

Министерство образования и науки Российской Федерации
Федеральное агентство по образованию

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
им А.Н. ТУПОЛЕВА

Л.М. УРМАНЧЕЕВ

ФИЗИЧЕСКИЕ ОСНОВЫ ЭЛЕКТРОНИКИ

Учебное пособие

Казань 2009

УДК 621.38

Урманчеев Л.М. Физические основы электроники: Учебное пособие.
Казань: Изд-во Казан. гос. техн. ун-та, 2009. 72 с.
ISBN

Рассмотрены виды и структура полупроводников, процессы генерации и рекомбинации носителей заряда, законы движения носителей в полупроводниках, вакууме и плазме, физические явления в электрических переходах, термоэлектрические и фотоэлектрические явления в полупроводниках и переходах, фотоэлектрические преобразователи, светодиоды, электровакуумные лампы и электронно-лучевые приборы. Предназначено для студентов направления 654400 – телекоммуникация.

Ил.60. Табл. 2. Библиогр.: 10 назв.

Рецензенты: доктор физико-математических наук, профессор А.В. Голенищев-Кутузов (Казанский государственный энергетический университет);
кафедра Информатики и информационно-управляющих систем (Казанский государственный энергетический университет)

ISBN

©: Изд-во Казан. гос. техн. ун-та, 2009
© Л.М. Урманчеев, 2008

ПРЕДИСЛОВИЕ

Данное учебное пособие составлено в соответствии со стандартом ЕН.Ф.9 на программу дисциплины “Физические основы электроники” для студентов направления 654400 – телекоммуникация.

Программа по дисциплине “Физические основы электроники” базируется на знаниях следующих дисциплин:

“Математика” (дифференцирование, интегрирование, решение дифференциальных уравнений);

“Физика” (электростатика, электродинамика, статистическая физика, квантовая механика, физика твердого тела);

Цель преподавания дисциплины “Физические основы электроники” – обеспечить базовую подготовку студентов, необходимую для успешного изучения электроники и специальных дисциплин.

В пособии рассмотрены виды и структура полупроводников, элементы зонной теории твердого тела, зонные диаграммы, равновесная и неравновесная концентрация носителей заряда, процессы генерации и рекомбинации носителей, законы движения носителей, дрейфовые и диффузионные токи, гальваномагнитный эффект Холла, физические явления в электрических переходах (электронно-дырочный переход, контакты металл-полупроводник, гетеропереходы), математическая модель идеализированного электронно-дырочного перехода, физические явления, вызывающие отклонения от идеализированной модели, виды пробоя *p-n*-перехода, инерционные свойства *p-n*-перехода, барьерная и диффузионная емкости, термоэлектрические явления (эффект Зеебека и Пельтье), фотоэлектрические явления в полупроводниках и переходах, виды электронной эмиссии, особенности движения носителей в вакууме и плазме.

Также рассмотрены устройство, принцип действия и характеристики фотоэлектрических преобразователей, светодиодов, электровакуумных ламп и электронно-лучевых приборов. Эти вопросы изучаются в данном курсе, поскольку они важны, но, согласно стандарту ОПД Ф.2, не включены в программу последующей дисциплины «Электроника» для данного направления.

ВВЕДЕНИЕ

Электроника – область науки и техники, охватывающая исследования, разработку и принципы применения электронных приборов.

Микроэлектроника – область электроники, предметом которой является принципиально новый вид электронных приборов – интегральные микросхемы.

В электронных приборах используются различные электромагнитные, оптические, акустические, тепловые, электрохимические явления в твердом теле, вакууме, газе, жидкости для преобразования информационных сигналов или энергии.

Электронные приборы классифицируются:

- по виду рабочей среды – полупроводниковые, электронные электровакуумные, ионные электровакуумные (газоразрядные), хемотронные;
- по виду преобразуемых сигналов – электропреобразовательные, электро-световые, фотоэлектрические, акустоэлектрические, термоэлектрические, механоэлектрические, электрохимические.

Они различаются также по другим признакам – назначению и выполняемым функциям, диапазону рабочих частот, мощности.

Дисциплина «Физические основы электроники» посвящена изучению физических явлений в полупроводниках и вакууме, обусловленных взаимодействием электронов с электромагнитным полем и веществом, принципам использования этих явлений для создания электронных приборов.

ЭЛЕКТРОФИЗИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

1. ПОЛУПРОВОДНИКИ

К полупроводникам относятся вещества с удельной электропроводностью σ в интервале $10^{-10} < \sigma < 10^4$ См/см. Вещества с меньшей электропроводностью относят к диэлектрикам, с большей – к металлам. Эти границы условны, особенно между диэлектриками и полупроводниками, между которыми по существу нет принципиальных различий. Главное различие между полупроводниками и металлами заключается в том, что у металлов удельное сопротивление возрастает с ростом температуры, а у полупроводников – падает.

К полупроводникам относится большая группа веществ, но в производстве приборов используются лишь немногие, главным образом, германий Ge, кремний Si, арсенид галлия GaAs и некоторые другие.

Структура полупроводников

Кристаллическая решетка. Применяемые в электронике полупроводники имеют, как правило, монокристаллическую структуру, т.е. во всем объеме вещества атомы размещены в узлах строго определенной единой кристаллической решетки. Кристаллическая решетка состоит из множества повторяющихся и примыкающих друг к другу *элементарных ячеек* той или иной формы и размера. В случае простейшей кубической решетки ребро элементарной ячейки – куба – является *постоянной решетки*.

Регулярность (периодичность) структуры кристалла приводит к *анизотропии* – зависимости его свойств от направления в кристаллической решетке. Различные плоскости сечения кристалла (*кристаллографические плоскости*) принято обозначать трехзначными *индексами Мюллера*, заключенными в круглые скобки: (111), (100) и т.п., а направления, перпендикулярные к ним – теми же индексами, заключенными в квадратные скобки: [100], [110] и т.п. На рис.1 показаны плоскости (100), (110) и (111) для простейшей кубической решетки. Кристаллографические плоскости имеют разную плотность атомов на единицу площади: (111) – наибольшую, (100) – наименьшую – и отличаются физическими и химическими свойствами.

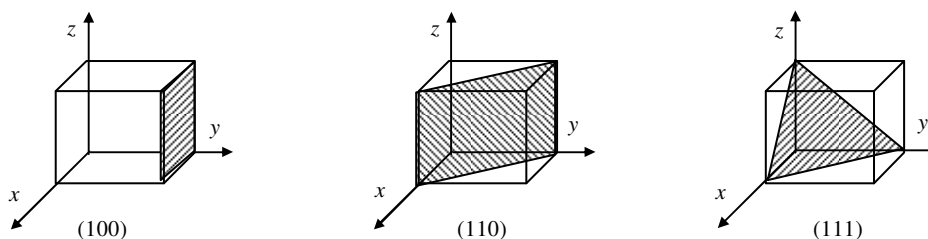


Рис.1

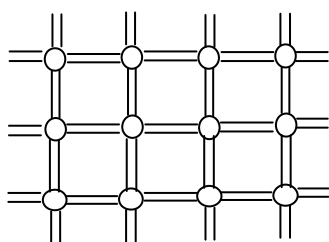


Рис.2

В германии и кремнии атомы образуют кубическую решетку типа алмаза, где каждый атом связан ковалентной связью с 4 ближайшими соседями. Для наглядности эту решетку можно представить в виде плоской модели (рис.2), где двойные черточки, связывающие атомы, отображают ковалентную химическую связь, образованную двумя электронами.

Дефекты решетки. Структура кристалла никогда не бывает идеальной, всегда имеются дефекты решетки и дислокации (смещения плоскостей решетки).

Дефекты точечного типа могут иметь вид пустого узла (*дефект по Шоттки*), совокупности пустого узла и междуузельного атома (*дефект по Френкелю*) или чужеродного атома – примеси. Любой реальный кристалл содержит примеси – либо паразитные, от которых не удастся избавиться при очистке, либо полезные, которые специально вводятся для получения нужных свойств кристалла. Примеси могут располагаться либо между узлами решетки (*примесь внедрения*), либо в самих узлах вместо основных атомов (*примесь замещения*).

Поверхность кристалла. На поверхности кристалла часть ковалентных связей оказывается разорванной, что приводит к нарушению энергетического равновесия. Равновесие восстанавливается различными путями: может измениться расстояние между атомами в приповерхностном слое; может произойти захват – *адсорбция* – чужеродных атомов или молекул; может образоваться химическое соединение (например, окисел), не имеющее незаполненных связей на поверхности, и т. п. В любом случае структура тонкого приповерхностного слоя (толщиной несколько нанометров) отличается от структуры основного объема кристалла и этот слой следует рассматривать как *особую область кристалла*.

Тепловые колебания решетки. Фононы

Атомы, находящиеся в узлах кристаллической решетки, могут совершать колебательные движения около положения равновесия. Тепловые колебания атомов кристалла не являются независимыми из-за сильного взаимодействия между ними. Смещение одного атома вызывает смещение соседних, и возникает коллективное смещение атомов, распространяющееся в кристалле в виде упругих волн. В области длинных волн при $\lambda \gg d$ (d – постоянная решетки) – это обычная звуковая волна, распространяющаяся со скоростью звука v_0 (порядка 10^5 м/с). Движение каждого атома описывается суперпозицией всех возбужденных волн.

В квантовой механике каждой волне смещения с определенной частотой ω и волновым вектором \vec{k} ставится в соответствие некая квазичастица, которая названа *фононом*. Распространение в твердом теле волны смещения можно рассматривать как движение одного или многих фононов. Каждый фонон переносит энергию $\epsilon_{\text{фон}} = \hbar\omega$ и *квазиимпульс* $\vec{p}_{\text{фон}} = \hbar\vec{k}$ (эти выражения идентичны выражениям для энергии и импульса *фотона* – кванта электромагнитного излучения). Использование понятия «квазиимпульс» не означает, что через кристалл передается «обычный» импульс с соответствующей скоростью; однако квазиимпульс – это величина, которая сохраняется в различных процессах взаимодействия, включающих в себя рождение и уничтожение фононов.

Фонон – квант тепловой энергии кристаллической решетки. Энергия тепловых движений кристалла равна сумме энергий всех возбужденных фононов. Можно сказать, что вся тепловая энергия кристалла содержится в «газе» фононов,

заполняющем объем кристалла. Плотность этого «газа» и средняя энергия фононов возрастают с ростом температуры.

Фононы делятся на акустические и оптические. Первые соответствуют синфазным колебаниям в элементарной ячейке кристаллической решетки и распространению в кристалле волн, подобных акустическим. Их спектр простирается от нуля до некоторой максимальной частоты (порядка 10^{13} Гц для кремния).

Оптические фононы соответствуют противофазным колебаниям в элементарной ячейке кристалла и существуют на частотах, больших некоторой граничной частоты ($1,2 \cdot 10^{13}$ Гц для кремния), поэтому для их возбуждения требуется значительная энергия.

Средняя энергия фононов $\approx kT$, где k – постоянная Больцмана, T – абсолютная температура, а при низких температурах и того меньше.

Носители заряда в полупроводниках

В полупроводниках существуют два вида носителей заряда – *электроны* и *дырки*.

Беспримесный и бездефектный полупроводник с идеальной кристаллической решеткой называется *собственным полупроводником*. При абсолютном нуле температуры в нем нет свободных носителей заряда – все электроны связаны с атомами и не могут перемещаться по кристаллу, вещество не проводит электрический ток, т.е. является диэлектриком.

С ростом температуры возрастает количество и энергия фононов и столкновение электрона с фононом, обладающим достаточной энергией, приводит к разрыву валентной связи. При столкновении фонон исчезает, его энергия передается электрону. Оторвавшийся электрон может свободно перемещаться по кристаллической решетке. В том месте, откуда выбит электрон, образуется некомпенсированный положительный заряд, равный заряду электрона. Это и есть дырка. Она также может свободно перемещаться по кристаллу за счет перескока

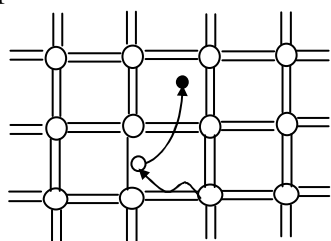


Рис.3

валентного электрона с соседнего атома, в результате дырка оказывается у соседнего атома и далее этот процесс повторяется. Дырка ведет себя подобно частице с положительным элементарным зарядом. Образование электронно-дырочной пары показано на рис.3, где дырка обозначена мелким светлым кружком, электрон – темным.

Процесс образования электронно-дырочных пар под действием теплового движения называется *термогенерацией*. Наряду с термогенерацией идет и обратный процесс – *рекомбинация*, когда свободный электрон соединяется с дыркой и восстанавливается валентная связь, пара носителей исчезает.

В собственном полупроводнике электроны и дырки всегда образуются парами и их концентрации равны:

$$n_i = p_i$$

Здесь n и p — концентрация электронов и дырок соответственно, индексом “ i ” обозначают свойство, относящееся к собственному полупроводнику.

Примесные полупроводники содержат атомы посторонних элементов, встроенные в кристаллическую решетку. Примеси специально вводят в полупроводник для изменения его электрофизических свойств (этот процесс называется *легированием*). В примесных полупроводниках концентрации электронов и дырок могут отличаться на много порядков.

Примеси бывают донорные, акцепторные и нейтральные.

Для четырех валентных элементарных полупроводников — германия и кремния — донорными примесями являются атомы пятивалентных элементов, таких как фосфор **P**, мышьяк **As**, сурьма **Sb**, акцепторными — атомы трех валентных элементов: бор **B**, индий **In**, галлий **Ga**, алюминий **Al**.

Встраиваясь в решетку, атомы пятивалентных элементов образуют четыре связи с ближайшими соседями, пятый электрон оказывается лишним. Он не участвует в образовании химической связи и слабо связан с атомом примеси, легко отрывается от него и становится свободным. В узле решетки остается положительно заряженный ион примеси. Он жестко закреплен в решетке и не может перемещаться по кристаллу. В полупроводнике с донорной примесью основными носителями заряда являются электроны. Его называют полупроводником с электронной проводимостью или полупроводником n -типа.

При введении трехвалентного атома у него не хватает одного электрона для образования четырех связей. Недостающий электрон может быть захвачен у соседнего атома, у которого образуется дырка. Атом примеси превращается в отрицательный ион. В полупроводнике с акцепторной примесью основными носителями заряда являются дырки и его называют полупроводником с дырочной проводимостью или полупроводником p -типа.

Нейтральные примеси не изменяют концентрацию носителей заряда.

Энергетические уровни и зоны

Основой теории полупроводников и полупроводниковых приборов является зонная теория твердого тела.

Согласно квантовой теории электроны в атоме имеют дискретный энергетический спектр, т.е. энергия электрона может принимать только вполне определенный ряд дискретных значений. На каждом энергетическом уровне может находиться только два электрона с противоположными спинами (принцип Паули).

Твердое тело представляет собой множество атомов, сильно взаимодействующих друг с другом. Это взаимодействие приводит к тому, что энергетический уровень отдельного атома расщепляется на множество уровней, т.е. вместо одного разрешенного уровня получается целая зона близко расположенных уровней. Из одного уровня образуется столько уровней, сколько атомов в решетке кристалла.

Во взаимодействии участвуют, главным образом, электроны внешней оболочки атома, т.е. валентные электроны, расщепление нижнего энергетического уровня этих электронов образует *валентную зону*. Уровни энергии в зонах

дискретны, но расположены настолько близко друг к другу, что спектр в зоне можно считать практически непрерывным.

В металлах валентная зона заполнена электронами не полностью в отличие от полупроводников и диэлектриков, в которых при абсолютном нуле температуры все уровни валентных электронов заполнены.

В полупроводниках валентная зона (ВЗ) отделена от следующей разрешенной зоны, называемой *зоной проводимости* (ЗП), *запрещенной зоной* (ЗЗ), в которой нет разрешенных уровней энергии. На рис.4 показана структура энергетических зон собственного (*i*-типа) и примесных (*n*- и *p*-типа) полупроводников в рабочей области температур, когда атомы примеси практически полностью ионизированы (\oplus \ominus – ионы примеси, \circ – дырки, \bullet – электроны, W – энергия электрона, W_c – уровень дна зоны проводимости, W_v – уровень потолка валентной зоны, ΔW – ширина запрещенной зоны, ΔW_d , ΔW_a – энергия ионизации донорных и акцепторных примесей, W_F – уровень Ферми). Уровни атомов примеси расположены в запрещенной зоне: донорных – близко к

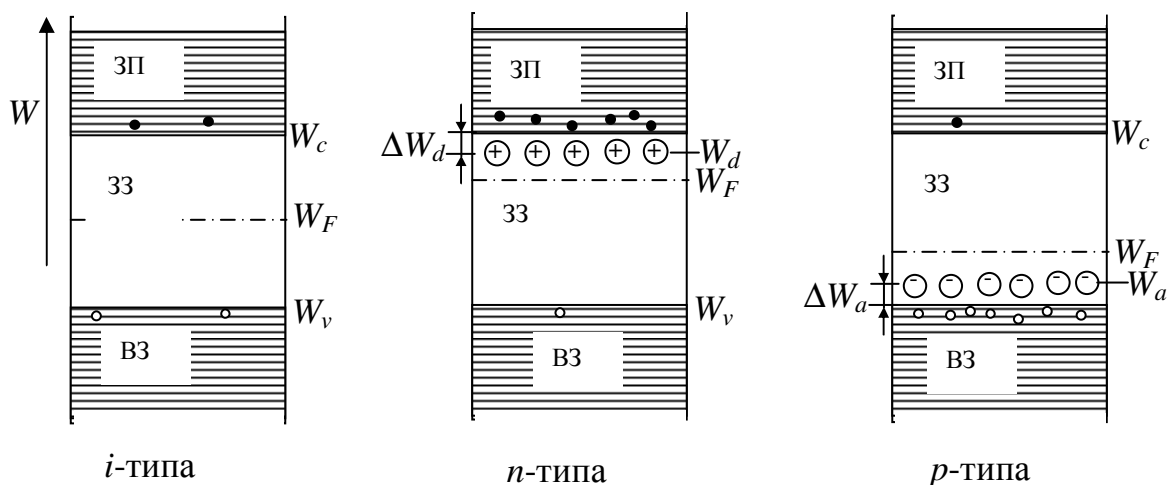


Рис.4

дну зоны проводимости, акцепторных – близко к потолку валентной зоны. Приведенные структуры относятся к *невырожденным* полупроводникам.

Невырожденным называется полупроводник с невысокой концентрацией примеси, в котором отсутствует взаимодействие между атомами примеси, поэтому нет расщепления их уровней.

Вырожденным называется полупроводник с высокой концентрацией примеси, в котором в результате взаимодействия между атомами примеси, их уровни расщепляются и образуют зону. При сильном вырождении эта зона сливается с ближайшей разрешенной зоной. Такие полупроводники называют также *полуметаллами*.

Уровень Ферми характеризует среднюю энергию электронов, это такой уровень энергии, вероятность заполнения которого равна одной второй. Это не означает, что в веществе реально существует такой разрешенный уровень. В полупроводнике уровень Ферми располагается, как правило, в запрещенной зоне, где нет разрешенных уровней. В собственном полупроводнике уровень Ферми находится примерно в середине запрещенной зоны, в полупроводнике

n -типа – ближе к дну зоны проводимости, в полупроводнике p -типа – ближе к потолку валентной зоны, в вырожденном полупроводнике – в зоне проводимости (n -тип) или в валентной зоне (p -тип).

Уровень Ферми – важнейшая характеристика системы. Его значение определяется тем, что во всех частях системы, находящейся в равновесии, уровень Ферми одинаков, какой бы разнородной система ни была, т.е. $W_F = \text{const}$. Это – одно из фундаментальных положений физики твердого тела.

Равновесная концентрация носителей в полупроводнике

Вероятность заполнения энергетического уровня определяется статистической функцией Ферми-Дирака:

$$s(W) = \frac{1}{\exp\left(\frac{W - W_F}{kT}\right) + 1}, \quad (1)$$

где k – постоянная Больцмана, T – абсолютная температура.

Концентрацию электронов в зоне проводимости можно найти, проинтегрировав удвоенное произведение вероятности заполнения на плотность энергетических уровней (каждый уровень может содержать два электрона):

$$n = \int_{W_c}^{\infty} 2s(W) \cdot dN(W) \quad (2)$$

Для невырожденных полупроводников экспонента в формуле (1) значительно больше единицы и распределение Ферми-Дирака можно заменить распределением Больцмана:

$$s(W) = \exp\left(-\frac{W - W_F}{kT}\right) \quad (3)$$

Вследствие быстрого спада экспоненты с увеличением W основной вклад в интеграл (2) вносят уровни находящиеся вблизи дна зоны проводимости, где $N(W) \propto \sqrt{W - W_c}$. Интегрирование с учетом распределения (3) приводит

к выражению:

$$n = N_c \exp\left(-\frac{W_c - W_F}{kT}\right) \quad (4)$$

где N_c – эффективная плотность состояний (энергетических уровней) в зоне проводимости (плотность разрешенных квантовых состояний в энергетической полосе шириной kT у границы зоны).

Аналогично концентрация дырок определяется выражением

$$p = N_v \exp\left(-\frac{W_F - W_v}{kT}\right), \quad (5)$$

где N_v – эффективная плотность состояний в валентной зоне.

Отсюда следует, что произведение концентраций электронов и дырок для данного невырожденного полупроводника и данной температуры не зависит от вида и степени легирования:

$$np = N_c N_v \exp\left(-\frac{\Delta W}{kT}\right), \quad (6)$$

где $\Delta W = W_c - W_v$ – ширина запрещенной зоны. Для собственной концентрации получаем

$$n_i = p_i = \sqrt{N_c N_v} \exp\left(-\frac{\Delta W}{2kT}\right). \quad (7)$$

Концентрация носителей в собственном полупроводнике сильно зависит от ширины запрещенной зоны и температуры. В табл. 1 приведены данные для наиболее распространенных полупроводников при комнатной температуре ($T=300$ К).

Таблица 1

| Полупроводник | ΔW , эВ | n_i , см ⁻³ |
|---------------|-----------------|--------------------------|
| Ge | 0,67 | $2,5 \cdot 10^{13}$ |
| Si | 1,11 | $2 \cdot 10^{10}$ |
| GaAs | 1,42 | $1,5 \cdot 10^6$ |

Формулу (6) можно записать в виде

$$np = n_i^2 \quad (8)$$

В объеме полупроводника выполняется условие электронейтральности, которое для примесных полупроводников можно записать в виде:

$$n_n = N_d^+ + p_n \quad \text{и} \quad p_p = N_a^- + n_p$$

для полупроводников n - и p -типа, соответственно (здесь N_d^+ и N_a^- – концентрации ионизированных донорных и акцепторных примесей, n_n и p_n – концентрация электронов и дырок в n -области, p_p и n_p – концентрация дырок и электронов в p -области).

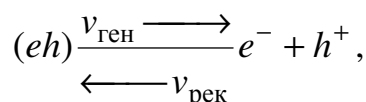
Донорные и акцепторные примеси образуют мелкие уровни и отстоят от соответствующих разрешенных зон на величину ΔW_d , $\Delta W_a \approx 0,01$ эВ, поэтому при комнатной температуре практически все примесные атомы ионизированы. Так как, как правило, $N_d \gg n_i$ и $N_a \gg n_i$, концентрация основных носителей практически равна концентрации легирующей примеси:

$$n_n \approx N_d \quad \text{и} \quad p_p \approx N_a \quad (9)$$

Концентрация неосновных носителей согласно формуле (8) на много порядков меньше:

$$p_n = n_i^2 / N_d \ll n_i \quad \text{и} \quad n_p = n_i^2 / N_a \ll n_i \quad (10)$$

Время жизни равновесных носителей. Установившаяся концентрация носителей заряда является результатом динамического равновесия между процессами генерации и рекомбинации. Эти процессы можно схематически представить в виде:



где (eh) отображает связанное состояние электрона, e^- и h^+ – свободный электрон и дырка.

Скорость термогенерации зависит от типа полупроводника и температуры:

$$v_{\text{ген}} = g_0(T)$$

Скорость рекомбинации пропорциональна произведению концентраций носителей:

$$v_{\text{рек}} = r_0 np$$

В равновесии $v_{\text{ген}} = v_{\text{рек}}$, отсюда

$$np = n_i^2 = g_0/r_0$$

Среднее время жизни равновесных носителей равно отношению концентрации этих носителей к скорости рекомбинации. Таким образом,

$$\tau_n = n/v_{\text{рек}} = 1/(r_0 p), \quad \tau_p = p/v_{\text{рек}} = 1/(r_0 n)$$

Отсюда следует, что время жизни неосновных носителей на несколько порядков меньше времени жизни основных, например,

$$(\tau_p)_n = 1/(r_0 N_d) \ll (\tau_n)_n = 1/(r_0 p_n) = N_d/(r_0 n_i^2)$$

Неравновесная концентрация носителей

Неравновесная концентрация носителей в полупроводнике может быть создана при различных энергетических воздействиях (облучение светом, импульсный нагрев, инжекция из соседних областей и др.). Неравновесные концентрации можно записать в виде

$$n = n_0 + \Delta n,$$

$$p = p_0 + \Delta p,$$

где Δn и Δp – отклонение концентраций от равновесных значений n_0 и p_0 .

