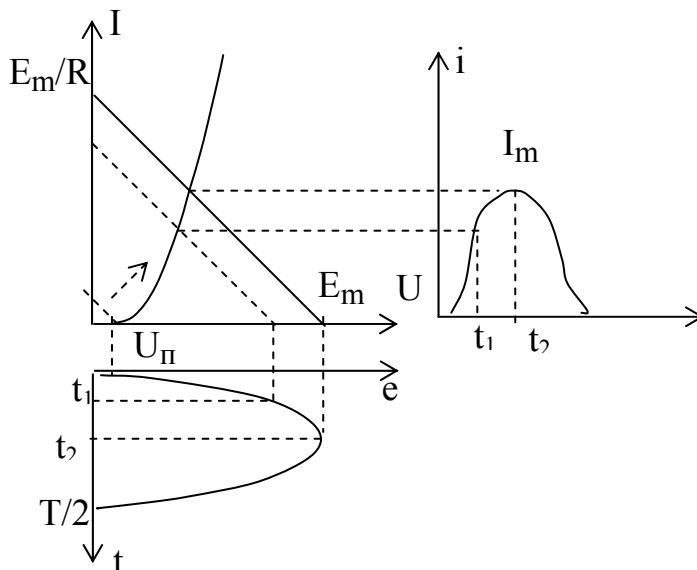


Рис. 2.18

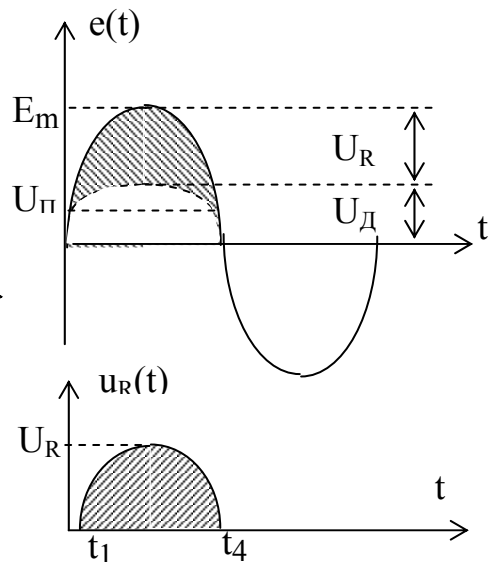


Рис. 2.19

жения на диоде u : $u_R(t) = e(t) - u(t)$. В свою очередь $u(t) = U_п + \Delta u(t)$ состоит из падения напряжения, вызванного пороговым эффектом $U_п$, а также переменной составляющей, величина которой определяется динамическим сопротивлением диода.

Для определения напряжения ΔU аппроксимируем реальную ВАХ диода кусочно-линейной характеристикой (см. рис. 2.18) и представим

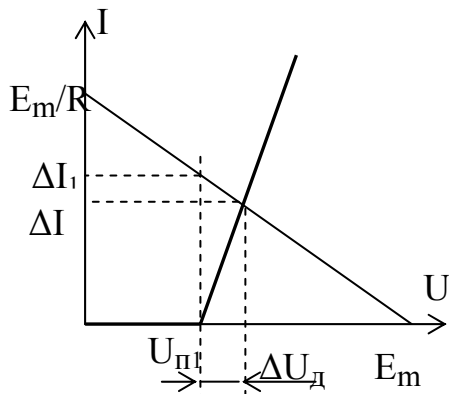


Рис. 2.20

ее на рис. 2.20, где $U_{п1}$ – пороговое напряжение аппроксимированной характеристики, ΔU – амплитудное значение переменной компоненты падения напряжения на диоде. Амплитуда импульса напряжения на нагрузке для идеального диода (пунктирная ВАХ) $\Delta U = E_m - U_{п1}$. Амплитуда импульса тока $\Delta I_1 = \Delta U/R$. Амплитуда выходного напряжения должна быть уменьшена на величину падения на-

пряжения, соответствующую аппроксимированной характеристике: $\Delta U_{\partial} = \Delta I R_{д} \approx \Delta I_1 R_{д} = (R_{д}/R) \Delta U$, т.е. $\Delta U_R = \Delta U - (R_{д}/R) \Delta U = \Delta U (1 - (R_{д}/R))$. Так как $R_{д}/R \ll 1$, то $\Delta U_R \approx \Delta U$. Поэтому приближенный расчет в этой цепи можно проводить по методике п.1.

3. Диод в смешанной цепи. В этой схеме наряду с источником постоянного напряжения смещения E_0 действует переменная ЭДС $e(t)$ (рис. 2.21). Линия нагрузки смещается в динамике в обе стороны относительно номинального положения, которое соответствует источнику постоянного смещения (рис. 2.22). Если в цепи действует только

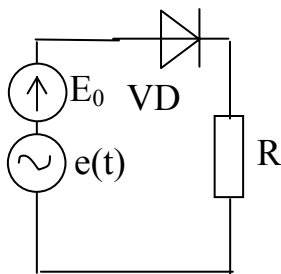


Рис. 2.21

источник постоянного смещения, то величина постоянного тока $I_0 = (E - U_{\partial})/R \approx E/R$. Динамическое сопротивление диода в данной рабочей точке составляет $R_{д} = \varphi_T/I_0$. Заменяем диод эквивалентным резистором $R_{д}$, тогда данная схема по переменной составляющей будет представлять собой делитель напряжения. Амплитуда переменной составляющей на диоде

$$\Delta U_{д} = \frac{E_m}{R + R_{д}} R_{д} \approx E_m \frac{R_{д}}{R}$$

невелика, так как $R_{д} \ll R$ (см. рис. 2.22). Заметим, что при работе на нелинейном участке ВАХ закон изменения тока не является гармоническим, что приводит к изменению постоянной составляющей тока на величину ΔI (см. рис. 2.22). Поэтому по изменению постоянной составляющей тока можно обнаруживать переменный сигнал. Этот эффект носит название детекторного эффекта диода.

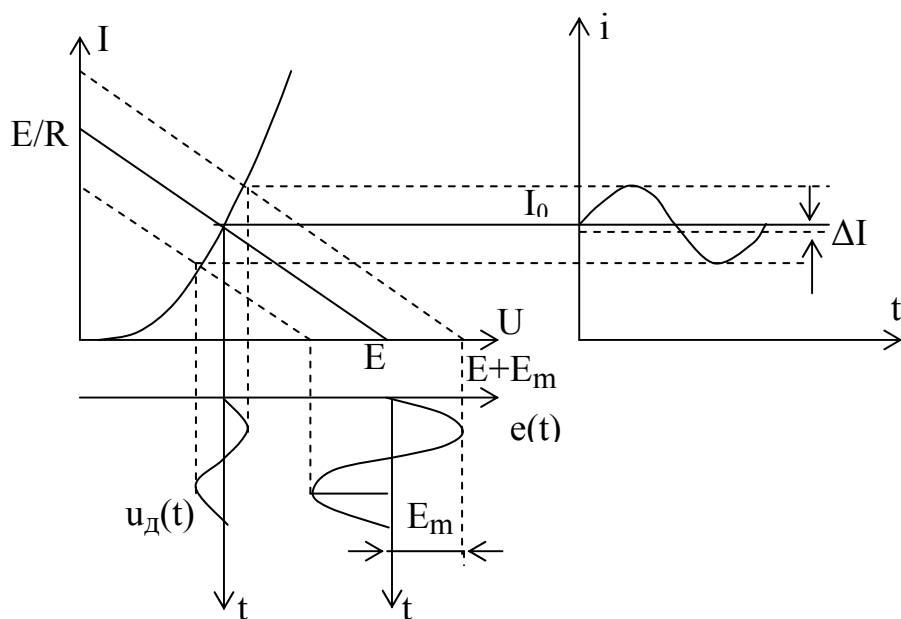


Рис. 2.22

2.2. Опорные диоды

Опорные диоды – это полупроводниковые приборы, в которых используются участки ВАХ с высокой крутизной, т.е. слабой зависимостью изменений напряжения от изменений тока.

Такие диоды осуществляют процесс стабилизации напряжения, т.е. поддержание практически неизменной величины напряжения в широком диапазоне изменений тока. На ВАХ диода имеются два участка с высокой крутизной как при прямом, так и при обратном смещении (рис. 2.23). Диод работает на участке стабилизации, если прямое напряжение превышает некоторое пороговое значение $U_{пр} > U_{п}$, а обратное – напряжение пробоя $U_{обр} > U_{проб}$. В соответствии с этим существуют две разновидности опорных диодов: 1) стабилитроны, использующие участок электрического пробоя (обратное смещение $U > 3$ В); 2) стабилоры, работающие на участке прямой ветви диода ($U_{ст} = 0,7...2$ В). Диапазоны стабилизируемого напряжения $U_{ст} = 0,7...2$ В – для стабилоры и $U_{ст} > 3$ В – для стабилитрона определяются, соответственно, падением

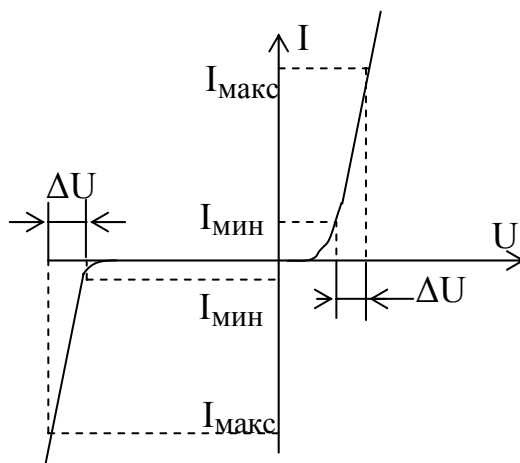


Рис. 2.23

напряжения на прямосмещенном кремниевом переходе и минимальным напряжением электрического пробоя.

Условные графические обозначения приведены на рис. 2.24.

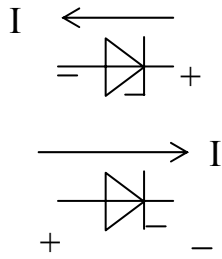


Рис. 2.24

Здесь же указаны полярности рабочих напряжений, а также направления рабочих токов. В маркировке опорных диодов для указания подкласса стабилизаторов или стабилиторов используется функциональный элемент обозначения С (например КС168А). Три последующих цифровых элемента обозначают мощность диода и напряжение стабилизации. Опорные диоды разбиты на три подгруппы по величине рассеиваемой мощности, а внутри каждой подгруппы проведена дифференциация по величине напряжения стабилизации (табл. 2.1).

Таблица 2.1

| $P < 0,3$ Вт | $P = 0,3...5$ Вт | $P > 5$ Вт | $U_{ст}, В$ |
|--------------|------------------|------------|-------------|
| 101...199 | 401...499 | 701...799 | 0,1...9,9 |
| 210...299 | 510...599 | 801...899 | 10...99 |
| 301...399 | 601...699 | 901...999 | 101...199 |

По двум последним цифрам определяется величина напряжения стабилизации, причем значащая разрядность цифр зависит от строки таблицы. В первой строке между двумя последними цифрами ставится запятая, во второй – непосредственно выделяются две цифры, а в третьей – к двум цифрам добавляется единица в старшем разряде. Так, указанный выше диод имеет $U_{ст} = 6,8$ В и соответствует мало-мощному диоду с мощностью $P < 0,3$ Вт. Для более поздних разработок диодов две последние цифры не определяют точное значение напряжения стабилизации, а только указывают на принадлежность к соответствующему диапазону. Так, например, стабилитроны средней мощности КС406А, КС509А имеют напряжения стабилизации, соответственно, $U_{ст1} = 8,2$ В, $U_{ст2} = 14,7$ В. Первый из указанных выше диодов принадлежит 1-ой строке таблицы, а второй – 2-ой строке. Для опорных диодов ранних выпусков присвоен числовой интервал 801...899 без градации по мощности и напряжению стабилизации. Так, мощный стабилитрон Д815А имеет $U_{ст} = 5,6$ В.

Первоначальный выбор требуемого диода проводится по двум основным классификационным параметрам: рассеиваемой мощности и напряжению стабилизации.

Основные параметры опорного диода характеризуют, в частности, «геометрические» особенности ВАХ (рис. 2.25). К таким параметрам относятся:

1) номинальное (среднее) значение напряжения стабилизации или диапазон его изменения ($U_{ст.мин} \dots U_{ст.макс}$);

2) диапазон изменения тока на участке стабилизации ($I_{ст.мин} \dots I_{ст.макс}$);

3) динамическое сопротивление R_d , которое определяет крутизну характеристики;

4) температурный коэффициент напряжения стабилизации

$$ТКН = \alpha = \frac{\Delta U}{U_{ст} \Delta T},$$

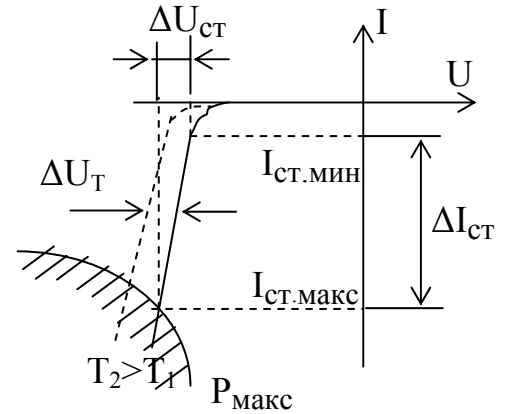


Рис. 2.25

где $\Delta T = T_2 - T_1$, что определяет температурную стабильность диода по основному параметру – напряжению стабилизации;

5) максимально допустимая рассеиваемая мощность

$P_{макс} = I_{ст. макс} U_{ст. макс}$, которая предохраняет переход диода в режим теплового пробоя.

К предельным эксплуатационным параметрам относятся: $P_{макс}$, $I_{ст.макс}$, $I_{ст. мин}$. Превышение первых двух параметров приводит к выходу прибора из строя, а при $I < I_{ст. мин}$ резко уменьшается крутизна ВАХ и увеличивается уровень шума на предпробойном участке. Минимальные значения тока стабилизации находятся в пределах 0,5...3 мА.

Динамическое сопротивление определяет качество стабилизации. Чем меньше этот параметр (выше крутизна ВАХ), тем в меньших пределах изменяется напряжение стабилизации при изменении тока. Это сопротивление составляет от единиц до десятков ом для маломощных диодов и увеличивается до сотен ом в диодах большей мощности. Это связано с тем, что напряжение стабилизации увеличивается с уменьшением степени легирования.

В зависимости от применения существуют следующие разновидности опорных диодов: общего назначения, импульсные, двуханодные, прецизионные (термокомпенсированные).