Объем полупроводника сохраняет электронейтральность, поэтому всегда

$$\Delta n = \Delta p$$

После прекращения воздействия избыточные носители рекомбинируют, и концентрации возвращаются к исходным значениям. Скорость рекомбинации пропорциональна отклонению:

$$\frac{d\Delta p}{dt} = -\frac{\Delta p}{\tau} \quad \text{и} \quad \Delta p = \Delta p(0) \cdot e^{-t/\tau}, \quad (11)$$

где τ – время жизни неравновесных носителей, $\Delta p(0)$ – начальное отклонение.

Когда приращение концентрации значительно меньше концентрации основных носителей, значение τ близко к времени жизни неосновных носителей.

Механизмы рекомбинации

Различают *непосредственную рекомбинацию* и *рекомбинацию на примесных центрах (ловушках)*.

При непосредственной рекомбинации электрон из зоны проводимости переходит в валентную зону, заполняя вакантный уровень, т.е. уничтожая дырку. При этом должна выделиться энергия, равная ширине запрещенной зоны ΔW . Энергия может выделяться либо в виде фотона (*излучательная рекомбинация*), либо в виде фонона (*безызлучательная рекомбинация*).

В большинстве полупроводников, в частности, в кремнии и германии, вероятность излучательной рекомбинации на несколько порядков меньше, чем безызлучательной. Однако вероятность безызлучательной рекомбинации также мала, поскольку сравнительно большая энергия редко может быть передана одному фонону, а ее распределение между двумя фононами маловероятно. Поэтому *непосредственная рекомбинация не является главным механизмом рекомбинации в полупроводниках.*

Главную роль играет рекомбинация на примесных центрах, когда атомы примеси образуют глубокие уровни вблизи середины запрещенной зоны, которые называются *ловушками*. В этом случае рекомбинация протекает в два этапа – сначала электрон переходит из зоны проводимости на уровень ловушки и находится там некоторое время, а затем переходит в валентную зону. На каждом этапе выделяется энергия, примерно вдвое меньшая, чем при непосредственной рекомбинации, что резко повышает вероятность передачи энергии фонону.

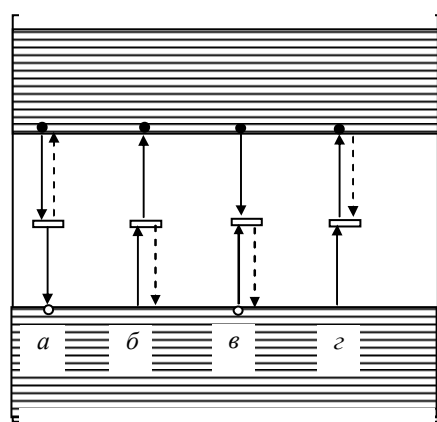


Рис.5

Столкновение с фононом может также перевести электрон из валентной зоны на уровень ловушки и на следующем этапе с этого уровня – в зону проводимости, что соответствует двухступенчатой генерации. В присутствии ловушек процессы генерации-рекомбинации идут значительно интенсивнее и время жизни носителей оказывается значительно меньше.

На рис.5 показаны двухступенчатые процессы. Электрон может перейти на уровень ловушки из зоны проводимости и затем соединиться с дыркой – рекомбинация (а) или из валентной зоны (оставляя дырку) и затем в зону проводимости – генерация (б). Ловушка может также захватить дырку, а затем на этот уровень может перейти свободный электрон (в) – рекомбинация. Наконец, атом примеси может отдать свой электрон в зону проводимости, а затем захватить электрон из валентной зоны (г) – генерация. Штриховые стрелки показывают, что захваченный носитель может также вернуться обратно.

Роль ловушек могут играть не только примесные атомы, но и различные дефекты кристаллической решетки. Из-за нарушения кристаллической решетки у поверхности полупроводника образуется множество дополнительных уровней в запрещенной зоне, которые и служат центрами рекомбинации. Поэтому время жизни носителей у поверхности оказывается значительно меньше, чем в объеме, и, если в объеме создана неравновесная концентрация, возникает поток носителей к поверхности, где они рекомбинируют.

2. ЗАКОНЫ ДВИЖЕНИЯ НОСИТЕЛЕЙ В ПОЛУПРОВОДНИКАХ

Движение электрона в кристалле подчиняется законам квантовой механики и описывается волновой функцией. В пространстве, свободном от полей, потенциальная энергия $U(x,y,z)$ имеет постоянное значение (ее можно положить равной нулю, так как энергия определена с точностью до постоянной) и энергия ε , импульс \vec{p} и скорость электрона \vec{v} определяются так же, как в классической механике:

$$\varepsilon = \frac{p^2}{2m_0}, \quad (12)$$

$$\vec{p} = m_0 \vec{v}, \quad (13)$$

где ε и m_0 – кинетическая энергия и масса электрона.

Волновая функция свободного электрона имеет вид плоской волны. Импульс электрона связан с волновым вектором \vec{k} соотношением

$$\vec{p} = \hbar \vec{k}, \quad (14)$$

где \hbar – постоянная Планка.

В кристалле потенциальная энергия $U(x,y,z)$ является периодической функцией, имеющей периодичность решетки. При сохранении полной энергии кинетическая энергия и импульс электрона не остаются постоянными, а испытывают периодические флуктуации. В кристалле роль импульса выполняет *квазиимпульс*, определяемый равенством, аналогичным (13):

$$\vec{P} = \hbar \vec{k} \quad (15)$$

Квазиимпульс обладает многими свойствами, присущими импульсу, и подчиняется закону сохранения. В отличие от электронов в вакууме связь энергии и квазиимпульса электронов в кристалле зависит от кристаллографического направления вследствие анизотропии кристалла. По отношению к внешним силам \vec{F} квазиимпульс ведет себя как импульс электрона:

$$\frac{d\vec{P}}{dt} = \vec{F} \quad (16)$$

Зависимость энергии электрона от квазиимпульса имеет достаточно сложный вид. Однако в полупроводниках в отличие от металлов валентная зона почти заполнена, а зона проводимости почти пустая. Электроны в зоне проводимости занимают уровни с наименьшей энергией, т.е. уровни вблизи абсолютного минимума (дна) зоны проводимости, а дырки – уровни вблизи абсолютного максимума (потолка) валентной зоны. Поэтому значение имеет в основном вид зависимости вблизи краев зон.

Зонная структура имеет самый простой вид, когда абсолютные экстремумы располагаются при $\vec{P} = 0$. В этом случае зависимость энергии от P_x будет иметь вид, показанный на рис.6; аналогичный вид имеет зависимость и от других компонент вектора \vec{P} .

Частные производные первого порядка в точке экстремума функции равны нулю, поэтому разложение функции $\varepsilon(\vec{P}) = \varepsilon(P_x, P_y, P_z)$ в ряд по степеням P_x, P_y, P_z вблизи минимума энергии для малых приращений имеет вид:

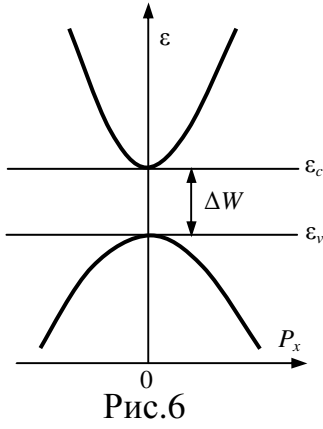


Рис.6

$$\varepsilon(P_x, P_y, P_z) = \varepsilon_c + AP_x^2 + BP_y^2 + CP_z^2, \quad (17)$$

$$\text{где } A = \frac{1}{2} \frac{\partial^2 \varepsilon}{\partial P_x^2}; \quad B = \frac{1}{2} \frac{\partial^2 \varepsilon}{\partial P_y^2}; \quad C = \frac{1}{2} \frac{\partial^2 \varepsilon}{\partial P_z^2}. \quad (18)$$

В кубическом кристалле можно принять $A=B=C$. Тогда, если отсчитывать энергию от дна зоны, выражение для энергии при малых значениях P можно записать в виде

$$\varepsilon = \frac{1}{2m_n} (P_x^2 + P_y^2 + P_z^2) = \frac{P^2}{2m_n}. \quad (19)$$

Это выражение по форме совпадает с (12), в котором импульс электрона заменен квазиимпульсом, а масса электрона в вакууме заменена величиной m_n , называемой *эффективной массой* электрона. В данном случае она не зависит от направления. Эффективная масса определяется по формуле

$$m_n = 1/(\partial^2 \varepsilon / \partial P^2) \quad (20)$$

Аналогично вблизи максимума энергии валентной зоны энергию электрона можно записать в виде

$$\varepsilon = \varepsilon_v - \frac{1}{2m_p} (P_x^2 + P_y^2 + P_z^2) = \varepsilon_v - \frac{P^2}{2m_p}. \quad (21)$$

Следовательно, электроны вблизи максимума энергии ведут себя как частицы с отрицательной эффективной массой, т.е. под действием внешней силы ускоряются в направлении противоположном ускорению свободного электрона. Это соответствует частице с положительным зарядом и с положительной эффективной массой m_p , т.е. дырке. Уменьшение энергии электрона в валентной зоне соответствует увеличению энергии дырки. Масса дырки определяется аналогично формуле (20):

$$m_p = -1/(\partial^2 \varepsilon / \partial P^2). \quad (22)$$

Скорости электронов и дырок связаны с квазиимпульсом соотношением, аналогичным (13):

$$\vec{v}_n = \vec{P} / m_n; \quad \vec{v}_p = \vec{P} / m_p. \quad (23)$$

Полупроводники, у которых абсолютные экстремумы зоны проводимости и валентной зоны соответствуют $\vec{P} = 0$, называются *прямозонными*. В них возможны прямые переходы из валентной зоны в зону проводимости с поглощением фотона и обратно с излучением фотона без участия фонона.

В непрямозонных полупроводниках дно зоны проводимости (абсолютный минимум) может располагаться на одном из кристаллографических направлений (и на эквивалентных направлениях) при определенном значении квазиимпульса $\vec{P}_0 \neq 0$. Тогда функция $\varepsilon(P_x, P_y, P_z)$ для малых приращений вблизи минимума в общем случае имеет вид

$$\varepsilon(P_x, P_y, P_z) = \varepsilon_c + \frac{1}{2m_x} (P_x - P_{x0})^2 + \frac{1}{2m_y} (P_y - P_{y0})^2 + \frac{1}{2m_z} (P_z - P_{z0})^2, \quad (24)$$

где m_x, m_y, m_z – эффективные массы электрона по соответствующим направлениям, которые могут быть различны.

Компоненты скорости электрона в этом случае равны

$$v_x = (P_x - P_{x0})/m_x; \quad v_y = (P_y - P_{y0})/m_y; \quad v_z = (P_z - P_{z0})/m_z \quad (25)$$

Движение электронов под действием внешних сил определяется средней эффективной массой, определяемой по формуле

$$\frac{1}{m_n} = \frac{1}{3} \left(\frac{1}{m_x} + \frac{1}{m_y} + \frac{1}{m_z} \right). \quad (26)$$

Аналогичные соотношения применимы для дырок.

На рис.7 показаны примеры зависимостей $\varepsilon(\vec{P})$ в кремнии (а) и арсениде

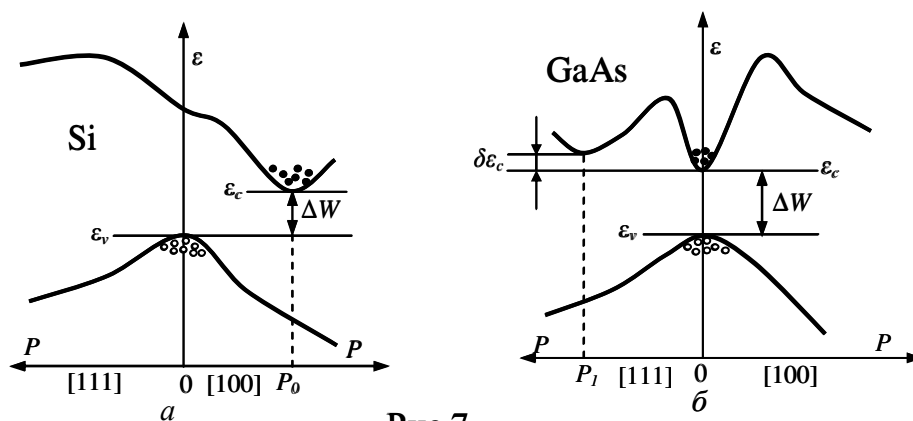


Рис.7

галлия (б) от кристаллографических направлений [100] и [111]. Верхние кривые соответствуют зоне проводимости, нижние – валентной зоне. Разность энергий между абсолютными экстремумами равна ширине запрещенной зоны.

Кремний относится к непрямозонным полупроводникам: дно зоны проводимости располагается в точке $(P_0, 0, 0)$ и других эквивалентных точках: $(-P_0, 0, 0)$, $(0, \pm P_0, 0)$ и $(0, 0, \pm P_0)$ – всего шесть минимумов. Средние эффективные массы электронов и дырок в кремнии: $m_n = 0,33m_0$, $m_p = 0,55m_0$ (m_0 – масса электрона в вакууме).

Арсенид галлия – прямозонный полупроводник. Абсолютный минимум располагается в точке $\vec{P} = 0$, но имеется еще один минимум при $\vec{P} = \vec{P}_1$, расположенный по оси энергии на $\delta\varepsilon = 0,36$ эВ выше первого. Вблизи первого минимума зависимость $\varepsilon(\vec{P})$ – резкая, и ей соответствует малая эффективная масса $m_{n1} = 0,07m_0$, второму более плавному минимуму соответствует большая эффективная масса $m_{n2} = 1,2m_0$. В равновесном состоянии электроны заполняют уровни в нижней долине вблизи дна зоны проводимости.

Тепловое движение. В идеальном бесконечном кристалле при отсутствии внешних сил свободные носители двигались бы беспрепятственно, сохраняя свою энергию и скорость. Реальный кристалл имеет конечные размеры и в нем всегда присутствуют дефекты кристаллической решетки, на которых происходит рассеяние носителей, т.е. изменение направления и скорости их движения при взаимодействии с дефектом. Важнейшими дефектами являются ионы донорных и

акцепторных примесей, поскольку практически в любом приборе применяются примесные полупроводники. Идеальность решетки нарушается также из-за тепловых колебаний атомов, которые вызывают рассеяние, называемое фононным, так как оно сопровождается возбуждением или поглощением фононов.

Свободные носители в полупроводниках можно рассматривать как электронный или дырочный газ, заполняющий объем кристалла, в котором частицы находятся в непрерывном хаотическом тепловом движении. Средняя скорость теплового движения определяется по известной формуле $v_T = \sqrt{3kT/m}$, где m – эффективная масса частицы. При комнатной температуре $v_T \approx 10^7$ см/с. Время пролета $t_{пр}$ и длина свободного пробега $l_{пр} = v_T \cdot t_{пр}$ определяется процессами рассеяния, т.е. зависит от концентрации центров рассеяния (в основном, ионизированных примесных атомов) и температуры.

При хаотическом движении направления скоростей носителей равновероятны и электрический ток равен нулю.

Дрейфовое движение. Если в полупроводнике создано электрическое поле, носители ускоряются этим полем на пути свободного пробега. Приобретенная скорость теряется при рассеянии, затем носители снова ускоряются до следующего столкновения. В результате хаотическое движение носителей приобретает составляющую, направленную вдоль вектора напряженности поля. Направленное движение носителей под действием электрического поля называется *дрейфом* (*дрейфовым движением*). Скорость направленного движения, усредненная по всем носителям одного знака, называется *скоростью дрейфа* $v_{др}$. Скорость дрейфа пропорциональна напряженности электрического поля E :

$$\vec{v}_{др} = \pm \mu \cdot \vec{E} \quad (27)$$

Коэффициент пропорциональности μ , см²/(В·с), называется подвижностью. Здесь знак «+» относится к дыркам, «-» – к электронам. Далее будем рассматривать одномерный случай и опускать обозначение векторной величины.

Направленное движение носителей заряда – это электрический ток. Плотности электронной и дырочной составляющих дрейфового тока равны произведению плотности соответствующих зарядов на скорость дрейфа:

$$(j_n)_{др} = -en(v_n)_{др} = e n \mu_n E \quad (28a)$$

$$(j_p)_{др} = ep(v_p)_{др} = e p \mu_p E, \quad (28б)$$

где μ_n и μ_p – подвижности электронов и дырок, e – элементарный заряд. Плотность дрейфового тока равна

$$j_{др} = (j_n)_{др} + (j_p)_{др} = e(n\mu_n + p\mu_p)E \quad (29)$$

Здесь коэффициент пропорциональности перед E – удельная электропроводность σ :

$$\sigma = e(n\mu_n + p\mu_p) \quad (30)$$

Диффузия. При неравномерном распределении частиц в объеме возникает направленное движение, стремящееся выровнять концентрацию, которое называется диффузией. Скорость диффузии пропорционально градиенту

концентрации с обратным знаком. В случае заряженных частиц это приводит к появлению диффузионных токов:

$$(j_n)_{\text{дф}} = eD_n(dn/dx) \quad \text{и} \quad (j_p)_{\text{дф}} = -eD_p(dp/dx), \quad (31)$$

где D_n и D_p – коэффициенты диффузии электронов и дырок, см²/с. Ток дырок направлен против градиента концентрации вследствие их положительного заряда, электронный ток противоположен направлению их движения.

Электронный и дырочный токи включают обе составляющие (28) и (31):

$$j_n = en\mu_n E + eD_n(dn/dx), \quad (32)$$

$$j_p = ep\mu_p E - eD_p(dp/dx). \quad (33)$$

Подвижность и коэффициент диффузии связаны соотношением Эйнштейна:

$$D = \varphi_T \cdot \mu \quad (34)$$

где φ_T называется температурным потенциалом и определяется формулой:

$$\varphi_T = \frac{kT}{e} \quad (35)$$

Уравнение непрерывности. Это уравнение связывает скорость изменения концентрации носителей с их распределением в пространстве.

Рассмотрим одномерную задачу. Выделим элемент объема полупроводника с единичной площадью толщиной dx (рис.8). Слева в этот объем втекает ток $j(x)$ и справа вытекает $j(x+dx)$. Разность токов, деленная на объем dx дает скорость изменения плотности заряда. Но эта величина есть частная производная $\partial j/\partial x$. Поделив ее на заряд носителя, получаем скорость изменения концентрации за счет разности токов. С учетом процессов генерации и рекомбинации можно записать

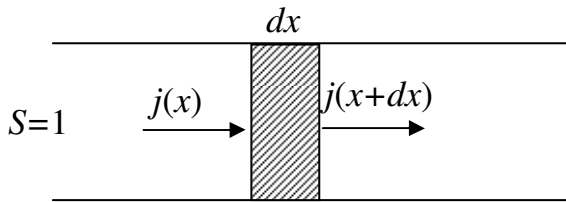


Рис.8

$$\frac{\partial p}{\partial t} = \Delta G - \frac{p - p_0}{\tau} - \frac{1}{e} \cdot \frac{\partial j_p}{\partial x};$$

$$\frac{\partial n}{\partial t} = \Delta G - \frac{n - n_0}{\tau} + \frac{1}{e} \cdot \frac{\partial j_n}{\partial x},$$

где ΔG – скорость генерации носителей за счет внешних факторов, вторые члены дают скорость рекомбинации избыточных

(неравновесных) носителей.

Когда внешние факторы отсутствуют ($\Delta G=0$), то, подставляя в эти уравнения выражения (32), (33) и дифференцируя их, получаем:

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau} + D_p \frac{\partial^2 p}{\partial x^2} - \mu_p E \frac{\partial p}{\partial x} - p\mu_p \frac{\partial E}{\partial x}; \quad (36a)$$

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau} + D_n \frac{\partial^2 n}{\partial x^2} + \mu_n E \frac{\partial n}{\partial x} + n\mu_n \frac{\partial E}{\partial x}. \quad (36b)$$

Это и есть уравнения непрерывности для дырок и электронов в общем виде. В объеме полупроводника сохраняется электронейтральность, поэтому, согласно уравнению Пуассона, $\partial E/\partial x=0$, уравнения упрощаются:

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau} + D_p \frac{\partial^2 p}{\partial x^2} - \mu_p E \frac{\partial p}{\partial x}; \quad (37a)$$

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau} + D_n \frac{\partial^2 n}{\partial x^2} + \mu_n E \frac{\partial n}{\partial x}. \quad (37б)$$

Наличие электрического поля E не приводит к нарушению условия электронейтральности. Оно может существовать в объеме полупроводника даже в равновесном состоянии при неоднородном легировании, и в общем случае его нужно учитывать. В частном случае ($E=0$) уравнение непрерывности превращается в уравнение диффузии:

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau} + D_p \frac{\partial^2 p}{\partial x^2}; \quad (38а)$$

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau} + D_n \frac{\partial^2 n}{\partial x^2}. \quad (38б)$$

Эффект Холла. Эффектом Холла называется возникновение поперечной эдс при прохождении электрического тока через проводник, помещенный в магнитное поле. Этот эффект был открыт Холлом в 1879 г., еще до открытия электрона, и впоследствии оказался тем ключом, который позволил раскрыть сущность процесса проводимости в полупроводниках и провести грань между полупроводниками и другими типами плохо проводящих веществ.

Рассмотрим проводник в виде прямоугольной пластины, ширина и толщина которой малы по сравнению с длиной (рис.9). В пластине в направлении x

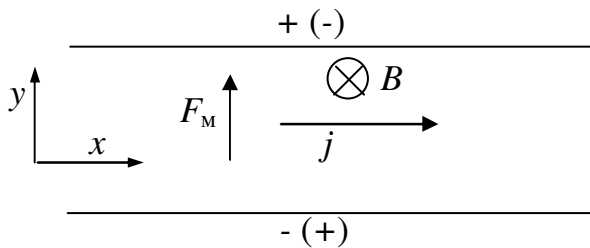


Рис.9.

протекает электрический ток плотностью j и приложено магнитное поле B под прямым углом к току так, чтобы оно уходило в плоскость чертежа. Движущиеся носители будут испытывать действие магнитной силы $\vec{F}_M = [e\vec{v}\vec{B}]$, в данном случае эта сила направлена вверх по оси y как для

дырок, так и для электронов (заряд электронов отрицательный, но скорость противоположна току). Верхняя поверхность будет заряжаться положительно относительно нижней, если носителями тока являются дырки, и отрицательно, если электроны. Заряды на верхней и нижней поверхностях будут накапливаться до тех пор, пока электрические силы не погасят действие магнитной силы:

$$e|E_y| = e|v_x|B, \quad (39)$$

где E_y – напряженность поперечного электрического поля. Разность потенциалов между верхом и низом кристалла равна этой напряженности умноженной на высоту пластины. Эффект Холла описывается с помощью коэффициента Холла R_H , определяемого равенством

$$E_H = R_H j B, \quad (40)$$

где для наших обозначений $E_H = E_y$. Выразив v_x в уравнении (39) через j согласно равенствам (28) и сравнив с уравнением (40), получаем для случая, когда основными носителями являются электроны:

$$R_H = -\frac{1}{ne}, \quad (41a)$$

а для случая, когда ток является дырочным

$$R_H = \frac{1}{pe}. \quad (41б)$$

Измерение коэффициента Холла позволяет определить знак и концентрацию основных носителей заряда в полупроводниках. Коэффициент Холла имеет размерность L^3Q^{-1} , где Q – электрический заряд, обычно используют единицы $\text{см}^3\text{Кл}^{-1}$.

Интересно отметить парадоксальный факт, что в некоторых металлах, например в бериллии, где, казалось бы, единственными носителями заряда являются электроны, знак коэффициента Холла оказался положительным. Таким образом, в некоторых случаях и в металлах объектами, ответственными за проводимость, оказываются дырки.

ЭЛЕКТРИЧЕСКИЕ ПЕРЕХОДЫ

1. ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

Электронно-дырочный переход (*p-n*-переход) – тонкий переходной слой, образующийся между двумя областями кристалла с разным типом проводимости. Этот слой обеднен носителями заряда и обладает высоким удельным сопротивлением.

Металлургической границей называется поверхность, разделяющая *p*- и *n*-области

Различают *резкий p-n-переход* и *плавный*. В резком переходе тип проводимости резко (ступенчато) меняется на металлургической границе. В плавном переходе один тип проводимости плавно переходит в другой, металлургической границей считается поверхность, где $N_a=N_d$.

Переходы бывают *симметричные* $(N_a)_p=(N_d)_n$ и *несимметричные* $(N_a)_p \neq (N_d)_n$. Когда концентрации отличаются на 2–3 порядка, переход называется *односторонним*, концентрация носителей в сильнолегированной области обозначается верхним индексом «+»: p^+n или n^+p . В этом случае сильнолегированная область называется *эмиттером*, слаболегированная – *базой*.

Электронно-дырочный переход в равновесном состоянии

Контактная разность потенциалов. При создании контакта двух областей с разной концентрацией носителей возникает диффузия носителей в область, где их концентрация меньше: дырки диффундируют из *p*-области в *n*-область, электроны – обратно. Здесь они рекомбинируют с основными носителями и исчезают. В результате образуется слой, обедненный носителями заряда, в котором заряд ионов акцепторных и донорных примесей не компенсирован

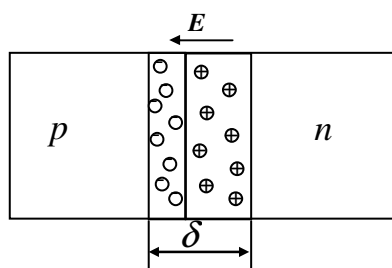


Рис.10

(рис.10). Этот объемный заряд создает электрическое поле E , которое препятствует дальнейшей диффузии носителей. При равновесии ток отсутствует, т.е. должны выполняться условия:

$$j_p = (j_p)_{дф} + (j_p)_{др} = 0 \text{ и } j_n = (j_n)_{дф} + (j_n)_{др} = 0 \quad (42)$$

При этом возникает *контактная разность потенциалов*, которую легко найти из этих условий с учетом выражений (32), (33):

$$-e\mu_p p \frac{dU}{dx} - eD_p \frac{dp}{dx} = 0$$

Используя соотношение Эйнштейна $D_p = \Phi_T \cdot \mu_p$, получаем

$$\frac{dU}{dx} = -\varphi_T \cdot \frac{dp}{dx}.$$

Интегрирование дает $U = \varphi_T \ln p + \text{const}$, отсюда контактная разность потенциалов φ_k равна

$$\varphi_k = U_n - U_p = \varphi_T \ln \frac{p_{p0}}{p_{n0}}, \quad (43a)$$

где нулевым индексом обозначены равновесные значения концентрации в соответствующих областях. Аналогично φ_k выражается через концентрацию электронов:

$$\varphi_k = \varphi_T \ln \frac{n_{n0}}{n_{p0}}. \quad (43b)$$

С учетом того, что в области рабочих температур $p_{p0} \approx N_a$, и $p_{n0} \approx n_i^2/N_d$, выражение (43a) можно записать в виде

$$\varphi_k = \varphi_T \ln \frac{N_a N_d}{n_i^2} \quad (44)$$

Следует отметить, что при выводе этих формул не использованы какие либо особенности структуры p - n -перехода, значение φ_k от них не зависит.

Распределение носителей, электрического поля и потенциала в резком p - n -переходе. Концентрация основных носителей заряда резко падает при движении вглубь p - n -перехода, сравнивается с собственной концентрацией на металлургической границе, которая принята за начало координат, и далее уменьшается до концентрации неосновных носителей. Распределение носителей показано на рис.11 в логарифмическом (*a*) и линейном (*b*) масштабе. Строго говоря, границы перехода несколько размыты, но обычно рассматривается идеальная модель с резкими границами, считая, что на этих границах концентрация носителей резко падает до нуля. Распределение плотности заряда ρ в этой модели показано на рис.11,в. Плотности заряда в p - и n -областях обедненного слоя (толщиной δ_p и δ_n , соответственно) постоянны и равны

$$\rho_p = -eN_a, \text{ в области } -\delta_p < x < 0; \quad (45a)$$

$$\rho_n = eN_d, \text{ в области } 0 < x < \delta_n, \quad (45b)$$

p - n -переход в целом электронейтрален, заряды в этих областях равны:

$$q = eN_a\delta_p = eN_d\delta_n \quad (46)$$

Отношение $\delta_p / \delta_n = N_d / N_a$, толщина p - n -перехода $\delta = \delta_p + \delta_n$, отсюда:

$$\delta_p = \frac{N_d}{N_d + N_a} \delta; \quad \delta_n = \frac{N_a}{N_d + N_a} \delta. \quad (47)$$

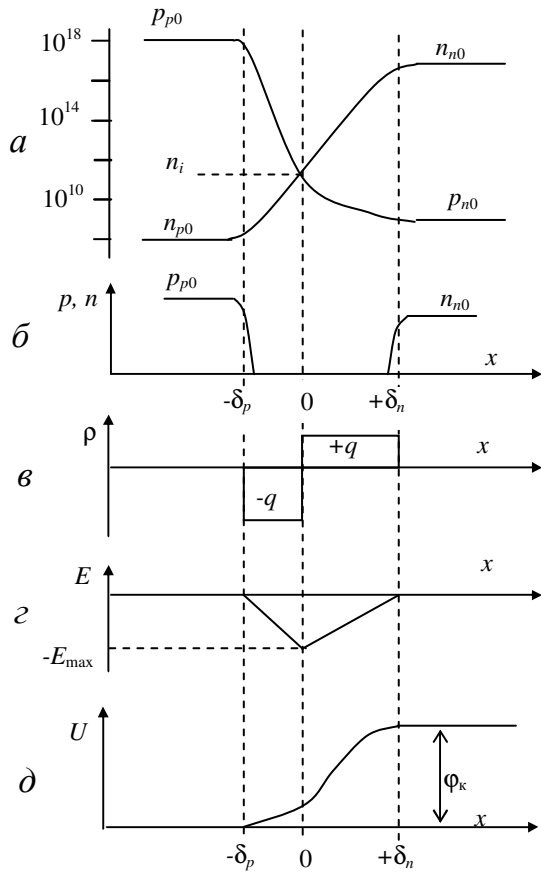


Рис.11

Напряженность электрического поля E можно найти из одномерного уравнения Пуассона

$$\frac{d^2U}{dx^2} = -\frac{\rho}{\epsilon\epsilon_0}, \quad (48)$$

где U – потенциал, ϵ – диэлектрическая проницаемость, ϵ_0 – диэлектрическая постоянная.

С учетом того, что $dU/dx = -E$, и выражений (45) это уравнение записывается в виде

$$\frac{dE}{dx} = -\frac{eN_a}{\epsilon\epsilon_0}, \dots (-\delta_p < x < 0) \quad \text{и}$$

$$\frac{dE}{dx} = \frac{eN_d}{\epsilon\epsilon_0}, \dots (0 < x < \delta_n). \quad (49)$$

Интегрируя при граничных условиях $E|_{x \leq -\delta_p} = E|_{x \geq \delta_n} = 0$, получаем для p -области линейную функцию:

$$E = -\frac{eN_a}{\epsilon\epsilon_0}(x + \delta_p), \dots (-\delta_p < x < 0) \quad (50)$$

Максимальная напряженность

достигается на металлургической границе:

$$-E_{\max} = \frac{eN_a}{\epsilon\epsilon_0}\delta_p = \frac{eN_aN_d}{\epsilon\epsilon_0(N_a + N_d)}\delta. \quad (51)$$

Знак «-» здесь показывает, что поле направлено противоположно координате x . Далее в n -области напряженность поля спадает по линейному закону до нуля (рис.11,г).

Потенциал $U(x)$ получается интегрированием функции $E(x)$ и описывается квадратичной функцией, контактная разность потенциалов, очевидно, равна площади треугольника на рис.11,г:

$$\varphi_k = \frac{1}{2}E_{\max} \cdot \delta = \frac{eN_aN_d}{2\epsilon\epsilon_0(N_a + N_d)}\delta^2. \quad (52)$$

Отсюда находим толщину p - n -перехода в равновесном состоянии:

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0\varphi_k}{e} \cdot \left(\frac{1}{N_a} + \frac{1}{N_d}\right)} \quad (53)$$

Для резко несимметричного перехода одной из составляющих δ_p и δ_n можно пренебречь, практически весь переход располагается в слабо-легированной области – базе, формула (53) при этом упрощается:

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0\varphi_k}{eN}}, \quad (53a)$$

где N – концентрация легирующей примеси в базе.

В таблице 2 приведены примеры, позволяющие судить о порядке величин φ_k , E_{\max} и δ ($T=300K$).

Таблица 2

| Полупроводник | n_i , см ⁻³ | N_a , см ⁻³ | N_d , см ⁻³ | φ_k , В | E_{\max} , кВ/см | δ , мкм |
|---------------|-----------------------------|-----------------------------|-----------------------------|--------------------|-----------------------|----------------|
| Ge | $2,4 \cdot 10^{13}$ | 10^{18} | 10^{14} | 0,36 | 3 | 2,4 |
| Si | $2 \cdot 10^{10}$ | 10^{19} | 10^{16} | 0,83 | 55 | 0,3 |

Энергетическая диаграмма p - n -перехода приведена на рис.12. Уровень Ферми одинаков для всей равновесной системы. Между областями существует

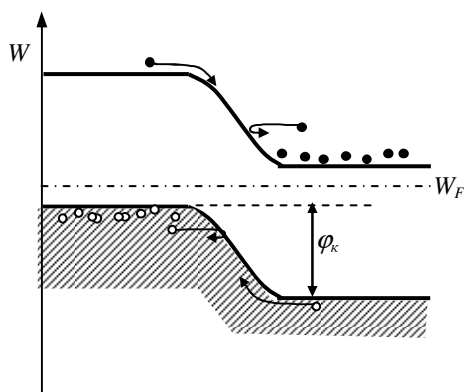


Рис.12

потенциальный барьер, равный $e\varphi_k$. Основные носители, движущиеся к переходу, отражаются обратно, так как поле в переходе для них является тормозящим. Лишь ничтожная часть их имеет энергию $W > e\varphi_k$, достаточную для преодоления потенциального барьера. Неосновные носители свободно проходят через переход, для них электрическое поле является ускоряющим. Для наглядности можно представить электроны как тяжелые шарики, катящиеся по

дну зоны проводимости, а дырки – как легкие пузырьки катящиеся, по потолку валентной зоны.

Электронно-дырочный переход в неравновесном состоянии

Если к электронно-дырочному переходу подключить источник напряжения, равновесное состояние нарушается и через переход протекает ток. Напряжение подводится с помощью двух металлических контактов, обладающих низким контактным сопротивлением (*омических контактов*), к p - и n -областям. Полученное устройство является *полупроводниковым диодом*.

Так как сопротивление обедненного слоя значительно больше сопротивления нейтральных областей, то при малом токе внешнее напряжение U практически полностью прикладывается к обедненному слою.

Под действием приложенного напряжения изменяется высота потенциального барьера $\varphi = \varphi_k - U$. Если плюс источника питания приложен к p -области, напряжение считается *прямым* ($U > 0$), при нем потенциальный барьер понижается, при *обратном* напряжении ($U < 0$) – повышается. На рис.13 приведены энергетические диаграммы перехода для двух этих случаев. При этом изменяется также толщина p - n -перехода и граничные концентрации в p - и n -областях:

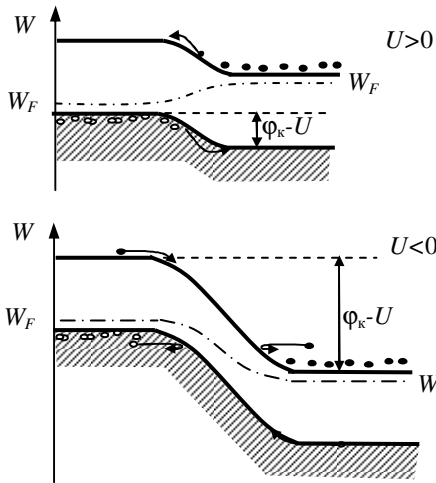


Рис.13

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0(\phi_k - U)}{e} \cdot \left(\frac{1}{N_a} + \frac{1}{N_d} \right)}, \quad (54)$$

$$\phi_k - U = \phi_T \ln \frac{p_p}{p_n} = \phi_T \ln \frac{n_n}{n_p} \quad (55)$$

Поскольку концентрация основных носителей значительно больше, чем концентрация неосновных, относительное отклонение их от равновесных значений пренебрежимо мало, изменяются, главным образом, концентрации неосновных носителей:

$$p_p \approx p_{p0}, \quad n_n \approx n_{n0}, \quad p_n \neq p_{n0}, \quad n_p \neq n_{p0}$$

С учетом этих соотношений, подставив в

(55) выражение (44), получаем

$$p_n = p_{n0} e^{U/\phi_T},$$

$$n_p = n_{p0} e^{U/\phi_T}.$$

На границах *p-n*-перехода образуется неравновесная концентрация носителей:

$$\Delta p_n(0) = p_n - p_{n0} = p_{n0} (e^{U/\phi_T} - 1); \quad (56)$$

$$\Delta n_p(0) = n_p - n_{p0} = n_{p0} (e^{U/\phi_T} - 1). \quad (57)$$

При прямом напряжении ($U > 0$) происходит перенос носителей в область, где они являются неосновными, – *инжекция* неосновных носителей: $\Delta p_n, \Delta n_p > 0$. При $U < 0$ происходит *экстракция* неосновных носителей – перенос их в область, где они являются основными: $\Delta p_n, \Delta n_p < 0$. В резко несимметричном переходе инжекция и экстракция носят односторонний характер – происходят в слаболегированной области – базе.

Вольтамперная характеристика идеализированного *p-n*-перехода

Идеализированный p-n-переход – это упрощенная модель реального *p-n*-перехода, в котором пренебрегаются процессами генерации и рекомбинации в обедненном слое и сопротивлением нейтральных областей, а также считается, что вне обедненного слоя электрическое поле отсутствует, и носители движутся только вследствие диффузии, носящей одномерный характер.

Инжектированные носители диффундируют в глубь полупроводника и рекомбинируют с основными носителями. Через *p-n*-переход протекает ток, компенсирующий отвод носителей от границы:

$$j = j_{дф} + j_{др}$$

Из уравнения диффузии (38) для установившегося режима получаем:

$$\frac{\partial p}{\partial t} = -\frac{\Delta p_n}{\tau_p} + D_p \frac{\partial^2 \Delta p_n}{\partial x^2} = 0$$

Опуская индекс при *p*, это уравнение можно записать в виде

$$\frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{L_p^2},$$

где $L_p = \sqrt{D_p \tau_p}$ – диффузионная длина, это среднее расстояние, на которое успевают диффундировать носители за время жизни τ_p .

Общее решение этого уравнения имеет вид

$$\Delta p = A_1 e^{x/L_p} + A_2 e^{-x/L_p}$$

Коэффициенты A_1 и A_2 находятся из граничных условий:

$$\Delta p|_{x=0} = \Delta p(0); \quad \Delta p|_{x \rightarrow \infty} = 0.$$

Отсюда

$$\begin{aligned} A_1 &= 0; \quad A_2 = \Delta p(0); \\ \Delta p &= \Delta p(0) \cdot e^{-x/L_p}. \end{aligned} \quad (58)$$

Градиент концентрации на границе перехода

$$\left. \frac{d\Delta p}{dx} \right|_{x=0} = -\frac{\Delta p(0)}{L_p}. \quad (59)$$

Плотность дырочного тока равна

$$j_p = -eD_p \left. \frac{d\Delta p}{dx} \right|_{x=0} = \frac{eD_p \Delta p(0)}{L_p}. \quad (60)$$

Аналогично для плотности электронного тока получаем

$$j_n = \frac{eD_n \Delta n(0)}{L_n}, \quad (61)$$

где D_n и L_n – коэффициент диффузии и диффузионная длина электронов в p -области.

Полный ток равен сумме токов электронов и дырок: $I = S(j_p + j_n)$, где S – площадь p - n -перехода. Подставив (56), (57) в (60), (61), вольт-амперную характеристику идеализированного p - n -перехода можно записать в виде

$$I = I_0 (e^{U/\Phi_T} - 1), \quad (62)$$

где коэффициент I_0 называется тепловым током:

$$I_0 = eS \left(\frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n} \right). \quad (63)$$

Подставив сюда p_{n0} и n_{p0} из (10), выражение (63) можно представить в виде:

$$I_0 = eS n_i^2 \left(\frac{D_p}{L_p N_d} + \frac{D_n}{L_n N_a} \right) \quad (64)$$

Тепловой ток сильно зависит от температуры. Основной вклад в эту зависимость вносит множитель n_i^2 . С учетом выражения (7) формулу (64) можно представить в виде

$$I_0 = I_{00} \exp\left(-\frac{\Delta W}{kT}\right), \quad (65)$$

где множитель I_{00} слабо зависит от температуры, и в первом приближении этим можно пренебречь.

Напряжение открытого р-п-перехода. Теоретическая ВАХ, описываемая

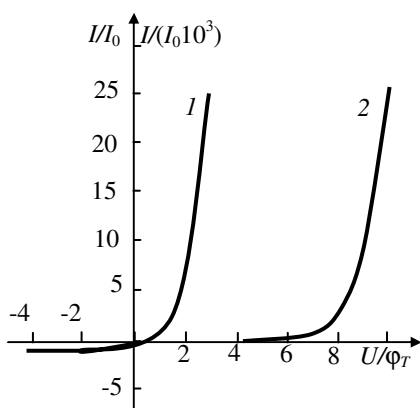


Рис.14

формулой (62) для малых прямых напряжений приведена на рис.14 (кривая 1, шкала I/I_0). При прямом напряжении ток резко возрастает, увеличиваясь на порядок при приращении напряжения на $2,3\phi_T$ (60 мВ при $T = 300$ К).

При обратных напряжениях, превышающих по модулю $2-3\phi_T$, ток не зависит от напряжения и равен I_0 . Рабочие прямые токи р-п-переходов на 3–4 и более порядков превышают I_0 . Кривая 2 на рис.14 отображает ВАХ в масштабе рабочих токов (шкала отсчета

$I/(10^3 I_0)$). В этом масштабе имеется «пятка» прямой ветви ВАХ, где ток очень мал и на графике не виден, затем характеристика круто идет вверх, прямое напряжение меняется незначительно и в первом приближении его можно считать постоянным. Это напряжение называется *напряжением открытого р-п-перехода* и обозначается U^* . *Напряжением отпираания р-п-перехода* можно считать величину $U^* \approx 0,1$ В, где ток составляет около 2% от значения рабочих токов (при $T = 300$ К). Тепловые токи кремниевых и германиевых р-п-переходов при прочих равных условиях отличаются на шесть порядков, поэтому при одинаковых рабочих токах их характеристики существенно отличаются. При работе в нормальном режиме для кремниевого р-п-перехода можно принять $U^* = 0,7$ В, для микрорежима – $U^* = 0,5$ В. Для германиевого р-п-перехода в нормальном режиме $U^* = 0,15-0,25$ В и напряжение отпираания близко к нулю.

Дифференциальные параметры идеализированного р-п-перехода.

Дифференциальными параметрами называются величины, связывающие малые приращения зависимых и независимых переменных, т.е. это частные производные зависимых переменных по независимым. Ток р-п-перехода является функцией напряжения и температуры, и для малых приращений можно записать

$$\Delta I = S \cdot \Delta U + S_{I(T)} \cdot \Delta T,$$

где $S = \left(\frac{\partial I}{\partial U} \right)_T$ – крутизна;

$S_{I(T)} = \left(\frac{\partial I}{\partial T} \right)_U$ – дифференциальная температурная чувствительность тока.

Часто вместо крутизны используется дифференциальное сопротивление

$$r_d = S^{-1} = \left(\frac{\partial U}{\partial I} \right)_T.$$

Важным параметром является также дифференциальная температурная чувствительность напряжения $S_{U(T)} = \left(\frac{\partial U}{\partial T} \right)_I$. Эти параметры связаны соотношением $S_{U(T)} = -S_{I(T)}/S$.

Для прямой ветви ВАХ при $I \gg I_0$ формулу (62) с учетом (65) можно записать в виде

$$I \approx I_{00} \exp\left(-\frac{eU - \Delta W}{kT}\right). \quad (66)$$

Дифференцируя это выражение, получаем

$$S = I/\varphi_T; \quad (67)$$

$$S_{I(T)} = -\frac{eU - \Delta W}{kT^2} \cdot I; \quad (68)$$

$$S_{U(T)} = \frac{U - \Delta W/e}{T}. \quad (69)$$

Поскольку $U < \Delta W$; $S_{I(T)} > 0$, $S_{U(T)} < 0$.

Вольтамперная характеристика реального p - n -перехода

Прямая ветвь ВАХ. При подаче прямого напряжения дырки и электроны движутся навстречу друг другу, и часть их рекомбинирует в p - n -переходе, образуя дополнительный прямой ток, называемый *током рекомбинации*. Полный ток складывается из тока инжекции, определяемого выражением (62), и тока рекомбинации. Следовательно, ток реального p - n -перехода больше, чем идеализированного. Ток рекомбинации пропорционален n_i и $\exp[U/(2\varphi_T)]$, тогда как ток инжекции пропорционален n_i^2 и $\exp(U/\varphi_T)$ и поэтому более резко зависит от температуры, ширины запрещенной зоны и приложенного напряжения. Ток рекомбинации следует учитывать в полупроводниках с широкой запрещенной зоной при работе в области малых токов и при пониженных температурах. При работе в нормальном режиме током рекомбинации можно пренебречь.

Одной из особенностей реальной ВАХ является падение напряжения на омическом сопротивлении базы r_6 . Приложенное внешнее напряжение не полностью падает на p - n -переходе, а распределяется между ним и слоем базы. Тогда

$$I = I_0 \left(\exp \frac{U - I \cdot r_6}{\varphi_T} - 1 \right). \quad (70)$$

Падение напряжения в слое базы сказывается при больших токах. Пренебрегая единицей в формуле (70) при $I \gg I_0$, можно записать

$$U = \varphi_T \cdot \ln(I/I_0) + I \cdot r_6 \quad (71)$$

Обратная ветвь ВАХ. Реальный обратный ток кроме теплового тока I_0 содержит еще две составляющие – ток *термогенерации* $I_{ген}$ и ток *утечки* $I_{ут}$:

$$I_{обр} = I_0 + I_{ген} + I_{ут} \quad (72)$$

Процессы генерации и рекомбинации носителей происходят во всем объеме полупроводника как в нейтральных областях, так и в области перехода. В равновесии их скорости равны, поэтому направленных потоков нет. Когда к переходу приложено обратное напряжение, область перехода дополнительно обедняется носителями, рекомбинация замедляется и равновесие нарушается. Неравновесные носители ускоряются электрическим полем перехода и электроны выбрасываются в n -область, дырки – в p -область. Ток термогенерации пропорционален объему перехода, поэтому растет с ростом обратного смещения.

Реальные p - n -переходы имеют участки, выходящие на поверхность полупроводникового кристалла. На поверхности вследствие загрязнений и влияния поверхностного заряда могут образоваться проводящие пленки и каналы, по которым протекает ток утечки. Он увеличивается пропорционально напряжению и при достаточно большом обратном напряжении может превысить тепловой ток и ток термогенерации. Специальные защитные покрытия поверхности позволяют подавить утечку. В кремниевых приборах поверхность кристалла покрыта защитным слоем оксида и ток утечки, как правило, пренебрежимо мал.

Пробой p - n -перехода.

Когда обратное напряжение превышает некоторое критическое значение, обратный ток p - n -перехода начинает резко возрастать. Это явление называется пробоем p - n -перехода. В зависимости от характера процессов в p - n -переходе различают *электрический* и *тепловой* пробой. Электрический пробой бывает двух видов – *лавинный* и *туннельный*.

Лавинный пробой. В основе этого пробоя лежит ударная ионизация и лавинообразное умножение носителей в p - n -переходе в сильном электрическом поле. Когда напряженность поля достигает значений 80–120 кВ/см, электроны и дырки на пути свободного пробега приобретают энергию, достаточную для того, чтобы выбить электрон из валентной связи. В результате рождается новая пара носителей, и далее этот процесс повторяется уже с участием новых носителей. Этот процесс характеризуется коэффициентом лавинного умножения:

$$M = n_2 / n_1,$$

где n_1 – число носителей поступающих в p - n -переход, n_2 – число носителей покидающих p - n -переход. Размноженные электроны выбрасываются в n -область, дырки – в p -область. Зависимость M от напряжения определяется выражением

$$M = \frac{1}{1 - (U / U_{\text{л}})^k}, \quad (73)$$

где $U_{\text{л}}$ – напряжение лавинного пробоя; $k=2-6$ в зависимости от материала и типа p - n -перехода; $M \rightarrow \infty$ при $U \rightarrow U_{\text{л}}$.