Диоды общего назначения используют для получения фиксированного (опорного) напряжения в схемах, для стабилизации постоянного напряжения в некотором диапазоне изменения тока. В цепях переменного тока – для ограничения мгновенных значений напряжения, т.е. получения неизменного напряжения в некотором малом временном интервале. В импульсных схемах они используются как быстродействующие пороговые элементы, так как в этих диодах нет инжекционных процессов накопления. Импульсные опорные диоды применяются для ограничения (фиксации) амплитуды импульсных сигналов.

В таких диодах минимизирована барьерная емкость, что позволяет получать малую длительность переходных процессов. Например, импульсный стабилитрон 2С175Е имеет общую емкость не более 7 пФ и время переключения порядка 1 нс.

Двуханодные диоды эквивалентны встречно-последовательному включению двух идентичных опорных диодов (рис. 2.26), поэтому

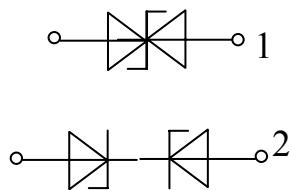


Рис. 2.26

имеют специфическое условное графическое обозначение. Эти диоды имеют симметричную ВАХ и применяются для стабилизации двухполярных напряжений или двухстороннего ограничения переменных, а также для защиты элементов цепи от перенапряжений обеих полярностей, причем в них нормируется несимметричность уровней напряжений стабилизации. Так, диод КС182А при $U_{ст} = 8,2$ В имеет несимметричность напряжения $\Delta U = 0,33$ В.

Прецизионные стабилитроны применяют как источники опорного напряжения, которые имеют малый разброс напряжения стабилизации, а также высокую температурную стабильность. В них дополнительно регламентируется временная нестабильность напряжения стабилизации, а также время выхода на режим (несколько десятков минут). Так, прецизионный диод КС166В имеет ТКН = $0,0005\%/^{\circ}\text{C}$ (уход напряжения стабилизации в широком диапазоне температур составляет 2 мВ), а временная нестабильность (за 5000 часов) равна 1,4 мВ. Конструктивно такой диод представляет собой три последовательно соединенных перехода: один, включенный в обратном направлении, – стабилизирующий и два прямосмещенных – термокомпенсирующие (рис. 2.27). Разные знаки ТКН для перехода, работающего в режиме лавинного пробоя, и для прямосмещенных переходов обеспечивают эффект термокомпенсации.

Зависимость ТКН от напряжения стабилизации приведена на рис. 2.28.

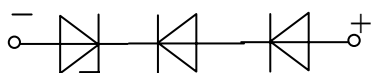


Рис. 2.27

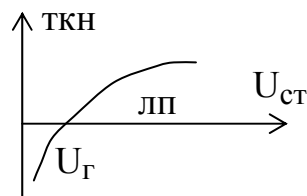


Рис. 2.28

Изменение знака ТКН происходит при граничном значении напряжения стабилизации 5-6 В при переходе в режим лавинного пробоя (ЛП).

2.2.1. Параметрический стабилизатор напряжения

Принцип стабилизации напряжения раскрывается на основе простейшей схемы (рис. 2.29). Она содержит два основных элемента: стабилитрон VD и балластный резистор R_B .

Резистор нагрузки R_H – это элемент цепи, на котором стабилизируется напряжение U_H . Схема обеспечивает практическую неизменность выходного напряжения при изменении входного ΔE относительно номинального E . Следует отметить, что диод подключается параллельно резистору нагрузки, так как именно при таком включении напряжение на этих двух элементах совпадает.

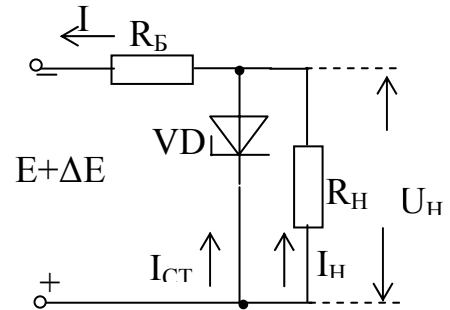


Рис. 2.29

Балластный резистор выполняет двойную роль: в статическом режиме ограничивает величину тока через диод, а в динамическом – обеспечивает эффект стабилизации.

Статический режим работы схемы анализируется при постоянном номинальном значении входного напряжения E . ЭДС входного источника распределяется между нагрузкой и балластным резистором:

$$E = U_H + U_B, \quad (2.4)$$

а общий ток I делится между диодом и резистором нагрузки:

$$I = I_{ст} + I_H. \quad (2.5)$$

Установим взаимосвязь между входным и выходным напряжениями. Тогда на основе (2.4) с учетом (2.5) имеем

$$E = U_H + (I_{ст} + I_H)R_B. \quad (2.6)$$

Выражая токи через выходное напряжение, получим

$$E = U_H + \left(\frac{U_H}{R_{ст}} + \frac{U_H}{R_H} \right) R_B,$$

или окончательно

$$E = U_H \left(1 + \frac{R_B}{R_{ст}} + \frac{R_B}{R_H} \right), \quad (2.7)$$

где $R_{ст}$ – статическое сопротивление диода.

Для анализа схемы в динамическом режиме заменим диод динамическим сопротивлением R_d . Тогда эквивалентная схема относительно изменений входного напряжения приведена на рис. 2.30. Эк-

вивалентное сопротивление параллельной цепи определяется по формуле

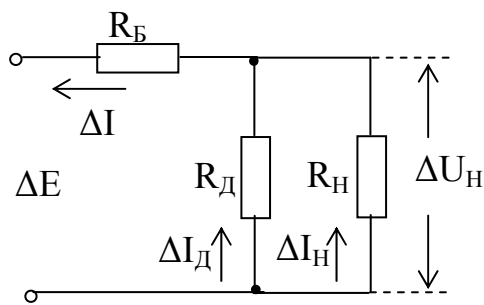


Рис. 2.30

$$R_{\text{э}} = \frac{R_{\text{д}} R_{\text{н}}}{R_{\text{д}} + R_{\text{н}}},$$

а так как $R_{\text{д}} \ll R_{\text{н}}$, то справедливо приближенное равенство

$$R_{\text{э}} = \frac{R_{\text{д}}}{\frac{R_{\text{д}}}{R_{\text{н}}} + 1} \approx R_{\text{д}}.$$

В этой схеме реализован принцип делителя напряжения относительно изменений входного напряжения, т.е. ΔE распределяется между $R_{\text{Б}}$ и $R_{\text{Д}}$. Так как $R_{\text{Д}} \ll R_{\text{Б}}$, то основное приращение напряжения падает на балластном резисторе, а изменение выходного напряжения незначительно:

$$\Delta U_{\text{Н}} = \Delta E \frac{R_{\text{Д}}}{R_{\text{Б}} + R_{\text{Д}}} \approx \Delta E \frac{R_{\text{Д}}}{R_{\text{Б}}},$$

т.е. $\Delta U_{\text{Н}} \ll \Delta E$. Отсюда становится очевидным назначение и название балластного резистора, он играет роль балласта, на котором падает основное изменение напряжения, поэтому только небольшая часть изменяющегося напряжения выделяется на резисторе нагрузки.

В этой схеме реализован также вспомогательный эффект – делителя приращений тока. Изменение общего тока, возникшее из-за нестабильности входного напряжения, делится между параллельными элементами $R_{\text{Д}}$ и $R_{\text{Н}}$. Так как $R_{\text{Д}} \ll R_{\text{Н}}$, то основное изменение тока пройдет через диод $\Delta I_{\text{Д}} \gg \Delta I_{\text{Н}}$, а поэтому $\Delta U_{\text{Н}} = I_{\text{Н}} R_{\text{Н}}$ также будет незначительным. Из эквивалентной схемы (см. рис. 2.30) следует

$$\Delta U_{\text{Н}} = \Delta I \frac{R_{\text{Д}} R_{\text{Н}}}{R_{\text{Д}} + R_{\text{Н}}},$$

откуда

$$\Delta I_{\text{Н}} = \Delta I \frac{R_{\text{Д}}}{R_{\text{Д}} + R_{\text{Н}}} \approx \Delta I \frac{R_{\text{Д}}}{R_{\text{Н}}}, \quad \text{т.е.} \quad \Delta I_{\text{Н}} \ll \Delta I.$$

В целом для динамического режима можно записать следующее соотношение по второму закону Кирхгофа (2.6):

$$\Delta E = \Delta U_{\text{Н}} + (\Delta I_{\text{СТ}} + \Delta I_{\text{Н}}) R_{\text{Б}},$$

или по аналогии с (2.7)

$$\Delta E = \Delta U_{\text{H}} \left(1 + \frac{R_{\text{Б}}}{R_{\text{H}}} + \frac{R_{\text{Б}}}{R_{\text{Д}}} \right). \quad (2.8)$$

Качество стабилизации всей схемы определяется коэффициентом стабилизации

$$K_{\text{СТ}} = \frac{\Delta E/E}{\Delta U_{\text{H}}/U_{\text{H}}}, \quad (2.9)$$

который показывает, во сколько раз относительные изменения входного напряжения больше, чем выходного. Используя взаимосвязь входных и выходных напряжений в статике (2.7) и динамике (2.8), получаем

$$K_{\text{СТ}} = \frac{1 + \frac{R_{\text{Б}}}{R_{\text{Д}}} + \frac{R_{\text{Б}}}{R_{\text{H}}}}{1 + \frac{R_{\text{Б}}}{R_{\text{СТ}}} + \frac{R_{\text{Б}}}{R_{\text{H}}}}. \quad (2.10)$$

Таким образом, здесь реализован схемотехнический принцип стабилизации напряжения, основанный на нелинейном делителе напряжения. Нелинейный элемент – диод, включенный в одно из плечей делителя, имеет существенные отличия в двух параметрах: статическом и динамическом сопротивлениях ($R_{\text{СТ}} \gg R_{\text{Д}}$). Благодаря этому он удерживает сравнительно большое статическое напряжение при малой величине падения напряжения относительно его изменений.

Наряду с физическим и аналитическим обоснованиями принципа работы схемы можно дать и геометрическую интерпретацию на основе нагрузочной прямой. Построим линию нагрузки, используя опыты холостого хода и короткого замыкания. Режим холостого хода соответствует разрыву ветви с диодом (см. рис. 2.29). В точках подключения диода напряжение холостого хода составляет

$$U_x = \frac{E}{R_{\text{H}} + R_{\text{Б}}} R_{\text{H}}.$$

Режим короткого замыкания $U = 0$ соответствует короткому замыканию ветви с диодом, тогда ток короткого замыкания в цепи $I_{\text{К}} = E/R_{\text{Б}}$. Отмечая точки холостого хода (1) и короткого замыкания (2) на осях координат ВАХ диода, получаем нагрузочную прямую (рис. 2.31). Нестабильность входного напряжения вызывает параллельное смещение нагрузочной прямой. Аналогично можно проанализировать изменение положения прямой при варьировании резисторов $R_{\text{Б}}$ и R_{H} .

Основная характеристика рассмотренной схемы, определяющая

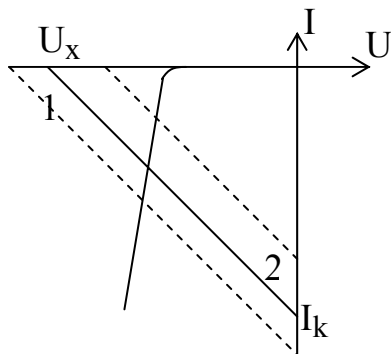


Рис. 2.31

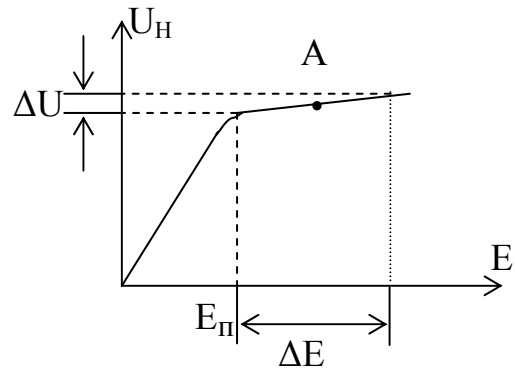


Рис. 2.32

взаимосвязь входного и выходного напряжений, - это передаточная динамическая характеристика $U_H = f(E)$ (рис. 2.32). При $E < E_П$ схема работает в режиме холостого хода, т.е. диод закрыт, а входное напряжение делится между R_B и R_H . Значение порогового входного напряжения определяется из условия перехода диода в режим пробоя $U_x = U_{CT}$, т.е.

$$E_П = U_{CT} \left(1 + \frac{R_B}{R_H} \right).$$

По этой характеристике находится экспериментальное значение коэффициента стабилизации (2.9). Номинальные значения напряжений E и U_H соответствуют рабочей точке А (см. рис. 2.32).

2.3. Варикапы

Варикап – это диод, принцип действия которого основан на зависимости барьерной емкости перехода от обратного напряжения.

Таким образом, варикап выполняет функции электрически пере-страиваемой емкости (variable capacitor). Из двух емкостей перехода: барьерной и диффузионной, для реализации емкостного элемента используется только барьерная емкость. В то же время диффузионная емкость соответствует прямому смещению перехода, что приводит к ее шунтированию малым динамическим сопротивлением перехода. При этом добротность емкости резко падает.

Условное обозначение варикапа показано на рис. 2.33. Здесь же указана полярность включения этого элемента, а также направление малого обратного тока. Так как идеальная емкость не пропускает постоянную составляющую тока (ток утечки равен нулю), то отсюда

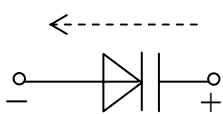


Рис. 2.33

становится понятным формальное правило включения варикапа. Напряжение на варикапе создает обратный ток, протекающий против направления стрелки в условном обозначении прибора.

В маркировке данного подкласса диодов указывается функциональный элемент (буква В), например, КВ109А или 2ВС118А – сборка из двух кремниевых варикапов с общим катодом, КВ132АР – комплект из двух варикапов (буква Р), К130АГ – комплект из четырех варикапов (буква Г). Для варикапов ранних выпусков выделены числа 901...950, например Д902.

Зависимость барьерной емкости от обратного напряжения $C_B = f(U_{обр})$ определяет вольт-фарадную характеристику (рис. 2.34), аналитическое представление которой задает следующее соотношение:

$$C_B = \frac{C_0}{\left(1 + \frac{|U_{обр}|}{\Phi_K}\right)^m}, \quad (2.11)$$

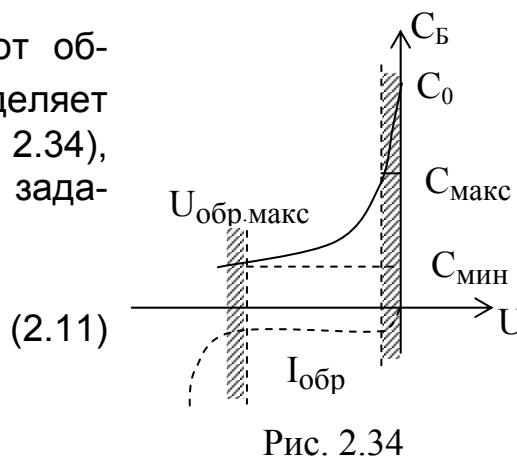


Рис. 2.34

где $m = 1/2$ – резкий переход, $m = 1/3$ – плавный, с линейным законом распределения примесей.