Напряжение $U_{\text{л}}$ зависит от материала и степени легирования базы, оно тем выше, чем выше удельное сопротивление базы ρ_6 :

$$U_{\text{л}} = a \rho_6^m; \quad m = 0,6-0,75.$$

Туннельный пробой. Это вид пробоя возникает в узком переходе при высокой степени легирования ($\rho_6 \approx 10^{-3}$ Ом·м), когда ширина перехода составляет доли мкм, а напряженность поля достигает значения 200–400 кВ/см. При туннельном пробое электроны «просачиваются» (туннелируют), не изменяя свою энергию, через тонкий потенциальный барьер из валентной зоны p -области в зону

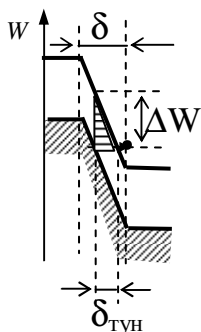


Рис.15

проводимости n -области (рис.15). Это квантово-механическое явление, обусловленное волновыми свойствами частицы. Туннельный переход возможен только в случае, если по другую сторону барьера есть разрешенные уровни с такой же энергией. Вероятность туннельного перехода зависит от высоты барьера ΔW и расстояния $\delta_{\text{тун}}$, которое при этом проходит частица. Из рис.15 видно, что $\delta_{\text{тун}}$ меньше ширины p - n -перехода δ . Напряжение туннельного пробоя $U_{\text{тун}}$ линейно зависит от удельного сопротивления базы.

Электрический пробой *обратим* и используется для стабилизации напряжения.

Тепловой пробой. В основе теплового пробоя лежит разогрев p - n -перехода при протекании обратного тока. С ростом температуры обратный ток резко возрастает и увеличивается мощность, рассеиваемая в p - n -переходе, что приводит к дальнейшему повышению температуры, росту тока и, в конечном итоге, к плавлению полупроводника и выходу прибора из строя. Тепловой пробой, как правило, *необратим* в отличие от электрического пробоя.

Мощность, подводимая к p - n -переходу, равна $P_{\text{подв}} = U_{\text{обр}} I_{\text{обр}}$, отводимая мощность определяется теплопроводностью:

$$P_{\text{отв}} = \frac{T_{\text{п}} - T_0}{R_T},$$

где $T_{\text{п}}$ – температура перехода, T_0 – температура поверхности кристалла, R_T – тепловое сопротивление (К/Вт).

С повышением напряжения растет как подводимая, так и отводимая мощность; пока напряжение остается меньше напряжения теплового пробоя U_T , сохраняется динамическое равновесие: $P_{\text{отв}} = P_{\text{подв}}$. При превышении критического значения U_T выделяющееся тепло не успевает отводиться от p - n -перехода, и температура начинает неограниченно возрастать, ток при этом растет, а напряжение падает. Таким образом, ВАХ при тепловом пробое имеет участок с отрицательным дифференциальным сопротивлением.

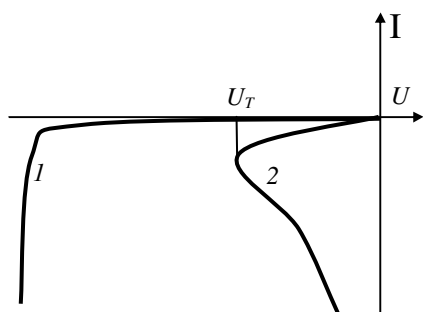


Рис.16

На рис.16 показан вид обратной ветви ВАХ при электрическом (кривая 1) и при тепловом пробое (кривая 2). Напряжение U_T зависит от ΔW , $I_{\text{обр}}$ и R_T . Германиевые p - n -переходы с малой шириной запрещенной зоны ΔW и с большим тепловым током склонны к тепловому пробое. В кремниевых p - n -переходах $U_{\text{п}}$, $U_{\text{тун}} < U_T$ и тепловой пробой самостоятельно возникнуть, как правило, не может. Он может произойти только после

электрического пробоя, когда ток превысит допустимое значение.

Емкость *p-n*-перехода

Диод обладает емкостными свойствами, т.е. способен накапливать и отдавать заряд при увеличении или уменьшении приложенного напряжения. Накопление заряда происходит в *p-n*-переходе и в базе диода, и в соответствии с этим различают две составляющие емкости – барьерную и диффузионную:

$$C_D = C_{\text{бар}} + C_{\text{дф}}$$

Барьерная емкость. При изменении напряжения меняется толщина (ширина) *p-n*-перехода и в соответствие с этим меняется некомпенсированный заряд ионов донорных и акцепторных примесей в переходе. Дифференциальную емкость $C_{\text{бар}}$ можно рассматривать как емкость плоского конденсатора с обкладками, расположенными на расстоянии, равном толщине перехода δ :

$$C_{\text{бар}} = \epsilon \epsilon_0 S / \delta, \quad (74)$$

где ϵ – диэлектрическая проницаемость, ϵ_0 – диэлектрическая постоянная, S – площадь перехода.

Подставляя значение δ для резкого одностороннего перехода, получаем

$$C_{\text{бар}} = S \sqrt{\frac{\epsilon \epsilon_0 N}{2(\phi_K - U)}} = \frac{C_0}{(1 - U / \phi_K)^{1/2}}, \quad (75)$$

где N – концентрация легирующей примеси в базе, C_0 – емкость при $U=0$.

В общем случае

$$C_{\text{бар}} = \frac{C_0}{(1 - U / \phi_K)^\gamma}, \quad (76)$$

где $\gamma=1/2$ для резкого перехода и $\gamma=1/3$ – для плавного.

Диффузионная емкость обусловлена накоплением неравновесных носителей в базе. Объем при этом сохраняет электронейтральность – заряд накапливается не в виде электрического заряда, а в виде избыточной концентрации.

Для стационарного режима нетрудно найти заряд, накопленный в базе, проинтегрировав функцию распределения неравновесных носителей (58) по x

$$Q_\delta = Se \int_0^\infty \Delta p(0) \cdot e^{-x/L_p} \cdot dx = Se \Delta p(0) L_p.$$

Выразив затем $\Delta p(0)$ через ток $I = e S j_p(0) = e S D_p \Delta p(0) / L_p$, получаем

$$Q_\delta = I \tau = I_0 \tau (e^{U/\phi_T} - 1), \quad (77)$$

где $\tau = L_p^2 / D_p$ – время жизни неравновесных носителей в базе (индекс p для общности опущен).

Эта формула относится к диодам с толстой базой, когда толщина базы (расстояние до омического контакта) $w \gg L$, где L – диффузионная длина неравновесных носителей. Дифференцируя (77) по U , получаем

$$C_{\text{дф}} = (I + I_0) \tau / \phi_T \quad (78)$$

При обратном соотношении $w \ll L$ распределение концентрации в базе является практически линейным, и $Q_\delta = I \cdot \tau_{\text{дф}}$, где $\tau_{\text{дф}}$ – среднее время диффузии через базу:

$$\tau_{дф} = w^2 / (2D). \quad (79)$$

В этом случае

$$C_{дф} = (I + I_0) \tau_{дф} / \varphi_T. \quad (80)$$

$C_{дф}$ и $C_{бар}$ проявляются в разных режимах работы: $C_{дф} \gg C_{бар}$ при $I \gg I_0$, и $C_{бар} \gg C_{дф}$ при $U < 0$. $C_{дф}$ нельзя считать «чистой» емкостью, так как она зависит от частоты или длительности импульса в импульсном режиме работы.

2. КОНТАКТЫ МЕТАЛЛ-ПОЛУПРОВОДНИК

Контакты металла с полупроводником могут быть выпрямляющими (называются также контактами Шоттки) и невыпрямляющими (омическими). Тип контакта зависит от разности *термоэлектронных работ выхода* электрона из металла A_M и из полупроводника A_S и от типа проводимости полупроводника. Термоэлектронная работа выхода – это разность между энергией электрона в вакууме и уровнем Ферми:

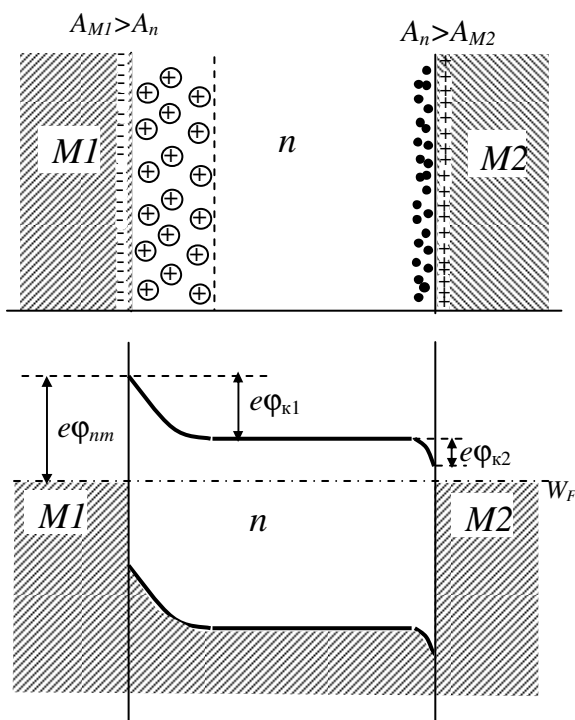
$$A = W_{вак} - W_F. \quad (81)$$

Чем больше A , тем меньше энергия электронов, поэтому электроны переходят в область с большей работой выхода: если $A_M > A_S$, электроны переходят из полупроводника в металл, если $A_M < A_S$, – из металла в полупроводник. В равновесии уровни Ферми выравниваются, между областями возникает контактная разность потенциалов:

$$\varphi_k = (A_M - A_S) / e \quad (82)$$

Слой полупроводника, прилегающий к контакту, обогащается или обедняется основными носителями и образуется омический или выпрямляющий контакт. Если $A_M > A_S$, то контакт $M-n$ – выпрямляющий, $M-p$ – омический. Если $A_M < A_S$, контакт $M-n$ – омический, $M-p$ – выпрямляющий.

На рис.17 показаны оба типа контакта с полупроводником n -типа (эта структура представляет собой диод Шоттки) и приведена энергетическая диаграмма.



В выпрямляющем контакте обедненный слой расположен целиком в полупроводнике. Он обладает высоким сопротивлением и нелинейной ВАХ. Толщина этого слоя определяется так же, как и толщина $p-n$ -перехода:

$$\delta = \sqrt{\frac{2\epsilon\epsilon_0\varphi_k}{eN}}, \quad (83)$$

где N – концентрация легирующей примеси в полупроводнике.

Потенциальный барьеры для электронов, движущихся из металла в

полупроводник и из полупроводника в металл различны и равны, соответственно, ϕ_{nm} и $\phi_{к1}$. Высота барьера ϕ_{nm} определяется выражением

$$\phi_{nm} = (A_M - \chi_n) / e, \quad (84)$$

где χ_n называется энергией электронного средства, или средством к электрону, и определяется как минимальная энергия, необходимая для удаления электрона со дна зоны проводимости полупроводника в вакуум.

На рис.18 приведены энергетические диаграммы выпрямляющего контакта

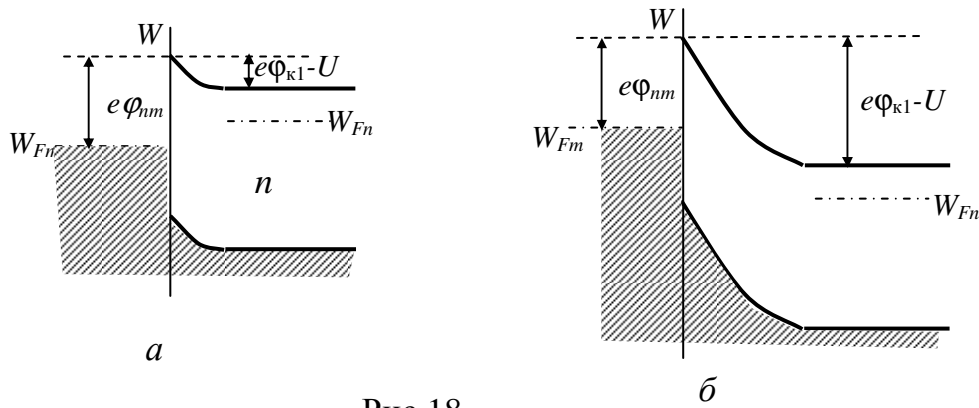


Рис.18

при подаче прямого («+» источника подан на металл, $U > 0$) (а) и обратного ($U < 0$) напряжения (б).

Приложенное внешнее напряжение изменяет потенциальный барьер для электронов, движущихся из полупроводника в металл, а для обратного направления он остается неизменным, поскольку все внешнее напряжение падает в обедненном слое полупроводника. При обратном включении ток обусловлен электронами, имеющими энергию выше потенциального барьера $e\phi_{nm}$.

Теория выпрямляющего контакта базируется на теории термоэлектронной эмиссии электронов. Вольт-амперная характеристика описывается такой же функцией, как ВАХ идеализированного $p-n$ -перехода, причем эта зависимость выполняется с большой точностью в широком интервале токов (до восьми порядков величины тока):

$$I = I_0 (e^{U/\phi_T} - 1) \quad (85)$$

Ток I_0 определяется током термоэлектронной эмиссии из металла в полупроводник:

$$I_0 = S \cdot A_R T^2 \exp(-\phi_{nm} / \phi_T), \quad (86)$$

где A_R – постоянная Ричардсона, для кремния n -типа $A_R = 110 \text{ A}/(\text{см}^2 \cdot \text{K})$, S – площадь контакта.

При равной площади, температуре и степени легирования полупроводника обратный ток контакта Шоттки значительно больше теплового тока кремниевого $p-n$ -перехода. Поэтому напряжение отпираия контакта Шоттки с кремнием ниже, чем кремниевого $p-n$ -перехода.

Омический контакт имеет линейную ВАХ и малое сопротивление. Особенностью этого контакта является также высокая скорость генерации и рекомбинации, поэтому концентрация носителей на границе сохраняется практически неизменной и при протекании тока.

Следует особо отметить, что контакты металл-полупроводник работают на основных носителях заряда, через эти контакты нет инжекции неосновных носителей, нет накопления неравновесного заряда в полупроводнике. Контакт Шоттки обладает только барьерной емкостью, поэтому диоды Шоттки обладают высоким быстродействием.

3. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называется переход, образованный между двумя полупроводниками с различной шириной запрещенной зоны. Если полупроводники имеют одинаковый тип проводимости, то они образуют *изотипный* гетеропереход. Если тип их проводимости различен, получается *анизотипный* гетеропереход. Принципиальным отличием гетеропереходов от *p-n*-переходов является различная высота потенциального барьера для электронов и дырок и наличие разрывов краев зон на металлургической границе, обусловленных различием диэлектрических проницаемостей.

На рис. 19,а приведена зонная диаграмма двух изолированных полупроводников. Они различаются шириной запрещенной зоны ΔW , диэлектрическими проницаемостями ϵ , работой выхода A и сродством к электрону χ . Различие в положении дна зоны проводимости обозначено ΔW_c , а различие в положении потолка валентной зоны ΔW_v .

Зонная диаграмма анизотипного гетероперехода, образованная такими полупроводниками, приведена на рис.19,б для равновесного состояния.

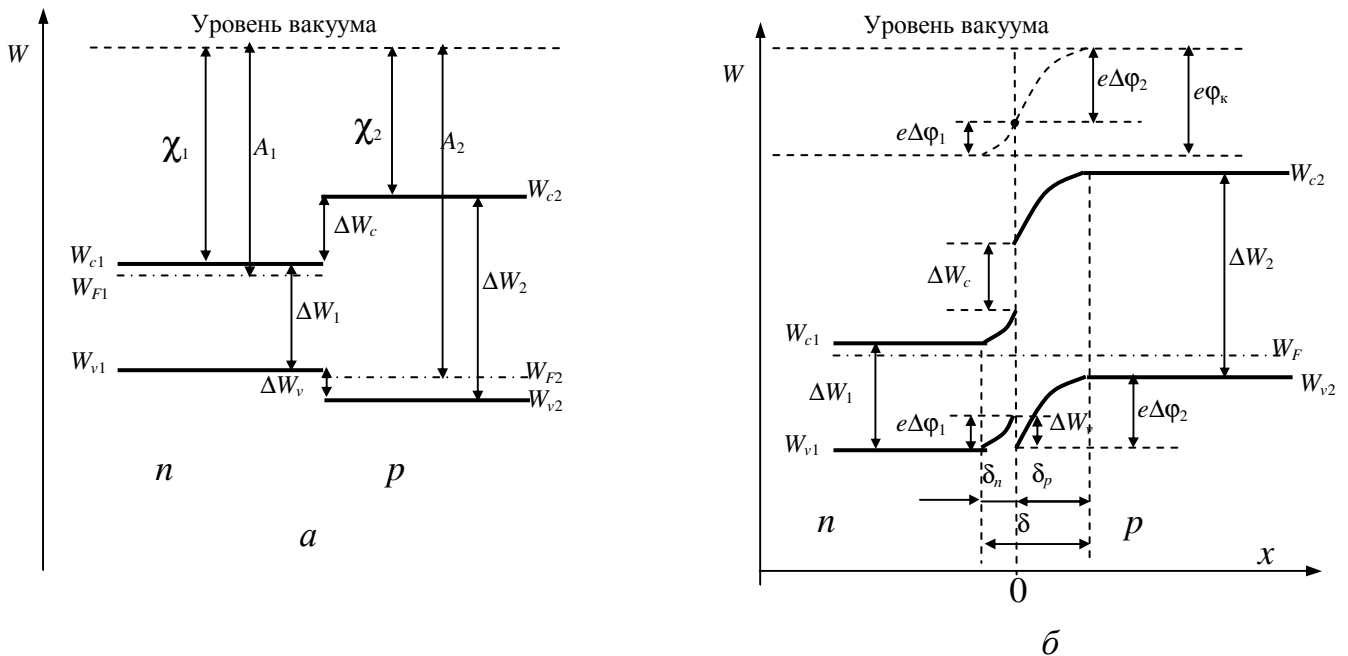


Рис.19

В данном случае полупроводник с широкой запрещенной имеет дырочную проводимость, а с узкой – электронную. Разрывы в положении краев зоны проводимости ΔW_c и краев валентной зоны ΔW_v на металлургической границе не связаны с уровнем легирования. Изгибы краев зон вблизи границы связаны с

образованием обедненных слоев толщиной δ_n и δ_p , содержащих объемные заряды ионов доноров и акцепторов. Значение изгибов ($e\Delta\phi_1$ и $e\Delta\phi_2$) равно внутренней разности потенциалов, образующейся в обедненных слоях. Сумма $\Delta\phi_1 + \Delta\phi_2$ представляет собой контактную разность потенциалов ϕ_k . Как видно из рис.19,б, высота потенциального барьера для электронов, движущихся из n -области в p -область, равна $\phi_k + \Delta W_c / e$ и значительно выше, чем для дырок, движущихся из p -области в n -область, где она равна $\phi_k - \Delta W_v / e$. Поэтому при приложении прямого напряжения будет преобладать инжекция дырок, даже если n -область легирована сильнее. Таким образом, можно получить коэффициент инжекции, близкий к единице (т.е. одностороннюю инжекцию), даже если база (в данном случае n -область) легирована сильнее эмиттера.

Зонная диаграмма p - n -гетероперехода для тех же полупроводников в случае, когда полупроводник с широкой запрещенной имеет электронную проводимость, а с узкой – дырочную, приведена на рис.20,а для равновесного состояния и на рис.20,б – при приложении прямого напряжения. Как видно из рисунка, в этом случае будет происходить односторонняя инжекция электронов в p -область.

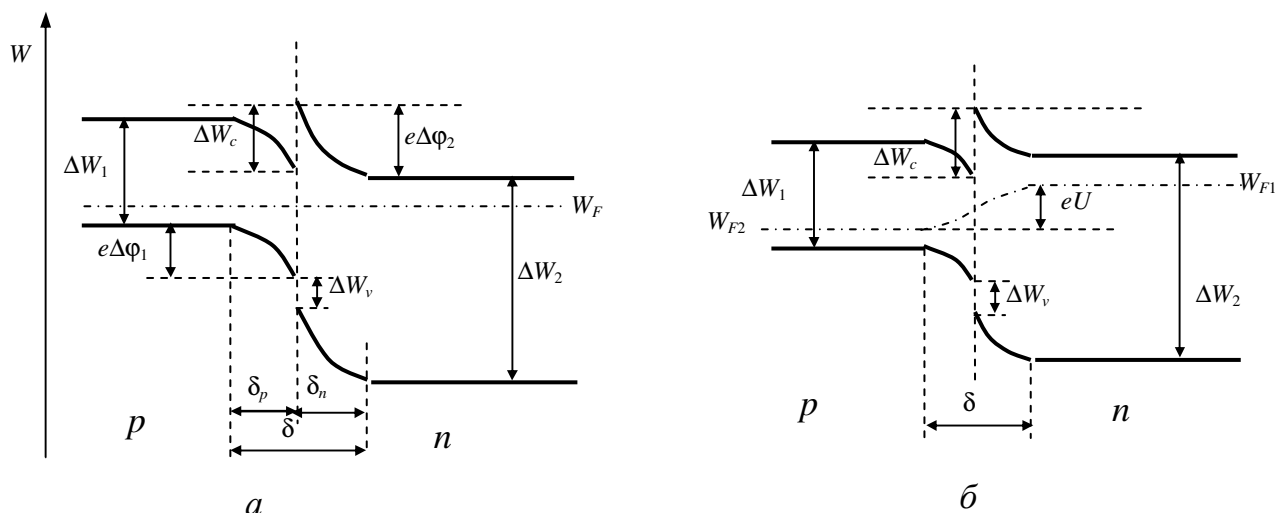


Рис.20

Вольт-амперная характеристика идеализированного p - n -гетероперехода описывается формулой вида (62). На ВАХ реального гетероперехода сильное влияние оказывают центры генерации-рекомбинации в обедненном слое, концентрация которых может быть велика из-за большого числа дефектов структуры вблизи металлургической границы, возникающих вследствие неполной согласованности кристаллических решеток разных полупроводников. Поэтому при прямом напряжении преобладает ток рекомбинации, при обратном – ток генерации.

В изотипных гетеропереходах потенциальный барьер для неосновных носителей значительно выше, чем для основных. Это используется на практике для ограничения области накопления неосновных носителей.

Гетеропереходы широко применяются в излучающих и фотоэлектрических приборах (светодиоды, фотодиоды, полупроводниковые лазеры и др.).

Использование гетеропереходов в биполярных и полевых транзисторах позволяет значительно увеличить их рабочие частоты.

4. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

Термоэлектрическими называются явления, связанные с преобразованием тепловой энергии в электрическую и обратным преобразованием.

Возникновение термо-эдс (эффект Зеебека). Явление термоэлектричества было открыто в 1823 г. Зеебеком и заключается в следующем. Если составить цепь из двух различных металлов или полупроводников и поддерживать температуру T_1 одного места соединения (спая) отличной от температуры T_2 другого (рис.21), то в цепи появится эдс, называемая термоэлектродвижущей силой (термо-эдс) E_T . Это явление называют *эффектом Зеебека*. Подобная цепь называется термоэлектрическим преобразователем (термоэлементом) или термопарой. Термо-

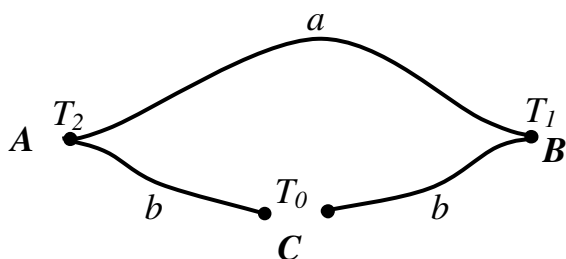


Рис.2.12

эдс зависит от температуры горячего и холодного спаев и от материалов, из которых оставлена цепь, и для однородных проводников, подчиняющихся закону Ома, не зависит от распределения температур между спаями. В замкнутой цепи протекает ток, направленный на выравнивание температуры цепи, при размыкании

одной из ветвей на концах этой ветви возникает разность потенциалов равная E_T , которую в небольшом интервале температур можно считать пропорциональной разности температур с некоторым коэффициентом $\alpha_{1,2}$, называемым коэффициентом термо-эдс:

$$E_T = \alpha_{1,2}(T_2 - T_1). \quad (87)$$

Эффект Зеебека в металлах обусловлен температурной зависимостью внутренней контактной разности потенциалов, возникающей при различной концентрации электронов в контактирующих металлах. При $T_2 - T_1 = 100^\circ\text{C}$ величина E_T в металлах не превосходит нескольких милливольт. Термоэлектрические явления в металлах используются, в основном, для измерения температуры (*термопары, термостолбики*).