Эта зависимость имеет нелинейный характер, а рабочий участок ограничен значениями $C_{макс}$ и $C_{мин}$. При выходе за эти пределы происходит шунтирование емкости малым динамическим сопротивлением перехода (см. ВАХ, отмеченную на рис. 2.34 пунктиром).

Основные параметры варикапов:

- 1) C_H – номинальное значение емкости при заданном обратном напряжении (от десятков до сотен пикофарад);
- 2) $K_C = C_{макс}/C_{мин}$ – коэффициент перекрытия по емкости, определяет диапазон изменения параметра (несколько единиц);
- 3) Q – добротность при заданной частоте, определяет качество прибора (от десятков до сотен единиц);
- 4) $TKE = \Delta C/(C\Delta T)$ – температурный коэффициент емкости, определяет ее температурную нестабильность $(2 \cdot 10^{-4} \dots 6 \cdot 10^{-4})K^{-1}$;
- 5) $U_{обр. макс}$ – максимально допустимое обратное напряжение, предотвращает переход диода в режим пробоя (в среднем десятки вольт);

б) $I_{обр. макс}$ – максимальный обратный ток, соответствующий $U_{обр. макс}$, используется при расчете статического режима (в основном от десятых долей до единиц микроампер).

1, 2-й параметры характеризуют основной функциональный параметр прибора (емкость), 3,4,6-й – различные стороны качества прибора, а 5-й – предельный параметр. Параметр добротности во многом определяет конструктивно-технологические особенности варикапов. Для анализа добротности приведем упрощенную эквивалентную схему прибора для малых переменных сигналов (рис. 2.35). Так как диод работает при обратных напряжениях, то $p-n$ -переход представлен в виде параллельного соединения барьерной емкости и динамического сопротивления. Резистор r_B определяет объемное сопротивление базовой области как имеющую наименьшую степень легирования.

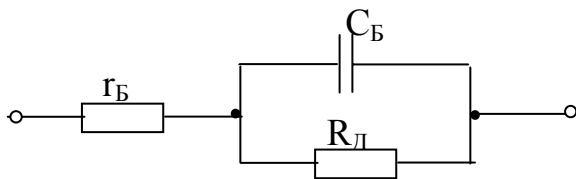


Рис. 2.35

Формально добротность рассчитывается как отношение реактивного сопротивления к активному. Записывая полное комплексное сопротивление варикапа

$$z = r_B + \frac{R_d / (j\omega C_B)}{R_d + \frac{1}{j\omega C_B}},$$

выделяя активную и реактивную (мнимую) компоненты и взяв их отношение, получим

$$Q = \frac{\omega C_B R_d^2}{r_B + r_B R_d^2 \omega^2 C_B^2 + R_d},$$

или в форме, удобной для анализа,

$$Q = \frac{\omega C_B R_d}{\frac{r_B}{R_d} + r_B R_d \omega^2 C_B^2 + 1}. \quad (2.12)$$

В этом соотношении $r_B \ll R_d$, кроме того, в области низких частот

$$\omega C_B \ll 1 / \sqrt{R_d r_B}, \quad \text{поэтому} \quad Q_{нч} \approx \omega C_B R_d, \quad (2.13)$$

а в области верхних частот

$$\omega C_B \gg 1/\sqrt{R_D r_B}, \quad \text{тогда} \quad Q_{\text{вч}} \approx \frac{1}{\omega C_B r_B}. \quad (2.14)$$

Эти же приближенные выражения можно получить, если рассмотреть эквивалентную схему (см. рис. 2.35) отдельно в разных частотных областях. На низких частотах емкостное сопротивление $X = 1/(\omega C_B)$ велико и соизмеримо с R_D , а $r_B \ll X$ и $r_B \ll R_D$, поэтому сопротивлением r_B в схеме можно пренебречь. В этой области частот эквивалентная схема представляет параллельное соединение C_B и R_D , из которой по определению добротности приходим к формуле (2.13). На высоких частотах $R_D \gg X$, поэтому в параллельном соединении следует пренебречь большим сопротивлением R_D . Тогда эквивалентная схема представляет собой последовательное соединение r_B и C_B , из которой находится величина добротности (2.14).

В полном диапазоне частот добротность имеет максимум в интервале 1...10 МГц – для низкочастотных варикапов и в интервале 10...30 МГц – для высокочастотных (рис. 2.36). Варикапы в основном применяют на высоких частотах, при этом рабочая формула добротности (2.14) соответствует последовательной схеме r_B и C_B . Поэтому для повышения добротности нужно уменьшать омическое сопротивление базовой области диода. Этот элемент структуры является также основным источником шумов. Для снижения базового сопротивления нужно уменьшать толщину базы и удельное сопротивление материала. Толщина не может быть меньше ширины области объемного заряда. Малое сопротивление базы получают выбором исходного полупроводникового материала с большой подвижностью носителей заряда (GaAs, Ge с n -типом проводимости). Повышать концентрацию носителей заряда в базе (увеличивать степень легирования) для снижения r_B нецелесообразно, так как при этом уменьшается подвижность носителей заряда и пробивное напряжение перехода.

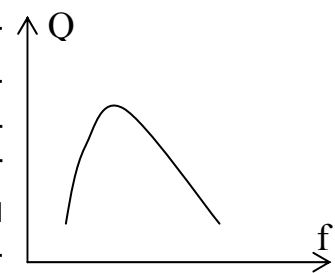


Рис. 2.36

Тонкую базу и одновременно большое пробивное напряжение получают в структуре $p^+ - n - n^+$, где $n - n^+$ – двухслойная базовая область. Низкоомный слой базы n^+ представляет собой довольно толстую подложку, а тонкий высоковольтный слой n непосредственно участвует в создании барьерной емкости.

К преимуществам варикапов можно отнести их практическую безинерционность, возможность электронной перестройки емкости, малые габариты и высокую надежность. Быстродействие варикапов оп-

ределяется временем перезарядки барьерной емкости при изменении обратного напряжения. При этом происходит смещение основных носителей заряда в прилегающих к переходу областях. Инерционность этого процесса определяется временем релаксации $T = \epsilon\epsilon_0\sigma$, что составляет для германия $1,4 \cdot 10^{-13}$ с.

Основные функциональные применения варикапов – это электронная перестройка частоты колебательных контуров и умножение частоты. Умножительные диоды ВЧ-диапазона получили название варакторных, например 3А614А – умножительный диод СВЧ-диапазона. Эти диоды используют нелинейное свойство вольт-фарадной характеристики. Тогда при синусоидальном управляющем напряжении форма тока будет негармонической, что дает возможность получения высших гармоник.

На рис. 2.37 приведена схема включения варикапа в колебательный контур. Емкость C_B в основном определяет резонансную частоту контура $\omega_0 \approx 1/\sqrt{LC_B}$, так как эквивалентная емкость контура

$$C_k = \frac{CC_B}{C + C_B} \approx C_B$$

при условии, что $C_B \ll C$. Разделительный конденсатор C «отделяет» индуктивность от варикапа по постоянной составляющей. В противном случае индуктивность зашунтирует (закоротит) диод и рабочая точка на ВАХ будет находиться в окрестности нулевого напряжения. Режим работы варикапа по постоянному току задает источник напряжения E (обратное смещение) и переменный резистор R_2 . Высокоомный резистор R_1 «отделяет» варикап от источника питания по переменной составляющей. В противном случае малое внутреннее сопротивление источника зашунтировало бы диод по переменной составляющей и резко уменьшило его добротность. Существенное увеличение этого резистора также нежелательно, так как обратный тепловой ток диода вызовет на высокоомном резисторе заметное падение напряжения, что приведет к нестабильности резонансной частоты контура.

Вариант включения варикапной сборки в контур приведен на рис. 2.38. В данной схеме увеличивается коэффициент перекрытия по

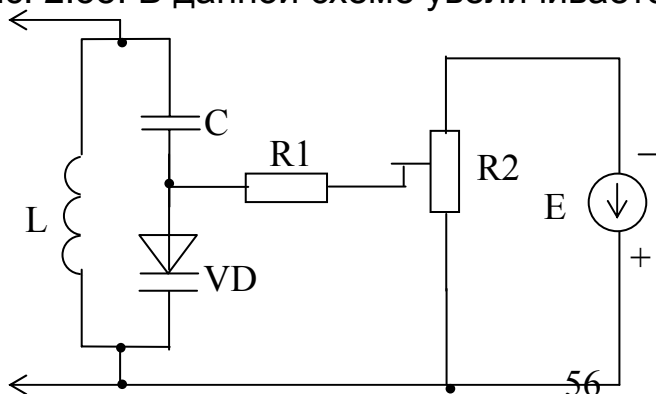


Рис. 2.37

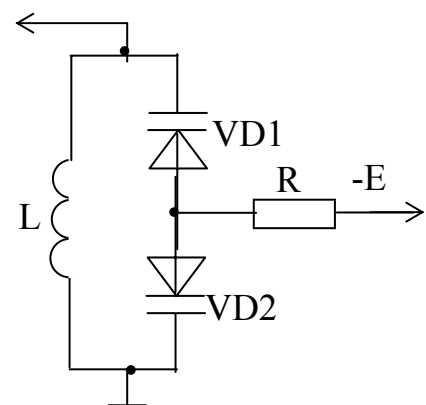


Рис. 2.38

емкости (диапазон перестройки) и улучшается линейность ВФХ.

2.4. Параметрические диоды

Параметрический диод – это варикап СВЧ-диапазона, предназначенный для усиления и генерирования сигналов.

Этот прибор относится к числу активных диодов. В параметрическом усилителе в колебательный контур включается варикап (параметрический диод). Параметр диода (емкость) дважды за период колебаний сигнала изменяется в нужной фазе с помощью внешнего вспомогательного генератора накачки. В результате достигается эффект параметрического усиления. С энергетической точки зрения усиление происходит за счет преобразования энергии генератора накачки при помощи нелинейной переменной емкости в энергию сигнала. Такое преобразование эквивалентно внесению в колебательный контур отрицательного сопротивления, которое компенсирует потери в нем. Основное преимущество таких усилителей – чрезвычайно высокая чувствительность вследствие малого уровня собственных шумов.

Принцип параметрического усиления основан на способности конденсатора изменять запасенную в нем энергию при изменении емкости. Так, если приложенному напряжению U_1 соответствует величина емкости C_1 , то накопленный заряд $q = C_1 U_1$, что соответствует запасенной энергии

$$W_1 = \frac{C_1 U_1^2}{2} = \frac{q^2}{2C_1}.$$

При резком уменьшении емкости $C_2 = C_1 - \Delta C$ в определенный момент времени переменного сигнала величина накопленного заряда не изменится, тогда напряжение на варикапе мгновенно увеличится до $U_2 = q/C_2$, что приведет к увеличению запасенной энергии $W_2 = q^2/(2C_2)$. Если в дальнейшем увеличить емкость в тот момент времени, когда переменное напряжение проходит через нулевое значение, то такое изменение не приведет к отбору энергии из контура. Таким образом, уменьшая емкость в момент максимума напряжения и увеличивая ее в момент прохождения напряжения через нуль (рис. 2.39), можно увеличивать ампли-

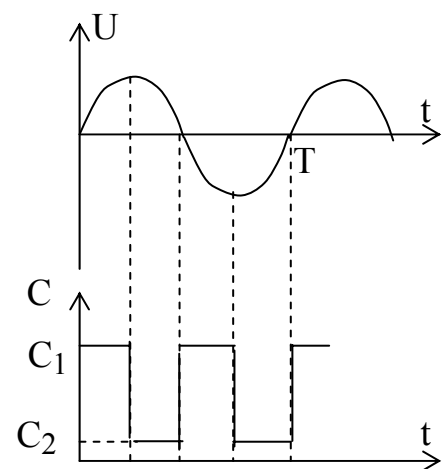


Рис. 2.39

туда переменного напряжения на контуре, т.е. осуществить усиление сигнала. Изменение емкости диода производится с помощью вспомогательного генератора накачки, причем его частота в два раза выше частоты сигнала. Время переключения емкости должно составлять сотые доли периода. Например, при частоте усиливаемого сигнала $1 \text{ ГГц} = 10^9 \text{ Гц}$ время переключения $T = 10^{-11} \text{ с} = 10 \text{ пс}$. Поэтому СВЧ-варикапы выделяют в особый класс быстродействующих приборов – параметрические диоды. Диоды изготавливаются на основе GaAs, Ge.

Подкласс ВЧ- и СВЧ-диодов при маркировке определяют буквой А с последующим цифровым элементом 4. Например 1А401А – германиевый параметрический диод, применяемый на длине волны, меньшей 3 см. Емкость перехода составляет 0,11 пФ при величине постоянной времени 0,85 пс.

Основные параметры параметрических диодов соответствуют параметрам варикапов. Дополнительно вводится величина постоянной времени – динамический параметр, определяющий их быстродействие. Основной предельный эксплуатационный параметр – рассеиваемая на диоде мощность.

2.5. Туннельные диоды (Езаки)

Туннельный диод – это диод на основе вырожденного полупроводника, в котором туннельный эффект приводит к появлению на прямой ветви участка с отрицательным динамическим сопротивлением.

Условное графическое обозначение диода показано на рис. 2.40. Прибор работает при прямом смещении перехода, полярность этого напряжения также указана на рисунке. Для изготовления этих диодов

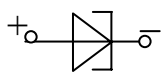


Рис. 2.40

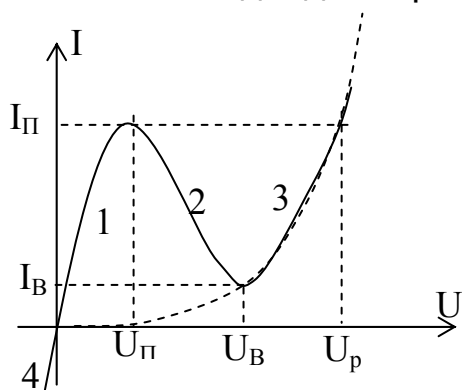


Рис. 2.41

применяются полупроводники с высокой концентрацией примесей в *p*- и *n*-областях. Вследствие сильной степени легирования переход имеет малую толщину, примерно на два порядка меньшую, чем в обычных диодах. Кроме того, в вырожденных полупроводниках изменяется структура энергетической диаграммы. В результате квантово-механического взаимодействия атомов примеси примесные уровни расщепляются в примесную зону, которая сливается с валентной зоной *p*-области и с зоной проводимости *n*-области.

Вольт-амперная характеристика диода показана на рис. 2.41. Эта зави-

симость имеет две особенности. В диапазоне прямых напряжений $U_{\Pi} < U < U_e$ («падающий» участок – 2) диод имеет отрицательное динамическое сопротивление, когда с увеличением напряжения ток уменьшается. Кроме того, этот диод не обладает свойством односторонней проводимости. При изменении полярности напряжения (участок 4) обратный ток имеет такую же величину, как и прямой (участок 1).

Для сравнения на рисунке пунктиром показана ВАХ диода, соответствующая обычному p - n -переходу. Превышение прямого тока для туннельного диода на участках 1,2 над диффузионным током свидетельствует о наличии дополнительного механизма токопереноса – туннельного тока.

Обоснование ВАХ туннельного диода может быть проведено на основании энергетической диаграммы перехода (рис. 2.42).

В исходном состоянии, при нулевом смещении перехода (рис. 2.42, а), происходит перекрытие валентной зоны p -области и зоны проводимости n -области как следствие сильного легирования этой структуры. Заштрихованным областям соответствуют уровни энергии, которые заняты электронами с наибольшей вероятностью. В незаштрихованной области «потолка» валентной зоны p -области ($\Delta\epsilon_v$) велика концентрация дырок, поэтому много вакантных энергетических уровней для электронов. Туннельный ток создается электронами, которые просачиваются сквозь потенциальный барьер без изменения их энергии. При отсутствии внешнего напряжения небольшие туннельные потоки электронов направлены навстречу друг другу и равны по величине, поэтому суммарный ток через переход равен нулю. При прямом напряжении уменьшается высота потенциального барьера ($\Delta\epsilon_b$), поднимается «дно» зоны проводимости n -области (ϵ_c) относительно p -области. В случае $U = U_{\Pi}$ наблюдается максимальное перекрытие соответствующих энергетических областей (рис. 2.42, б), что соответствует экстремальному значению туннельного тока I_{Π} . Этот ток создается туннелированием электронов (основных

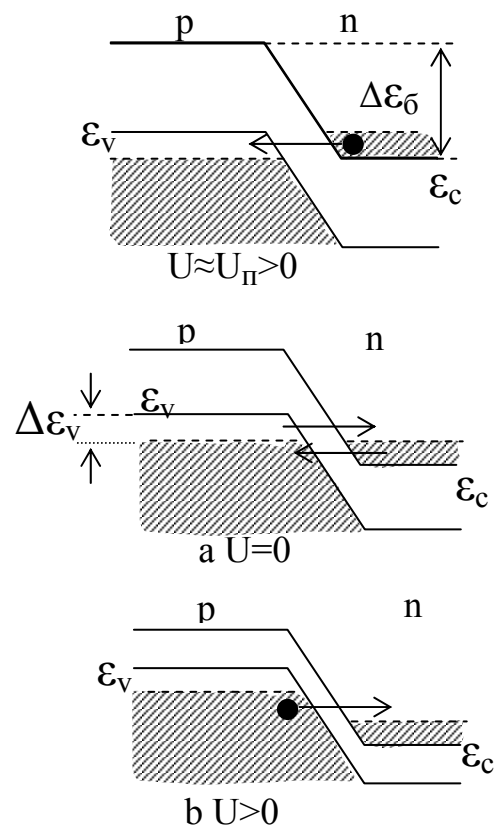


Рис. 2.42

носителей заряда) из n -области в p -область. При дальнейшем увеличении прямого напряжения уменьшается перекрытие зон, туннельный ток будет уменьшаться, что приведет к появлению участка 2 на ВАХ. Если $U > U_{\text{в}}$, то туннельный ток отсутствует, на участке 3 ВАХ протекает прямой диффузионный ток перехода.

Обратное смещение (рис. 2.42, в) изменяет направление туннельного тока. Электроны из валентной зоны p -области туннелируют на свободные энергетические уровни зоны проводимости n -области.

Основные параметры туннельных диодов целесообразно разбить на следующие группы:

1. Статические параметры, отражающие геометрические особенности ВАХ (см. рис. 2.41):

- ток пика $I_{\text{п}}$ - прямой ток в точке максимума характеристики (от десятых долей до десятков миллиампер);
- ток впадины $I_{\text{в}}$ - прямой ток в точке минимума ВАХ;
- отношение токов пика и впадины $I_{\text{п}}/I_{\text{в}}$ косвенно определяет крутизну падающего участка ВАХ (Ge - $I_{\text{п}}/I_{\text{в}} = 4...8$, GaAs - $I_{\text{п}}/I_{\text{в}} = 6...10$);
- напряжение пика $U_{\text{п}}$ - прямое напряжение, соответствующее пиковому току (Ge - $U_{\text{п}} = 60...100$ мВ, GaAs - $U_{\text{п}} = 160...300$ мВ);
- напряжение впадины $U_{\text{в}}$ - прямое напряжение, соответствующее току впадины (GaAs - $U_{\text{в}} = 350...600$ мВ, Ge - $U_{\text{в}} = 250...350$ мВ);
- напряжение раствора $U_{\text{р}}$ - прямое напряжение большее $U_{\text{в}}$, при котором ток равен пиковому (Ge - $U_{\text{р}} = 400...630$ В, GaAs - $U_{\text{р}} = 0,65...1,3$ В).

2. Динамические параметры:

- резонансная частота диода f_0 - это частота, на которой реактивная составляющая полного сопротивления диода обращается в нуль;
- предельная резистивная частота $f_{\text{пр}}$ - это частота, на которой активная составляющая полного сопротивления диода обращается в нуль;
- емкость диода $C_{\text{Б}}$;
- отрицательное сопротивление $R_{\text{д}}$.

3. Предельные эксплуатационные параметры:

- максимальное значение прямого тока $I_{пр. макс}$ (или $U_{пр. макс}$);
- максимальное значение обратного тока $I_{обр. макс}$;
- максимальная рассеиваемая мощность $P_{макс}$.

Туннельный диод – это прибор СВЧ-диапазона, используемый в частотном интервале от единиц до нескольких десятков гигагерц. Эквивалентная схема диода для малых сигналов на участке отрицательного динамического сопротивления приведена на рис. 2.43. Она состоит из отрицательного динамического сопротивления R_d , барьерной емкости C_b , общего сопротивления потерь r и индуктивности выводов L . Полное сопротивление диода

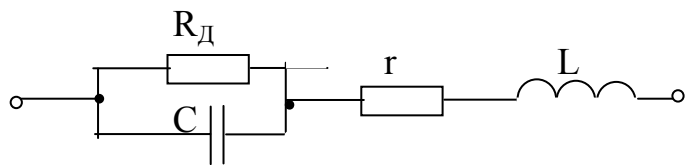


Рис. 2.43

$$z = r + j\omega L + \frac{1}{g + j\omega C},$$

где дифференциальная проводимость $g = 1/R_d$.

Выделяя активную и реактивную компоненты, получим

$$z = r + \frac{R_d}{1 + \omega^2 C^2 R_d^2} + j \left(\omega L - \frac{\omega C R_d^2}{1 + \omega^2 C^2 R_d^2} \right).$$

Приравнивая к нулю действительную часть полного сопротивления, определяем предельную частоту:

$$\omega_{пр} = \frac{1}{|R_d|C} \sqrt{\frac{|R_d|}{r} - 1}. \quad (2.15)$$

Диод проявляет себя как элемент с отрицательным динамическим сопротивлением пока $\omega < \omega_{пр}$. В этом случае компенсируются потери и возможен процесс усиления или генерирования колебаний.

Резонансную частоту диода получаем, приравнивая к нулю мнимую часть полного сопротивления, т.е.

$$\omega_0 = \frac{1}{|R_d|C} \sqrt{\frac{|R_d|^2 C}{L} - 1}. \quad (2.16)$$

В режиме усиления не должно происходить самовозбуждения, т.е. резонансная частота должна находиться за предельной частотой $\omega_0 > \omega_{пр}$. Из этого неравенства находим условие усиления:

$$L < |R_{\text{д}}|Cr. \quad (2.17,а)$$

Для возникновения самовозбуждения, т.е. генерирования колебаний, справедливо обратное условие:

$$L > |R_{\text{д}}|Cr. \quad (2.17,б)$$

В усилительных диодах для уменьшения индуктивности выводы выполняются в виде пластин, мембран и т.д.

Генераторные диоды должны обеспечивать максимальную мощность в нагрузке, которая определяется следующим соотношением:

$$P = 0,12(U_{\text{в}} - U_{\text{п}})(I_{\text{п}} - I_{\text{в}}).$$

Туннельные диоды изготавливаются на основе германия или арсенида галлия. Разность напряжений $U_{\text{в}} - U_{\text{п}}$ больше у материала с более широкой запрещенной зоной. Поэтому арсенид-галлиевые диоды обеспечивают получение большей мощности. Большее соотношение $I_{\text{п}}/I_{\text{в}}$ этих диодов также способствует увеличению мощности.

Форма ВАХ определяет применение туннельных диодов в быстродействующих импульсных и переключательных схемах. Различие двух уровней сигнала обеспечивают параметры $I_{\text{п}}/I_{\text{в}}$, а также $U_{\text{п}}, U_{\text{в}}, U_{\text{р}}$. При этом линия нагрузки должна пересекать ВАХ на трех участках прямой ветви диода, тогда нужно выбирать $R_{\text{н}} > R_{\text{д}}$. При работе на падающем участке характеристики (режим усиления и генерирования) $R_{\text{н}} < R_{\text{д}}$, поэтому нагрузочный резистор не включают, его роль выполняет сопротивление потерь диода.

Подкласс туннельных диодов при их маркировке отмечают элементом – буквой И. Третий элемент маркировки диода (цифра) определяет функциональное назначение: 1 - усилительный (1И104А), 2 - генераторный (3И203А), 3 – переключательный (АИ301А).

2.6. Обращенные диоды

Обращенный диод – это разновидность туннельного диода, проводимость которого при обратном напряжении больше, чем при прямом.

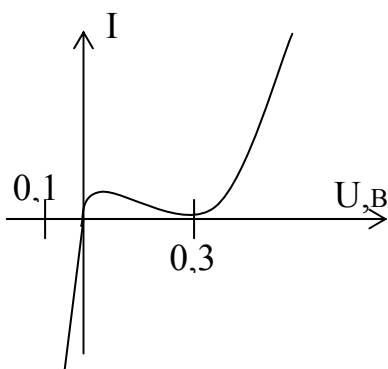


Рис. 2.44

ВАХ этого диода является обратной (перевернутой) по отношению к характеристике классического $p-n$ -перехода и занимает меньший интервал напряжений (рис. 2.44). Обратная ветвь $0 < U_{\text{обр}} < 0,1$ В соответствует проводящему состоянию диода, а участок прямой ветви $U < 0,3$ В – непроводящему. В интервале напряжений

$-0,1 < U < 0,3$ В диод обладает свойством односторонней проводимости в обратном проявлении.

Условное обозначение диода показано на рис. 2.45. Здесь же указана полярность напряжения, соответствующая проводящему состоянию (ток направлен против стрелки диода). В маркировке диода 1И401А использованы следующие основные элементы:

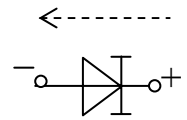


Рис. 2.45

И – подкласс туннельных диодов;

4 – обращенный диод (как разновидность туннельного).

Принцип действия прибора основан на туннельном эффекте в обратносмещенном переходе, выполненном из полупроводника, находящегося на границе вырождения. В отличие от туннельного диода степень легирования здесь ниже. На энергетической диаграмме в исходном состоянии нет перекрытия зоны проводимости n -области и валентной зоны p -области. Поэтому туннельный ток появляется только при обратном смещении, когда возникает перекрытие соответствующих зон. Этот ток создается электронами, туннелирующими из валентной зоны p -области в зону проводимости n -области. При прямых напряжениях туннельный ток отсутствует, так как нет взаимного соответствия энергетических уровней в p - и n -областях структуры. Для больших значений прямых напряжений появляются условия возникновения прямого инжекционного тока. Основные носители заряда преодолевают потенциальный барьер и диффундируют через переход. Однако этот участок прямой ветви не является рабочим.

Для практической реализации обращенного диода нужно фактически уменьшить концентрацию примесей только в n -области (при сохранении критической в p -области). Тогда прямой туннельный ток уменьшится из-за уменьшения концентрации электронов в n -области. Значение пикового тока тоже резко уменьшится и на начальном участке ВАХ прямой ветви останется лишь небольшая неравномерность.

Основные параметры обращенных диодов:

1. Постоянное прямое напряжение (при заданной малой величине прямого тока) $U_{пр}$ ограничивает рабочий участок непроводящей прямой ветви (Ge – 0,35 В, GaAS – 0,6 В).

2. Постоянное обратное напряжение (при заданном обратном токе) определяет падение напряжения для проводящего состояния $U_{обр}$ (около 0,1 В);

3. Постоянный обратный ток является предельным параметром, ограничивающим рабочий участок проводящего состояния $I_{обр. макс}$ (1...8 мА).

4. Постоянный прямой ток является предельным параметром для непроводящего состояния $I_{пр. макс}$ (0,3...0,8 мА).

5. Емкость диода определяет инерционные свойства C (единицы пикофарад).

Из анализа параметров следуют основные преимущества обращенных диодов: малое падение напряжения делает их незаменимыми при обработке слабых сигналов, высокое быстродействие диода определяет малая величина его емкости, так как проводящая обратная ветвь диода соответствует туннельному току, поэтому отсутствует эффект накопления неосновных носителей ($C_{диф} = 0$), а туннелирование – это малоинерционный процесс. Следует обратить внимание также на малую величину максимально допустимого напряжения для непроводящего состояния.

Диоды применяются в быстродействующих импульсных схемах, а также для нелинейных преобразований сигналов СВЧ-диапазона (при детектировании и преобразовании частоты).

2.7. Диоды Шотки

Диод Шотки – это диод, выпрямительное свойство которого основано на использовании электрического перехода металл-полупроводник.

Электрический переход между металлом и полупроводником может обладать свойством односторонней проводимости (выпрямляющий переход) или подчиняться линейному закону Ома (омический переход). Омический переход применяется для формирования внешних выводов от полупроводниковых областей.

В рассматриваемом приборе используется электрический переход Шотки. Этот переход возникает между металлом и полупроводником, если они имеют разные величины работ выхода электронов. Для полупроводника n -типа работа выхода из него должна быть меньше, чем работа выхода из металла, т.е. $A_n < A_m$ (p -тип полупроводника удовлетворяет обратному неравенству $A_n > A_m$).