Значительно сильнее эффект Зеебека проявляется в полупроводниковых элементах, состоящих из ветвей с дырочной и электронной проводимостью.

Термо-эдс в полупроводниковых структурах состоит из трех составляющих:

- первая вызвана диффузией дырок в p -ветви и электронов в n -ветви от горячего спая к холодному. В результате на холодных концах ветвей образуется избыток основных носителей, на горячих концах остается избыточный заряд ионов легирующей примеси, холодный спай n -ветви заряжается отрицательно, p -ветви – положительно;

– вторая обусловлена температурной зависимостью контактной разности потенциалов. Из выражений (44) и (7) следует, что контактная разность потенциалов уменьшается с температурой и дает вклад в термо-эдс с той же полярностью, что и первая составляющая;

– третья возникает вследствие увлечения носителей заряда фононами. При наличии градиента температуры возникает направленное движение фононов от горячего конца к холодному (это движение обуславливает решеточную теплопроводность). Сталкиваясь с носителями заряда, фононы увлекают за собой в *n*-ветви электроны, в *p*-ветви – дырки.

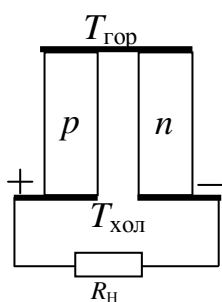


Рис.22

Полупроводниковые термоэлементы используются для непосредственного преобразования тепловой энергии в электрическую. На рис.22 приведена схема полупроводникового термоэлемента с сопротивлением нагрузки. Ветви термоэлемента соединены между собой металлическими пластинами. Термоэлементы могут быть соединены между собой параллельно, последовательно и смешанно в термо-электрические батареи. КПД таких источников тока может достигать до 10%.

Поглощение и выделение тепла в контактах (эффект Пельтье). При пропускании постоянного тока через термоэлемент в спаях его ветвей в единицу времени поглощается или выделяется (в зависимости от направления тока) некоторое количество тепла, пропорциональное времени, току и некоторому коэффициенту $\Pi_{1,2}$, названному коэффициентом Пельтье

$$Q_{\Pi} = \pm \Pi_{1,2} I t. \quad (88)$$

Особенно сильно эффект Пельтье, как и эффект Зеебека, проявляется в полупроводниковых элементах, состоящих из ветвей с дырочной и электронной проводимостью. Если направление тока в цепи термоэлемента такое, как показано на рис.23, электроны и дырки отводятся от спае 1, их убыль восполняется термогенерацией в переходе. На

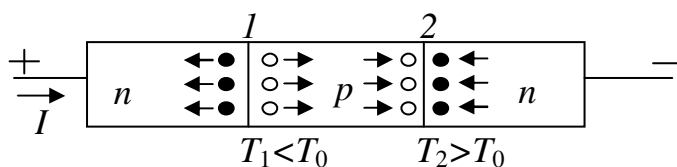


Рис.23

переброс электрона из валентной зоны в зону проводимости затрачивается некоторая тепловая энергия кристаллической решетки, поэтому спай 1 охлаждается. К спаю 2 с разных сторон подводятся электроны и дырки, и энергия, выделяющаяся при их рекомбинации, передается в виде тепла кристаллической решетке, спай 2 будет нагреваться.

В результате при прохождении тока термоэлемент работает как своеобразный тепловой насос, забирая тепловую энергию на спае 1 и выделяя ее на спае 2.

Эффект Пельтье является обратным эффекту Зеебека. Поэтому для одного и того же термоэлемента существует связь между коэффициентами $\alpha_{1,2}$ и $\Pi_{1,2}$:

$$\Pi_{1,2} = \alpha_{1,2} T. \quad (89)$$

Эффект Пельтье используется в полупроводниковых холодильниках для охлаждения небольших объемов. Это явление удобно для использования в термостатах и кондиционерах: простое изменение направления тока позволяет

либо охлаждать, либо нагревать спаи элементов, находящиеся внутри термостата или помещения.

ОПТОЭЛЕКТРОННЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

1. ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

Оптоэлектронными называются явления, возникающие в веществе в результате взаимодействия электронов с электромагнитным излучением оптического диапазона. На этих явлениях основан принцип действия фотоэлектрических приборов, преобразующих энергию оптического излучения в электрическую энергию, и излучающих приборов, осуществляющих обратное преобразование.

Оптический диапазон занимает область длин электромагнитных волн λ от 1 мм до 1 нм. Эти границы довольно условны. Часто область $\lambda=1...0,1$ мм относят к радиоволнам СВЧ диапазона, а область $\lambda=80...1$ нм – к мягкому рентгеновскому излучению. Оптический диапазон разбивается на три под-диапазона: инфракрасный ($\lambda = 0,78...1000$ мкм), видимый ($\lambda = 0,38...0,78$ мкм) и ультрафиолетовый ($\lambda = 0,001...0,78$ мкм).

Световой поток Φ может измеряться в люменах (лм) или определяться мощностью P_Φ в ваттах (Вт), т.е. количеством энергии, протекающим через данную поверхность за 1 с. Единица измерения люмен ориентирована на восприятие света человеческим глазом на каждой длине волны излучения. Обе величины связаны между собой коэффициентом, называемым спектральной видностью: $V_\lambda = \Phi/P_\Phi$ (лм/Вт).

Взаимодействие света с веществом можно характеризовать коэффициентами отражения R , поглощения α и пропускания T . Интенсивность света, проникшего внутрь тела, уменьшается с глубиной x по экспоненциальному закону:

$$\Phi = (1-R) \cdot \Phi_0 \cdot e^{-\alpha x}$$

где Φ_0 – интенсивность падающего света. Зависимость коэффициента поглощения от энергии кванта света $\alpha(h\nu)$ или от длины волны $\alpha(\lambda)$ называется спектром поглощения. Поглощение кванта света (фотона) определяется квантовыми переходами частиц из одного энергетического состояния в другое и приводит к изменению энергетического состояния свободных и связанных электронов или изменению колебательной энергии атомов.

Фоторезистивный эффект. Существует несколько механизмов поглощения света, в полупроводниковых фотоэлектрических приборах используется в основном собственное, или фундаментальное, поглощение.

Собственное (фундаментальное) поглощение связано с разрывом валентной связи, в результате которого образуется электронно-дырочная пара. Этот процесс возможен лишь в случае, когда энергия фотона превышает ширину запрещенной

зоны: $h\nu > \Delta W$, поэтому спектр фундаментального поглощения ограничен минимальной частотой $\nu_{кр} = \Delta W/h$, или максимальной длиной волны $\lambda_{кр} = hc/\Delta W$, которая называется «красной границей». Процесс генерации характеризуется *квантовым выходом* (*квантовой эффективностью*) η – отношением числа генерированных электронно-дырочных пар к общему числу падающих фотонов. При собственном поглощении возрастает концентрация носителей в полупроводнике и, соответственно, уменьшается его сопротивление, возникает *фотопроводимость*. Это, так называемый, *фоторезистивный эффект*. На этом эффекте основан принцип действия *фоторезисторов*.

Фотогальванический эффект. В полупроводниковых структурах, содержащих *p-n*-переход, неравновесные носители, возникшие в *p-n*-переходе и прилегающих к нему областях, толщиной порядка диффузионной длины, разделяются электрическим полем перехода – дырки выбрасываются в *p*-область, электроны – в *n*-область, возникает *фототок* I_ϕ , текущий из *n*-области в *p*-область. В результате *p*-область приобретает положительный заряд, *n*-область – отрицательный, *p-n*-переход смещается в прямом направлении, возникает фото-ЭДС; это явление называется *фотогальваническим эффектом*. Снижение потенциального барьера *p-n*-перехода приводит к возникновению диффузионного тока $I_{дф}$, текущего навстречу фототоку. При разомкнутой внешней цепи эти два тока уравновешивают друг друга. В общем случае внешний ток равен сумме этих токов:

$$I = I_{дф} + I_\phi = I_0 [\exp(U/\varphi_T) - 1] - I_\phi \quad (90)$$

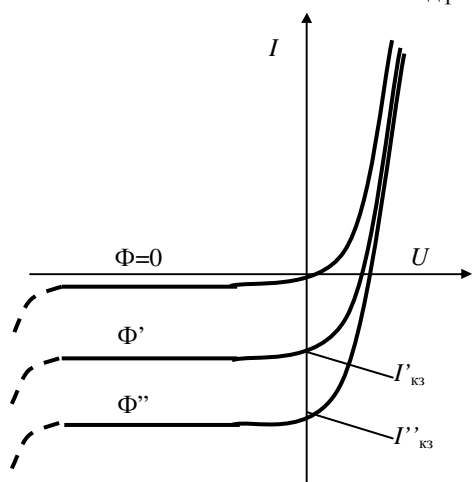


Рис.24

На рис.24 приведена вольт-амперная характеристика *p-n*-перехода при облучении. Обратный ток растет пропорционально световому потоку, ток короткого замыкания $I_{кз}$ (при $U=0$) и напряжение холостого хода $U_{хх}$ (при $I=0$), согласно формуле (90), определяются выражениями:

$$I_{кз} = -(I_0 + I_\phi); \quad U_{хх} = \varphi_T \ln(1 + I_\phi / I_0) \quad (91)$$

Третий квадрант ВАХ *p-n*-перехода используется в фотодиодах. В четвертом квадранте ток принимает отрицательные значения, а напряжение – положительные, т.е. *p-n*-переход является источником энергии, он преобразует энергию излучения в электрическую энергию. В этом режиме работают, в частности, солнечные преобразователи.

Фотогальванический эффект возникает и в других выпрямляющих переходах.

2. ФОТОЭЛЕКТРИЧЕСКИЕ ПРИБОРЫ

Фотоэлектрические приборы разделяются на фотоприемники, преобразующие световой сигнал в электрический, и фотоэлементы, преобразующие

энергию излучения в электрическую энергию (это, главным образом, солнечные преобразователи).

Фоторезисторы

Фоторезистор (рис.25) представляет собой тонкую пластинку или пленку полупроводника 1, нанесенную на диэлектрическую (обычно стеклянную) подложку 3 с двумя омическими контактами 2, к которым подключается электрический источник питания. Падающее на поверхность излучение генерирует в нем свободные носители и с ростом их концентрации увеличивается проводимость фоторезистора:

$$\sigma = \sigma_0 + \Delta\sigma; \quad \sigma_0 = e(\mu_n n_0 + \mu_p p_0); \quad \Delta\sigma = e(\mu_n \Delta n + \mu_p \Delta p),$$

где σ_0 – темновая проводимость, Δn и Δp – приращение концентрации электронов и дырок; μ_n и μ_p – подвижности электронов и дырок.

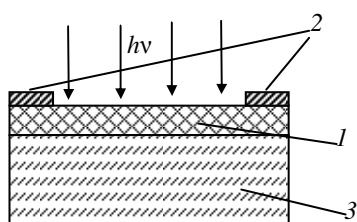


Рис.25

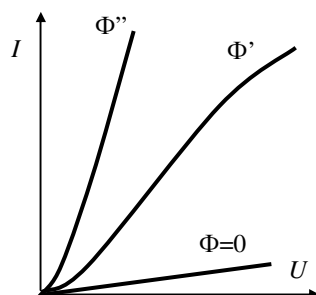


Рис.26

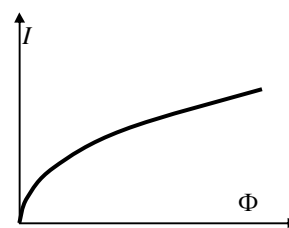


Рис.27

Вольт-амперная характеристика фоторезисторов линейна при заданном световом потоке в ограниченной области напряжений, с наклоном, зависящим от интенсивности светового потока (рис.26).

Передающая характеристика – зависимость тока от светового потока при постоянном напряжении – приведена на рис.27. Под действием света непрерывно идет генерация носителей, их концентрация возрастает до тех пор, пока скорость генерации не сравняется со скоростью рекомбинации:

$$V_{\text{ген}} = a \cdot \Phi, \quad V_{\text{рек}} = b \cdot \Delta n \cdot \Delta p.$$

Отсюда в установившемся режиме $\Delta n = \Delta p = \sqrt{a/b} \cdot \sqrt{\Phi}$ и, соответственно, ток возрастает пропорционально $\sqrt{\Phi}$.

Спектральная характеристика – зависимость чувствительности от частоты

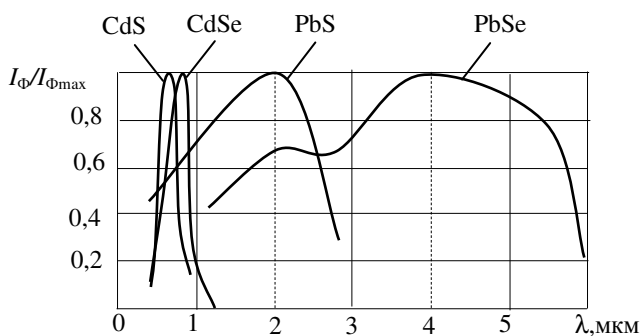


Рис.28

или длины волны излучения – приведена на рис.28 для некоторых фоторезисторов. Чувствительность растет с уменьшением длины волны, начиная с красной границы, затем падает. Это объясняется тем, что с увеличением энергии фотонов резко увеличивается коэффициент поглощения и световой поток поглощается в основном в тонком слое вблизи поверхности, где

велика скорость рекомбинации, и, соответственно, уменьшается избыточная концентрация носителей.

Фоторезисторы характеризуются следующими параметрами:

– темновое сопротивление – $R_{\text{тем}}$;

– монохроматическая токовая чувствительность – $S_{I\lambda} = I_{\text{ф}} / P_{\lambda}$, где P_{λ} – мощность монохроматического излучения с длиной волны λ ;

– интегральная токовая чувствительность по всему спектру излучения $S_{I\Sigma} = I_{\text{ф}} / P_{\text{ф}}$. Поскольку ток зависит от приложенного напряжения, чаще используется более удобный параметр – удельная интегральная чувствительность – $S_{\Sigma\text{уд}} = I_{\text{ф}} / (P_{\text{ф}} \cdot U)$;

– время нарастания $t_{\text{нр}}$ и спада $t_{\text{сп}}$ фототока при воздействии на фоторезистор прямоугольным импульсом излучения. Эти параметры обычно определяются по уровням 0,9 и 0,1 амплитуды импульса фототока;

– граничная частота $f_{\text{гр}}$, соответствующая частоте модуляции светового потока, при котором чувствительность уменьшается до уровня 0,707 от статического значения.

Высокая чувствительность фоторезисторов достигается в результате достаточно длительного накопления неравновесных носителей, поэтому, чем больше время жизни, тем выше чувствительность и тем ниже граничная частота. Ток фоторезистора может во много раз превышать ток генерации из-за излучения.

Наиболее распространены фоторезисторы на основе CdS и CdSe, спектральные характеристики которых располагаются на красном краю видимой области спектра. В инфракрасном диапазоне работают фоторезисторы на основе PbS и PbSe а также изготовленные из Ge, Si, GaAs, InSb и ряд других.

Фоторезисторы на основе CdS и CdSe имеют $R_{\text{тем}} = 10^7 \dots 10^{14}$, $S_{I\Sigma} = 0,5 \dots 20$ мСм/лм и $f_{\text{гр}} \approx 100$ Гц.

Фотодиоды

Фотодиод (ФД) – полупроводниковый диод, обратный ток которого может управляться излучением. В ФД имеется специальное окно, через которое свет попадает на выпрямляющий контакт. Принцип действия ФД основан на фотогальваническом эффекте при обратном смещении выпрямляющего перехода, используется третий квадрант ВАХ. В семейство фотодиодов входят приборы с p - n -переходом, с p - i - n -структурой, с контактом металл-полупроводник и с гетеропереходом.

На рис.29 приведены: *а* – схема включения; *б* – семейство ВАХ при заданных световых потоках; *в* – характеристика передачи ФД. Фототок линейно зависит от светового потока.

Германиевый ФД имеет узкую спектральную характеристику, лежащую в ИК области с максимумом чувствительности при $\lambda = 1,55$ мкм. Спектральная характеристика кремниевых ФД шире, максимумом чувствительности лежит в области $\lambda = 0,6 \div 1$ мкм.

Основные параметры ФД – такие же как и фоторезисторов: монохроматическая токовая чувствительность $S_{I\lambda} = I_{\text{ф}} / P_{\lambda}$, интегральная токовая чувствительность $S_{I\Sigma} = I_{\text{ф}} / P$, время фотоотклика ($t_{\text{нр}}$ и $t_{\text{сп}}$), граничная частота $f_{\text{гр}}$.

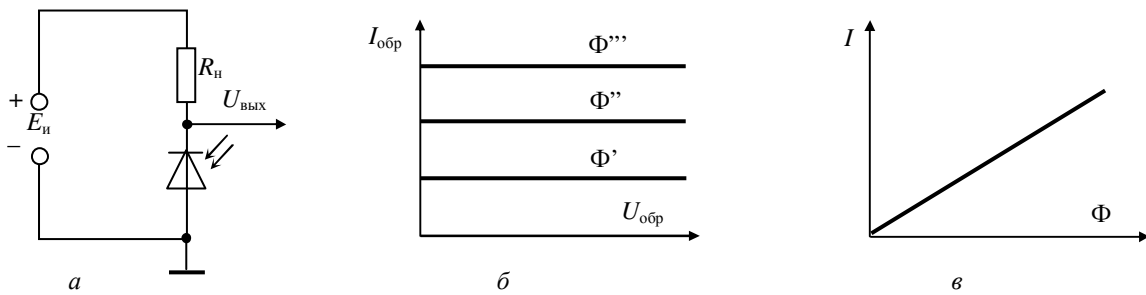


Рис.29

Инерционность фотодиодов определяется временем диффузии к обедненной области $t_{дф}$, временем дрейфа $t_{др}$ через обедненную область и емкостью обедненной области $C_{бар}$. Типичные времена переключения для фотодиодов порядка 10^{-8} с.

Высокое быстродействие достигается в $p-i-n$ -фотодиоде. В этом диоде сильно легированные p^+ - и n^+ -области разделены слоем полупроводника с собственной проводимостью (i -слой). При обратном смещении обедненная область распространяется на весь i -слой и генерация носителей происходит именно в этом слое. Тем самым исключается $t_{дф}$. Толщина i -слоя существенно больше, чем толщина $p-n$ -перехода в обычных ФД, поэтому $C_{бар}$ мало. Граничная частота $p-i-n$ -фотодиодов может достигать 10 ГГц.

В фотодиодах с барьером Шоттки выпрямляющий контакт образуется между полупроводником и тонкой полупрозрачной пленкой металла, пропускающей большую часть падающего излучения. Эти диоды обладают высокой эффективностью в видимой и ультрафиолетовой областях спектра. В этих областях коэффициент поглощения в полупроводниках очень велик (больше 10^5 см^{-1}) и эффективная глубина поглощения мала ($1/\alpha < 0,1 \text{ мкм}$), поэтому падающее излучение в основном поглощается вблизи поверхности полупроводника. Квантовая эффективность может достигать 70%, инерционность этих приборов определяется временем пролета носителей через область объемного заряда ($10^{-10} \div 10^{-11}$ с).

Фотодиоды на гетероструктурах позволяют получить квантовую эффективность, близкую к 100%.

Если к $p-n$ -переходу приложить напряжение, близкое к напряжению лавинного пробоя, можно получить усиление фототока лавинным умножением генерированных светом носителей. Лавинные фотодиоды применяются при обнаружении слабых оптических сигналов.

Солнечные преобразователи

Фотогальванический эффект используется в полупроводниковых фотоэлементах для преобразования энергии излучения в электрическую энергию. Их основное применение – для преобразования солнечной энергии. В преобразователях с $p-n$ -переходом применяются кремний и арсенид галлия, но предпочтение отдано кремнию.

На рис.30 приведены ВАХ солнечных преобразователей с площадью $p-n$ -перехода $S=1\text{см}^2$ при плотности мощности $P_{\text{пл}}=0,2\text{ Вт/см}^2$ и $T=20^\circ\text{C}$. Они

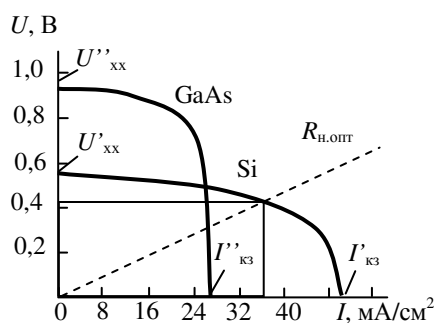


Рис.30

соответствуют четвертому квадранту ВАХ освещенного $p-n$ -перехода. Преобразователь на основе GaAs имеет большее значение фото-ЭДС (вследствие большей ширины запрещенной зоны) но меньший ток короткого замыкания. Диффузионная длина неосновных носителей в кремнии больше, чем в арсениде галлия, поэтому выше квантовый выход $\eta_{\text{ф}}$ и ток $I_{\text{кз}}$. Предельный расчетный КПД кремниевого преобразователя 23% получен без учета потерь при $\eta_{\text{ф}} = 1$. Реальные КПД кремниевых преобразователей 14-

18%, арсенид-галлиевых – 11%.

При подключении нагрузки ток $I_{\text{н}}$ и напряжение $U_{\text{н}}$ во внешней цепи определяются точкой пересечения ВАХ и линии нагрузки (рис.30). Площадь прямоугольника $I_{\text{н}}U_{\text{н}}$ пропорциональна мощности, отдаваемой преобразователем в нагрузку; при оптимальном сопротивлении нагрузки достигается максимальная мощность.

На рис.31 показана структура кремниевого преобразователя: 1 – фронтальный металлический электрод на освещаемой поверхности; 2 – тыльный металлический электрод, полностью покрывающий поверхность кристалла с неосвещаемой стороны; 3 – просветляющее покрытие из монооксида кремния SiO . Чтобы не затенялась освещаемая поверхность, ширина фронтального электрода не должна превышать 1мм.

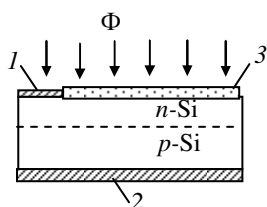


Рис.31

В эквивалентной схеме (рис.32) генератор тока $I_{\text{ф}}$ моделирует фототок, диод VD – диффузионный ток перехода, резистор $R_{\text{ут}}$ имитирует сопротивление

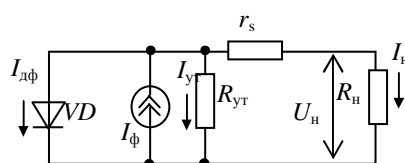


Рис.32

утечки перехода, резистор r_s – потери в объеме p - и n -областей и контактах металл-полупроводник. Омические потери в резисторах $R_{\text{ут}}$ и r_s уменьшают выходную мощность и КПД преобразователя и предпринимаются

специальные меры для уменьшения их влияния.

Отдельные элементы соединяются в солнечные батареи. Кремниевые преобразователи позволяют получать мощность до 100 Вт с квадратного метра площади при удельной мощности 50 Вт/кг.

3. СВЕТОДИОДЫ

Светодиод – излучающий полупроводниковый прибор с одним или несколькими *p-n*-переходами, преобразующий электрическую энергию в энергию некогерентного оптического излучения.

Светодиоды применяются в оптических линиях связи, в индикаторных устройствах, в оптронах и т.д.

Принцип действия светодиодов основан на самопроизвольной излучательной рекомбинации носителей заряда, инжектируемых через *p-n*-переход. Рекомбинация сосредоточена в *p-n*-переходе и в прилегающих к нему слоях полупроводника. Основную роль играет межзонная рекомбинация: электрон переходит из зоны проводимости в валентную зону, соединяясь с дыркой, энергия фотона определяется шириной запрещенной зоны. Внутренний квантовый выход $\eta_{\text{ф}}$ определяется как отношение числа излученных фотонов к числу рекомбинировавших пар.

Вероятность излучательной рекомбинации зависит от особенности структуры энергетических зон полупроводника. В процессах излучения и поглощения света должны выполняться законы сохранения энергии и импульса (квазиимпульса). Импульсом фотона обычно можно пренебречь, поскольку его величина мала для видимого и инфракрасного света.

В прямозонных полупроводниках, у которых дно зоны проводимости и потолок валентной зоны располагаются при значении квазиимпульса $\vec{P} = 0$, разрешены прямые оптические переходы, при которых квазиимпульс электрона сохраняется. Такие переходы не требуют участия третьей частицы (кроме электрона и дырки), и вследствие этого вероятность таких переходов высока. Каждая рекомбинация при прямом переходе сопровождается излучением фотона. Непрямые переходы происходят с участием фонона и вероятность излучения фотона мала.