Процесс образования перехода Шотки и его энергетическая диаграмма в равновесном состоянии показаны на рис. 2.46, где \mathcal{E}_F - общий энергетический уровень Ферми для данной структуры, \mathcal{E}_C - нижняя граница зоны проводимости, \mathcal{E}_V - верхняя граница валентной зоны полупроводника. Двойной штриховкой отмечены энергетические состояния, занятые электронами.

Этот процесс состоит из следующих этапов:

1. В начальный момент времени поток электронов из полупроводника в металл больше, чем в обратном направлении, так как $A_{\text{П}} < A_{\text{М}}$.

2. Движение электронов вызывает перераспределение зарядов в приграничных областях, что приводит к нарушению электрической нейтральности. В полупроводниковой области «обнажается» нескомпенсированный заряд положительных ионов донорной примеси, а в тонком приповерхностном слое металла появляется избыточный отрицательный заряд электронов.

3. Между противоположными зарядами в приконтактной области возникает внутреннее электрическое поле. Из-за высокой концентрации электронов в металле это поле практически не проникает в него. Поэтому ширина перехода d определяется слоем положительных ионов.

4. Поле перехода представляет собой потенциальный барьер для электронов полупроводника. Только небольшая часть высокоэнергетичных электронов способна преодолеть этот барьер и перейти в область металла.

5. С уменьшением электронного потока из полупроводника устанавливается динамическое равновесие между двумя разнонаправленными потоками электронов. В структуре устанавливается равновесное состояние, которое характеризуется толщиной перехода d , напряженностью поля E и высотой энергетического $q\phi_K$ или потенциального барьера

$$\phi_K = \frac{A_{\text{М}} - A_{\text{П}}}{q} \quad (2.18)$$

Эта величина еще называется контактной разностью потенциалов.

Наличие контактной разности потенциалов создает изгиб зон в приграничной области полупроводника. Чтобы электрону попасть в эту область, ему нужно преодолеть силу электрического противодействия. Поэтому чем ближе к границе контакта, тем большей энергией должен обладать электрон, чтобы попасть в эту область (см. изгиб зоны в сторону больших энергий).

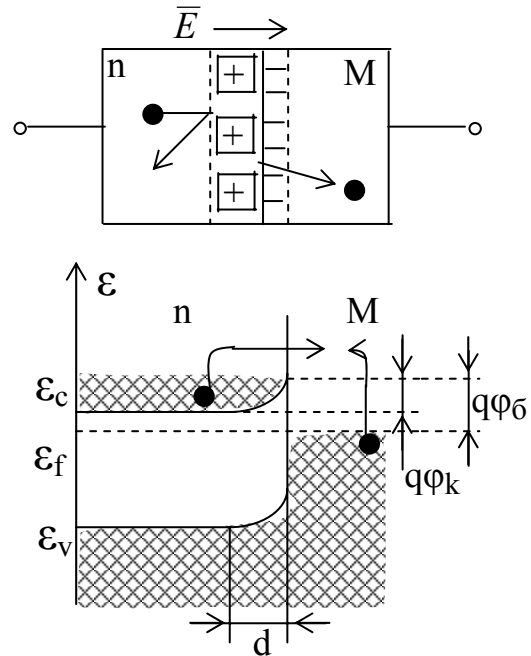


Рис. 2.46

Электроны из металла и полупроводника должны преодолевать разную высоту барьера при переходе через границу. Величина $q\phi_K$ – это высота энергетического барьера для электронов, переходящих из полупроводника в металл, а $q\phi_B$ – для электронов, переходящих в обратном направлении.

Прямое смещение на переходе уменьшает потенциальный барьер для электронов полупроводника, что приводит к увеличению этой компоненты тока (рис. 2.47). Встречный поток электронов из металла

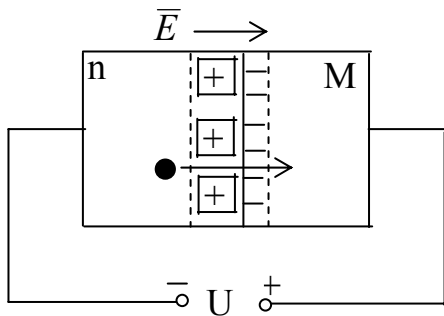


Рис. 2.47

при этом остается прежним, так как металл экранирует поле перехода и высота барьера для этих электронов не изменяется. Следовательно, прямой ток создается потоком электронов через переход, переходящих из полупроводника в металл. На энергетической диаграмме прямосмещенного перехода поднимаются энергетические уровни полупроводника ϵ_F и ϵ_C на величину qU , что приводит к уменьшению соответствующего энергетического барьера (рис. 2.48), т.е. $q(\phi_K - U)$. При обратном смещении

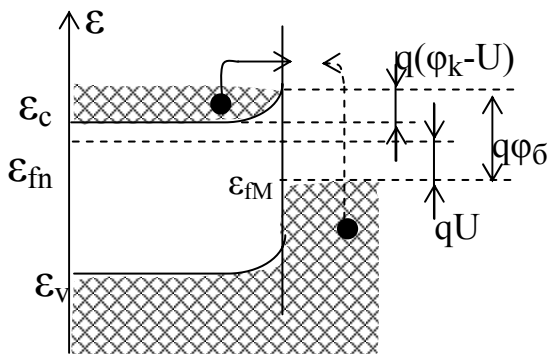


Рис. 2.48

перехода остается только компонента малого электронного потока из металла в полупроводник, что связано с увеличением барьера для прямого потока. Обратный ток является током насыщения, так как не зависит от величины обратного напряжения.

ВАХ диода формально описывается зависимостью, аналогичной p - n -переходу

$$I = I_s \left(e^{\frac{U}{\Phi T}} - 1 \right), \quad (2.19)$$

причем экспоненциальный характер сохраняется в очень широком диапазоне изменения тока. В реальных переходах с учетом дополнительных эффектов нарушается постоянство тока I_s при обратных смещениях.

Для диодов Шотки можно привести следующие отличительные особенности:

1. Отсутствует диффузионная емкость, связанная с процессами накопления и рассасывания неосновных носителей в базовой области. Поэтому высокое быстродействие определяется только малым временем перезаряда барьерной емкости.

2. Величины прямого падения напряжения меньше, чем в кремниевых *p-n*-переходах. Типовые значения составляют 0,4...0,5 В.

3. Способность пропускать большие величины прямых токов. Это связано с хорошим теплоотводом от металлической области структуры и малым прямым напряжением.

4. Сравнительно небольшие величины максимально допустимых обратных напряжений (несколько десятков вольт). Это определяется краевыми эффектами при пробое. Кроме того, для арсенида галлия не удается получить малую концентрацию дефектов, что также уменьшает пробивное напряжение.

Диоды изготавливают на основе кремния или арсенида галлия с электропроводностью *n*-типа из-за большей подвижности носителей заряда. Условное обозначение диода приведено на рис. 2.49. Маркировка диода определяется областью его применения.

Диоды Шотки предназначены для быстродействующих импульсных схем, выпрямительных устройств и нелинейных преобразований сигнала в ВЧ- и СВЧ-диапазонах.

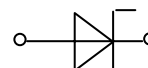


Рис. 2.49

2.8. P-i-n-диоды

P-i-n-диод – это диод с трехслойной структурой, состоящей из высокоомного слоя собственного полупроводника, расположенного между двумя сильнолегированными областями *p*- и *n*-типов (рис. 2.50).

Область полупроводника с собственной электропроводностью значительно увеличивает толщину *p-n*-перехода d , что позволяет получить малые емкости диода. При этом одновременно повышается пробивное напряжение. Напряженность поля внутри перехода уменьшается при той же величине разности потенциалов (см. рис. 2.50):

$$E = \Delta\phi/\Delta x.$$

Сопротивление обедненной области обычно велико (порядка нескольких десятков килоом), так как наличие внутреннего поля еще более обедняет *i*-слой, т.е. число собственных носителей заряда уменьшается. Такие диоды еще называют регулируе-

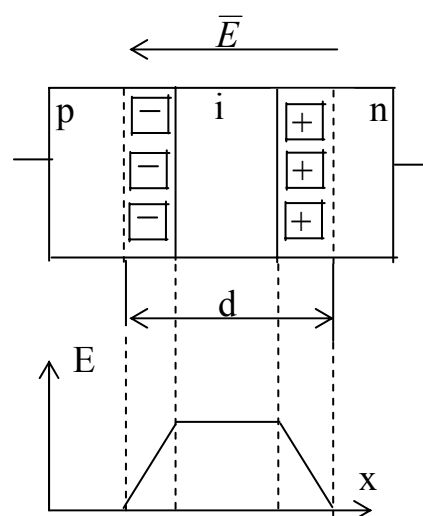


Рис. 2.50

мыми резистивными диодами, так как его сопротивление в значительной мере определяется концентрацией зарядов в i -области. Величина сопротивления может меняться в широких пределах в зависимости от протекающего через диод тока.

Внешнее напряжение, приложенное в прямом направлении, уменьшает высоту потенциального барьера для основных носителей заряда. В результате чего происходит инжекция дырок из p -области и электронов из n -области внутрь слоя i . Инжектированные избыточные носители заряда диффундируют навстречу друг другу в промежуточной области и частично рекомбинируют друг с другом. Оставшаяся часть инжектированных зарядов переходит в соответствующие p - и n -области, где завершается второй более быстрый этап рекомбинации. Количество зарядов, накопленных в i -области, определяется величиной тока инжекции и временем жизни неравновесных носителей, а проводимость прибора - плотностью зарядов в промежуточном слое.

Следует отметить, что прямой ток для $p-i-n$ -структуры меньше, чем для $p-n$ -перехода. Это связано с тем, что на $p-i$ - и $i-n$ -переходах падает только часть приложенного напряжения. Другая часть напряжения будет падать на высокоомной области i -слоя вследствие малой концентрации носителей заряда в нем, особенно при малых прямых токах. Поэтому вентильное свойство в таких структурах выражено слабее.

При использовании $p-i-n$ -диода в аналоговых устройствах ток изменяется от значений порядка 10 мА при прямых смещениях до значений, близких к нулю, при обратных.

Под действием обратного напряжения из i -слоя почти полностью удаляются носители заряда и через него будет протекать только небольшой ток утечки. Диоды изготавливаются на исходном кремниевом кристалле, что еще более минимизирует величину обратного тока.

Диоды с $p-i-n$ -структурой в основном применяют в схемах переключения и аналоговых устройствах управления ВЧ- и СВЧ-сигналов.

Работа диода в режиме переключения основана на резком отличии концентрации свободных носителей заряда в i -области при прямом и обратном смещениях. Высокое пробивное напряжение позволяет коммутировать большие уровни СВЧ-мощности. Быстродействие процесса ограничено временем рассасывания носителей заряда из i -слоя при переключении с прямого напряжения на обратное. Предельные величины переключаемой импульсной мощности составляют сотни киловатт до частоты порядка 40 ГГц. При параллельном включении диода в линию передачи он имеет при прямом смещении малое внутреннее сопротивление, которое шунтирует линию. Тогда ос-

новная часть СВЧ-мощности отражается, а на диоде рассеивается ее незначительная часть.

В аналоговом режиме $p-i-n$ -диоды на высоких частотах применяются в качестве сопротивления, близкого к активному, величиной которого можно управлять в широких пределах в зависимости от величины тока, протекающего через него. Нижняя частота, определяющая применимость прибора в этом режиме, связана с временем жизни приближенным соотношением $f_0 = (2\pi\tau)$. На низких частотах ($f < f_0$) диод обладает выпрямительными свойствами и подобен обычному $p-n$ -переходу. В окрестности граничной частоты ($f = f_0$) он эквивалентен линейному активному сопротивлению с небольшой нелинейной составляющей. На высоких частотах ($f \gg f_0$) диод ведет себя как линейное активное сопротивление, величиной которого можно управлять постоянным током или низкочастотным управляющим сигналом. При действии источника смещения в прямом направлении переменный сигнал накладывается на ток, обусловленный смещением. Изменяя напряжение смещения, можно управлять переменным сигналом (модулировать его).

Прибор выпускается в корпусах различных типов, в том числе в бескорпусном варианте. Диоды с балочными выводами предназначены для монтажа в микрополосковые схемы методом термокомпрессии или сварки. Полосковые корпуса содержат встроенные цепи согласования и применяются в широкополосных схемах до частот 10 ГГц. В корпусе прибора может содержаться либо одиночная структура ($p-i-n$) – 2A503A, либо пара с общим анодом ($n-i-p-i-n$) – 2A505A.

Основные параметры диодов:

1. Емкость диода C (0,2...1) пФ.
2. Время обратного восстановления $T_{обр}$ (60...2 мкс, для СВЧ-диодов уменьшается до 5...100 нс).
3. Прямое сопротивление потерь при заданном токе определяет активную составляющую полного сопротивления диода (единицы ом).
4. Максимально допустимая величина обратного напряжения предотвращает переход диода в режим пробоя (100...200) В.
5. Максимальная величина рассеиваемой мощности (от единиц до десятков ватт).
6. Максимальная величина прямого тока (100...500) мА.

Три последних параметра относятся к предельно допустимым, первые два определяют инерционные свойства, причем за время $T_{обр}$ происходит рассасывание зарядов из i -области, т.е. восстанавливается обратное сопротивление. Время прямого восстановления на один-

два порядка меньше. Эффективность переключательного диода определяет обобщенный параметр – критическая частота

$$f_k = \frac{1}{2\pi C \sqrt{r_{пр} r_{обр}}}, \quad (2.20)$$

где $r_{пр}$, $r_{обр}$ – сопротивления потерь диода соответственно в прямом и обратном смещениях.

2.9. Диоды СВЧ-диапазона

СВЧ-диод – это диод, предназначенный для нелинейных преобразований сверхвысокочастотных сигналов. Эти приборы работают в диапазоне частот от 300 МГц до 300 ГГц, что соответствует интервалу длин волн от одного метра до одного миллиметра.

В зависимости от вида нелинейного преобразования, т.е. функционального назначения, СВЧ диоды делятся на следующие типы: детекторные, модуляторные, смесительные, умножительные, ограничительные, переключательные. Диоды со специфическими принципами действия относятся к особым типам: туннельные и обращенные диоды, параметрические, лавинно-пролетные, диоды Ганна.

Детекторный диод предназначен для детектирования сигнала, т.е. выделения сигнала более низкой частоты из модулированного по амплитуде высокочастотного колебания. Полезный эффект детектирования определяется чувствительностью по току: S_i – это отношение приращения выпрямленного тока к поданной на диод СВЧ-мощности, вызвавшей это приращение. Наибольшая чувствительность соответствует прямому току смещения в несколько десятков микроампер. Коэффициент качества (добротность) является обобщенным параметром диода

$$K = \frac{S_i r_d}{\sqrt{n_{ш} r_d + r_{ш}}}, \quad (2.21)$$

где r_d – дифференциальное сопротивление при заданном токе, $r_{ш}$ – эквивалентное шумовое сопротивление (примерно 1 кОм), $n_{ш}$ – шумовое отношение, или коэффициент шума. Хорошие диоды имеют $K > 100 \text{ Вт}^{-1/2}$. Коэффициент стоячей волны по напряжению (КСВН) определяет степень согласования диодного блока с волновым сопротивлением тракта (составляет несколько единиц). Чем меньше этот коэффициент, тем лучше согласование, т.е. меньше потери сигнала. Основной предельный эксплуатационный параметр – величина рассеиваемой мощности (десятки милливатт).

Смесительный диод – это диод, предназначенный для преобразования высокочастотного сигнала в сигнал более низкой промежуточной (разностной) частоты. Преобразование осуществляется с помощью вспомогательного сигнала – напряжения гетеродина.

Основной параметр диода, определяющий эффективность преобразования высокочастотного сигнала в колебание промежуточной частоты, – потери преобразования, выраженные в децибелах

$$L = 20 \lg \frac{P_{\text{ВХ}}}{P_{\text{ПЧ}}},$$

где $P_{\text{вх}}$ – мощность СВЧ-сигнала, подаваемого на диод, $P_{\text{пч}}$ – мощность сигнала промежуточной частоты. Возможность различать полезный сигнал на фоне шумов определяет такие шумовые параметры диода:

1. Шумовое отношение $n_{\text{ш}}$ – отношение номинальной мощности шумов диода в рабочем режиме к номинальной мощности тепловых шумов соответствующего активного сопротивления.

2. Коэффициент шума – отношение сигнал-шум на входе и выходе:

$$F = (P_c/P_{\text{ш}})_{\text{ВХ}} / (P_c/P_{\text{ш}})_{\text{ВЫХ}}. \quad (2.22)$$

3. Нормированный коэффициент шума как обобщенный параметр диода:

$$F_{\text{н}} = L + 10 \lg (n_{\text{ш}} + F_{\text{упч}} - 1), \quad (2.23)$$

где $F_{\text{упч}}$ – коэффициент шума усилителя промежуточной частоты, равный 1,5 дБ.

Вспомогательным параметром диода является выпрямленный ток (постоянная составляющая), протекающий в выходной цепи диода в рабочем режиме. Он характеризует потери преобразования: чем больше этот ток, тем меньше потери. Этот параметр используется для контроля исправности диода при подаче определенной СВЧ-мощности.

Модуляторные диоды предназначены для модуляции амплитуды переменного тока, а умножительные диоды (варакторные) – для умножения частоты. Принцип действия умножительных диодов основан на нелинейной зависимости барьерной емкости от обратного напряжения.

СВЧ-диоды реализуются на основе p - n -перехода, перехода Шотки, p - i - n -структур. Отсутствие эффекта накопления в диодах Шотки обеспечивает лучшие частотные характеристики и более высокую скорость переключения, чем диоды с p - n -переходами. Полная емкость диода составляет 1...2 пФ и уменьшается до 0,1...0,04 пФ для более

высокочастотных диодов. Диоды Шотки применяют в смесителях, детекторах, ограничителях, быстродействующих переключателях. В смесителях они обеспечивают лучшие шумовые характеристики.

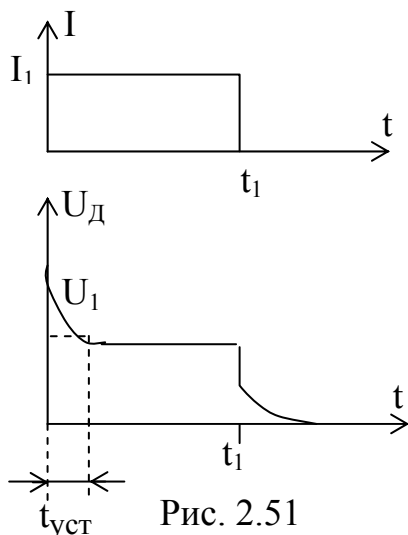
В СВЧ-диодах на основе $p-n$ -перехода минимизируют время жизни неосновных носителей в базе, уменьшают барьерную емкость перехода, а также сопротивление базы для уменьшения потерь мощности в диоде. Для увеличения удельной проводимости базы применяют материал с высокой подвижностью носителей заряда, например, арсенид галлия с n -типом проводимости. Диоды имеют также малые геометрические размеры перехода для уменьшения барьерной емкости.

СВЧ-диоды имеют отличительное конструктивное исполнение. Корпус диода обладает минимальной собственной емкостью и индуктивностью. Металлические части покрывают тонким слоем серебра или золота для уменьшения сопротивления контакта и повышения коррозионной стойкости. В деци- и сантиметровом диапазоне волн (до частот 12 ГГц) применяют корпус патронного типа. В конструкции коаксиального типа отсутствуют выступающие детали корпуса, которые применяют до частот около 30 ГГц. Для более коротковолнового диапазона используется конструкция в виде волноводной вставки. В микроразмерных линиях применяют бескорпусные диоды.

Для обозначения подкласса СВЧ-диодов при маркировке указывают букву А, стоящую на второй позиции, последующие цифровые позиции определяют функциональное назначение прибора. Например, 3А110А – смесительный, 2А201А – детекторный, 1А401А – параметрический, 2А505А – переключательный, АА603А – умножительный.

2.10. Импульсные диоды

Импульсный диод – это диод, имеющий малую длительность переходных процессов и предназначенный для применения в импульсных режимах работы.



Переходные процессы в диоде возникают при включении или переключении его с прямого смещения на обратное. Физическими причинами инерционности являются: эффект накопления и рассасывания неравновесных носителей заряда в базе (эффект диффузионной емкости), а также перезаряд барьерной емкости перехода.

Переходный процесс при прохождении через диод импульса прямого тока большой амплитуды показан на рис. 2.51. В

первый момент на диоде падает относительно большое напряжение U_1 , которое постепенно уменьшается до установившегося значения. Этот переходной процесс включения связан с эффектом модуляции проводимости базы диода. По мере накопления избыточных носителей заряда в базе при инжекции уменьшается ее объемное сопротивление, что приводит к уменьшению падения напряжения. Время установления прямого напряжения $t_{yсм}$ – это интервал времени от начала импульса прямого тока до момента достижения заданного значения прямого напряжения на диоде ($1,2U_{yсм}$).

С момента окончания импульса прямого тока t_1 происходит медленное уменьшение напряжения на диоде, что соответствует этапу послеинжекционной ЭДС. Концентрация избыточных носителей заряда в базе не может измениться мгновенно, поэтому напряжение уменьшается по мере рекомбинации зарядов в базе. Скачок напряжения в момент времени t_1 обусловлен исчезновением падения напряжения на объемном сопротивлении базы при нулевом значении тока.

Переходный процесс переключения диода приведен на рис. 2.52.

Диод управляется генератором ЭДС - E с малым внутренним сопротивлением. В момент времени t_1 происходит коммутация диода с прямого смещения на обратное. Проводящее состояние диода $t < t_1$ характеризуется постоянным прямым током I_1 , величина которого ограничивается сопротивлением источника. В этом случае в базовой области устанавливается динамическое равновесие между процессами инжекции неравновесных носителей заряда и их рекомбинацией, что определяет величину накопленного заряда

$Q_H = I_1 T$, где T - время жизни избыточных зарядов. Распределение избыточных зарядов в базе для разных моментов времени показано на рис. 2.53. Переходный процесс переключения состоит из следующих этапов:

1. Этап высокой обратной проводимости $t_1 < t < t_2$. Ток изменяет направление, так как происходит экстракция неосновных неравновесных

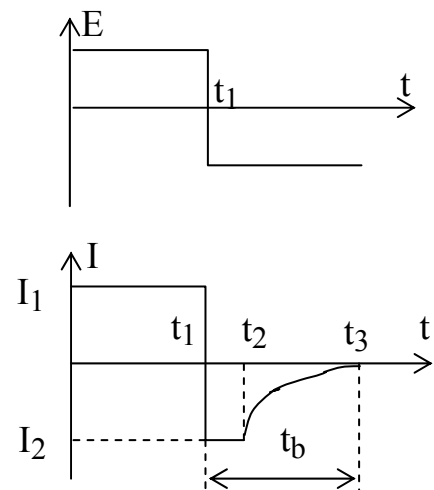


Рис. 2.52

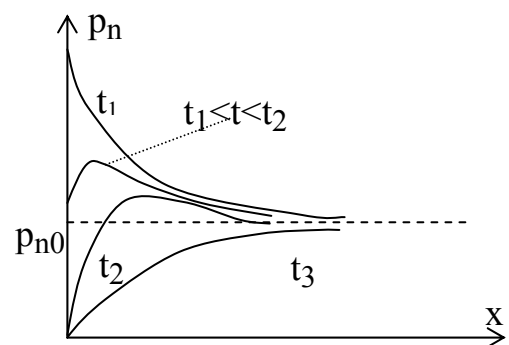


Рис. 2.53

носителей заряда через переход в эмиттерную область. В соответствии с обратным направлением тока меняется знак градиента концентраций при $X = 0$ (см. рис. 2.53). Величина тока неизменна, так как процесс рассасывания происходит при постоянном значении градиента концентраций. Значение тока определяется теми же сопротивлениями, что и при прямом смещении.

2. Этап восстановления обратного сопротивления $t > t_2$. После уменьшения концентрации избыточных зарядов до нуля вблизи перехода происходит экспоненциальное падение тока вследствие уменьшения градиента концентрации. Длительность переходного процесса t_6 определяется двумя вышеуказанными этапами. Время восстановления обратного сопротивления – это интервал времени, в течение которого обратный ток после смены полярности напряжения уменьшается до заданного значения.

Длительность плоской части импульса обратного тока определяется временем жизни неравновесных зарядов и соотношением прямого и обратного токов:

$$\Delta t = t_2 - t_1 \approx 0,2\tau (I_1/I_2).$$

Поэтому для изготовления быстродействующих диодов уменьшают время жизни, легируя базу золотом, которое выполняет роль ловушек рекомбинации. Однако с увеличением степени легирования увеличивается обратный ток диода и уменьшается напряжение пробоя.

Скорость уменьшения обратного тока при $t > t_2$ обусловлена не только рассасыванием избыточных зарядов в базе, но и перезарядом барьерной емкости. С увеличением этой емкости замедляется скорость спада тока. Поэтому для повышения быстродействия уменьшают площадь перехода. В диодах с тонкой базой снижается величина накопленного заряда, а также уменьшается время его рассасывания.

Параметры импульсных диодов могут быть сгруппированы следующим образом:

1. Динамические параметры, определяющие быстродействие диода:

- время установления прямого сопротивления T_y ;
- время восстановления обратного сопротивления T_6 ;
- емкость диода C_d (от долей до единиц пикофарад).

Параметр T_y определяет время включения диода, которое ограничивает минимальную длительность импульса. Параметр T_6 задает время переключения диода с прямого смещения на обратное, т.е. определяет минимальный период следования импульсов. В импульсных диодах Шотки отсутствует эффект накопления неосновных носителей

заряда, поэтому они имеют наименьшее время переключения. В качестве критерия быстродействия может указываться также заряд переключения. Это величина заряда, который выводится из базы для перевода диода в закрытое состояние, т.е. связан с отрицательным выбросом обратного тока. Заряд переключения меньше накопленного, так как часть избыточных зарядов рекомбинирует внутри базы.

2. Статические параметры, определяющие степень идеальности диода как прибора с односторонней проводимостью:

– постоянное прямое падение напряжения $U_{пр}$ (при заданном прямом токе);

– постоянный обратный ток $I_{обр}$ (при заданном обратном напряжении).

3. Предельные параметры, предотвращающие пробой или тепловое разрушение структуры:

– постоянное или импульсное обратное напряжение $U_{обр}$ (от десятков до единиц вольт);

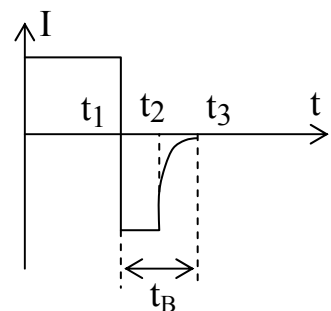
– максимальное значение прямого тока $I_{пр.макс}$ (постоянного, импульсного или среднего).

Импульсные диоды изготавливаются на основе германия, кремния и арсенида галлия. Наряду с дискретными вариантами (одиоными диодами) выпускаются наборы диодов: сборки или матрицы (КДС525А, 2ДС627А, 2Д917А). Это совокупность идентичных диодных структур, выполненных на одном кристалле, которые не соединены электрически или объединены по заданной схеме. Число диодов в такой конструкции от 4 до 16.

2.11. Диоды с накоплением заряда

Диод с накоплением заряда (ДНЗ) – это диод, принцип действия которого основан на эффекте накопления неосновных носителей заряда в базовой области.

В этом диоде в качестве рабочего участка используется импульс обратного тока, который формируется в процессе рассасывания накопленных ранее избыточных носителей заряда, причем технологически обеспечиваются условия получения короткого прямоугольного импульса с большой амплитудой. За счет внутреннего встроенного электрического поля уменьшается также длительность спада обратного тока $\Delta t = t_3 - t_2$ (рис. 2.54), появление которого связано с рассасыванием ННЗ из внутренних областей базы. Поэтому этот прибор называют также диодом



с резким восстановлением обратного сопротивления (т.е. с малым временем t_g).

Работа диода состоит из двух этапов. В течение первого ($t > t_1$), подготовительного, этапа происходит накопление зарядов в базовой области за счет инжекции. Во время второго ($t_1 < t < t_3$) - формируется короткий рабочий импульс. Встроенное электрическое поле, уменьшающее длительность обратного импульса, получают за счет неравномерного распределения примесей в базовой области. Если исходный полупроводник с n -типом проводимости с концентрацией примесей N_D легировать с одной стороны акцепторной примесью N_A , то в глубине полупроводника она распределится неравномерно

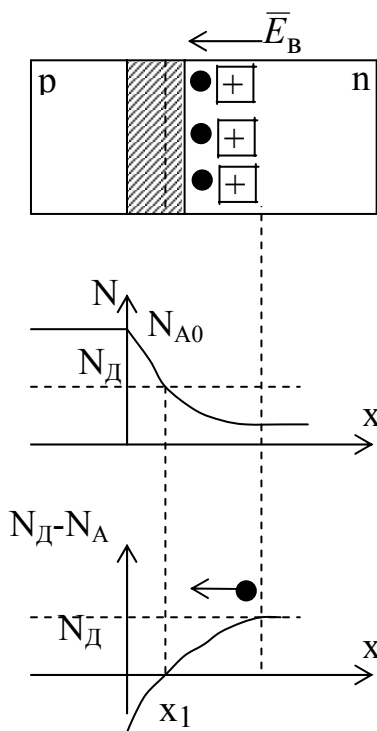


Рис. 2.55

(рис. 2.55). Координата X_1 , где $N_A(X_1) = N_D$, соответствует условной границе перехода.