Для излучения видимого света ширина запрещенной зоны должна быть в пределах $1,8 \text{ эВ} \leq \Delta W = h\nu \leq 3,2 \text{ эВ}$. Это требование сразу исключает использование германия и кремния в индикаторных светодиодах. Кроме того, они являются непрямозонными полупроводниками.

В качестве основных полупроводниковых материалов для светодиодов применяются арсенид галлия GaAs, фосфид галлия GaP, нитрид галлия GaN, карбид кремния SiC, трехкомпонентный твердый раствор фосфида и арсенида галлия $\text{GaAs}_{1-x}\text{P}_x$ ($0 \leq x \leq 1$) и ряд других двойных и многокомпонентных полупроводниковых соединений.

К прямозонным полупроводникам относятся GaAs и $\text{GaAs}_{1-x}\text{P}_x$ при $0 \leq x \leq 0,45$. При возрастании x от 0 до 0,45 ширина запрещенной зоны соединения увеличивается от 1,42 до 1,98 эВ и преобладают прямые переходы. Дальнейшее увеличение содержания фосфора приводит к непрямым переходам и уменьшению квантового выхода. Чистый фосфид галлия относится к непрямозонным полупроводникам ($\Delta W = 2,26 \text{ эВ}$) и квантовый выход незначителен.

Для увеличения квантового выхода в непрямозонных полупроводниках внедрением примесей создаются рекомбинационные центры-ловушки вблизи дна зоны проводимости (донорные) и/или вблизи потолка валентной зоны (акцепторные). Инжектированные электроны попадают сначала на уровни

донорных ловушек, которые затем захватывают дырку, излучая фотоны. Закон сохранения импульса выполняется при передаче части импульса ловушке. Излучательные переходы между донорными и акцепторными ловушками позволяют получить генерацию света на различных длинах волн.

Качество излучающего диода характеризуется внешним квантовым выходом. Даже при большом значении внутреннего квантового выхода внешний квантовый выход может оказаться малым вследствие низкого вывода излучения во внешнюю среду. Часть излучения теряется из-за самопоглощения, энергия фотонов при этом может быть передана электронам валентной зоны с переводом их в зону проводимости или свободным носителям в пределах одной зоны, может происходить поглощение фотонов кристаллической решеткой и некоторые другие виды поглощения.

Из-за высокого коэффициента преломления исходного материала большая часть излучения, падающее на плоскую границу раздела полупроводник-воздух, претерпевает полное внутреннее отражение и, претерпев многократные отражения внутри кристалла, в конечном счете, теряется из-за самопоглощения. Применение сферической границы раздела позволяет уменьшить эти потери. На рис.33 приведено устройство такого светодиода, у которого база 3 выполнена в виде полусферического монокристалла полупроводника, 4 – эмиттер, 1 и 2 – контакты. Для увеличения КПД светодиодов применяют также прозрачные полусферические покрытия из стекла и пластмасс с высоким показателем преломления, просветляющие покрытия внешней поверхности и т.д.

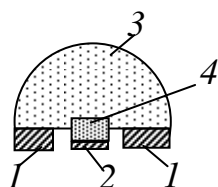


Рис.33

Основные характеристики светодиодов:

- яркостная – зависимость яркости от тока $B(I)$;
- спектральная – зависимость яркости, или силы света, или мощности от длины волны $B(\lambda)$, $\Phi(\lambda)$, $P_{\phi}(\lambda)$;
- вольт-амперная $I(U)$.

Примеры вольт-амперных характеристик светодиодов приведены на рис.34. Различия прямых ветвей ВАХ связаны с различной шириной запрещенной зоны полупроводниковых материалов.

Начальный участок яркостной характеристики нелинеен (рис.35). При малых напряжениях и соответственно при малых токах на светодиоде излучение не возникает. Рабочее напряжение должно превышать пороговое значение, которое определяется энергией излучаемого фотона: $U_{\text{пор}}=h\nu/e$. Яркость излучения светодиодов достигает значений в несколько сот кд/м².

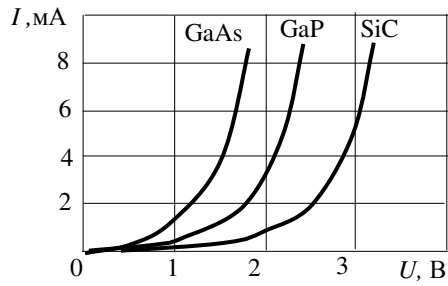


Рис.34

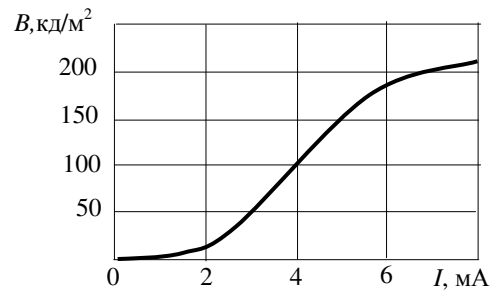


Рис.35

Примеры спектральных характеристик светодиодов на основе различных материалов приведены на рис.36. Переход электронов при рекомбинации носителей происходит не между двумя определенными уровнями, а между двумя

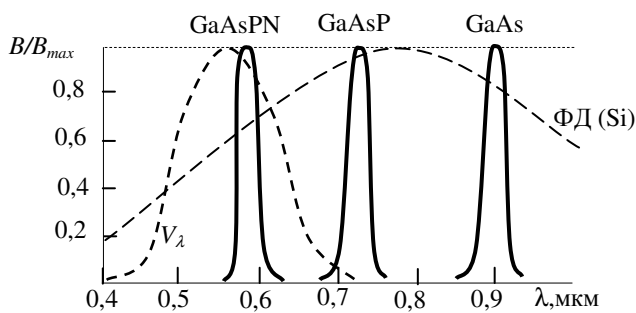


Рис.36

группами близко расположенных уровней, поэтому спектр излучения оказывается размытым. Спектральный диапазон светодиода характеризуется шириной спектральной характеристики, определяемой на высоте 0,5 от максимума. На рис.36 штриховыми линиями показаны также спектральные характе-

ристики чувствительности человеческого глаза V_λ и кремниевого фотодиода.

Спектр излучения арсенид-галлиевого светодиода лежит в инфракрасной области, поэтому он не может применяться в индикаторных устройствах. Однако он широко применяется в оптических линиях связи, в оптронах и т.д.

Инерционность светодиодов характеризуется временами нарастания t_{nr} и спада t_{sp} импульса излучения, которые измеряются при подаче прямоугольного импульса тока по уровням 0,9 и 0,1 от максимальной яркости. Эти параметры определяются временем перезаряда емкости диода и составляют доли микросекунд.

Кпд светодиода – отношение мощности излучения к потребляемой мощности – обычно не превышает нескольких процентов.

ВАКУУМНАЯ ЭЛЕКТРОНИКА

1. ОБЩИЕ СВЕДЕНИЯ

Принцип действия электровакуумных приборов (ЭВП) основан на явлениях, связанных с движением заряженных частиц в вакууме или газе. В зависимости от характера рабочей среды они разделяются на электронные и ионные (газоразрядные) ЭВП.

В электронных ЭВП рабочей средой является высокий вакуум, единственными носителями заряда являются свободные электроны, поэтому в межэлектродном пространстве имеется объемный заряд. В процессе движения нет столкновений, электроны не теряют энергию, скорость электрона полностью определяется электрическим потенциалом в данной точке, если начальная скорость равна нулю. Движение происходит исключительно под действием электрического и магнитного полей или по инерции, диффузия отсутствует. Скорость электронов, как правило, значительно превышает скорость движения носителей в полупроводниках, где она обычно не превышает $10^7 \text{ см} \cdot \text{с}^{-1} = 100 \text{ км} \cdot \text{с}^{-1}$:

$$v = \sqrt{\frac{2eU}{m}} = 600\sqrt{U} \text{ (км} \cdot \text{с}^{-1} \cdot \text{В}^{-1/2}\text{)}, \quad (92)$$

где e и m – заряд и масса электрона; U – потенциал в данной точке.

В ионных ЭВП в межэлектродном пространстве образуется квазинейтральная плазма – смесь электронов, ионов и нейтральных молекул, обладающая высокой электропроводностью, – падение напряжения в объеме малое, основное падение происходит вблизи поверхности электродов на расстоянии порядка длины свободного пробега, где сосредоточен объемный заряд. В объеме плазмы сохраняется электронейтральность – заряд электронов скомпенсирован зарядом ионов. При этом практически весь ток переносится электронами, поскольку масса ионов в десятки тысяч раз больше массы электронов.

В ЭВП рабочая среда отделена от внешней среды герметичным баллоном, внутри которого размещены электроды, предназначенные для управления физическими процессами. В любом ЭВП имеется по крайней мере два электрода: катод (источник электронов) и анод, на который подается положительное напряжение относительно катода и который собирает электроны.

Электронная эмиссия

При комнатной температуре электрон не может выйти за пределы твердого тела, так как на границе раздела существует потенциальный барьер. Высота этого барьера определяется термоэлектронной работой выхода электрона A – разностью между энергией электрона в вакууме W_0 и энергией уровня Ферми в твердом теле W_F :

$$A = W_0 - W_F$$

Для преодоления потенциального барьера электрон должен получить дополнительную энергию $\Delta W \geq A$. Испускание электронов твердым телом называется *эмиссией*, она осуществляется путем подведения энергии к твердому

телу от внешнего источника. В зависимости от вида подводимой энергии различают несколько видов эмиссии:

- *термоэлектронная эмиссия* осуществляется при нагреве вещества до достаточно высокой температуры. Плотность тока термоэлектронной эмиссии j_e от температуры T определяется формулой Ричардсона–Дэшмана:

$$j_e = A_R T^2 \exp[-A/(kT)],$$

где $A_R=10\div 300$ А/(см²К²) – постоянная Ричардсона, зависящая от материала катода, k – постоянная Больцмана;

- *фотоэлектронная эмиссия* возникает при облучении вещества световым потоком с энергией квантов $h\nu$, превышающих работу выхода электронов. Кинетическая энергия эмитированных электронов определяется выражением:

$$mv^2/2 = h\nu - A;$$

- *автоэлектронная (электростатическая) эмиссия* возникает при наличии сильного электрического поля у поверхности твердого тела. При этом резко сужается ширина энергетического барьера и становится возможным туннельный переход электронов из катода во внешнее пространство. Для автоэлектронной эмиссии необходимы напряженности электрических полей 10^6 В/см и более;

- *вторичная электронная эмиссия* происходит при бомбардировке поверхности тела потоком электронов с большой кинетической энергией. Этот процесс характеризуется коэффициентом вторичной эмиссии σ – отношением числа электронов n_2 , выбитых с поверхности, к числу первичных электронов n_1 , бомбардирующих поверхность: $\sigma=n_2/n_1$;

- *ионно-электронная эмиссия* возникает при бомбардировке поверхности тела потоком ускоренных ионов. Этот вид эмиссии используется в газоразрядных приборах.

В ЭВП наиболее часто используется термоэлектронная эмиссия.

Термоэлектронные катоды характеризуются следующими параметрами:

- рабочая температура $T_{\text{раб}}$;
- предельный катодный ток $I_e=j_e S$, где j_e – плотность тока, S – площадь катода;
- эффективность – отношение тока эмиссии к мощности накала: $H=I_e/P_{\text{нак}}$
- долговечность (срок службы).

Термоэлектронные катоды различаются:

– по способу накала:

- прямонакальные, нагреваемые пропусканием тока через катод, изготовленный в виде нити;
- косвенного накала, или подогревные, в которых катод изолирован от нагревателя;

– по материалу и структуре:

- металлические (применяется исключительно вольфрам);
- активированные металлические (пленочные), у которых на металлическую основу (W, Mo) нанесены атомы металла с меньшей работой выхода (Ba, Sr, Th);
- полупроводниковые – оксидный катод – смесь оксидов щелочноземельных металлов (BaO, SrO, CaO) на металлической подложке.

Оксидные катоды нашли наиболее широкое применение в ЭВП. Они имеют наиболее низкую рабочую температуру ($\leq 1000^\circ\text{C}$), $j_e=20-90 \text{ мА/см}^2$, $H=80-200 \text{ мА/Вт}$.

2. ЭЛЕКТРОННЫЕ ЛАМПЫ

Электронные лампы – это ЭВП с термоэлектронным катодом и электростатическим управлением электронным потоком, которые служат для усиления, детектирования, генерации и преобразования электрических сигналов.

Диод

Вакуумный диод – простейшая лампа, имеющая только анод и катод. Электроды могут быть плоскими или (чаще) цилиндрической формы. Диод обладает односторонней проводимостью и используется для выпрямления тока. Условное изображение диода приведено на рис.37 для ламп с косвенным (а) и прямым накалом (б).

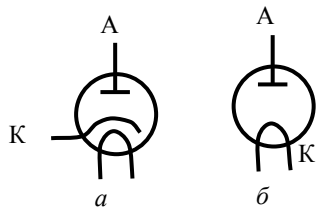


Рис.37

При включении накала около катода образуется облако эмитированных электронов, создающих отрицательный объемный заряд, потенциал достигает минимума на некотором расстоянии x' от поверхности катода. При подаче положительного напряжения на анод электроны устремляются к аноду и протекает анодный ток. Если анодное напряжение достаточно велико, минимум потенциала исчезает, все эмитированные электроны достигают анода и ток достигает максимального значения, равного току эмиссии. Этот режим называется режимом *насыщения*. Если анодное напряжение недостаточно для насыщения, только часть наиболее быстрых электронов преодолевает потенциальный барьер, созданный объемным зарядом, остальные отбрасываются обратно к катоду, анодный ток меньше тока эмиссии. Этот режим называется режимом *объемного заряда*.

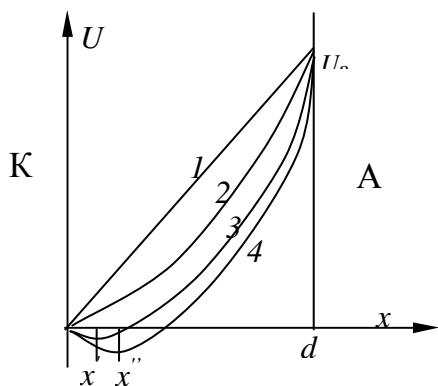


Рис.38. Распределение потенциала в плоском диоде при различных напряжениях накала: $U_{n1}=0, U_{n4}>U_{n3}>U_{n2}$

На рис.38 показано распределение потенциала в плоском диоде при различных напряжениях накала и постоянном анодном напряжении. В отсутствие эмиссии ($U_n=0$) потенциал линейно растет от катода к аноду (прямая 1), при протекании тока зависимость потенциала от координаты нелинейная вследствие объемного заряда, создаваемого электронами. Кривая 2 соответствует режиму насыщения, кривые 3 и 4 – режиму объемного заряда. Чем выше ток эмиссии, тем минимум потенциала глубже и дальше отстоит от поверхности катода.

Теоретическую вольт-амперную характеристику диода для режима объемного заряда можно вывести, используя уравнение Пуассона. Для диода с плоскопараллельными электродами (плоскостной диод) это уравнение имеет вид:

$$d^2U/dx^2 = -en/\epsilon_0, \quad (93)$$

где x – расстояние от катода, n – плотность электронов, ϵ_0 – диэлектрическая постоянная. Полагая, что скорость электрона у поверхности катода равна нулю, запишем уравнения для плотности тока j и скорости электронов v :

$$j = -env, \quad (94)$$

$$v^2 = 2eU/m, \quad (95)$$

где e и m – заряд и масса электрона. Предполагается также, что минимум потенциала достигается на поверхности катода. Граничными условиями задачи, таким образом, являются:

$$U=0 \text{ и } dU/dx=0 \text{ при } x=0; \quad U=U_a \text{ при } x=d,$$

где d – межэлектродное расстояние.

Решение системы уравнений (93)–(95) имеет вид

$$I = G_d U_a^{3/2} \quad (96)$$

где коэффициент G_d называется первеансом диода. Для плоскостного диода

$$G_d = \frac{4\epsilon_0}{9d^2} \sqrt{\frac{2e}{m}} = 2,33 \cdot 10^{-6} \frac{S}{d^2}, \text{ А} \cdot \text{В}^{-3/2}, \quad (97)$$

где S – площадь сечения диода.

Выражение (96) известно как “Закон степени три вторых”, оно применимо лишь для режима объемного заряда. При достижении насыщения анодный ток перестает зависеть от анодного напряжения. При отрицательных анодных напряжениях анодный ток практически равен нулю.

На рис.39 штриховой линией показана теоретическая ВАХ вакуумного диода и сплошной линией – реальные ВАХ при двух значениях напряжения накала. Реальные ВАХ идут более полого и плавно переходят в режим насыщения. В приборах с оксидными катодами, которые имеют большое удельное сопротивление, режим насыщения не наблюдается, что объясняется уменьшением работы выхода под действием внешнего электрического поля и дополнительным разогревом катода протекающим током.

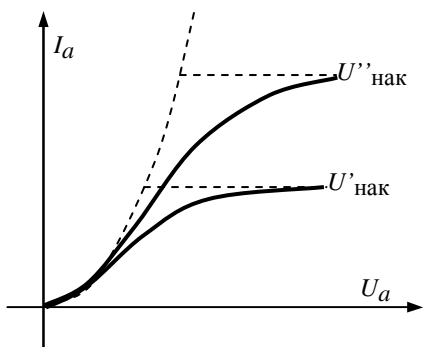


Рис.39

Применение вакуумных диодов основано на использовании униполярной проводимости – для выпрямления, детектирования и других нелинейных преобразований. В большинстве случаев они вытеснены полупроводниковыми диодами и сохраняют значение как высоковольтные выпрямители (кенотроны) и как СВЧ диоды.

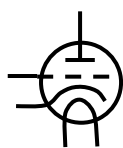


Рис.40

Триод

В этой лампе между анодом и катодом имеется еще один электрод – сетка, предназначенная для управления анодным током.

Сетка выполняется обычно в виде проволочной спирали, реже в виде проволочной сетки. Электронный поток может проникать сквозь сетку к аноду. Анодный ток I_a является функцией двух переменных – анодного U_a и сеточного U_c напряжений, причем потенциал сетки сильнее влияет на анодный ток, так как сетка расположена гораздо ближе к катоду. Условное изображение триода приведено на рис.40.

На сетку относительно катода может подаваться как положительный, так и отрицательный потенциал. Сетка может приводить к сильному изменению электрического поля (потенциала) и соответственно объемного заряда в межэлектродном пространстве, особенно в области между сеткой и катодом. При больших отрицательных напряжениях на сетке даже при положительном потенциале анода во всем пространстве между катодом и сеткой существует тормозящее для электронов электрическое поле, все электроны возвращаются к катоду, анодный ток равен нулю. Наименьшее по модулю значение отрицательного потенциала сетки, при котором $I_a = 0$, называется *напряжением запираения* U_{c0} для заданного напряжения анода. Если отрицательное напряжение на сетке становится по модулю меньше напряжением запираения U_{c0} , то между витками сетки поле становится ускоряющим, часть электронов, эмитированных катодом, устремляются к аноду и возникает анодный ток. Чем менее отрицательным становится потенциал сетки, тем сильнее возрастает анодный ток. При положительных потенциалах на сетке часть электронов, движущихся вблизи витков, притягиваются сеткой, и возникает сеточный ток I_c , а основная доля электронного потока попадает на анод, образуя анодный ток I_a .

Электрическое поле в области сетки является неоднородным. По мере удаления от витков неоднородность поля быстро ослабевает и вблизи катода и анода поле становится практически однородным (при достаточно густой сетке). Поэтому можно полагать, что оно создается не совместным действием анода и сетки, а действием одного сплошного электрода, расположенного в плоскости сетки. Таким образом, для расчета катодного тока триод можно заменить эквивалентным диодом, анод которого располагается на месте сетки (рис.41), здесь же показаны емкости между электродами.

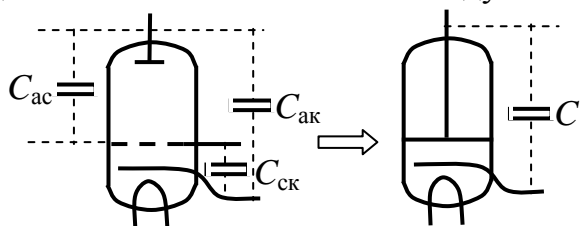


Рис.41

Анодное напряжение эквивалентного диода, при котором катодные токи ламп равны, называется *действующим напряжением* U_d . Равенство катодных токов означает, что напряженности электрического поля у поверхности катода у триода и эквивалентного диода равны, что в свою очередь означает, что

равны и заряды, наведенные на катоде напряжениями, приложенными к электродам. Эти заряды равны, соответственно:

$$Q_{тр} = C_{ск} U_c + C_{ак} U_a,$$

$$Q_d = C U_d,$$

где $C_{ск}$, $C_{ак}$ – емкости между сеткой и катодом и между анодом и катодом триода, C – емкость между электродами диода. Отсюда

$$C_{ск} U_c + C_{ак} U_a = C U_d$$

Пренебрегая емкостью C_{ac} , можно положить $C \approx C_{ck} + C_{ак}$ и, обозначив $C_{ca}/C_{ck} = D$, выразить действующее напряжение как функцию анодного и сеточного напряжений триода:

$$U_d = \frac{U_c + DU_a}{1 + D} \quad (98)$$

Параметр D называется *проницаемостью* лампы. Он характеризует степень проникновения поля анода в пространство сетка-катод и показывает, во сколько раз анодное напряжение слабее влияет на потенциальный барьер у катода по сравнению с напряжением сетки. Обычно $D \ll 1$ и можно положить

$$U_d \approx U_c + DU_a \quad (99)$$

Приведение триода к эквивалентному диоду позволяет применить закон степени три вторых к триоду:

$$I_k = G_T U_d^{3/2} = G_T (U_c + DU_a)^{3/2} \quad (100)$$

где G_T – первеанс триода. Из выражения (100) находится напряжение запирания лампы, при котором $I_k = 0$:

$$U_{c0} = -DU_a \quad (101)$$

Статические характеристики триода. При постоянном напряжении накала анодный I_a и сеточный I_c токи триода являются функциями двух переменных – анодного и сеточного напряжений:

$$I_a(U_a, U_c), \quad I_c(U_a, U_c).$$

Практически эти зависимости удобно представить в виде функции одной переменной, считая вторую независимую переменную постоянным параметром режима. Соответственно, имеется четыре семейства характеристик:

$I_a(U_a), U_c = \text{const}$ – анодная (выходная) характеристика,

$I_a(U_c), U_a = \text{const}$ – анодно-сеточная (передаточная) характеристика,

$I_c(U_c), U_a = \text{const}$ – сеточная (входная) характеристика,

$I_c(U_a), U_c = \text{const}$ – сеточно-анодная характеристика.

На рис.42,а приведены анодно-сеточные (сплошные линии) и сеточные (штриховые линии) характеристики, а на рис.42,б – анодные (сплошные линии) и сеточно-анодные (штриховые линии) характеристики, бóльший индекс соответствует большему (по модулю) значению параметра.

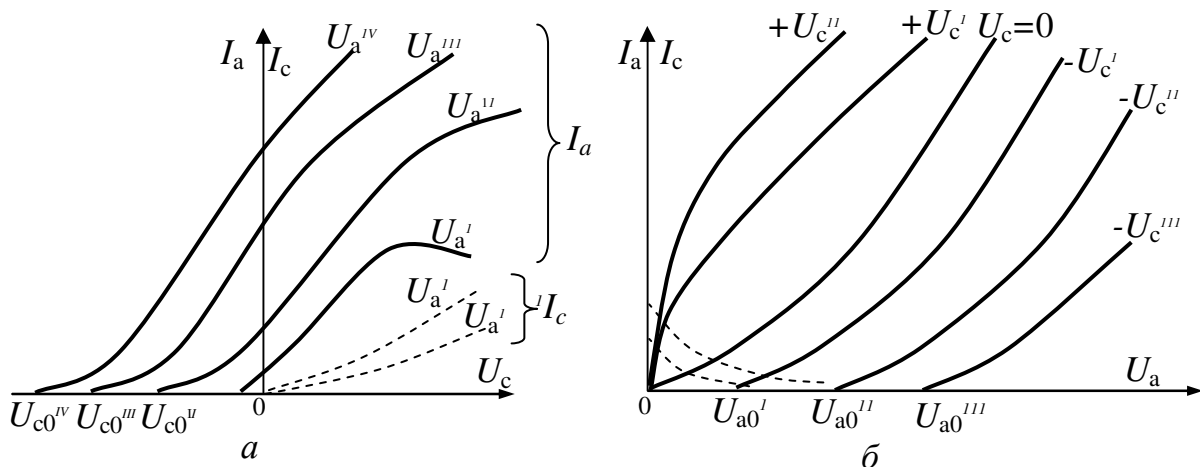


Рис.42

При отрицательных напряжениях сетки сеточный ток отсутствует, все электроны, преодолевшие минимум потенциала у катода, устремляются на анод. Анодный ток равен катодному и может быть вычислен по формуле (100):

$$I_a = I_k = G_T U_d^{3/2} = G_T (U_c + D U_a)^{3/2} \quad (102)$$

При $U_c = U_{c0}$, когда $I_a = 0$, лампа заперта. Увеличение анодного напряжения вызывает смещение начала анодно-сеточных характеристик влево в соответствии с соотношением (101). Анодные характеристики при $U_c < 0$ начинаются с некоторого положительного потенциала $U_{a0} = -U_c/D$.

При положительных напряжениях сетки катодный ток распределяется между сеткой и анодом:

$$I_k = I_a + I_c$$

При этом существуют два режима токораспределения.

При напряжениях сетки, меньших анодного, электроны, пролетевшие между витками сетки, попадают в ускоряющее поле и поэтому все попадают на анод. Ток сетки образуется электронами, которые перехватываются витками сетки на пути от катода к аноду. Этот режим работы называется *режимом перехвата электронов* (рис.43,а). В режиме перехвата отношение I_a/I_c зависит от соотношения площадей поверхности сетки и анода, и обычно ток анода превышает ток сетки. С увеличением U_c одновременно увеличиваются оба тока, что обусловлено в основном ростом катодного тока из-за увеличения действующего напряжения.

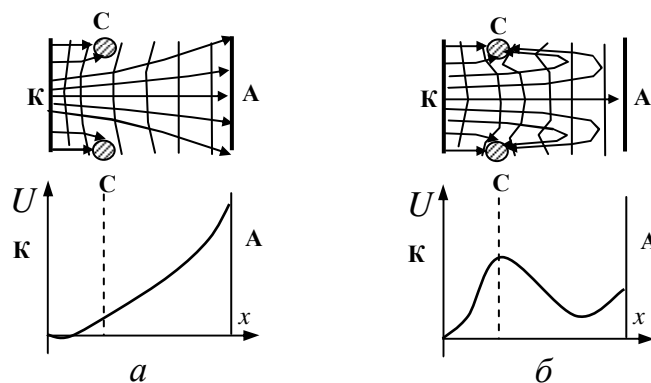


Рис.43

Когда напряжение на сетке становится больше анодного, наступает *режим возврата электронов к сетке* (рис.43,б). В этом режиме часть электронов перехватывается сеткой, другая часть, пролетев между витками сетки по искривленной траектории, попадает в тормозящее поле между сеткой и катодом, поскольку $U_a < U_c$. Это тормозящее поле преодолевают только электроны, имеющие достаточно большую продольную составляющую скорости, которые, попав на анод, образуют ток I_a . Остальные электроны, получившие сильное отклонение в неоднородном поле сетки, возвращаются к сетке. В режиме возврата при увеличении U_c ток I_a уменьшается, а ток I_c возрастает. Возвращающиеся электроны создают большой объемный заряд между сеткой и анодом, что приводит к появлению дополнительного потенциального барьера для электронов,двигающихся к аноду, следовательно, возрастает возврат электронов. С ростом

анодного тока минимум потенциала может снизиться до нуля. Поверхность нулевого потенциала называют виртуальным катодом.

При положительных напряжениях сетки анодные характеристики (см. рис.42,б) начинаются от нуля анодного напряжения и имеют крутой начальный участок, соответствующий режиму возврата электронов, и более пологий, соответствующий режиму перехвата. Сеточный ток уменьшается с ростом анодного напряжения.

Режим возврата для практического использования лампы неблагоприятен, так как при этом резко возрастает сеточный ток и работа лампы становится неэффективной.

Дифференциальные параметры триода. Дифференциальные параметры – это величины, связывающие малые приращения токов и напряжений, для которых эта связь является линейной. Для приращения анодного тока можно записать:

$$\Delta I_a = S \Delta U_c + G_i \Delta U_a, \quad (103)$$

где коэффициенты S и G_i являются частными производными функции $I_a(U_c, U_a)$:

$$S = \left(\frac{\partial I_a}{\partial U_c} \right)_{U_a} \quad \text{– крутизна анодно-сеточной характеристики;}$$

$$G_i = \left(\frac{\partial I_a}{\partial U_a} \right)_{U_c} \quad \text{– выходная дифференциальная проводимость.}$$

Кроме этих параметров используются также:

$$R_i = G_i^{-1} = \left(\frac{\partial U_a}{\partial I_a} \right)_{U_c} \quad \text{– выходное дифференциальное сопротивление (или}$$

внутреннее сопротивление лампы);

$$\mu = - \left(\frac{\partial U_a}{\partial U_c} \right)_{I_a} \quad \text{– статический коэффициент усиления.}$$

Чтобы анодный ток оставался постоянным, необходимо чтобы напряжения U_a и U_c изменялись в противоположных направлениях, μ считается положительной величиной, поэтому в ее определении стоит знак минус. Полагая в выражении (103) $\Delta I_a = 0$, находим связь между параметрами:

$$\mu = S/G_i = SR_i \quad (104).$$

При отрицательных напряжениях сетки $\mu = 1/D$.

Обычно крутизна триодов составляет единицы мА/В, у некоторых типов ламп до 20...50 мА/В. Внутреннее сопротивление R_i имеет значения единицы–десятки кОм.

Тетроды

Недостатками триода являются относительно малый коэффициент усиления μ и большая проходная емкость C_{ac} между анодом и сеткой, которая создает нежелательную обратную связь и ограничивает частотный диапазон.

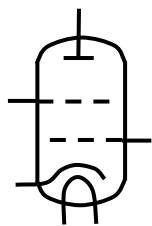


Рис.44

В тетроде (рис.44) для уменьшения этой емкости между анодом и управляющей сеткой устанавливают еще одну сетку, которая называется второй или экранирующей сеткой. Эта сетка выполняет роль электростатического экрана, поэтому емкость C_{ac} резко снижается. Для эффективного экранирования вторая сетка должна быть достаточно густой, но в то же время достаточно прозрачной для электронов, движущихся к аноду. На экранирующую сетку подается положительное напряжение U_{c2} , соизмеримое с напряжением

источника питания анода.

Тетрод имеет существенный недостаток, связанный с наличием динаatronного эффекта, возникающего при $U_a < U_{c2}$. Динаatronным эффектом называется переход электронов вторичной эмиссии на другой электрод.

На рис.45 приведено распределение потенциала по линии, проходящей между витками сеток при $U_a < U_{c2}$. Вторичная эмиссия с анода возникает при $U_a > 15-20$ В, электрическое поле для вторичных электронов при $U_a < U_{c2}$ является ускоряющим и они устремляются к экранирующей сетке. В результате этого анодный ток уменьшается, ток экранирующей сетки увеличивается.

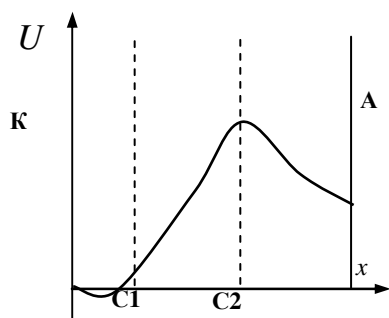


Рис.45

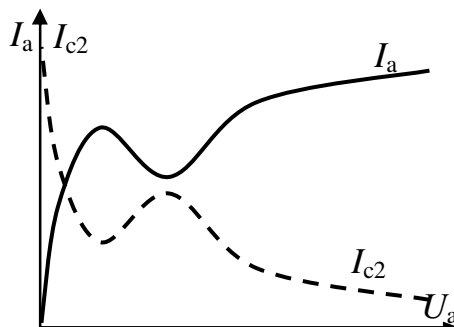


Рис.46

Коэффициент вторичной эмиссии с ростом U_a сначала увеличивается, затем начинает уменьшаться, анодный ток достигает минимального значения и начинает возрастать. При $U_a > U_{c2}$ вторичные электроны, выбитые с анода, возвращаются на анод, динаatronный эффект исчезает.

На рис.46 приведены зависимости анодного тока и тока второй сетки от анодного напряжения при $U_{c1} < 0$. Наличие падающего участка на анодной характеристике может привести к существенным нелинейным искажениям усиливаемого сигнала и самопроизвольной генерации колебаний.

Лучевой тетрод

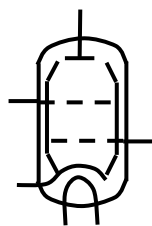


Рис.47

В лучевых тетродах (рис.47) динаatronный эффект подавляется за счет формирования плотных потоков (лучей) первичных электронов, создающих в пространстве А-С₂ объемный заряд, достаточный для образования потенциального барьера, препятствующего попаданию вторичных электронов с анода на экранирующую сетку. Формирование лучей осуществляется фокусировкой электронного потока в двух плоскостях.

В этих лампах витки управляющей и экранирующей сеток

делают с одинаковым шагом и размещают строго против друг друга. Проходя между витками отрицательно заряженной управляющей сетки, электроны собираются в узкие плоские пучки и пролетают мимо находящихся в «тени» витков экранирующей сетки, почти не оседая на них. В другой плоскости фокусировка осуществляется с помощью специальных лучеобразующих электродов, на которые подается нулевой потенциал (эти электроды показаны в условном изображении на рис.47).

Для создания требуемой величины минимума потенциала необходим значительный анодный ток, поэтому лучевые тетроды применяют, главным образом, как мощные выходные лампы.

Пентод

В пентодах (рис.48) для устранения динатронного эффекта между анодом и экранирующей сеткой размещается третья – защитная, или антидинатронная, – сетка. Третья сетка соединяется с катодом, т.е. имеет нулевой потенциал.

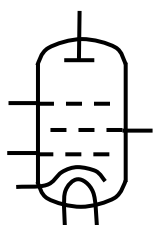


Рис.48

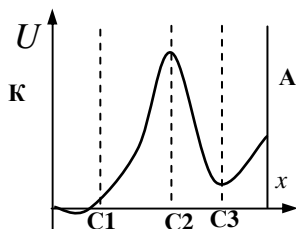


Рис.49

Распределение потенциала в пентоде показано на рис. 49, из которого видно, что вторичные электроны попадают в замедляющее поле между анодом и третьей сеткой, и поэтому они не достигают экранной сетки и возвращаются на анод. Так устраняется динатронный эффект.

Приводя пентод последовательно к эквивалентному тетроду, триоду и диоду, получаем действующее напряжение пентода:

$$U_d = U_{c1} + D_1 U_{c2} + D_1 D_2 U_{c3} + D_1 D_2 D_3 U_a, \quad (105)$$

где D_1, D_2, D_3 – проницаемости соответствующих сеток. С учетом того, что $U_{c3} = 0, D_1 D_2 D_3 \ll D_1$, получаем

$$U_d \approx U_{c1} + D_1 U_{c2} \quad (106)$$

Закон степени три вторых для пентода можно записать в виде

$$I_k = G_{п} (U_{c1} + D_1 U_{c2})^{3/2}, \quad (107)$$

где $G_{п}$ – первеанс пентода.

Из выражения (107) видно, что напряжение запириания по первой сетке $U_{c0} = -D_1 U_{c2}$ зависит от напряжения второй сетки и не зависит от анодного напряжения. Это результат того, что поле анода дополнительно экранировано от катода защитной и экранной сетками.

При $U_{c1} < 0$ катодный ток распределяется между экранирующей сеткой и анодом:

$$I_k = I_a + I_{c2}$$

При этом существуют два режима токораспределения. При $U_a < (0,1-0,5)U_{c2}$ – режим возврата электронов ко второй сетке, при $U_a > (0,1-0,5)U_{c2}$ – режим прямого перехвата, когда все электроны, пролетевшие между витками второй сетки, достигают анода. В режиме прямого перехвата коэффициент токораспределения $K_I = I_a / I_{c2}$ линейно зависит от отношения U_a / U_{c2} .

На рис. 50 приведены анодно-сеточные характеристики пентода для двух значений анодного напряжения и двух значений напряжения экранирующей сетки, на рис. 51 показано семейство выходных характеристик; больший индекс

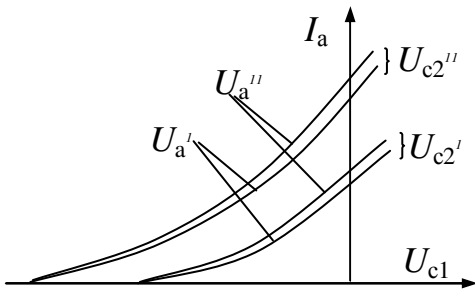


Рис.50

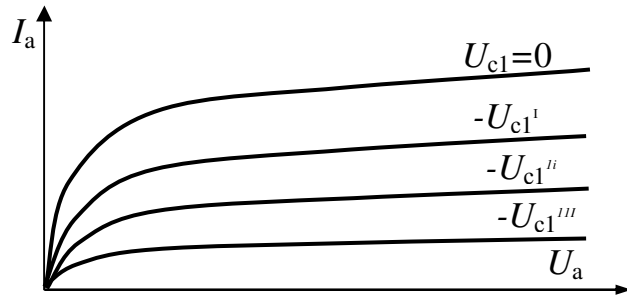


Рис.51

соответствует большему (по модулю) значению параметра.

Зависимость анодно-сеточных характеристик от анодного напряжения обусловлена увеличением K_I при увеличении U_a . Той же причиной обусловлен наклон пологих участков выходных характеристик, соответствующих режиму прямого перехвата. Крутые участки выходных ВАХ соответствуют режиму возврата электронов.

Дифференциальные параметры пентода. Основные дифференциальные параметры пентода – те же, что и триода: крутизна S , внутреннее сопротивление R_i и статический коэффициент усиления μ . При $U_{c3}=0$, $U_{c1}<0$ можно записать:

$$I_a = \frac{K_I}{1 + K_I} \cdot I_{\kappa}(U_a, U_{c1}, U_{c2});$$

$$S = \left(\frac{\partial I_a}{\partial U_{c1}} \right)_{U_a, U_{c2}} = \frac{K_I}{1 + K_I} \cdot S_{\kappa}, \quad (108)$$

где S_{κ} – крутизна катодно-сеточной характеристики;

$$G_i = \frac{1}{R_i} = \left(\frac{\partial I_a}{\partial U_a} \right)_{U_{c1}, U_{c2}} = \frac{K_I}{1 + K_I} \cdot \frac{\partial I_{\kappa}}{\partial U_a} + I_{\kappa} \cdot \frac{\partial}{\partial U_a} \left(\frac{K_I}{1 + K_I} \right); \quad (109)$$

$$\mu = - \left(\frac{\partial U_a}{\partial U_c} \right)_{I_a, U_{c2}} = S \cdot R_i < \frac{1}{D_1}. \quad (110)$$

Из выражений (108) видно, что при одинаковой конструкции узла катод–управляющая сетка крутизна пентода меньше, чем триода.

Первое слагаемое в формуле (109) отражает проникновение поля, создаваемого анодом, к катоду, оно практически равно нулю. Внутреннее сопротивление определяется вторым слагаемым, т.е. зависимостью коэффициента токораспределения от анодного напряжения.

Пентод характеризуется высоким внутренним сопротивлением и высоким статическим коэффициентом усиления.

Эквивалентные схемы ламп

На рис.52 приведена эквивалентная схема лампы для малого сигнала (линейного режима) в области частот, в которой еще не сказывается индуктивность выводов и конечная скорость пролета электронов от катода к аноду. Для мощных ламп эта область ограничена частотами порядка 10 МГц, для маломощных – десятков МГц.

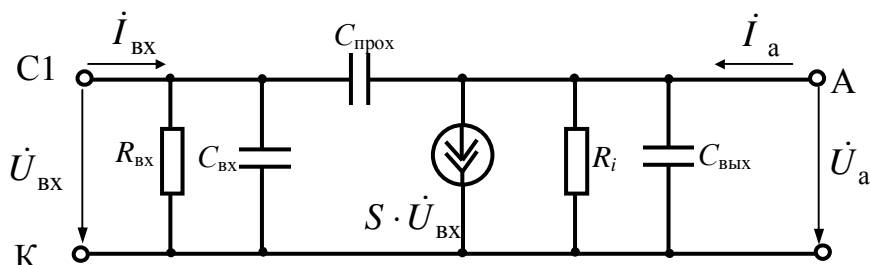


Рис.52

Для пентода

$$C_{\text{вх}} = C_{c1,k} + C_{c1,c2} + C_{c1,c3} \approx C_{c1,k} + C_{c1,c2}$$

$$C_{\text{вых}} = C_{a,k} + C_{a,c3} + C_{a,c2} \approx C_{a,c3} + C_{a,c2}$$

$$C_{\text{прох}} \approx C_{a,c1}$$

Для триода и тетрода в этих формулах исключаются соответствующие межэлектродные емкости. При $U_{c1} < 0$ можно пренебречь активной проводимостью во входной цепи ($R_{\text{вх}} = \infty$).

Проходная емкость $C_{\text{прох}}$ создает обратную связь между входом и выходом. В нагруженном (усилительном) режиме это приводит к увеличению входной емкости (эффект *Миллера*) и снижению граничной частоты:

$$C_{\text{вх.н}} = C_{c1,k} + C_{c1,c2} + (1 + K_U) C_{\text{прох}},$$

где K_U – коэффициент усиления. При наличии паразитных индуктивностей может возникнуть также самопроизвольная генерация (самовозбуждение лампы).

В маломощных высокочастотных пентодах проходная емкость снижена до $10^{-2} \div 10^{-3}$ пФ.

Области применения ламп

В настоящее время одним из основных применений ламп является использование их в мощных усилителях, генераторах и модуляторах.

Маломощные лампы применяются главным образом лишь в аппаратуре, работающей в тяжелых условиях эксплуатации: высокая температура, высокий уровень радиации.

3. ЭЛЕКТРОННО-ЛУЧЕВЫЕ ПРИБОРЫ (ЭЛП)

Электронно-лучевые приборы (ЭЛП) – электронные ЭВП, в которых используется поток электронов, сконцентрированный в виде одного или нескольких лучей. Приборы, у которых баллон имеет форму трубки, называются *электронно-лучевыми трубками (ЭЛТ)*.

ЭЛТ, предназначенные для получения видимого изображения на экране, получили название электронно-графических. К ним относятся приемные ЭЛТ: осциллографические, индикаторные, кинескопы, дисплеи ЭВМ и др.

В передающих трубках оптическое изображение преобразуется в последовательность электрических сигналов.

В запоминающих трубках возможно преобразование как электрических сигналов в изображение и изображения в последовательность электрических сигналов, так и оба вида преобразований.

Основными элементами ЭЛТ являются электронный прожектор, отклоняющая система и экран (или мишень). Электроды ЭЛТ располагаются внутри стеклянного баллона, в котором создан высокий вакуум.

Электронный прожектор (ЭП) формирует тонкий электронный луч.

Отклоняющая система (ОС) направляет этот луч в нужную точку экрана.

Экран представляет собой слой люминофора, нанесенного на внутреннюю поверхность дна баллона трубки. *Люминофор* – вещество, способное светиться при бомбардировке его электронами, в месте попадания луча появляется светящаяся точка, видимая через стекло.

Фокусировка и отклонение луча могут осуществляться электрическими и магнитными полями.

Электронные прожекторы ЭЛТ

Электронные прожекторы с электростатической фокусировкой.

Триодные прожекторы. Устройство простейшего электронного прожектора (ЭП) показано на рис.53. Все электроды ЭП имеют цилиндрическую форму и питаются от одного источника через делитель напряжения.

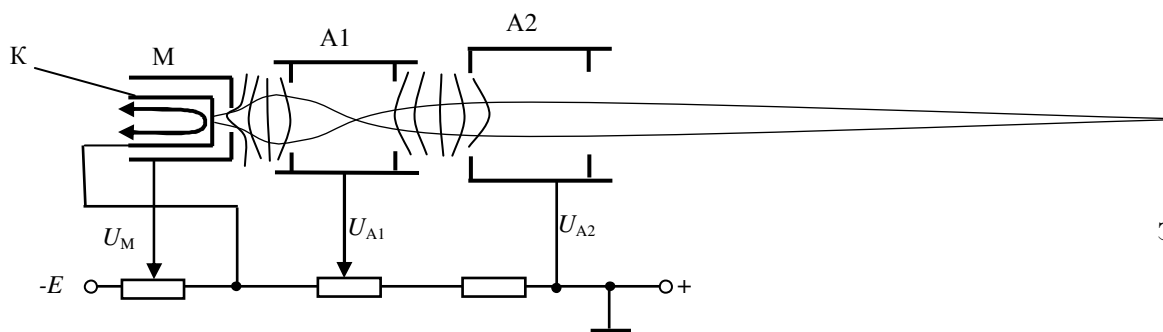


Рис.53

Катод (К) выполняется в виде небольшого цилиндра, внутри которого помещен подогреватель. Эмитирующей частью является дно цилиндра, покрытое люминофором.

оксидным слоем. Катод размещен внутри модулятора (М) – цилиндра с отверстием, который управляет током луча, аналогично управляющей сетке лампы. На модулятор подается отрицательное напряжение относительно катода. Электроны ускоряются и фокусируются напряжениями, приложенными к первому (А1) и второму (А2) анодам. На первый анод подается напряжение несколько сот вольт, на второй – несколько киловольт.

Электронные линзы. Между модулятором и первым анодом и между первым и вторым анодом образуются неоднородные аксиально-симметричные электрические поля, обладающие фокусирующим действием на электронный поток – *электронные линзы*. ЭП с двумя линзами называются триодными.

Принцип действия электростатических электронных линз поясняется на рис.54. Вблизи оси системы эквипотенциальные поверхности близки к сфери-

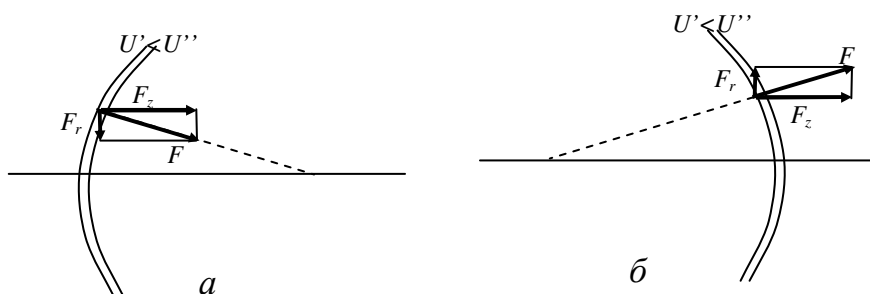


Рис.54

ческим и напряженность электрического поля направлена перпендикулярно к этим поверхностям к центру или от центра кривизны. Поэтому сила, действующая на электрон, имеет составляющие, направленные как вдоль оси F_z , так и перпендикулярно оси F_r . В неоднородном ускоряющем поле электроны получают ускорение в направлении оси, если эквипотенциальные поверхности обращены выпуклостью навстречу потоку электронов (рис.54,а), и линза является собирающей. Если эквипотенциальные поверхности обращены выпуклостью в направлении движения электронов, линза является рассеивающей (рис.54,б).

Фокусное расстояние линзы зависит от скорости электронов – чем больше скорость, тем больше фокусное расстояние. Как видно на рис.53, обе линзы в ЭП можно представить как совокупность собирающей и рассеивающей линз. В целом они, однако, являются собирающими, так как скорость электронов в собирающей области меньше.

Первая короткофокусная линза на рис.53 дает «изображение» пятна поверхности катода, с которого эмитируются электроны, вторая, длинно-фокусная, проектирует это «изображение» на экран. Фокусировка луча на экране осуществляется регулировкой напряжения первого анода.

Электронные линзы, как и оптические, обладают сферической aberrацией, заключающейся в том, что электроны, входящие в линзу под разными углами, фокусируются на несколько различных расстояниях. В результате вместо светящейся точки на экране получается небольшое пятно. Для уменьшения этого эффекта внутри анодов устанавливаются диафрагмы, отсекающие электроны, летящие под слишком большими углами к оси.

Яркость изображения управляется напряжением, подаваемым на модулятор. Зависимость тока катода от напряжения модулятора называется *модуляционной характеристикой*.

Недостатком простейшего ЭП является то, что при изменении яркости нарушается фокусировка, а при фокусировке луча изменяется яркость. Этот недостаток устраняется в ЭП с ускоряющим электродом, который применяется в большинстве современных ЭЛТ.

Устройство ЭП с ускоряющим электродом показано на рис.55. На ускоряющий электрод подается напряжение второго анода. Этот электрод является эффективным электростатическим экраном, поэтому потенциал первого анода не влияет на яркость.