При $X < X_1$ концентрация $N_A > N_D$, что соответствует дырочной проводимости, а при $X > X_1$ имеем полупроводник n -типа. Наличие градиента концентрации примесей при $X > X_1$ определяет неравномерное распределение электронов в этой части структуры. Вследствие местной диффузии электронов происходит перераспределение зарядов в приграничной области. Появляется двойной слой разноименных зарядов, состоящий из избыточных электронов с одной стороны и нескомпенсированных положительных ионов примеси с другой. Между ними возникает встроенное электрическое поле E_g . Напряженность поля определяется из условия динамического равновесия между процессами диффузии и дрейфа в этой области в установившемся состоянии:

$$j_n = qn\mu_n E + qD_n \text{grad}(n) = 0,$$

откуда

$$E = -\frac{D_n}{\mu_n} \cdot \frac{\text{grad}(n)}{n} = -\varphi_T \frac{\text{grad}(N_D)}{N_D}.$$

Это поле препятствует диффузии дырок вглубь базы во время инжекции, поэтому избыточные заряды накапливаются в приграничной области p - n -перехода. При переключении диода с прямого смещения на обратное это поле ускоряет процесс рассасывания дырок, обеспечивая их движение в сторону перехода.

ДНЗ относятся к импульсным диодам, что и определяет их основные токовые параметры этого общего класса. Кроме того, эти диоды имеют такие специфические параметры:

1. Время выключения $t_{\text{вык}}$ определяет длительность формируемого обратного импульса.

2. Заряд переключения $Q_{\text{пк}}$ – это часть накопленного заряда, которая вытекает во внешнюю цепь при изменении направления тока с прямого на обратный.

3. Эффективное время жизни неосновных носителей заряда $T_{\text{эфф}}$ определяет длительность рекомбинации избыточных неосновных носителей заряда в базе диода. Формально это время соответствует отношению заряда, переносимого переходным обратным током, к значению прямого тока $T_{\text{эфф}} = Q/I_{\text{пр}}$.

4. Максимально допустимый импульсный обратный ток $I_{\text{макс. имп}}$ является предельным параметром, а также определяет амплитуду импульса обратного тока, формируемого диодом.

Например, время выключения диода 2Д528А составляет 50 пс, импульсный прямой ток – 200 мА, а 2Д524Б – 200 пс и 1 А соответственно.

Одно из основных назначений ДНЗ – это формирование коротких импульсов большой амплитуды с длительностью в несколько десятков пико- или наносекунд. Они находят применение также в высокочастотных умножителях частоты. Короткие импульсы большой амплитуды имеют высокочастотные гармоники, которые могут быть выделены, что позволяет осуществить умножение частоты. ДНЗ используют в СВЧ-диапазоне для умножения частоты в 6 и более раз.

2.12. Лавинно-пролетные диоды

Лавинно-пролетный диод (ЛПД) – это диод с удлиненной базой, работающий в режиме лавинного размножения носителей заряда и предназначенный для генерирования СВЧ-колебаний.

Этот прибор относится к классу активных диодов с эквивалентным отрицательным сопротивлением. Одна из возможных структур диода приведена на рис. 2.56. Она состоит из p - n -перехода (заштрихованная область), а также удлиненной базы, состоящей из высоколегированной электронной области (n^+) и полупроводника с собственной проводимостью (i). Название диода включает в себя два основных явления, отражающих его принцип действия. Ударное лавинное образование

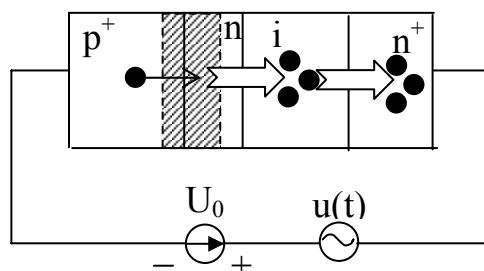


Рис. 2.56

электронно-дырочных пар внутри перехода и наличие конечного времени пролета носителей заряда через i -область приводит к фазовому сдвигу тока через диод относительно приложенного напряжения. Вследствие этого возникает эффект отрицательного динамического сопротивления. Величина этого сопротивления стабильна, если время пролета не меняется при изменении напряжения на ЛПД. Это возможно при больших значениях напряженности электрического поля, когда дрейфовая скорость становится постоянной и не зависит от величины поля.

При подключении обратного смещения U_0 , меньшего напряжения пробоя, через диод протекает небольшой ток неосновных носителей заряда. Так как толщина n -слоя мала, то весь он практически занят областью объемного заряда. Пролетный слой собственной проводимости также имеет высокое сопротивление. Поэтому напряженность поля вдоль данной структуры распределяется в соответствии с омическим сопротивлением областей (рис. 2.57). Лавинный процесс об-

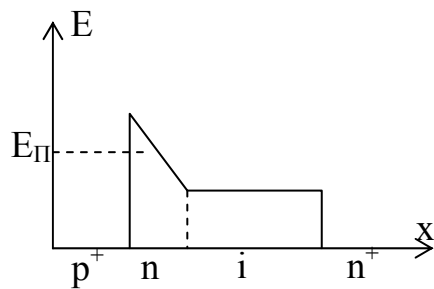


Рис. 2.57

разования электронно-дырочных пар внутри перехода происходит в том случае, если напряженность поля превысит некоторое пороговое значение $E_П$.

Подадим на диод постоянное обратное напряжение U_0 и переменный сигнал $u(t)$, причем $U_0 < U_{проб}$ (меньше напряжения лавинного пробоя), а амплитуда

переменного сигнала такова, что в момент максимума суммарное напряжение превосходит напряжение пробоя (рис. 2.58). В определен-

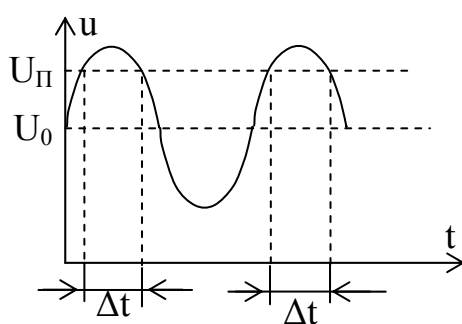


Рис. 2.58

ные интервалы времени Δt происходит периодически повторяющийся лавинный пробой перехода. Дырки, появившиеся вследствие лавинного размножения, дрейфуют в p^+ -область, а электроны – к n^+ -области. Таким образом, через слой i проходит модулированный по плотности поток электронов, затрачивая на это некоторое время пролета (дрейфа). В связи с конечным временем пролета ток во внешней цепи

отстает по фазе от изменения напряжения. В пролетном слое происходит энергетический взаимообмен между электронным потоком и изменяющимся электрическим полем. При правильном выборе ширины обедненного слоя, т.е. времени пролета, возникает режим автогенерации, приводящий к росту амплитуды переменного сигнала.

Процесс энергообмена поясняет пространственно-временная диаграмма, приведенная на рис. 2.59. Толщина пролетной области выбрана таким образом, что электроны достигают n^+ -области (координата x_0) в конце периода высокочастотного колебания T . Если пренебречь инерционностью процесса лавинообразования, то поток электронов движется по прямой 1, затрачивая время пролета $T = (3/4)T$. В течение положительной части периода высокочастотного поля $\Delta t = t_2 - t_1 = T/4$ электроны ускоряются, а во время последующего интервала времени $T/2$ замедляются и отдают энергию полю. Среднее значение эквивалентного дифференциального сопротивления за период получается отрицательным. Если процесс лавинообразования запаздывает на $T/4$ (прямая 2), то электроны сразу попадают в замедляющий интервал, что соответствует максимальной величине R_d .

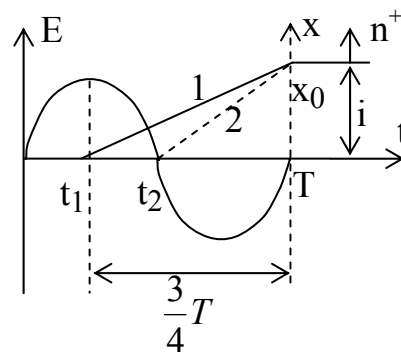


Рис. 2.59

Если поместить ЛПД в резонансную камеру, то процесс генерации происходит без дополнительного внешнего источника $u(t)$. Этот режим работы получил название IMPATT-режим (*Impact Avalanche Transit Time*), т.е. ударная лавинная ионизация плюс пролетное время. При больших токах ЛПД работают в ключевом режиме с «захваченной плазмой», который получил название TRAPATT-режим (*Traped Plasma Avalanche Triggered Transit*). В этом случае частота генерируемых колебаний становится меньше, а их амплитуда и КПД диода увеличиваются. Этот режим реализуется в p^+n-n^+ -структуре при подаче напряжения значительно превышающего пробивное. Развитие процесса лавинного пробоя последовательно перемещается из части n -слоя, прилегающего к p^+ -области, к противоположной границе n -области. Это связано с перераспределением падения напряжения (напряженности поля) между частями n -слоя с резко отличающимися концентрациями электронов и дырок. Поэтому слой умножения перемещается по области объемного заряда, а вся n -область быстро заполняется электронно-дырочной плазмой. Необходимая задержка здесь возникает из-за быстрого движения фронта волны, ионизации и медленного дрейфа носителей заряда при их экстракции из области объемного заряда.

Увеличение мощности можно получить также в p^+p-n-n^+ -структуре, т.е. в ЛПД с двойным пролетным пространством. Здесь кроме дрейфа электронов в n -области происходит синфазный дрейф дырок в p -области.

ЛПД находят широкое применение в качестве генераторов СВЧ-сигналов в диапазоне 0,5...50 ГГц и выше. Диапазон частот, в которых данная структура имеет отрицательное динамическое сопротивление, определяется малым временем пролета.

Этот твердотельный генераторный прибор см- и мм-диапазонов относится к числу наиболее мощных: от десятков милливатт до 1...5 Вт в непрерывном режиме. Большие значения мощности и КПД соответствуют более низким частотам.

В усилительном режиме диод не используется, что связано с высоким уровнем собственных шумов, так как лавинный пробой всегда сопровождается дробовым шумом.

Маркировка ЛПД соответствует системе маркировки СВЧ-диодов – 1A704A (1 – германиевый, A – СВЧ-диод, 7 – генераторный режим).

К основным параметрам ЛПД относятся:

- выходная непрерывная СВЧ-мощность при заданном рабочем токе P ;
- рабочий диапазон частот Δf ;
- постоянное рабочее обратное напряжение $U_{обр}$ (60...120 В);
- коэффициент полезного действия (3%...15%, максимум до 40%);
- максимальное значение постоянного рабочего тока $I_{макс}$ как предельный эксплуатационный параметр.

Конструктивно ЛПД оформляется в специальных корпусах патронного или коаксиального типа.

2.13. Диоды Ганна

Диод Ганна – это активный прибор, действие которого основано на появлении отрицательного дифференциального сопротивления под влиянием сильного электрического поля.

В зарубежной литературе для этого диода используют сокращенное наименование – TED (*Transferred – Electron Devices*). Особенность структуры прибора – отсутствие выпрямляющего электрического перехода. Он представляет собой кристалл арсенида галлия n -типа с омическими контактами с обеих сторон.

Эффект Ганна возникает в полупроводниках с особой многодолинной структурой зоны проводимости. При этом энергетическая диаграмма должна быть описана в пространстве квазиимпульсов $\varepsilon = f(\bar{k})$. Так, например, энергия свободного электрона массой m_0 , находящегося в вакууме, описывается следующим выражением:

$$\varepsilon = \frac{m_0 \bar{v}^2}{2} = \frac{m_0^2 \bar{v}^2}{2m_0} = \frac{\bar{P}^2}{2m_0},$$

где $\bar{P} = m_0^2 \bar{v}^2$ – векторный импульс электрона. Тогда энергетическая диаграмма свободного электрона в вакууме, в пространстве волнового вектора \bar{k} определяется зависимостью, представленной на рис. 2.60. Переход от векторного импульса к новой координате осуществляется по следующей формуле:

$$\bar{P} = \frac{h}{2\pi} \bar{k},$$

где h – постоянная Планка.

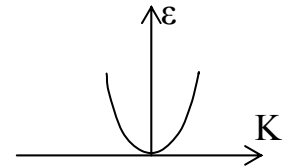


Рис. 2.60

Увеличение этого импульса приводит к квадратичному возрастанию энергии электрона. Энергетическая диаграмма арсенида галлия n -типа имеет в зоне проводимости два минимума (в кристаллографическом направлении $[100]$), которые показаны на рис. 2.61. Области энергий в зоне проводимости около каждого минимума называются долинами.

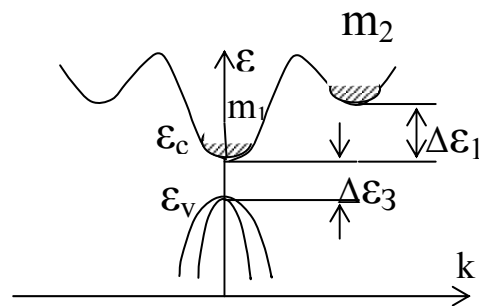


Рис. 2.61

В слабых электрических полях электроны проводимости имеют малые дрейфовые скорости, квазиимпульсы и поэтому находятся в центральной долине. В сильных электрических полях электроны получают дополнительную энергию $\Delta\varepsilon_1 = 0,36$ эВ и переходят в боковые долины. Причем электроны, занимающие уровни в центральной долине, имеют малую эффективную массу ($m_1 = 0,072m_0$ – «легкие» электроны) и высокую подвижность. В боковой долине находятся «тяжелые» электроны с $m_2 = 1,2m_0$, поэтому $\mu_2 < \mu_1$, т.е. подвижности отличаются в десятки раз. Таким образом, с увеличением напряженности электрического поля средняя подвижность свободных электронов уменьшается.

Плотность тока в общем виде можно определить так:

$$j = q(n_1\mu_1 + n_2\mu_2)E,$$

где n_1 и n_2 – концентрации, соответственно, «легких» и «тяжелых» электронов. Зависимость плотности тока от напряженности электрического поля показана на рис. 2.62. Участок 1 при малых полях $E < E_{\Pi}$ соответствует «легким» элек-

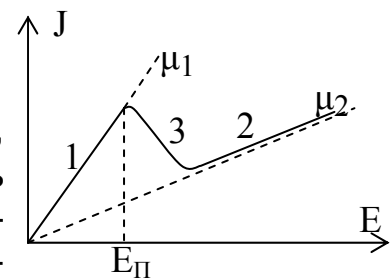


Рис. 2.62

тронам, а участок 2 – «тяжелым». Переходный участок 3 определяется соотношением концентраций «легких» и «тяжелых» электронов.