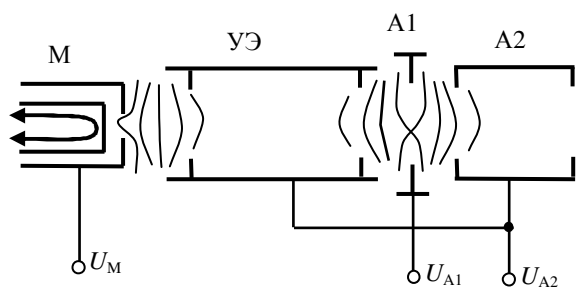


Рис.55

В этом прожекторе первый анод выполнен в виде диафрагмы с отверстием в 2–3 раза превышающим диаметр пучка, что исключает попадание электронов на этот электрод. Ток в цепи А1 равен нулю, поэтому изменение напряжения на нем при фокусировке луча не влияет на распределение тока и напряжения на элементах делителя, т.е. не вызывает изменения потенциала модулятора, а,

значит, и яркости.

Недостатком этого прожектора является то, что при больших ускоряющих напряжениях затруднено управление яркостью, поскольку запирающее напряжение получается слишком большим. Поэтому эти прожекторы применяются до напряжений $U_{A2}=5-8$ кВ. В высоковольтных ЭЛТ применяются тетродные прожекторы.

Тетродный прожектор. В тетродном прожекторе имеется два ускоряющих электрода и три электронные линзы (рис.56). На первый относительно короткий ускоряющий электрод УЭ1 подается напряжение $U_{УЭ1}=200-500$ В, что

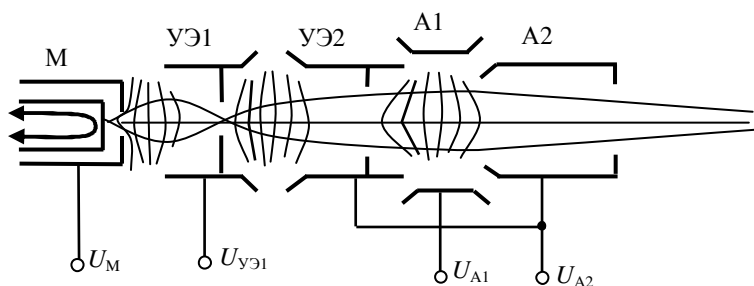


Рис.56

образованную в области УЭ2-А1-А2.

обеспечивает легкость управления яркостью, на второй – полное ускоряющее напряжение U_{A2} (до 15–20кВ). Первая, короткофокусная, линза между модулятором и УЭ1 фокусирует луч на оси, вторая уменьшает угол вхождения электронов в главную длиннофокусную линзу,

Электронный прожектор с магнитной фокусировкой.

В ЭЛТ с магнитной фокусировкой в качестве главной линзы, фокусирующей луч на экране, используется неоднородное аксиально-симметричное магнитное поле (магнитная линза) короткой катушки, у которой диаметр соизмерим с ее

длиной. Катушка надевается на горловину трубки. Фокусировка осуществляется регулированием тока катушки. ЭП может быть выполнен по триодной или тетродной схеме.

Принцип действия магнитной линзы поясняется на рис.57. На электрон в магнитном поле действует сила Лоренца, пропорциональная векторному произведению скорости \vec{v} и магнитной индукции \vec{B} :

$$\vec{F} = -e[\vec{v}\vec{B}] = e[\vec{B}\vec{v}]$$

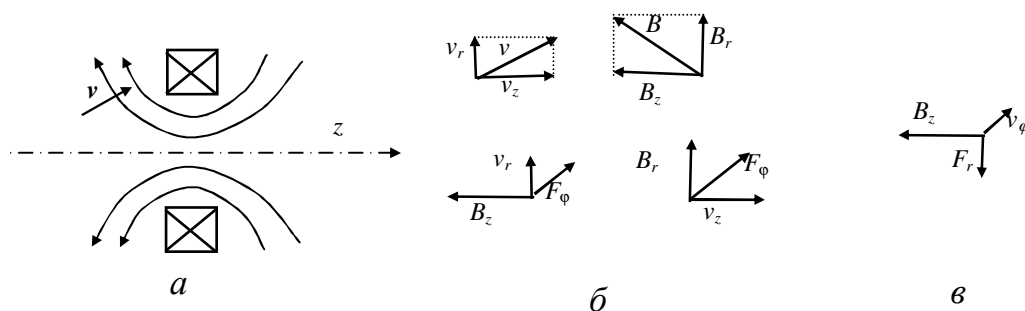


Рис.57

Векторы \vec{v} и \vec{B} лежат в плоскостях, включающих ось симметрии, сила направлена перпендикулярно этим плоскостям, т.е. закручивает электрон вокруг оси. В частности, на рис.57 сила Лоренца направлена перпендикулярно плоскости от читателя. На рис.57,б показано разложение векторов на радиальные и осевые компоненты и силы, обусловленные действием B_r на v_z и B_z на v_r . Обе силы направлены в одну сторону и закручивающая сила F_ϕ равна их сумме:

$$F_\phi = eB_z v_r + eB_r v_z$$

Электрон приобретает азимутальную составляющую скорости v_ϕ и возникает сила F_r , отклоняющая электрон к оси (рис.57,в). Электроны в магнитной линзе движутся по суживающейся спирали и вылетают из нее, приобретая составляющую скорости, направленную к оси. Регулируя ток катушки можно сфокусировать луч на экране.

Отклоняющие системы ЭЛТ

Электростатическая отклоняющая система. Простейшая электростатическая отклоняющая система (ЭОС) представляет собой две пары пластин, повернутых относительно друг друга на 90° . Каждая пара отклоняет луч по одной из координатных осей.

На рис.58 показано отклонение луча по оси y.

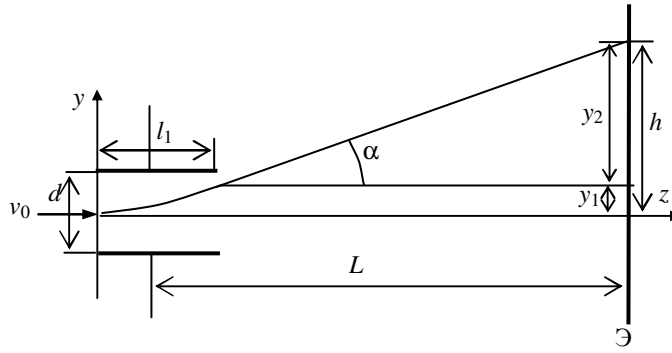


Рис.58

Между пластинами на электрон действует сила $F_y = -eE = eU_y/d$ и он движется с ускорением $a = eU_y/(d \cdot m)$, где e и m – заряд и масса электрона, E – напряженность электрического поля, U_y – разность потенциалов между пластинами, d – расстояние между пластинами. Электрон влетает в ОС со скоростью $v_z = v_0 = \sqrt{2eU_{A2}/m}$ и, двигаясь по параболической траектории, пролетает ОС за время $\tau = l_1/v_0$, приобретая к концу ОС компоненту скорости $v_{y1} = a\tau$ и отклонение $y_1 = a\tau^2/2$. Подставляя a , τ , v_0 , получаем

$$y_1 = \frac{l_1^2}{4d \cdot U_{A2}} \cdot U_y.$$

Вылетев из ОС под углом α к оси ($\text{tg}\alpha = v_{y1}/v_0$), электрон на пути до экрана приобретает отклонение

$$y_2 = l_2 \text{tg}\alpha = l_2 v_{y1} / v_0 = \frac{l_1 \cdot l_2}{2d \cdot U_{A2}} \cdot U_y.$$

Суммируя y_1 и y_2 , получаем полное отклонение на экране:

$$h = y_1 + y_2 = \frac{l_1 L}{2d \cdot U_{A2}} U_y,$$

где $L = l_2 + l_1/2$ – расстояние от центра ОС до экрана. Чувствительностью электростатической ОС (ЭОС) называется отношение величины отклонения к отклоняющему напряжению:

$$S_{\text{э}} = \frac{l_1 L}{2d \cdot U_{A2}} \quad (111)$$

Магнитная отклоняющая система. Магнитная ОС (МОС) обычно содержит две пары катушек, надеваемых на горловину трубки и создающих магнитные поля во взаимно перпендикулярных направлениях. Каждая пара отклоняет луч по одной из координатных осей.

На рис.59 показано отклонение луча в вертикальной плоскости (по оси y). Вектор \mathbf{B} направлен по оси x вглубь рисунка ($B_y = B_z = 0$). Электрон, влетая в однородное магнитное поле со скоростью v_0 , под действием силы Лоренца получает центростремительное ускорение a и движется по дуге окружности с радиусом r :

$$a = \frac{eBv_0}{m} = \frac{v_0^2}{r}; \quad r = \frac{mv_0}{eB}$$

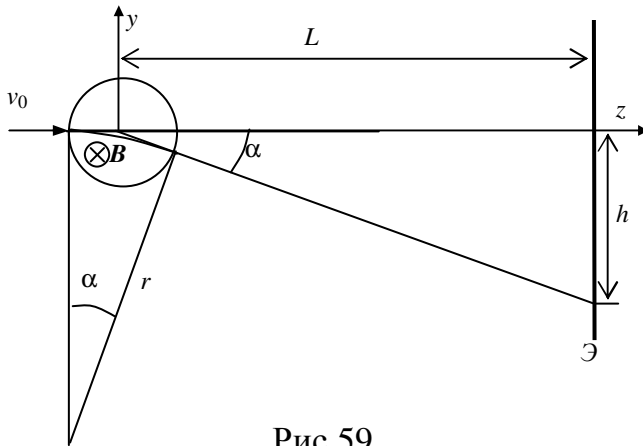


Рис.59

Электрон вылетает из магнитного поля под углом α к оси и далее движется прямо-линейно. Полное отклонение на экране

$$h = L \operatorname{tg} \alpha$$

Радиус окружности r , как правило, значительно больше диаметра отклоняющих катушек l_1 , поэтому путь, пройденный электроном в ОС $s = \alpha \cdot r \approx l_1$, и $\alpha \approx l_1/r$. При малых углах $\operatorname{tg} \alpha \approx \alpha \approx l_1/r$. Тогда

$$h \approx$$

$$L l_1 / r = \frac{e l_1 L}{m v_0} B = \frac{\sqrt{e / 2 m} \cdot l_1 L}{\sqrt{U_A}} B.$$

Магнитная индукция пропорциональна числу витков w и току катушки I :

$$B = k_1 w I,$$

где k_1 – коэффициент пропорциональности, определяемый конструкцией катушки, произведение wI называется магнитодвижущей силой. Чувствительность магнитной ОС определяется как отношение величины отклонения к магнитодвижущей силе:

$$S_M = k_1 \frac{\sqrt{e / 2 m} \cdot l_1 L}{\sqrt{U_A}}. \quad (112)$$

Сравнивая формулы (111) и (112), видим, что S_M существенно слабее зависит от ускоряющего напряжения, поэтому применяется в трубках с высоким анодным напряжением. МОС обладает бóльшей линейностью и допускает отклонение луча на бóльшие углы по сравнению с ЭОС. Достоинством МОС является также внешнее расположение, дающее возможность их вращения. В отличие от отклоняющих пластин обе пары катушек можно совместить по длине трубки и уменьшить общую длину ЭЛТ.

Недостатком МОС является значительно бóльшая потребляемая мощность и бóльшая инерционность. ЭОС могут работать на частотах до нескольких сотен мегагерц, МОС – до нескольких десятков килогерц.

Экраны ЭЛТ.

Экран представляет собой тонкий слой люминофора, нанесенный на внутреннюю поверхность торцевой части баллона. Различают катодолюминофоры – вещества способные интенсивно светиться при бомбардировке их электронами, и фотолюминофоры, которые светятся после облучения светом.

Характеристики экранов:

- цвет изображения или спектральная характеристика;
- яркость свечения, которая зависит от плотности тока луча j и потенциала экрана $U_{\text{Э}}$ согласно закону Ленарда:

$$B = A j (U_{\text{Э}} - U_0)^n, \quad (113)$$

где A и n – константы, зависящие от люминофора, U_0 – минимальный потенциал начала свечения;

- разрешающая способность, которая может измеряться числом различных точек на 1 см^2 , числом различных линий на 1 см высоты или на всю высоту экрана, шириной линий;
- длительность послесвечения – время спада яркости после прекращения бомбардировки до уровня 1%. Экраны разделяются на экраны с очень коротким (менее 10^{-5} с), коротким ($10^{-5} - 10^{-2} \text{ с}$), средним ($10^{-2} - 10^{-1} \text{ с}$), длительным ($10^{-1} - 16 \text{ с}$), и очень длительным (более 16 с) послесвечением.

Потенциал экрана.

Люминофор является диэлектриком и электроны, попадающие на экран, понижают его потенциал. Когда потенциал экрана сравняется с потенциалом катода, изображение полностью исчезнет. Отвод заряда от экрана возможен лишь за счет вторичной эмиссии. Для отвода вторичных электронов внутреннюю поверхность трубки вблизи экрана покрывают проводящим слоем – аквадагом (коллоидный графит), который соединяется со вторым анодом.

На рис.60 показана зависимость коэффициента вторичной эмиссии σ от ускоряющего напряжения. В области потенциалов между первым $U_{\text{кр1}}$ и вторым $U_{\text{кр2}}$ критическими потенциалами $\sigma > 1$, и

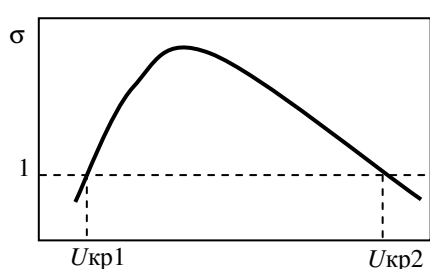


Рис.60

ускоряющее напряжение $U_{\text{А2}}$ должно лежать в этом интервале. Для разных экранов $U_{\text{кр1}} = 300 \div 500 \text{ В}$, $U_{\text{кр2}} = 5 \div 40 \text{ кВ}$.

При $U_{\text{А2}} < U_{\text{кр1}}$ потенциал экрана близок к нулю, при $U_{\text{А2}} > U_{\text{кр2}}$ потенциал экрана принимает значение $U_{\text{Э}} = U_{\text{кр2}}$ и яркость далее не увеличивается. В области $U_{\text{кр1}} < U_{\text{А2}} < U_{\text{кр2}}$ потенциал экрана устанавливается автоматически, когда число уходящих с экрана электронов станет равно числу падающих. При этом $U_{\text{Э}} \approx U_{\text{А2}}$.

В высоковольтных трубках часто применяется алюминированный экран: тонкий слой алюминия (1 мкм) наносится на слой люминофора и соединяется со вторым анодом ($U_{\text{Э}} = U_{\text{А2}}$). Высокая яркость достигается из-за высокого напряжения и из-за отражения света от слоя алюминия.

Особенности ЭЛТ различного назначения

Осциллографические трубки. Осциллографические ЭЛТ предназначены для получения графических изображений временных или функциональных зависимостей электрических сигналов.

Для наблюдения зависимости сигнала от времени на пластины, отклоняющие луч по вертикали, подается исследуемое напряжение, а на пластины, отклоняющие по горизонтали, – линейно изменяющееся во времени

(пилообразное) напряжение, называемое напряжением развертки. Луч пробегает с постоянной скоростью по горизонтальной оси, одновременно отклоняясь по вертикали пропорционально напряжению сигнала, таким образом на экране получается график зависимости напряжения от времени. Затем луч быстро возвращается в исходную точку (луч при этом гасится) и начинается следующий цикл развертки.

В осциллографических трубках (ОТ) используются ЭП с электростатической фокусировкой, обеспечивающие простоту и экономичность схем питания, и электростатические ОС, имеющие малую емкость и индуктивность. Обычно применяется триодный прожектор с ускоряющим электродом и напряжением питания U_{A2} до нескольких кВ.

Экраны ОТ обычно имеют зеленое свечение – для визуального наблюдения и синее или голубое – для фотографирования.

Большинство ОТ предназначены для регистрации процессов с частотой до 25–30 МГц. Для работы с более высокочастотными сигналами применяются специальные *широкополосные* трубки.

Регистрация быстропротекающих процессов требует увеличения скорости развертки. При низкой частоте повторения это приводит к падению яркости, для сохранения которой необходимо повысить U_{A2} , при этом, однако, падает чувствительность ОС. Эта проблема решается в трубках с послеускорением, в которых луч после прохождения ОС дополнительно ускоряется напряжением $U_{A3} > U_{A2}$ третьего анода. При необходимости, может быть несколько ступеней ускорения. Дополнительные аноды изготавливаются в виде колец из проводящего покрытия (аквадага).

Разрывы между анодами образуют электростатические линзы, что может приводить к искажению осциллограмм. Для уменьшения искажений дополнительные аноды изготавливают достаточно большого диаметра в расширенной цилиндрической части трубки и напряжение анодов повышают постепенно (не более чем в два раза на каждой ступени). Таким способом можно довести ускоряющее напряжение до 25–30 кВ и скорость развертки – до 50 м/мкс. Луч в трубках с послеускорением можно включать лишь при наличии развертки, иначе произойдет сильный локальный разогрев экрана и выгорание люминофора.

Чувствительность ОС падает, когда время пролета ОС становится сравнимым с периодом колебаний. Для расширения частотного диапазона отклоняющие пластины разделяются на несколько более коротких пар, индуктивно связанных между собой, т.е. пластины заменяются отрезком длинной линии с сосредоточенными параметрами. Скорость распространения сигнала в этой линии согласована со скоростью электронов. Для дальнейшего расширения частотного диапазона отклоняющие пластины заменяются отрезком спиральной замедляющей системы.

Радиолокационные трубки (РЛТ). В радиолокации применяются индикаторные ЭЛТ с яркостной отметкой, которая сигнализирует о появлении отраженного сигнала и указывает координаты цели в полярных координатах.

В этих трубках применяется круговая развертка и радиальное отклонение луча. При выключенной развертке луч фокусируется в центре экрана. При

включении развертки луч пробегает по радиусу, одновременно равномерно отклоняясь в азимутальном направлении. Затем луч быстро возвращается в центр и начинается следующий цикл развертки. В результате за период круговой развертки след луча покрывает всю площадь экрана.

Такой вид развертки обычно осуществляется вращением одной пары отклоняющих катушек. На катушки подается ток пилообразной формы, который создает линейно растущее магнитное поле, вызывающее отклонение луча по радиусу. Поворот катушек приводит к повороту плоскости отклонения, в результате реализуется круговая развертка. Вращение катушек синхронизировано с вращением антенны.

В отсутствие сигнала на модулятор подается запирающее напряжение – луч выключен. Приходящий сигнал подводится к модулятору и отпирает луч. На экране появляется светящаяся точка. По положению этой точки на экране определяется азимутальный угол и расстояние до цели, которое пропорционально расстоянию от центра экрана до светящейся точки.

Для обеспечения точности определения координат требуется высокое качество фокусировки и высокая линейность отклонения луча.

Для обеспечения достаточной яркости при малой длительности отпирающего импульса необходима большая мощность луча, т.е. достаточно большой ток луча и высокое ускоряющее напряжение.

Вследствие малой скорости круговой развертки нужно длительное послесвечение экрана.

Длительное послесвечение экрана позволяет проследить за траекторией цели и, в частности, определить ее скорость.

Всем этим требованиям наиболее полно можно удовлетворить в случае применения магнитных систем фокусировки и отклонения луча. МОС применяется во всех РЛТ, МФС – в большинстве. В РЛТ используется высокое ускоряющее напряжение (до 20 кВ).

Экраны РЛТ имеют длительное и очень длительное послесвечение и высокую яркость изображения. Применяются также экраны, которые не светятся под действием электронного луча, а темнеют. Трубки с таким экраном называются *скиатронами*. В этих трубках вместо люминофора используется напыленный слой хлористого калия.

Кинескопы черно-белого телевидения. Кинескоп – это ЭЛТ, предназначенная для преобразования телевизионного сигнала в световое изображение.

По стандарту, принятому в России и большинстве других стран, кадр телевизионного изображения разбивается на 625 строк. Телевизионный сигнал представляет собой последовательность импульсов, каждый из которых соответствует одной строке изображения. Амплитуда импульса изменяется во времени в соответствии с яркостью передаваемого изображения в данной строке.

Телевизионный сигнал подается на модулятор ЭЛТ, одновременно включается развертка луча, который прописывает строку изображения, модулированную по яркости в соответствии с амплитудой сигнала. Затем луч смещается на шаг вниз и прописывается следующая строка и т.д., все 625 строк.

После этого луч возвращается к первой строке и начинает прописываться изображение следующего кадра.

В кинескопах применяется электростатическая фокусировка луча и магнитная отклоняющая система. Электронный прожектор – тетродный, с напряжениями $U_{УЭ1} = 200\text{--}600\text{ В}$ и $U_{A2} = 10\text{--}20\text{ кВ}$. Высокое ускоряющее напряжение необходимо для обеспечения достаточной яркости. Наличие дополнительного УЭ1 с небольшим напряжением облегчает управление яркостью. Этот прожектор вполне удовлетворяет требованиям к качеству фокусировки, значительно менее жестким, чем в РЛТ.

МОС обеспечивает высокое качество изображения без искажений и дефокусировки, позволяет отклонить луч на большие углы (до 110°) и уменьшить длину трубки.

Экраны кинескопов имеют прямоугольную форму стандартных размеров и покрыты смесью люминофоров с желтым и голубым свечением, в итоге цвет излучения получается близким к белому. Экран имеет среднее послесвечение. Часто применяется алюминированный экран, позволяющий повысить яркость изображения и предохранить люминофор от бомбардировки отрицательными ионами, образующимися внутри трубки.

Цветные кинескопы. Действие цветных кинескопов основано на особенностях восприятия цвета человеческим глазом, который имеет цветочувствительные элементы трех типов, каждый из которых реагирует лишь на один из трех цветов: синий, зеленый и красный. Все разнообразие цвета воспринимается как результат смешения трех основных цветов в различном количественном соотношении.

Цветной кинескоп имеет три электронных прожектора, расположенных в вершинах равностороннего треугольника (дельтавидное расположение) или на одной линии в плоскости (планарное расположение). Каждый прожектор управляется сигналом, соответствующим одному из трех основных цветов.

Экран содержит зерна люминофора, светящиеся синим, зеленым и красным цветом с дельтавидным или линейчатым расположением, соответственно. На небольшом расстоянии (12 мм) от экрана размещена маска – тонкая (0,15 мм) металлическая пластина с множеством ($3 \cdot 10^5\text{--}6 \cdot 10^5$) отверстий диаметром 0,25 мм. Отклонение всех трех лучей по строкам осуществляется одновременно одной и той же МОС. Независимо от угла отклонения все три луча пересекаются в плоскости маски и проходят через одно отверстие, далее лучи расходятся и каждый луч попадает на люминофор «своего» цвета. Экран и маска являются единым целым, даже небольшое смещение маски может исказить передачу цвета.

СПИСОК ЛИТЕРАТУРЫ

1. *Аваев Н.А., Шишкин Г.Г.* Электронные приборы: Учебник для вузов. Под ред. проф. Г.Г. Шишкина. М.: Изд-во МАИ, 1996.
2. *Степаненко И.П.* Основы микроэлектроники: Учебное пособие для вузов. М.: Лаборатория Базовых Знаний, Изд-е 3-е, 2004.
3. *Тугов Н.М. и др.* Полупроводниковые приборы: Учебник для вузов / Под ред. В.А. Лабунцова. М.: Энергоатомиздат, 1990.
4. *Бобровников Л.З.* Электроника: Учебник для вузов. Изд-е 5-е, перераб. и доп. СПб.: Питер, 2004.
5. Электронные, квантовые приборы и микроэлектроника: Учебное пособие для вузов/ Ю.Л.Бобровский и др.; Под. ред. проф. Н.Д.Федорова. М.: Радио и связь, 1998.
6. *Зи С.* Физика полупроводниковых приборов: В 2-х книгах. Пер. с англ. – Изд-е 2-е, перераб. и доп. – М.: Мир, 1984.
7. *Викулин И.М., Стафеев В.И.* Физика полупроводниковых приборов. Изд-е 2-е, перераб. и доп. М.: Радио и связь, 1990.
8. *Щука А.А.* Электроника: Учебное пособие / Под ред. проф. А.С. Сигова. – СПб.: БХВ-Петербург, 2005.
9. *Блейкмор Дж.* Физика твердого тела. Пер. с англ. М.: Мир, 1988.
10. *Смит Р.* Полупроводники. Пер. с англ. М.: Мир, 1982.

Урманчиев Линард Мазитович

Физические основы электроники

Учебное пособие

Ответственный за выпуск Н.И. Данич
Технический редактор С.В. Фокеева
Компьютерная верстка –