Условие получения отрицательной проводимости $dj/dE < 0$ находим при дифференцировании предыдущего выражения. Пренебрегая членами с параметром μ_2 ($\mu_1 \gg \mu_2$) и учитывая $dn_1/dE = -dn_2/dE$, получим

$$\frac{dn_2}{dE} > \frac{n_1}{E}.$$

Следовательно, на участке отрицательной проводимости небольшое изменение напряженности поля должно вызывать интенсивный переход электронов из одной долины в другую.

Таким образом, в исходном состоянии, при отсутствии внешнего напряжения, почти все электроны находятся в зоне с малым уровнем энергии (центральная долина), имеют высокую подвижность, т.е. являются «легкими» (малоинерционными).

Пусть к диоду приложено внешнее напряжение, которое создает внутри структуры электрическое поле, несколько меньшее порогового значения $E < E_{\text{п}}$ (рис. 2.63). Названия внешних выводов определяются

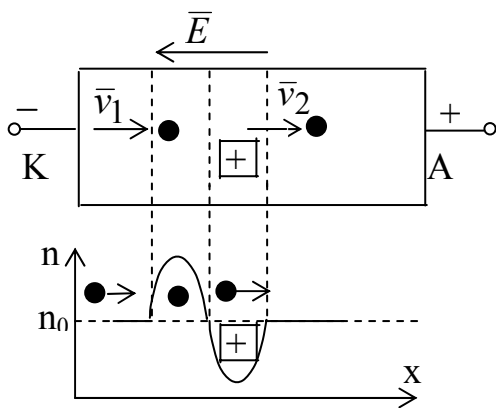


Рис. 2.63

знаком потенциала: положительный А – анод, отрицательный К – катод. Вначале в структуре происходит формирование особой области объемного заряда – домена (*domain* – область), что связано с перераспределением электронов в некоторой локальной области. Дрейфовая скорость электронов определяется напряженностью поля $v = \mu E$, поэтому зависимость $v = v(E)$ аналогична характеристике $j(E)$ (см. рис. 2.62).

Допустим, что вблизи катода в результате флуктуации локально уменьшилась концентрация электронов. Это приведет к уменьшению проводимости $\sigma = qn\mu$, увеличению падения напряжения на этом участке, а также к возрастанию напряженности поля. Если $E > E_{\text{п}}$, то в соответствии с зависимостью $j(E)$ уменьшится скорость электронов, так как появятся тяжелые носители заряда с малой подвижностью, что опять приведет к уменьшению проводимости. Развитие процесса соответствует следующей логической цепочке:

$$n \downarrow \rightarrow \sigma \downarrow \rightarrow U \uparrow \rightarrow E > E_{\text{п}} \rightarrow \mu \downarrow \rightarrow v \downarrow.$$

Следовательно, скорость движения электронов внутри слоя меньше, чем вне его. Происходит перераспределение зарядов. Около катода образуется область накопления «тяжелых» электронов, которые медленно движутся к аноду и создают избыточный отрицательный заряд. Перед фронтом скопления «тяжелых» электронов ближе к аноду появляется положительный объемный заряд нескомпенсированных ионов донорной примеси. Это связано с быстрым уходом «легких» электронов из этой области (см. рис. 2.63).

Область, содержащая двойной слой электрических зарядов из избыточных электронов и положительных ионов, называется доменом. Домен обладает собственным электрическим полем E_D , направление которого совпадает с внешним полем, причем по мере формирования домена поле в нем увеличивается, а за его пределами уменьшается. Это приводит к росту скорости «тяжелых» электронов, снижению скорости «легких». Когда скорости электронов в домене и вне его сравняются, дальнейшее накопление заряда прекращается. В сформированном домене устанавливается динамическое равновесие между электронами, притекающими в него со стороны катода и уходящими со стороны анода. При этом весь домен дрейфует от катода к аноду с постоянной скоростью.

Процесс периодического формирования и исчезновения доменов приводит к появлению импульсов тока во внешней цепи. Домен появляется у катода, движется с постоянной скоростью и, достигнув анода, рассасывается. После этого напряженность поля во всем кристалле повышается, что приводит к зарождению нового домена и повторению всего цикла.

Устойчивое состояние соответствует наличию только одного домена, так как после его появления напряженность поля в остальной части структуры падает и при $E < E_D$ второй домен образоваться не может.

Изменения тока, проходящего через диод, показаны на рис. 2.64. При образовании домена происходит быстрый спад тока ($0 < t < t_1$), так как этот процесс сопровождается уменьшением напряженности поля вне домена ($j = qnv = qn\mu E$). В течение дрейфа ($t_1 < t < t_2$) ток не изменяется и имеет минимальное значение, так как вне домена в пролетной области напряженность поля также минимальна. Рассасывание домена ($t_2 < t < t_3$) сопровождается увеличением напряженности поля, что приводит к увеличению тока до максимального значения j_1 .

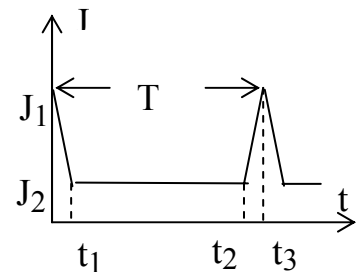


Рис. 2.64

Таким образом, эффект генерации в однородной полупроводниковой структуре вызван тем, что в ней периодически появляется, перемещается и исчезает область сильного поля – домен, причем он выполняет роль своеобразного клапана, управляющего величиной тока в диоде.

Колебания тока в генераторе возникают в том случае, если время пролета домена от катода к аноду $T = l/v$ больше времени его формирования. Это время определяется временем релаксации $\tau = \epsilon\epsilon_0/\sigma$. Отсюда условие возникновения колебаний $n_0 l > \epsilon\epsilon_0 v / q\mu_2$.

Максимальная частота генерации определяется минимальным значением длины кристалла. При обратном неравенстве колебания самопроизвольно не возникают, но диод можно использовать для усиления СВЧ-сигналов.

Кроме рассмотренного выше пролетного режима диод Ганна может работать в режиме ограниченного накопления объемного заряда (ОНОЗ) – в английской транскрипции LSA – *Limited Spacecharge Accumulation*. В этом режиме на диод подается постоянное напряжение смещения $U_0 = E_0 l = (2...3)E_{\text{п}} l$ и синусоидальное напряжение большей амплитуды (рис. 2.65). Тогда в течение части периода общая

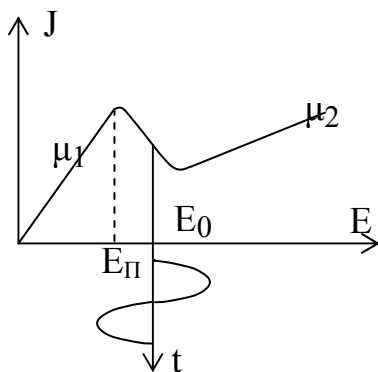


Рис. 2.65

напряженность поля меньше порогового значения. Пока $E > E_{\text{п}}$, у катода формируется домен, а в той части периода, когда $E < E_{\text{п}}$, он рассасывается. При малом периоде колебаний домен рассасывается, не успев сформироваться. При формировании домена ток через прибор уменьшается, а во время рассасывания увеличивается. Таким образом, возникают колебания тока, период которых определяется не временем пролета домена, а внешним резонатором.

Условия реализации режима следующие:

1. Домен не должен успеть сформироваться, пока $E > E_{\text{п}}$ ($\mu = \mu_2$), т.е. период колебаний $T < (2...3)\tau$ меньше времени релаксации $\tau = \epsilon\epsilon_0/\sigma$.

2. Домен должен успеть исчезнуть за время, пока $E < E_{\text{п}}$ ($\mu = \mu_1$), т.е. $T \gg \epsilon\epsilon_0 / q\mu_1$, откуда рабочий интервал частот

$$\epsilon\epsilon_0 / (q\mu_1) \ll \frac{n}{f} < (2...3)\epsilon\epsilon_0 / (q\mu_2).$$

Первое преимущество режима ОНОЗ заключается в независимости рабочей частоты от длины кристалла. Это позволяет работать на более высоких частотах, а также изменять частоту генерации путем перестройки внешнего резонатора. Второе преимущество – в возможности получения больших мощностей, так как к кристаллам большей длины можно подавать большие напряжения.

Диоды, реализуемые на основе арсенида галлия, работают в частотном диапазоне 1...50 ГГц. Верхняя граница диапазона ограничивается минимальной длиной кристалла в пролетном режиме. Уменьшение размера образца ограничено толщиной домена. Кроме того, для увеличения энергии электрона на 0,3 эВ (боковая долина) при пороговой величине поля 3 кВ/см необходимо пролетное расстояние не менее 1 мкм. Нижняя граница частотного диапазона (1 ГГц) соответствует длине образца $l < 100$ мкм. При увеличении его размеров колебания носят неопределенный шумовой характер, что связано с внутренними неоднородностями кристалла, на которых зарождаются домены.

На диодах из фосфида индия могут быть получены большие частоты со значительно большими мощностями. Преимущество этого материала связано с возможностью получения больших скоростей дрейфа электронов, а также большим значением пороговой напряженности поля. В настоящее время качество приборов из этого материала ограничено технологическими возможностями.

Генераторы с ОНОЗ увеличивают верхний частотный диапазон до 100 ГГц при КПД около 20% и мощности порядка десятых долей ватта.

Основные параметры диодов Ганна:

1. Выходная непрерывная мощность P . В пролетном режиме составляет 10...500 мВт.
2. Рабочая частота f . В пролетном режиме типовой частотный диапазон 4...26 ГГц.
3. Постоянный рабочий ток I (от сотен миллиампер до единиц ампер).
4. Постоянное напряжение U (3...12) В.
5. Максимальная рассеиваемая мощность $P_{\text{макс}}$.
6. Коэффициент полезного действия η (1...30%).

Срок службы диодов сравнительно невелик, что связано с перегревом кристалла из-за выделяющейся мощности и воздействием сильных электрических полей. Так, например, величина рассеиваемой мощности для диода АА726А составляет 17 Вт. Шумы диодов Ганна значительно меньше, чем у ЛПД. Причиной амплитудных шумов являются флуктуации скорости движения электронов. Однако эти

шумы малы, так как приборы работают в режиме насыщения дрейфовой скорости, которая почти не меняется при изменении электрического поля.

У диодов в пролетном режиме практически отсутствует возможность электронной перестройки частоты. С изменением внешнего напряжения немного изменяется толщина домена при почти постоянной скорости дрейфа. Изменение частоты составляет десятые доли процента. Частотно-перестраиваемые генераторы реализуются в режиме ОНОЗ. В генераторах Ганна используются резонаторы волноводного, коаксиального или полоскового типа.

Библиографический список

Викулин И.М., Стафеев В.И. Физика полупроводниковых приборов. – М.: Радио и связь, 1990.

Крутякова М.Г., Чарыков Н.А., Юдин В.В. Полупроводниковые приборы и основы их проектирования. – М.: Радио и связь, 1983.

Пасынков В.В., Чиркин Л.К. Полупроводниковые приборы. – М.: Высш.шк., 1987.

Полупроводниковые приборы. Диоды выпрямительные, стабилитроны, тиристоры: Справочник/А.Б. Гитцевич, А.А. Зайцев, В.В. Мокряков и др./Под ред. А.В. Голомедова. – М.: Радио и связь, 1988.

Полупроводниковые приборы. Диоды высокочастотные, импульсные, оптоэлектронные приборы: Справочник/А.Б. Гитцевич, А.А. Зайцев, В.В. Мокряков и др./Под ред. А.В. Голомедова. - М.: КУБК-а, 1994.

Тугов Н.М., Глебов Б.А., Чарыков Н.А. Полупроводниковые приборы. – М.: Энергоатомиздат, 1990.

Физический энциклопедический словарь. – М.: Сов. энциклопедия, 1983.

Электронные приборы/В.Н. Дулин, Н.А. Аваев, В.П. Демин и др./Под ред. Г.Г. Шишкина. – М.: Энергоатомиздат, 1989.

Оглавление

| | |
|---|----|
| 1. Физические основы полупроводниковых приборов..... | 3 |
| 1.1. Электропроводность полупроводников..... | 3 |
| 1.2. Электронно-дырочный переход..... | 14 |
| 1.2.1. Электронно-дырочный переход в равновесном состоя- нии..... | 14 |
| 1.2.2. Основное свойство р-п-перехода – свойство односто- ронней проводимости..... | 17 |
| 1.2.3. Вольт-амперная характеристика идеализированного р-п-перехода..... | 20 |
| 1.2.4. ВАХ реальных р-п-переходов..... | 24 |
| 1.2.5. Свойства р-п-перехода..... | 25 |
| 2. Полупроводниковые диоды..... | 34 |
| 2.1. Выпрямительные диоды..... | 34 |
| 2.1.1. Применение выпрямительных диодов..... | 38 |
| 2.1.2. Особенности расчета простейших схем с диодами..... | 41 |
| 2.2. Опорные диоды..... | 45 |
| 2.2.1. Параметрический стабилизатор напряжения..... | 49 |
| 2.3. Варикапы..... | 52 |
| 2.4. Параметрические диоды..... | 57 |
| 2.5. Туннельные диоды (Езаки)..... | 58 |
| 2.6. Обращенные диоды..... | 62 |
| 2.7. Диоды Шотки..... | 64 |
| 2.8. Р-і-п-диоды..... | 67 |
| 2.9. Диоды ВЧ- и СВЧ-диапазона..... | 70 |
| 2.10. Импульсные диоды..... | 72 |
| 2.11. Диоды с накоплением заряда..... | 75 |
| 2.12. Лавинно-пролетные диоды..... | 77 |
| 2.13. Диоды Ганна..... | 80 |
| Библиографический список..... | 86 |

Барсуков Сергей Николаевич

ЭЛЕМЕНТНАЯ БАЗА РАДИОЭЛЕКТРОНИКИ
Часть 1
ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

Редактор А.Н. Емленинова

Св. план, 2002

Подписано в печать 13.09.2002

Формат 60x84 1/16. Бум. офс. №2. Офс. печ.

Усл.-печ. л. 4,9. Уч.-изд. л. 5,5. Т. 100 экз. Заказ 440. Цена свободная

Национальный аэрокосмический университет им. Н.Е. Жуковского

«Харьковский авиационный институт»

61070, Харьков-70, ул. Чкалова, 17

<http://www.khai.edu>

Издательский центр «ХАИ»

61070, Харьков-70, ул. Чкалова, 17

izdat@khai.edu